

Evaluation of vegetation cover and land-surface albedo in MPI-ESM CMIP5 simulations

V. Brovkin,¹ L. Boysen,¹ T. Raddatz,¹ V. Gayler,¹ A. Loew,¹ and M. Claussen^{1,2}

Received 29 May 2012; revised 2 November 2012; accepted 27 November 2012.

[1] In recent generation Earth system models (ESMs), land-surface grid cells are represented as tiles covered by different plant functional types such as trees or grasses. Here, we present an evaluation of the vegetation-cover module of the ESM developed at the Max Planck Institute for Meteorology in Hamburg, Germany (MPI-ESM) for present-day conditions. The vegetation continuous fields (VCF) product that is based on satellite observations in 2001 is used to evaluate the fractional distributions of woody vegetation cover and bare ground. The model performance is quantified using two metrics: a square of the Pearson correlation coefficient, r^2 , and the root-mean-square error (RMSE). On a global scale, r^2 and RMSE of modeled tree cover are equal to 0.61 and 0.19, respectively, which we consider as satisfactory values. The model simulates tree cover and bare ground with r^2 higher for the Northern Hemisphere (0.66) than for the Southern Hemisphere (0.48–0.50). We complement this analysis with an evaluation of the simulated land-surface albedo using the difference in net surface radiation. On a global scale, the correlation between modeled and observed albedos is high during all seasons, whereas the main disagreement occurs in spring in the high northern latitudes. This discrepancy can be attributed to a high sensitivity of the land-surface albedo to the simulated snow cover and snow-masking effect of trees. By contrast, the tropics are characterized by very high correlation and relatively low RMSE (5.4–6.5 W/m²) during all seasons. The presented approach could be applied for an evaluation of vegetation cover and land-surface albedo simulated by different ESMs.

Citation: Brovkin, V., L. Boysen, T. Raddatz, V. Gayler, A. Loew, and M. Claussen (2013), Evaluation of vegetation cover and land-surface albedo in MPI-ESM CMIP5 simulations, *J. Adv. Model. Earth Syst.*, 5, doi:10.1029/2012MS000169.

1. Introduction

[2] Terrestrial ecosystems substantially affect near-surface heat and moisture fluxes as well as the exchange of greenhouse gases between the land surface and the atmosphere. Land-surface models (LSMs) are considered nowadays as a standard part of atmospheric models. LSMs have been constantly developed and updated from very simplified concepts used in the 1980s toward current generation models that include more ecological processes such as the effect of climate and CO₂ changes on land vegetation composition [Arora, 2002; Sellers *et al.*, 1997]. Models that include the transient response of vegetation cover to climate changes (dynamic global vegetation models (DGVMs)) were first developed as stand-alone models without feedback to climate [Cramer

et al., 2001; Sitch *et al.*, 2003]. DGVM approaches were quickly adopted into LSMs to calculate interactive vegetation cover according to climate change simulated by atmospheric models. Changing fractions of woody and herbaceous vegetation cover, as well as of bare land, affect land-surface albedo and evapotranspiration, which, in turn, modify near-ground temperature and precipitation. These land-atmosphere interactions are pronounced in several large-scale “hot spots” [Claussen *et al.*, 2004]. These areas include high northern latitudes where the snow-masking effect of forests leads to additional warming in comparison with herbaceous cover; tropical forests, where extensive transpiration enhances atmosphere moisture recycling; and subtropical deserts, where the presence of vegetation cover shifts climate toward moister conditions. Some local-scale feedbacks such as the formation of vegetation patterns in semideserts are not yet represented in the global models but proposed for inclusion [Rietkerk *et al.*, 2011].

[3] The vegetation-cover composition affects not only land biophysics but also land-atmosphere exchange of CO₂ and other greenhouse gases such as CH₄ and N₂O. In particular, changes in tree cover strongly affect the

¹Max Planck Institute for Meteorology, Hamburg, Germany.

²Meteorological Institute, Klima Campus, University of Hamburg, University of Hamburg, Germany.

amount of carbon stored in biomass and soil. This has an effect on the atmospheric CO₂ concentration and operates as a biogeochemical feedback between vegetation dynamics and climate [Bathiany et al., 2010; Port et al., 2012]. Feedbacks between forests and atmospheric chemistry through emissions of volatile organic compounds such as isoprene are an emerging research field as well [e.g., Arneth et al., 2010].

[4] Through biophysical and biogeochemical effects, changes in vegetation cover affect the simulated climate in future climate projections, such as performed in the framework of the Climate Model Intercomparison Project 5 (CMIP5). Several CMIP5 Earth system models (ESMs) include dynamic vegetation-cover models [Collins et al., 2011; Watanabe et al., 2011]. These models simulate fractions of plant functional types (PFTs) in response to the climate simulated by the atmospheric models. Since ESMs inevitably simulate temperature and precipitation fields that differ from observations, vegetation patterns in ESMs are different from the ones simulated by vegetation models driven by observed climate data (stand-alone models). The latter also have biases, as the parameterizations of vegetation dynamics in these models are far from being perfect.

[5] These biases need to be quantified. Quantitative assessment of LSMs, as a way to evaluate their performance, is part of the long-term modeling strategy. Several projects, such as Project for Intercomparison of Landsurface Parameterization Schemes, are dedicated to the intercomparison of LSMs [Schlosser et al., 2000; Slater et al., 2001]. A quantitative assessment of LSMs has been done recently in the form of benchmarking that involves scoring metrics aiming to express the simulation quality of numerous aspects of the vegetation cover in terms of one scalar value. Several scoring metrics have been suggested for physical characteristics of the LSMs [Abramowitz et al., 2008], the terrestrial carbon cycle and hydrology [Blyth et al., 2011; Randerson et al., 2009], and atmospheric CO₂ observations [Cadule et al., 2010].

[6] The aim of this study is to evaluate the vegetation cover and land-surface albedo simulated by the ESM developed at the Max Planck Institute for Meteorology in Hamburg, Germany (MPI-ESM). As the outcome of a model evaluation is always specific to the model version and resolution, we focus here on the CMIP5 model version at spatial resolution T63 (1.9° × 1.9°), which was used for future climate and vegetation-cover projections.

2. Methods

2.1. Atmosphere-Ocean Model

[7] We used the MPI-ESM in low resolution. It includes the atmospheric model ECHAM6 in T63 resolution with 47 vertical levels described by Stevens et al. [2013], the oceanic model MPI-OM at approximately 1.6° resolution with 40 vertical layers [Jungclaus et al., 2006], and the LSM JSBACH [Raddatz et al., 2007] sharing the horizontal grid of the atmospheric model. All MPI-ESM modules interact directly without flux adjustments. A detailed description of the model and an evaluation of the model

performance regarding temperature and precipitation fields is given by M. A. Giorgetta et al. (Climate change from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project 5, submitted to *Journal of Advances in Modeling Earth Systems*, 2012).

2.2. Land-Surface Model

[8] The LSM of MPI-ESM, JSBACH, simulates fluxes of energy, water, momentum, and CO₂ between land and atmosphere. The modeling concept is based on a tiled (fractional) structure of the land surface. Each land grid cell is divided into tiles covered with eight natural PFTs (i.e., different types of trees, shrubs, and grasses) and four anthropogenic PFTs (crop and pasture types; see Table 1). Two types of bare surface are taken into account, seasonally bare soil and permanently bare ground, i.e., desert.

[9] The calculation of land-surface albedo is done separately for visible and near-infrared solar radiation. It considers the fractional cover of each PFT, the desert fraction, the leaf area index of each PFT, and the snow on soil as well as on the vegetation canopy. The effect that forests mask underlying snow is accounted for. A detailed description of the albedo scheme is presented by Otto et al. [2011], except for the albedo of snow-covered surfaces, which is described by Dickinson et al. [1993].

[10] The vegetation model in JSBACH includes an efficient module for vegetation dynamics [Brovkin et al., 2009]. It is based on the assumption that competition between different PFTs is determined by their relative competitiveness expressed in annual net primary productivity, as well as natural and disturbance-driven mortality (fire and wind disturbance).

[11] Anthropogenic land use is predetermined. In particular, the extent of pasture and cropland is prescribed. Transitions from natural to anthropogenic land cover and vice versa follow the New Hampshire harmonized land use protocol by Hurtt et al. [2011]. The dynamic vegetation model of JSBACH only affects the natural vegetation distribution and defines the type of pasture (C₃ or C₄). The implementation of the harmonization protocol into MPI-ESM is described by C. Reick et al. (The representation of natural and anthropogenic land

Table 1. A Correspondence Between JSBACH and VCF Vegetation Classes

JSBACH PFT	VCF Type
Tropical broadleaf evergreen trees	Tree
Tropical broadleaf deciduous trees	
Extratropical evergreen tree	
Extratropical deciduous tree	
Deciduous shrubs	Grass
C3 grass	
C4 grass	
C3 crop	
C4 crop	
C3 pasture	
C4 pasture	

cover change in MPI-ESM, submitted to *Journal of Advances in Modeling Earth Systems*, 2012).

2.3. Observations of Vegetation Cover Used for the Model Evaluation

[12] We chose the vegetation continuous fields (VCF) data set [Hansen et al., 2003, 2007] derived from the moderate resolution imaging spectroradiometer (MODIS) sensor data for comparison with the model results. Two reasons for selecting the VCF product were decisive. First, the product has a global coverage on $1 \text{ km} \times 1 \text{ km}$ for a relatively recent time period (the inputs date from 31 October 2000 to 9 December 2001). Second, the VCF data set describes the land surface in fractions of vegetation-cover types: woody vegetation, herbaceous vegetation, and bare ground. This is very similar to the DGVM approach describing the vegetation cover in fractions of PFTs. The only difference is that DGVMs have usually more than two PFTs per grid cells. For comparison purposes, the model output is lumped into three broad vegetation classes: woody PFT (trees and shrubs), herbaceous PFT (grasses and crops), and a bare (nonvegetated) ground (see Table 1).

2.4. Observations of Albedo Used for the Model Evaluation

[13] MODIS surface albedo (MCD43C3, ver5) observations of 10 years (2001–2010) [Schaaf et al., 2002] are used for comparison with JSBACH results. The MODIS surface albedo has an absolute error in the order of 0.02, whereas it is slightly higher over snow-covered areas. A brief summary of different validation studies for the MODIS albedo product can be found in Liu et al. [2009]. The albedo observations are filtered in accordance with the product quality flags to ensure that only best quality observations are considered in the reference data set. The data are then reprojected to the Gaussian T63 grid of the LSM. The monthly mean surface albedo and its variance are calculated from the 10 year time series for each grid cell. Ten years of observations is a relatively short period in estimating the climatological mean of the surface albedo, but the MODIS observations started only in 2001. On the other hand, changes in vegetation cover and characteristics might already occur at subdecadal timescales [e.g., de Jong et al., 2012; Fensholt and Proud, 2012]. Such changes might significantly change the surface albedo on decadal timescales and therefore affect climate [Govaerts and Lattanzio, 2008; Knorr et al., 2001; Loew and Govaerts, 2010; Myhre et al., 2005]. We analyzed the effect of sampling the MODIS observations as well as the MPI-ESM simulations on shorter timescales (5 years) and found very similar results to those based on the 10 year time series. For the simulated land-surface albedo, we were not limited by the length of the record and used the climatological mean of a 30 year period (1971–2000).

3. Evaluation of the Vegetation Cover

[14] Here, we evaluate vegetation cover simulated by the fully interactive MPI-ESM in the ensemble mean of three CMIP5 historical simulations from 1850 to 2005.

For evaluation, we use the VCF product that is based on the MODIS satellite data for the year 2001 [Hansen et al., 2007]. Consequently, we selected the vegetation cover simulated for the year 2001 from the historical CMIP5 simulation (1860–2005). The comparison of matching years is especially important, as anthropogenic land use was prescribed for the historical period, which significantly affects vegetation cover in many regions. An interannual variability in tree cover and bare ground fraction due to interannual variability of climate is relatively small because of slow vegetation dynamics.

3.1. Evaluation Metrics

[15] Agreement between simulated and observed vegetation cover could be evaluated using several metrics. To quantify the spatial correlation between vegetation patterns from the model and from observations, a square of the Pearson correlation coefficient (r^2) is used. In a linear approximation, this metric quantifies a fraction of variation explained by the model:

$$r^2 = \frac{\left[\sum_{i=1}^N w_i (M_i - \bar{M})(O_i - \bar{O}) \right]^2}{\sum_{i=1}^N w_i (M_i - \bar{M})^2 \sum_{i=1}^N w_i (O_i - \bar{O})^2},$$

where M_i and O_i are the vegetation fractions simulated by the model or observed in the grid cell i , respectively; w_i is a normalized weight (area) of grid cell i ($\sum_{i=1}^N w_i = 1$), and N is the total number of grid cells under evaluation.

[16] The amplitude of the difference between two data sets is measured using the root-mean-square error (RMSE):

$$\text{RMSE} = \sqrt{w_i \sum_{i=1}^N (M_i - O_i)^2}. \quad (1)$$

[17] The two metrics, r^2 and RMSE, are calculated separately for each vegetation class as well as for different regions.

[18] As the sum of all cover fractions (trees, grass, and bare ground, or nonvegetated fraction) is equal to one, only two out of the three vegetation classes are independent. For the evaluation, we chose the two classes: (i) fraction of trees and (ii) fraction of bare ground. The rationale for considering bare ground instead of grass cover is that bare ground plays an important role in the albedo-based feedback between vegetation cover and climate.

3.2. Global Evaluation

[19] The simulated tree cover fraction is in a good overall agreement with the observations (Figure 1). The main patterns of boreal forest match the data. A noticeable disagreement is in the northern polar regions (above

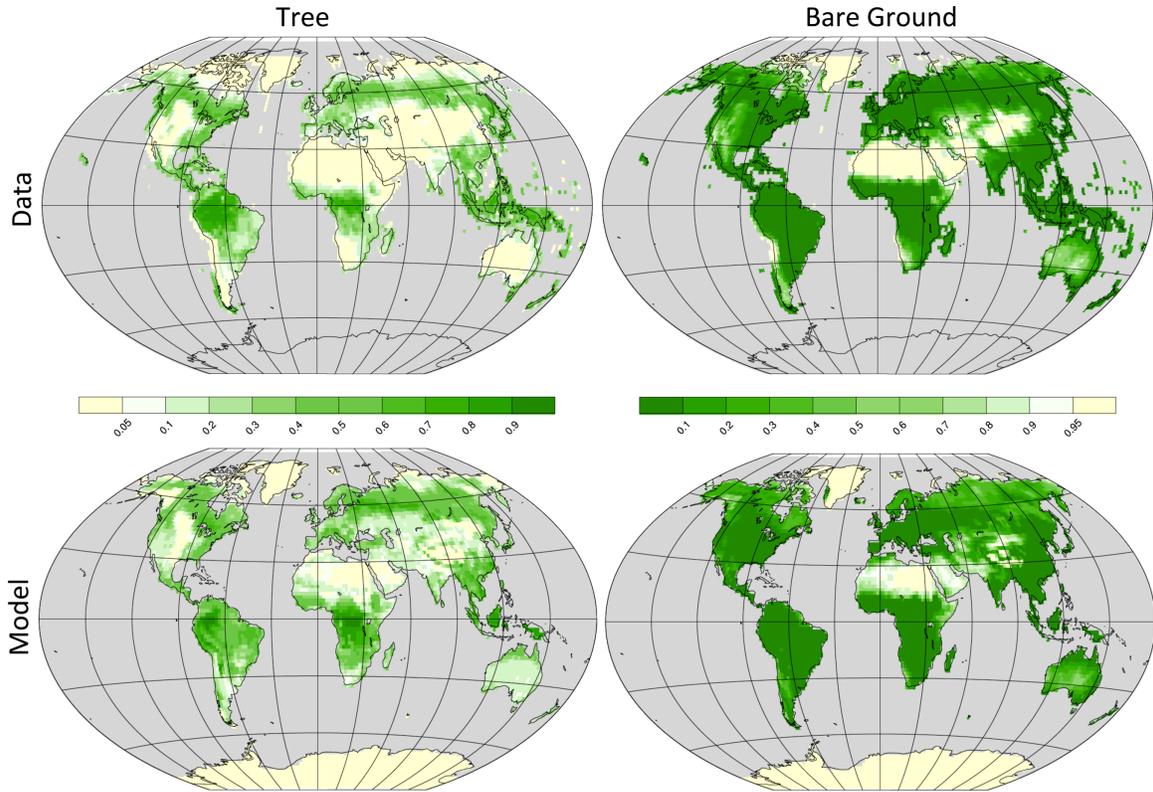


Figure 1. Vegetation classes simulated by the MPI-ESM for the (top) year 2001 and (bottom) VCF [Hansen *et al.*, 2007] upscaled to the model resolution. (left) tree cover and (right) bare ground as grid cell fractions.

60°N), where the simulated tree fraction is too high in West and East Siberia as well as in the northwest territories of Canada (Figure 2, top). This mismatch could be explained by two reasons. First, the vegetation model parameterizations, in particular, the simple disturbance module and the absence of permafrost, cause too much tree cover in high latitudes. Second, the simulated climate is too warm in the northwest territories of Canada (except British Columbia, where the modeled annual mean surface-air temperature is by 2°C–4°C lower than in the observations). The bare ground fraction is overestimated in the regions above 60°N likely due to low productivity and reduced canopy cover simulated by the model.

[20] Tropical forests are simulated reasonably by the model as well. The main deficiencies are too low tree cover fractions in the Amazon region and too extended tree cover in northeastern Brazil, west Sahel, tropical and southern Africa, as well as in Australia (Figure 2). The bare ground is underestimated in subtropical deserts, especially in central Asia (CEAS). A comparison of the zonal vegetation-cover distribution shows that the model slightly overestimates tree cover in all latitudes and underestimates bare ground in subtropical and temperate regions (Figure 3), but the general agreement between model and data is remarkable. The overestimation of the tree cover is likely due to the simplified parameterization of tree-grass competition in the model, and missing processes, such as permafrost.

[21] The quality of the model simulations is quantified in Table 2. On a global scale, r^2 and RMSE of tree cover are equal to 0.61 and 0.19, respectively, which we consider as satisfactory values. The model simulates tree cover and bare ground with r^2 higher for the Northern Hemisphere (0.66) than for the Southern Hemisphere (0.48–0.50). Grass cover, as an intermediate class between tree cover and bare ground, is reproduced less reliably in both Northern Hemisphere and Southern Hemisphere (r^2 of 0.44 and 0.17, respectively).

[22] Furthermore, we calculated an overall score including tree cover and bare ground for the tropics and extratropics (Table 3). This is done by taking the mean of the scores for both vegetation types in both tropics and extratropics. For scoring, we use the 1-RMSE metric, since this metric is increasing with less error. We applied equal weights of 100 for both metrics and received scores of 62 and 80 for r^2 and 1-RMSE, respectively. An overall score of 71 out of 100 is calculated for MPI-ESM (Table 3). This relative score could be used for comparison with different models or model versions, regarding the model performance for these two vegetation types.

3.3. Regional Evaluation

[23] For regional evaluation, we used the world subdivision in accordance with 14 geographical regions of the Global Fire Emission Database (GFED, see *van der Werf et al.* [2006]). The metrics values, r^2 and RMSE,

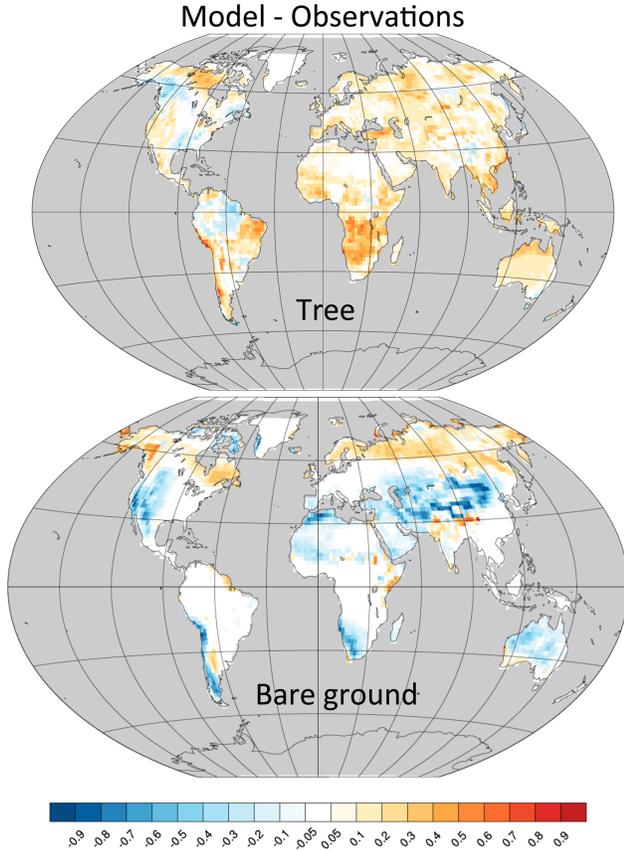


Figure 2. Differences between the model and observations in (top) tree and (bottom) bare ground fractions.

were calculated separately for each region. They are presented in Figure 4 and Figure S1 in the auxiliary material¹ for tree and bare ground cover, respectively. The tree cover in the boreal Asia is reproduced well in the model with r^2 of 0.67. The North African tree cover is represented with highest r^2 score of 0.78. RMSE is very low for bare ground for two regions (6 and 10, Equatorial Asia and Northern Hemisphere South America, respectively); however, this variable is not representative because of very low bare ground fraction in these regions (Figure 4).

4. Evaluation of the Land-Surface Albedo

[24] Evaluation of the land-surface albedo, α_s , in the framework of the climate model should account not for the absolute values of albedo but for its relative significance for the radiation budget at the surface. For example, in winter time in the polar regions, there is little incoming solar radiation, SW_{down} , and, therefore, the significance of the albedo changes in these regions is highly uncertain. Hence, it is more appropriate to analyze the net surface solar radiation:

$$Q_{\text{sw}} = SW_{\text{down}} - SW_{\text{up}} = SW_{\text{down}} \times (1 - \alpha_s), \quad (2)$$

¹Auxiliary materials are available in the HTML. doi:10.1029/2012MS000169.

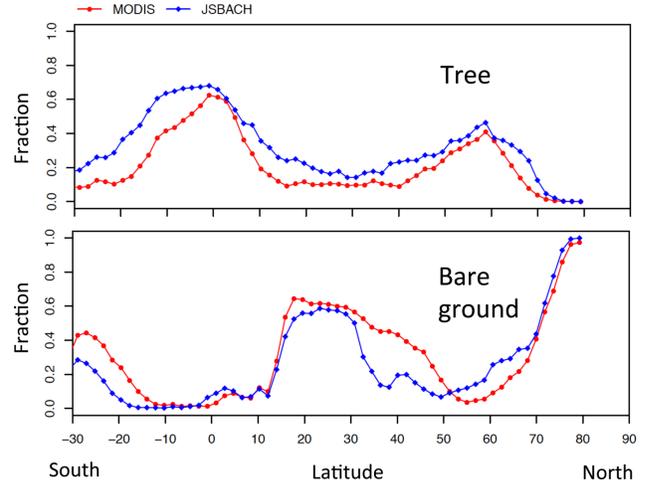


Figure 3. Comparison of zonal averages between MPI-ESM results for the year 2001 (blue) and VCF (red). (top) Tree fraction and (right) bare ground fraction.

where SW_{up} is the upward flux of solar radiation. A climatological mean seasonal cycle of SW_{down} derived from MPI-ESM simulations for the period of 2001–2005 was used to weight the surface albedo. Instead of using SW_{down} from the MPI-ESM model output, one might have also taken SW_{down} estimates from long-term satellite records. *Hagemann et al.* [2013] analyze in detail the accuracy of the simulated SW_{down} fields and conclude that the MPI-ESM does not show significant differences to existing observational data sets. Therefore, we used simulated SW_{down} fields for both observed and modeled albedo values.

4.1. Evaluation Metrics

[25] As for vegetation cover, r^2 and RMSE were used for the analysis of spatial correlation and amplitude of

Table 2. Evaluation of Simulated Vegetation Cover in Terms of r^2 and RMSE

Region	Vegetation Class	r^2	RMSE	N (Number of Grid Cells)
Global	Tree	0.61	0.19	4354
	Grass	0.39	0.24	4354
	Bare ground	0.65	0.22	4169
Northern Hemisphere	Tree	0.66	0.16	3514
	Grass	0.44	0.23	3514
	Bare ground	0.66	0.21	3413
Southern Hemisphere	Tree	0.48	0.25	840
	Grass	0.17	0.24	840
	Bare ground	0.48	0.21	756
Tropics	Tree	0.63	0.21	1556
	Grass	0.57	0.20	1556
	Bare ground	0.86	0.16	1429
Extratropics	Tree	0.57	0.16	2798
	Grass	0.24	0.26	2798
	Bare ground	0.43	0.27	2740

Table 3. Evaluation of Vegetation Cover on the Global Scale

Region	Vegetation Class	r^2	RMSE
Tropics	Tree	0.63	0.21
	Bare ground	0.86	0.16
Extratropics	Tree	0.57	0.16
	Bare ground	0.43	0.27
Weighted global		0.62	0.20
Score (max: 100)		62	80
Total score (max: 100)		71	

difference between the model and observations. The surface albedo shows a higher temporal variability than the vegetation cover. Therefore, an additional evaluation metric is applied, which takes into account the different temporal dynamics of surface albedo in different regions. A normalized error variance (e^2) is calculated according to Reichler and Kim [2008] as follows:

$$e^2 = \sum_{i=1}^N \left(\frac{w_i (\overline{M}_i - \overline{O}_i)^2}{\sigma_{\text{obs},i}^2} \right), \quad (3)$$

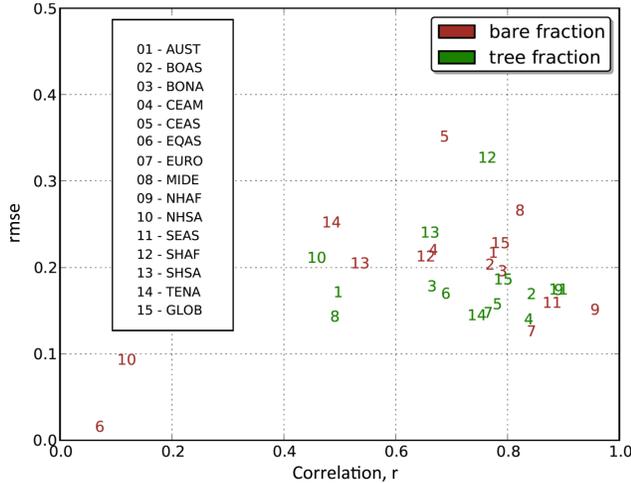


Figure 4. A diagram of model-data agreement for tree and bare ground fractions. Statistics for 15 regions were computed, and a number was assigned to each region considered. The position of each number on the plot quantifies how closely the modeled vegetation patterns of the different regions match the observations ($r = 1$, $\text{RMSE} = 0$). The region acronyms are as follows: Australia and New Zealand (AUST), boreal Asia (BOAS), boreal North America (BONA), central Asia (CEAM), central Asia (CEAS), Equatorial Asia (EQAS), Europe (EURO), Middle East (MIDE), Northern Hemisphere Africa (NHAF), Northern Hemisphere South America (NHSA), Southeast Asia (SEAS), Southern Hemisphere Africa (SHAF), Southern Hemisphere South America (SHSA), Temperate North America (TENA), and Global (GLOB).

where w_i is an area weight of grid cell i ($\sum_i w_i = 1$) of the grid cell i ; $\sigma_{\text{obs},i}^2$ is the interannual variance of the observations in the grid cell i ; \overline{M}_i and \overline{O}_i are the values of the variable (net surface solar radiation) averaged over the analysis time period in the grid cell i in the model and the observation, respectively; and N is the total number of land grid cells.

4.2. Global Evaluation

[26] The maps of difference between model and observation in net surface solar radiance Q_{SW} for four seasons are presented in Figure 5. Bluish (reddish) colors correspond to regions where simulated surface albedo is higher (lower) than in the observations. In the high latitudes of the Northern Hemisphere, the differences between the model and the data are most pronounced during the boreal winter (December–January–March (DJF)) and spring (March–April–May (MAM)) seasons (Figure 5, top). The albedo in this period is dominated by the snow cover and its masking by the tree cover [e.g., Bonan, 2008; Essery et al., 2009]. In the northwestern part of North America (British Columbia), the model underestimates tree cover fraction (Figure 2, top). This is a likely explanation of too high albedo in the model that results in an underestimation of net surface radiation in this region in all seasons (Figure 5). In the regions above 60°N , the model generally overestimates tree cover (Figure 2, top). In boreal winter season, this does not influence the surface radiation budget because of too little solar radiation. In the MAM season, the flux of radiation is much more significant. The patterns of too low-surface albedo in North America (the Nunavut province of Canada) and Eurasia (West Siberia) are complemented by the regions with too high-surface albedo (Alaska, East Siberia). These mixed patterns could be explained by the interplay of patterns of too high tree cover (West Siberia, Nunavut) and too high albedo of snow cover (e.g., eastern Siberia). In the summer season, the land-surface albedo is overestimated by the model almost everywhere except in the Sahel and some desert regions in Asia and Australia.

[27] The correlation between simulated and observed albedos on global and hemispherical scales is high (Table 4). The tropics are characterized by high correlation and relatively low RMSE (5.4 – 6.5 W/m^2) during all seasons. The albedo has highest score during boreal winter on the global scale and in the Southern Hemisphere. During boreal spring season (MAM), the correlation is the lowest on both global and hemispheric scales. This is presumably an indication of the biases in the albedo induced by biases in snow and tree cover in high and temperate northern latitudes discussed above. The boreal autumn (September–October–November (SON)) is characterized by lowest values of RMSE in the Northern Hemisphere as well as on the global scale and highest value of correlation in the Northern Hemisphere.

4.3. Regional Evaluation

[28] Our model generally overestimates albedo during all seasons. The box-and-whisker plots of Figure 6, therefore, show a slightly negative shortwave net radiation

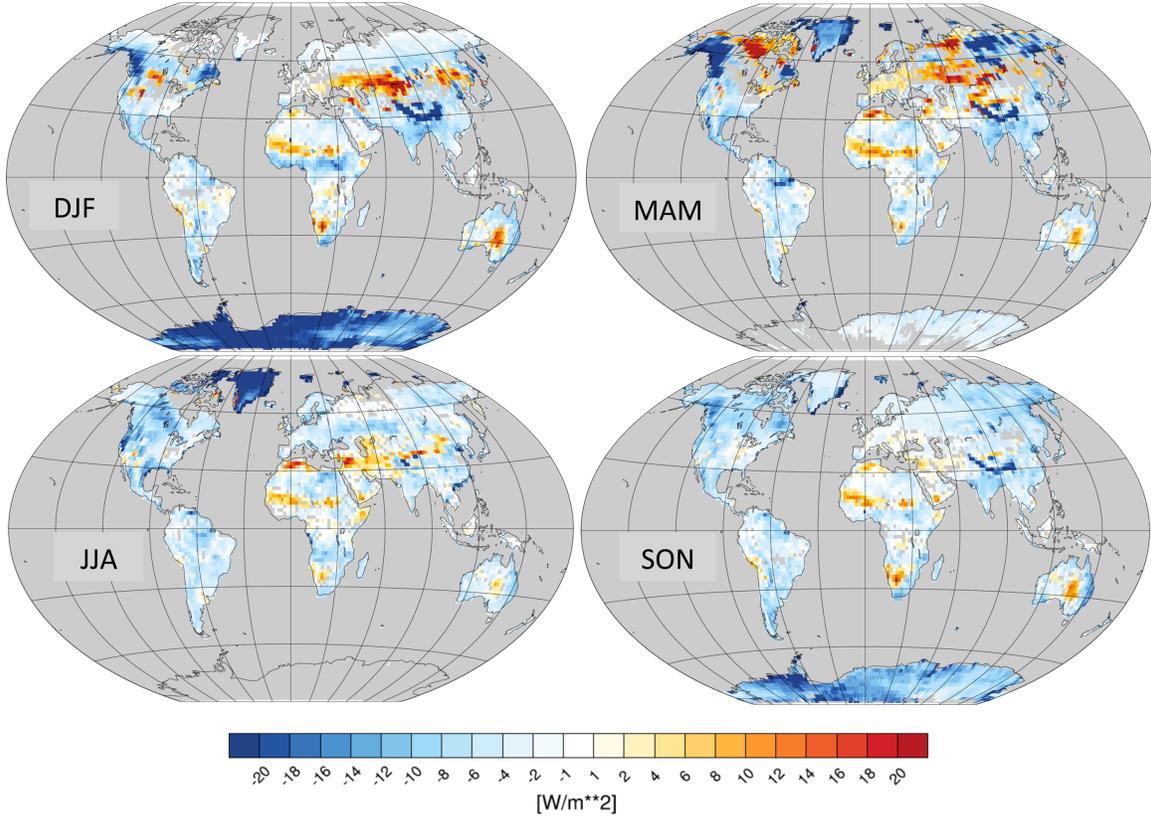


Figure 5. The difference between model and observations in the net surface radiation flux Q_{SW} (W/m^2) averaged for different seasons (top left, DJF; top right, MAM; bottom left, June–July–August (JJA); bottom right, SON). Areas with statistically significant changes ($p < 0.05$) are stippled.

bias for most regions. The net radiation bias is most pronounced in boreal North America (BONA) in spring (MAM). In this season, the model substantially underestimates the net radiation flux (Figure 6). The score of the 14 GFED regions is presented in a *Taylor* [2001] plot in Figure 7 and in a table format in Figure S2. The variance of the model is hereby normalized to the variance of the observations to make the results of the different regions better comparable. It appears that the BONA region has the lowest r^2 and the highest RMSE, both in absolute units (W/m^2) and relative to the net radiation flux, $RMSE/Q_{SW}$. The region has very low correlation coefficients r during all seasons, although CEAS, Southeast Asia (SEAS), and Europe show even lower correlation coefficients for single seasons. The e^2 metric provides a

single error for each region and allows a relative ranking between them. The results shown in Figure 8 indicate that four regions (BONA, central America, CEAS, and SEAS) have relatively high errors, which is due to pronounced differences in these regions throughout the annual cycle. BONA, for instance, shows maximum e^2 in summer. During this season, the observed interannual albedo variability is small, resulting in relatively large values for e^2 compared to other regions.

5. Discussion and Conclusions

[29] The performance of vegetation-cover dynamics simulated by MPI-ESM is evaluated for present-day conditions. The patterns generated by the model

Table 4. Evaluation of Land-Surface Albedo on the Global Scale

Region	r^2				RMSE (W/m^2)			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
Global	0.99	0.86	0.90	0.98	11.1	12.3	12.6	7.4
Northern Hemisphere	0.88	0.84	0.89	0.97	9.2	14.3	14.3	6.7
Southern Hemisphere	1.00	0.81	0.85	0.99	14.4	5.2	5.4	8.7
Tropics	0.93	0.97	0.98	0.96	6.5	6.2	5.5	5.4

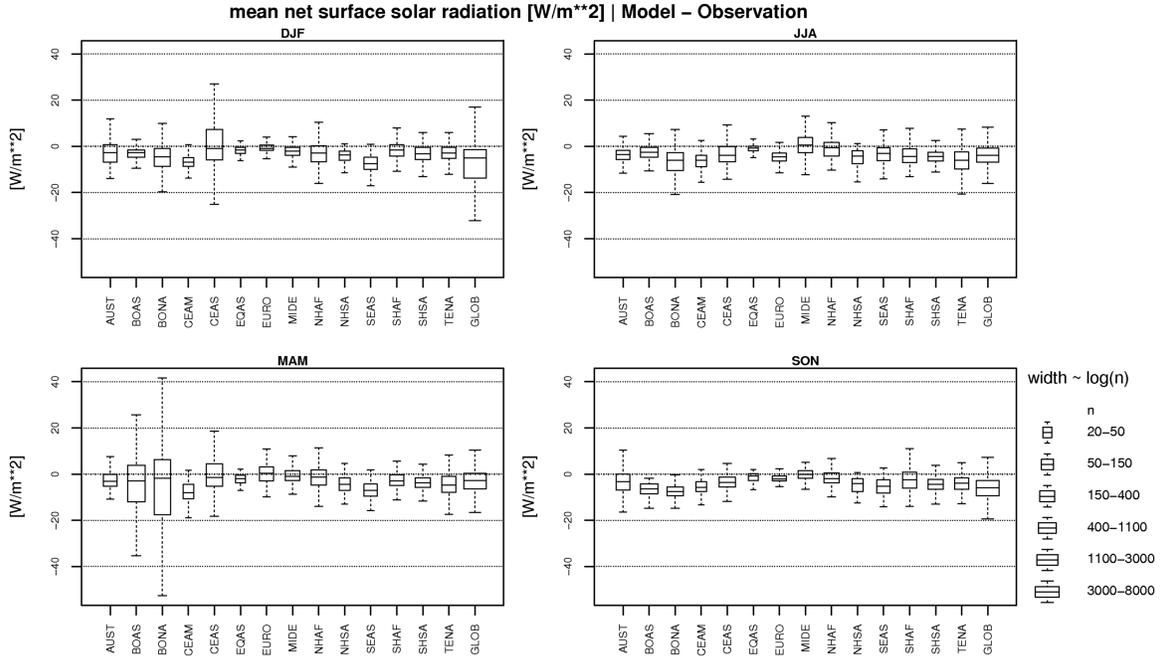


Figure 6. Box-and-whisker plots for the difference between model and observations in the net surface radiation flux Q_{SW} (W/m^2) for different regions. For season acronyms, see Figure 5; and for region acronyms, see Figure 4.

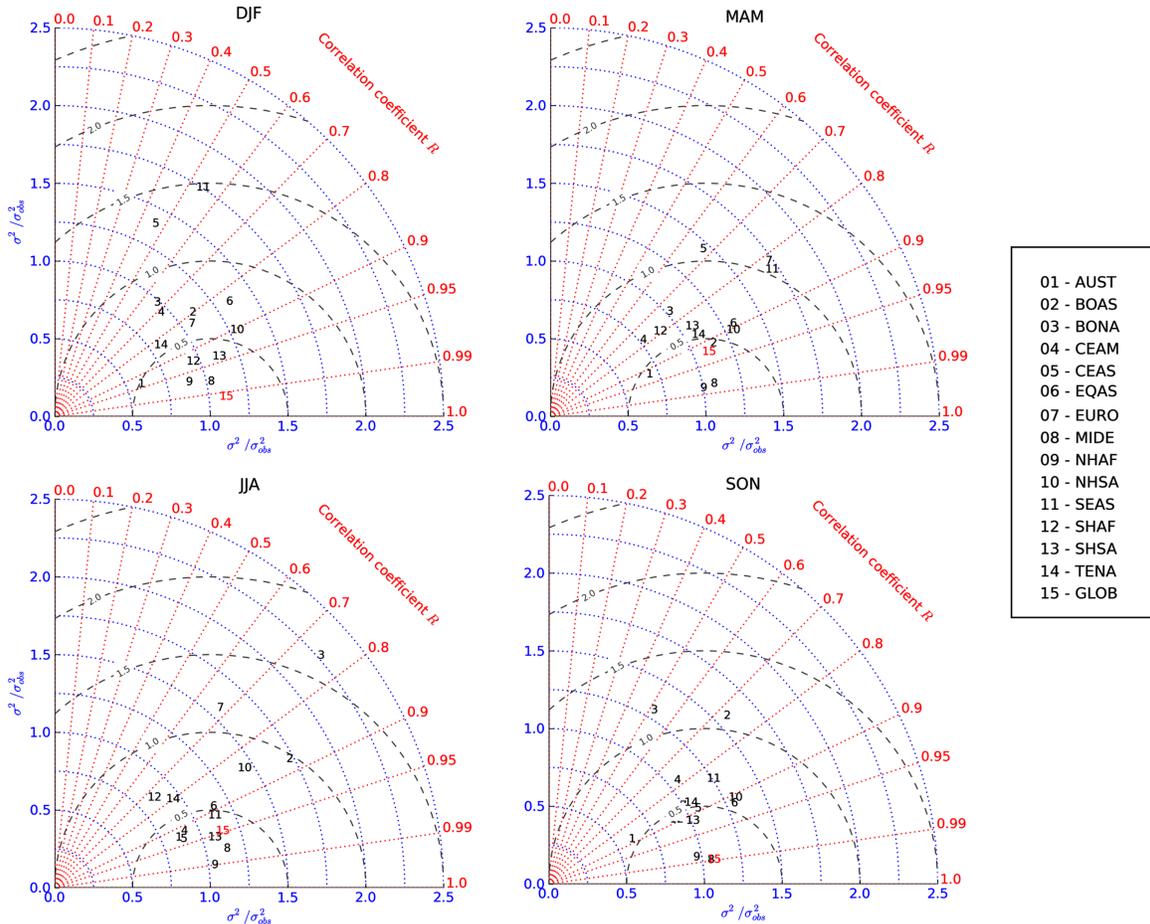


Figure 7. Taylor diagram of model-data agreement for the net surface radiation flux Q_{SW} (W/m^2) for different regions and seasons. The region and season acronyms are as in Figures 4 and 5, respectively. Global average is shown in red.

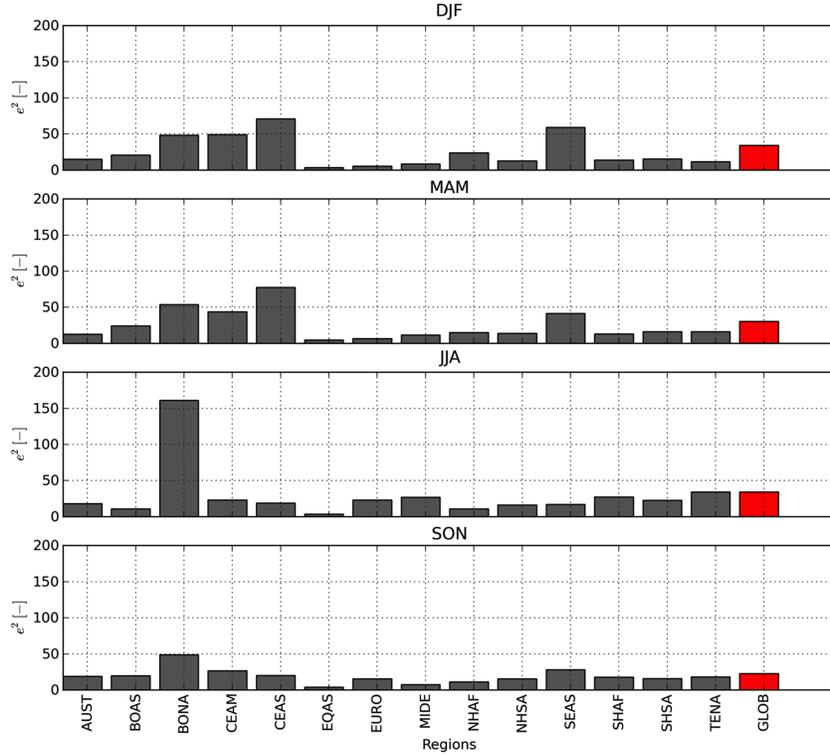


Figure 8. Normalized error (e^2) for the net surface radiation flux Q_{SW} (W/m^2) for different regions and seasons. The region and season acronyms are as in Figures 4 and 5, respectively. Global average is shown in red.

generally coincide well with the data; however, the model tends to overestimate tree cover and to underestimate bare ground. The model simulates tree cover and bare ground with r^2 higher for the Northern Hemisphere (0.66) than for the Southern Hemisphere (0.48–0.50). Grass cover, as an intermediate class between tree cover and bare ground, is reproduced less reliably in both Northern Hemisphere and Southern Hemisphere (r^2 of 0.44 and 0.17, respectively). In the vegetation model parameterization, we assume that trees are dominant over grasses. This simple approach partly explains the lower correlation for grass cover.

[30] The global-scale patterns of land-surface albedo simulated by the model and weighted by the incoming solar radiation match the observations well. In comparison with the vegetation cover, land-surface albedo generally shows higher correlation with observations. At the same time, the agreement between model and data differs depending on regions and seasons. The main disagreement between model and data occurs during boreal spring in the Northern Hemisphere, where biases in simulation of snow cover and tree cover cause substantial error and increased variance. In general, the model exhibits a slight overestimation of land-surface albedo, which is pronounced in all seasons. *Hagemann et al.* [2013] provide a more detailed assessment of MPI-ESM surface albedo and show that the CMIP5 version of MPI-ESM has considerably improved compared with the previous model versions.

[31] Vegetation cover is evaluated using continuous metrics (r^2 and RMSE). Another approach recently used by *Poulter et al.* [2011] for the evaluation of simulated vegetation cover is based on a β -diversity metric (mean Euclidean distance) that allows simultaneously accounting for more than one vegetation class. We have focused here on two classes, trees and bare ground, because they have the strongest effect on biophysical properties of land surface such as albedo and evapotranspiration. Vegetation models that use discrete vegetation classes could be evaluated using a discrete metric such as κ statistics [*Monserud and Leemans*, 1992]. The application of vegetation-cover metrics, therefore, depends on the model setup and on the research question.

[32] The metrics applied here are useful as relative measures of mismatches between model and observations. They quantify general performance of the model and are useful for regional and seasonal differences between model and observations. While we found that general patterns of albedo and vegetation cover are well represented by the model, we also identified regions where the model performance could be considerably improved. In addition, these metrics could be applied for benchmarking different models or model versions. In particular, such studies could help to identify differences between versions of the model driven by observed and simulated climatology and allow a deeper analysis of mechanisms behind mismatches between data and observations.

[33] **Acknowledgments.** This work was partly supported through the Cluster of Excellence “CliSAP” (EXC177), University of Hamburg, funded by the German Science Foundation (DFG). This study is contributed to the International Land Model Benchmarking Project (ILAMB). The authors thank Soenke Zaehle and two anonymous reviewers for their constructive comments.

References

- Abramowitz, G., R. Leuning, M. Clark, and A. Pitman (2008), Evaluating the performance of land surface models, *J. Clim.*, 21(21), 5468–5481.
- Armeth, A., et al. (2010), Terrestrial biogeochemical feedbacks in the climate system, *Nat. Geosci.*, 3(8), 525–532.
- Arora, V. (2002), Modeling vegetation as a dynamic component in soil-vegetation-atmosphere transfer schemes and hydrological models, *Rev. Geophys.*, 40(2), 1006, doi:10.1029/2001RG000103.
- Bathiany, S., M. Claussen, V. Brovkin, T. Raddatz, and V. Gayler (2010), Combined biogeophysical and biogeochemical effects of large-scale forest cover changes in the MPI earth system model, *Biogeosciences*, 7(5), 1383–1399.
- Blyth, E., D. B. Clark, R. Ellis, C. Huntingford, S. Los, M. Pryor, M. Best, and S. Sitch (2011), A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, *Geosci. Model Dev.*, 4(2), 255–269.
- Bonan, G. B. (2008), Forests and climate change: Forcings, feedbacks, and the climate benefits of forests, *Science*, 320(5882), 1444–1449.
- Brovkin, V., T. Raddatz, C. H. Reick, M. Claussen, and V. Gayler (2009), Global biogeophysical interactions between forest and climate, *Geophys. Res. Lett.*, 36, L07405, doi:10.1029/2009GL037543.
- Cadule, P., P. Friedlingstein, L. Bopp, S. Sitch, C. D. Jones, P. Ciais, S. L. Piao, and P. Peylin (2010), Benchmarking coupled climate-carbon models against long-term atmospheric CO₂ measurements, *Global Biogeochem. Cycles*, 24, GB2016, doi:10.1029/2009GB003556.
- Claussen, M., P. M. Cox, X. Zeng, P. Viterbo, A. C. M. Beljaars, R. A. Betts, H.-J. Bolle, T. Chase, and R. Koster (2004), The global climate, in *Vegetation, Water, Humans and the Climate, Part A*, edited by P. Kabat, et al., Springer, Heidelberg.
- Collins, W. J., et al. (2011), Development and evaluation of an Earth-System model—HadGEM2, *Geosci. Model Dev.*, 4(4), 1051–1075.
- Cramer, W., et al. (2001), Global response of terrestrial ecosystem structure and function to CO₂ and climate change: Results from six dynamic global vegetation models, *Global Change Biol.*, 7(4), 357–373.
- de Jong, R., J. Verbesselt, M. E. Schaepman, and S. de Bruin (2012), Trend changes in global greening and browning: Contribution of short-term trends to longer-term change, *Global Change Biol.*, 18(2), 642–655.
- Dickinson, R., A. Henderson-Sellers, and P. Kennedy (1993), Biosphere-Atmosphere Transfer Scheme (BATS) version 1e as coupled to the NCAR Community Climate Model, *NCAR Tech. Note NCAR/TN-387+STR*, Boulder, Colo.
- Essery, R., N. Rutter, J. Pomeroy, R. Baxter, M. Stahli, D. Gustafsson, A. Barr, P. Bartlett, and K. Elder (2009), An evaluation of forest snow process simulations, *Bull. Am. Meteorol. Soc.*, 90(8), 1120–1135.
- Fensholt, R., and S. R. Proud (2012), Evaluation of Earth Observation based global long term vegetation trends—Comparing GIMMS and MODIS global NDVI time series, *Remote Sens. Environ.*, 119, 131–147.
- Govaerts, Y., and A. Lattanzio (2008), Estimation of surface albedo increase during the eighties Sahel drought from Meteosat observations, *Global Planet. Change*, 64(3–4), 139–145.
- Hagemann, S., A. Loew, and A. Andersson (2013), Combined evaluation of MPI-ESM land surface water and energy fluxes, *J. Adv. Model. Earth Syst.*, doi:10.1029/2012MS000173, in press.
- Hansen, M., R. S. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. A. Sohlberg (2003), Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS Vegetation Continuous Fields algorithm, *Earth Interact.*, 7(10), 1–15.
- Hansen, M., R. DeFries, J. R. Townshend, M. Carroll, C. Dimiceli, and R. Sohlberg (2007), Vegetation Continuous Fields MOD44B, 2001 percent tree cover, collection 4, Univ. of Maryland, College Park, Md.
- Hurt, G. C., et al. (2011), Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, *Clim. Change*, 109(1–2), 117–161.
- Jungclaus, J. H., N. Keenlyside, M. Botzet, H. Haak, J. J. Luo, M. Latif, J. Marotzke, U. Mikolajewicz, and E. Roeckner (2006), Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM, *J. Clim.*, 19(16), 3952–3972.
- Knorr, W., K. G. Schnitzler, and Y. Govaerts (2001), The role of bright desert regions in shaping North African climate, *Geophys. Res. Lett.*, 28(18), 3489–3492.
- Liu, J. C., C. Schaaf, A. Strahler, Z. T. Jiao, Y. M. Shuai, Q. L. Zhang, M. Roman, J. A. Augustine, and E. G. Dutton (2009), Validation of moderate resolution imaging spectroradiometer (MODIS) albedo retrieval algorithm: Dependence of albedo on solar zenith angle, *J. Geophys. Res.—Atmos.*, 114, D01106, doi:10.1029/2008JD009969.
- Loew, A., and Y. Govaerts (2010), Towards multidecadal consistent Meteosat surface albedo time series, *Remote Sens.*, 2, 957–967.
- Monserud, R. A., and R. Leemans (1992), Comparing global vegetation maps with the kappa-statistic, *Ecol. Model.*, 62(4), 275–293.
- Myre, G., Y. Govaerts, J. M. Haywood, T. K. Berntsen, and A. Lattanzio (2005), Radiative effect of surface albedo change from biomass burning, *Geophys. Res. Lett.*, 32(20), L20812, doi:10.1029/2005GL022897.
- Otto, J., T. Raddatz, and M. Claussen (2011), Strength of forest-albedo feedback in mid-Holocene climate simulations, *Clim. Past*, 7, 1027–1039.
- Port, U., V. Brovkin, and M. Claussen (2012), The role of dynamic vegetation cover in future climate change, *Earth Syst. Dyn. Discuss.*, 3, 485–522.
- Poulter, B., P. Ciais, E. Hodson, H. Lischke, F. Maignan, S. Plummer, and N. E. Zimmermann (2011), Plant functional type mapping for earth system models, *Geosci. Model Dev.*, 4(4), 993–1010.
- Raddatz, T. J., C. H. Reick, W. Knorr, J. Kattge, E. Roeckner, R. Schnur, K. G. Schnitzler, P. Wetzol, and J. Jungclaus (2007), Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, *Clim. Dyn.*, 29(6), 565–574.
- Randerson, J. T., et al. (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, *Global Change Biol.*, 15(10), 2462–2484.
- Reichler, T., and J. Kim (2008), How well do coupled models simulate today’s climate?, *Bull. Am. Meteorol. Soc.*, 89(3), 303–312.
- Rietkerk, M., et al. (2011), Local ecosystem feedbacks and critical transitions in the climate, *Ecol. Complex.*, 8(3), 223–228.
- Schaaf, C. B., et al. (2002), First operational BRDF, albedo nadir reflectance products from MODIS, *Remote Sens. Environ.*, 83(1–2), 135–148.
- Schlosser, C. A., A. G. Slater, A. Robock, A. J. Pitman, K. Y. Vinnikov, A. Henderson-Sellers, N. A. Speranskaya, K. Mitchell, and P. D. Contributors (2000), Simulations of a boreal grassland hydrology at Valdai, Russia: PILPS phase 2(D), *Mon. Weather Rev.*, 128(2), 301–321.
- Sellers, P. J., et al. (1997), Modeling the exchanges of energy, water, and carbon between continents and the atmosphere, *Science*, 275(5299), 502–509.
- Sitch, S., et al. (2003), Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Global Change Biol.*, 9(2), 161–185.
- Slater, A. G., et al. (2001), The representation of snow in land surface schemes: Results from PILPS 2(d), *J. Hydrometeorol.*, 2(1), 7–25.
- Stevens, B., et al. (2013), The atmospheric component of the MPI-M earth system model: ECHAM6, *J. Adv. Model. Earth Syst.*, doi:10.1002/jame.20015, in press.
- Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.—Atmos.*, 106(D7), 7183–7192.
- van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, 6, 3423–3441.
- Watanabe, S., et al. (2011), MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments, *Geosci. Model Dev.*, 4(4), 845–872.

Corresponding author: V. Brovkin, Max Planck Institute for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany. (victor.brovkin@zmaw.de)