Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

Intercomparison of Terrestrial Carbon Fluxes and Carbon Use Efficiency Simulated by CMIP5 Earth System Models

Dongmin Kim[1], Myong-In Lee[1*], Su-Jong Jeong[2], Jungho Im[1], Dong Hyun Cha[1] and Sanggyun Lee[1]

[1]School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Korea

[2]School of Environmental Science and Engineering, Southern University of Science and Technology, Nanshan, Shenzhen, Guangdong, China

---

Corresponding author: Dr. Myong-In Lee
School of Urban and Environmental Engineering
Ulsan National Institute of Science and Technology,
50 UNIST-gil, Ulsan 44919, Korea
Email: milee@unist.ac.kr

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

Abstract

This study compares historical simulations of the terrestrial carbon cycle produced by 10 Earth System Models (ESMs) that participated in the fifth phase of the Coupled Model Intercomparison Project (CMIP5). Using MODIS satellite estimates, this study validates the simulation of gross primary production (GPP), net primary production (NPP), and carbon use efficiency (CUE), which depend on plant function types (PFTs). The models show noticeable deficiencies compared to the MODIS data in the simulation of the spatial patterns of GPP and NPP and large differences among the simulations, although the multi-model ensemble (MME) mean provides a realistic global mean value and spatial distributions. The larger model spreads in GPP and NPP compared to those of surface temperature and precipitation suggest that the differences among simulations in terms of the terrestrial carbon cycle are largely due to uncertainties in the parameterization of terrestrial carbon fluxes by vegetation. The models also exhibit large spatial differences in their simulated CUE values and at locations where the dominant PFT changes, primarily due to differences in the parameterizations. While the MME-simulated CUE values show a strong dependence on surface temperatures, the observed CUE values from MODIS show greater complexity, as well as non-linear sensitivity. This leads to the overall underestimation of CUE using most of the PFTs incorporated into current ESMs. The results of this comparison suggest that more careful and extensive validation is needed to improve the terrestrial carbon cycle in terms of ecosystem-level processes.

Keywords: earth system models, carbon use efficiency, CMIP5, MODIS, GPP, NPP

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

1. Introduction

Earth system models (ESMs) have been developed in the past several decades to simulate vegetation changes in space and time through carbon cycle-related interactions between the biosphere and the atmosphere. The temporal variations in atmospheric CO_2 in the models are driven by CO_2 emissions from natural and anthropogenic sources, as well as uptake by vegetated land surfaces and the ocean. Net imbalances in carbon fluxes drive the secular trend in CO_2. The magnitude of the imbalance is model-dependent and results in differences in the future warming projected by various ESMs. Previous studies showed that the observed trend of atmospheric CO_2 was not reproduced correctly during the past century, given the historical record. There was also substantial spread among models, even though they were forced by identical anthropogenic emissions (Friedlingstein et al., 2006, 2014; Hoffman et al., 2013; Zhao and Zeng, 2014). The model bias persists into their future projections. Hoffman et al. (2013) pointed out that the spread of projected CO_2 concentrations among fifteen Coupled Model Intercomparison Project (CMIP5; Taylor et al., 2012) ESMs in 2100 was approximately 20 % of their multi-model average. Friedlingstein et al. (2014) showed that the degree of surface temperature warming by 2100 was different by more than a factor of two, depending on the models and representative concentration pathway (RCP) 8.5 scenarios used.

Previous studies (Friedlingstein et al., 2006; Booth et al., 2012; Hoffman et al., 2013; Anav et al., 2013; Aroa et al., 2013; Friedlingstein et al., 2014) have suggested that the uncertainty in CO_2 concentrations simulated by ESMs should be largely attributed to the terrestrial carbon uptake, rather than to the uptake by ocean. Hoffman et al. (2013) and Friedlingstein et al. (2014) compared the carbon uptake by land and ocean, simulated by ESMs and found that the amount of carbon accumulated by the ocean is positive in all models by 2100, whereas the models exhibited a large spread in the amount of carbon taken up by the land; the results even had

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

different signs. Aroa et al. (2013) indicated that the simulated sensitivity of terrestrial carbon storage to the atmospheric CO_2 concentration was 3-4 times larger than that of ocean. This suggests that the terrestrial carbon cycle is one of the important factors that need improvement for minimizing uncertainty in future climate predictions.

It is generally recognized that changes in the carbon pools in the biosphere should play a key role in determining atmospheric CO_2 concentration levels in the future. Shao et al. (2013) showed that the net biome production (NBP) simulated by CMIP5 ESMs is enhanced in the 21st century and that the biomass particularly increases over tropical rainforests and vegetated surfaces in the mid-latitudes through the CO_2 fertilization effect. Not only long-term increases in biomass but also future changes in its seasonal cycle would significantly affect CO_2 concentrations. Zhao and Zeng (2014) indicated that the amplitude of the seasonal cycle of atmospheric CO_2 tends to increase in the future, due to an increase of 68 % in the seasonal cycle of NBP during the growing season in their future simulations. Comprehensive model intercomparisons on the simulation of biome production at various ecosystem levels are needed to explain the differences among simulations and minimize projection uncertainties.

The exchange of carbon between the atmosphere and terrestrial ecosystems consists of complicated biogeochemical processes operating over a heterogeneous surface, and the quality and the performance of the global model simulations is often diagnosed using carbon cycle variables such as gross primary production (GPP) and autotrophic respiration (Ra) by plants. Net primary production (NPP) is defined as GPP minus Ra. Heterotrophic respiration (Rh), involving the decomposition of soil litter, is also an important process involved in the carbon cycle. By validation using ground and satellite observational data, previous studies identified the systematic biases of ESMs and discussed the possible reasons for these biases. Anav et al. (2013) indicated that current ESMs tend to overestimate terrestrial biomass and global GPP

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

(Anav et al., 2013). Shao et al. (2013) showed that ESMs exhibit large disagreements in the relationship between carbon cycle variables and hydrological variables, such as precipitation and soil moisture, emphasizing the importance of the hydrological cycle in terms of its effects on the terrestrial carbon cycle. The simulated soil carbon amount in the subsurface root zone, which is the major source of plant growth, showed systematic biases and large model spread, from 40 to 240 %, compared with observational data (Todd-Brown et al., 2013). That study suggested that it might be responsible for the large spread of atmospheric CO2 concentrations simulated by the models.

While most previous intercomparison studies involving ESMs have focused on the validation of the global mean budget of terrestrial carbon pools and fluxes (Anav et al., 2013; Shao et al., 2013; Todd-Brown et al., 2013), which is useful for evaluating the overall performance of ESMs and quantifying simulation uncertainties, more detailed analyses addressing regional scales and different vegetation types are needed to identify the key sources of systematic biases in the models. Anav et al. (2013) evaluated regional changes in biogeochemical variables for two hemispheres and the tropical region separately. In particular, an investigation of systematic biases in different types of ecosystems is required to improve the existing parameterizations of terrestrial carbon fluxes by vegetation. In contrast to the many observational studies in biology that address various plant function type (PFT) levels (De Lucia et al., 2007; Zhang et al., 2009; Zhang et al., 2014), studies that benchmark model simulations of PFT levels have obtained less attention, and this is one of the primary motivations of this study.

For a better elucidation of systematic biases in the models, this study focuses particularly on the comparison of carbon use efficiency (CUE), which is sensitive to the various PFTs. For the short-term carbon cycle, Ra is a primary measure of the release of carbon to the atmosphere,

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

and its magnitude is known to be about half of GPP for most vegetated surfaces (King et al., 2006; Piao et al., 2010). CUE is defined as the ratio of NPP to GPP, which is a useful diagnostic measure for the comparison of parameterizations for the terrestrial carbon fluxes driven by vegetation that are implemented differently in current ESMs. The absolute magnitudes of the production terms are the results of feedbacks between climate and vegetation. Normalized flux terms can highlight the differences among simulations driven by parameterization differences in terrestrial carbon fluxes. Previous studies based on in situ (De Lucia et al., 2007) and satellite (Zhang et al., 2009) data analyses have indicated that CUE is not a constant with a value of approximately 0.47 (Gifford, 1994; Dewar et al., 1999) but varies depending on climatic conditions and PFTs. In this regard, the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data provide the global coverage of GPP and NPP as a useful reference for the model validation for CUE at the PFT level. Zhang el al. (2014) suggested observed CUE by MODIS tends to slightly increase in the recent years.

The purpose of this study is the intercomparison of CMIP5 ESMs in terms of their simulations of the terrestrial carbon cycle, based on a quantitative evaluation of the performance of terrestrial carbon flux parameterizations in their land surface models (LSM). This analysis specifically focuses on the assessment of CUE at the PFT level and makes an effort to provide useful suggestions to the modeling community for reducing systematic biases in the terrestrial carbon cycle in current ESMs. This study consists of following sections: Section 2 describes the observational data and model output used in this study. Section 3 compares the model simulations in terms of their climate and terrestrial carbon cycle variables, comparing first the multi-model ensemble (MME) average to diagnose common and systematic biases in the current models and then identifies differences among simulations across the ESMs in their simulated climates and carbon fluxes. The comparison of CUE at various PFT levels is

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

169 followed by more comprehensive comparisons for identifying differences among simulations
170 driven by model parameterizations. Finally, Section 4 provides a summary and conclusions.
171
172 **2. Data and Analysis Methods**
173 **2.1 Observational data**
174 This study used GPP and NPP as primary variables to validate the global carbon cycle as
175 simulated by various ESMs. Reference observational data were obtained from the NASA
176 MODIS MOD17 data product, which includes the first satellite-driven estimates of carbon
177 fluxes on vegetated surfaces on a global scale (Running and Gower, 1991; Zhao et al., 2005).
178 The MODIS algorithm uses a data model based on the radiation use efficiency logic of
179 Monteith (1972) to estimate GPP, which is basically a linear function of the amount of
180 Photosynthetically Active Radiation (PAR) absorbed. The fraction of PAR and the leaf area
181 index (LAI) are provided to the model by the MODIS MOD15 products. A conversion
182 efficiency parameter relating absorbed radiation to the actual productivity depends on
183 vegetation type and climate condition. The upper limit of conversion efficiency uses the Biome
184 Parameter Lookup Table (BPLUT) for different vegetation types. The vegetation types include
185 evergreen needleleaf forest (ENF), evergreen broadleaf forest (EBF), deciduous needleleaf
186 forest (DNF), deciduous broadleaf forest (DBF), mixed forests (MF), open and closed
187 shrublands (SHR), grasslands (GRA), and croplands (CROP), which are based on the land
188 cover classification from the MODIS MCD12Q1
189 (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q1). Figure 1
190 shows the horizontal distribution of vegetation types from MODIS. The conversion efficiency
191 is modified by climate conditions such as incoming solar radiation, temperature, and vapor
192 pressure deficit, which are obtained from atmospheric reanalyses developed by NASA's Global

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

193 Modeling and Assimilation Office and the NCEP/NCAR Reanalysis II. The NPP estimation
194 by MODIS calculates daily leaf and fine root maintenance respiration, annual growth
195 respiration, and annual maintenance respiration of live cells in woody tissue, which are
196 subtracted from the GPP. Biome-specific physiological parameters are also specified by
197 BPLUT for respiration calculations.

198 The MOD17 dataset provides 8-day, monthly, and annual mean GPP and NPP for 2000-2012.
199 This study used the gridded GPP and NPP products, which have a spatial resolution of 30
200 arcsec (0.0083 degree), provided by the Numerical Terradynamic Simulation Group (NTSG)
201 of the University of Montana (NTSG MOD17 v55).

202 Although MODIS is affected by uncertainties in biomass types and meteorological data sets
203 (Zhao et al. 2005), the derived GPP and NPP values are able to capture realistic spatial and
204 temporal variations over different biomes and climate regimes. Zhao et al. (2005) and Heinsch
205 et al. (2006) demonstrated that the data are consistent with ground-based flux tower
206 measurements of GPP and field-observed NPP estimates with high correlation (r=0.859).

207 For comparison with MODIS, this study also used GPP estimates from FLUXNET-MTE
208 (Multi-Tree Ensemble; Jung et al., 2011), which is an upscaled data set providing global
209 coverage that is derived from 178 surface flux tower observations using a machine learning
210 technique. FLUXNET-MTE provides an explicit estimate of carbon fluxes over vegetated
211 surfaces. The dataset provides monthly data at a 0.5° × 0.5° (latitude × longitude) spatial
212 resolution and covers the period 1982 – 2007. Although this gridded global dataset is useful
213 for validation of ESMs, its key limitations are also discussed in the literature (Jung et al., 2011).
214 Wide geographical regions are not represented by measurement stations; for example, there is
215 a lack of samples over Siberia, Africa, South America and tropical Asia compared with North
216 America and Europe. Estimates of annual-mean upscaled ecosystem respiration have higher

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

certainty than the anomalies and show approximately 5-10 % underestimation. Additionally, the data have limitations in accounting for disturbances due to land use changes, given that unchanged land cover data from the International Geosphere-Biosphere Program (IGBP) satellite are used for all periods. This may introduce spurious trends into the GPP estimates from the FLUXNET-MTE project. The dataset does not provide estimates of Ra, but instead provides the summation of Ra and Rh. The geographical distribution of satellite-derived GPP from MODIS shows a high degree of consistency with that from in situ FLUXNET observations. Figure 2 compares the annual GPP distributions from MODIS and FLUXNET for the same period, 2000-2005. A notable difference between the two appears in the Amazon, where MODIS tends to underestimate the productivity significantly. In the remaining regions, MODIS tends to produce slight underestimates in the tropics and overestimates in the high latitudes when compared with FLUXNET. The annual GPP values from MODIS and FLUXNET are 108.76 GtC and 107.41 GtC, respectively, for the averaging period of 2000-2005, with a small difference that is no more than 1 % of the total value. The pattern of differences did not change significantly even if the FLUXNET data were averaged over a longer period (1983-2005). In fact, the interannual variation did not modify the global-mean annual GPP value significantly when the reference period was extended to 1983-2005, which yielded a small reduction to 106.55 GtC using the FLUXNET data.

This study also used the observed surface air temperature and precipitation data from the Institute for Climate Impact Research based on the CRU (Climate Research Unit) meteorological dataset (Harris et al., 2014). In this data product, temperature and precipitation at stations worldwide were interpolated to a horizontal resolution of 0.5° × 0.5° (latitude × longitude) covering the global land surface.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

240

241 **2.2 Model Data**

242 Historical simulations performed using 10 ESMs were used in this study. Brief descriptions
243 of these models is provided in Table 1. The historical simulations (that is, experiment 5.2 or
244 the ESM historical 1850–2005 simulation; Taylor et al., 2012) were forced by gridded CO_2
245 emissions data for fossil fuel consumption from Andres et al. (2011). While conventional CO_2
246 concentration-driven runs have no vegetation feedback on atmospheric CO_2, these emissions-
247 driven runs enables climate-carbon cycle feedbacks via changes in vegetation. Note that three
248 models – GFDL-ESM2M, GFDL-ESM2G, and MPI-ESM LR – of them enabled the dynamic
249 vegetation model in their historical simulations for 1850 – 2005, which model was able to
250 consider dynamic change of PFT boundaries by climate conditions (Table 1). Atmospheric CO_2
251 concentrations are simulated prognostically from the net budget of natural and anthropogenic
252 carbon fluxes to and from the atmosphere. The simulation of GPP is directly controlled by the
253 formulae representing photosynthesis in the models. As shown in Table 1, the parameterization
254 of photosynthesis by vegetation is formulated similarly in the 10 ESMs. This parameterization
255 is mostly based on Farquhar et al. (1980) for C3 plants in cold climates, with revisions for C4
256 plants in warm climates by Collatz et al. (1992). Leaf photosynthesis in CLM4 is proportional
257 to the concentration of carbon dioxide in the atmosphere, as well as the temperature and
258 moisture surrounding leaves. It adjusted the minimum rate among the light-use, water-use and
259 carbon assimilation approaches in CLM4.

260 NPP is diagnosed in ESMs by subtracting Ra from GPP. Parameterizations for Ra are more
261 diverse in formulation across the models compared to that of photosynthesis. Note that
262 CESM1-BGC and NorESM-ME1 incorporate identical land surface models, in which the
263 nitrogen cycle is allowed to limit plant assimilation for the parameterization of carbon fluxes

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

by terrestrial vegetation, so called the interactive carbon-nitrogen (CN) cycle. Respiration is proportional to temperature and nitrogen concentration. The models without interactive nitrogen cycles diagnose nitrogen concentrations from the carbon concentration in each carbon pool, whereas the models with interactive nitrogen cycles predict the nitrogen concentrations. The only exception is MRI-ESM, which uses an empirical formula for estimating NPP based on Obata (2007). In the model, the monthly NPP is empirically derived from physical variables such as temperature and precipitation from the Miami model (Lieth, 1975; Friedlingstin et al., 1995).

The model data were obtained from the Earth System Grid Federation (ESGF), an international network of distributed climate data servers (Williams et al., 2011). For the purposes of comparison, the model outputs, as well as the MODIS data, were interpolated onto the same 1° × 1° grid (latitude × longitude).

2.3 Analysis Methods

In Section 3.3, CUE is diagnosed at the ecosystem level for the MODIS observations and the various ESM simulations. For simplicity, an identical distribution of vegetated surfaces based on to the MODIS classification (Figure 1) was applied to both the observed and the simulated fluxes. This is because each model has their own vegetation classifications, which are not available from the CMIP5 data archive.

It is noted that the deficiency in the simulation of CUE by individual models is not only caused by deficiencies in the parameterization of carbon fluxes due to vegetation but also by differences in the classifications of PFTs, which are specified differently in each model. For example, LM3.0 in GFDL ESM2 M and ESM2G simulate 5 PFTs (i.e., 3 types of trees and 2 types of grasses), while NCAR and NorESM's CLM4.0 specifies the PFTs in much greater

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

288 detail by including 17 different types (i.e., 8 types of trees, 3 types of shrubs, 3 types of grasses
289 and 3 types of crops). Although referencing PFTs from the observations instead of using own
290 PFTs in each model might not be a perfect comparison, it is still meaningful to identify the first
291 order differences driven by parameterization method and the classification difference as well
292 where the latter is regarded as the model bias too.
293
294
295 **3. Results**
296 **3.1. Systematic Biases in the Multi-Model Ensemble**
297 Systematic biases in the ESM simulations are examined first by taking multi-model
298 ensemble averages (MME) for simulated surface air temperature and precipitation, respectively
299 (Figure 3). Despite the realistic representation of annual-mean surface temperatures, MME
300 exhibits systematic biases with significant hemispheric differences. Warm biases are seen in
301 the Northern Hemisphere, particularly in northeastern Asia and North America, whereas there
302 exists a cold bias in most of the Southern Hemisphere. MME generally shows wet biases in
303 precipitation, except over South America. Wet biases seem to be consistent with cold biases in
304 the tropical regions, where the deep convective rainfall tends to produce deep clouds that
305 attenuate incoming solar radiation at the surface.
306 The annual GPP, NPP and Ra values from the MODIS observations and the MME are
307 compared in Figure 4. The observed GPP values from MODIS are generally high in areas of
308 EBF in tropical regions, such as Amazon, South Asia, and Central Africa, and in areas of DBF,
309 such as those in Indochina, China, India, Europe and the southeastern part of North America.
310 GPP is observed to be small in areas of SHR in Australia and in boreal regions of MF and GRA
311 in northern Eurasia. GPP is close to zero over dry and non-vegetated surfaces, such as the

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

312 Sahara Desert and central Australia. The MME of the ESMs tends to reproduce these
313 geographical differences realistically, although the estimated magnitudes are too large over
314 most of the globe. Although Ra tends to be overestimated as well, MME shows a net positive
315 bias in NPP in most terrestrial regions, suggesting that the MME should underestimate the
316 observed trend of atmospheric CO_2 increase.

317 The global-mean values of GPP, NPP, and Ra are compared in Figure 5. Note that spread of
318 the simulations is large, particularly due to the outlier value produced by MRI-ESM1. The
319 median value of GPP simulated by ESMs is centered slightly above the value from MODIS
320 and is approximately 20 % higher (+18 GtC). The median value of NPP is also overestimated
321 by 10.2 GtC compared with the 52.1 GtC NPP from MODIS. The median value of Ra is
322 underestimated.

323 The formulations of GPP and Ra are closely related to temperature and precipitation
324 (Rahman et al., 2005; Yang et al., 2006), and, the model biases in those carbon fluxes might be
325 driven both by systematic biases in climate conditions such as temperature and precipitation
326 and the uncertainty in the parameterization formulations themselves. The Taylor diagram is a
327 common and useful measure for simulated spatial distributions that calculates spatial
328 correlation coefficients between observed and simulated values and the normalized standard
329 deviation of simulated values from the global mean over the whole domain of comparison.
330 Figures 6a and 6b show Taylor diagrams for the annual mean surface air temperature and
331 precipitation, respectively. The MME simulation of temperature by the CMIP5 ESMs is quite
332 close to the CRU observations. The spatial correlations are greater than 0.95 in all models. The
333 normalized standard deviations are within the range of 0.8 to 1.5, which is relatively small
334 compared with other simulated variables. The Taylor diagram of precipitation shows less
335 accuracy and more model spread than that of SATs. The spatial correlation of the MME is

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

336 approximately 0.76; the MME also shows higher normalized standard deviations compared
337 with temperature, suggesting that current ESMs exhibit relatively larger discrepancies in
338 precipitation and the terrestrial water cycle. Spatial patterns of GPP simulated by the ESMs
339 (Figure 6c) show even larger systematic biases with lower spatial correlations and larger spatial
340 changes (i.e., higher normalized standard deviations) than the observed values. Model spread
341 becomes much larger than that of temperature and precipitation. The simulated pattern
342 correlations from the ESMs are lowest for NPP (Figure 6d). The correlation for the MME is
343 slightly higher than 0.5. The models also exhibit much higher spatial variation than the
344 observed values for both GPP and NPP.

345 The Taylor diagram analysis suggests that the systematic biases in the ESMs may be
346 successively amplified by deficiencies in the simulation of climate and the terrestrial carbon
347 cycle. Regarding the climate conditions that affect the terrestrial carbon cycle, particularly the
348 distribution of precipitation and the water cycle seem to contribute more to the bias than does
349 temperature. In addition, the much larger spread in GPP and NPP simulated by the ESMs
350 compared to that in temperature and precipitation suggests that there should be much larger
351 uncertainty in the parameterization of terrestrial carbon cycle in the current ESMs. Biases and
352 model spread are even larger in NPP compared with GPP, implying that the simulation
353 uncertainty is much larger when the photosynthesis and the respiration are combined. The
354 performance of the MME in terms of GPP and NPP is not necessarily higher than that of the
355 individual models in this case, due to the presence of persistent and large deficiencies in the
356 individual models.

357

358 **3.2. Model Dependences**

359 The simulation of annual GPP values shows significant model dependence as shown in

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

Figure 5. MRI-ESM1 shows the largest value among the models. The three models, ESM2G, ESM2 M, and MPI-ESM-LR, simulate relatively larger values of GPP than the rest of the models. As the simulation of Ra shows relatively small model dependence, models that simulate larger GPP values tend to produce larger NPP in general. MRI-ESM1 is an exception, and the simulated GPP of this model is significantly reduced by its large Ra, leading to an NPP value close to the median value. The two models, CESM1-BGC and NorESM1-ME, that share the same land surface model simulate the smallest NPP values, which is a significant underestimation relative to the MODIS estimate.

To examine further what causes the global bias in carbon fluxes, the spatial distribution of the GPP bias pattern in carbon fluxes simulated using each model is compared in Figure 7. Each model exhibits its own systematic biases. MRI-ESM1 shows a significant positive bias in most vegetated regions, which is particularly pronounced in tropical rainforests. The group of models with higher global-mean GPP values in Figure 5 (i.e., MPI-ESM1-LR, ESM2 M, and ESM2G) shows GPP bias patterns that are remarkably similar to each other. GPP is overestimated in most regions in these models except for the upper inland region of the Amazon. The rest of the models show mixed spatial patterns of positive and negative biases. The large negative GPP bias in part of the Amazon is primarily responsible for the lowest global-mean GPP values, which are simulated by CanESM2 and BCC_CSM1 M. The negative bias is clear in the boreal high-latitude regions above 40 N in the CESM1-BGC and NorESM1-ME models. The systematic biases in the models reflect the uncertainties in the parameterized carbon cycles, as well as in the simulated climates. Most models simulate larger production in the tropics, due to abundant rainfall and high temperatures, and smaller production in high latitudes due to less precipitation and low temperatures. As GPP is much larger in magnitude than Ra, the NPP bias pattern in each model is mostly dominated by that of GPP rather than Ra, leading to consistent

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

patterns (cf. Figure 7 and Figure 8). The two GFDL models implemented with the same LM3 land surface model (i.e., ESM2M and ESM2G) and the other two models that use CLM4 (CESM1-BGC and NorESM1-ME) show NPP biases with opposite signs in the boreal regions above 40 N, highlighting significant model differences in parameterizations of carbon fluxes due to vegetation.

3.3. Carbon Use Efficiency

The bias patterns of GPP and NPP simulated by the various ESMs presented in Figure 7 and 8 are the result of complicated feedbacks between the carbon cycle (mostly by terrestrial vegetation) and climate. As the magnitude of the bias is also a function of biomass, this study further compared carbon use efficiency by dividing NPP by GPP. This normalized carbon flux ratio can highlight the difference among simulations driven by parameterization differences in terrestrial carbon fluxes by vegetation. The spatial pattern of CUE obtained by MODIS shows significant variations (Figure 9). In MODIS, most tropical areas with high GPP values generally show low CUE values below 0.4, particularly over the Amazon, central Africa and Southeast Asia. In contrast, CUE is in general greater than 0.5 over wide areas in high latitudes and a few low-latitude, high-elevation regions. The spatial distribution of CUE apparently depends on climate conditions such as precipitation and temperature in that regions with large amounts precipitation and warm climates show low CUE values, while regions experiencing dry and cold climates show high CUE values. Overall, the MME of 10 ESMs tends to reproduce the observed distribution from MODIS reasonably well. However, the MME values are lower than the observed values in most regions, which can largely be attributed to the underestimation of CUE values by MRI-ESM1. The bias pattern of CUE differs strongly among the models. Note that the bias pattern of CUE tends to characterize the parameterization differences in the

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

408 terrestrial carbon fluxes used in the ESMs. The bias patterns of CUE are almost identical to
409 each other for models that share the same land surface model, such as BCC_CSM1 and
410 BCC_CMS1 M, and ESM2 M and ESM2G, and CESM1-BGC and NorESM1-ME,
411 respectively. The two BCC models tend to overestimate CUE in Eurasia, North America, and
412 Africa, while they produce underestimates in Australia and South America. CanESM2 shows
413 a similar pattern as the two BCC models. MPI-ESM1-LR shows a similar bias structure except
414 in that it produces overestimates in South America. CESM1-BGC, NorESM1-ME, and MRI-
415 ESM1 exhibit an underestimation of CUE over most terrestrial regions.

416 The model dependence is depicted better by the zonal mean CUE distribution (Figure 10).
417 The observed CUE values show a clear latitudinal dependence and generally increases with
418 latitude. The zonal mean of CUE from MODIS ranges from 0.3 to 0.7, with a global average
419 of 0.49. It indicates that the biomass in high latitudes tends to take up atmospheric carbon more
420 efficiently compared with that in tropics. Even though the model spread is larger, the zonal
421 mean MME is able to reproduce the observed relationship between CUE and latitude. Some
422 models, such as CESM1-BGC, NorESM1-ME and MRI-ESM1, are notably different from the
423 other models, as well as from MODIS, and simulate low values, particularly at middle to high
424 latitudes. These results are consistent with those in Shao et al. (2013). They suggested that
425 respiration decreases more rapidly than production in response to latitudinal decreases in mean
426 temperature in all models expect NorESM1-ME and CESM1-BGC. The reason for the
427 underestimation of CUE in the two models are caused by their low estimates of NPP. Using the
428 same data from MODIS, Zhang et al. (2009) suggested that there exists a clear relationship
429 between CUE and climate conditions, such as surface air temperature and precipitation, that
430 are critical for biomass growth.

431 Figure 11 compares the relationship from MODIS with the model simulations. The observed

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

CUE from MODIS is more influenced by temperature than precipitation, as is particularly clear in dry regions with precipitation below 50 mm yr^{-1}. In general, the observed CUE decreases with increasing temperature. Moreover, observed CUE values show the sensitivity of CUE to precipitation in the tropics, where plant growth is more sensitive to precipitation compared with high latitudes. The MME basically follows this temperature sensitivity, although it tends to underestimate CUE. It is caused by the overestimation of Ra in most models compared with the MODIS estimates (Figure S3). Individual models show their own deficiencies. For example, the GFDL models (ESM2 M and ESM2G) tend to overestimate the sensitivity of CUE to precipitation in tropical regions compared with MODIS. It indicates that the gradients in CUE with temperature in the GFDL models are weaker than those in MODIS. In contrast, the models based on CLM4.0, such as CESM1-BGC, NorESM1-ME and MRI-ESM1, show a weaker sensitivity of CUE to both temperature and precipitation than the other models. This result might be caused by other limiting and trigger processes, such as nitrogen limitation, which are larger than the sensitivity to temperature and precipitation. This large divergence in the model sensitivity of CUE to temperature and precipitation induces differences in the atmospheric CO2 concentrations in the future among the full coupled ESMs.

Figure 12 compares the observed values and differences among simulations in terms of CUE depending on the dominant PFTs according to the classification in Figure 1. In the MODIS observations, the CUE values over broadleaf forests (DBF and EBF) are generally lower than over needleleaf forests (DNF and ENF), implying that dense forests tend to not only take up large amounts of atmospheric carbon for photosynthesis but also release large amounts of carbon to the atmosphere though respiration. In this regard, the efficiency of carbon uptake by the broadleaf forests is smaller than that of needleleaf forests.

The observed variations in CUE depending on the PFTs are reproduced realistically by the

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

456 MME. The differences between MODIS and the MME is large in areas of DNF and DBF, but
457 those vegetation types occupy relatively small fractions of the vegetated surface. The model
458 spread is large, regardless of plant function types. This is primarily due to the low CUE values
459 produced by three of the models, CESM1-BGC, MRI-ESM1 and NorESM1-ME, for all of the
460 plant function types. These three ESMs have their own unique formulations in parameterizing
461 terrestrial carbon fluxes. In the case of MRI-ESM1, it determines the monthly Ra empirically
462 based on a function of the surface air temperature and precipitation (Obata, 2007). The
463 simulated NPP in MRI-ESM1 is the residual term between GPP and Ra that is evidently
464 different from that of the other ESMs. The two CLM 4.0-based models, CESM1-BGC and
465 NorESM1-ME, include coupled carbon and nitrogen (CN) cycles, which seems to lead to
466 dramatic differences in CUE compared with the other models that do not represent interactions
467 between the carbon and nitrogen cycles. Inclusion of the nitrogen cycle in the models tends to
468 constrain the amount of carbon uptake in vegetated land surface (Zaehle et al., 2010;
469 Friedlingstein et al., 2014) and produces higher simulated growth respiration than in other
470 models (Shao et al., 2013).

471 To examine the impact of the CN cycle in the model further, this study conducted two
472 additional sensitivity experiments using CESM1-BGC, one with interactive carbon-nitrogen
473 cycle (CN) and the other with no nitrogen cycle (Only C). Figure 13 shows that CN tends to
474 decrease GPP in most of areas compared with Only C, whish suggests that the implementation
475 of nitrogen cycle in this model reduces the amount of carbon uptake by vegetation drastically
476 as a limiting factor. Accordingly NPP also tends to decrease in most of the regions at the
477 decrease of GPP. It is interesting to see that CUE decrease is particularly significant in mid- to
478 high-latitudes rather than in the tropics. This result is quite consistent with the simulation
479 difference between the CN models (CESM1-BGC and NorESM1-ME) and the rest of ESMs

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

(e.g., the zonal mean CUE shown in Figure 10).

This study further compares the observed and the MME-simulated CUE sensitivity to the surface temperature for each plant function type (Figure 14). The MODIS observations show more scatter in CUE values for a given temperature, suggesting that the natural carbon cycle is not simply determined by temperature, but is also controlled by other factors. In most PFTs, the observed CUE is maintained close to or even higher than 0.6, particularly in low canopy plants such as SHR, CROP and GRA, for surface temperatures lower than 10 ℃. CUE tends to decrease significantly at temperatures higher than 10 ℃. This observed feature may be interpreted based on the ecological significance of the resistance to low temperatures by plants (Allen et al., 2010). Low temperatures tend to reduce biosynthetic production by plants and can even disturb vital functions to cause permanent injuries and death. The survival capacity of plants tries to make its metabolic processes continue to function under low temperature stresses and using its cold resistance (Larcher, 1968). It suggests that the CUE values of vegetation may be lowered in favorable environmental conditions, such as warm temperatures and abundant precipitation, as there is plenty of production and plant growth. Vegetation experiencing cold temperatures and insufficient precipitation adapts to survive by increasing CUE.

In contrast, even though the multi-model ensemble average is taken for the various ESMs, the simulated CUE variation shows a clearer change with temperature, suggesting that the parameterization of the terrestrial carbon cycle in current ESMs depends too much on temperature conditions. A decreasing trend is clear in the MME regardless of PFTs in response to an increase in temperature. From the MME simulation results, CUE values in all PFTs shows a clear linear change in response to temperature variation. This implies that the current models

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

503 do not adequately consider the observed ecological resistance to temperature, and the balance
504 between respiration and production in the models is more simplified than the observations. In
505 fact, the parameterizations of most land surface models are based on conceptual leaf-level
506 formulations, such as those used in the calculation of biochemical photosynthesis processes
507 and the dependence of CO_2 exchange on stomatal conductance, which use temperature and soil
508 moisture explicitly in their formulations. The comparison results in this study suggest that the
509 models might need to consider ecosystem-level parameterizations which simulate carbon and
510 nitrogen fluxes and vegetation and soil pools and are estimated at a long (e.g., monthly) time
511 step based on spatially explicit information on climate, ecosystem type, soil type, and elevation
512 (Zhu and Zhuang, 2015) to reflect the nonlinear relationship for the interaction between climate
513 condition and vegetation.
514

515 **4. Summary and Concluding Remarks**

516 The simulations of climate and the terrestrial carbon cycle have been examined by comparing
517 surface temperatures and precipitation, as well as GPP, Ra, and NPP values, simulated by 10
518 different CMIP5 ESMs with the CRU surface observational data for climate-related variables
519 and the MODIS satellite estimates for the carbon cycle over 6 years (2000-2005).

520 Despite the systematic biases with significant hemispheric differences, the spatial
521 distributions of temperature and precipitation, which are closely related to biogeochemical
522 variables (Rahman et al., 2005; Yang et al., 2006), are relatively similar when compared with
523 observations. More model discrepancies appeared in the simulation of the carbon cycle, which
524 reflects overestimation of GPP over most of the globe. The terrestrial carbon fluxes simulated
525 by the ESMs are diverse, and the models exhibit large spread, even though the multi-model
526 ensemble mean (MME) shows strong resemblance in terms of its spatial distribution to the

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

observed pattern by cancelling out the systematic biases in each model. The results show that the biases of terrestrial carbon fluxes are due less to the bias in the spatial distribution of climate conditions but more to the larger uncertainty in their parameterizations.

We also analyzed carbon use efficiency (CUE) by dividing NPP by GPP, which is a physiological parameter defined as the proportion of carbon acquisition (e.g., GPP) to vegetation growth (NPP). Analyzing CUE help us to understand the carbon storage in simulated terrestrial ecosystem in ESMs. At first, the spatial distribution of observed CUE from space (e.g., MODIS) depends on climate condition such as precipitation and temperature. For example, the regions of large precipitation and warm climate show low CUE, while the regions of dry and cold climate show high CUE. It indicates that CUE at the regions with warm temperature and abundant precipitation could be lowered as there is a plenty of production and plant growth. The vegetation in cold temperature and insufficient precipitation adapts to the environmental condition for survival by increasing CUE.

In different with MODIS, we found clear difference of CUE between ESMs. The bias pattern of two ESMs from BCC showed the hemispheric contrast to positive in NH and negative in SH. The strong negative bias of CUE over southern hemisphere is shown in GFDL's models. The CUE in ESMs based on CLM4 (e.g., CESM-BGC and NorESM-ME) are significantly underestimated globally. This large uncertainty of CUE in individual models is influenced by biogeochemical parameterization of land surface model. In the MME, the spatial distribution of CUE is reasonably simulated. However, Strong negative bias is found over Amazon. It is caused that unbalanced ratio of GPP and Ra in the terrestrial carbon fluxes over tropical forest such as evergreen broadleaf forest the most models. The inverse relationship between temperature and CUE is reasonably simulated in the MME over dry regions. Generally, Ra is more sensitive to temperature than GPP in the real world over a certain range of temperatures

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

(Woodwell et al., 1990; Ryan, 1991; Piao et al., 2010). It means that the sensitivity of temperature to photosynthesis is weaker than that of respiration (Arnone and Korner, 1997; Enquist et al., 2007).

The CUE variation depending to the PFTs, MME is realistically reproduced in every PFTs. The model spread is large. It indicates a wide spread due to the different PFTs in each land models and systematic bias such as failure of PFT description in land models. The observed CUE values show a reasonable degree of non-linearity in terms of its response to temperature. In contrast, the stronger sensitivity of CUE to temperature increases in the MME is reflected by the systematic biases of simulated biogeochemical processes which depends on temperature conditions strongly in every PFTs.

However, most of the advanced ESMs have adopted leaf-scale biogeochemistry which involves parameterizations of photosynthesis and respiration based on small spatio-temporal scales that depend on laboratory experiments and limited in situ studies. It makes up one of the major uncertainties of carbon cycle processes in future climate change simulations from recent advanced ESMs. Atkin et al. (2008) suggested that most biogeochemical models are adjusted and incomplete parameterizations of biogeochemical processes. Due to the lack of observational data, many biogeochemical studies have focused on the total amount of primary production and respiration. Therefore, understanding and evaluating the global-scale ecosystem is challenging, based on the leaf scale biogeochemical parameterization used in the models. This leaf-level parameterization for biogeochemical processes is insufficient for long-term simulations (Zaehle et al., 2014). For realistic long-term simulations, such as climate change experiments including the carbon cycle and feedback processes, parameterizations representing idealized and generalized ecosystem-level processes are needed, rather than site-specific and leaf-level processes.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

575

576

577 **Acknowledgement**

578 This study is supported Basic Science Research Program through the National Research
579 Foundation of Korea (NRF), funded by the Ministry of Education, Science and Technology
580 (2012M1A2A2671851) and the Supercomputing Center/Korea Institute of Science and
581 Technology Information with supercomputing resources including technical support (KSC-
582 2015-C3-035).

583

584

585

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

References

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., and Zhu, Z.: Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth System Models, J. Clim., 26, 6801–6843, doi:10.1175/JCLI-D-12-00417.1, 2013.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J. H., Allard, G., Running, S. W., Semerci, A., Cobb, N.: A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests, For. Ecol. Manage., 259, 660–684, doi:10.1016/j.foreco.2009.09.001, 2010.

Andres, R. J., Gregg, J. S., Losey, L., Marland, G., and Boden: Monthly, global emissions of carbon dioxide from fossil fuel consumption, Tellus B, 63, 309–327, doi:10.1111/j.1600-0889.2011.00530.x, 2011.

Arnone III, J. A., Körner, C.: Temperature adaptation and acclimation potential of leaf dark respiration in two species of Ranunculus from warm and cold habitats, Arct. Antract. Alp. Res., 29, 122-125, doi:10.2307/1551842, 1997.

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., Wu, T.: Carbon–concentration and carbon–climate feedbacks in CMIP5 earth system models, J. Clim., 26, 5289-5314, doi:10.1175/JCLI-D-12-00494.1, 2013.

Atkin, O. K., Atkinson, L. J., Fisher, R. A., Campbell, C. D., Zaragoza-Castells, J., Pitchford, J. W., Woodward, F. I., Hurry, V.: Using temperature-dependent changes in leaf scaling relationships to quantitatively account for thermal acclimation of respiration in a coupled global climate–vegetation model, Glob. Change Biol., 14, 2709–2726, doi:10.1111/j.1365-

2486.2008.01664.x, 2008.

Booth, B. B. B., Jones, C. D., Collins, M., Totterdell, I. J., Cox, P. M., Sitch, S., Huntingford, C., Betts, R. A., Harris, G. R., and Lloyd, J.: High sensitivity of future global warming to land carbon cycle processes, Environ. Res. Lett., 7, 024002, doi:10.1088/1748-9326/7/2/024002, 2012.

Collatz, G. J., Ribas-Carbo, M., and Berry, J. A.: Coupled photosynthesis-stomatal conductance model for leaves of C4 plants, Aust. J. Plant Physiol., 19, 519–538, doi:10.1071/PP9920519, 1992.

De Lucia, E. H., Drake, J. E., Thomas, R. B., and Gonzalez-Meler, M.: Forest carbon use efficiency: is respiration a constant fraction of gross primary production? Glob. Chang. Biol., 13, 1157–1167, doi:10.1111/j.1365-2486.2007.01365.x, 2007.

Dewar, R. C., Medlyn, B. E., and McMurtrie, R. E.: Acclimation of the respiration photosynthesis ratio to temperature: insights from a model, Glob. Chang. Biol., 5, 615–622. doi:10.1046/j.1365-2486.1999.00253.x, 1999.

Enquist B. J., Kerhoff, A. J., Stark, S. C., Swenson, N. G., McCarthy M. C., and Price, C. A.: A general integrative model for scaling plant growth, carbon flux, and functional trait spectra, Nat. 449, 218-222, doi:10.1038/nature06061, 2007.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 149, 78–90, doi:10.1007/BF00386231, 1980.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., 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., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate–carbon

634 cycle feedback analysis: Results from the C4MIP model intercomparison, J. Clim., 19, 3337–
635 3353, doi:10.1175/JCLI3800.1, 2006.
636 Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K.,
637 and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, J.
638 Clim., 27, 511-525, doi:10.1175/JCLI-D-12-00579.1, 2014.
639 Friedlingstein, P., Fung, I., Holland, E., John, J., Brasseur, G., Erickson, D., and Schimel, D.:
640 On the contribution of CO2 fertilization to the missing biospheric sink, Global Biogeochem.
641 Cycles, 9, 541–556, doi:10.1029/95GB02381, 1995.
642 Gifford, R.M.: The global carbon-cycle – a viewpoint on the missing sink, Aust. J. Plant
643 Physiol., 21, 1–15, doi:10.1071/PP9940001, 1994.
644 Harris. I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of
645 monthly climatic observations – the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623–642,
646 doi:10.1002/joc.3711, 2014.
647 Heinsch, F. A., Zhao, M. S., Running, S. W., Kimball, J. S., Nemani, R. R., Davis, K. J.,
648 Bolstad, P. V., Cook, B. D., Desai, A. R., Ricciuto, D. M., Law, B. E., Oechel, W. C., Kwon,
649 H., Luo, H. Y., Wofsy, S. C., Dunn, A. L., Munger, J. W., Baldocchi, D. D., Xu, L. K., Hollinger,
650 D. Y., Richardson, A. D., Stoy, P. C., Siqueira, M. B. S., Monson, R. K., Burns, S. P., and
651 Flanagan, L. B.: Evaluation of remote sensing based terrestrial productivity from MODIS using
652 regional tower eddy flux network observations, IEEE Trans. Geosci. Remote Sens., 44,
653 1908–1925, doi:10.1109/TGRS.2005.853936, 2006.
654 Hoffman, F. M., Randerson, J. T., Arora, V. K., Bao, Q., Cadule, P., Ji, D., Jones, C. D.,
655 Kawamiya, M., Khatiwala, S., Lindsay, K., Obata, A., Shevliakova, E., Six, K. D., Tjiputra, J.
656 F., Volodin, E. M., and Wu, T.: Causes and implications of persistent atmospheric carbon
657 dioxide biases in Earth System Models, J. Geophys. Res. Biogeosci., 119, 141-162, doi:

10.1002/2013JG002381, 2013.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J. Q., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res, 116, G00J07, doi:10.1029/2010JG001566, 2011.

King, A. W.: Atmosphere: Plant respiration in a warmer world, Sci., 312, 536, doi:10.1126/science.1114166, 2006.

Larcher, W. and Mair, B.: Das Kälteresistenzverhalten von Quercus pubescens, Ostrya carpinifolia und Fraxinus ornus auf drei thermisch unterschiedlichen Standorten, Oecol. Plant., 3, 255–270, 1968.

Leith, C. E.: Climate response and fluctuation dissipation, J. Atmos. Sci., 32, 2022-2026, doi:10.1175/1520-0469(1975)032<2022:CRAFD>2.0.CO;2, 1975.

Monteith, J.: Solar radiation and productivity in tropical ecosystems, J. Appl. Ecol., 9, 747-766, doi:10.2307/2401901, 1972.

Obata, A.: Climate carbon cycle model response to freshwater discharge into the North Atlantic, J. Clim., 20, 5962–5976, doi:10.1175/2007JCLI1808.1, 2007.

Piao, S., Luyssaert, S., Ciais, P., Janssens, I. A., Chen, A., Cao, C., Fang, J., Friedlingstein, P., Luo, Y., and Wang, S.: Forest annual carbon cost: a global-scale analysis of autotrophic respiration, Ecol., 91, 652–661, doi:10.1890/08-2176.1, 2010.

Rahman, A. F., Sims, D. A., Cordova, V. D., and El-Masri, B. Z.: Potential of MODIS EVI and surface temperature for directly estimating per-pixel ecosystem C fluxes, Geophys. Res. Lett., 32, L19404, doi:10.1029/2005GL024127, 2005.

Running, S.W. and Gower, S.T.: FOREST-BGC, a general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets, Tree Physiol., 9, 147–160, doi:10.1093/treephys/9.1-2.14, 1991.

Ryan, M. G.: Effects of climate change on plant respiration. Ecol. Appl., 1, 157–167, doi: 10.2307/1941808, 1991.

Shao P., Zeng, X., Sakaguchi, K., Monson, R. K., and Zeng, X.: Terrestrial carbon cycle: climate relations in eight CMIP5 earth system models, J. Clim., 26, 8744-8764, doi:10.1175/JCLI-D-12-00831.1, 2013.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93, 485–498, doi:0.1175/BAMS-D-11-00094.1, 2012.

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 earth system models and comparison with observations, Biogeosci., 10, 1717–1736, doi:10.5194/bg-10-1717-2013, 2013.

Williams, D. N., Lawrence, B. N., Lautenschlager, M., Middleton, D., and Balaji, V.: The earth system grid federation: delivering globally accessible petascale data for CMIP5, Proc. of the 32nd Asia-Pacific Advanced Network Mtg., doi:10.7125/APAN.32.15, 2011.

Woodwell, G. M.: The effects of global warming, In J. Leggett, editor, Global warming: the Greenpeace report, Oxford Univ. Press, Oxford, UK, 116-132, 1990.

Yang, W., Shabanov, N. V., Huang, D., Wang, W., Dickinson, R. E., Nemani, R. R., Knyazikhin, Y., and Myneni, R. B.: Analysis of leaf area index products from combination of MODIS Terra and Aqua data, Remote Sens. Environ., 104, 297-312, doi:10.1016/j.rse.2006.04.016, 2006.

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., Luo,

706 Y., Wang, Y. P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Warlind, D., Weng, E., Parton,
707 W., Iversen, C. M., Gallet-Budynek, A., McCarthy, H., Finzi, A., Hanson, P. J., Prentice, I. C.,
708 Oren, R., and Norby, R. J.: Evaluation of 11 terrestrial carbon–nitrogen cycle models against
709 observations from two temperate Free-Air CO2 Enrichment studies, New phytol., 202, 803-
710 822, doi:10.1111/nph.12697, 2014.

711 Zhang, Y., Yu, J. G., Yang, J., Wimberly, M. C., Zhang, X., Tao, J., Jiang, Y., and Zhu, J.:
712 Climate-driven global changes in carbon use efficiency, Glob. Ecol. Biogeogr., 23, 144–155,
713 doi:10.1111/geb.12086, 2014.

714 Zhang, Y., Xu, M., Chen, H., and Adams, J.: Global pattern of NPP to GPP ratio derived
715 from MODIS data: effects of ecosystem type, geographical location and climate, Glob. Ecol.
716 Biogeogr., 18, 280–290, doi:10.1111/j.1466-8238.2008.00442.x, 2009.

717 Zhao, F. and Zeng, N.: Continued increase in atmospheric CO2 seasonal amplitude in the
718 21st century projected by the CMIP5 Earth System Models, Earth Syst. Dynam., 5, 423-439,
719 doi:10.5194/esd-5-423-2014, 2014.

720 Zhao, M. S., Heinsch, F. A., Nemani, R. R., and Running, S.W.: Improvements of the
721 MODIS terrestrial gross and net primary production global data set, Remote Sens. Environ.,
722 95, 164–176, doi:10.1016/j.rse.2004.12.011, 2005.

723 Zhu Q. and Zhuang Q.: Ecosystem biogeochemistry model parameterization: Do more flux
724 data result in a better model in predicting carbon flux? Ecosphere, 6, 283, doi:10.1890/ES15-
725 00259.1, 2015.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

726

727 **Table 1. List of ESMs used in this study and their features**

Number	Models	Modeling center	Horizontal resolution	ESM Reference	Land model	Photosynthesis	Autotropic Respiration	Nitrogen Cycle	Dynamic Vegetation
1	BCC-CSM 1	Beijing Climate Center, China	2.812° × 2.812°	Wu et al. (2013)	BCC-AVIM1	Farquhar et al., (1980)	Foley et al. (1996)	No	No
2	BCC-CSM 1M	Beijing Climate Center, China	1.125° × 1.125°	Wu et al. (2013)	BCC-AVIM1	Farquhar et al., (1980) Collatz et al. (1992)	Foley et al. (1996)	No	No
3	CanESM2	Canadian Centre for Climate Modeling and Analysis, Canada	2.812° × 2.812°	Arora et al. (2011)	CTEM	Farquhar et al., (1980) Collatz et al. (1992)	Ryan (1991)	No	No
4	CESM1-BGC	Community Earth System Model Contributors, NSF-DOE-NCAR, USA	1.25° ×0.9°	Long et al. (2013)	CLM4	Farquhar et al., (1980) Collatz et al. (1992)	Foley et al. (1996)	Yes	No
5	GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5° ×2°	Dunne et al. (2013)	LM3	Farquhar et al., (1980) Collatz et al. (1992)	Foley et al. (1996)	No	Yes
6	GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5° ×2°	Dunne et al. (2013)	LM3	Farquhar et al., (1980) Collatz et al (1992)	Ryan (1991)	No	Yes

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

7	MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, and National Institute for Environmental Studies, Japan	2.812° × 2.812°	Watanabe et al. (2011)	MATSIRO+ SEIB-DGVM	Farquhar et al., (1980)	Ryan (1991)	No	No
8	MPI-ESM LR	Max Planck Institute for Meteorology, Germany	2.812° × 2.812°	Ilyina et al. (2013)	JSBACH	Farquhar et al., (1980)	Obata (2007)	No	Yes
9	MRI-ESM1	Meteorological Research Institute, Japan	1.125° ×1.125°	Yukimoto et al. (2011)	HAL	Farquhar et al., (1980) Collatz et al. (1992)	Ryan (1997)	No	No
10	NorESM1-ME	Norwegian Climate Centre, Norway	2.5° ×1.875°	Tjiputra et al. (2013)	CLM4	Farquhar et al., (1980) Collatz et al. (1992)	Foley et al. (1996)	Yes	No

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

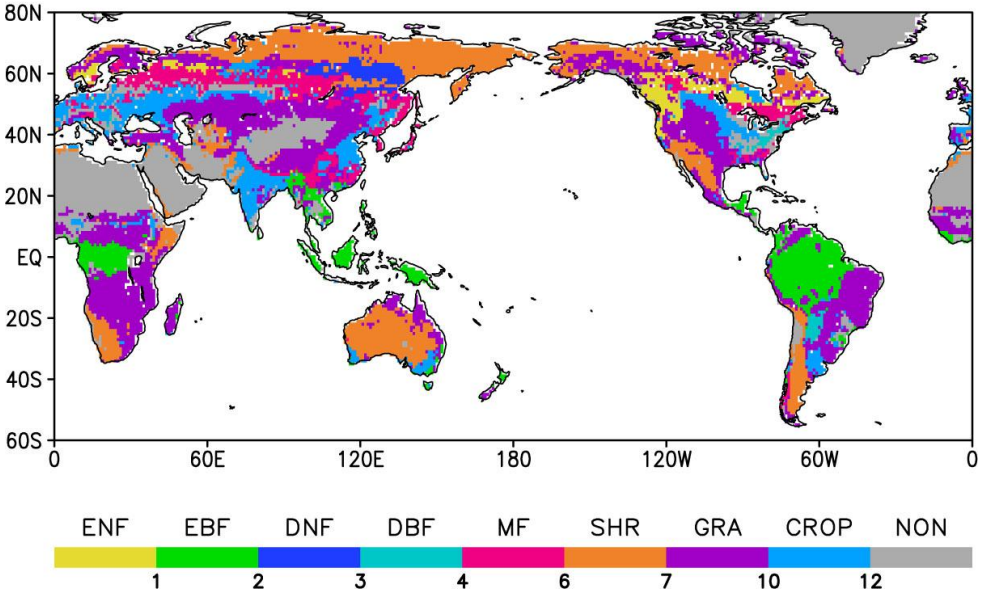


Figure 1. Horizontal distribution of dominant plant function types (PFTs) using the MODIS land cover data that include evergreen needleleaf forest (ENF), evergreen broadleaf forest (EBF), deciduous needleleaf forest (DNF), deciduous broadleaf (DBF), mixed forest (MF), shrub land (SHR), grass (GRA), cropland (CROP) and non-vegetated area (NON).

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

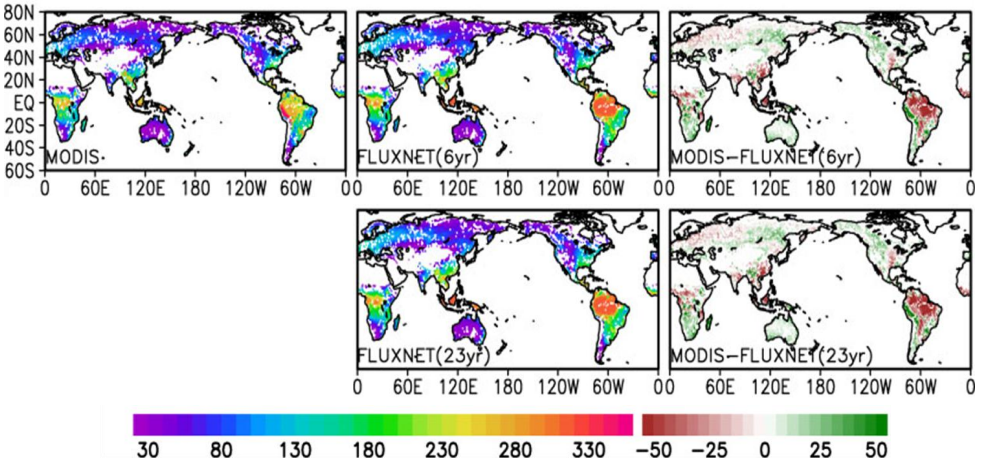


Figure 2. Spatial distributions of annual-mean GPP from MODIS (upper left), FLUXNET (upper middle), and MODIS minus FLUXNET (upper right) averaged for 6 years (2000-2005). Bottom panels show the GPP from FLUXNET averaged for 23 years (1983-2005, bottom left), and its difference from MODIS averaged for 6 years (bottom right). The unit is gC m^2 mon^{-1}.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

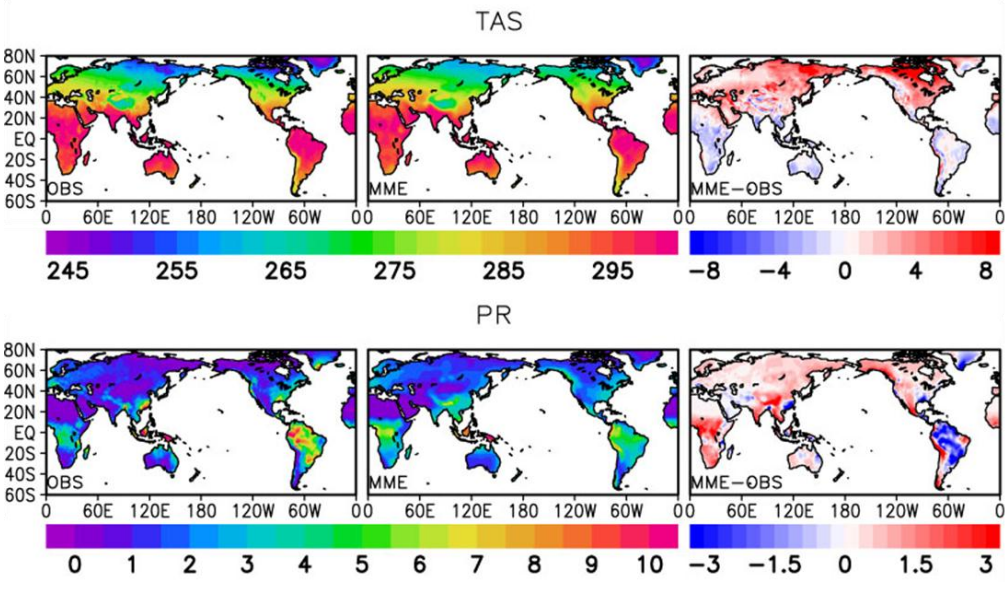


Figure 3. Annual-mean surface air temperature (top panels, unit: K) and precipitation (bottom panels, mm d^{-1}) averaged for 2000-2005 from the CRU observations (left), and the multi-model ensemble (MME) mean (middle), and the model biases (MME minus CRU, right).

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

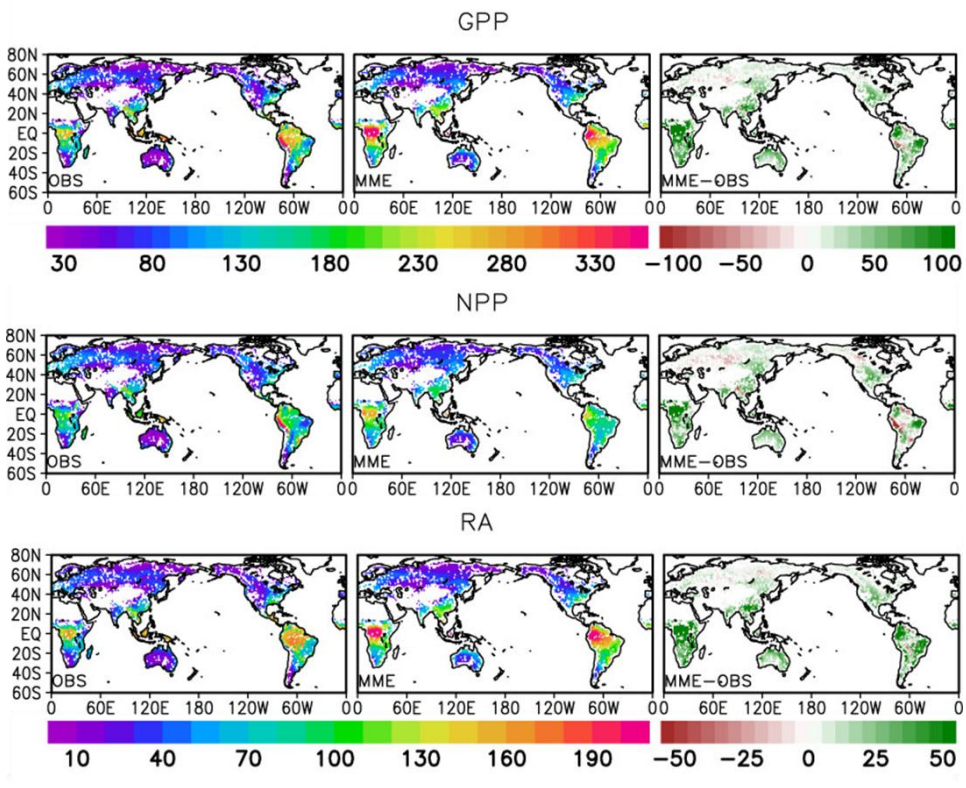


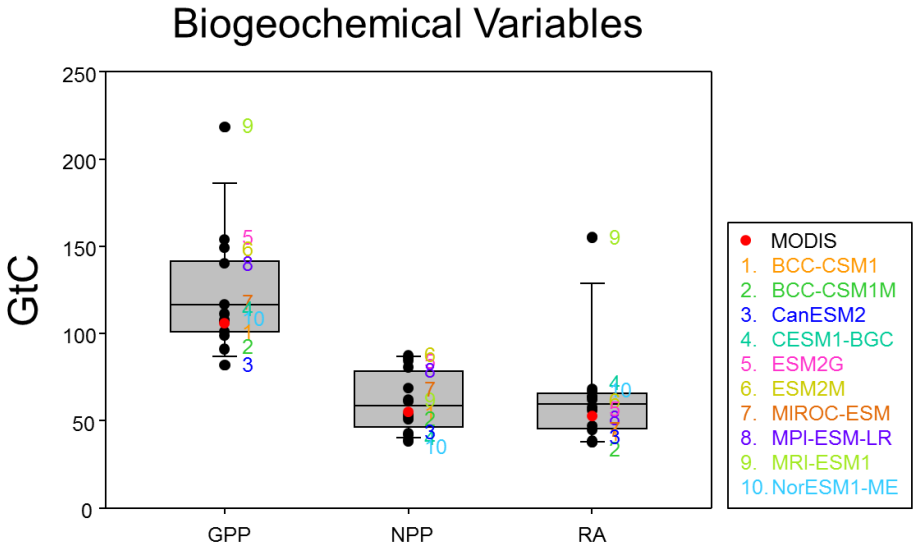
Figure 4. Same as in Figure. 3 except GPP (top), NPP (middle), and Ra (bottom) from the MODIS observations and MME. The unit is gC m^2 mon^{-1}.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

783

784

785



786

787 **Figure 5**. Global-mean values of GPP, NPP and Ra from MODIS and CMIP5 ESMs. The

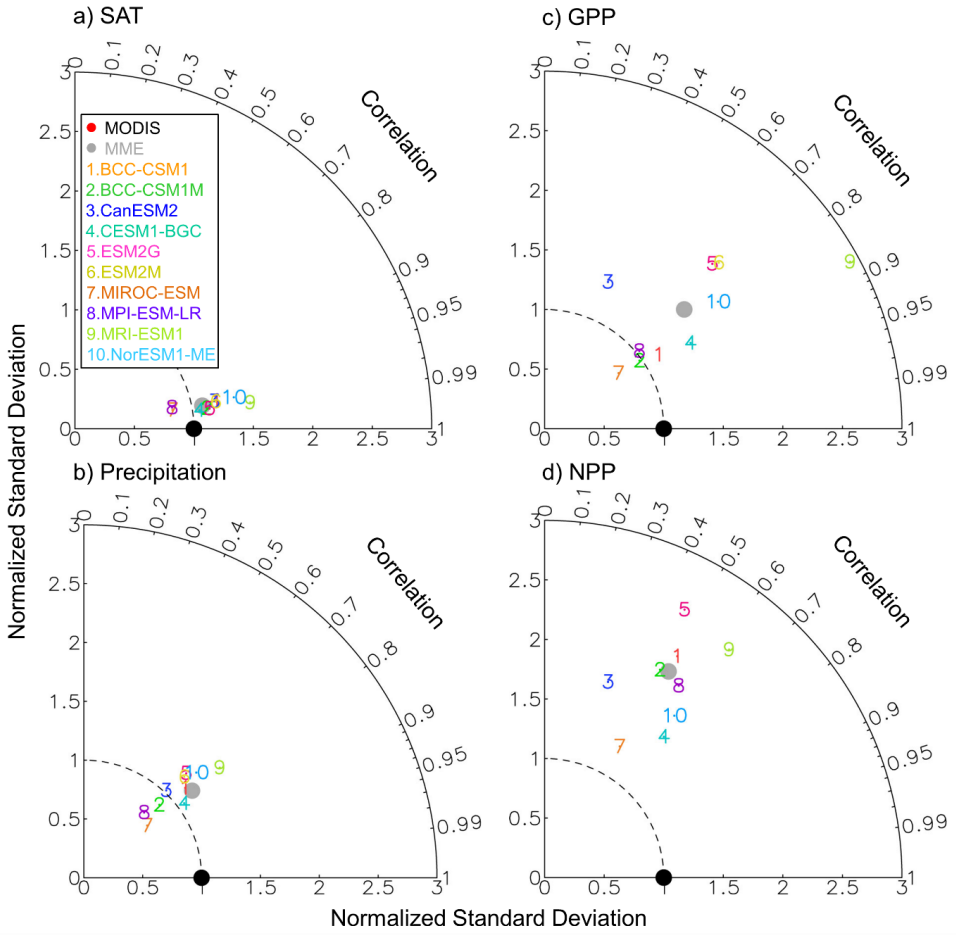
788 values are the average over the land grids only with latitude weighting for the period of 2000

789 – 2005.

790

791

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016



792

793 **Figure 6**. Taylor diagram of CMIP5 ESMs for annual-mean distribution of (a) surface air
794 temperature, (b) precipitation, (c) gross primary production (GPP) and (d) net primary
795 production (NPP) with respect to the corresponding observations for 6 years (2000-2005). Only
796 the vegetated grid points were included. The observed values are from CRU for temperature
797 and precipitation are MODIS for GPP and NPP.

798

799

800

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

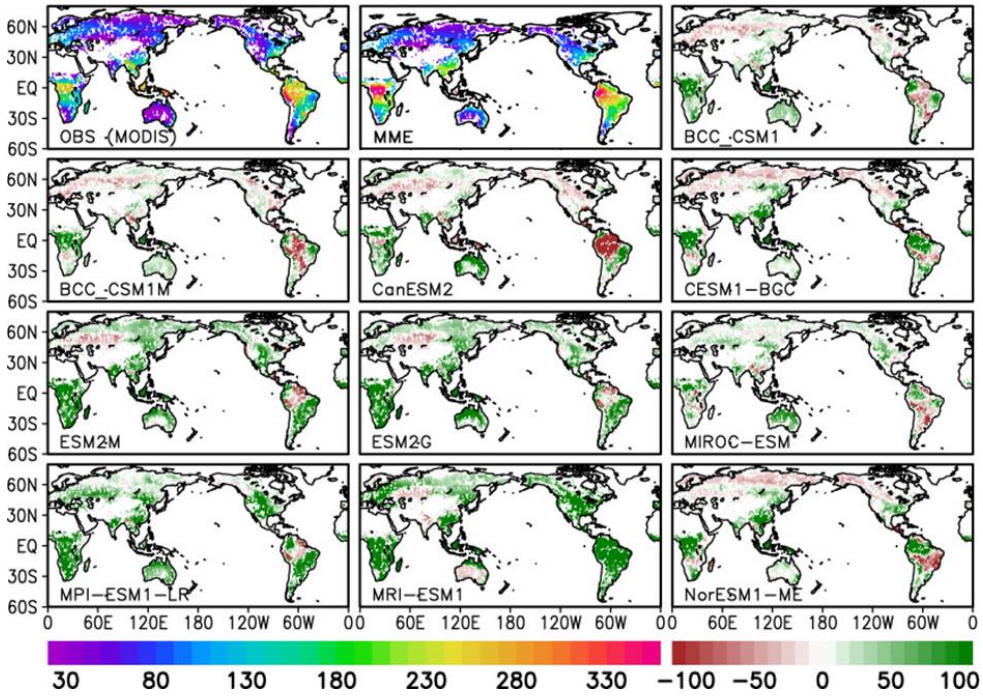


Figure 7. Spatial distribution of annual GPP from the MODIS observation (top left), MME (top middle) and the simulation bias in each model (model minus MODIS). The unit is gC m^2 mon^{-1}.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

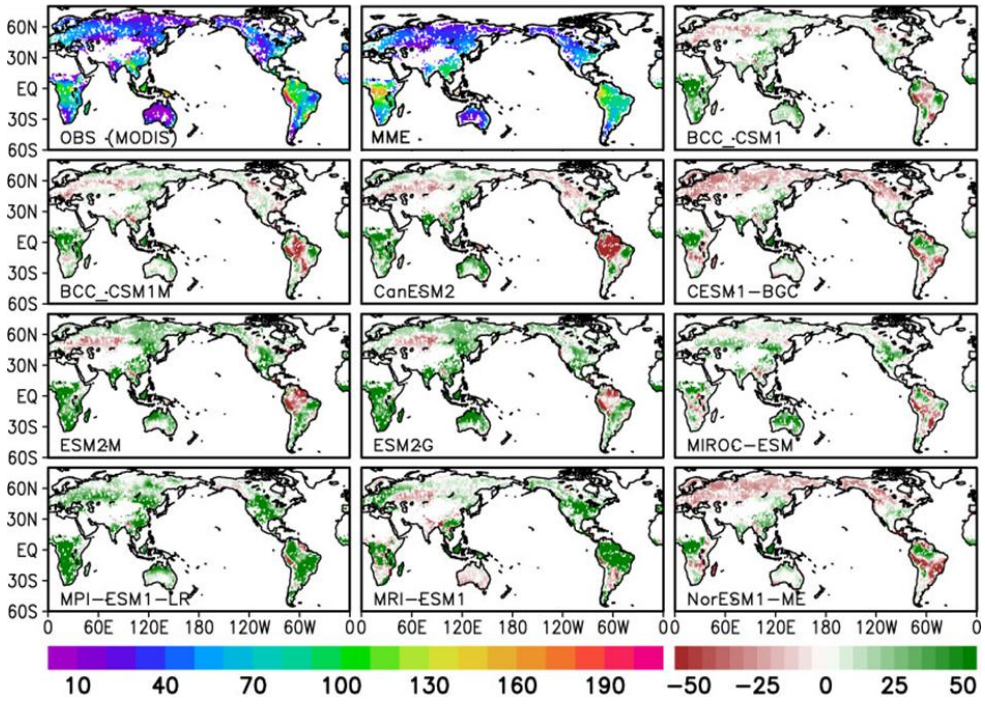


Figure 8. Spatial distribution of annual NPP from the MODIS observation (top left), MME (top middle) and the simulation bias in each model (model minus MODIS). The unit is gC m^2 mon^{-1}.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

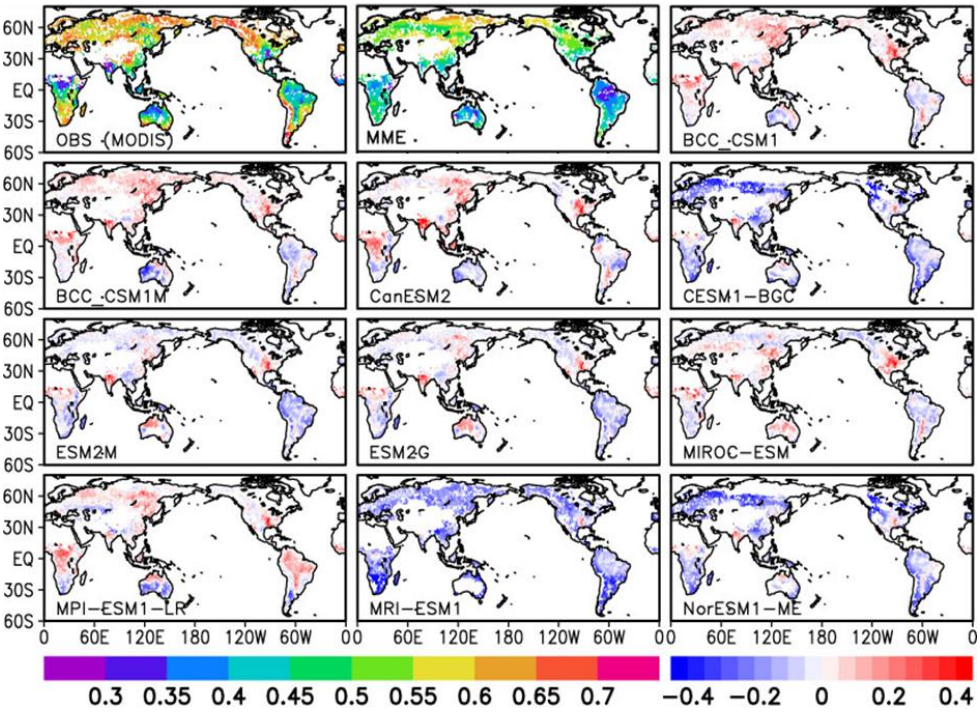


Figure 9. Spatial distribution of annual CUE from the MODIS observation (top left), MME (top middle) and the simulation bias in each model (model minus MODIS). CUE is a positively-defined ratio as NPP divided by GPP and less than or equal to 1.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

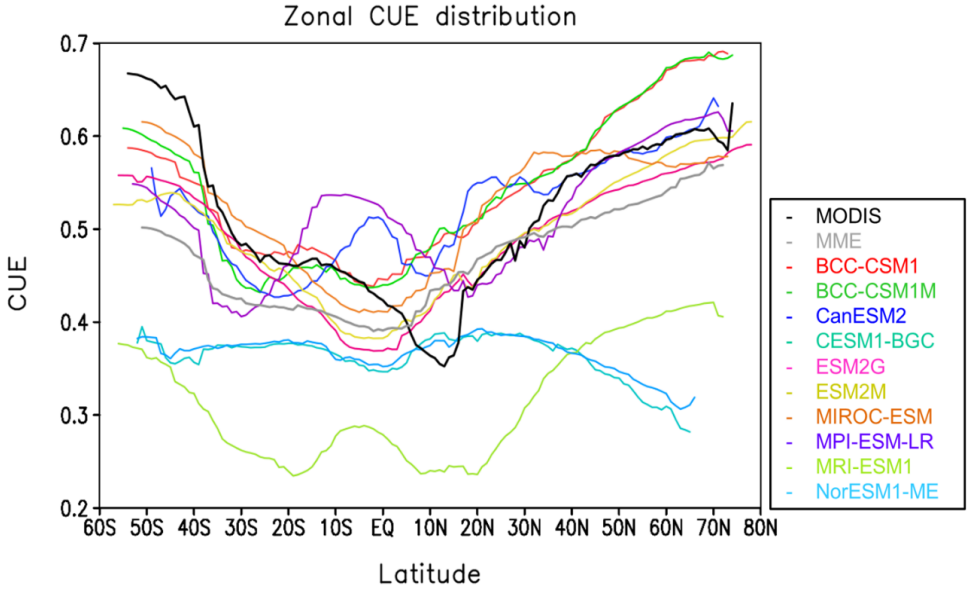


Figure 10. The zonal mean CUE from MODIS (black), MME (grey), and 10 ESMs (grey circles with number).

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

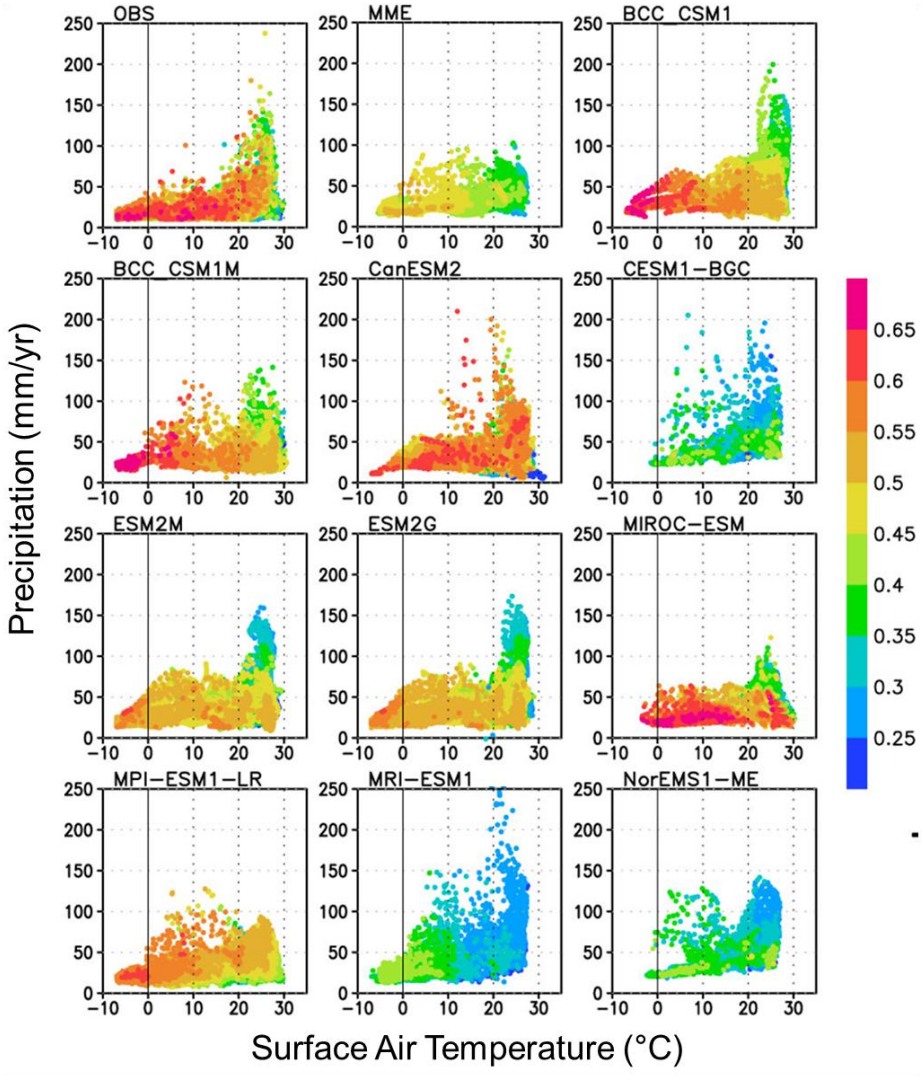


Figure 11. Scatter plot of CUE with the variation of surface air temperature (x-axis) and precipitation (y-axis). Color indicates CUE.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

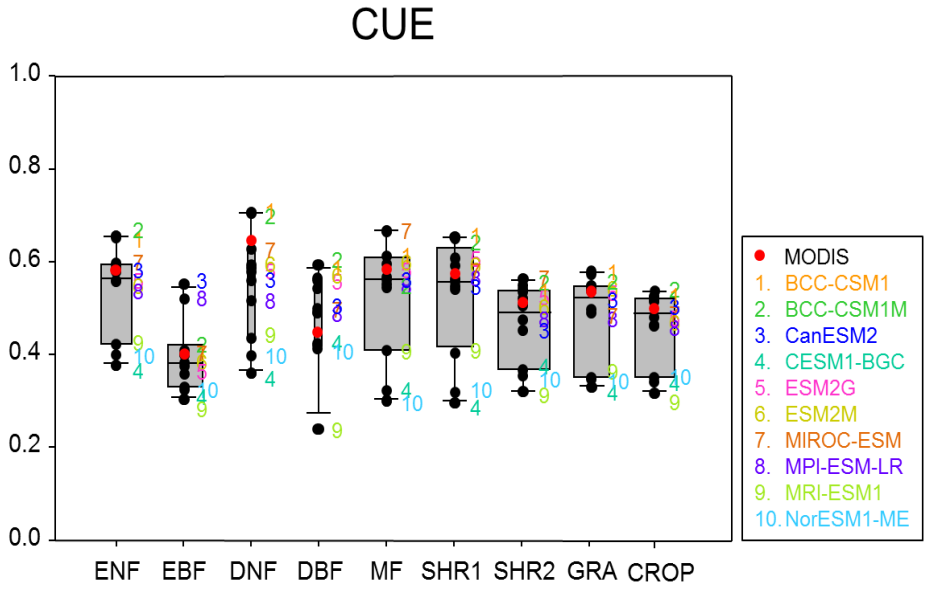


Figure 12. CUE averaged for each PFT. The box widths are proportional to the root mean square of number of grids. The coefficients of proportionality box widths in each PFTs are: ENF (0.80), EBF (0.48), DNF (0.12), DBF (0.11), MF (1.25), SHR1 (0.91), SHR2 (1.78), GRA (0.70) and CROP (0.73).

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

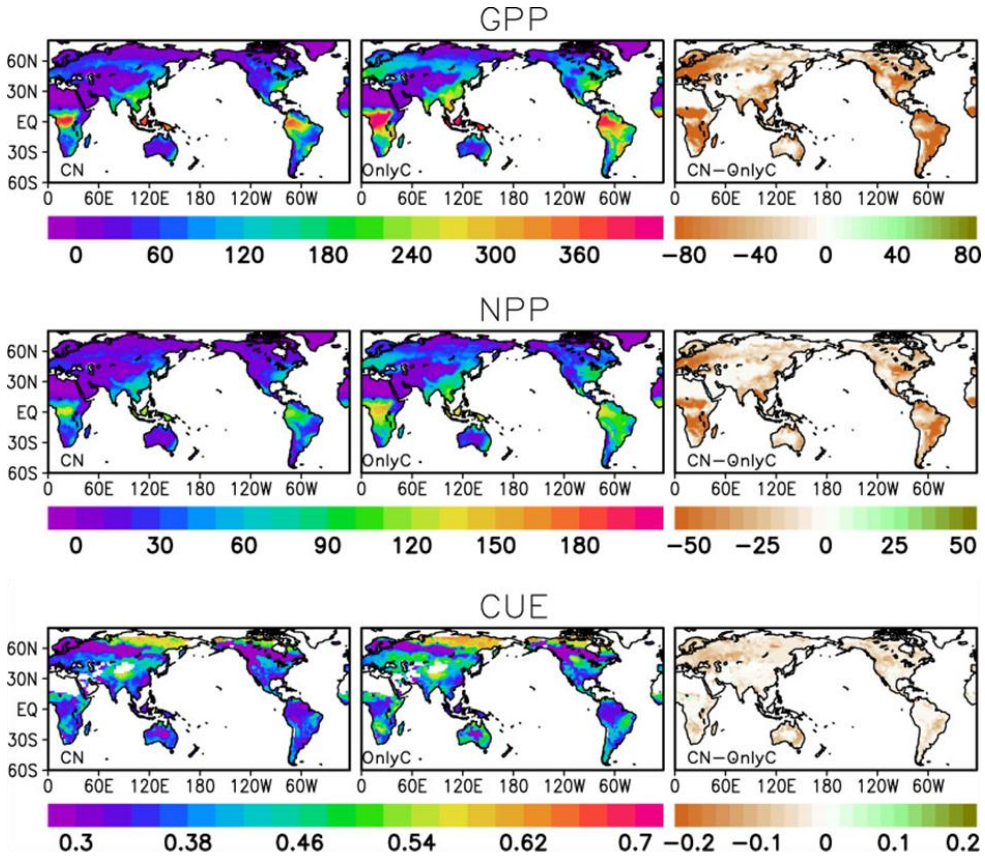


Figure 13. Spatial distributions of annual GPP, NPP and CUE and their differences from the interactive carbon-nitrogen cycle simulation (CN) and the run with no nitrogen cycle (Only C) by CESM-BGC. The units of GPP and NPP are gC m^2 mon^{-1}. CUE is a positively-defined ratio as NPP divided by GPP and less than or equal to 1.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

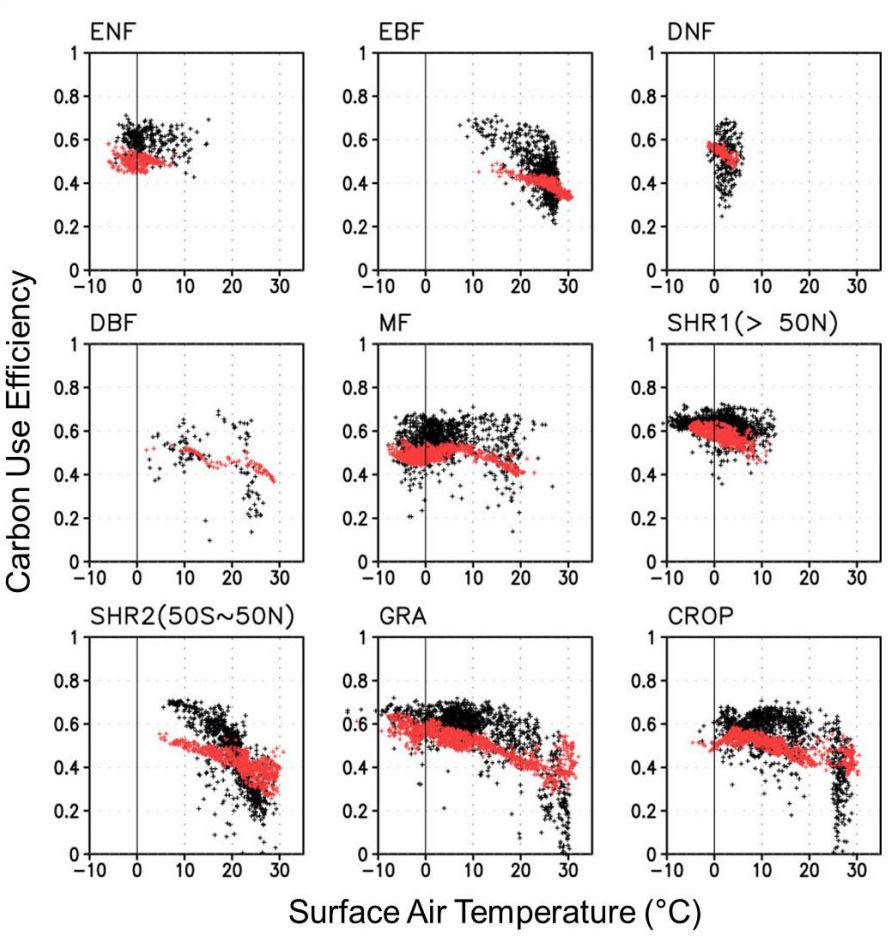


Figure 14. Scatter plots of CUE (y-axis) as a function of temperature (x-axis). Each panel shows the plot for different PFT. Satellite-derived values from MODIS are presented with black dots and the multi-model ensemble (MME) means by 10 ESMs are with red dots.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

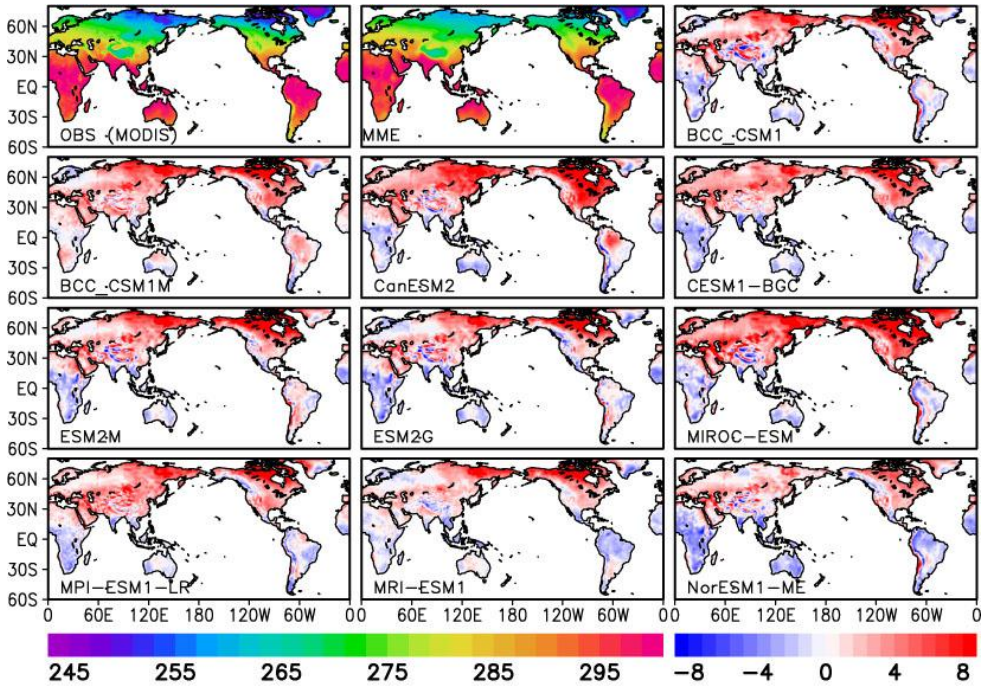


Figure S1. Spatial distribution of annual-mean surface air temperature from the CRU observation (top left), MME (top middle) and the simulation bias in each model (model minus CRU). The unit is K.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

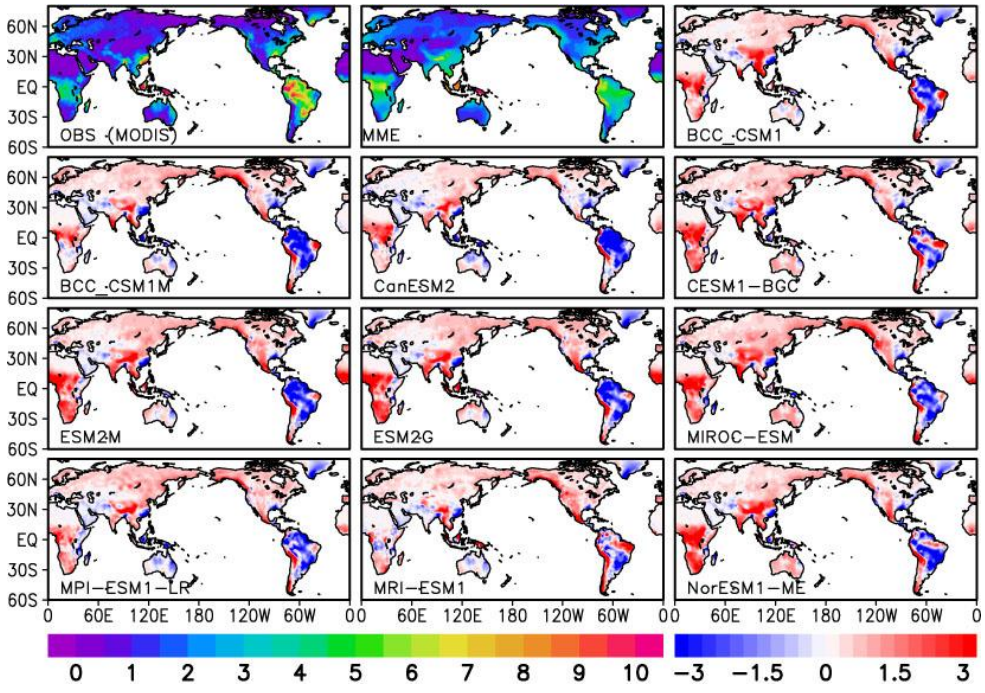


Figure S2. Spatial distribution of annual-mean precipitation from the CRU observation (top left), MME (top middle) and the simulation bias in each model (model minus CRU). The unit is mm d^{-1}.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

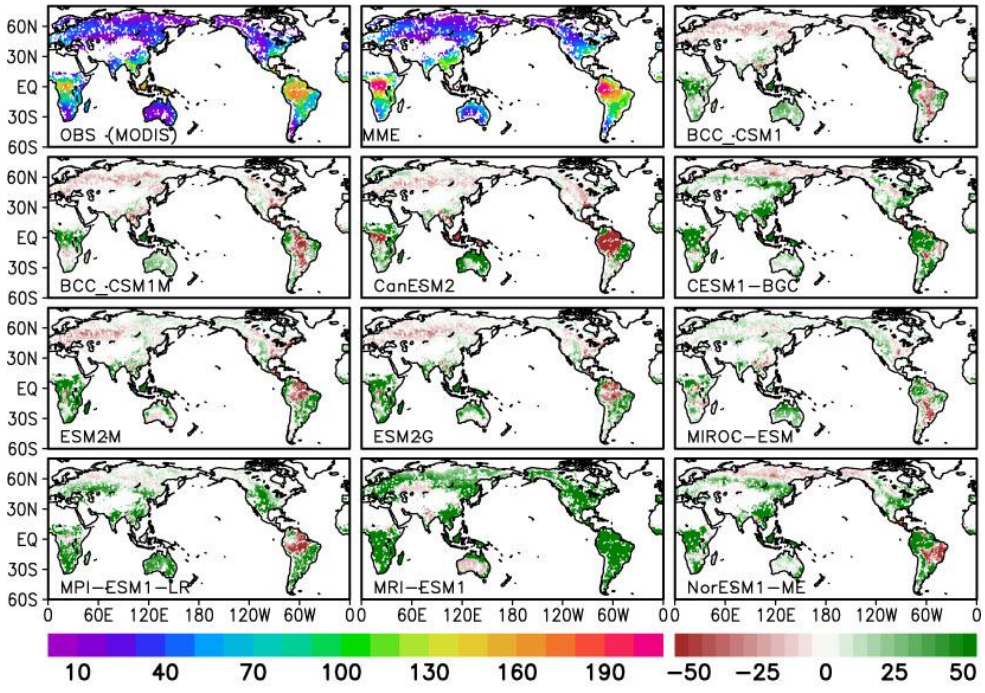


Figure S3. Spatial distribution of annual Ra from the MODIS observation (top left), MME (top middle) and the simulation bias in each model (model minus MODIS). The unit is gC m^2 mon^{-1}.

Biogeosciences Discuss., doi:10.5194/bg-2016-536, 2016
Manuscript under review for journal Biogeosciences
Published: 16 December 2016

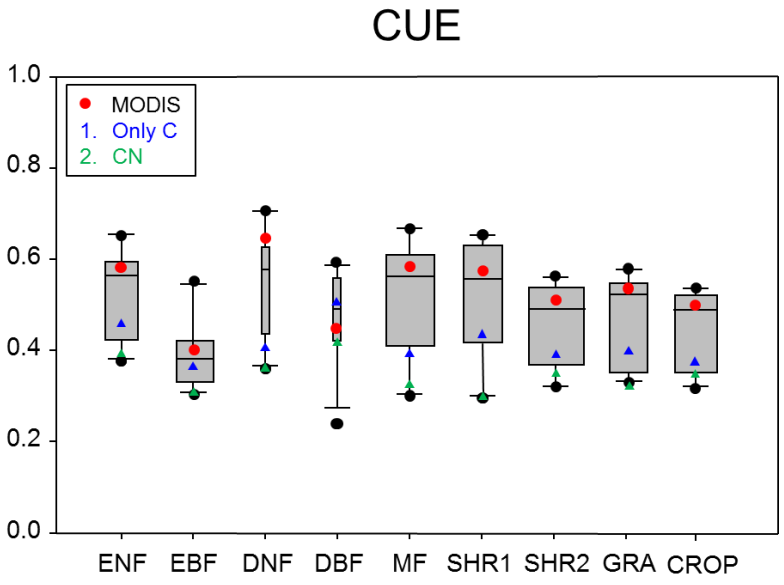


Figure S4. Simulated CUE averaged for each PFT in the two model sensitivity experiments using NCAR CEMS-BGC. CN (green triangles) indicates the run with interactive CN cycle and Only C (blue triangles) indicates the run that the nitrogen limitation effect is disabled. MODIS is also shown in red dots. The box widths are proportional to the root mean square of number of grids. The coefficients of proportionality box widths in each PFTs are: ENF (0.80), EBF (0.48), DNF (0.12), DBF (0.11), MF (1.25), SHR1 (0.91), SHR2 (1.78), GRA (0.70) and CROP (0.73).