Biogeosciences
Discussions

1 **An assessment of geographical distribution of different plant functional types**

2 **over North America simulated using the CLASS-CTEM modelling**

3 **framework**

4

5 **Rudra K. Shrestha[1], Vivek K. Arora[1], Joe R. Melton[2], and Laxmi Sushama[3]**

6 [1]Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada,

7 University of Victoria, Victoria, BC, V8W 2Y2, Canada

8 [2]Climate Research Division, Environment and Climate Change Canada, Toronto, Ontario, Canada

9 [3]Département des sciences de la Terre et de l'atmosphère, Université du Québec à Montréal, Canada

10

11 *Correspondence to: V. K. Arora (Vivek.Arora@canada.ca)*

12

Abstract

14

15 The performance of the competition module of the CLASS-CTEM (Canadian Land Surface

16 Scheme and Canadian Terrestrial Ecosystem Model) modelling framework is assessed at 1°

17 spatial resolution over North America by comparing the simulated geographical distribution of

18 plant functional types (PFTs) with two observation-based estimates. The model successfully

19 reproduces the broad geographical distribution of trees, grasses and bare ground although

20 limitations remain. In particular, compared to the two observation-based estimates, the simulated

21 fractional vegetation coverage is lower in the arid south-west North American region and higher

22 in the Arctic region. The lower than observed simulated vegetation coverage in the south-west

23 region is attributed to lack of representation of shrubs in the model and plausible errors in the

24 observation-based data sets. The observation-based data indicates vegetation fractional coverage

25 of more than 60% in this arid region, despite only 200-300 mm of precipitation that the region

26 receives annually and observation-based leaf area index (LAI) in the region are lower than one.

27 The higher than observed vegetation fractional coverage in the Arctic is due to the lack of

28 representation of moss and lichen PFTs and also likely because of inadequate representation of

29 permafrost in the model as a result of which the C_3 grass PFT performs overly well in the region.

30 The model generally reproduces the broad spatial distribution and the total area covered by the

31 two primary tree PFTs (needleleaf evergreen and broadleaf cold deciduous trees) reasonably

32   well. The simulated fractional coverage of tree PFTs increases after 1960s in response to the CO_2

33   fertilization effect and climate warming. Differences between observed and simulated PFT

34   coverages highlight limitations in the model and provide insight into physical and structural

35   processes that need improvement.

36

37

Biogeosciences

Discussions

1   Introduction

The terrestrial ecosystem plays an important role in regulating climate and weather through land-atmosphere exchange of water and energy (Cramer et al., 2001; Garnaud et al., 2015; Pielke et al., 1998; Ran et al., 2016) and in mitigating climate change by sequestering atmospheric CO_2 (Bonan, 2008; Timmons et al., 2016). The projected sink of atmospheric CO_2 is uncertain due to disagreements among the Earth system models (ESMs) (Arora et al., 2013; Friedlingstein et al., 2006) primarily due to differing responses of their terrestrial ecosystem modules to future changes in atmospheric CO_2. This uncertainty arises primarily because of the differences in the strength of the CO_2 fertilization effect on the land carbon cycle components (Arora et al., 2013; Cramer et al., 2001; Friend et al., 2013) but also because of differences in the response of vegetation. Models differ in how the spatial distribution of vegetation, and its composition, changes in response to changing climate and increasing CO_2 (Cramer et al., 2001). These differences are also resolution dependent. For example, models with coarse grid resolutions cannot explicitly resolve climatic niches, which in turn potentially contributes to biases in simulated vegetation distribution (Melton and Arora, 2016; Shrestha et al., 2016).

Vegetation responds to changes in climate and atmospheric CO_2 concentration by changing its structural attributes including leaf area index (LAI), rooting depth, vegetation height, and canopy mass, as well as its areal extent. Structural vegetation changes generally occur over seasonal to decadal time scales (Kramer and Kozlowski, 1979), while the slower areal extent changes typically occur on decadal to centennial time scales (Ritchie and Macdonald, 1986). The dynamic behavior of vegetation affects weather and climate due to its strong control over biophysical processes. At hourly to daily timescales, vegetation affects the exchange of water and energy between the land surface and the atmosphere primarily through the control of leaf stomata. At longer timescales from seasonal, annual to decadal timescales, vegetation affects components of energy and water balance through its structure (LAI, rooting depth, etc.) and its areal extent and thereby land surface albedo. Conversely, dynamics of vegetation is directly influenced by climate and the competitive ability of the plants. In this way vegetation responds to climate by changing its structure and areal extent depending on the colonization ability of plants.

Biogeosciences
Discussions

Open Access

68   These climate-vegetation interactions have been well documented (e.g. Gobron et al., 2010;

69   Wang et al., 2011).

70

71   Natural vegetation is typically characterized in dynamic global vegetation models (DGVMs)

72   based on a limited number of PFTs (Sitch et al., 2003) because it is impossible to represent

73   thousands of species in a model. Species characterized by similar attributes, mainly based on

74   their form and interactions with the environment (Box, 1996), are grouped together as a single

75   PFT. For example, tree species with similar leaf form such as fir *(Abies)*, spruce *(Picea)* and pine

76   *(Pinus)* are classified as needleleaf evergreen trees. The geographical distribution of the PFTs in

77   DGVMs is determined by their ability to grow and increase their areal extent given certain

78   climate and soil conditions and their competitive ability.

79

80   One way of representing competition between PFTs in DGVMs is through the use of the Lotka-

81   Volterra (LV) equations. While originally developed for predator-prey competition, the LV

82   equations have been used in a number of DGVMs (Arora and Boer, 2006; Brentnall et al., 2005;

83   Cox, 2001; Zhang et al., 2015). The use of the classical form of the LV equations for modelling

84   competition between PFTs, however, leads to an amplified expression of dominance in that the

85   dominant PFT ends up occupying a disproportionately large fraction of a grid cell leading to

86   little co-existence between PFTs. Arora and Boer (2006) proposed changes to the classical

87   implementation of the LV equations for modelling competition between PFTs to reduce this

88   amplified expression of dominance. Their approach, which has been implemented in the CLASS-

89   CTEM modelling framework and which allows improved co-existence of PFTs compared to the

90   classical LV equations, has been shown to simulate vegetation distribution reasonably well at the

91   global (Melton and Arora, 2016) as well as point (Shrestha et al., 2016) scales. Both these

92   studies used climate averaged over ~3.75° spatial resolution. The CLASS-CTEM framework

93   consists of the Canadian Land Surface Scheme (CLASS) coupled to the Canadian Terrestrial

94   Ecosystem Model (CTEM) which is a dynamic vegetation model.

95

96   In this paper, we evaluate the competition module of the CLASS-CTEM modelling framework at

97   the regional scale over the North American domain at 1° spatial resolution. This resolution is

98   much finer than the 3.75° resolution used in the Melton and Arora (2016) study and therefore in

Biogeosciences
Discussions

99   principle should allow a more realistic simulation of geographical distribution of PFTs as climate

100   niches are resolved.

101

102   The rest of this paper is organized as follows: Section 2 describes the CLASS-CTEM modelling

103   framework, details of the observation-based data and the experimental setup. Results are

104   presented in section 3 and a discussion follows in section 4. Finally, a summary and conclusions

105   are provided in section 5.

106

107   **2   Model, data and methods**

108

109   **2.1   CLASS-CTEM model**

110

111   The results presented here are obtained by coupling version 2.0 of CTEM (Melton and Arora,

112   2016), which dynamically simulates fractional coverage of its PFTs, to version 3.6 of CLASS

113   (Verseghy et al., 1993). CTEM simulates terrestrial processes for seven non-crop and two crop

114   PFTs (Table 1) and prognostically tracks carbon in three living vegetation components (leaves,

115   stems and roots) and two dead carbon pools (litter and soil). The terrestrial ecosystem processes

116   simulated in this study include photosynthesis, autotrophic respiration, heterotrophic respiration,

117   dynamic leaf phenology, allocation of carbon from leaves to stem and root components, fire, and

118   competition between PFTs which dynamically determines the fractional coverage of each PFT.

119   The amount of carbon in the leaf, stem and root components is used to estimate structural

120   attributes of vegetation. LAI is calculated from leaf biomass using PFT-dependent specific leaf

121   area (SLA) which determines area of leaves that can be constructed per kg C of leaf biomass

122   (Arora and Boer, 2005); vegetation height is calculated based on stem biomass for tree PFTs and

123   LAI for grass PFTs; and rooting depth is calculated based on root biomass (Arora and Boer,

124   2003). CTEM operates at a time step of one day except for photosynthesis and leaf respiration

125   which are calculated every 30 minutes for consistency with CLASS' energy and water balance

126   calculations which require stomatal resistance calculated by the photosynthesis module of

127   CTEM.

128

129   CLASS simulates the energy and water balance components at the land surface and operates at a

130   30 minutes time step. Liquid and frozen soil moisture and soil temperature are evaluated for

131   three soil layers (with maximum thicknesses of 0.1, 0.25 and 3.75 m). The actual thicknesses of

132   these permeable soil layers are determined by the depth to bedrock, which is specified on the

133   basis of the global data set of Zobler (1986). CLASS distinguishes four PFTs (needleleaf trees,

134   broadleaf trees, crops and grasses) which map directly to the nine PFTs represented in CTEM as

135   shown in Table 1. Needleleaf trees in CTEM are divided into deciduous and evergreen types,

136   broadleaf trees are divided into cold and drought deciduous and evergreen types, and crops and

137   grasses are divided into C_3 and C_4 types based on their photosynthetic pathways. In coupled

138   mode, CLASS uses the dynamically simulated vegetation attributes (including LAI, vegetation

139   height, canopy mass and rooting depth) and stomatal resistance calculated by CTEM, and CTEM

140   uses the soil moisture, soil temperature and net shortwave radiation calculated by CLASS. The

141   coupling frequency between CLASS and CTEM is one day.

142

143   2.1.1   **Competition parameterization**

144

145   Competition between PFTs in CTEM is parameterized following Arora and Boer (2006) who

146   presented a modified version of the LV equations. The approach is described in detail by Melton

147   and Arora (2016) and briefly summarized here. Consider, for simplicity, two PFTs that exist in a

148   grid cell with fractional coverages f_1 and f_2. Let PFT 1 represent a tree PFT and PFT 2 represent

149   a grass PFT. The bare fraction of grid cell not covered by any vegetation is represented by f_B. As

150   a result, $f_1 + f_2 + f_B = 1$. The rate of change of fractional coverages of the two PFTs and bare

151   fraction, for this example, are given by,

152

153   $$\frac{df_1}{dt} = c_1 f_1^{\beta}(1 - f_1) - m_1 f_1 \tag{1}$$

154

155   $$\frac{df_2}{dt} = c_2 f_2^{\beta}(1 - f_1 - f_2) - c_1 f_1^{\beta} f_2 - m_2 f_2 \tag{2}$$

156

157   $$\frac{df_B}{dt} = -c_1 f_1^{\beta} f_B - c_2 f_2^{\beta} f_B + m_1 f_1 + m_2 f_2 \tag{3}$$

158

159   where c_1, c_2 and m_1, m_2 are the colonization and mortality rates for PFT 1 and PFT 2,

160   respectively. Colonization and mortality rates cannot be negative. Equations (1) and (2) show

161   that PFT 1 can invade the fraction covered by PFT 2 and the bare fraction; and that PFT 2 can

162   only invade the bare fraction. PFT 2 is not allowed to invade the fraction covered by PFT 1

163   because it is ranked lower than PFT 1. In CTEM, the superiority or ranking of the seven natural

164   non-crop PFTs is based on the tree-grass distinction and their colonization rates. Trees are always

165   considered to be superior than grasses because of their ability to shade them (Siemann and

166   Rogers, 2003). Within the tree and grass PFTs the dominance is determined dynamically based

167   on the colonization rate. The exponent β $(0 \le \beta \le 1)$, an empirical parameter, controls the

168   behaviour of the LV equations. For $\beta = 1$, the equations represent the classical form of the LV

169   equations. The equilibrium fractional coverages for PFT 1 and 2 and bare fraction for this

170   classical form of the LV equations, denoted by $\tilde{f}_1, \tilde{f}_2$ and are given by,

171

172
$$\tilde{f}_1 = \max\left\{\left(\frac{c_1 - m_1}{c_1}\right), 0\right\} \tag{4}$$

173

174
$$\tilde{f}_2 = \max\left\{\left(\frac{(c_2 - m_2) - \left(1 + \frac{c_2}{c_1}\right)(c_1 - m_1)}{c_2}\right), 0\right\} \tag{5}$$

175

176
$$\tilde{f}_B = \frac{(m_1 \tilde{f}_1 + m_2 \tilde{f}_2)}{(c_1 \tilde{f}_1 + c_2 \tilde{f}_2)} \tag{6}$$

177

178   In equations (1) and (2), if the fractional coverages of PFT 1 and PFT 2 are initially zero then the

179   PFTs cannot expand for $\beta = 1$, implying that a minimum seeding fraction is always required.

180   Furthermore, in equation (5) as long as $(c_1 - m_1)$ is greater than $(c_2 - m_2)$ then the equilibrium

181   solution for f_2 will always be zero and PFT 2 will not be able to coexist with PFT 1. These

182   features of the classical form of the LV equations are avoided when $\beta = 0$, following Arora and

183   Boer (2006). The equilibrium fractional coverages for PFT 1 and 2 and bare fraction for the case

184   with $\beta = 0$ are given by,

185

186
$$\tilde{f}_1 = \left(\frac{c_1}{c_1 + m_1}\right) \tag{7}$$

187
$$\tilde{f}_2 = \frac{c_2(1 - \tilde{f}_1)}{(c_1 + c_2 + m_2)} = \left(\frac{c_2 m_1}{(c_1 + m_1)(c_1 + c_2 + m_2)}\right) \tag{8}$$

Biogeosciences
Discussions

Open Access

EGU

$$\tilde{f}_B = \frac{(m_1 \tilde{f}_1 + m_2 \tilde{f}_2)}{(c_1 + c_2)} \tag{9}$$

188

189 Unlike the classical version of the LV equations, the modified version of the equations with $\beta = 0$

190 does not require a minimum seeding fraction, and PFTs are able to increase their areal extent as

191 long as the climate is favorable and c_i is positive. Also, as long as $m_1 > 0$ and $c_2 > 0$ then PFT 2

192 is able to coexist at equilibrium with PFT 1. Other values of β between 0 and 1 give the dominant

193 PFT varying levels of access to sub-dominant PFTs but coexistence is most possible in the case

194 with $\beta = 0$.

195

196 The calculations of colonization and mortality rates are described in detail in Melton and Arora

197 (2016). Briefly, the colonization rate depends on the net primary productivity of a PFT. The

198 better a PFT performs for given climatic and soil conditions; the higher is its colonization rate.

199 The mortality rate represents the combined effect of four different processes: intrinsic or age-

200 related mortality, growth or stress mortality, mortality due to disturbance, and mortality due to

201 adverse climate which ensures that tree PFTs do not venture outside their bioclimatic zones.

202
203 **2.2   Forcing data**

204

205 The Climate Research Unit − National Centre for Environmental Prediction (CRU-NCEP)

206 reanalysis dataset (Viovy, 2012), is used to drive the model. The meteorological variables

207 (surface temperature, pressure, precipitation, wind, specific humidity, and incident short-wave

208 and long-wave radiation fluxes) are available at a spatial resolution of $0.5° \times 0.5°$ and at a six

209 hourly time interval for the period 1901-2010. These data are interpolated to $1°$ resolution

210 spatially, and disaggregated to half-hourly time resolution, a standard CLASS-CTEM model

211 integration time step. Temperature, pressure, wind, specific humidity, and long-wave radiation

212 are linearly interpolated in time while short-wave radiation is assumed to change with the solar

213 zenith angle with maximum radiation occurring at solar noon. Following Arora (1997) , the six-

214 hourly precipitation amount (P, mm/6-hour) is used to estimate the number of wet half-hours

215 (w_h) in a given six-hour period for $P > 0$ as

216

217
$$w_h = \text{integer}(\max[1, \min(12, 2.6 \log(6.93\, P))]). \tag{10}$$

218

219   The total precipitation amount is then distributed randomly but conservatively over these wet

220   half-hours. For instance, if seven out of 12 half hours intervals are calculated to be wet using

221   equation (10) then seven random numbers varying between 0 and 1 are generated and the six-

222   hourly precipitation amount is divided into seven parts in proportion to their respective random

223   numbers

224

225   Figure 1 shows the spatial distribution of mean annual precipitation and surface temperature over

226   the North American domain considered in this study. Mean annual precipitation values range

227   from less than 200 mm in the arid south-west United States and the high Arctic to more than

228   1500 mm on the Pacific coast. Mean annual temperature varies from around 24° C near the

229   southern limit of the domain in Mexico to less than -20° C in the Arctic tundra.

230

231   **2.3   Observation-based data**

232   2.3.1   **Fractional coverage of PFTs**

233

234   Observation-based estimates of fractional coverages of PFTs are based on a modified version of

235   the Wang et al. (2006) data set (hereafter WANG06) and the Moderate Resolution Imaging

236   Spectroradiometer land cover product (Friedl et al., 2013) (hereafter MODIS). These data are

237   used to evaluate the model results.

238

239   The WANG06 data set was developed for use by CTEM in simulations in which competition is

240   turned off and prescribed fractional coverage of PFTs is used. It combines observation- and

241   model-based data to estimate the annual change in fractional coverage of CTEM's nine PFTs

242   from 1850 to 2000. The Global Land Cover for the year 2000 (GLC2000), which is considered

243   as a base year for environmental assessment, divides the global land cover in 22 types is

244   available at 1 km resolution. WANG06 (their Table 2) mapped the GLC2000 data to CTEM's

245   nine PFTs aggregated to 0.5° resolution. The GLC2000 data were then extrapolated back to 1850

246   by adjusting the changes in crop area based on the then available Ramankutty and Foley (1999)

247   rop data set. Here, we use a modified version of the WANG06 data set which is based on the

Biogeosciences
Discussions

Open Access
EGU

248 HYDE v.3.1 crop data set (Hurtt et al., 2011) and generate an estimate of fractional coverage of

249 CTEM PFTs for the period 1850-2012.

250

251 The MODIS data set is based on the International Geosphere-Biosphere Programme (IGBP)

252 global vegetation data and University of Maryland's Science Data Set classification schemes at

253 0.25° spatial resolution. The data are derived from NASA HDF-EOS MODIS/Terra land cover

254 type. The data set is for the period 2001 to 2014 and contains 17 land cover types which we map

255 to CTEM's nine PFTs following the logic used in Wang et al. (2006) as shown in Table 2. The

256 fractional coverage of each of the nine CTEM PFT is first obtained at 0.25 degree resolution for

257 each year using the mapping scheme described in Table 2. These fractional coverages are then

258 re-gridded to the 1° spatial resolution for individual years. Finally, the data are averaged over the

259 period 2001-2014 to evaluate model results. MODIS data are known to exhibit substantial

260 interannual variability. Broxton et al. (2014), for instance, report that globally 40% of land pixels

261 show land cover change one or more times during 2001–2010 period. This does not necessarily

262 indicate changes in land cover but rather these differences are due to low accuracy in

263 categorizing the remotely sensed vegetation into one of the 17 MODIS land cover types, as

264 Broxton et al. (2014) note. This low accuracy is itself attributed to the fact that many landscapes

265 include mixtures of vegetation classes. Our re-gridding of fractional coverages to 1° spatial

266 resolution and averaging over the 2001-2014 time period to obtain climatology of land cover

267 alleviates some of the uncertainty since the effect of inaccurately classified land cover categories

268 is reduced due to both spatial and temporal averaging.

269

270 The separation of the broadleaf deciduous PFT into its drought and cold deciduous components

271 is performed via the approach used by WANG06. They assumed that below 24 °N deciduousness

272 is caused by soil moisture limitation and hence all broadleaf deciduous trees below this latitude

273 are drought deciduous, and above 34 °N deciduousness is caused by low temperatures and so all

274 broadleaf deciduous trees above this latitude are cold deciduous. Between 24 °N and 34 °N,

275 following WANG06 we assume a linear transition from drought deciduous to cold deciduous

276 trees. Finally, the separation of grasses into their C_3 and C_4 components is based on the

277 geographical distributions of the C_3 and C_4 fractions in the WANG06 data set.

278   2.3.2   **Gross primary productivity and LAI**

279

280   Observation-based estimates of gross primary productivity (GPP) are based on Beer et al. (2010).

281   These data are based on the ecosystem level GPP obtained using eddy covariance measurements

282   from more than 250 stations across the globe. Beer et al. (2010) extrapolated these eddy

283   covariance flux data of GPP to the global scale using diagnostic models for the period 1982 –

284   2008, and the average over this time period is used to evaluate the model results. LAI data used

285   for validation are the same as those used by Anav et al. (2013) and are based on Zhu et al. (2013)

286   who use normalized difference vegetation index (NDVI) data from the Advanced Very High

287   Resolution Radiometer (AVHRR) satellite to calculate average LAI for the period 1981 – 2010.

288

289   **2.4   Experimental setup**

290

291   2.4.1   **Equilibrium pre-industrial simulation**

292

293   The equilibrium pre-industrial simulation was initialized from zero biomass and zero fractional

294   coverage for all non-crop PFTs. The fractions of C_3 and C_4 crop PFTs in each grid cell are

295   specified corresponding to year 1850 based on the HYDE 3.1 dataset. The model was then run

296   for 600 years driven by 1901-1925 CRU-NCEP climate data cycled repeatedly. These data do

297   not show any warming trend (Wen et al., 2011) as opposed to the later part of the 20th century.

298   Atmospheric CO_2 concentration was set to 285 ppm corresponding to the pre-industrial 1850

299   level. This pre-industrial equilibrium simulation yields initial conditions including fractional

300   coverages of PFTs and carbon in all the live and dead pools for the transient 1850-2010

301   simulation. The 600 years simulation is sufficient for fractional vegetation cover and carbon

302   pools to reach equilibrium.

303   2.4.2   **Transient historical simulation**

304

305   The transient historical simulation is performed for the period 1851-2010 and its carbon pools

306   and fractional coverage of non-crop PFTs are initialized from the equilibrium pre-industrial

307   simulation as mentioned above. The years 1851 to 1900 of this historical simulation are driven

308   with CRU-NCEP climate data corresponding to the period 1901-1925, cycled twice. For the

309   period 1901-2010 the climate data corresponding to each year are used. Time varying

310   concentrations of atmospheric CO_2 are supplied for the period 1851-2010 based on the values

311   used in the fifth Coupled Modelling Intercomparison Project (CMIP5,

312   http://tntcat.iiasa.ac.at/RcpDb/) which are extended past 2005 to 2010 based on data from the

313   National Oceanic and Atmospheric Administration

314   (ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_annmean_gl.txt). The annual time-varying

315   fractional coverages of C_3 and C_4 crop PFTs in each grid cell are based on the HYDE 3.1 dataset.

316   The crop fractions in a grid cell are not available for colonization and neither are they subject to

317   disturbance by fire. Competition between PFTs occurs over the remaining non-crop fraction of a

318   grid cell. As total crop fraction in a grid cell changes over time (based on the HYDE 3.1 dataset)

319   the fractional area available for competition also changes.

320

321   The simulated results are evaluated against their observation-based counterparts using averaged

322   values over the last 30 years of the simulation corresponding to the period 1981-2010. This is the

323   same and/or very close to the time period for modified WANG06 land cover data set (1981-

324   2010), Beer et al. (2010) GPP (1982-2008), and Zhu et al. (2013) LAI (1981-2010). The only

325   exception is the MODIS-based land cover data which are available for the 2001-2014 period.

326   **3   Results**

327

328   **3.1   Continental scale values of PFT coverage**

329

330   Figures 2a compares the simulated vegetation areas summed over our North American domain

331   with the WANG06 and MODIS observation-based estimates. In the absence of another measure

332   of uncertainty, we use the range between these two observation-based estimates and assess if

333   simulated areal coverage of a given land cover type lies within or outside this range. The

334   simulated total vegetated area over North America (14.8×10^6 km^2) is very similar to the

335   modified WANG06 (14.4×10^6 km^2) and MODIS derived (14.2×10^6 km^2). At the most basic

336   tree-grass-bare ground level, the simulated areas are closer to the MODIS-based estimates, than

337   to the estimate based on the modified WANG06 data. The simulated area covered by tree PFTs

338 $(7.8 \times 10^6 \text{ km}^2)$ is 6% lower than the MODIS derived estimate $(8.2 \times 10^6 \text{ km}^2)$ and 21% lower

339 than WANG06 $(9.7 \times 10^6 \text{ km}^2)$. The simulated grass coverage $(4.7 \times 10^6 \text{ km}^2)$ is 35% higher

340 than the MODIS derived estimate $(3.5 \times 10^6 \text{ km}^2)$. Both simulated and MODIS-based estimates

341 of area covered by grass PFTs are, however, substantially higher than the WANG06 $(2.4 \times 10^6$

342 $\text{km}^2)$ estimate. Averaged over the North American region, the simulated partitioning of land area

343 (excluding cropland area) covered by trees, grasses and bare ground (45%, 27%, 28%) is much

344 closer to the MODIS based data (48%, 20% and 32%) than to the modified WANG06 based data

345 (56%, 14%, 30%).

346

347 Figure 2b shows a comparison of simulated areas of individual PFTs with observation-based

348 estimates. This is a more stringent test of the performance of the competition module of CTEM.

349 The observation-based estimates of areas of all individual PFTs are available for the modified

350 WANG06 dataset. The MODIS based estimates were derived based on the mapping of MODIS'

351 17 land cover types to CTEM PFTs as shown in Table 2, which itself is mostly based on

352 WANG06. In Figure 2b, the observation-based estimates show that needleleaf evergreen (NDL

353 EVG) and broadleaf cold deciduous (BDL DCD CLD) are the dominant tree PFTs across North

354 America which is also shown by the model. The simulated total area of the NDL EVG tree PFT

355 $(3.9 \times 10^6 \text{ km}^2)$ is 28% less than WANG06 $(5.3 \times 10^6 \text{ km}^2)$ and 15% less than the MODIS based

356 estimate $(4.7 \times 10^6 \text{ km}^2)$. The simulated total area of BDL DCD CLD tree PFT $(3 \times 10^6 \text{ km}^2)$ is

357 13% less than WANG06 $(3.4 \times 10^6 \text{ km}^2)$ and 3% greater than MODIS based $(2.9 \times 10^6 \text{ km}^2)$

358 estimate. Overall, the model is able to capture the areas covered by individual PFTs reasonably

359 well. However, differences remain between observations-based and simulated estimates

360 especially the larger simulated area for C_3 grasses than both observation-based estimates.

361 Reasons for these differences include limitations in the model but also the manner in which

362 remotely-sensed vegetation is categorized into broad-scale vegetation types and then mapped

363 onto CTEM's nine PFTs, as discussed later.

364

365 In both Figures 2a and 2b although simulated areal coverages at the basic tree-grass-bare ground

366 level and for individual PFTs (except for C_3 grasses) are comparable to observation-based

13

Biogeosciences
Discussions

Open Access

EGU

367   estimates they are outside the range defined by difference of the WANG06 and MODIS based

368   estimates.

369

370   Figure 2c shows the time series of simulated areas summed over the domain covered by tree and

371   grass PFTs, the total vegetated area and the remaining bare ground. The specified area covered

372   by crop PFTs, based on the HYDE 3.1 data set, is also shown and first increases over the

373   historical period and then stabilizes and in fact somewhat decreases in association with cropland

374   abandonment over the north-eastern United States. The increase in the crop area results in a

375   decrease in the area covered by tree and grass PFTs up until the time when the crop area

376   stabilizes around 1970. In the model, this causes land use change emissions associated with

377   deforestation. After this time, as vegetation productivity responds to increasing atmospheric CO_2

378   concentration, the area covered by tree PFTs increases somewhat and colonizes available bare

379   areas and those covered by grass PFTs. This leads to a small reduction in the area covered by

380   grass PFTs as well as bare ground and the associated increase in the total vegetated area.

381

382   **3.2   Geographical distribution of PFTs**

383

384   3.2.1   **Total vegetated and bare ground fractions**

385

386   Figures 3 and 4 compare the geographical distribution of simulated total vegetated and bare

387   fractions across North America with the two observation-based estimates derived from the

388   modified WANG06 and MODIS data sets. The two observation-based estimates are also

389   compared amongst themselves. The metrics used are averaged root mean square difference

390   (RMSD) and spatial correlations (R^2).

391

392   The observation-based geographical distribution of vegetated fraction in Figure 3 (middle

393   column) shows densely vegetated land over the eastern part of the continent and less vegetation

394   coverage over colder regions in the North and drier regions in the south-central and south-west

395   United States. These broad scale patterns are consistent with the precipitation and temperature

396   climatologies of the region (Figure 1). The model reasonably reproduces the observed vegetation

397   distribution (left panel) with some obvious limitations. Simulated vegetation cover is

398   underestimated across the arid south-west United States, Great Plains and part of the Canadian

399   Prairies (right panel) due to lower simulated fractional coverage of tree and grass PFTs over

400   these regions (which are shown in the next section). The model overestimates vegetation

401   coverage in Northern Canada because of higher simulated grass cover in the Arctic as discussed

402   below in more detail. The spatial correlation and RMSD when comparing simulated vegetated

403   fraction to both observation-based estimates are 0.79 and around 18%, respectively. The spatial

404   correlation and RMSD between the two observation-based estimates themselves are 0.86 and

405   around 14%, respectively.

406

407   The simulated and observation-based bare ground fractions across North America are compared

408   in Figure 4. The observation-based estimates show that bare ground fraction is higher in Arctic

409   Canada and Alaska where, of course, cold temperatures limit vegetation growth and in the south-

410   west United States, Great Plains and the Prairies where low rainfall limits vegetation growth

411   (Figure 1). The biases in simulated bare ground fraction mirror those in the simulated vegetated

412   fraction but in an opposite manner. The model underestimates bare ground fraction across Arctic

413   Canada due to higher simulated grass cover as discussed in the next section. The model

414   overestimates the bare ground fraction generally across the arid and semi-arid south-west United

415   States, Great Plains and the Prairies. The spatial correlation and RMSD when comparing

416   simulated bare ground fraction to both observation-based estimates are around 0.46 and around

417   18%, respectively. The spatial correlation and RMSD between the two observation-based

418   estimates themselves are 0.68 and around 14%, respectively.

419

420   3.2.2   **Tree and grass cover**

421

422   Figure 5 compares the simulated tree cover with the two observation-based estimates. The model

423   reasonably reproduces the broad scale patterns including the Canadian boreal forest and the

424   temperate forests across the southeastern United States. However, the model simulates lower tree

425   cover across the western part of the continent compared to both observation-based estimates

426   particularly over the southwestern United States which is characterized by arid climate (Figure

427   1). The observation-based estimates do not particularly well agree over this region either. The

428   MODIS derived estimate suggests around 25% tree cover in the southwestern United States

429   while the WANG06 derived estimate suggests a tree cover of around 60% over a large area in

430   the region. The spatial correlation and RMSD when comparing simulated tree cover to both

431   observation-based estimates are around 0.68 and around 17%, respectively. The spatial

432   correlation and RMSD between the two observation-based estimates themselves are 0.75 and

433   around 15%, respectively. Possible reasons for differences between simulated and observation-

434   based estimates are discussed in detail in the discussion section and include the fact that the

435   CLASS-CTEM framework does not currently represent shrubs and there are limitations in the

436   observation-based data sets themselves. Shrubs are more prevalent in arid and semi-arid regions

437   where they are better suited to grow compared to both trees and grasses.

438

439   Figure 6 compares the geographical distribution of the simulated grass cover with the two

440   observation-based estimates. The broad geographical distribution of simulated grass cover

441   compares well with the two observation-based estimates with the notable exception of the Arctic

442   region including Alaska and northern Canada, where the model overestimates grass cover. This

443   overestimation of grass cover in the Arctic region is also the reason for the overestimation of

444   total vegetation fraction and the underestimation of bare fraction that was seen earlier in Figures

445   3 and 4 respectively.

446

447   As shown in Figure 6, the spatial correlation and RMSD when comparing simulated grass cover

448   to both observation-based estimates lie between 0.33 and 0.38 and between around 15-17%,

449   respectively. The spatial correlation and RMSD between the two observation-based estimates

450   themselves are 0.54 and around 9%, respectively. The two observation-based estimates disagree

451   most markedly over the western half of the United States where the MODIS derived estimates of

452   grass cover are higher.

453

454   3.2.3   **Needleleaf evergreen and broadleaf cold deciduous trees**

455

456   Figures 7a and 7b compare the geographical distribution of NDL EVG and BDL DCD CLD

457   trees, respectively, with their observation-based estimates. These two are the primary tree PFTs

458   which exist in the North American domain considered here.

459

460   In Figure 7a, the overall simulated coverage of NDL EVG trees is lower than both observation-

461   based estimates as was also seen in Figure 2b. The simulated values are primarily lower in

462   western Canada and over a large area in the western United States according to estimates based

463   on the modified WANG06 data set. This is also the case along the wide swath of the Canadian

464   boreal forest. The model overestimates the coverage of NDL EVG trees in the eastern United

465   States. The spatial correlation and RMSD when comparing simulated coverage of NDL EVG

466   trees to both observation-based estimates lie between 0.36 and 0.40 and between around 16-17%,

467   respectively. The spatial correlation and RMSD between the two observation-based estimates

468   themselves are 0.52 and around 16%, respectively.

469

470   The geographical distribution of BDL DCD CLD trees is compared with its observation-based

471   estimates in Figure 7b. Although the simulated domain summed area of BDL DCD CLD trees (3

472     $\times 10^6$ km^2) is comparable to estimates based on the modified WANG06 (3.4×10^6 km^2) and

473   MODIS (2.9×10^6 km^2) data sets, there are two primary limitations in its simulated geographical

474   distribution. First, the simulated values are generally overestimated in Canadian boreal forests

475   and underestimated in the eastern United States. Second, the model simulates near zero coverage

476   in the arid south-western United States. The spatial correlation and RMSD when comparing

477   simulated coverage of BDL DCD CLD trees to both observation-based estimates are around 0.3

478   and around 12%, respectively. The spatial correlation and RMSD between the two observation-

479   based estimates themselves are 0.60 and around 8%, respectively.

480

481   ### 3.2.4 C_3 and C_4 grasses

482

483   Figures 8a and 8b compare the simulated geographical distribution of C_3 and C_4 grasses with

484   observation-based estimates.

485

486   In Figure 8a, the most obvious limitation of the model is its excessive simulated grass coverage

487   in Alaska and in Arctic Canada. Other than this, the model reproduces the broad geographical

488   distribution of C_3 grasses including the Great Plains of United States and the Canadian Prairies,

489   where a large extent of grasslands is observed. The overestimated grass coverage at high

490   latitudes leads to a total simulated C_3 grass area (4.4×10^6 km^2) that is higher than estimates

491   based on the modified WANG06 (1.9×10^6 km^2) and MODIS (2.8×10^6 km^2) data sets. The

492   spatial correlation and RMSD when comparing simulated coverage of C$_3$ grasses to both

493   observation-based estimates lie between 0.34-0.38 and between around 15-17%, respectively.

494   The spatial correlation and RMSD between the two observation-based estimates themselves are

495   0.54 and around 12%, respectively.

496

497   Figure 8b shows the distribution of C$_4$ grasses which mostly occur in the tropics and do not

498   occupy large areas in North America (as was also seen in Figure 2b). The modelled geographical

499   distribution of C$_4$ grasses is larger than observation-based estimates but the absolute fractions

500   remain small so that the simulated area covered over the whole domain (0.35×10^6 km^2) is

501   actually smaller than estimates based on the modified WANG06 (0.45×10^6 km^2) and MODIS

502   (0.7×10^6 km^2) data sets. The spatial correlation and RMSD when comparing simulated

503   coverage of C$_4$ grasses to both observation-based estimates lie between 0.12-0.16 and between

504   around 3-5%, respectively. The spatial correlation and RMSD between the two observation-

505   based estimates themselves are 0.62 and around 5%, respectively.

506

507   ### 3.2.5   Broadleaf evergreen and drought deciduous trees

508

509   The least prevalent PFTs in the North American domain considered here are broadleaf evergreen

510   (BDL EVG) and broadleaf drought deciduous (BDL DCD DRY) trees. As they are represented

511   in the model these are primarily tropical PFTs and hence generally do not exist above around 30

512   °N (see Figure 9), according to the bioclimatic limits used in the model for tree PFTs (Melton

513   and Arora, 2016). In our simulations, these PFTs therefore exist near the southern edge of the

514   United States. We do not evaluate spatial correlation and RMSD for these PFTs compared to the

515   two observation-based estimates for three reasons: 1) the geographical distribution of these PFTs

516   is limited to a small total area, 2) the geographical distribution of the BDL EVG tree PFT based

517   on observations cannot be directly compared to simulated values because, when mapping land

518   cover types to CTEM PFTs in WANG06, evergreen shrubs (which exist much farther north than

519   30 °N) are assigned to the the BDL EVG tree PFT, and 3) the geographical distribution of the

520   BDL DCD DRY tree PFT in the observation-based data sets is based on the arbitrary latitudinal

521   thresholds of 24 °N and 34 °N as mentioned earlier.

522

3.3 LAI and GPP

524

525 Figure 10 compares the geographical distribution of simulated LAI and GPP with observation-

526 based estimates for the present day. In Figure 10a, the simulated geographical distribution of

527 LAI compares well with the observation-based estimates. The spatial correlation and RMSD

528 between simulated and observation-based estimates are 0.74 and 0.81 m^2/m^2, respectively. The

529 domain averaged simulated LAI of 2.5 m^2/m^2 is higher than the observation-based estimate of

530 2.1 m^2/m^2. The model captures the broad geographical patterns with higher LAI over the boreal

531 forest region in Canada and also in the eastern United States similar to observations. However,

532 some differences remain particularly over the drier southwest United States where the model

533 simulates bare ground with negligible LAI but observations suggest a small LAI of around 1

534 m^2/m^2. In contrast, the model slightly overestimates LAI over northern and Arctic Canada where

535 it simulates a higher fractional coverage of C_3 grasses, as seen earlier.

536

537 Consistent with the geographical distribution of LAI, the simulated GPP is overestimated in the

538 eastern United States and the Canadian boreal forest (Figure 10b). The broad geographical

539 distribution of GPP, similar to LAI, is consistent with the observation-based estimates. The

540 spatial correlation and RMSD between simulated and observation-based estimates are 0.78 and

541 225 gC/m^2.year, respectively. The domain averaged simulated GPP of 737 gC/m^2.year is higher

542 than the observation-based estimate of 628 gC/m^2.year. As with LAI, the simulated GPP is lower

543 than observations over the drier southwest region of the United States where the model simulates

544 more bare ground than observation-based estimates, and the model overestimates GPP over the

545 northern and Arctic Canada.

546

547 Figure 11 shows the time series of annual domain averaged GPP, LAI, net primary productivity

548 (NPP) and domain summed net biome productivity (NBP). The NBP term is essentially the net

549 atmosphere-land CO_2 flux which is the result of all terrestrial ecosystem processes including

550 photosynthesis, autotrophic and heterotrophic respiration, fire and land use change. NBP values

551 of zero indicate that the system is in equilibrium such that carbon gained by photosynthesis is

552 equal to carbon lost by respiration and other processes. Simulated GPP, LAI and NPP all show

Biogeosciences
Discussions

Open Access

553 an increase over the 20th century due to the increase in atmospheric CO_2 concentration and the

554 associated change in climate. The increase in CO_2 drives the increase in GPP and subsequently in

555 NPP and LAI through the CO_2 fertilization effect. The net result of this gradually increasing NPP

556 is that the terrestrial ecosystems become a sink of carbon and this is seen in the resulting positive

557 values of NBP. The simulated sink over the North American domain for the periods 1990-2000

558 and 2000-2010 is around 0.4 and 0.5 Pg C/year, respectively. Crevoisier et al. (2010) compare

559 the carbon sink over the North American region from five studies (their Table 1) for time periods

560 in the 1990s and 2000s. These reported sinks vary from 0.81±0.72 to 1.26±0.23 Pg C/year for the

561 period 1992-1996, 0.58 Pg C/yr for the period 2001-2006 and Crevoisier et al. (2010) themselves

562 estimate a value of 0.51 ± 0.41 Pg C/yr for the period 2004-2006. The sinks simulated by

563 CLASS-CTEM over the 1990s and 2000s are broadly consistent with these estimates.

564

565 **4   Discussion**

566

567 Allowing a terrestrial ecosystem model to simulate fractional coverages of its PFTs adds another

568 degree of freedom to the model compared to the case where the fractional coverages of its PFTs

569 are specified. This is a more stringent test of a model's performance. Errors in the simulated

570 geographical distribution of PFTs will, of course, lead to corresponding errors in the

571 geographical distribution of primary terrestrial ecosystem carbon pools and fluxes. Yet, the

572 CLASS-CTEM model is broadly able to reproduce the geographical distributions of GPP and

573 LAI. Limitations, of course, remain. In particular, the simulated LAI and GPP are high in Alaska

574 and in northern and Arctic Canada, and these variables are lower than their observation-based

575 estimates in arid regions of the western United States. The simulated fractional vegetation

576 coverage reflects these patterns.

577

578 It is difficult to conclusively determine whether these model limitations are due to the limitations

579 in the biogeochemistry parameterizations of the model for its existing PFTs or the simple

580 structural limitation that the model does not represent shrub, moss and lichen PFTs. Shrubs are

581 adapted to grow in arid and semi-arid regions, whether in cold or hot climates (where neither

582 grasses nor trees are able to grow) and their representation in the model would likely help to

583 increase the fractional vegetation cover in arid regions including those in the western United

584   States. At high latitudes grass growth is inhibited by mosses and lichens which flourish in cold

585   and damp conditions. A representation of moss and lichen PFTs and improved representation of

586   permafrost in the model would likely help to decrease simulated grass coverage in Arctic

587   regions. In the current version of the CLASS-CTEM model bioclimatic limits are used only for

588   tree PFTs to ensure that these PFTs do not venture outside their pre-determined bioclimatic

589   zones. In the model, bioclimatic limits are not used for grasses and their geographical

590   distribution is entirely the result of plant physiological processes and their competitive

591   interactions with the tree PFTs and amongst themselves. Since, in the Arctic region, grasses do

592   not face competition from tree PFTs, and moss and lichen PFTs are not represented in the model,

593   they are free to increase their expanse – climate permitting, of course. Another possible reason

594   for higher than observed grass coverage in the Arctic region is that in the current implementation

595   of CLASS only three permeable soil layers with maximum thicknesses of 0.1, 0.25 and 3.75 m

596   are represented and a boundary condition of zero heat flux is assumed across the bottommost

597   layer. This simple representation does not allow to model permafrost realistically. Permafrost is

598   more realistically modelled with multiple permeable and impermeable (extending into the bed

599   rock) layers that go sufficiently deep (> 30 m at least) to capture the slow evolution of soil

600   temperatures in response to climate warming (Teufel et al., 2017). The current set up of three

601   layers that go only 4.1 m deep produces soil temperatures that are warmer than in the set up

602   when permeable and impermeable layers are sufficiently deep and produces permafrost extent

603   that is lower than observation-based estimates (Koven et al., 2013). It is likely that warmly

604   biased soil temperatures in the current set up contribute to promote grass growth and allow it to

605   cover a larger area in the Arctic region than would be the case when permafrost is more

606   realistically modelled.

607

608   The lower than observed fractional vegetation cover in the arid and semi-arid regions of the

609   western United States, however, may not solely be due to model limitations alone. Here, we

610   argue that the manner in which remotely sensed land cover types are mapped to CTEM PFTs,

611   and the errors in calculating bare ground fraction in remotely sensed products also contribute to

612   mismatch between modelled and observation-based values of fractional vegetation cover. We

613   illustrate this by comparing the functional relationship between LAI and total vegetation cover.

614   Figure 12a shows this relationship for model simulated values. As expected, as LAI increases so

Biogeosciences
Discussions

615   does the total vegetation cover. The relationship between these two variables is fairly tight in the

616   model and the green line is an exponential fit. The red dots in the figure correspond to grid cells

617   that lie in the region identified in the inset in Figure 12d and broadly correspond to the western

618   half of the United States. Figures 12b and 12c show the same relationship but between the

619   observation-based estimate of LAI from Zhu et al. (2013) (as mentioned in Section 2.3.2) and the

620   total vegetation cover based on the WANG06 and MODIS derived land cover data sets,

621   respectively. The blue and magenta lines in Figures 12b and 12c are the corresponding

622   exponential fits. When compared with Figure 12a, Figures 12b and 12c show much more scatter

623   around the fitted curves, and the overall relationship appears to break down for the red dots

624   corresponding to the grid cells in the western United States. A careful look at the red dots in

625   Figures 12b and 12c shows that the observation-based vegetation cover in the Western United

626   States for a large fraction of grid cells is around 60% regardless of the observation-based LAI

627   which ranges between 0.1 and 1.5 m^2/m^2. Clearly, it is physically unrealistic to achieve fractional

628   vegetation coverage of 60% below LAI values of 0.6 m^2/m^2 (the m^2/m^2 unit implies m^2 of leaf

629   area per m^2 of ground area) and this indicates that the fractional vegetation cover in this region is

630   likely overestimated in both observation-based data sets.

631

632   There are at least two ways in which errors in total vegetation cover can occur. The first relates

633   to the method by which the fractional vegetation cover is calculated for the land cover types in

634   the original remotely sensed land cover products: that is, for the 22 land cover types in the

635   GLC2000 data set upon which the WANG06 data are based and the 17 land cover types in the

636   MODIS data set. An example of such an error for arid regions is illustrated by Lawley et al.

637   (2014) who suggest that the MODIS soil fractional cover product, at least in its present form, is

638   unsuited to monitoring sparsely vegetated arid landscapes and generally unable to separate soil

639   from vegetation in situations where normalized difference vegetation index (NDVI) is low. The

640   second way in which errors are introduced is through the mapping of the remotely sensed land

641   cover types to the CTEM PFTs following Table 2 of WANG06 for the GLC2000 land cover

642   types, and following Table 2 in this manuscript for the MODIS land cover types. This mapping

643   is based on available information in the literature but is also based on expert judgement which

644   introduces subjectiveness. For instance, it is debatable what fraction of the "open shrublands"

645   MODIS land cover type, which exists over much of the arid southwestern United States, is in

646 fact bare ground. In Table 2, we have allocated a value of 0.4 to this fraction following

647 WANG06. Had WANG06 allocated a higher value than this to bare ground, our simulated values

648 would have compared better with the observation-based values of bare ground fraction over arid

649 regions. Nevertheless this would not have changed the relationship, or rather the lack thereof,

650 between the observation-based estimates of LAI and the total vegetation cover in the western

651 half of the United States seen in Figures 12b and 12c.

652

653 The simulated areas covered by the primary two tree PFTs (NDL EVG and BDL DCD COLD)

654 have their weaknesses but large differences also exist between the two observation-based

655 estimates especially for the NDL EVG PFT. Modelling competition between two tree PFTs is

656 much more difficult than between trees and grasses. In the latter case trees are always considered

657 superior to grasses, but in the case of competition between two tree PFTs the superiority is based

658 on parameterized colonization rates which depend on simulated NPP. Based on comparisons

659 with observation-based estimates, the main limitation in model results here is that the model

660 overestimates the coverage of NDL EVG trees, and underestimates the coverage of BDL DCD

661 COLD trees in the eastern United States, while the opposite is true in western Canada. The

662 model, of course, does not represent individual species, while in the real world competition

663 occurs at the species level. One example of how additional PFTs in the CLASS-CTEM

664 framework can lead to improved model performance is illustrated by Peng et al., (2014). This

665 application of the model shows how sub-dividing the NDL EVG PFT into coastal and interior

666 types for the province of British Columbia in Canada leads to improvement in simulated LAI and

667 GPP. A recent attempt to explicitly represent physiological process in a model to simulate

668 competition between needleleaf and broadleaf cold deciduous trees at a regional scale is

669 illustrated in (Fisher et al., 2015) who incorporated the concepts from the Ecosystem

670 Demographics (ED) model into the community land model – dynamic vegetation model (CLM-

671 DGVM). Their results provide some interesting insights; however, validation of this approach at

672 the global scale over a wide range of PFTs remains challenging.

673

674 Finally, one of the objectives of this study was to evaluate if resolving climate niches by

675 performing CLASS-CTEM simulation at a finer resolution of 1° in this study allowed improved

676 simulation of geographical distribution of PFTs than in the Melton and Arora (2016) study that

23

677   evaluated the competition module of the CLASS-CTEM model at 3.75° spatial resolution at the

678   global scale. Figures 2 through 4 of Melton and Arora (2016) compare simulated geographical

679   distributions of PFTs with WANG06 data. Comparing their results over North America with

680   ones obtained here we note that the primary model limitations remain unchanged in the

681   application of the model at both spatial resolutions. These include lower simulated fractional

682   vegetation coverage in the arid south-west North American region and higher in the Arctic

683   region (due to higher grass coverage). In addition, in both applications of the model the

684   differences in simulated geographical distribution of NDL EVG and BDL DCD CLD PFTs,

685   compared to the WANG06 land cover data, are also similar. Model differences, compared to the

686   WANG06 data, therefore remain more or less similar in the application of the model at both

687   spatial resolutions.

688

689   The comparison between observation-based and simulated fractional coverages is the most

690   robust at the basic tree-grass-bare ground level. The subjectiveness introduced in the process of

691   mapping remotely sensed land cover types to the PFTs represented in a model, as mentioned

692   above, makes the comparison of simulated and observation-based fractional coverages for

693   individual PFTs less robust. Nevertheless, comparisons with observations allow useful insights

694   into model limitations as we have seen here.

695

696

697   ## 5   Summary and conclusions

698

699   This study evaluates the CLASS-CTEM simulated fractional coverages of PFTs, when driven

700   with observed climate forcing, against the observation-based estimates from MODIS and the

701   modified WANG06 data sets over the North American region. In the past, performance of the

702   competition module of the CLASS-CTEM modelling framework has been assessed at global

703   scale, at a coarse spatial resolution of 3.75° (Melton and Arora, 2016), as well as at point scale,

704   for a range of locations across the globe (Shrestha et al., 2016). Our objective here was to assess

705   the performance of the CLASS-CTEM competition module at a higher spatial resolution of 1°

706   over North America. To achieve this objective we compared simulated present day geographical

distributions of fractional coverages of PFTs, but also LAI and GPP with their observation-based estimates.

The CLASS-CTEM modelling framework is generally able to reproduce the dominant features of the geographic distribution of PFT coverage, and LAI and GPP over the North American region. After 1960, the model simulates increasing GPP and LAI in response to changing climate as well as increased atmospheric CO_2 concentrations and the resulting sink for the 1990s and 2000s is broadly consistent with other estimates.

The simulated geographical distribution of PFTs, when compared to observation-based estimates, show two primary limitations which are excessive grass cover in the Arctic region and low vegetation cover in the arid western United States, although for the latter the observation-based estimates themselves may have their own weaknesses. There are three main factors in the CLASS-CTEM modelling framework that may have contributed to these differences: 1) the absence of a shrub PFT, which we believe is the reason for low simulated vegetation coverage in the arid to semi-arid western United States, 2) the absence of moss and lichen PFTs that may inhibit the establishment of grasses, and 3) probably a lack of sensitivity of C_3 grasses to high latitude climate and an inadequate representation of permafrost. Future model developments will focus on these aspects with a view to improving model performance.

Biogeosciences
Open Access
Discussions
EGU

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. Climate, 26, 6801-6843, doi:10.1175/JCLI-D-12-00417.1, 2013.

Arora, V. K.: Land surface modelling in general circulation models: a hydrological perspective, PhD Thesis, University of Melbourne, Australia, 1997.

Arora, V. K. and Boer, G. J.: A representation of variable root distribution in dynamic vegetation models, Earth Interactions, 7, 1-19, 2003.

Arora, V. K. and Boer, G. J.: A parameterization of leaf phenology for the terrestrial ecosystem component of climate models, Glob. Change Biol., 11, 39-59, 2005.

Arora, V. K. and Boer, G. J.: Simulating Competition and Coexistence between Plant Functional Types in a Dynamic Vegetation Model, Earth Interactions, 10, 1-29, 2006.

Arora, V. K., Boer, G. J., and Friedlingstein, P. E. A.: Carbon-concentration and carbon-climate feedbacks in CMIP5 earth system models, J. Climate, 26, 5289-5314, 2013.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rodenbeck, C., Arain, M. A., Baldocchih, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Vivoy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate, Science, 329, 834-838, doi:10.1126/science.1184984, 2010.

Bonan, G. B.: Forests and climate change: Forcings, feedbacks, and the climate benefits of forests, Science, 320, 1444-1449, doi:10.1126/science.1155121, 2008.

Box, E. O.: Plant functional types and climate at the global scale, J. Veg. Sci., 7, 309-320, doi:10.2307/3236274, 1996.

Brentnall, S. J., Beerling, D. J., Osborne, C. P., Harland, M., Francis, J. E., Valdes, P. J., and Wittig, V. E.: Climatic and ecological determinants of leaf lifespan in polar forests of the high CO2 Cretaceous 'greenhouse' world, Glob. Change Biol., 11, 2177-2195, 2005.

Broxton, P., X. Zeng, D. Sulla-Menashe, and P. Troch, 2014: A Global Land Cover Climatology Using MODIS Data. J. Appl. Meteor. Climatol., 53, 1593–1605, doi: 10.1175/JAMC-D-13-0270.1.

Cox, P. 2001: Description of the "TRIFFID" Dynamic Global Vegetation Model. Tech. Note 24.

Biogeosciences
Discussions

Open Access

EGU

770   Cramer, W., Bondeau, A., Woodward, F. I., Prentice, I. C., Betts, R. A., Brovkin, V., Cox, P. M.,
771   Fisher, V., Falloon, P. D., Foley, J., Friend, A. D., Kucharik, C., Lomas, M. R.,
772   Ramankutty, N., Sitch, S., Smith, B., White, A., and Molling-Young, C.: Global response
773   of terrestrial ecosystem structure and function to CO2 and climate change: results from
774   six dynamic global vegetation models, Global Change Biology, 7, 357-373, 2001.
775
776   Crevoisier, C., Sweeney, C., Gloor, M., Sarmiento, J. L., and Tans, P. P.: Regional US carbon
777   sinks from the three-dimensional atmospheric CO2 sampling, PANAS, 107, 18348-
778   18353, DOI: 10.1073/pnas.0900062107, 2010.
779
780   Dai, Y., Zeng, X., Dickinson, R. E., and Coauthors: Common Land Model (CLM), Technical
781   documentation and user's guide. [Available online at
782   [http://climate.eas.gatech.edu/dai/clmdoc.pdf.], 2001.
783
784   Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., Mcdowell, N. G., Knox, R.
785   G., Koven, C., Holm, J., Rogers, B. M., Spessa, A., Lawrence, D., and Bonan, G.: Taking
786   off the training wheels: the properties of a dynamic vegetation model without climate
787   envelopes, CLM4.5(ED), Geosci. Model Dev., 8, 3593-3619, doi:10.5194/gmd-8-3593-
788   2015, 2015.
789
790   Friedl, M., Strahler, A., Schaaf, C., Hodges, J. C. F., and Salomon, J., 2013:.: Binary MODIS
791   MOD12C1 0.25 Degree Land Cover Climate Modeler Grid. Available at
792   [http://duckwater.bu.edu/lc/] from the Department of Geography, Boston University,
793   Boston, Massachusetts, USA., 2013.
794
795   Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney,
796   S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M.,
797   Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner,
798   E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and
799   Zeng, N.: Climate–carbon cycle feedback analysis: Results from the C4MIP model
800   intercomparison, J. Climate, 19, 3337-3353, 2006.
801
802   Friend, A. D., Chard, Lucht, W., Rademacher, T., Keribin, R., Betts, R., Cadule, P., Ciais, P.,
803   Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R.,
804   Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N.,
805   Warszawski, L., Wiltshire, A., and Woodward, F. I.: Carbon residence time dominates
806   uncertainty in terrestrial vegetation response to future climate and atmospheric CO2,
807   Proc. Natl. Acad. Sci. USA, 111, 3280-3285, 2013.
808
809   Garnaud, C., Sushama, L., and Verseghy, D.: Impact of interactive vegetation phenology on the
810   Canadian RCM simulated climate over North America, Climate Dynamics, 45, 1471-
811   1492, doi:10.1007/s00382-014-2397-9, 2015.
812
813   Gobron, N., Belward, A., and Knorr, W.: Monitoring biosphere vegetation 1998-2009, Geophys.
814   Res. Lett., 37, L15402, 2010.
815

Biogeosciences
Discussions

816 Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P.,
817   Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T.,
818   Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A.,
819   Thornton, P., Van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios
820   for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood
821   harvest, and resulting secondary lands, Climate Change, 109, 117-161, 2011.
822
823 Kramer, P. J. and Kozlowski, T. T., 1979:.: Physiology of woody plants. Academic press, 1979.
824
825 Koven, C. D., Riley, W. J. and Stern, A.: Analysis of Permafrost Thermal Dynamics and
826   Response to Climate Change in the CMIP5 Earth System Models, J. Clim., 26(6), 1877–
827   1900, 2013.
828
829 Lawley, E. F., Lewis, M. M., and Ostendorf, B.: Evaluating MODIS soil fractional cover for arid
830   regions, using albedo from high-spatial resolution satellite imagery, Int. J. Remote Sens.,
831   35, 2028-2046, 2014.
832
833 Melton, J. R. and Arora, V. K.: Competition between plant functional types in the Canadian
834   Terrestrial Ecosystem Model (CTEM) v. 2.0, Geosci. Model Dev., 9, 323-361, 2016.
835
836 Peng, Y., Arora, V. K., Kurz, W. A., Hember, R. A., Hawkins, B. J., Fyfe, J. C., and Werner, A.
837   T.: Climate and atmospheric drivers of historical terrestrial carbon uptake in the province
838   of British Columbia, Canada, Biogeosciences, 11, 635-649, doi:10.5194/bg-11-635-2014,
839   2014.
840
841 Pielke, R. A., Avissar, R., Raupach, M., Dolman, A. J., Zeng, X., and Denning, S.: Interactions
842   between the atmosphere and terrestrial ecosystem: influence on weather and climate,
843   Global Change Biology, 4, 461-475, 1998.
844
845 Ramankutty, N. and Foley, J. A.: Estimating historical changes in global land cover: Croplands
846   from 1700 to 1992, Global Biogeochem. Cycles, 13, 997-1027, 1999.
847
848 Ran, L., Pleim, J., Gilliam, R., Binkowski, F. S., Hogrefe, C., and Band, L.: Improved
849   meteorology from an updated WRF/CMAQ modeling system with MODIS vegetation
850   and albedo, J. Geophys. Res. Atmos., 121, 2393-2415, doi:10.1002/2015JD024406, 2016.
851
852 Ritchie, T. C. and Macdonald, G. M.: The patterns of post-glacial spread of White Spruce, J.
853   Biogeogr., 13, 527-540, 1986.
854
855 Shrestha, R. K., Arora, V. K., and Melton, J. R.: The sensitivity of simulated competition
856   between different plant functional types to sub-grid-scale representation of vegetation in
857   a land surface model, J. Geophys. Res. Biogeosci., 121, doi:10.1002/2015JG003234,
858   2016.
859
860 Siemann, E. and Rogers, W. E.: Changes in light and nitrogen availability under pioneer trees
861   may indirectly facilitate tree invasion of grasslands, J. Ecology, 91, 923-931, 2003.

Biogeosciences
Discussions

Open Access

EGU

862

863   Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Lucht,
864   W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics,
865   plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation
866   model, Glob. Change Biol., 9, 161-185, 2003.
867

868   Teufel, B., Sushama, L., Arora, V., and Verseghy, D.: Impact of dynamic vegetation phenology
869   on the simulated pan-Arctic land surface state, submitted to Climate Dynamics, 2017.
870

871   Timmons, D., Buchholz, T., and Veeneman, C. H.: Forest biomass energy: assessing
872   atmospheric carbon impacts by discounting future carbon flows, GCB Bioenergy, 8, 631-
873   643, 10.1111/gcbb.12276, 2016.
874

875   Verseghy, D. L., Mcfarlane, N. A., and Lazare, M.: CLASS - A Canadian land surface scheme
876   for GCMs, II. Vegetation model and coupled runs, Int. J. Climatol., 13, 347-370, 1993.
877

878   Viovy,   N.:   CRU-NCEP   reanalysis   data   version   4.   Available   at
879   [http://dods.extra.cea.fr/store/p529viov/cruncep/V4_1901_2012/], (Accessed May 2015),
880   2012.
881

882   Wang, A., Price, D. T., and Arora, V. K.: Estimating changes in global vegetation cover (1850-
883   2100) for use in climate models, Global Biogeochem. CycesCycles, 20, GB3028, 2006.
884

885   Wang, G., Sun, S., and Mei, R.: Vegetation dynamics contributes to the multi-decadal variability
886   of precipitation in the Amazon region, Geophys. Res. Lett., 38, L19703,
887   doi:10.1029/2011GL049017, 2011.
888

889   Wen, X., Tang, G., Wang, S., and Huang, J.: Comparison of global mean temperature series,
890   Advances in Climate Change Research, 2, 187-192, 2011.
891

892   Zhang, Z., Xue, Y., Macdonald, G., Cox, P. M., and Collatz, G. J.: Investigation of North
893   America vegetation variablilityvariability under recent climate: A study using the
894   SSiB4/TRIFFID biophysical/dynamic vegetation model, J. Geophys. Res. Atmos., 120,
895   1300-1321, 2015.
896

897   Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R., and
898   Myneni, R. B.: Global data sets of vegetation leaf area index (LAI)3g and fraction of
899   photosynthetically active radiation (FPAR)3g derived from global inventory modeling
900   and mapping studies (GIMMS) normalized difference vegetation index (NDVI3g) for the
901   period 1981- to 2011, Remote Sens., 5, 927-948, doi:10.3390/rs5020927, 2013.
902

903   Zobler, L.: A world soil file for global climate modelling, Tec. Rep. NASA TM-87802, 14-32,
904   1986.
905

906

Biogeosciences
Discussions

Open Access

EGU

907
908
909   Table 1: Plant functional types (PFTs) represented in CTEM and their relation to CLASS PFTs.
910

911

CLASS PFTs	CTEM PFTs	CTEM PFT Symbol
Needleleaf trees	Needleleaf Evergreen trees	NDL-EVG
	Needleleaf Deciduous trees	NDL-DCD
Broadleaf trees	Broadleaf Evergreen trees	BDL-EVG
	Broadleaf Cold Deciduous trees	BDL-DCD-CLD
	Broadleaf Drought/Dry Deciduous trees	BDL-DCD-DRY
Crops	C_3 Crops	CROP-C3
	C_4 Crops	CROP-C4
Grasses	C_3 Grasses	GRASS-C3
	C_4 Grasses	GRASS-C4

Table 2: Reclassification of the 17 MODIS land cover classes into the nine CTEM PFTs

SN	Items	Tree				Crop	Grass	Bare	Reference
		NDL EVG	NDL DCD	BDL EVG	BDL DCD				
1	Woody Savanna			0.1	0.4		0.25	0.25	Dai et al. (2001)
2	Water bodies							1	
3	Urban built up areas	0.05			0.05		0.1	0.8	Dai et al. (2001)
4	Savanna			0.05	0.3		0.4	0.25	Wang et al. (2006)
5	Permanent Wetlands						0.25	0.75	Dai et al. (2001)
6	Permanent snow and ice							1	Wang et al. (2006)
7	Open Shurblands	0.1			0.15		0.35	0.4	Wang et al. (2006)
8	Needleleaf evergreen	1							Wang et al. (2006)
9	Needleleaf deciduous		0.8				0.1	0.1	Wang et al. (2006)
10	Mixed forest	0.45			0.45		0.1		Wang et al. (2006)
11	Grasslands						0.65	0.35	Wang et al. (2006)
12	Croplands					0.9		0.1	Wang et al. (2006)
13	Cropland natural veg. mosaic			0.2		0.5	0.2	0.1	Wang et al. (2006)
14	Closed shrublands	0.2	0.2		0.4		0.2		Wang et al. (2006)
15	Broadleaf evergreen			1					Wang et al. (2006)
16	Broadleaf deciduous				1				Wang et al. (2006)
17	Bare ground							1	Wang et al. (2006)

1

2

3
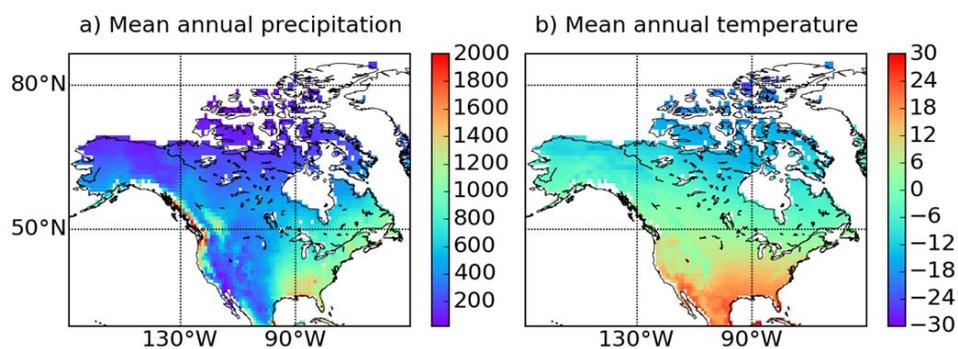
4

5   Figure 1. Spatial distribution of mean annual a) precipitation (mm), and b) temperature (°C)

6   across North America. Grid cells with permanent ice/glaciers have been masked out.

7
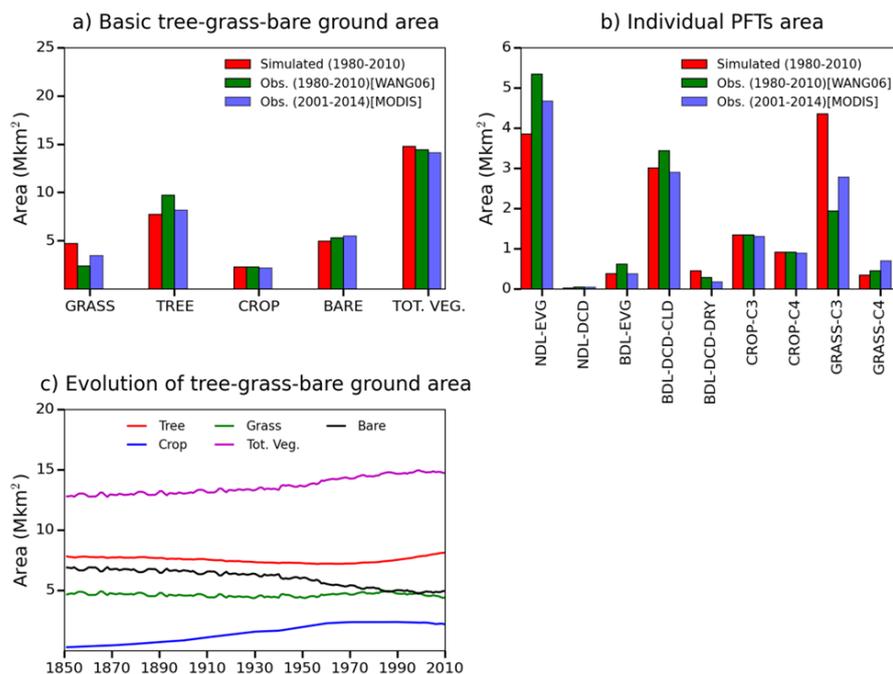
Figure 2. Comparison of observation-based and simulated vegetation areas summed over the North American domain a) grass, treed, crop, bare ground and total vegetated area, b) individual PFT areas, and c) evolution of simulated vegetation areas summed over the domain.

14
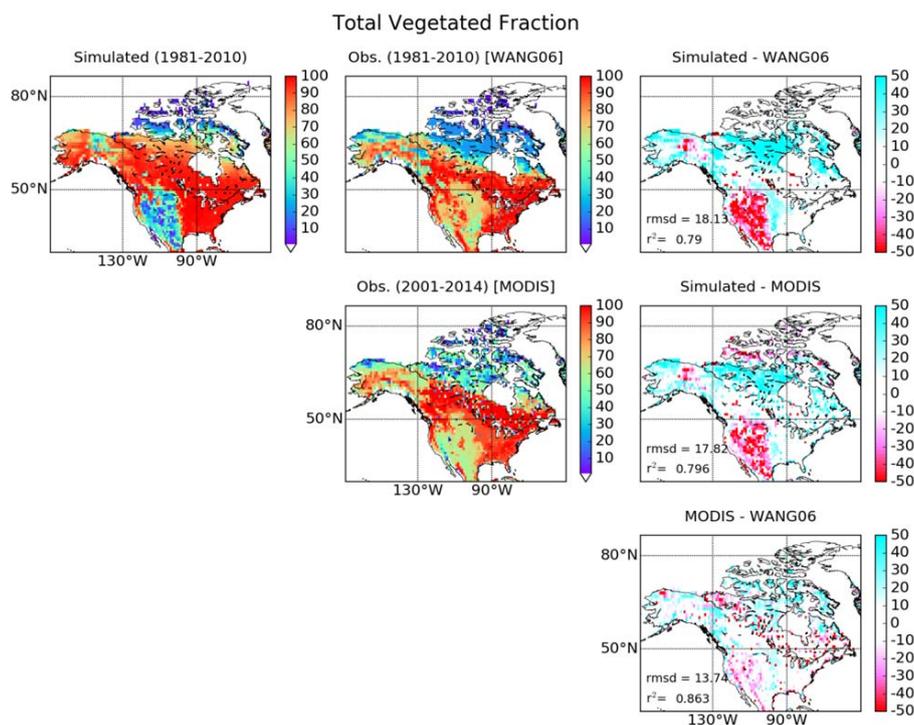
15

16   Figure 3. Spatial distribution of total vegetated coverage across North America. Simulated,

17   observation-based, and differences are presented in the left, middle and right columns,

18   respectively. The differences column includes model biases with respect to WANG06 (top panel)

19   and MODIS (middle panel), and the difference between the two observation-based estimates

20   (bottom panel). Root mean square difference (rmsd) and coefficient of determination (r^2) are also

21   shown in each case.

22

23
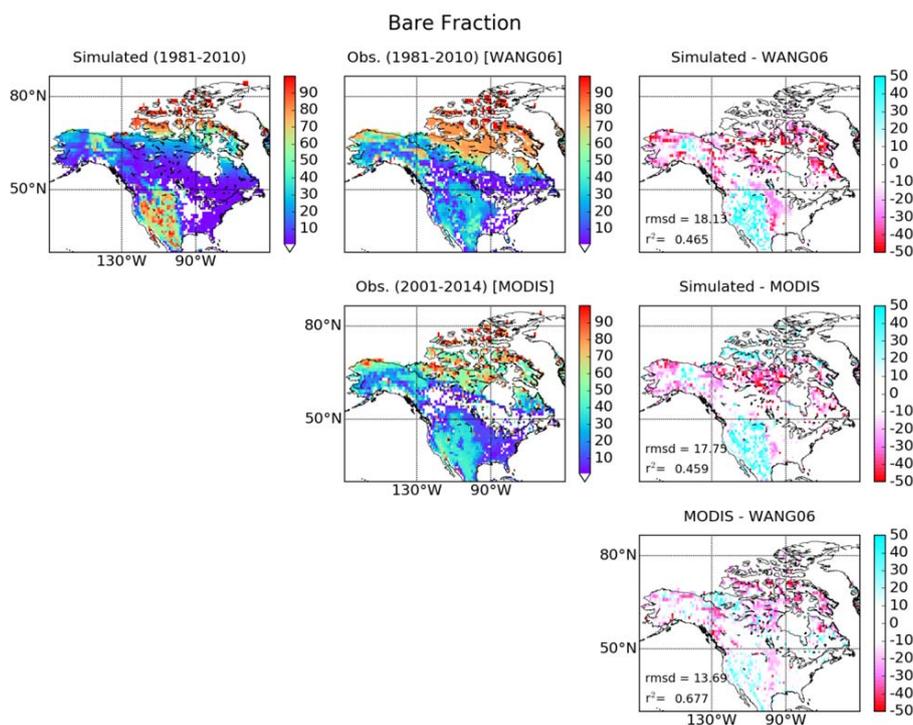
24

25   Figure 4. Spatial distribution of bare ground coverage across North America. Simulated,

26   observation-based, and differences are presented in the left, middle and right columns,

27   respectively. The differences column includes model biases with respect to WANG06 (top panel)

28   and MODIS (middle panel), and the difference between the two observation-based estimates

29   (bottom panel). Root mean square difference (rmsd) and coefficient of determination (r^2) are also

30   shown in each case.

31

32
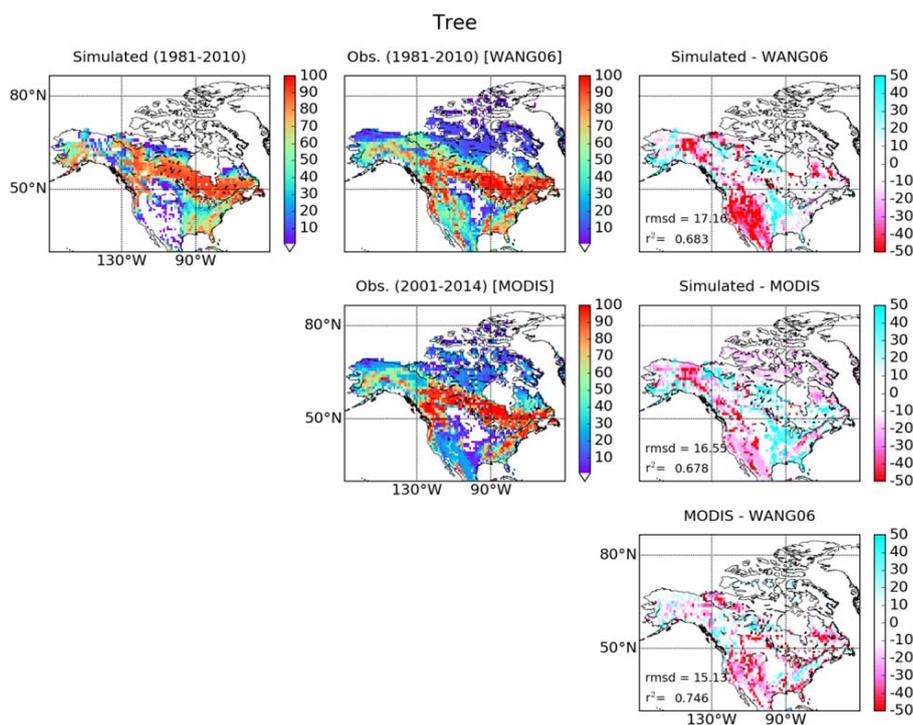
Figure 5. Spatial distribution of tree coverage across North America. Simulated, observation-based, and differences are presented in the left, middle and right columns, respectively. The differences column includes model biases with respect to WANG06 (top panel) and MODIS (middle panel), and the difference between the two observation-based estimates (bottom panel). Root mean square difference (rmsd) and coefficient of determination (r^2) are also shown in each case.

41

42
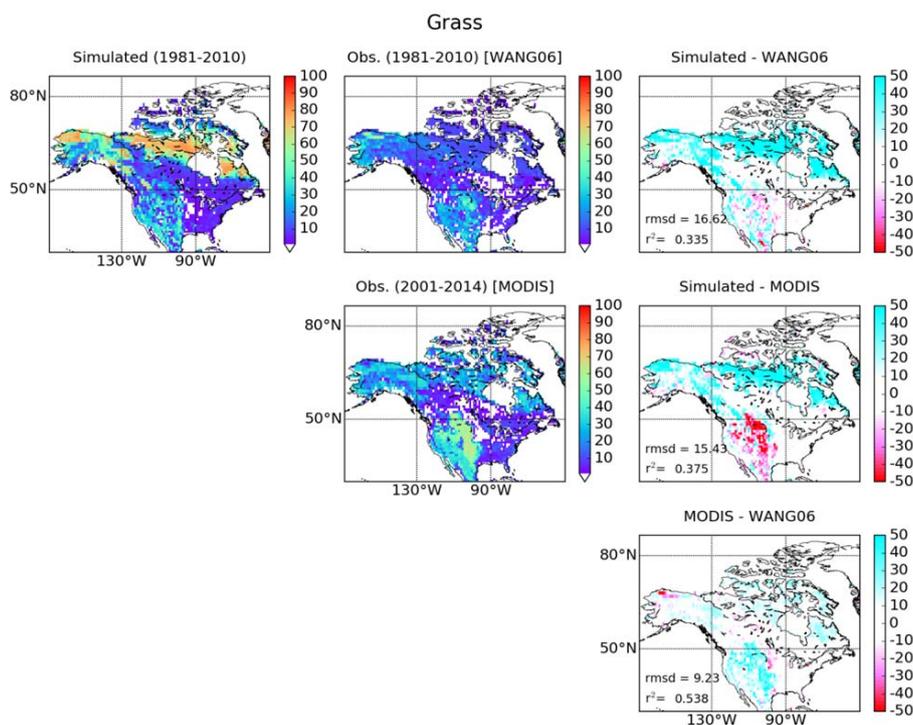
Figure 6. Spatial distribution of grass coverage across North America. Simulated, observation-based, and differences are presented in the left, middle and right columns, respectively. The differences column includes model biases with respect to WANG06 (top panel) and MODIS (middle panel), and the difference between the two observation-based estimates (bottom panel). Root mean square difference (rmsd) and coefficient of determination (r^2) are also shown in each case.

49

50

51

52
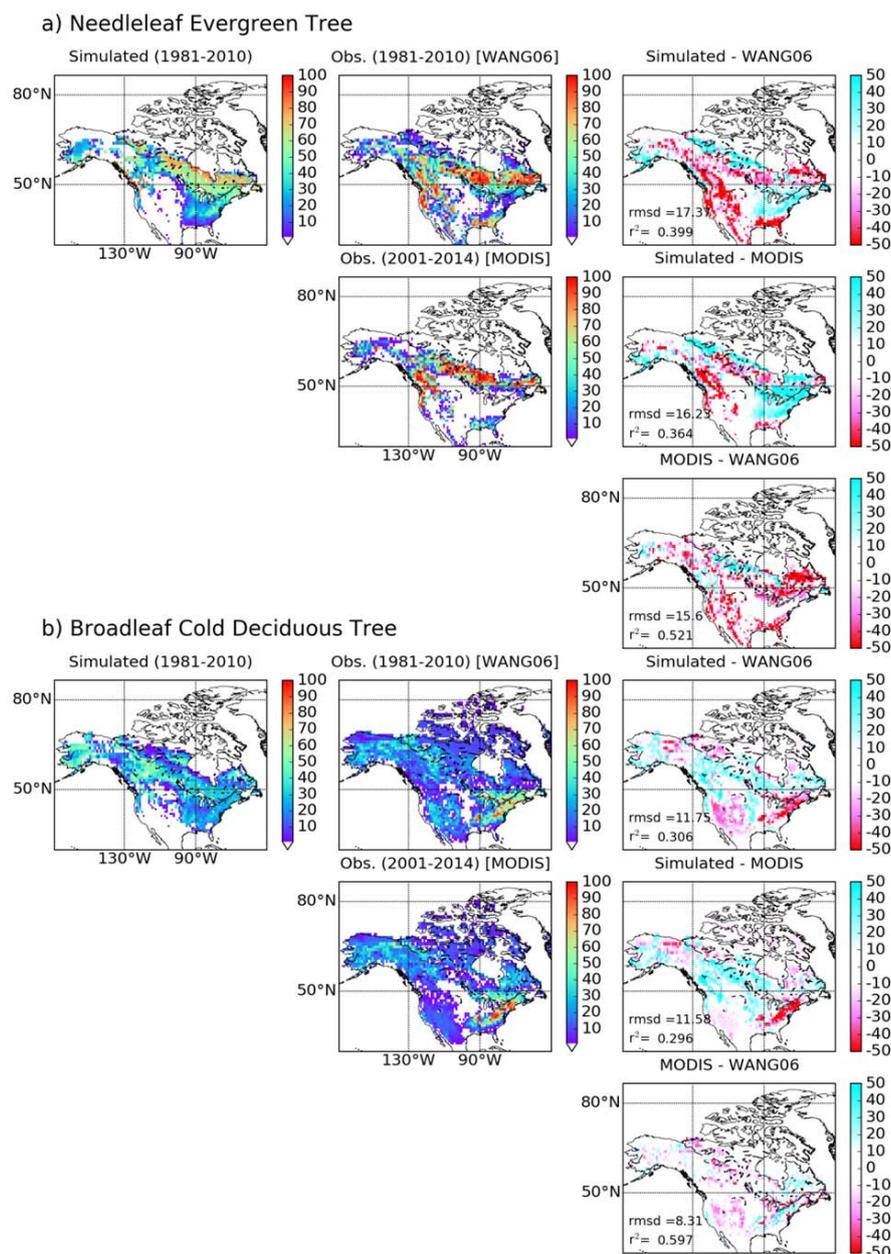
Figure 7. Spatial distribution of a) needleleaf evergreen tree, and b) broadleaf cold deciduous tree across North America. Simulated, observation-based, and differences are presented in the left, middle and right columns, respectively. The differences column includes model biases with respect to WANG06 (top panel) and MODIS (middle panel), and the difference between the observation-based estimates (bottom panel). Root mean square difference (rmsd) and coefficient of determination (r^2) are also shown in each case.
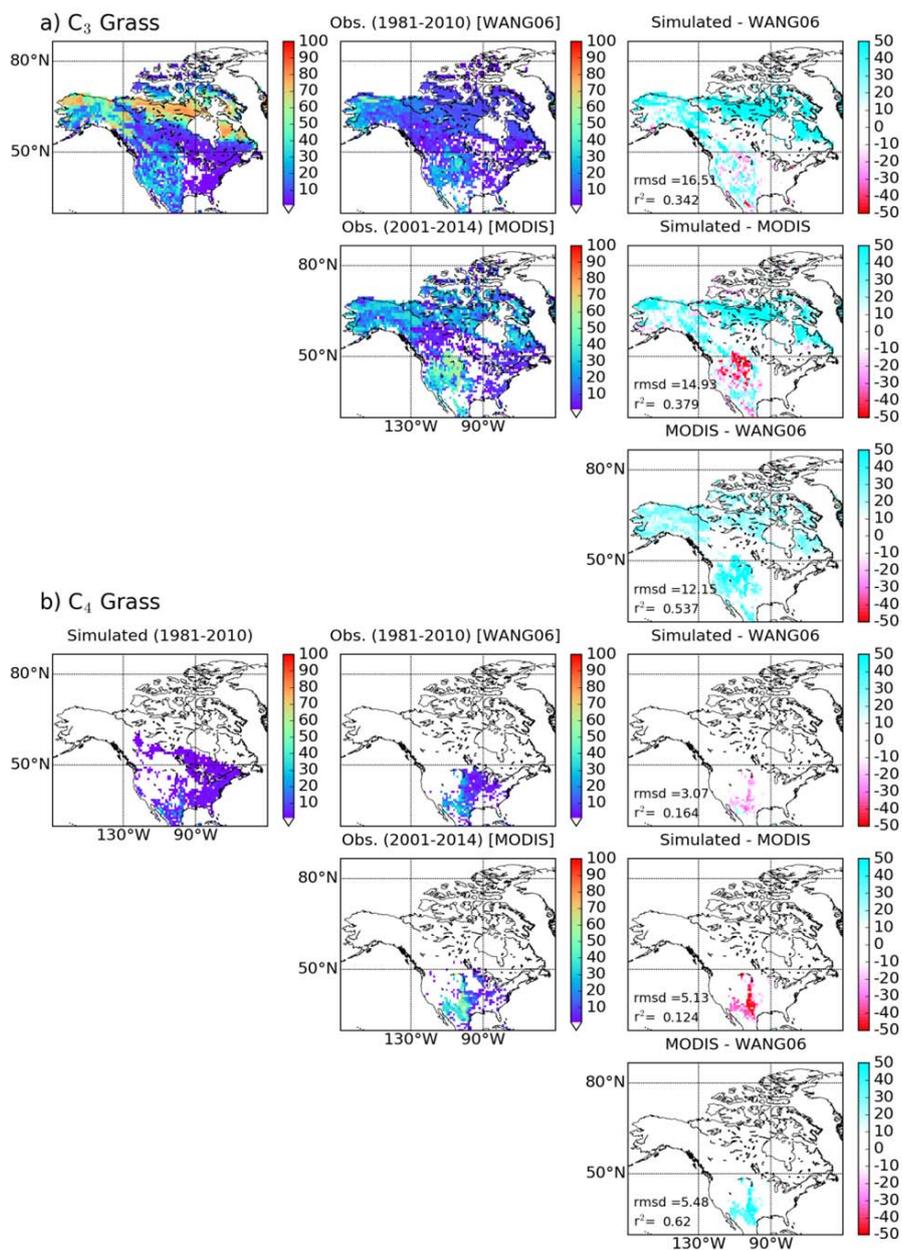
Figure 8. Spatial distribution of a) C_3 grasses, and b) C_4 grasses across North America. Simulated, observation-based, and differences are presented in the left, middle and right columns, respectively. The differences column includes model biases with respect to WANG06 (top panel) and MODIS (middle panel), and the difference between the the observation-based estimates (bottom panel). Root mean square difference (rmsd) and coefficient of determination (r^2) are also shown in each case.

66
67
68
69

70

71
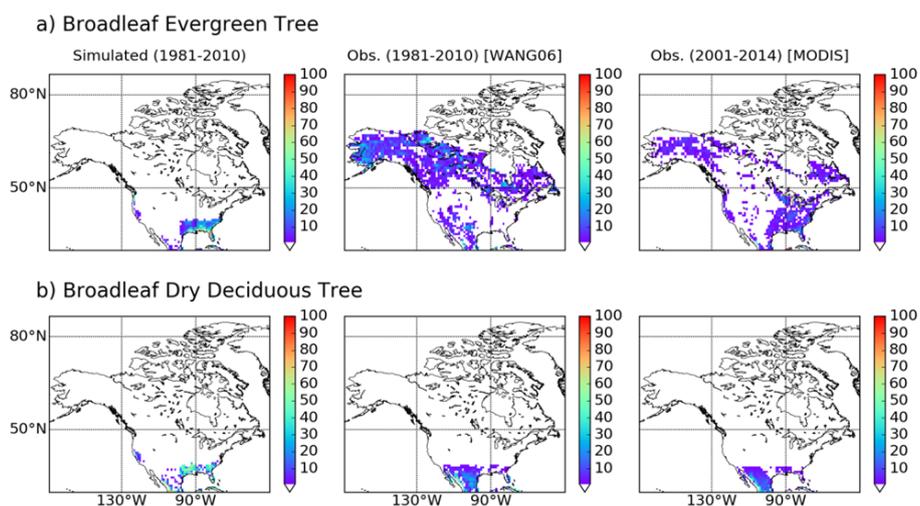
Figure 9. Spatial distribution of a) broadleaf evergreen tree, and b) broadleaf dry deciduous tree across North America. Simulated, WANG06 and MODIS distribution are presented in the left, middle and right columns, respectively.

75

76

77

78

79

80

81

82

83
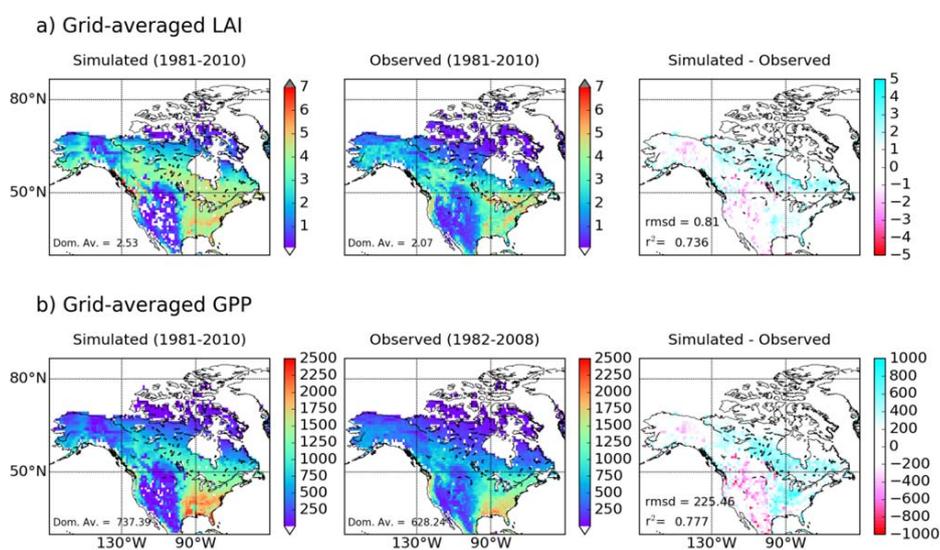
Figure 10. Spatial distribution of a) grid averaged maximum LAI (m^2 m^{-2}), and b) grid averaged GPP (g C m^2 y^{-1}) across North America. Simulated, observation-based, and differences between them are presented in the left, middle and right columns, respectively. Root mean square difference (rmsd) and coefficient of determination (r^2) are also shown in each case.

88

89
90

91
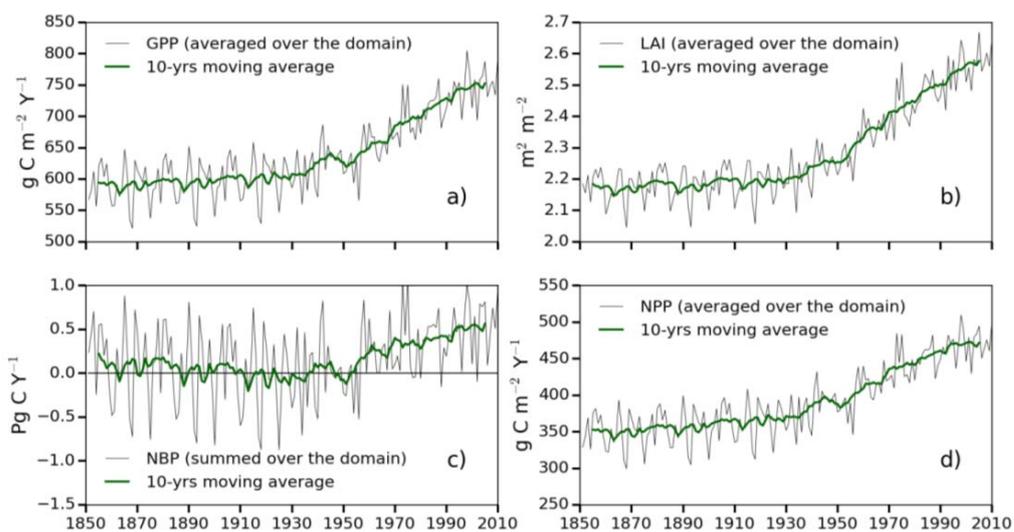
Figure 11. Time series evolution of a) domain averaged GPP (g C m^{-2} y^{-1}), b) domain averaged

LAI (m^2 m^{-2}), c) domain total NBP (Pg C m^{-2} y^{-1}), and d) domain averaged NPP (g C m^{-2} y^{-1}).

94

Biogeosciences

Discussions

Open Access

EGU

95

96
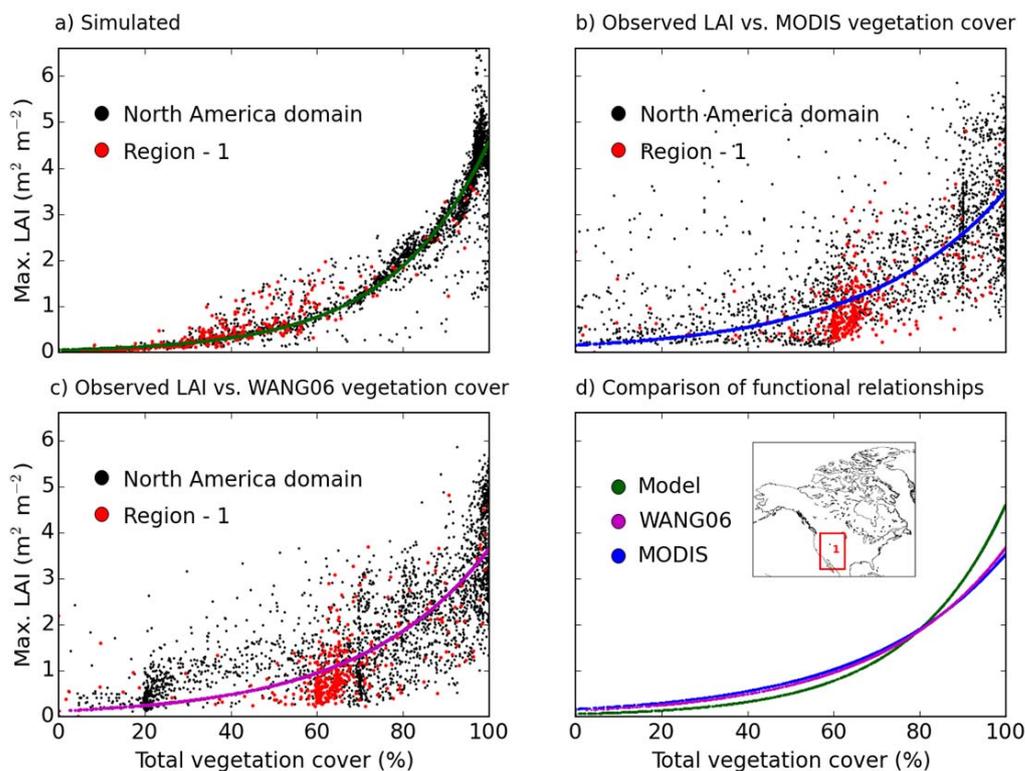
Figure 12. Scatter plots of a) simulated LAI vs. simulated total vegetation coverage, b) observed LAI vs. MODIS-derived total vegetation coverage, c) observed LAI vs. WANG06 total vegetation coverage. Plot d) shows a comparison of the fitted curves represented by solid lines, with an inset map of North America showing the sub-domain of interest bounded by a red rectangle.