Earth Syst. Dynam. Discuss., 6, 761–818, 2015
www.earth-syst-dynam-discuss.net/6/761/2015/
doi:10.5194/esdd-6-761-2015

Earth System Dynamics Discussions
Open Access

Leaf Area Index in Earth System Models: evaluation and projections

N. Mahowald[1], F. Lo[1], Y. Zheng[1], L. Harrison[2], C. Funk[2], and D. Lombardozzi[3]

[1]Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY, 14853, USA
[2]Department of Geography, University of California at Santa Barbara, Santa Barbara, CA 93106, USA
[3]Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, CO, 80307, USA

Received: 3 March 2015 – Accepted: 17 March 2015 – Published: 15 April 2015

Correspondence to: N. Mahowald (mahowald@cornell.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The amount of leaves in a plant canopy (measured as leaf area index, LAI) modulates key land–atmosphere interactions, including the exchange of energy, moisture, carbon dioxide (CO_2), and other trace gases, and is therefore an essential variable in predicting terrestrial carbon, water, and energy fluxes. The latest generation of Earth system models (ESMs) simulate LAI, as well as provide projections of LAI in the future to improve simulations of biophysical and biogeochemical processes, and for use in climate impact studies. Here we use satellite measurements of LAI to answer the following questions: (1) are the models accurately simulating the mean LAI spatial distribution? (2) Are the models accurately simulating the seasonal cycle in LAI? (3) Are the models correctly simulating the processes driving interannual variability in the current climate? And finally based on this analysis, (4) can we reduce the uncertainty in future projections of LAI by using each model's skill in the current climate? Overall, models are able to capture some of the main characteristics of the LAI mean and seasonal cycle, but all of the models can be improved in one or more regions. Comparison of the modeled and observed interannual variability in the current climate suggested that in high latitudes the models may overpredict increases in LAI based on warming temperature, while in the tropics the models may overpredict the negative impacts of warming temperature on LAI. We expect, however, larger uncertainties in observational estimates of interannual LAI compared to estimates of seasonal or mean LAI.

Future projections of LAI by the ESMs are largely optimistic, with only limited regions seeing reductions in LAI. Future projections of LAI in the models are quite different, and are sensitive to climate model projections of precipitation. They also strongly depend on the amount of carbon dioxide fertilization in high latitudes. Based on comparisons between model simulated LAI and observed LAI in the current climate, we can reduce the spread in model future projections, especially in the tropics, by taking into account model skill. In the tropics the models which perform the best in the current climate tend to project a more modest increase in LAI in the future compared to the average

ESDD
6, 761–818, 2015

of all models. These top performing models also project an increase in the frequency of drought in some regions of the tropics, with droughts being defined as minus one standardized deviation events.

1 Introduction

Providing future projections of climate change impacts is one of the goals motivating the development of Earth system models (ESMs). Included in many of these models are carbon cycle modules, which include simulations of land vegetation (Friedlingstein et al., 2006). These models predict leaf area index (LAI) and other carbon cycle variables. LAI represents the amount of leaf area per unit land area, and is an important land carbon attribute. Many ESMs calculate leaf-level carbon and water fluxes, which are then scaled regionally and globally based on LAI (e.g. Oleson et al., 2013). The surface energy budget, as well as plant-based emissions and deposition of aerosols and chemically and radiatively important gases, are also sensitive to predicted LAI (e.g. Oleson et al., 2013). Therefore, small errors in simulated LAI can become large errors in many ESMs' biophysical and biogeochemical processes, and changes in LAI alone can change climate (e.g. Bounoua et al., 2000; Ganzeveld et al., 1998; Lawrence and Slingo, 2004; Oleson et al., 2013). Additionally, LAI can be observed from satellite (Zhu et al., 2013), and thus represents one of the few land carbon variables that can be directly evaluated in coupled models (e.g. Randerson et al., 2009). Finally LAI (and the related normalized difference vegetation index, NDVI) changes can indicate the health of our ecosystems and the availability of natural resources. As such, LAI is widely used in the famine prediction community (Funk and Brown, 2006; Groten, 1993) and represents a variable that is easy to use in climate impacts studies. Thus it is important to evaluate the current generation of coupled models for their ability to simulate LAI and consider the 21st century projections for LAI in Earth System Models.

The current generation of Earth System Models have prepared historical and future scenario simulations, as part of the preparation for the Intergovernmental Panel on Cli-

ESDD
6, 761–818, 2015

mate Change (IPCC), organized in phase 5 of the Coupled Modeling Intercomparison Project (CMIP5), part of the Working Group on Coupled Models, of the World Climate Research Program (Taylor et al., 2009). There have been extensive evaluations and comparisons of the carbon cycle in the earth system models in the latest CMIP5 (e.g. Arora et al., 2013; Friedlingstein et al., 2013; Jones et al., 2013). Most of this analysis has focused on the ability of the models to predict the land, ocean and atmospheric carbon cycle (e.g. Arora et al., 2013; Friedlingstein et al., 2013; Jones et al., 2013). There has also been some evaluation of seasonal variability in LAI in high latitudes compared to satellites for the ESMs (Anav et al., 2013; Murray-Tortarolo et al., 2013). Additionally, Shao et al. (2013) evaluated the relationship between the carbon cycle and other variables, for example, temperature, over decadal and longer time scales. It should also be noted that there is a long history of model evaluation of vegetation outside of earth system models as well (e.g. Cramer et al., 1999).

Here, we evaluate the models using the satellite-derived 30 year products that have recently become available (Zhu et al., 2013) and have been used to identify trends in NDVI and LAI (Forkel et al., 2013; Jong et al., 2013; Mao et al., 2013; Vrieling et al., 2013). While satellite derived LAI estimates are known to have systematic and random errors, they have been usefully employed to evaluate the relative importance of different climate factors (e.g. temperature, precipitation) for vegetation productivity (Zeng et al., 2013). We expand on previous studies that evaluated simulated LAI (e.g. Anav et al., 2013) by looking at LAI means and variability across all latitudes, and considering what climate factors impact interannual variability.

There are several potential drivers of LAI changes in the future, such as temperature, precipitation, as well as carbon dioxide fertilization, which can impact future projections. Based on the interannual variability of LAI and climate drivers in the current climate, we consider whether the models can reproduce the observed relationships, suggesting they have the correct sensitivity to such important drivers of LAI as temperature or precipitation (e.g. Fung et al., 2005; Lobell et al., 2011; Zeng et al., 2013). The main questions we seek to answer in this paper are (1) are the models accurately simu-

ESDD
6, 761–818, 2015

lating the mean LAI spatial distribution? (2) Are the models accurately simulating the seasonal cycle in LAI? (3) Are the models correctly simulating the processes driving interannual variability in the current climate? And finally based on this analysis, (4) can we reduce the uncertainty in future projections of LAI by using each model's skill in the current climate? (e.g. Cox et al., 2013; Hoffman et al., 2014; Steinacher et al., 2010)? To this end, we develop several metrics to evaluate the models' ability (similar to that done for other climate variables, e.g. Taylor, 2001), some of which could be used in future model intercomparisons (e.g. Luo et al., 2012; Randerson et al., 2009).

In this paper we present in our methods Sect. 2 and a comparison of LAI variability in space and time between observations and the models in Sect. 3. In Sect. 4 projections of climate change in temperature, precipitation and LAI are shown, while Sect. 5 presents our summary and conclusions.

2 Methods and datasets

2.1 Observational data

Leaf Area Index (LAI) data derived from satellite over the 30 year period 1981–2010 are used to evaluate the CMIP5 models. This observational dataset is derived using neural network algorithms using the Global Inventory Modeling and Mapping (GIMMS) Normalized Difference Vegetation Index (NDVI3g) and the Terra Moderate Resolution Spectroradiameter (MODIS) LAI (Zhu et al., 2013). A detailed description of the algorithm and comparison to ground-truth observations are shown in Zhu et al. (2013). The satellite data are only available over regions with green vegetation, and thus are lacking over desert and arid regions. Root Mean Squared Errors (RMSE) in the LAI estimates, derived from comparison with land-based observations are estimated to be approximately 0.68 LAI (Zhu et al., 2013); the new LAI product seems to be able to capture observed interannual variability patterns based on comparisons to ground based data (Zhu et al., 2013).

ESDD
6, 761–818, 2015

Gridded temperature data for the period 1981–2010 were derived from the GHCN_CAMS 2 m temperature dataset (Fan and Dool, 2008), while the precipitation was derived from the Global Climatological Precipitation Project (Adler et al., 2003). Estimates of the uncertainty in gridded datasets suggest that the uncertainty in temperatures at a grid box level is estimated to be 0.2 °C (Jones et al., 1997). Uncertainty in precipitation uncertainty is larger and can be as large as 45 % in poorly observed regions (e.g. Dai et al., 1997).

2.2 Model datasets

The Climate Model Intercomparison Project (CMIP5), as part of the Working Group on Coupled Models of the World Climate Resource Program, organized a set of experiments which were assessed as part of the 5th Assessment of the Intergovernmental Panel on Climate Change (Taylor et al., 2009). Coupled carbon model experiments were included in the CMIP5 (e.g. Friedlingstein et al., 2006). The historical simulations and Representative Concentration Pathway for 8.5 (RCP8.5; van Vuuren et al., 2011, 2007), using fixed carbon dioxide concentrations, were analyzed here for the models (Table 1). We chose to focus on the RCP8.5 scenario as it has the largest changes in carbon dioxide and climate.

Model variables analyzed included monthly-mean precipitation, surface temperature, and LAI. Only models which had data for all these variables, for historical and RCP8.5 scenarios, were included in this study. Some models submitted multiple versions, at different resolutions or with slightly different physics (Table 1). Even though some of the models are closely related (e.g. CESM1-BGC and NorESM-ME), we include different configurations of the same model.

2.3 Methodology for evaluation of LAI's relationship with climate variables

In order to assess the ability of the earth system models to simulate the temperature and precipitation dependence of LAI in the future, we use current relationships in the

ESDD
6, 761–818, 2015

observations of LAI and climate. We want to evaluate whether the models simulate the correct temperature and precipitation impacts on vegetation. To do so, we develop several metrics.

For the models and the observations, we show results based on annual averages. We also considered a more complicated time period, where the LAI and temperature are based on growing seasons. The growing season is defined as the monthly maximum LAI and its two adjacent months. Because previous studies (e.g. Zeng et al., 2013) have shown that precipitation shows the highest correlation at 1–3 months ahead of vegetation, we use average precipitation for the month of the maximum LAI and the three previous months. This implies that pre-maximum LAI precipitation is most important for soil moisture during the growing season (Funk and Budde, 2009). Zeng et al., 2013 showed that temperature correlations are highest with vegetation during the month of maximum LAI, and thus temperature during the growing season is used. Results obtained using the growing season were quantitatively different from using an annual average, but qualitatively similar, and with similarly strong correlations. Thus for simplicity we present only results using the annual time period metrics in this paper.

Results for the model and observations are evaluated on a $2.5^\circ \times 2.5^\circ$ grid based on the observed temperature and precipitation data (see Sect. 2.1). For the analysis here, the averages shown are grid-box means, not areal averages. This allows us to use similar weighting for both the averages and the rank correlation coefficients, and tends to weight the global analysis towards high latitudes. However, most of the analysis focuses on regional areas (tropical ($< 30^\circ$), mid-latitudes ($> 30^\circ$ and $< 60^\circ$) and high-latitudes ($> 60^\circ$)), where the differences between weighting by area and weighting by grid box will be reduced.

The metrics used in this study are summarized in Table 2. We examine observed and model-simulated current climate mean of LAI (similar to previous studies e.g. Randerson et al., 2009). The annual mean LAI in the models and observations are averaged over different areas: global, tropical ($< 30^\circ$), mid-latitudes ($> 30^\circ$ and $< 60^\circ$) and high-latitudes ($> 60^\circ$) and compared (Table 2: mean LAI: model/obs.). A second metric

ESDD
6, 761–818, 2015

evaluates the models' ability to capture spatial variations in LAI, using the spatial correlation across the grid-boxes of the annual mean LAI in the model compared to the observations. (Table 2: Mean: Corr.).

Important for this study is the consideration of the temporal variability simulated in the model. The magnitude of the seasonal cycle is calculated as the SD (standard deviation) of the climatological month means over 30 years at each grid box. Metrics for the seasonal cycle were computed by a spatial average over each region (Table 2: SD Seasonal: Model/obs.). This allows us to consider over broad regions whether the model has too strong or weak of a seasonal cycle. For the seasonal cycle, the ability to capture the timing of phenology can be important (e.g. Anav et al., 2013). To analyze this ability, we computed the temporal correlation of observed and model-simulated monthly means at every grid box, and then averaged over each region (Table 2: Seasonal Avg. Corr.).

To evaluate interannual variability (IAV), we consider the magnitude of the interannual variability, which is calculated as the SD across years at each grid box. The IAV is then spatially averaged and compared between the model and observations (Table 2: SD IAV: Model/obs.). This comparison will show if the models exhibit too strong or weak of IAV across wide regions.

Previous studies have examined correlations between temperature and precipitation and satellite NDVI (normalized difference vegetation index) (e.g. Zeng et al., 2013), which is used to derive LAI (e.g. Zhu et al., 2013). The observed LAI at high latitudes tend to be dominated by changes in temperature, while the tropics are more dominated by moisture (Zeng et al., 2013), which is also seen in coupled-carbon climate models for carbon cycle variables (e.g. Fung et al., 2005). In order to understand what may be driving the IAV in the LAI, we calculate metrics to look at the correlation between anomalies in LAI and anomalies in temperature, precipitation, as well as trends with time. Of course, correlations do not show causation, but can show relationships that are consistent with certain causations. In an ideal world, we would like the models to emulate the same sensitivity of LAI to temperature or precipitation as in the real world.

ESDD
6, 761–818, 2015

Similar to Zeng et al. (2013) we assess the rank correlation of temperature and precipitation to leaf area index in the observations but we also look at trends over the last 30 years. At each grid box, a correlation between the annual mean of temperature and precipitation and time (e.g. the trend with time) against leaf area index are obtained for each model and observation. Because the models do not simulate exactly the same climate as observed, we cannot expect the models to simulate the same LAI at each grid box. However, we expect that across multiple grid boxes that the relationships should be similar between the model and observations. Thus, we calculate the temporal correlation between LAI and temperature, LAI and precipitation, and LAI and time (e.g. incrementing the year, to see if there is a trend in time), at every grid point, and calculate the spatial average of each correlation across the different regions (Table 2: LAI vs. T: Avg. Corr. for example).

We also perform a comparison of LAI simulated in a fully coupled model simulation to LAI simulated by a land model driven by observationally derived datasets (called CLM-obs) (e.g. Qian et al., 2006), but extended using CRU data through 2010 (Harris et al., 2013) for the Community Land Model (Lawrence et al., 2012; Lindsay et al., 2014). For this model, substantial computer software development has occurred so that the same model can be similarly driven by observation-based meteorology ("offline) or the simulated meteorology ("online") (Oleson et al., 2013). The reason for including this simulation is to test the sensitivity of the results to different driving meteorology. If we compare the same model driven by different meteorological data, we can isolate metrics that can identify model traits, from those that are dependent on meteorology, or simply are not strong enough to be used as metrics. Of course this analysis is dependent on the model and datasets used, but can be used as a sensitivity study to suggest how important meteorological factors are in the analysis.

Other metrics were also considered for this paper, including the overlap in the probability density functions (e.g. Maxino et al., 2008) and root mean squared differences, but these did not provide additional information that would justify the additional complexity in the paper.

ESDD
6, 761–818, 2015

Land use can heavily perturb the mean and evolution of the seasonal cycle and interannual variability in LAI, so we perform sensitivity tests where we exclude grid boxes with more than 50 % of agriculture based on Ramankutty et al. (2008). Results were quantitatively and qualitatively similar to those presented here.

There are several ways to present metrics to ease the evaluation and interpretation of models (e.g. Taylor, 2001; Gleckler et al., 2008). Here we chose to present the metrics described above in one figure (Fig. 6) with higher numbers representing a better simulation, for ease of interpretation. For correlations, this representation is straightforward, since 1 is a perfect correlation and lower values represent a worse simulation, so we chose to convert the model mean/observational metric to have the same approximate range for ease of display. We convert the mean model bias metric (model/obs) to a value that varies between 0 and 1, with 1 being close to perfect. We want to penalize models which have too high of a mean equally with model that have too low of a mean. In order to do this, we use the following formula to display the values in Fig. 6:

$$\text{Model Evaluation Value} = \frac{2}{\left\{\frac{\text{Model Mean}}{\text{Observed Mean}} + \frac{\text{Observed Mean}}{\text{Model Mean}}\right\}} \quad (1)$$

We use this same way to convert mean biases and SD biases, which is slightly different than previous studies (e.g. Gleckler et al., 2008), since it does not square the SDs. Since we use ranks and rank correlations, the difference between these methods is unlikely to be significant, and allows us to use a similar ranking method for mean and SD comparisons.

2.4 Methodology for future projections

We evaluate the mean change in the future using the RCP8.5 scenarios for temperature, precipitation, and LAI (e.g. Tables 3 and 4) using the models listed in Table 1. RCP8.5 is the most extreme scenario in the CMIP5 archive. We use it to identify regions that are most at risk in the future. Areas with mean changes in LAI, precipita-

ESDD
6, 761–818, 2015

tion, and temperature that are larger than the historical variability indicate statistically changes to climate and vegetation from current climate. To highlight where models predict these areas will be, mean changes over 20 year time periods are divided by the SD over the current climate (1981–2000) and shown in terms of SD units (e.g. Mahlstein et al., 2012; Tebaldi et al., 2011). The spatial and temporal scale we use to define these changes can be important for whether these signals are statistically significant (Lombardozzi et al., 2014) and we calculate this using a 20 year time scale at the grid level.

In addition, there could be changes in LAI variability, which may also be important for understanding the impact of climate change. For many regions we are concerned about the incidence of time periods with low precipitation and/or high temperatures causing low vegetative productivity, which we will refer to here as drought. In terms of drought, the length and frequency matter, so the percent of the time during which the variable is in drought is also calculated. We define drought as time periods where LAI is one SD (evaluated during the current climate) below that of the mean during the current climate. By definition, if the variables have a Gaussian distribution, each gridbox would be in "drought" for 1/6 (16 %) of the time. Thus we seek to estimate the fraction of the time in the future that this condition exists, and specifically whether it increases in the future.

3 Evaluation of model LAI

Our goal in this section is to explore the value of a new 30 year satellite LAI record for evaluating LAI simulations in the current generation of CMIP5 models. Anav et al. (2013) evaluated the LAI seasonal cycle and interannual variability for current climate high latitude Northern Hemisphere simulations. Here we look across all regions and also look at the temperature and precipitation as potential drivers of interannual variability.

ESDD
6, 761–818, 2015

3.1 Climatological comparison

The observed mean LAI has the largest values of leaf area index in the tropics (Fig. 1a). The largest seasonal cycle tends to be in mid-latitude regions, although there is still a signal in some parts of the tropics (Fig. 1b). The interannual variability tends to be much smaller than the seasonal cycle, and is equally large in tropics, mid-latitudes and high latitudes (Fig. 1c). One should note that there are many possible errors in the observational datasets, although the latest versions used here tend to have smaller biases than previous versions (Zhu et al., 2013). In this study, we compare model simulations against the satellite data, recognizing that these data are not perfect, and thus our conclusions are sensitive to potential biases in the observational data.

The models tend to overestimate the mean LAI compared to the observations (Fig. 2, Table 3), and this is true at all latitudes (Fig. 2, Table 4). Several models have a large overestimates (> 50 % too high), including bcc-csm1, bcc-csm1-1, BNU-ESM, GFDL-ESM2G, GFDL-ESM2M, MIROC-ESM. The overpredictions relative to the satellite data tend to be larger in tropical regions for most models, but are also larger in the high latitudes for the GFDL model versions (Fig. 2, Table 4).

Some models also tend to over predict the strength of the seasonal cycle (e.g. bcc-csm1, BNU-ESM, MIROC-ESM) (Fig. 2b, h, and k; Table 4), where the strength of the seasonal cycle is measured by the globally averaged SDs of the monthly mean climatology. But the region in which they overpredict the strength of the seasonal cycle differ. Several models underpredict the seasonal cycle at high-latitudes (e.g. CanESM2, CESM1-BGC, GFDL-ESM2G, GFDL-ESM2M, INMCM4, IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR) (Fig. 2b, h and k; Table 4; Anav et al., 2013). The magnitude and direction of bias in model projections also vary by region. For example, one set of models overpredicts the strength of the seasonal cycle in the tropics and mid-latitudes, but underpredicts in the high-latitudes (e.g. bcc-csm1, bcc-csm1_1), while one set of models overpredicts the seasonal cycle at mid latitudes, but underpredicts in the tropics (e.g. MIROC models). Another set of models underpredict the seasonal cycle across

ESDD
6, 761–818, 2015

all latitudes, but especially the tropics and high-latitudes (e.g. HadGEM2 models). The spatially averaged correlations between the seasonal cycle in the observations and models show a range of between 0.2 to 0.58 (Table 4), suggesting the need for substantial improvement in the timing of the seasonal cycle. Of course, there is not a strong seasonal cycle in the tropics, where the lowest correlations tend to occur (Table 4a). A smaller seasonal signal in this region could lead to larger relative errors in the observational estimates, and the smaller seasonal signal could be harder for models to simulate accurately. In mid-latitudes, where the seasonal cycle is likely to be more robust, the correlation coefficient averages above 0.5 for most models, except for the GFDL models and the INMCM4 (Table 4b).

The interannual variability tends to be overpredicted in some of the models (e.g. bcc-csm1, bcc-csm1_1, BNU-ESM, CESM1-BGC, GFDL-ESM2G, GFDL-ESM2M, MIROC-ESM, MIROC-ESM_CHEM) (Fig. 2c, f, i, and l; Table 4). For this calculation, the interannual variability (IAV) is calculated as the SD across the years of the annual average. The models tend to overpredict the IAV magnitude in the tropics more than other locations (Fig. 2c, f, i and l; Table 4).

We also consider the sensitivity of CESM models to different meteorological forcing data. Simulations by the land model within the CESM, but driven by observational-based meteorology, are very similar in the mean, seasonal cycle and interannual variability strength to those using meteorology simulated within the earth system model (Fig. 2 and Tables 3 and 4: CLMobs vs. CESM row of Table 2). This suggests that these metrics are more model dependent than meteorology dependent. Of course, with another model we might obtain a different result, but this result suggests that these model tests are not too dependent on the simulation of meteorology to be used in the ESM framework.

3.2 LAI–climate relationships

Next we explore the observed and model relationship between LAI and climate variables on the interannual time scale. Our goal is to assess whether the models can

ESDD
6, 761–818, 2015

simulate the observed relationships between interannual variation in temperature and precipitation and the interannual variability in LAI. In addition, we also consider whether there is trend in the LAI (thus a correlation between advancing time and the LAI in the observations and the model). The observations suggest statistically significant relationships between interannual variability in LAI and temperature, precipitation, and time trends (Fig. 3). As previously shown, there are strong positive relationships between LAI and temperature in high latitudes (Fig. 3a; e.g. Anav et al., 2013; Ichii et al., 2002; Zeng et al., 2013). In the tropics ($< 30^{\circ}$), the relationship can be positive or negative but some regions tend towards a negative relationship (Fig. 3a). This is consistent with our understanding that many places in the tropics are close to the optimal growing temperature already, and increases may lead to reduced productivity (Lobell et al., 2011), although this also could be related to moisture stress (Fung et al., 2005). Precipitation patterns tend to show positive correlations in many regions (Fig. 3b), but with some high latitude regions exhibit a negative correlation of precipitation with LAI. These relationships highlight the regional nature of sensitivity to temperature or precipitation, as seen in previous studies (e.g. Anav et al., 2013; Ichii et al., 2002; Wang et al., 2013; Zeng et al., 2013). High latitude regions, furthermore, may depend on snowfall, which will be poorly captured by our precipitation compositing procedure.

Note that there are regions with substantial trends over time (1981–2010) in the LAI (Fig. 3c), especially in high latitudes in the Northern Hemisphere. This could be associated with the longer growing season due to warming (e.g. Lucht et al., 2002; Zeng et al., 2013). It is also possible that this trend is due to carbon dioxide fertilization effects (e.g. Friedlingstein and Prentice, 2010). For high latitudes, we calculate a rank correlation of 0.58 across the models between the carbon dioxide fertilization factor on land for the Earth system models (called the beta land factor in Arora et al., 2013) and the average correlation of observed LAI with time, suggesting that there may be a component of carbon dioxide fertilization in the models' temporal trends. These trends are stronger in the models than the obsevations, which may be related to an overestimate of the fertilization effect.

ESDD
6, 761–818, 2015

Notice that while considering LAI interannual variability correlations with temperature, precipitation or time, that there are also strong correlations between temperature, precipitation and time themselves (e.g. IPCC, 2007). Here we do not attempt to differentiate these signals because of the statistical complexity and the shortness of the time record. There is also the possibility that the shortness of the record considered here will lead to aliasing of the real variability, especially in regions like the Sahel with strong decadal scale variations (e.g. Loew, 2014).

Model simulations can capture many of these relationships (e.g. Fig. 4 vs. Fig. 3), but with varying strengths (Table 3; Fig. 5). Because the coupled models are not intended to predict specific events or decadal variability, we want to evaluate the broad pattern of these relationships, instead of specific details. Thus, we consider the spatial mean of the temporal correlation between LAI and climate variables in the observations and model (as described in Sect. 2.3; Tables 3 and 4; Fig. 5). This tests, for example, whether the models capture the mainly positive correlation between LAI and temperature in high latitudes and mixed but more negative correlation in the tropics (Figs. 3a, 4a and 5).

Notice that for one land model (the CLM), simulated interactively within an earth system model (CESM-BGC) (Fig. 4a) presents more similar results of the LAI–temperature correlation to the CLM driven by observed-based datasets (CLM-obs) (Fig. 4b) than to either the observations (Fig. 3a), or other models (Fig. 4d or e). Thus, LAI correlations with temperature indicate a metric that appears to be intrinsic to the CLM model. Both the CLM and CLM-obs simulations (Fig. 4a and b) exhibited stronger negative LAI–temperature correlations over the tropics than seen in the observations (Fig. 3a). In general, almost all of the models exhibited a much stronger negative correlation between temperature and LAI in the tropics (Table 4a; Fig. 5) than that observed.

The relationship between CESM and CESM-obs LAI–precipitation relationships (e.g. Fig. 4c and d compared to Fig. 4a and b) are weaker. In fact, LAI–precipitation relationships do not appear to be more similar when the CLM within the CESM is compared against the CLM driven by observed-meteorology (Tables 3 and 4), suggesting

ESDD
6, 761–818, 2015

that LAI–precipitation relationships are very sensitive to the meteorology used, and not a good metric to be tested in a coupled earth system environment. This could be due to the fact that precipitation has a weaker relationship with LAI, except in an occurrence of rare but strong drought (e.g. Funk and Brown, 2006). Or this could be due to more complicated time lags that need to be considered or because random chance becomes too important. Growth in many tropical regions may be radiation limited, rather than water limited. The LAI–precipitation correlation metric may be more useful for inter-model comparison when used for offline-model tests when observed meteorology is used to force the models (e.g. Murray-Tortarolo et al., 2013).

Most of the models have too strong a negative relationship between LAI and temperature in the tropics, and too strong of a positive relationship in the high latitudes (Fig. 5, Table 4a–c). In the tropics, only the BNU-ESM model does not have too strong of a negative impact of temperature, while in the high latitudes, the CanESM2, HadGEM2-CC, HadGEM2-ES, MPI-ESM-MR tend to have twice the spatial average correlation of the observations.

For precipitation, there is a less clear relationship across latitudes, although one model has a tendency towards higher correlations with precipitation (INM-CM4; Figs. 4f and 5). Similarly, the correlation between LAI and time is variable between models, but there is no strong relationship across latitudes (Fig. 5c, f, i and l). A summary of the ability of the models to capture these metrics suggests that all the models could be improved in their simulation of LAI, but that many of the models are roughly doing a similarly good job at simulating LAI, depending on which metric is considered.

Note that there is measurement noise in the observational datasets, while the model values do not have an equivalent random noise added. We expect the measurement noise to reduce the correlations of LAI with the environmental variables in the observations relative to the true values, as seen compared to many models (Fig. 5). Thus, our metrics for interannual variability are likely to be more impacted by uncertainty in the observations than for the annual mean or seasonal cycle, and thus they may be less useful for evaluation of the models, although potentially interesting.

ESDD
6, 761–818, 2015

Figure 6 summarizes our comparisons of the models to the observations for LAI for the different metrics in Tables 2 and 4. Note that in order to show both correlations and model mean biases in the same figure, we have converted the model-data comparisons into Model Evaluation Values using Eq. (1) in Sect. 2.4, where 1 is a perfect model simulation and lower values represent worse model simulations. Overall none of the models do a perfect job, and improved simulation of LAI for all models is important. In addition, as discussed above, some models perform better in some regions than others. In order to more easily see how the models compare, we also show the ranking of the different models in each region (Table 5). For this comparison, we exclude the magnitude of the IAV, because the observational estimates for this are more likely to be in error than for the annual mean and seasonal analysis. In the tropics the top three models are the INMCM4, the IPSL-CM5A-LR and the IPSL-CM5B-LR. For the mid-latitudes the top models are the CanESM2, IPSL-CM5A-MR and the HADGEM2-ES. For high-latitudes the top models are the BNU-ESM, bcc-csm1 and the MIROC-ESM_CHEM (Table 5; Fig. 6).

4 Future projections

4.1 East African example

Previous studies have suggested increases in biomass and carbon in East Africa, because of wetter conditions, despite increases in temperature (Doherty et al., 2010). An example set of model time series of temperature, precipitation, net primary production and leaf area index from 1900–2100 shows that, similar to previous modeling studies (Doherty et al., 2010; Meehl et al., 2007), there is a mean increase in precipitation, as well as leaf area index in this region (Fig. 7). However, one model (IPSL-CMA-LR) shows an increase LAI variability (Fig. 7a), which could have large negative impacts to the local population despite an increased mean LAI. As this simple example suggests, for studies on the impact of climate change, we should look not only at the mean

ESDD
6, 761–818, 2015

change in leaf area index, but also look at changes in the variability of leaf area index, in order to identify regions which may be at risk for famine in the future.

The projections for future LAI are quite variable across different models for this region (Fig. 8), although generally quite optimistic in most of the CMIP5 models. For example, one model (BNU-ESM) predicts very large increases in LAI (> 4; Fig. 8a), while another (MPI-ESM) predicts modest increases and decreases (< 0.5; Fig. 8b). Normalizing by the SD ensures we only interpret results that are more than one SD from the current climate and statistically significant (e.g. Tebaldi et al., 2011), shows similar patterns. Notice that if the SD in the current climate is zero (i.e. there is not interannual variability reported at this grid box), the normalized difference is not finite, and is removed from further analysis. Some models are quite optimistic in East Africa, while others are less so (Fig. 8c vs. Fig. 8d).

Usually we consider the multi-model mean for future projections (e.g. Meehl et al., 2007). Model mean climate projections for the next century suggest statistically significant increases in LAI (the mean change divided by current variability) over most of East Africa (Fig. 9a, c and e). But some regions see a reduction in mean LAI after the mid-century. If we consider also the possibility of an increased mean, and also increased variability, which may indicate more frequent drought, we see that broad areas of East Africa are at a higher risk for drought by 2090 (Fig. 9e), despite a higher mean LAI. Here we define drought as the percent of time LAI is one or more SDs below the mean (as defined in the current climate), and thus any non-gray colored area indicates a higher drought risk than in the current climate (Fig. 9b, d, and f). While the areas with increased drought tend to be in regions with reductions in mean LAI, or smaller increases in LAI, the projections for increased drought show large regions at risk and thus maybe a more conservative metric for future vulnerability studies.

To consider the question of whether models project an increased risk of drought in East Africa in the future, one must also keep in mind that there are larger uncertainties in projections at smaller scales (e.g. Hawkins and Sutton, 2009), and thus one should not believe that in Kenya, for example, there will be an increase in LAI, while neigh-

ESDD
6, 761–818, 2015

boring countries will necessarily see a decrease. The ability of the models to resolve and project at such small scales is not strong enough (e.g. Hawkins and Sutton, 2009). Broader scale patterns considered in the next section are likely to be more robust. In addition, the projection of precipitation estimated in climate models for this region, which tend to be optimistic, is quite different than statistical studies which suggest less precipitation in a warming climate (e.g. Jury and Funk, 2013). Many of the observed drying trends in this region are linked to the sea surface temperature gradient between the equatorial western and central Pacific (Funk et al., 2014). While this gradient has strengthened, causing an intensification of the Walker circulation and drying, the future state of this gradient is uncertain.

4.2 Global projections

At the global level, there is also variability in the projections of future LAI changes. Some models project large increases while other project more modest increases (e.g. Fig. 10a and b; notice the different scale). After normalizing by the SD to highlight the results that are statistically significant (Fig. 10c and d), we still see large variations in the projections, especially in the tropics.

The 21st century projections of mean model LAI (normalized by the SD in each model) suggest a statistically significant increase in leaf area index over much of the globe, especially high latitudes (Fig. 11). Some tropical regions are seen to be at risk for reductions in mean LAI, such as in Central America and the Amazon basin. These regions are also at risk of more frequent drought, as identified by the percent of the time their LAI is below one SD of the current mean (Fig. 11b, d and f). More frequent drought is also projected for large areas of the tropics and subtropics where projected increases to mean LAI are small in magnitude or negligible (Fig. 11a vs. Fig. 11b, for example).

Models vary in how much change they project in the future (Fig. 12). The model projections tend to have larger increases than decreases in absolute magnitude of the

ESDD
6, 761–818, 2015

LAI (Fig. 12), and some models project very large increases, while there are only very small decreases predicted in a small number of cases.

4.3 Drivers of LAI projections

Next we consider what drives the differences in model projections for LAI, using the example of RCP8.5 at 2080–2100. Notice that in this study we neglect the possibly important process of direct human land management, which could dominate the changes in the land surface in some regions. By correlating at each grid box, across the models, the temperature and LAI projections, we can look for relationships between model projections of temperature and LAI, which may be causal (Fig. 13a). There are strong positive correlations between model simulated changes in temperature and LAI in some regions, especially the northern high latitudes, suggesting that models with a projected larger warming in the high latitudes also simulate larger increases in LAI. On the other hand, there are strong negative correlations in the tropics, for example the Amazon (Fig. 13a), suggesting that models that simulate higher tropical temperature changes tend to have lower LAI projections in the future. Notice that while higher temperatures may drive higher LAI, higher LAIs may also be driving higher temperatures because of the importance of LAI in changing surface energy fluxes (e.g. Lawrence and Slingo, 2004).

The projected changes in precipitation are strongly correlated with projected changes in LAI, when we correlate across models (Fig. 14a), suggesting that changes in precipitation across the models drive much of the difference between models in many regions. In addition, if we look spatially at where the lower LAIs occur, it is where the precipitation has decreased. If a region has a model mean lower precipitation in the in the future (Fig. 13c), it also has lower LAI predicted by the model mean (Fig. 11e).

Another important potential contributor to the future projections of LAI is the effectiveness of the carbon fertilization in the models (e.g. Arora et al., 2013). Using the carbon dioxide fertilization factor (β-land) from the Arora et al. (2013) study we use a rank correlation to explore the importance of the carbon dioxide fertilization strength

ESDD
6, 761–818, 2015

for predicting future LAI across the models. We would expect models that respond more strongly with increased carbon uptake under higher CO_2 conditions (i.e. larger β-land) to have greater LAI in the future. Globally the correlation is 0.34, suggesting that some of the differences in future LAI projections across models is due to differences in the model simulation of CO_2 fertilization. The value is −0.36, 0.26 and 0.58, for tropical, mid-latitude and high latitude, regions, respectively. Thus for high latitudes, especially, the projections of LAI appear to be dependent on the way the models' simulate the carbon dioxide fertilization in the different models. This could also be, however, an artifact that the two models with the lowest carbon dioxide effect (CESM-BGC and NOR-ESM) use the same land carbon model, which predicts low values of LAI in high latitudes for present day and does not tend to increase LAI much in the future (Thornton et al., 2009). These models also have low carbon dioxide fertilization effects, because of their nitrogen colimitation, which could be driving the correlation between model projections of LAI and carbon dioxide fertilization in the high latitudes. It is interesting that in the tropics the carbon dioxide fertilization is negatively correlated to future LAI changes. Again, this could be an artifact of having only two related low carbon fertilization models, as these models see a strong increase in nitrogen mineralization in the tropics in a warming climate, which allows an increase in productivity in the future tropics (Thornton et al., 2009). In other words, the strong negative correlation in the tropics between LAI projections and CO_2 fertilization could be due to the smaller temperature impact on carbon cycle in the N-limited models (the β-land and γ-land (climate impact on carbon cycle) are negatively correlated in Table 2 of Arora et al., 2013).

Finally, the disconnect between carbon dioxide fertilization effect and future LAI in the mid-latitudes and tropics could also be due to the way that carbon is allocated among different biomass pools in models. For example, in the CLM, the land model for the CESM-BGC, CO_2 fertilization causes a larger increase to wood allocation (62 %) than to leaf allocation (21 %) in the Southeastern US (D. Lombardozzi, personal communication, 2015). Thus, the issue of how LAI responds in different models is interesting and should be considered in future studies.

ESDD
6, 761–818, 2015

4.4 Reducing spread in the future projections

There are large differences between the different models' projections of future LAI (e.g. Figs. 12 and 14c). Previous studies have tried to reduce the uncertainty in future projections by looking for relationships between model metrics and future projections, and then choosing the models which best match the observations in the current climate (e.g. Cox et al., 2013; Hoffman et al., 2014) or by subsampling models for different regions by their performance (e.g. Steinacher et al., 2010). In this section we use both approaches to try to reduce the spread in projections at the end of the 21st century (2081–2100). In essence, we are looking for a correspondence between current model performance and the future projection, in order to reduce the uncertainty in the future projections. In many cases in climate modeling and projections, there is no correlation between current climate skill and projections (e.g. Cook and Vizy, 2006), however in some limited cases there is a correlation between metric score and a projection, and one is able to constrain future projections (e.g. Cox et al., 2013; Steinacher et al., 2010). Here we consider whether such a case applies here.

Across broad regions, we evaluate which metrics are the most useful for potentially constraining future climate projections by considering how the metric is correlated with the projections (Tables 2–4). We consider 4 regions: the globe, tropics (latitudes $< 30^{\circ}$), mid-latitudes (latitudes between 30 and 60°), and high latitudes (latitudes $> 60^{\circ}$). For the first approach, we look for the metrics that have the highest correlation coefficient to constrain the future estimate (similar to Cox et al., 2013) (Tables 2 and 4; Fig. 15a and b). For the globe these metrics are the average LAI vs. date correlation, and the global mean LAI ratio of model to observation. This analysis suggests that models with the largest relative drift in LAI over the last 30 years, will have the largest change in LAI in the future (Fig. 15a; Table 3). It also suggests that models with the highest LAI now, will have a bigger change in the future (Fig. 15b; Table 3). In Fig. 15a and b, the observational based estimates are indicated by the gray line. Notice that the value that the models match best with the observations are different for different metrics, and thus

ESDD
6, 761–818, 2015

does not allow us to uniquely constrain the future projections. Also notice that there is one model with a very large change in LAI in the future, which can drive much of the correlation. We use rank correlations instead of simple correlations, however, so that this result is largely insensitive to the removal of one model.

For the tropical region, (as also seen in the global analysis) the drift with time (LAI correlation with date) and the mean model/observation have the largest correlations (Table 2; Fig. 15c and d). Again this suggests that the models with the largest drift now, continue that drift in the future, and that models with larger LAI will have a larger change in the future. Unfortunately, for the tropics as well as the globe, these two metrics would constrain our future projections to two different LAI values.

For mid-latitudes, the best correlation between a metric and future projections of LAI comes from the spatial correlation of LAI with precipitation (Fig. 15e and f). Unfortunately very few of these correlations are actually statistically significant (> 0.36), and this metric is sensitive to whether the meteorology driving the model is from the earth system model or the reanalysis dataset (see discussion in Sect. 4.1). The metric with the second highest correlation comes from the model predicted change in precipitation between the present and the future. Thus mid-latitude projections of LAI are difficult to constrain based on model metrics, but are sensitive to changes in precipitation (as seen also in Fig. 13b).

For high latitudes there are three metrics with similar correlation coefficients: the average temporal correlation in the seasonal cycle, the size of the interannual variability in LAI model/observations and the size of the seasonal cycle in LAI model observations. The value of this metric that is ideal will allow us to constrain the future projections (Fig. 15g, h and i). Unfortunately again, these three metrics suggest a different projected change in LAI when the observed value is used to identify the models that are most realistic (grey line in Fig. 15g, h and i).

Overall, this analysis of multiple metrics suggests that there is no single metric available that is the most important in all circumstances for improving our estimates for the

ESDD
6, 761–818, 2015

changes in LAI. Thus, deduction of one more probable future is not available to us in this case (as opposed to Cox et al., 2013, where only one metric is presented).

The second approach for reducing spread in the future projections follows the ideas of Steinacher et al. (2010). Here for each region, we chose the models that did the best job for several metrics (i.e. using the rankings in Table 5), instead of just one metric at a time (as above). For this study, we chose to use the top half of the models, based on their performance for each region (Table 5), so instead of including 18 models, we include 9 models for each region. Using this approach does change the mean future projections, especially for the tropics and high latitudes (Table 6; Fig. 14), and does reduce the spread in the model values in the tropical region, but does not reduce the mean spread in mid-latitudes or high latitudes (Table 6). In the tropics, the top models tend to have lower future projections of LAI than the average of all the models (0.07 instead of 0.16 $m^2 m^{-2}$). This is actually consistent with the analysis in Fig. 15, since the models with the higher skill (close to grey line) would tend to have a lower or medium values of future LAI projection (Fig. 15c and d). For the mid-latitudes, there is not much difference between using all models or the top performing models (Table 6), while for high latitudes, the top models tend to project higher LAI in the future, also consistent with Fig. 15g, h, and i, where the observations tend to suggest higher LAI projections are more consistent for the metrics with the highest correlation.

The spatial distribution of the change in the future projections using the all models vs. the top models (Fig. 14a vs. Fig. 14b) is consistent with the mean over the regions, with the largest change being seen across the tropics, with a reduction in both the mean LAI projection as well as the SD. The changes in mid-latitudes and high latitudes from subsampling only the top performing models are not very large in most locations (Fig. 14a vs. Fig. 14b). Only in the tropics is the spread in the models reduced in the future projections (Fig. 14c vs. Fig. 14d). And the percent drought in the future is increased in the tropics, because of the sub-sampling of the top models (Fig. 14e vs. Fig. 14f).

ESDD
6, 761–818, 2015

Our analysis suggests that using multiple metrics does provide information that allows us in some cases (especially the tropics) to change our mean future projection, and reduce the spread between models predictions.

5 Summary and conclusions

LAI is a key variable in ESM simulations because it is the primary way that leaf-level biogeophysical and biogeochemical processes are scaled to regional and global areas. While previous studies focused on evaluating and comparing land carbon modules in the CMIP5 have largely considered the carbon cycle, here we expand the study of Anav et al. (2013) to look at LAI in all latitudes as well as future projections. We consider this the first study of future projections of LAI in the ESMs, and encourage future work in this area because of the importance of LAI in driving the physical and biogeochemical interchanges between land and atmosphere.

We examined simulations of LAI in current and future climates and current climate simulations are compared to available observations. Here we use the newly developed 30 year LAI satellite dataset (Zhu et al., 2013) and the CMIP5 archive to answer the following questions: (1) are the models accurately simulating the mean LAI spatial distribution? (2) Are the models accurately simulating the seasonal cycle in LAI? (3) Are the models correctly simulating the processes driving interannual variability in the current climate? And finally based on this analysis, (4) can we reduce the uncertainty in future projections of LAI by using each model's skill in the current climate? Many metrics that could be used in future model intercomparisons (e.g. Randerson et al., 2009; Luo et al., 2012) are considered (Table 2). Previous satellite derived LAI data were thought to be biased low, but the latest version resolves many of these issues (Zhu et al., 2013). Since LAI is one of the few land variables that is directly observed across the globe, it is appropriate for large scale climate model evaluation (e.g. Gleckler et al., 2008; Luo et al., 2012).

ESDD
6, 761–818, 2015

Models are able to simulate many, but not all the features of the observed mean, seasonal cycle and interannual variability LAI. The models tend to have too large values of mean, seasonal cycle strength and interannual variability strength in LAI, relative to the satellites (similar to previous studies e.g. Anav et al., 2013). Correlations with the seasonal cycle timing, as well as the spatial distribution of mean and the strength of the interannual variability suggest models could be improved.

We also considered the correlation of the interannual variability in the LAI with temperature, precipitation and date, and these suggest that the models could improve their simulation of these responses, and tend to have too strong a correlation (either negative or positive) with temperature. The modeled LAI trends were much stronger than the observed LAI trends in high latitudes, potentially indicating too strong of CO_2 fertilization effect. The evaluation of interannual variability relationships between LAI and climate variables is likely to be the most sensitive to observational error, however, and need to be interpreted with caution.

Over much of the globe in the future there is an increase in mean LAI, except where there is predicted a mean decrease in precipitation (Figs. 11c and 13c). But how large the increase is varies substantially by model (Fig. 12).

Using the example of East Africa, we propose that not only mean LAI, but also the variability in LAI is important in identifying vulnerability in future projections. One way to calculate this is to calculate the percentage of the time in which the LAI is below one SD of the mean from the current climate, which we define as drought conditions. This increase in drought frequency indicates more at-risk regions despite higher mean LAIs being projected. The most notable future change is the increase in drought conditions in parts of the tropics (Fig. 14e and f), suggesting a higher vulnerability for food security in these regions.

We use two different methods for reducing the large spread in future projections, and find that combining multiple metrics to chose better models (e.g. similar to Steinacher et al., 2010) seems to work more robustly than simply correlating one metric against future projections (e.g. Cox et al., 2013; Hoffman et al., 2014), because the differ-

ESDD
6, 761–818, 2015

ent metrics suggest different values (Fig. 15). Overall, our method suggests that the top performing models (top half of the models from Table 6) suggest a smaller future increase in LAI in the tropics, and more regions with future reductions in LAI than assessments based on including all the models. This approach also reduces the spread between models in the tropics. Using only the top models, however, did not make such a large difference in projections in the mid- and high latitudes (Fig. 14). Our results suggest that the better performing models tend to project lower LAIs in the future in the tropics in contrast to Cox et al. (2013), which focused on carbon–temperature relationships in the Amazon and which showed that observational constraints on the models tend to suggest less loss in carbon under higher temperatures. However these results may not be inconsistent as they consider different metrics in different regions, and LAI is not necessarily linearly related to carbon uptake in the models (see discussion in Sect. 4.2; D. Lombardozzi, personal communication, 2015), suggesting that more analysis of how allocation is parameterized in the land carbon models is warranted.

Finally, the spread between models' projections of LAI was correlated with model's projections of precipitation (Fig. 14b and c vs. Fig. 12c). Thus our projections of LAI rest finally on the ability to project future precipitation, which in many regions the changes in precipitation are not large enough to be statistically significantly outside natural variability (e.g. Tebaldi et al., 2011) and there are discrepancies between climate model and statistical model predictions (e.g. Jury; Funk, 2013). In addition, increasing temperatures are likely to stress systems, even if there is additional rainfall (e.g. Lobell et al., 2011), expanding the regions at risk to increased drought (Fig. 11).

Acknowledgements. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We acknowledge NSF-0832782 and 1049033 and assistance from C. Barrett and S. Schlunegger. We acknowledge the assistance of the LAI development group for making the LAI 3g product available, and

ESDD
6, 761–818, 2015

the NOAA/OAR/ESRL PSD group for making the GPCP and GHCN gridded products available online at http://www.esrl.noaa.gov/psd/. This work was made possible, in part, by support provided by the US Agency for International Development (USAID) Agreement No. LAG-A-00-96-90016-00 through Broadening Access and Strengthening Input Market Systems Collaborative Research Support Program (BASIS AMA CRSP). All views, interpretations, recommendations, and conclusions expressed in this paper are those of the authors and not necessarily those of the supporting or cooperating institutions.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., and Arkin, P.: The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), J. Hydrometeorol., 4, 1147–1167, 2003.

Anav, A., Murray-Tortarolo, G., Friedlingstein, P., Stich, S., Piao, S., and Zhu, Z.: Evaluation of land surface models in reproducing satellite derived leaf area index over the high latitude-Northern Hemisphere. Part II: Earth system models, Remote Sensing, 5, 3637–3661, 2013.

Arora, V. K., Scinocca, J., Boer, G. J., Christian, J., Denman, K. L., Flato, G., Kharin, V., Lee, W., and Merryfield, W.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophys. Res. Lett., 38, L05805, doi:10.1029/2010GL046270, 2011.

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C., Christian, J., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon–concentration and carbon–climate feedbacks in CMIP5 earth system models, J. Climate, 26, 5289–5314, 2013.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjánsson, J. E.: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate, Geosci. Model Dev., 6, 687–720, doi:10.5194/gmd-6-687-2013, 2013.

Bounoua, L., Collatz, G., Los, S. O., Sellers, P., Dazlich, D., Tucker, C., and Randall, D.: Sensitivty of climate to changes in NDVI, J. Climate, 13, 2277–2292, 2000.

ESDD
6, 761–818, 2015

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, Geosci. Model Dev., 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Cook, K. and Vizy, E.: Coupled model simulations of the West African Monsoon System: twentieth- and twenty-first-century simulations, J. Climate, 19, 3681–3703, 2006.

Cox, P., Pearson, D., Booth, B., Friedlingstein, P., Huntingford, C., Jones, C., and Luke, C.: Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability, Nature, 494, 341–344, 2013.

Cramer, W., Kicklighter, D. W., Bondeau, A., Moore, B., Churkina, G., Nemry, B., Ruimy, A., and Schloss, A. E. A.: Comparing global models of terrresrial net primary production (NPP): overview and key results, Glob. Change Biol., 5, 1–15, 1999.

Dai, A., Fung, I. Y., and Del Genio, A. D.: Surface observed global land precipitation variations during 1900–88, J. Climate, 10, 2943–2962, 1997.

Doherty, R., Stich, S., Smith, B., Lewis, S., and Thornton, P.: Implications of future climate and atmospheric CO_2 content for regional biogeochemistry, biogeography and ecosystem services across East Africa, Glob. Change Biol., 16, 617–640, doi:10.1111/j.1365-2486.2009.01997.x, 2010.

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P., Ethe', C., Fairhead, L., Fichefet, T., Flavoni, S., Freidlingstein, P., Lefebvre, M., Lefrevre, F., Levy, C., Li, Z., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Merurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth system modl: from CMIP3 to CMIP5, Clim. Dynam., 40, 2123–2165, 2013.

Dunne, J., John, J., Sheviliakova, E., Stouffer, R. J., Krasting, J., Malyshev, S., Milly, P., Sentman, L., Adcroft, A., Cooke, W., Dunne, K., Harrison, M., Krasting, J., Malyshev, S., Milly, P., Phillips, P., Sentman, L., Samuels, B., Spelman, M., Winton, M., Wittenberg, A., and Zadeh, N.: GFDL's ESM2 global cupoled climate-carbon Earth system models. Part II: Carbon system formation and baseline simulation characteristics, J. Climate, 26, 2247–2267, 2013.

ESDD
6, 761–818, 2015

Fan, Y. and Dool, V. D.: A global monthly land surface air temperature analysis for 1948–present, J. Geophys. Res., 113, D01103, doi:10.1029/2007JD008470, 2008.
Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M., Neigh, C., and Reichstein, M.: Trend change detection in NDVI time series: effects of inter-annual variability and methodology, Remote Sensing, 5, 2113–2144, doi:10.3390/rs5052113, 2013.
Friedlingstein, P. and Prentice, I. C.: Carbon–climate feedbacks: a review of model and observation based estimates, Curr. Opinion Environ. Sustain., 2, 251–257, 2010.
Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Bloh, W. v., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Mathews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnurr, R., Strassmen, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate–carbon cycle feedback analysis, results from the C4MIP Model intercomparison, J. Climate, 19, 3337–3353, 2006.
Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, A., Anav, A., Liddicoat, S., and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, J. Climate, 27, 511–526, doi:10.1175/JCLI-D-1112-00579.00571, 2013.
Fung, I., Doney, S., Lindsay, K., and John, J.: Evolution of carbon sinks in a changing climate, P. Natl. Acad. Sci. USA, 102, 11201–11206, 2005.
Funk, C. and Brown, M.: Intra-seasonal NDVI change projections in semi-arid Africa, Remote Sens. Environ., 101, 249–256, 2006.
Funk, C. and Budde, M.: Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe, Remote Sens. Environ., 113, 115–125, 2009.
Funk, C., Hoell, A., Shukla, S., Bladé, I., Liebmann, B., Roberts, J. B., Robertson, F. R., and Husak, G.: Predicting East African spring droughts using Pacific and Indian Ocean sea surface temperature indices, Hydrol. Earth Syst. Sci., 18, 4965–4978, doi:10.5194/hess-18-4965-2014, 2014.
Ganzeveld, L., Lelieveld, J., and Roelofs, G.-J.: A dry deposition parameterization for sulfur oxides in a chemistry and general circulation model, J. Geophys. Res., 103, 5679–5694, 1998.
Gleckler, P., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, J. Geophys. Res., 113, D06104, doi:10.1029/2007JD008972, 2008.
Groten, S.: NDVI-crop monitoring and early yield assessment of Burkino Faso, Int. J. Remote Sens., 14, 1495–1515, 1993.

ESDD
6, 761–818, 2015

Harris, I., Jones, P., Osborn, T., and Lister, D.: Updated high-resolution grids of monthly climatic obsevations-the CRU TS3.10 dataset, Int. J. Climatol., 34, 623–642, doi:10.1002/joc.3711, 2013.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, B. Am. Meteorol. Soc., 90, 1095–1107, doi:10.1175/2009BAMS2607.1171, 2009.

Hoffman, F., Randerson, J., Arora, V. K., Bao, Q., Cadule, P., Ji, D., Jones, C., Kawamiya, M., Khatiwala, S., Lindsay, K., Obata, A., Sheviliakova, E., Six, K., Tjiputra, J. F., Volodin, E., and Wu, T.: Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models, J. Geophys. Res.-Biogeo., 119, 141–162, doi:10.1002/2013JG002381, 2014.

Ichii, K., Kawabata, A., and Yamaguchi, Y.: Global correlation analysis for NDVI and climatic variables and NDVI trends: 1982–1990, Int. J. Remote Sens., 23, 3873–3878, 2002.

IPCC: Summary for policymakers, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Avery, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, UK, New York, NY, USA, 2007.

Jones, C., Robertson, E., Arora, V. K., Friedlingstein, P., Sheviliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., Kawamiya, M., Liddicoat, S., Lindsay, K., Reick, C., Roelandt, C., Segschneider, J., and Tjiputra, J. F.: 21st Century compatible CO_2 emissions and airborne fraction simulated by CMIP5 Earth System models under 4 representative concentration pathways., J. Climate, 26, 4398–4413, doi:10.1175/JCLI-D-12-00554.1, 2013.

Jones, P., Osborn, T., and Briffa, K.: Estimating sampling errors in large-scale temperature averages, J. Climate, 10, 2548–2568, 1997.

Jong, R., Verbesselt, J., Zeileis, A., and Schaepman, M.: Shifts in global vegetation activity trends, Remote Sensing, 5, 1117–1133, doi:10.3390/rs5031117, 2013.

Jury, M. and Funk, C.: Climate trends over Ethiopia: regional signals and drivers, Int. J. Climatol., 33, 1924–1935, 2013.

Lawrence, D. and Slingo, J.: An annual cycle of vegetatio in a GCM. Part I: Implementation and impact on evaporation, Clim. Dynam., 22, 87–105, 2004.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Fletcher, C. G., Lawrence, P. J., Levis, S., Swenson, C., and Bonan, G. B.: The CCSM4 land simulation, 1850–2005: assessment of surface climate and new capabilities, J. Climate, 25, 2240–2260, 2012.

ESDD
6, 761–818, 2015

Lindsay, K., Bonan, G., Doney, S., Hofffman, F., Lawrence, D., Long, M. C., Mahowald, N., Moore, J. K., Randerson, J. T., and Thornton, P.: Preindustrial and 20th century experiments with the Earth System Model CESM1-(BGC), J. Climate, 27, 8981–9005, 2014.

Lobell, D., Schlenker, W., and Costa-Roberts, J.: Climate trends and global crop production since 1980, Science, 333, 616–620, 2011.

Loew, A.: Terrestrial satellite records for climate studies: how long is long enough? A test case for the Sahel, Theor. Appl. Climatol., 115, 427–440, doi:10.1007/s00704-00013-00880-00706, 2014.

Lombardozzi, D., Bonan, G., and Nychka, D.: The emerging anthropogenic signal in the land–atmosphere carbon cycle, Nat. Clim. Change, 4, 796–800, doi:10.1038/NCLIMATE2323, 2014.

Lucht, W., Prentice, I. C., Myneni, R., Stich, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W., and Smith, B.: Climate control of the high-latitude vegetation greening trend and the Pinatubo effect, Science, 296, 1687–1689, 2002.

Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J. B., Fisher, R., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., Jones, C. D., Koven, C., Lawrence, D., Li, D. J., Mahecha, M., Niu, S. L., Norby, R., Piao, S. L., Qi, X., Peylin, P., Prentice, I. C., Riley, W., Reichstein, M., Schwalm, C., Wang, Y. P., Xia, J. Y., Zaehle, S., and Zhou, X. H.: A framework for benchmarking land models, Biogeosciences, 9, 3857–3874, doi:10.5194/bg-9-3857-2012, 2012.

Mao, J., Shin, X., Thornton, P., Hoffman, F., Zhu, Z., and Myneni, R.: Global latitudinal-asymetric vegetation growth trends and their driving mechanisms: 1982–2009, Remote Sensing, 5, 1484–1497, doi:10.3390/rs5031484, 2013.

Maxino, C., McAvaney, B., Pitman, A., and Perkins, S.: Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation, Int. J. Climatol., 28, 1097–1112, 2008.

Meehl, G., Stocker, T., Collins, W., Friedlingstein, P., Gaye, A., Gregory, J. M., Kitoh, A., Knutti, R., Murphy, J., Noda, A., Raper, S., Watterson, I., Weaver, A., and Zhao, Z.-C.: Global climate projections, in: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Avert, K., Tignor, M., Miller, H., Cambridge University Press, Cambridge, UK, 2007.

ESDD
6, 761–818, 2015

Murray-Tortarolo, G., Anav, A., Friedlingstein, P., Stich, S., Piao, S., Zhu, Z., Poulter, B., Zaehle, S., Alhstrom, A., Lomas, M., Levis, S., Viovy, N., and Zeng, N.: Evaluation of land surface models in reproducing satellite-derived LAI over the high-latitude Northern Hemisphere. Part I: Uncoupled DGVMs, Remote Sensing, 5, 4819–4838, doi:10.3390/rs5104819, 2013.

Oleson, K., Lawrence, D., Bonan, G., Drewniak, B., Huang, M., Koven, C., Levis, S., Li, F., Riley, W., Subin, Z., Swensen, S., Thornton, P., Bozbiyik, A., Fisher, R., Kluzek, E., Lamarque, J. F., Lawrence, P., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D., Sacks, W., Sun, Y., Tang, J., and Yang, Z.-L.: Technical Description of version4.5 of the Community Land Model (CLM), NCAR, Boulder, CO, 2013.

Qian, T., Dai, A., Trenberth, K., and Oleson, K.: Simluation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations, J. Hydrometeorol., 7, 953–975, 2006.

Raddatz, T., Reick, C. H., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclaus, J.: Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, Clim. Dynam., 29, 565–574, 2007.

Ramankutty, N., Evan, A., Monfreda, C., and Foley, J.: Farming the planet: the geographic distribution of global agricultural lands in the year 2000, Global Biogeochem. Cy., 22, BG1003, doi:10.1029/2007GB002952, 2008.

Randerson, J., Hoffman, F., Thornton, P., Mahowald, N., Lindsay, K., Lee, Y.-H., Nevison, C. D., Doney, S., Bonan, G., Stockli, R., Covey, C., Running, S., and Fung, I.: Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, Glob. Change Biol., 15, 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x, 2009.

Shao, P., Zeng, X., Sakaguchi, K., Monson, R., and Zeng, X.: Terrestrial carbon cycle: climate relations in eight CMIP5 Earth System Models, J. Climate, 26, 8744–8764, 2013.

Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Cocco, V., Doney, S. C., Gehlen, M., Lindsay, K., Moore, J. K., Schneider, B., and Segschneider, J.: Projected 21st century decrease in marine productivity: a multi-model analysis, Biogeosciences, 7, 979–1005, doi:10.5194/bg-7-979-2010, 2010.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, 2001.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: A summary of the CMIP5 Experimental Design, available at: http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf (last access: 8 April 2015), 2009.

ESDD
6, 761–818, 2015

Thornton, P. E., Doney, S. C., Lindsay, K., Moore, J. K., Mahowald, N., Randerson, J. T., Fung, I., Lamarque, J.-F., Feddema, J. J., and Lee, Y.-H.: Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, Biogeosciences, 6, 2099–2120, doi:10.5194/bg-6-2099-2009, 2009.

van Vuuren, D., Elzen, M. G. D., Lucas, P., Eickhout, B., Strengers, B., Ruijven, B. V., Wonink, S., and Houdt, R. V.: Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs, Climatic Change, 81, 119–159, doi:10.1007/s10584-10006-19172-10589, 2007.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G., Kram, T., Krey, V., Nakicenovic, N., Smith, S., and Rose, S.: The representative concentration pathways: an overview, Climatic Change, 109, 5–31, 2011.

Volodin, E., Dianskii, N., and Gusev, A.: Simulating present day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, Izv. Ocean. Atmos. Phy.+, 46, 414–431, 2010.

Vrieling, A., de Leeuw, J., and Said, M.: Length of growing period over Africa: variability and trends from 30 years of NDVI time series, Remote Sensing, 5, 982–1000, doi:10.3390/rs5020982, 2013.

Wang, W., Ciais, P., Nemani, R., Canadell, J., Piao, S., Stich, S., White, M., Hashimoto, H., Milesi, C., and Myneni, R.: Variations in atmospheric CO_2 growth rates coupled with tropical temperature, P. Natl. Acad. Sci. USA, 110, 13061–13066, doi:10.11073/pnas.1219683110, 2013.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and Kawamiya, M.: MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments, Geosci. Model Dev., 4, 845–872, doi:10.5194/gmd-4-845-2011, 2011.

Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., Wei, M., and Shi, X.: Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century, J. Geophys. Res., 118, 4326–4347, doi:10.1002/jgrd.50320, 2013.

Zeng, F.-W., Collatz, G., Pinzon, J., and Ivanoff, A.: Evaluating and quantifying the climate-driven interannual variability in Global Inventory Modeling and Mapping Studies (GIMMS) normalized difference vegetation index at global scales, Remote Sensing, 5, 3918–3950, doi:10.3390/rs508918, 2013.

ESDD
6, 761–818, 2015

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R., and Myneni, R.: Global data sets of vegetation leaf area index (LAI)3g and fraction of photosynthetically active radiation (FPAR)3g derived from global inventory modeling and mapping studies (GIMMS) Normalized difference vegetation index (NDVI3g) for the period 1981 to 2011, Remote Sensing, 5, 927–948, 2013.

ESDD
6, 761–818, 2015

Table 1. Model simulation from the Climate Modeling Intercomparison Projection (CMIP5) included in this study.

Model	Land Model	Land Resolution	N-Cycle	Dynamic Veg.	Citation
BCC-CSM1	BCC-AVIM1.	2.8° × 2.8°	N	Y	Wu et al. (2013)
BCC-CSM1-M	BCC-AVIM1.	1.1° × 1.1°	N	Y	Wu et al. (2013)
BNU-ESM	CoLM + BNU-DGVM	2.8° × 2.8°	N	Y	(BNU-ESM, http://esg.bnu.edu.cn/BNU_ESM_webs/htmls/index.html)
CanESM2	CLASS2.7 + CTEM1	2.8° × 2.8°	N	N	Arora et al. (2011)
CESM1-BGC	CLM4	0.9° × 1.2°	Y	N	Lindsay et al. (2014)
GFDL-ESM2G	LM3	2.5° × 2.5°	N	Y	Dunne et al. (2013)
GFDL-ESM2M	LM3 (uses different physical ocean model)	2.5° × 2.5°	N	Y	Dunne et al. (2013)
HadGEM2-CC	JULES + TRIFFID	1.9° × 1.2°	N	Y	Collins et al. (2011)
HadGEM2-ES	JULES + TRIFFID (includes chemistry)	1.9° × 1.2°	N	Y	Collins et al. (2011)
INM-CM4	Simple model	2° × 1.5°	N	N	Volodin et al. (2010)
IPSL-CM5A-LR	ORCHIDEE	3.7° × 1.9°	N	N	Dufresne et al. (2013)
IPSL-CM5A-MR	ORCHIDEE	2.5° × 1.2°	N	N	Dufresne et al. (2013)
IPSL-CM5B-LR	ORCHIDEE (improved parameterization)	3.7° × 1.9°	N	N	Dufresne et al. (2013)
MIROC-ESM_	MATSIRO + SEIB-DGVM	2.8° × 2.8°	N	Y	Watanabe et al. (2011)
MIROC-ESM-CHEM	MATSIRO + SEIB-DGVM (adds chemistry)	2.8° × 2.8°	N	Y	Watanabe et al. (2011)
MPI-ESM-LR	JSBACH + BETHY	1.9° × 1.9°	N	Y	Raddatz et al. (2007)
MPI-ESM-MR	JSBACH + BETHY (ocean model higher resolution)	1.9° × 1.9°	N	Y	Raddatz et al. (2007)
NorESM1-ME	CLM4	2.5° × 1.9°	Y	N	Bentsen et al. (2013)

ESDD
6, 761–818, 2015

Table 2. Table of Metrics for LAI comparisons between model and observation used in the following tables. More description of these metrics are provided in Sect. 2.3.

Metrics		Description
Mean	Model/obs	Ratio of mean LAI from the model and observations
	Corr.	Spatial correlation of Mean LAI
SD Seasonal	Model/obs	Ratio of seasonal cycle strength: Ratio of SD of the climatological monthly mean LAI from the model and observations
	Avg. Corr.	Avg. Corr. of the temporal evolution of the climatological seasonal cycle in the model vs. observations at each grid box
SD IAV	Model/obs	Ratio of IAV strength: ratio of SD of the annual mean LAI from the model and observations
IAV LAI vs. T	Avg. Corr.	Avg. Corr. between LAI and temperature in IAV
IAV LAI vs. P	Avg. Corr.	Avg. Corr. between LAI and precipitation in IAV
IAV LAI vs. date	Avg. Corr.	Avg. Corr. between LAI and date in IAV
Other variables	ΔT	Change in temperature (2081–2100 minus current)
	ΔP	Change in precipitation (2081–2100 minus current)

ESDD
6, 761–818, 2015

Table 3. Evaluation of LAI over globe. Metrics are described in text and Table 2, models in Table 1. The CESM vs. CLM column indicates the value of the comparison between the CESM1-BGC and the CLM-obs simulations (which use the same land model, but different meteorology). The Corr Δ LAI row indicates the correlation coefficient across models between the model value of this metric (this column) against the change in LAI in 2080–2100 (last column).

	Mean LAI		Seasonal		SD IAV	LAI IAV correlations			ΔT	Δprecip	ΔLAI
Models	Model/ obs	Corr.	SD Model/ obs	Avg. Corr.	Model/ obs	LAI vs. Ts	LAI vs. Precip	LAI vs. time	(K)	(mm day^{-1})	(m^2 m^{-2})
Obs.						0.11	0.11	0.21			
bcc-csm1	1.74	0.70	1.28	0.54	1.64	−0.07	0.34	0.26	5.13	0.18	0.29
bcc-csm1-1	1.52	0.67	1.27	0.55	1.43	−0.13	0.38	0.28	4.88	0.17	0.35
BNU-ESM	2.12	0.56	1.47	0.48	1.79	0.27	0.00	0.32	6.39	0.28	1.01
CanESM2	1.05	0.66	0.75	0.40	1.15	0.02	0.18	0.17	5.55	0.24	0.09
CESM1-BGC	1.49	0.64	0.70	0.48	1.86	0.00	0.13	0.24	5.32	0.23	0.32
GFDL-ESM2G	2.27	0.45	0.78	0.18	1.64	−0.06	0.21	0.29	2.95	0.10	0.19
GFDL-ESM2M	2.35	0.39	0.78	0.18	1.93	−0.13	0.20	0.28	3.19	0.12	0.20
HadGEM2-CC	1.44	0.76	0.58	0.46	0.92	0.15	0.27	0.32	6.77	0.17	0.48
HadGEM2-ES	1.52	0.77	0.58	0.46	1.00	0.17	0.31	0.35	6.88	0.17	0.44
inmcm4	0.97	0.61	0.93	0.42	0.86	0.05	0.53	0.04	4.02	0.12	0.16
IPSL-CM5A-LR	1.44	0.67	0.98	0.49	1.21	0.03	0.27	0.08	6.73	0.22	0.10
IPSL-CM5A-MR	1.44	0.68	0.97	0.50	1.21	0.04	0.28	0.14	6.52	0.20	0.08
IPSL-CM5B-LR	1.33	0.60	0.95	0.50	1.36	0.02	0.24	0.17	5.39	0.10	0.18
MIROC-ESM	1.64	0.44	1.17	0.56	3.23	−0.08	0.12	0.11	7.13	0.27	0.22
MIROC-ESM-CHEM	1.62	0.44	1.11	0.53	3.23	−0.05	0.12	0.14	7.55	0.29	0.22
MPI-ESM-LR	1.32	0.59	0.83	0.45	0.85	−0.04	0.18	0.14	5.11	0.10	0.13
MPI-ESM-MR	1.36	0.60	0.86	0.26	0.85	0.02	0.19	0.20	4.63	0.10	0.12
NorESM1-ME	1.61	0.54	0.82	0.44	2.50	−0.05	0.16	0.17	4.18	0.19	0.12
CLMobs	1.44	0.73	0.71	0.53	2.08	0.01	0.22	0.23			
CESMvs.CLM	1.08	0.89	0.98	0.76	1.18	0.00	0.13	0.24			
Corr Δ LAI	0.49	0.10	0.06	0.30	0.27	0.02	−0.04	0.67	0.24	0.16	

ESDD
6, 761–818, 2015

Table 4a. Tropical LAI evaluation and projection. As in Table 3, but for tropical region (< 30°).

Models	Mean LAI		Seasonal		SD IAV	LAI IAV correlations			ΔT	Δprecip	ΔLAI
	Model/ obs	Corr.	SD Model/ obs	Avg. Corr.	Model/ obs	LAI vs. Ts	LAI vs. Precip	LAI vs. time	(K)	(mm day^{-1})	(m^2 m^{-2})
Obs.						0.02	0.20	0.20			
bcc-csm1	1.69	0.82	1.79	0.24	2.63	−0.45	0.38	0.10	4.18	0.12	0.18
bcc-csm1-1	1.44	0.78	1.66	0.27	2.06	−0.44	0.41	0.13	4.36	0.05	0.34
BNU-ESM	2.45	0.63	0.85	0.17	1.25	0.22	−0.04	0.33	4.07	0.27	1.29
CanESM2	1.23	0.54	0.74	0.13	1.81	−0.40	0.34	0.05	4.07	0.10	0.06
CESM1-BGC	1.72	0.72	0.91	0.38	2.69	−0.29	0.17	0.14	4.13	0.27	0.57
GFDL-ESM2G	1.80	0.64	0.96	0.17	2.06	−0.37	0.22	0.04	2.79	0.09	0.22
GFDL-ESM2M	1.74	0.63	0.92	0.16	2.50	−0.39	0.24	0.09	2.81	0.10	0.21
HadGEM2-CC	1.71	0.81	0.47	0.29	0.88	−0.08	0.28	0.15	5.81	−0.03	0.25
HadGEM2-ES	1.76	0.81	0.47	0.28	0.94	−0.15	0.33	0.11	5.95	−0.01	0.21
inmcm4	1.00	0.83	0.83	0.36	0.69	−0.19	0.68	−0.02	3.44	0.00	−0.04
IPSL-CM5A-LR	1.21	0.80	1.09	0.36	1.38	−0.25	0.39	−0.07	5.90	0.26	−0.03
IPSL-CM5A-MR	1.20	0.75	1.09	0.35	1.44	−0.24	0.41	0.01	6.05	0.26	−0.02
IPSL-CM5B-LR	1.09	0.70	1.02	0.33	1.63	−0.19	0.36	0.05	4.22	−0.06	−0.01
MIROC-ESM	1.61	0.53	0.64	0.35	5.06	−0.37	0.14	0.01	5.89	0.13	−0.08
MIROC-ESM_CHEM-	1.61	0.53	0.65	0.33	5.00	−0.38	0.15	0.05	6.18	0.11	−0.14
MPI-ESM-LR	1.41	0.75	1.04	0.15	1.19	−0.51	0.46	−0.04	5.47	−0.07	−0.14
MPI-ESM-MR	1.42	0.75	1.06	0.02	1.13	−0.50	0.45	−0.09	4.96	−0.02	−0.12
NorESM1-ME	1.73	0.58	0.98	0.32	3.44	−0.34	0.24	0.15	2.80	0.16	0.16
CLMobs	1.64	0.82	0.90	0.45	2.88	−0.19	0.26	0.10			
CESMvs.CLM	1.09	0.85	1.02	0.68	1.13	−0.29	0.17	0.14			
Corr Δ LAI	0.64	0.11	−0.08	−0.06	0.02	0.37	−0.40	0.80	−0.43	0.26	

ESDD
6, 761–818, 2015

Table 4b. Mid-latitude LAI evaluation and projection. As in Table 3, but for mid-latitude region (between 30 and 50°).

Models	Mean LAI		Seasonal		SD IAV	LAI IAV correlations			ΔT	Δprecip	ΔLAI
	Model/ obs	Corr.	SD Model/ obs	Avg. Corr.	Model/ obs	LAI vs. Ts	LAI vs. Precip	LAI vs. time	(K)	(mm day^{-1})	(m^2 m^{-2})
Obs.						0.11	0.18	0.22			
bcc-csm1	1.90	0.60	1.33	0.75	1.50	−0.14	0.43	0.23	4.71	0.03	0.50
bcc-csm1-1	1.61	0.52	1.34	0.74	1.50	−0.18	0.50	0.31	4.72	0.07	0.62
BNU-ESM	2.20	0.48	1.80	0.61	2.80	0.22	0.02	0.07	5.76	0.01	0.80
CanESM2	0.90	0.75	0.88	0.56	1.00	0.03	0.17	0.12	4.34	0.12	0.11
CESM1-BGC	1.86	0.73	0.82	0.51	2.80	−0.03	0.17	0.09	4.69	0.15	0.55
GFDL-ESM2G	1.37	0.52	1.46	0.17	2.20	−0.07	0.25	0.22	2.45	−0.06	0.20
GFDL-ESM2M	1.36	0.42	1.37	0.20	2.40	−0.11	0.25	0.15	2.64	−0.04	0.24
HadGEM2-CC	1.20	0.68	0.86	0.56	1.10	0.02	0.31	0.19	6.40	0.13	0.63
HadGEM2-ES	1.21	0.69	0.88	0.57	1.20	0.06	0.33	0.23	6.29	0.11	0.62
inmcm4	1.28	0.70	0.83	0.33	1.20	0.15	0.57	0.07	3.79	0.01	0.17
IPSL-CM5A-LR	1.68	0.71	1.05	0.53	1.90	0.00	0.29	0.05	6.16	−0.04	0.11
IPSL-CM5A-MR	1.62	0.76	1.00	0.57	2.00	−0.04	0.35	0.09	6.23	−0.11	0.07
IPSL-CM5B-LR	1.74	0.69	1.05	0.58	1.80	0.07	0.27	0.21	4.85	0.07	0.24
MIROC-ESM	1.89	0.44	1.79	0.71	2.80	−0.09	0.19	0.08	7.30	0.14	0.31
MIROC-ESM_CHEM-	1.88	0.46	1.75	0.69	2.80	−0.09	0.19	0.08	7.48	0.16	0.34
MPI-ESM-LR	1.42	0.41	0.94	0.71	0.70	0.11	0.15	0.07	5.08	0.01	0.19
MPI-ESM-MR	1.47	0.38	0.94	0.52	0.70	0.16	0.11	0.15	4.58	0.03	0.20
NorESM1-ME	2.19	0.68	1.09	0.50	3.80	−0.01	0.19	0.18	3.32	0.07	0.15
CLMobs	1.69	0.82	0.80	0.57	2.80	−0.05	0.21	0.01			
CESMvs.CLM	1.18	0.88	1.02	0.76	1.17	−0.03	0.17	0.09			
Corr Δ LAI	0.16	−0.31	0.19	0.38	0.12	−0.11	−0.00	0.35	0.34	0.46	

ESDD
6, 761–818, 2015

Table 4c. High-latitude LAI evaluation and projections. As in Table 3, but for high-latitude region (> 50°).

Models	Mean LAI		Seasonal		SD IAV	LAI IAV correlations			ΔT	Δprecip	ΔLAI
	Model/ obs	Corr.	SD Model/ obs	Avg. Corr.	Model/ obs	LAI vs. Ts	LAI vs. Precip	LAI vs. time	(K)	(mm day^{-1})	(m^2 m^{-2})
Obs.						0.20	−0.02	0.21			
bcc-csm1	1.73	0.58	0.69	0.91	0.79	0.36	0.25	0.43	5.71	0.26	0.27
bcc-csm1-1	1.61	0.57	0.77	0.91	0.86	0.21	0.28	0.41	5.16	0.25	0.27
BNU-ESM	1.45	0.35	1.81	0.90	1.93	0.36	0.02	0.48	7.69	0.38	0.95
CanESM2	0.85	0.63	0.69	0.74	0.43	0.42	0.03	0.31	7.33	0.40	0.10
CESM1-BGC	0.84	0.48	0.35	0.69	0.50	0.31	0.06	0.44	6.01	0.23	0.14
GFDL-ESM2G	3.67	0.22	0.21	0.22	1.00	0.26	0.18	0.59	3.33	0.19	0.17
GFDL-ESM2M	4.03	0.21	0.23	0.24	1.14	0.11	0.11	0.54	3.75	0.22	0.18
HadGEM2-CC	1.13	0.61	0.39	0.70	0.85	0.48	0.23	0.58	7.31	0.27	0.54
HadGEM2-ES	1.31	0.60	0.40	0.72	0.92	0.57	0.28	0.66	7.47	0.27	0.48
inmcm4	0.71	0.14	1.21	0.79	0.93	0.21	0.36	0.08	4.80	0.32	0.36
IPSL-CM5A-	1.68	0.42	0.83	0.71	0.79	0.32	0.15	0.27	7.92	0.36	0.23
IPSL-CM5A-	1.74	0.44	0.82	0.72	0.64	0.36	0.12	0.31	7.16	0.35	0.19
IPSL-CM5B-	1.50	0.40	0.77	0.75	0.86	0.19	0.10	0.25	6.90	0.29	0.32
MIROC-ESM_	1.52	0.31	1.13	0.83	1.00	0.22	0.05	0.24	7.62	0.38	0.32
MIROC-ESM-	1.46	0.28	0.97	0.81	1.07	0.29	0.06	0.26	8.16	0.40	0.35
MPI-ESM-LR	1.09	0.49	0.56	0.75	0.36	0.32	−0.07	0.37	4.96	0.19	0.23
MPI-ESM-MR	1.17	0.51	0.56	0.38	0.36	0.43	−0.02	0.52	4.50	0.17	0.20
NorESM1-ME	0.95	0.38	0.42	0.66	0.71	0.20	0.05	0.19	5.63	0.26	0.08
CLMobs	0.92	0.50	0.44	0.66	0.50	0.26	0.18	0.51			
CESMvs.CLM	0.93	0.97	0.78	0.96	1.17	0.31	0.06	0.44			
Corr Δ LAI	−0.03	−0.01	0.49	0.57	0.52	0.23	0.30	0.06	0.44	0.37	

ESDD
6, 761–818, 2015

Table 5. Model ranking based on performance on mean annual and seasonal cycle metrics for each region (see description in Sect. 2.3).

	Tropical	Mid-latitude	High-latitude
bcc-csm1	10	10	2
bcc-csm1-1	9	8	11
BNU-ESM	18	18	1
CanESM2	17	1	16
CESM1-BGC	6	11	17
GFDL-ESM2G	14	15	17
GFDL-ESM2M	16	17	6
HadGEM2-CC	10	5	7
HadGEM2-ES	14	3	11
inmcm4	1	8	13
IPSL-CM5A-LR	2	5	13
IPSL-CM5A-MR	4	1	9
IPSL-CM5B-LR	3	4	5
MIROC-ESM	12	15	4
MIROC-ESM_CHEM-	13	14	2
MPI-ESM-LR	5	7	9
MPI-ESM-MR	7	12	15
NorESM1-ME	8	13	7

ESDD
6, 761–818, 2015

Table 6. Mean and SD across models for future projections (LAI change in $m^2 m^{-2}$) (2081–2100) for all models and for the top half of the models.

	Tropics	Mid-latitude	High-latitude
Mean Change (all models)	0.16	0.35	0.31
Mean Change (top models)	0.07	0.31	0.37
SD across models (all models)	0.35	0.23	0.20
SD across models (top models)	0.25	0.24	0.24

ESDD
6, 761–818, 2015

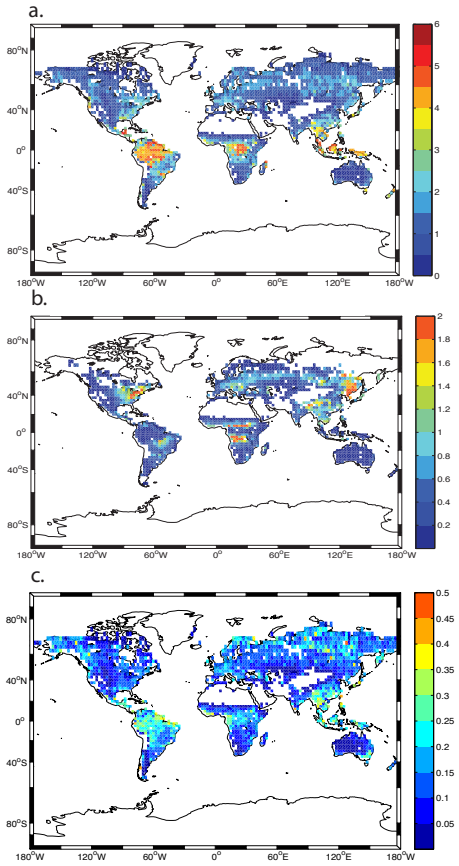

Figure 1. Observed distributions of leaf area index (LAI) (units of $m^2\,m^{-2}$) from satellite (Zhu et al., 2013) mean **(a)**; SD of the climatological observed monthly mean LAI, showing strength of the seasonal cycle **(b)**; SD of annual mean LAI, showing the strength of interannual variability **(c)**.

ESDD
6, 761–818, 2015

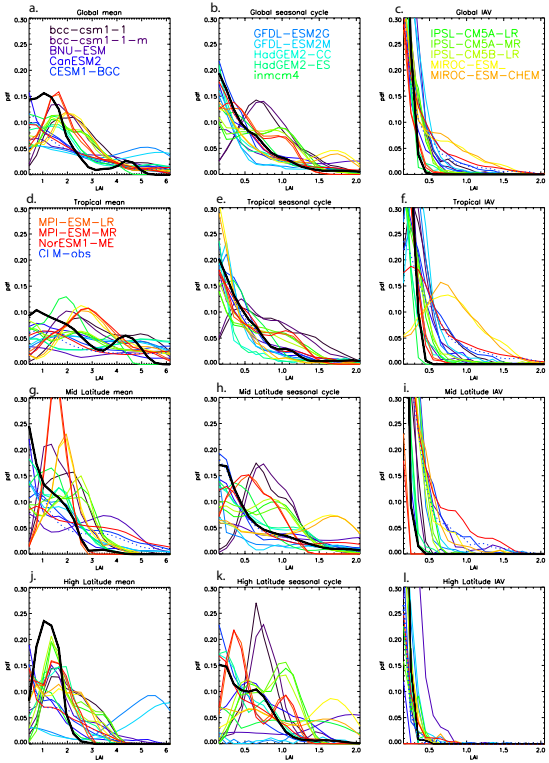

Figure 2. Probability density function (pdf) of observed (thick black line) and modeled LAI for mean (first column), seasonal cycle SD (middle column) and interannual variability SD (right column) for global **(a–c)**, tropical ($< 30^{\circ}$) **(d–f)**, mid-latitude (between 30 and 50°) **(g–i)** and high-latitudes ($> 50^{\circ}$) **(j–l)**, respectively. Probability density functions are smoothed using an Epanechnikov smoothing kernel. Models are show as colored lines, as indicated on legend in figure. CLM-obs (driven by observational-derived dataset, with the same land carbon model as CESM-BGC) is shown as a dotted line.

ESDD
6, 761–818, 2015

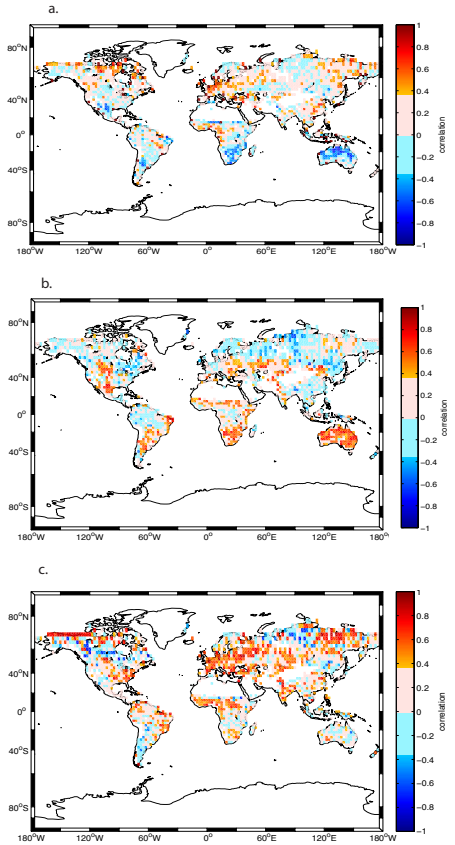

Figure 3. Rank correlation between observationally-derived interannual variability in LAI and temperature **(a)** and precipitation **(b)**, and year **(c)**. Correlations above an absolute value of 0.36 are significant at the 95 % and are shown in darker colors. Observations are derived from satellite retrievals (Zhu et al., 2013) for LAI and gridded datsasets GHCN (Fan and Dool, 2008) for temperature and GCPC (Adler et al., 2003) for precipitation.

ESDD
6, 761–818, 2015


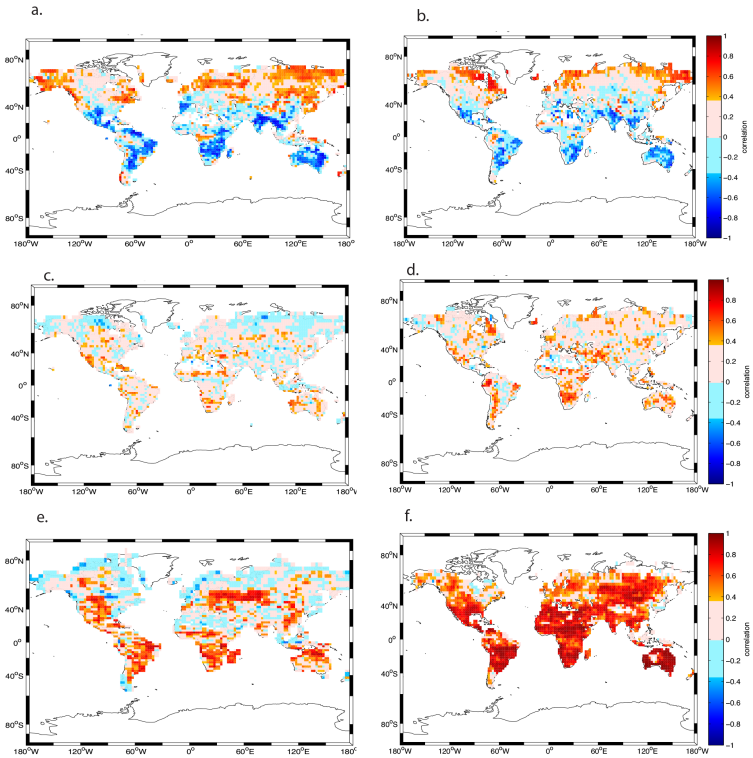

Figure 4. Rank correlation between model derived LAI and temperature (**a** and **b**) and precipitation (**c** and **d**) for the CESM-BGC (**a** and **c**) and for the CLM-obs (**b** and **d**). Both models have the same land model, but the difference is that the CESM-BGC meteorology is from the coupled climate model, while the CLM-obs is driven by datasets constrained by observations (Harris et al., 2013; Qian et al., 2006). The rank correlation between model derived LAI and precipitation are shown for the IPSL-CM5A-LR andINM-CM4 models are shown in (**e** and **f**), respectively.

ESDD
6, 761–818, 2015

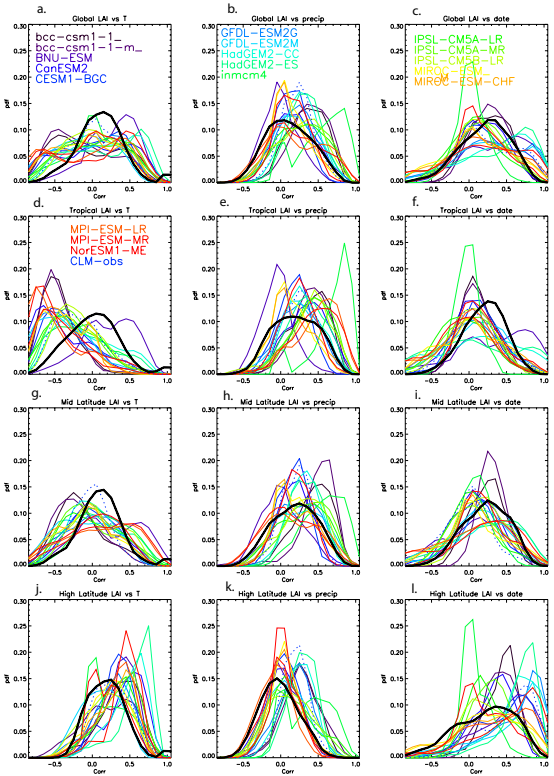

Figure 5. Probability density function (pdf) of rank correlations for the rank correlation between temperature (first column), precipitation (middle column) and date (right column) for global **(a–c)**, tropical (< 30˚) **(d–f)**, mid-latitude (between 30 and 50˚) **(g–i)** and high-latitudes (> 50˚) **(j–l)**, respectively. Probability density functions are smoothed using an Epanechnikov smoothing kernel. Models are show as colored lines, as indicated on legend in figure. CLM-obs (driven by observational-derived dataset, with the same land carbon model as CESM-BGC) is shown as a dotted line.

ESDD
6, 761–818, 2015

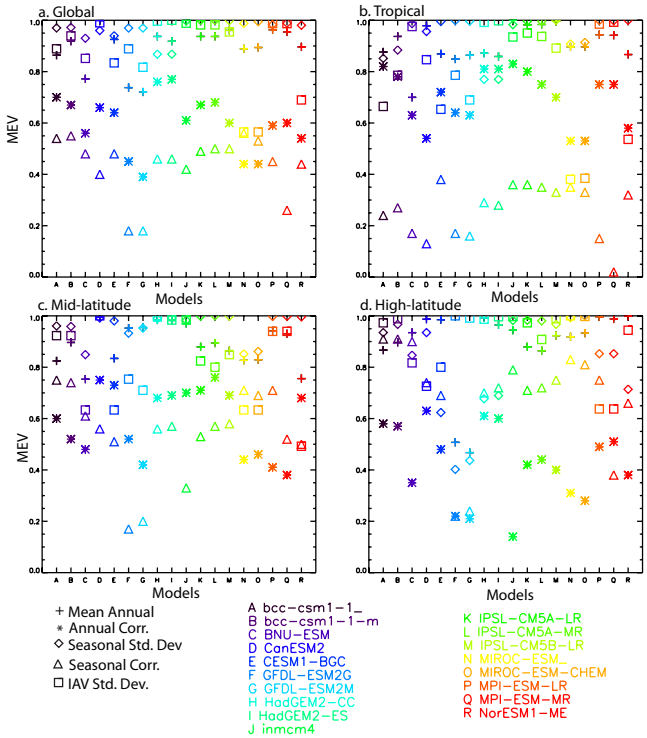

Figure 6. Comparison of model metrics for the annual mean and seasonal metrics from Table 2 across the models for **(a)** global, **(b)** tropical, **(c)** mid-latitude and **(d)** high-latitude regions. Similar information is shown in Tables 3 and 4a–c, but here converted to the Model Evaluation Value (Eq. 1) so that 1 is a perfect model simulation and lower values indicate worse simulations. Models are shown in Table 1, and listed in the figure. Metrics are mean annual (+), spatial correlation of mean annual (∗), seasonal cycle SD(diamond), mean seasonal cycle correlation (triangle) and interannual variability (IAV) SD (square).

ESDD
6, 761–818, 2015

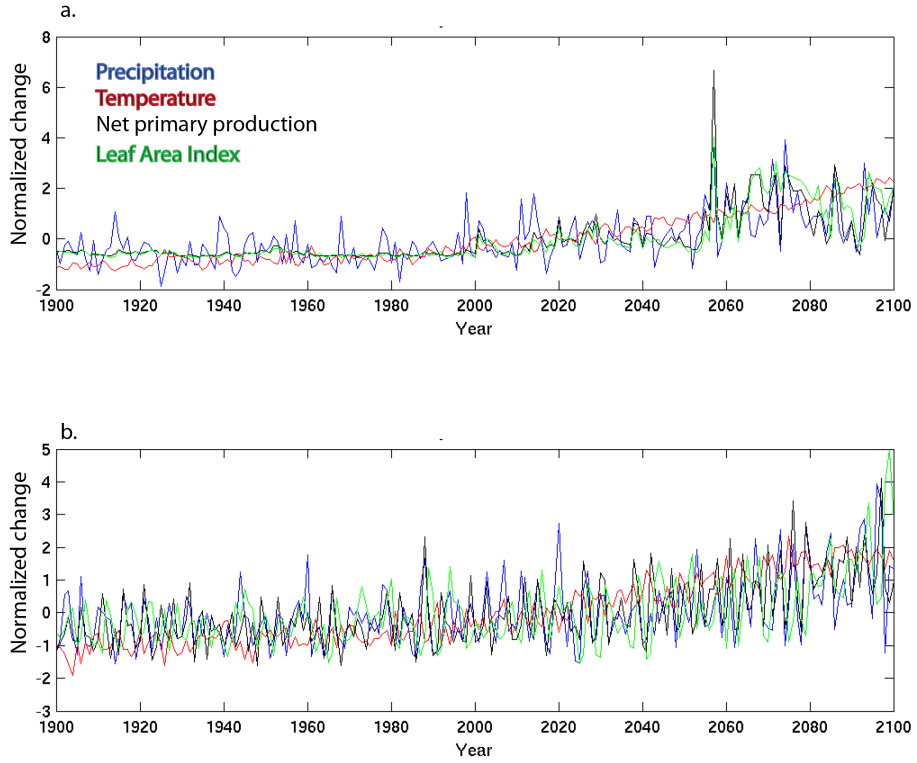

Figure 7. Time series of a location in Kenya (−1° N, 37° E) of the modeled precipitation, temperature, net primary productivity of carbon and leaf area index for the IPSL-CM5a-LR **(a)** and CESM-BGC **(b)** models. Each time series is normalized by removing the mean and dividing by the SD over the 1900–2100 time period.

ESDD
6, 761–818, 2015

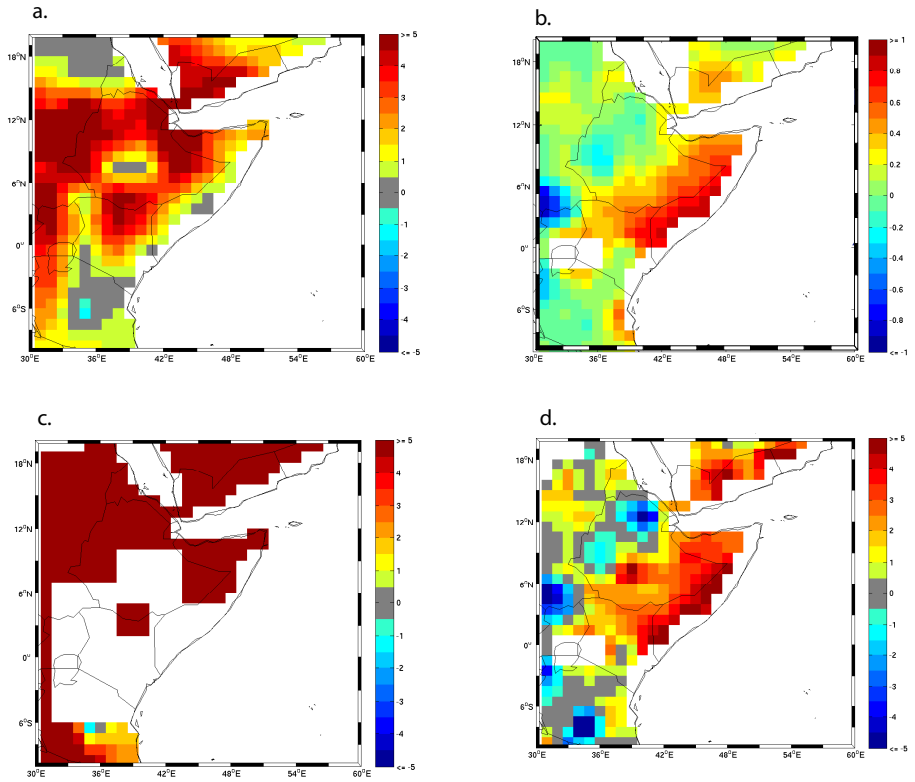

Figure 8. East African differences in mean LAI between future climate (2081–2100) and current climate (1981–2000) from two different models: BNU-ESM **(a)** and MPI-ESM **(b)** (notice the different scale). Difference in mean LAI (2081–2100 minus 1981–2000) divided by the SD in LAI (1981–2000) for BNU-ESM **(c)** and MPI-ESM **(d)**. Regions with a zero SD are left blank. Regions with absolute value > 1 are more than one SD away from current climate.

ESDD
6, 761–818, 2015

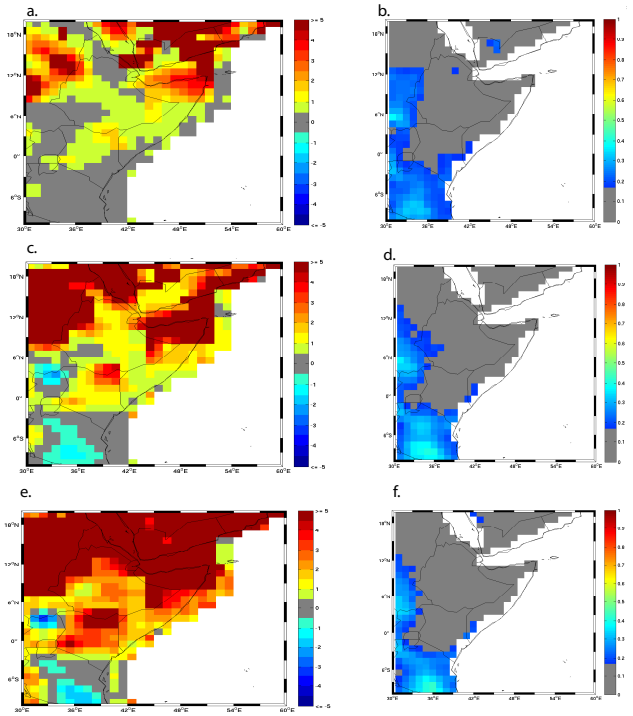

Figure 9. Mean of all models for the annual mean change in LAI over time relative to current period (1981–2000) focused on East Africa, normalized by each model's current (1981–2000) SD at each grid point, for 2011–2030 **(a)**, 2041–2060 **(c)** and 2081–2100 **(e)**. Drought frequencies based on the modeled mean percent of time that LAI is more than one SD below the current mean LAI for 2011–2030 **(b)**, 2041–2060 **(d)** and 2081–2100 **(f)**, where the current mean and SD are defined for each grid box for 1981–2000. For the current climate, the percentage of time below one SD will be 16 %, which is colored in grey, so all colors represent an increase in drought.

ESDD
6, 761–818, 2015

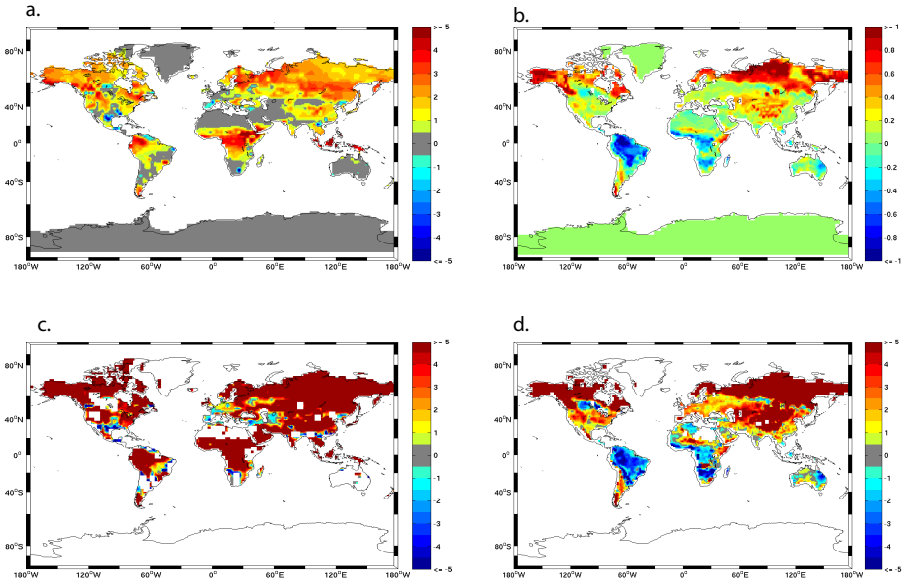

Figure 10. Global differences in mean LAI between future climate (2081–2100) and current climate (1981–2000) from two different models: BNU-ESM **(a)** and MPI-ESM **(b)** (notice the different scale). Difference in mean LAI (2081–2100 minus 1981–2000) divided by the SD in LAI (1981–2000) for BNU-ESM **(c)** and MPI-ESM **(d)**. Regions with a zero SD are left blank. Regions with absolute value > 1 are more than one SD away from current climate (as in Fig. 7, but global).

ESDD
6, 761–818, 2015

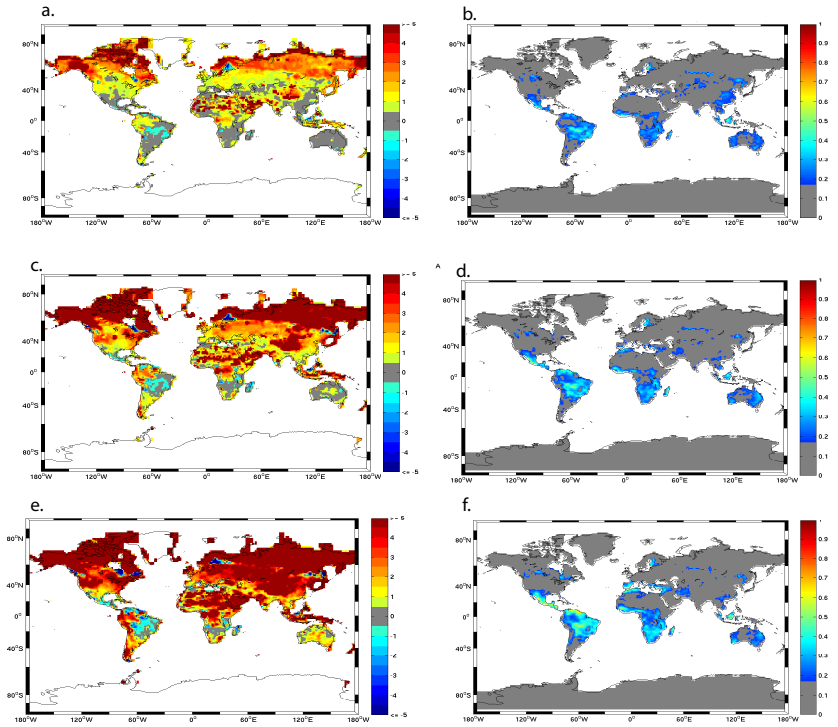

Figure 11. Mean of all models for the annual mean change in LAI over time relative to current (1981–2000) (same as Fig. 7, but for globe), normalized by each model's current (1981–2000) SD at each grid point, for 2011–2030 **(a)**, 2041–2060 **(c)** and 2081–2100 **(e)**. Indication of drought is the model mean percent of time that LAI is more than one SD below the current mean LAI and is shown for 2011–2030 **(b)**, 2041–2060 **(d)** and 2081–2100 **(f)**, where the current mean and SD are defined for each grid box for 1981–2000. For the current climate, the percentage of time below one SD will be 16 %, which is colored in grey, so all colors represent an increase in drought.

ESDD
6, 761–818, 2015

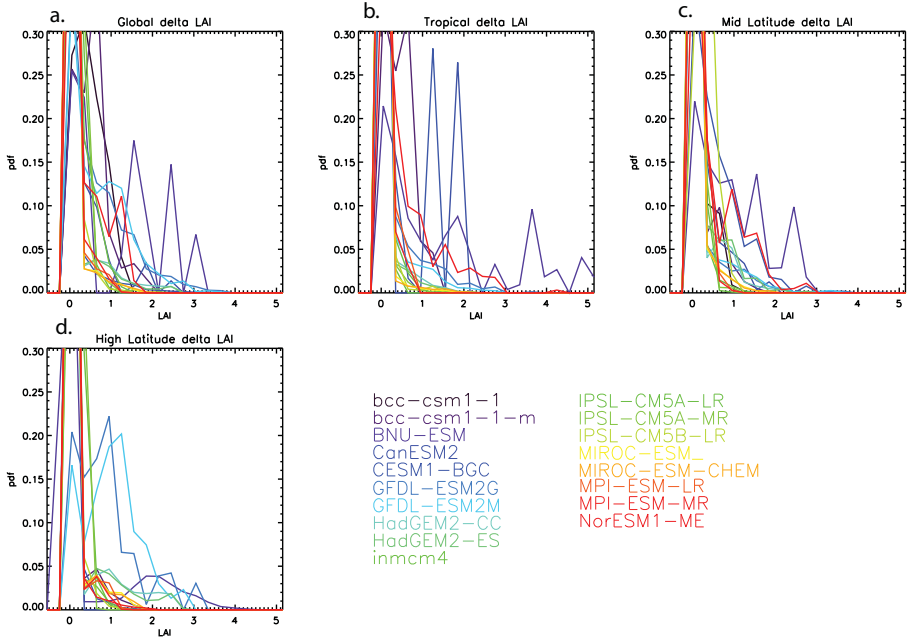

Figure 12. Probability density function of the change in LAI between 2081–2100 at each grid box for each model for the globe **(a)**, tropics (< 30˚) **(b)**, mid-latitudes (between 30 and 50˚) **(c)** and high-latitudes (> 50˚) **(d)**. Probability density functions are smoothed using an Epanechnikov smoothing kernel. Models are shown as colored lines, as indicated on legend in figure.

ESDD
6, 761–818, 2015

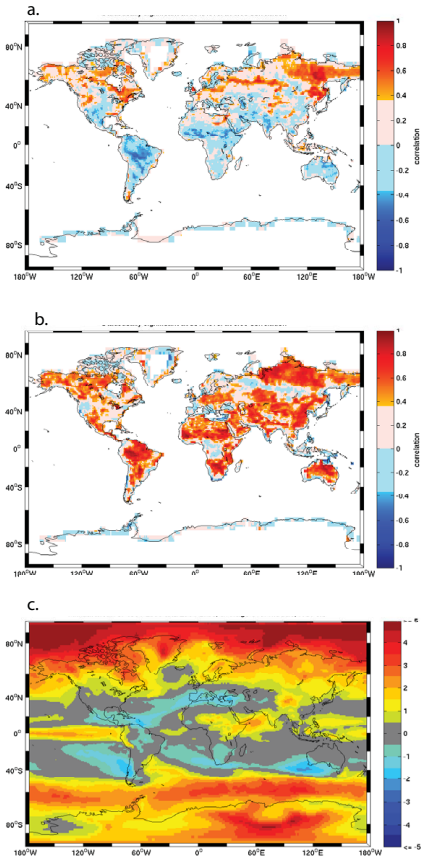

Figure 13. Rank correlation across models at every grid box of the mean model change in LAI (2081–2100 minus 1981–2000) against the model change over the same time period of temperature **(a)** and precipitation **(b)**. The mean model change (2081–2100 minus 1981–2000) in precipitation, normalized by the current SD (1981–2000) in precipitation at each grid cell **(c)**.

ESDD
6, 761–818, 2015




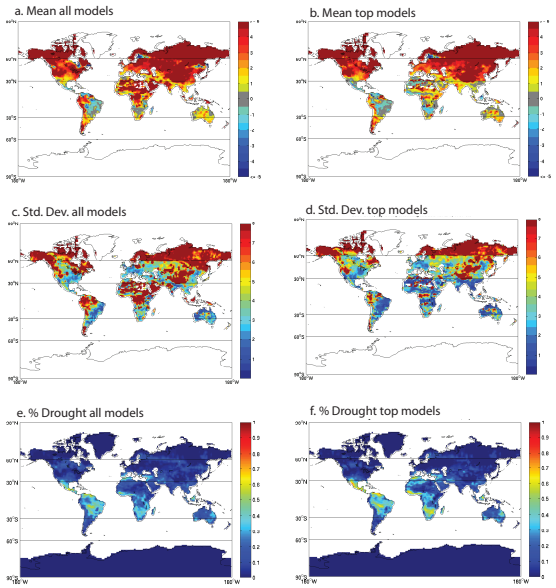

Figure 14. Mean of all models for the annual mean change in LAI over time (2081–2100) relative to current (1981–2000), normalized by each model's current (1981–2000) SD at each grid point **(a)** for all models (same as Fig. 11e) and **(b)** for the top models, defined as the models performing in the top half (Table 6) for each region, tropical, mid-latitude or high-latitude. Because different models are included in different regions, there can be discontinuities at the boundaries in Fig. 14b, d, f (e.g. 30 and 60° latitude). The SD in the mean future projection at 2081–2100 across the models at each grid point are shown for **(c)** all models and **(d)** top models. Indication of drought is the model mean percent of time that LAI is more than one SD below the current mean LAI and is shown for **(e)** all models (same as Fig. 11f) and **(f)**, top models for the period 2081–2100, where the current mean and SD are defined for each grid box for 1981–2000. For the current climate, the percentage of time below one SD will be 16 %, which is colored in grey, so all colors represent an increase in drought.

ESDD
6, 761–818, 2015

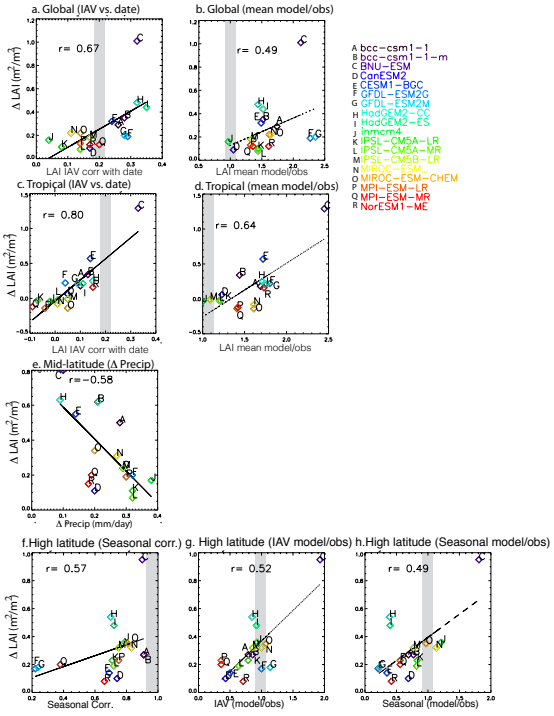

Figure 15. Scatterplot of the metrics with the highest absolute value of the correlation between the metric and future LAI changes across the globe (LAI correlated with date **(a)** and mean LAI model/obs **(b)**) tropics (< 30°) (LAI correlated with date **(c)** and mean LAI model/obs **(d)**), mid-latitudes (between 30 and 60°) projected change in precipitation **(e)**) and high-latitudes (> 60°) (seasonal cycle average correlation **(f)**, strength of IAV model/obs **(g)**, and seasonal cycle strength model/obs **(h)**). The symbols are in the shown colors for each model. The grey represents the value an ideal model would have based on the observations. The black line is the line which results from a linear regression of the *x* and *y* axis.

ESDD
6, 761–818, 2015