Pairing FLUXNET Sites to Validate Model Representations of Land

Use/Land Cover Change

Liang Chen[1], Paul A. Dirmeyer[1], Zhichang Guo[1], Natalie M. Schultz[2]

1 Center for Ocean-Land-Atmosphere Studies, George Mason University, Fairfax, Virginia, USA

2 School of Forestry and Environmental Studies, Yale University, New Haven, Connecticut, USA

* Corresponding author: Liang Chen

Center for Ocean-Land-Atmosphere Studies, George Mason University, Mail Stop 6C5, Fairfax, VA, 22030 USA

Email: lchen15@gmu.edu

23   **Abstract**

24   Land surface energy and water fluxes play an important role in land-atmosphere interactions,

25   especially for the climatic feedback effects driven by land use/land cover change (LULCC).

26   These have long been documented in model-based studies, but the performance of land surface

27   models in representing LULCC-induced responses has not been well investigated. In this study,

28   measurements from proximate paired (open versus forest) flux tower sites are used to represent

29   observed deforestation-induced changes in surface fluxes, which are compared with simulations

30   from the Community Land Model (CLM) and the Noah Multi-Parameterization (Noah-MP) land

31   model. Point-scale simulations suggest CLM can represent the observed diurnal and seasonal

32   changes in net radiation (R_{net}) and ground heat flux (G), but difficulties remain in the energy

33   partitioning between latent (LE) and sensible (H) heat flux. CLM does not capture the observed

34   decreased daytime LE, and overestimates the increased H during summer. These deficiencies are

35   mainly associated with models' greater biases over forest land-cover types and the

36   parameterization of soil evaporation. Global gridded simulations with CLM show uncertainties

37   in the estimation of LE and H at the grid level for regional and global simulations. Noah-MP

38   exhibits a similar ability to simulate the surface flux changes, but with larger biases in H, G, and

39   R_{net} change during late winter and early spring, which are related to a deficiency in estimating

40   albedo. Differences in meteorological conditions between paired sites is not a factor in these

41   results. Attention needs to be devoted to improving the representation of surface heat flux

42   processes in land models to increase confidence in LULCC simulations.

1. Introduction

Earth system models (ESMs) have long been used to investigate the climatic impacts of land use/land cover change (LULCC) (cf. Pielke et al. 2011; Mahmood et al. 2014). Results from sensitivity studies largely depend on the land surface model (LSM) that is coupled to the atmospheric model within ESMs. In the context of the Land-Use and Climate, Identification of Robust Impacts (LUCID) project, Pitman et al. (2009) found disagreement among the LSMs in simulating the LULCC-induced changes in summer latent heat flux over the Northern Hemisphere. de Noblet-Ducoudré et al. (2012) and Boiser et al. (2012) argued that the inter-model spread of LULCC sensitivity (especially regarding the partitioning of available energy between latent and sensible heat fluxes within the different land-cover types) highlights an urgent need for a rigorous evaluation of LSMs. From Phase 5 of the Coupled Model Inter-comparison Project (CMIP5), Brovkin et al. (2013) also found different climatic responses to LULCC among the participating models, and the diverse responses are associated with different parameterizations of land surface processes among ESMs. To deal with the uncertainties in LULCC sensitivity among models, the Land Use Model Inter-comparison Project (LUMIP) has been planned, with a goal to develop metrics and diagnostic protocols that quantify LSM performance and related sensitivities with respect to LULCC (Lawrence et al. 2016).

However, a paucity of useful observations has hindered the assessment of the simulated impacts of LULCC and limited the understanding of the discrepancies among models. In-situ and satellite observations make it possible to quantify the impacts of LULCC on land surface variables. Satellite-derived datasets have been used to explore the albedo, evapotranspiration (ET), and

68   land surface temperature changes due to historical LULCC (Boisier et al. 2013, 2014) and the

69   climatic effects of forest (Li et al. 2015).

70

71   Meanwhile, the development of FLUXNET (Baldocchi et al. 2001) enables the study of land

72   surface responses to different land-cover types based on paired field observations from

73   neighboring flux towers over forest and open land (Juang et al. 2007; Lee et al. 2011; Luyssaert

74   et al. 2014; Teuling et al. 2010; Williams et al. 2012). In terms of LSM evaluation, the paired site

75   observations have been mainly used to simulated impacts of LULCC on land surface temperature

76   (Chen and Dirmeyer 2016; Lejeune et al. 2016; Vanden Broucke et al. 2015). However, a more

77   fundamental question, "whether a model can well represent the observed LULCC-induced

78   changes in surface energy fluxes", has not been thoroughly investigated, even though we know

79   that the turbulent fluxes are tightly associated with both energy and water exchange between the

80   land surface and atmosphere.

81

82   In this study, we evaluate the performance of the Community Land Model (CLM) version 4.5

83   and the Noah Multi-Parameterization (Noah-MP) LSM in simulating the impacts of LULCC on

84   surface energy fluxes based on observations from FLUXNET sites. CLM and Noah-MP

85   represent perhaps the two most readily available and widely used state-of-the-art community

86   land models developed in the U.S. CLM is chosen because, as the land component for

87   Community Earth System Model (CESM), it prioritizes the simulation of biogeophysical and

88   biogeochemical processes for climate applications (Oleson et al. 2013). Much effort has gone

89   into improving the representation of the land-atmosphere interactions among different biomes

90   (Bonan et al. 2011), and the model itself has been used for many LULCC sensitivity studies

91   (e.g., Chen and Dirmeyer 2016, 2017; Schultz et al. 2016; Lejeune et al. 2017; Lawrence et al.

92   2012). Noah-MP has found use mainly in shorter time-scale, limited area applications, such as

93   weather and hydrologic forecasting, and as a LSM run at very high resolution coupled to

94   mesoscale models (e.g., WRF-Hydro, Gochis et al. 2015). It is planned to become the LSM used

95   in global weather and seasonal forecasting applications at the National Centers for

96   Environmental Prediction (NCEP). Its performance over varying land cover types has direct

97   consequences for its use in forecast models.

98

99   The rest of this paper is structured as follows. Section 2 describes the datasets used in the study

100   and experimental design. Section 3 presents comparison between observations and model

101   simulations in surface latent and sensible heat flux, ground heat flux, and net radiation. Section 4

102   shows the uncertainties within the FLUXNET pairs and model simulations. Sections 5 and 6

103   include discussion and conclusions, respectively.

104

105   **2. Methodology**

106

107   *2.1 Observational data*

108

109   We use half-hourly observations from 24 selected pairs of flux sites from the FLUXNET2015

110   Tier 1 dataset (http://fluxnet.fluxdata.org/data/fluxnet2015-dataset) and 4 pairs from the

111   AmeriFlux dataset (Baldocchi et al. 2001). These observations include meteorological forcings

112   for the LSM, and surface flux measurements for model validation, which include latent heat flux

113   (LE), sensible heat flux (H), ground heat flux (G), and net radiation (R_{net}). All of these variables

114   have been gap-filled (Reichstein et al. 2005; Vuichard and Papale 2015). Table 1 shows the

115   variable names and gap-filling algorithms used in FLUXNET2015. Because there is no directly

116   measured humidity variable reported, which is needed as a meteorological forcing for the LSMs,

117   relative humidity is calculated based on the reported vapor pressure deficit and surface air

118   temperature (Equation 1-2).

$$e_s = 6.11 \, exp \left(17.26938818 \frac{T_a}{237.3 + T_a} \right) \tag{1}$$

$$RH = \left(1 - \frac{VPD}{e_s} \right) \times 100 \tag{2}$$

119   in which T_a is air temperature (°C), e_s is saturation vapor pressure (hPa), VPD is vapor pressure

120   deficit (hPa), and RH is relative humidity (%). Additionally, for the turbulent flux measurements

121   over 18 pairs, FLUXNET2015 provides "corrected" fluxes based on an energy balance closure

122   correction factor, which is calculated for each half-hour as $(R_{net} - G) / (H + LE)$. More details

123   about the data processing can be found on the FLUXNET2015 website

124   (http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/data-processing/).

125

126   To simulate local land cover change for each pair, one flux tower is located in forest (deciduous,

127   evergreen or mixed; broadleaf or needleleaf) and the other is in a nearby open land cover type

128   (grassland, cropland or open shrub). Figure 1 shows the locations of the paired sites. Their

129   general characteristics are listed in Table S1. The median linear distance between the paired sites

130   is 21.6 km, and the median elevation difference is 20.0 m. Because of their proximities, the

131   paired sites share similar atmospheric background conditions, however they are not identical

132   (Chen and Dirmeyer 2016). Below we show that the differences in meteorology are usually small

133   and not likely a dominant factor in simulated surface flux differences in most of the pairs. We

134   consider the differences (open minus forest) in observed surface fluxes to be representative of the

135   effects of LULCC (deforestation in this case).

136

137   *2.2 Model simulations*

138

139   We have run the offline version of CLM 4.5 and Noah-MP at the point-scale for individual sites.

140   The forcing data, described below, includes downwelling long-wave radiation (W/m^2),

141   downwelling short-wave radiation (W/m^2), air temperature (K), precipitation (mm/s), relative

142   humidity (%), surface pressure (Pa), and wind speed (m/s) at half-hourly time steps. The plant

143   functional type (PFT) in CLM for each site is identified based on its reported land cover type

144   (Table S1) with prescribed climatological satellite phenology (Lawrence and Chase, 2010).

145   Because of the focus on biogeophysical impacts of LULCC in this study, the biogeochemistry

146   Carbon-Nitrogen module has been disabled in our simulations. The initial conditions for each

147   site are generated by cycling through available atmospheric forcings for about 40 years until soil

148   moisture and temperature reach quasi-equilibrium.

149

150   The differences in simulated surface fluxes between the paired sites are compared against the

151   observations, so that the performance of CLM in representing LULCC-induced surface flux

152   changes can be evaluated. In the single-point simulations, two types of forcing data are used for

153   each site: 1) measurements at this site; 2) measurements at the neighboring paired site.

154   Consequently, three types of differences in simulated surface fluxes can be calculated: 1) the

155   difference derived from individual forcings; 2) the difference from identical "forest forcings"

156   (both of the paired sites use the same forcings measured at the forest site); 3) the difference from

157 identical "open forcings" (both of the paired sites use the same forcings measured at the open

158 sites). Such an experimental design can well eliminate the influence from the uncertainties of

159 forcing data and the difference in atmospheric background of the paired sites.

160

161 The ultimate goal of evaluating CLM's performance at single-point scale is to assess its

162 capability to be used in global LULCC sensitivity simulations in both offline and coupled modes.

163 The paired sites are close enough that they are typically located within a single grid cell of

164 CESM. Moreover, the sub-grid heterogeneity of CLM allows the biogeophysical processes to be

165 calculated at the individual PFT level (15 PFTs available), and makes it possible to output

166 surface fluxes for individual land cover types. The paired sites can be presented as paired PFTs

167 within a single grid of CESM. They then share the same atmospheric forcings, and their

168 differences can be considered as the impacts of LULCC. It should be noted that the PFT-level

169 calculation is independent of the percentage of individual PFTs in the grid cell. Therefore, the

170 coverage of the PFTs in the shared grid cell does not influence the flux difference between the

171 paired PFTs in the global simulations.

172

173 We run CLM offline, globally driven by the CRUNCEP forcings from 1991 to 2010 (Viovy

174 2011) and present land cover conditions (Lawrence et al. 2012) at a horizontal resolution of

175 0.9°×1.25°. The paired PFTs are identified based on the locations and land cover types of the

176 FLUXNET paired sites, to ensure the single-point and global simulations are comparable.

177 Schultz et al. (2016) found the shared-soil-column configuration for vegetated land units in CLM

178 caused issues with PFT-level ground heat fluxes. They propose an individual-soil-column

179 scheme (PFTCOL) to better represent the PFT-level energy fluxes, so we also extract and

180 examine the output for the paired PFTs from the PFTCOL model configuration. Details about the

181 PFTCOL simulations can be found in Schultz et al. (2016). Additionally, a coupled simulation

182 with Community Atmosphere Model (CAM) has also been conducted. It shows very similar

183 results to the offline simulations, because the paired PFTs in a single model grid box always

184 share the same atmospheric forcings no matter if CLM is run offline or coupled with CAM.

185 Therefore, results from the coupled simulation are not included in this study.

186

187 Furthermore, we compare the performance of CLM with Noah-MP (Niu et al. 2011), which

188 serves as a participant model in Land Data Assimilation Systems (LDAS, Cai et al. 2014).

189 Single-point Noah-MP simulations are conducted in the same way as CLM simulations to ensure

190 their comparability. The monthly leaf area index (LAI) of each site is identical to the prescribed

191 satellite-based LAI in the corresponding CLM simulation. Table S2 shows selected options for

192 various physical processes in Noah-MP. Information about all model simulations is summarized

193 in Table 2.

194

195 **3. Surface energy fluxes and their changes**

196

197 First, we analyze the diurnal and seasonal cycles of surface energy fluxes and the LULCC-

198 induced changes. The diurnal cycle analysis is primarily focused on summer (DJF for the two

199 austral sites and JJA for the other sites). The seasonal cycle for the austral sites is shifted by 6

200 months to keep summer in the middle of the time series when comparing or compositing with the

201 Northern Hemisphere sites. The results shown below are composites averaged over all open (or

202 forest) sites or open-forest pairs. Not all sites have energy-balance corrected fluxes available;

203 exclusion of those sites shows very similar results for uncorrected fluxes to the average over all

204 sites (or pairs, not shown). There are also some pairs with relatively large changes in surface

205 fluxes. Exclusion of those pairs shows very consistent patterns with the results including all sites,

206 even though there is a slight influence on the magnitude of the changes (Figure S1). Therefore,

207 all sites are included in our analyses for each variable.

208

209 *3.1 Latent heat flux (LE)*

210

211 Figure 2a-b shows the diurnal cycle of *LE* averaged over all the open sites and forest sites during

212 summer. Compared with the observations without energy-balance correction, single-point CLM

213 simulations overestimate *LE* for the open sites with both their actual meteorological forcings and

214 the nearby forest forcings, but underestimate *LE* over the forest sites. The extracted PFT-level

215 output from the global simulations also exhibit similar biases. Relative to CLM, Noah-MP

216 simulations show better agreement with observations over the open sites, but a greater

217 underestimation over forest. The energy-balance correction tends to increase the values of *LE*.

218 Therefore, both CLM and Noah-MP have negative biases compared to the corrected fluxes

219 (except LE_CORR_25 over the open sites).

220

221 Figure 2c shows the difference in the diurnal cycle of *LE* due to LULCC (deforestation). It

222 should be noted that there is a substantial spread among the pairs in model simulations and

223 especially observations, indicating the diverse geographical backgrounds and specific vegetation

224 changes of these paired sites. The observations suggest an overall lower summer daytime *LE*

225 over the open land compared to forest. In spite of the considerable spread among the energy-

226   balance corrected *LE* observations (Figure 2ab), the differences between the forest and open

227   lands show consistent signals. However, both CLM and Noah-MP single-point simulations fail to

228   represent the observed decreased daytime *LE* as a result of deforestation. The simulated *LE* over

229   the open land is usually slightly greater than the forest from 10:00 to 16:00 at local time. Such a

230   discrepancy may be attributed to the large underestimation of daytime forest *LE* in the models.

231   Meanwhile, simulations by different forcings of the paired sites show robust signals, implying

232   that the bias of the simulated *LE* sensitivity should not be attributed to the uncertainties of the

233   forcing data. For the CLM global simulations, the PFTCOL case exhibits a similar diurnal

234   pattern to the single-point simulations, while decreased daytime *LE* is found consistently only in

235   the PFT simulations. As CLM-PFT is less physically realistic than CLM-PFTCOL from a soil

236   hydrologic perspective, its superior performance needs further investigation.

237

238   To explore the mechanism of the *LE* changes within CLM, we examine the changes in the three

239   components of evapotranspiration; namely canopy evaporation, canopy transpiration, and ground

240   evaporation (Figure 3). Unfortunately, these separate components are not measured and cannot

241   be directly validated. The CLM, PFT and PFTCOL simulations show an agreement in decreased

242   canopy evaporation after deforestation with the greatest decrease during the early morning.

243   There also is an agreement in an overall decreased canopy transpiration, but CLM simulations do

244   not exhibit an obvious change during the morning when greatly decreased canopy transpiration

245   can be found in the PFT and PFTCOL simulations. The main discrepancy among model versions

246   is found in ground evaporation, which increases after deforestation in the CLM and PFTCOL

247   simulations. The increased ground evaporation has exceeded the decreased canopy evaporation

248   and transpiration, resulting in slightly increased *LE* (Figure 2c). Interestingly, the PFT

249   simulations, which have known issues with PFT-level ground heat flux (Schultz et al. 2016),

250   show decreased daytime ground evaporation. Along with decreased canopy evaporation,

251   transpiration, and ground evaporation, the total LE decreases sharply after deforestation in the

252   PFT simulations, which agrees better with the observations than other simulations (Figure 2c).

253   However, the decreased ground evaporation may be associated with a problematic soil-column

254   scheme at sub-grid scale, which undermines the credibility of the agreement between the

255   observations and PFT simulations.

256

257   Figure 4 shows the changes in monthly LE after deforestation across the annual cycle. There is

258   clear and consistent seasonality in the LE changes from the observations. The four types of

259   observations show decreased LE (up to -24.0 W/m^2) during local summer. There is little change

260   in LE in the uncorrected observations during the winter season. However, there is significantly

261   increased LE (up to +17.9 W/m^2) in the energy-balance corrected observations in late winter and

262   early spring. Neither CLM nor Noah-MP capture the observed seasonality of LE change. As

263   found in the change in the diurnal cycle of the LE, the PFTCOL simulations exhibit a similar

264   pattern to the single-point simulations, while the PFT simulations show decreased LE throughout

265   the year with the maximum from May to August, and the best correlation (R = 0.81, P < 0.01)

266   with observations.

267

268   *3.2 Sensible heat flux (H)*

269

270   Figure 5a-b shows the diurnal cycle of H averaged over all open and forest sites during local

271   summer. Generally, the models overestimate H throughout the day, with the largest positive bias

272   during midday. Compared with the observations without energy-balance correction, the

273   overestimation can be up to 86.5 W/m^2 from CLM over the forest during noon and 46.4 W/m^2

274   over the open sites. The difference in H between the forest and open sites is shown in Figure 5c.

275   Robust signals are found among the four types of observations, so results from the energy-

276   balance corrected observations are not included hereafter, but are shown in Figure S2. Both

277   observations and models exhibit a clear diurnal pattern of change in H after deforestation – a

278   small nighttime increase and a large daytime decrease. Observations show a large spread among

279   the 28 pairs, which is much greater than that from the CLM simulations, indicating uncertainties

280   and variability among the observed fluxes and the robustness of simulated H sensitivity to

281   LULCC in the LSM. Compared with the observations, CLM shows a greater H decrease, which

282   is twice as much as in the observations. The overestimated H decrease may be related to the large

283   positive bias in H over the forest sites (Figure 5b). Additionally, the PFT simulations show the

284   largest H decrease, which may be associated with the ground heat issues in the shared-soil-

285   column scheme.

286

287   Seasonally, decreased H is found throughout the year after deforestation in both observations and

288   models (except for the same-forest-forcing CLM simulations in winter, Figure 6). The greatest

289   decrease is observed during spring, when both of the single-point CLM and PFTCOL

290   simulations show good agreement. However, CLM and Noah-MP simulations also show a large

291   decrease during summer, which has not been observed in the FLUXNET dataset. Again, the PFT

292   simulations show the greatest H decrease among the simulations and the largest bias compared

293   with the observations during the warm season.

294

295 Additionally, evaporative fraction (*EF*), which is defined as the ratio of *LE* to the available

296 energy (*LE+H*), is a useful diagnostic of the surface energy balance (Gentine et al. 2011).

297 Meanwhile, most of the correction methods to solve the imbalance issue of surface energy

298 budget assume the Bowen ratio for small- and large-scale eddies are similar or even equal

299 (Wilson et al. 2002; Foken 2008; Zhou and Wang 2016). Under such an assumption, *EF* can be

300 independent of energy closure issue, because EF is related to the Bowen ratio (B) as:

$$EF = (1 + B)^{-1} \qquad\qquad (3)$$

301 Figure 7 shows the change in the diurnal (summer only) and seasonal cycle of *EF* due to LULCC

302 from forest to open land. During summer, there are small changes in observed daytime *EF*

303 (Figure 7a) because of the decreases in both *LE* and *H*. However, both CLM and Noah-MP show

304 increased daytime *EF* due to the decreased *H* and slightly increased *LE* after deforestation.

305 Seasonally, the models show year-around increased *EF*, however, which is not observed in

306 FLUXNET from June to September, further demonstrating the models' deficiencies in

307 representing energy partitioning during summer.

308

309 *3.3 Diurnal and seasonal cycle of ground heat flux (G) and net radiation (R_{net})*

310

311 Figure 8a shows the change in the diurnal cycle of *G* after deforestation. Both the observations

312 and models exhibit increased *G* during the day and decreased *G* during the night. However,

313 models overestimate the magnitude of the *G* change, and discrepancies also exist in the timing of

314 maximum change. The greatest increase in *G* is observed during early afternoon, while the

315 greatest increase in simulated *G* occurs at noon in CLM (single-point and PFTCOL) and during

316 morning in Noah-MP. Because *G* is strongly correlated with R_{net} (Santanello and Friedl 2003),

317   we examine the timing of maximum observed G and R_{net} during summer. There are some sites

318   showing about a 1-hour lag between maximum R_{net} and G (not shown). Therefore, the lag

319   between simulated and observed peaks in G change can be partially attributed to the uncertainties

320   in G measurements that are commonly estimated with heat flux plates installed at some depth

321   (e.g., 5~10 cm) below the surface (Wang and Bou-Zeid 2012), while the LSM simulated G is

322   calculated at the surface. Meanwhile, the G changes (in both the diurnal and seasonal cycle) in

323   the PFT simulations are further from the observations than the other simulations. Such

324   disagreement further confirms the issues with the sub-grid soil column scheme in CLM, which is

325   discussed in the following section. The changes in observed G also have a clear seasonal pattern

326   – an increase during the warm season and a decrease during the cold season (Figure 8b). This

327   seasonality is well captured by the CLM simulations (especially the simulations with identical

328   forcings for the paired sites) in both magnitude and timing, but not evident in Noah-MP

329   simulations.

330

331   After exploring the three flux components of the surface energy balance, it is worthwhile to

332   examine the change in R_{net} after deforestation. During summer, the observations show that R_{net}

333   slightly increases during the night, and decreases considerably (up to -65.7 W/m^2) during the

334   day, which can be attributed to the increased albedo after deforestation (Figure 9a). Decreased

335   daytime R_{net} is also found in the CLM simulations, but with a slightly smaller magnitude.

336   Seasonally, there is a good agreement between the observations and CLM simulations, showing

337   a large R_{net} decrease during spring and summer but a relatively small decrease during autumn and

338   winter (Figure 9b). The Noah-MP simulations are comparable to CLM, but with a notable

339   deficiency in simulating the R_{net} change during late winter and early spring.

340

4. Uncertainty Analysis

342

4.1 Uncertainties among the FLUXNET pairs

344

The results discussed above are based on composites averaged over all forest and open sites. It is

worthwhile to examine the uncertainties in surface flux changes among different paired sites.

Figure 10a shows the changes in summer daytime (8:00 ~ 16:00) *LE* from the observations and

model simulations across the 28 pairs. This time period is chosen because it is the time of

greatest differences in surface energy fluxes (Figure 2c, 5c, 7a, 8a). The observations show

decreased *LE* associated with deforestation over 23 pairs, among which the pairs of evergreen

needleleaf forest and open shrub (No. 16~25) exhibit consistent decreases and the pairs of

deciduous broadleaf forest and crops (No. 1~4) show the overall greatest decrease. However,

both CLM and Noah-MP show relatively weak increases over most of the pairs, which further

demonstrate their deficiency in simulating *LE* change. Additionally, for both CLM or Noah, the

choice of forcings does not exert much influence on the simulated change in summer daytime

LE.

357

The changes in R_{net} over individual pairs are shown in Figure 10b. There are 27 pairs (all except

number 21) showing decreased R_{net} after deforestation, with the greatest decreases over the pairs

of evergreen needleleaf forest and grassland. Both CLM and Noah-MP well captures the

observed decreases in R_{net} over most of the pairs.

362

363   It should be noted that pair 15 shows large *LE* and R_{net} changes in Figure 10. This pair consists of

364   a site over valley grassland and the other site over mountain evergreen needleleaf forest with

365   60.29 km separation and 1186 m elevation difference. There are significantly different air

366   temperature and downwelling longwave radiation measurements between the sites (Figure S3).

367   Such large differences in *LE* and R_{net} here are likely associated with the distinct although

368   proximate geographical sites. Even though the exclusion of this site does not make a significant

369   change to the composite analysis in section 3 (not shown), it may raise another question if the

370   simulated sensitivity of surface energy fluxes is associated with the inconsistencies of

371   atmospheric forcings of LSMs at the single pair level.

372

373   *4.2 Uncertainties within the forcings for LSMs*

374

375   Based on the composite analysis in section 3, we have found that the simulated changes in

376   surface energy fluxes with identical forcings (either from forest or open sites) are consistent with

377   the simulations with individual forcings, demonstrating that the overall sensitivities of surface

378   energy fluxes are robust among the choices of different forcings. In this sub-section, we explore

379   the uncertainties of the simulated surface flux changes due to the different forcings for individual

380   pairs, especially with the focus on the roles of separation and elevation difference in the

381   simulated sensitivity of surface energy fluxes.

382

383   Since we have simulations with identical forcings for the paired sites, the difference in surface

384   flux changes between "forest forcings" and "open forcings" can be considered as the simulated

385   sensitivity of surface energy fluxes to variation in the atmospheric forcings. Figure 11 shows the

386   relationship with separation and elevation difference for individual pairs. Overall, the flux

387   changes are not associated with the separation and elevation difference between the paired sites,

388   further confirming the robustness of simulated signals from paired-site simulations.

389   Nevertheless, some "outliers" are identified. In the CLM simulations, only pair 15 shows large

390   differences in LE and H change. However, pairs 3, 7, and 12 also exhibit large differences in

391   Noah-MP simulations. The uncertainties in pairs 12 and 15 may be attributed to their large

392   elevation differences. For pair 7 in Australia, Noah-MP shows greater sensitivity of H and R_{net} to

393   atmospheric forcings over the evergreen broadleaf forest than grassland (not shown), leading to

394   large differences in the surface flux changes. However, this is the only pair with evergreen

395   broadleaf forest, and its behavior in Noah-MP needs further investigation. Even though the pair 3

396   sites are close with small elevation difference, we found considerably different downwelling

397   shortwave and longwave radiation between the two sites (not shown), which may explain the

398   uncertainties in the Noah-MP simulations.

399

400   **5. Discussion**

401

402   This study has examined simulated changes in the surface energy budget in response to local

403   land cover change based on paired proximate FLUXNET sites with differing land cover. Our

404   results suggest that CLM well represents the observed changes in R_{net} and G; but there remain

405   issues in simulating the energy partitioning between LE and H, which also further confirms the

406   large uncertainties in simulated ET responses to LULCC revealed in several recent studies (e.g.,

407   Pitman et al. 2009; Boisier et al. 2012, 2014; de Noblet-Ducoudré et al. 2012; Vanden Broucke

408   et al. 2015). Based on the observations, deforestation generally leads to a decrease in summer

409   daytime R_{net}, accompanied by decreased LE and H. On one hand, CLM captures the observed

410   signal of H change, but overestimates the decrease due to its large overestimation of H over the

411   forest. On the other hand, the model underestimates the LE over the forest, leading to an opposite

412   signal (a slight increase) of LE change comparing to the observations. Simulations in Noah-MP

413   show similar biases. Therefore, uncertainties in current LULCC sensitivity studies may persist

414   specifically in the representation of turbulent fluxes over forest land-cover types.

415

416   Scrutinizing the three components of ET suggests that the simulated increase in summer daytime

417   LE is mainly attributable to a large increase in ground evaporation, which counteracts the

418   decreased canopy evaporation and transpiration. This may raise another issue about the soil

419   resistance parameterization in CLM4.5. Previous studies indicate that the model generates

420   excessive ground evaporation when the canopy is sparse or absent (Swenson and Lawrence

421   2014; Tang et al. 2015). If there is overestimated ground evaporation over the open land, such a

422   bias can also contribute to the disagreement in the LULCC-induced ET changes. Swenson and

423   Lawrence (2014) have implemented a dry surface layer for the soil resistance parameterization to

424   solve this issue for the upcoming CLM5. An extension of the evaluation with CLM5 would be

425   useful to examine if the issue within the soil resistance parameterization is responsible for the

426   uncertainties in ET changes.

427

428   Besides the uncertainties in estimating turbulent fluxes over different land cover types, the

429   simulations show that differences in the meteorological forcings between nearby paired sites

430   seem to have little impact on the simulation of surface flux changes due to LULCC. Many

431   LSMs besides CLM employ a sub-grid tiling parameterization where multiple land surface types

432   exist within a single grid box, each maintaining a separate set of surface balances and returning a

433   weighted average set of fluxes to the atmosphere based on areal coverage of each surface type.

434   In this arrangement, each land surface type within a grid box receives the same meteorological

435   forcing from the overlying atmospheric model. It appears from our forcing-sensitivity studies

436   that this arrangement does not significantly impact the simulation of surface flux changes

437   associated with LULCC on the grid scale.

438

439   That said, the sub-grid comparison between different land cover types may yet be problematic

440   due to the shared soil column issue for vegetated land units in CLM (Schultz et al. 2016). Both

441   the single-point observations and simulations show significant differences in surface soil

442   moisture between most of the paired sites, even though no clear drying or wetting pattern is

443   found (Figure S4). The differences between the paired sites suggests that the shared soil column

444   for vegetated land in CLM may not well represent soil moisture and temperature at the sub-grid

445   scale, which may influence the simulations of land surface energy and water fluxes. We find an

446   unreasonably large change in PFT-level G between forest and open land especially for the

447   seasonal cycle in PFT simulations, while both observations, single-point and PFTCOL

448   simulations show a seasonal change with a very small range (within $\pm$ 3W/m^2). As G is the

449   calculated as the residual of the surface energy budget in CLM (Oleson et al. 2013), this sub-grid

450   G issue may cast even more uncertainties on the calculation of LE and H at the PFT level, as well

451   as their aggregated values at the grid level for regional or global simulations. Therefore, caution

452   should be taken when examining the LULCC sensitivity which involves sub-grid PFT changes.

453

454   Compared with CLM, Noah-MP exhibits a similar ability to simulate surface flux changes,

455   except for a deficiency in simulating H and R_{net} changes during late winter and early spring. We

456   have examined the daytime albedo change after deforestation, calculated from available

457   shortwave radiation terms, from observations and model simulations during local late

458   winter/early spring (February ~ April, FMA) and summer (Figure 12). Both CLM and Noah-MP

459   agree with the observations during summer. However, Noah-MP does not capture the observed

460   albedo increase over nearly half of the pairs during late winter/early spring. Greater disagreement

461   is also found during the local winter season (DJF, not shown), suggesting a deficiency in

462   snowmelt timing or snow albedo sensitivity to LULCC, despite improvement in the snow surface

463   albedo simulations by implementation of the Canadian Land Surface Scheme (CLASS;

464   Verseghy, 1991) in Noah-MP (Niu et al. 2011).

465

466   Finally, it should be recognized that the observational data are not perfect. In particular, there

467   may be systematic biases or even trends in specific instruments that contribute to the perceived

468   differences between paired sites (e.g., site 3). Ideally, redundant instrumentation at sites, or in

469   this case the rotation of an extra set of instruments among nearby paired sites, could be used to

470   identify, quantify and account for significant systematic biases in measurements for suspicious

471   variables. Furthermore, footprints of the flux towers may bias the comparison of surface fluxes

472   between the open and forest sites (Baker et al. 2003; Griebel et al. 2016). In other words, the

473   observed differences between sites can only be partially attributed to LULCC because their

474   environmental conditions may also be different. As most of current studies using paired sites to

475   represent LULCC, we have assumed that the paired sites share the similar background

476   atmospheric conditions, and any observed differences in surface climate conditions can be

477   attributed to LULCC (e.g., Lejeune et al. 2017; Luyssaert et al. 2014; Teuling et al. 2010;

478   Vanden Broucke et al 2015). Meanwhile, model simulations with the different forcings can

479   effectively examine the effects of the local environment of individual sites, because their

480   footprints can also be taken by the meteorological measurements. Our results show robust signals

481   of LULCC-induced changes in surface fluxes, implying that impacts of footprints at individual

482   sites are probably trivial.

483

484   **6. Conclusions**

485

486   This study has evaluated the performance of two state-of-the-art LSMs in simulating the

487   LULCC-induced changes in surface energy fluxes. Observations from 28 FLUXNET pairs (open

488   versus forest) are used to represent the observed flux changes following deforestation, which are

489   compared with the LSM simulations forced with meteorological data from the observation sites.

490   Diurnal and seasonal cycles of the flux changes have been investigated.

491

492   The single-point simulations in CLM and Noah-MP show the greatest bias in simulating LE

493   change. Significantly decreased daytime LE is observed during local summer, but not captured

494   by the models. The observed LE changes also exhibit an evident seasonality, which is not

495   represented in the model. The energy partitioning between LE and H might be a common issue

496   within the LSMs. Other studies have noted problems in the simulation of surface fluxes by

497   LSMs, including poor performance relative to non-physical statistical models (Best et al. 2015,

498   Haughton et al. 2016).

499

500   The sub-grid comparison from the global simulations in CLM yields unrealistic changes in *G* and

501   *H* when the soil column is shared among vegetated land units, even though there is a better

502   agreement in *LE* change with the observations. The individual-soil-column scheme improves the

503   representation of the PFT-level energy flux changes, but uncertainties still remain as with the

504   point-scale simulations. Therefore, these uncertainties must be considered when interpreting

505   global experiments of LULCC sensitivity studies with current LSMs.

506

507   Consistent aggregate performance across many paired sites suggests the problems in these LSMs

508   may not lie primarily with parameter selection at individual sites, but with more fundamental

509   issues of the representation of physical processes in LSMs. The simulation of LULCC may or

510   may not have become more consistent among models since LUCID (de Noblet-Ducoudré et al.

511   2012), but consistency with observed biophysical responses appears to be lacking. LUMIP

512   (Lawrence et al. 2016) will be a step toward better LSM simulation of LULCC responses, and

513   ultimately better simulations of the response of climate to LULCC.

514

515

516

517   *Acknowledgements:*

518   This study was supported by the National Science Foundation (AGS-1419445). This work used

519   eddy covariance data acquired and shared by the FLUXNET community, including these

520   networks: AmeriFlux, CarboEuropeIP, CarboItaly, CarboMont, Fluxnet-Canada, GreenGrass,

521   ICOS, and OzFlux-TERN. The ERA-Interim reanalysis data are provided by ECMWF and

522   processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was

523   carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and

524   Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic

525   Center, and the OzFlux, ChinaFlux and AsiaFlux offices. We thank all site investigators and flux

526   networks for their work to make our model evaluation possible. The authors wish to thank

527   Ahmed Tawfik at the National Center for Atmospheric Research for his assistance of preparing

528   forcing datasets for the AmeriFlux sites. Computing resources for the CLM and Noah-MP

529   experiments were provided by the NSF/CISL/Yellowstone supercomputing facility.

530

531

532

533

534 **REFERENCES**

535

536 Baker, I., A. S. Denning, N. Hanan, L. Prihodko, M. Uliasz, P. Vidale, K. Davis, and P. Bakwin,

537   2003: Simulated and observed fluxes of sensible and latent heat and CO2 at the WLEF-

538   TV tower using SiB2.5. Global Change Biol., **9**, 1262-1277, doi:10.1046/j.1365-

539   2486.2003.00671.x.

540 Baldocchi, D. and Coauthors, 2001: FLUXNET: A New Tool to Study the Temporal and Spatial

541   Variability of Ecosystem–Scale Carbon Dioxide, Water Vapor, and Energy Flux

542   Densities. *Bull. Amer. Meteor. Soc.,* **82**, 2415-2434, doi:10.1175/1520-

543   0477(2001)082<2415:FANTTS>2.3.CO;2.

544 Betts, A. K., 2004: Understanding Hydrometeorology Using Global Models. *Bull. Amer. Meteor.*

545   *Soc.,* **85**, 1673-1688, doi:10.1175/BAMS-85-11-1673.

546 Boisier, J., N. de Noblet-Ducoudré, and P. Ciais, 2014: Historical land-use-induced

547   evapotranspiration changes estimated from present-day observations and reconstructed

548   land-cover maps. *Hydrology and Earth System Sciences,* **18**, 3571-3590,

549 Boisier, J., N. de Noblet-Ducoudré, and P. Ciais, 2013: Inferring past land use-induced changes

550   in surface albedo from satellite observations: a useful tool to evaluate model

551   simulations. *Biogeosciences,* **10**, 1501-1516,

552 Boisier, J. P. and Coauthors, 2012: Attributing the impacts of land-cover changes in temperate

553   regions on surface temperature and heat fluxes to specific causes: Results from the first

554  LUCID set of simulations. *Journal of Geophysical Research: Atmospheres,* **117**, n/a-n/a,

555  doi:10.1029/2011JD017106.

556 Bright, R. M., K. Zhao, R. B. Jackson, and F. Cherubini, 2015: Quantifying surface albedo and

557  other direct biogeophysical climate forcings of forestry activities. *Global Change*

558  *Biol.,* **21**, 3246-3266, doi:10.1111/gcb.12951.

559 Brovkin, V. and Coauthors, 2013: Effect of Anthropogenic Land-Use and Land-Cover Changes

560  on Climate and Land Carbon Storage in CMIP5 Projections for the Twenty-First

561  Century. *J. Climate,* **26**, 6859-6881, doi:10.1175/JCLI-D-12-00623.1.

562 Cai, X., Z. Yang, Y. Xia, M. Huang, H. Wei, L. R. Leung, and M. B. Ek, 2014: Assessment of

563  simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the

564  NLDAS test bed. *Journal of Geophysical Research: Atmospheres,* **119**, 13,751-13,770,

565  doi:10.1002/2014JD022113.

566 Chen, L., and P. A. Dirmeyer, 2016: Adapting observationally based metrics of biogeophysical

567  feedbacks from land cover/land use change to climate modeling. *Environmental Research*

568  *Letters,* **11**, 034002, doi:10.1088/1748-9326/11/3/034002.

569 Chen, L., and P. A. Dirmeyer, 2016: Impacts of Land Use/Land Cover Change on Afternoon

570  Precipitation over North America. *J. Climate,* doi:10.1175/JCLI-D-16-0589.1.

571 de Noblet-Ducoudré, N. and Coauthors, 2012: Determining Robust Impacts of Land-Use-

572  Induced Land Cover Changes on Surface Climate over North America and Eurasia:

573   Results from the First Set of LUCID Experiments. *J. Climate,* **25**, 3261-3281,

574   doi:10.1175/JCLI-D-11-00338.1.

575 Dirmeyer, P. A., R. D. Koster, and Z. Guo, 2006: Do Global Models Properly Represent the

576   Feedback between Land and Atmosphere? *J. Hydrometeor.,* **7**, 1177-1198,

577   doi:10.1175/JHM532.1.

578 Foken, T., 2008: THE ENERGY BALANCE CLOSURE PROBLEM: AN OVERVIEW. *Ecol.*

579   *Appl.,* **18**, 1351-1367, doi:10.1890/06-0922.1.

580 Gochis, D.J., W. Yu, and D.N. Yates, 2015. The WRF-Hydro Model Technical Description and

581   User's Guide, Version 3.0. NCAR Technical Document, 120 pp., [Available at

582   https://www.ral.ucar.edu/sites/default/files/public/images/project/WRF_Hydro_User_Gui

583   de_v3.0.pdf].

584 Griebel, A., L. T. Bennett, D. Metzen, J. Cleverly, G. Burba and S. K. Arndt, 2016: Effects of

585   inhomogeneities within the flux footprint on the interpretation of seasonal, annual, and

586   interannual ecosystem carbon exchange. Agricultural and Forest Meteorology, **221**, 50-

587   60, doi:10.1016/j.agrformet.2016.02.002.

588 Janowiak, J. E., V. E. Kousky, and R. J. Joyce, 2005: Diurnal cycle of precipitation determined

589   from the CMORPH high spatial and temporal resolution global precipitation

590   analyses. *Journal of Geophysical Research: Atmospheres,* **110**, n/a-n/a,

591   doi:10.1029/2005JD006156.

592   Juang, J., G. Katul, M. Siqueira, P. Stoy, and K. Novick, 2007: Separating the effects of albedo

593   from eco-physiological changes on surface temperature along a successional

594   chronosequence in the southeastern United States. *Geophys. Res. Lett.,* **34**, n/a-n/a,

595   doi:10.1029/2007GL031296.

596   Lawrence, D. M. and Coauthors, 2016: The Land Use Model Intercomparison Project (LUMIP)

597   contribution to CMIP6: rationale and experimental design. *Geoscientific Model*

598   *Development,* **9**, 2973-2998,

599   Lawrence, D., and K. Vandecar, 2015: Effects of tropical deforestation on climate and

600   agriculture. *Nature climate change,* **5**, 27-36,

601   Lawrence, P. J., and T. N. Chase, 2010: Investigating the climate impacts of global land cover

602   change in the community climate system model. *Int. J. Climatol.,* **30**, 2066-2087,

603   doi:10.1002/joc.2061.

604   Lawrence, P. J. and Coauthors, 2012: Simulating the Biogeochemical and Biogeophysical

605   Impacts of Transient Land Cover Change and Wood Harvest in the Community Climate

606   System Model (CCSM4) from 1850 to 2100. *J. Climate,* **25**, 3071-3095,

607   doi:10.1175/JCLI-D-11-00256.1.

608   Lee, X. and Coauthors, 2011: Observed increase in local cooling effect of deforestation at higher

609   latitudes. *Nature,* **479**, 384-387, doi: 10.1038/nature10588.

610 Lejeune, Q., S. I. Seneviratne, and E. L. Davin, 2016: Historical land-cover change impacts on

611   climate: comparative assessment of LUCID and CMIP5 multi-model experiments. *J.*

612   *Climate,* doi:10.1175/JCLI-D-16-0213.1.

613 Li, Y., M. Zhao, S. Motesharrei, Q. Mu, E. Kalnay, and S. Li, 2015: Local cooling and warming

614   effects of forests based on satellite observations. *Nature communications,* **6**, doi:

615   10.1038/ncomms7603.

616 Luyssaert, S. and Coauthors, 2014: Land management and land-cover change have impacts of

617   similar magnitude on surface temperature. *Nature Climate Change,* **4**, 389-393, doi:

618   10.1038/nclimate2196.

619 Mahmood, R. and Coauthors, 2014: Land cover changes and their biogeophysical effects on

620   climate. *Int. J. Climatol.,* **34**, 929-953, doi:10.1002/joc.3736.

621 Niu, G. and Coauthors, , 2011: The community Noah land surface model with

622   multiparameterization options (Noah-MP): 1. Model description and evaluation with

623   local-scale measurements. *Journal of Geophysical Research: Atmospheres,* **116**, D12109,

624   doi:10.1029/2010JD015139.

625 Oleson, K. W., and co-authors, 2013: Technical Description of version 4.5 of the Community

626   Land Model (CLM), NCAR Technical Note, TN-503+STR, National Center for

627   Atmospheric Research, Boulder, CO, USA, 434pp., [Available at

628   http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM45_Tech_Note.pdf].

629   Pielke, R. A. and Coauthors, , 2011: Land use/land cover changes and climate: modeling analysis

630   and observational evidence. *Wiley Interdisciplinary Reviews: Climate Change,* **2**, 828-

631   850, doi:10.1002/wcc.144.

632   Pitman, A. J. and Coauthors, , 2009: Uncertainties in climate responses to past land cover

633   change: First results from the LUCID intercomparison study. *Geophys. Res. Lett.,* **36**, n/a-

634   n/a, doi:10.1029/2009GL039076.

635   Portmann, F. T., S. Siebert, and P. Döll, 2010: MIRCA2000—Global monthly irrigated and

636   rainfed crop areas around the year 2000: A new high-resolution data set for agricultural

637   and hydrological modeling. *Global Biogeochem. Cycles,* **24**, L14814,

638   doi:10.1029/2008GB003435.

639   Reichstein, M. and Coauthors, , 2005: On the separation of net ecosystem exchange into

640   assimilation and ecosystem respiration: review and improved algorithm. *Global Change*

641   *Biol.,* **11**, 1424-1439, doi:10.1111/j.1365-2486.2005.001002.x.

642   Santanello, J. A., and M. A. Friedl, 2003: Diurnal Covariation in Soil Heat Flux and Net

643   Radiation. *J. Appl. Meteor.,* **42**, 851-862, doi:10.1175/1520-

644   0450(2003)042<0851:DCISHF>2.0.CO;2.

645   Schultz, N. M., X. Lee, P. J. Lawrence, D. M. Lawrence, and L. Zhao, 2016: Assessing the use

646   of subgrid land model output to study impacts of land cover change. *Journal of*

647   *Geophysical Research: Atmospheres,* **121**, 6133-6147, doi:10.1002/2016JD025094.

648   Swenson, S. C., and D. M. Lawrence, 2014: Assessing a dry surface layer-based soil resistance

649   parameterization for the Community Land Model using GRACE and FLUXNET-MTE

650   data. *Journal of Geophysical Research: Atmospheres,* **119**, 10,299-10,312,

651   doi:10.1002/2014JD022314.

652   Tang, J., W. J. Riley, and J. Niu, 2015: Incorporating root hydraulic redistribution in CLM4.5:

653   Effects on predicted site and global evapotranspiration, soil moisture, and water

654   storage. *Journal of Advances in Modeling Earth Systems,* **7**, 1828-1848,

655   doi:10.1002/2015MS000484.

656   Teuling, A. J. and Coauthors, 2010: Contrasting response of European forest and grassland

657   energy exchange to heatwaves. *Nature Geoscience,* **3**, 722-727, doi:10.1038/ngeo950.

658   Vanden Broucke, S., S. Luyssaert, E. L. Davin, I. Janssens, and N. van Lipzig, 2015: New

659   insights in the capability of climate models to simulate the impact of LUC based on

660   temperature decomposition of paired site observations. Journal of Geophysical Research:

661   Atmospheres, 120, 5417-5436, doi:10.1002/2015JD023095.

662   Verseghy, D. L., 1991: Class—A Canadian land surface scheme for GCMS. I. Soil model. *Int. J.*

663   *Climatol.,* **11**, 111-133, doi:10.1002/joc.3370110202.

664   Viovy, N., 2011. CRUNCEP data set for 1901–2008, [Available at

665   https://www.earthsystemgrid.org/dataset/ucar.cgd.ccsm4.CRUNCEP.v4.html].

666   Vuichard, N., and D. Papale, 2015: Filling the gaps in meteorological continuous data measured

667   at FLUXNET sites with ERA-Interim reanalysis. *Earth System Science Data,* **7**, 157-171,

668   doi:10.5194/essd-7-157-2015.

669   Williams, C. A. and Coauthors, 2012: Climate and vegetation controls on the surface water

670   balance: Synthesis of evapotranspiration measured across a global network of flux

671   towers. *Water Resour. Res.,* **48**, W06523, doi:10.1029/2011WR011586.

672   Wilson, K. and Coauthors, 2002: Energy balance closure at FLUXNET sites. *Agric. For.*

673   *Meteorol.,* **113**, 223-243, doi:10.1016/S0168-1923(02)00109-0.

674   Xu, Z., R. Mahmood, Z. Yang, C. Fu, and H. Su, 2015: Investigating diurnal and seasonal

675   climatic response to land use and land cover change over monsoon Asia with the

676   Community Earth System Model. *Journal of Geophysical Research: Atmospheres,* **120**,

677   1137-1152, doi:10.1002/2014JD022479.

678   Zhang, L. and Coauthors, , 2016: Evaluation of the Community Land Model simulated carbon

679   and water fluxes against observations over ChinaFLUX sites. *Agric. For. Meteorol.,* **226–**

680   **227**, 174-185, doi:10.1016/j.agrformet.2016.05.018.

681   Zhou, C., and K. Wang, 2016: Biological and Environmental Controls on Evaporative Fractions

682   at AmeriFlux Sites. J. Appl. Meteor. Climatol., **55**, 145-161, doi:10.1175/JAMC-D-15-

683   0126.1.

684

685

686   **Table 1.** Information about the variables used from FLUXNET2015. The marginal distribution

687   sampling (MDS) filling method is based on Reichstein et al. (2005); and the ERA-interim filling

688   method can be found in Vuichard and Papale (2015).

Name	Gap-filling	Description
SW_IN_F	MDS and ERA-interim	downwelling shortwave radiation
LW_IN_F	MDS and ERA-interim	downwelling longwave radiation
PA_F	MDS and ERA-interim	atmospheric pressure
TA_F	MDS and ERA-interim	air temperature
VPD_F	MDS and ERA-interim	vapor pressure deficit
P_F	ERA-interim	precipitation
WS_F	ERA-interim	wind speed
LE_F_MDS	MDS	latent heat flux
H_F_MDS	MDS	sensible heat flux
G_F_MDS	MDS	ground heat flux
NETRAD	n/a	net radiation
LE_CORR	n/a	corrected LE_F_MDS by energy balance closure correction factors. LE_CORR_25, LE_CORR, and LE_CORR_75 are calculated based on 25, 50, and 75th percentiles of the factors, respectively.
H_CORR	n/a	corrected H_F_MDS by energy balance closure correction factors. H_CORR_25, H_CORR, and H_CORR_75 are calculated based on 25, 50, and 75th percentiles of the factors, respectively.

689

690

691   **Table 2**. Information about model simulations. "Nearby" observations indicate that the paired

692   sites have the identical forcings either from the companion forest or open sites.

Name	Forcings	Description
CLM	observations from individual sites	single-point CLM simulations with its own observations
CLM_forest	observations only from forest sites	single-point CLM simulations with the (nearby) forest observations
CLM_open	observations only from open sites	single-point CLM simulations with the (nearby) open land observations
CLM-PFT	CRUNCEP	global CLM simulations with default soil-column scheme with PFT-level output
CLM-PFTCOL	CRUNCEP	global CLM simulations with default individual-soil-column scheme scheme with PFT-level output
NOAH-MP	observations from individual sites	single-point NOAH-MP simulations with its own observations
NOAH-MP_forest	observations only from forest sites	single-point NOAH-MP simulations with the (nearby) forest observations
NOAH-MP_open	observations only from open sites	single-point NOAH-MP simulations with the (nearby) open land observations

693

694

695

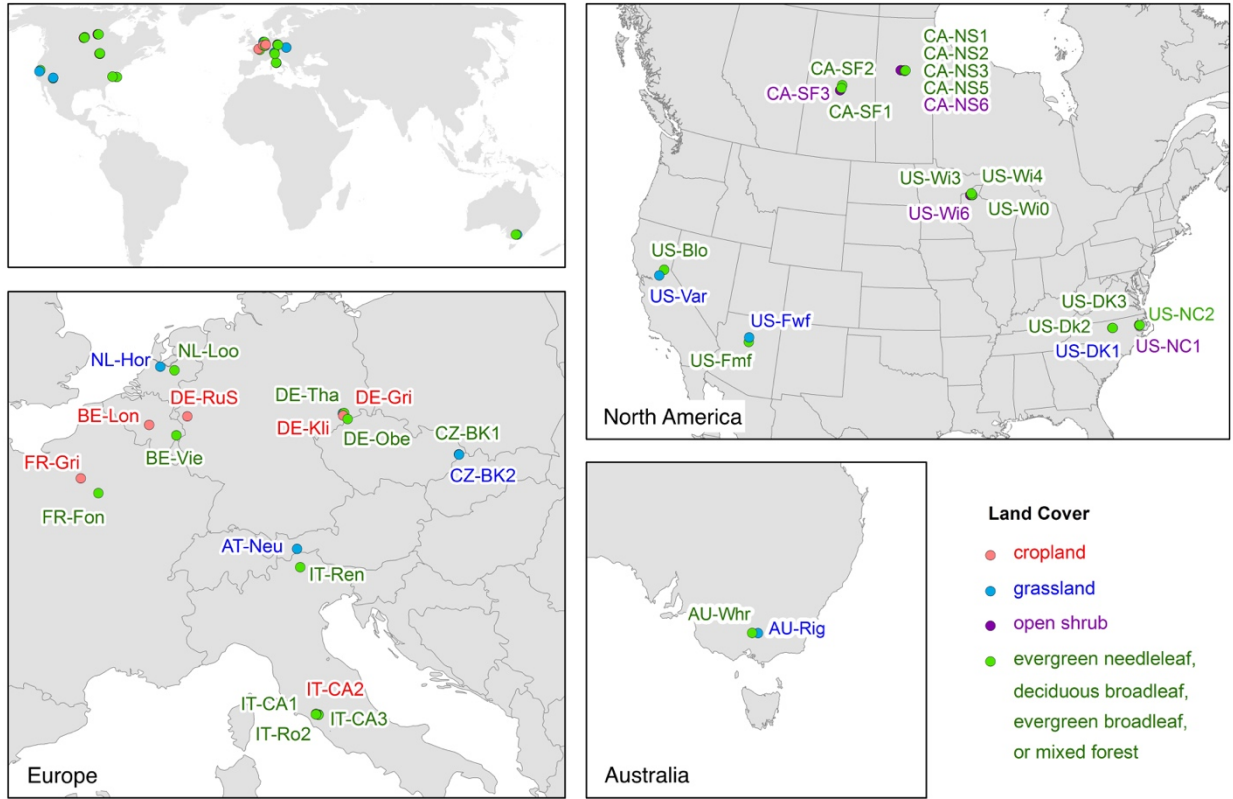



Figure 1. Location and land cover type of the paired sites. The land cover type of each site is

based on the reported land cover in FLUXNET database.

698

699

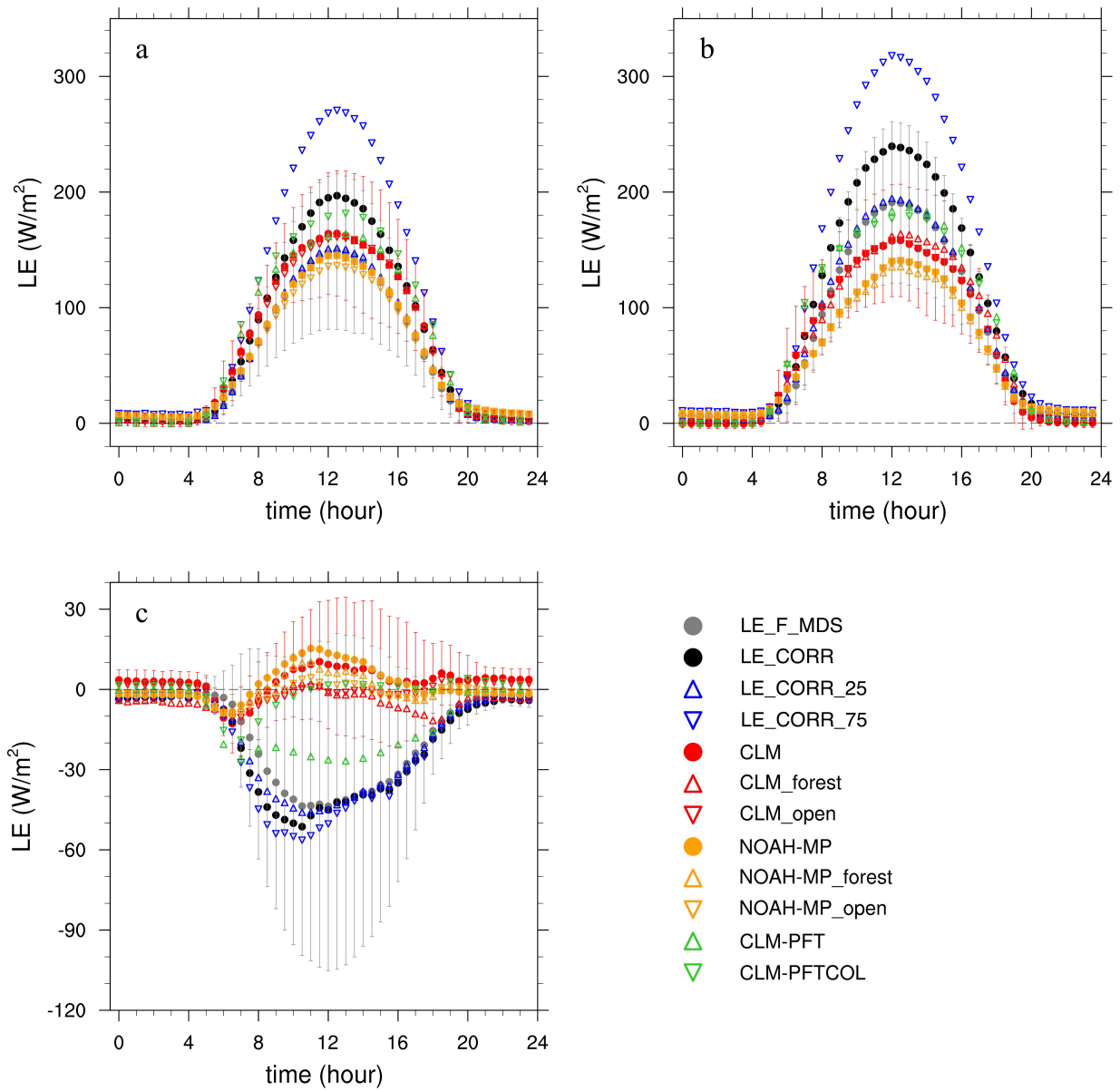
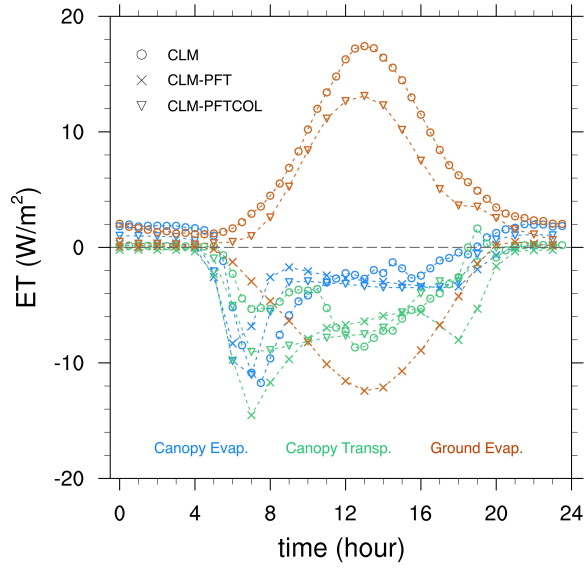
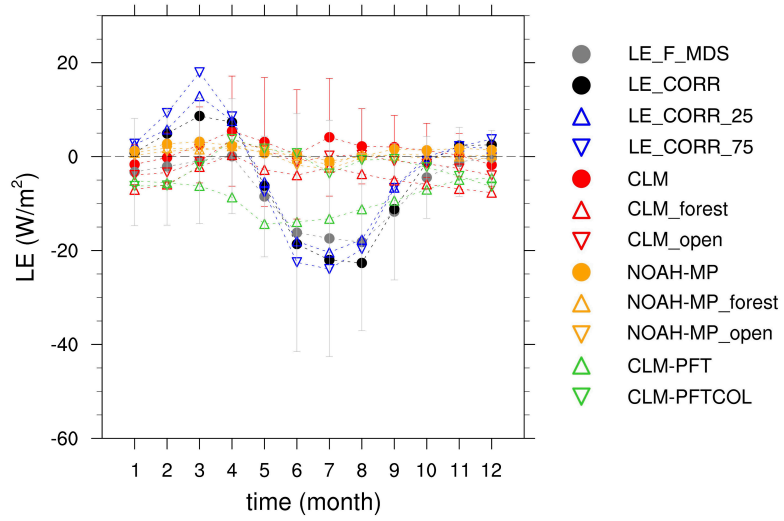



Figure 2. The diurnal cycle of *LE* (W/m²) averaged over all the open sites (a) and forest sites (b) and their difference (open – forest, c) during the summer. The gray error bars indicate the standard deviation of the observed *LE* (MDS) among the sites; the red error bars are for the simulated *LE* in the CLM case. Details about the four types of FLUXNET observations can be found in Table 1. Information about model simulations in CLM and Noah is described in Table 2.

706

707 **Figure 3**. Change in the diurnal cycle of components (colors) of evapotranspiration (canopy

708 evaporation, canopy transpiration, and ground evaporation) due to LULCC from forest to open

709 land (open – forest).

710

711

712   **Figure 4**. Change in the seasonal cycle of *LE* (W/m^2) due to LULCC from forest to open land

713   (open – forest).

714

715

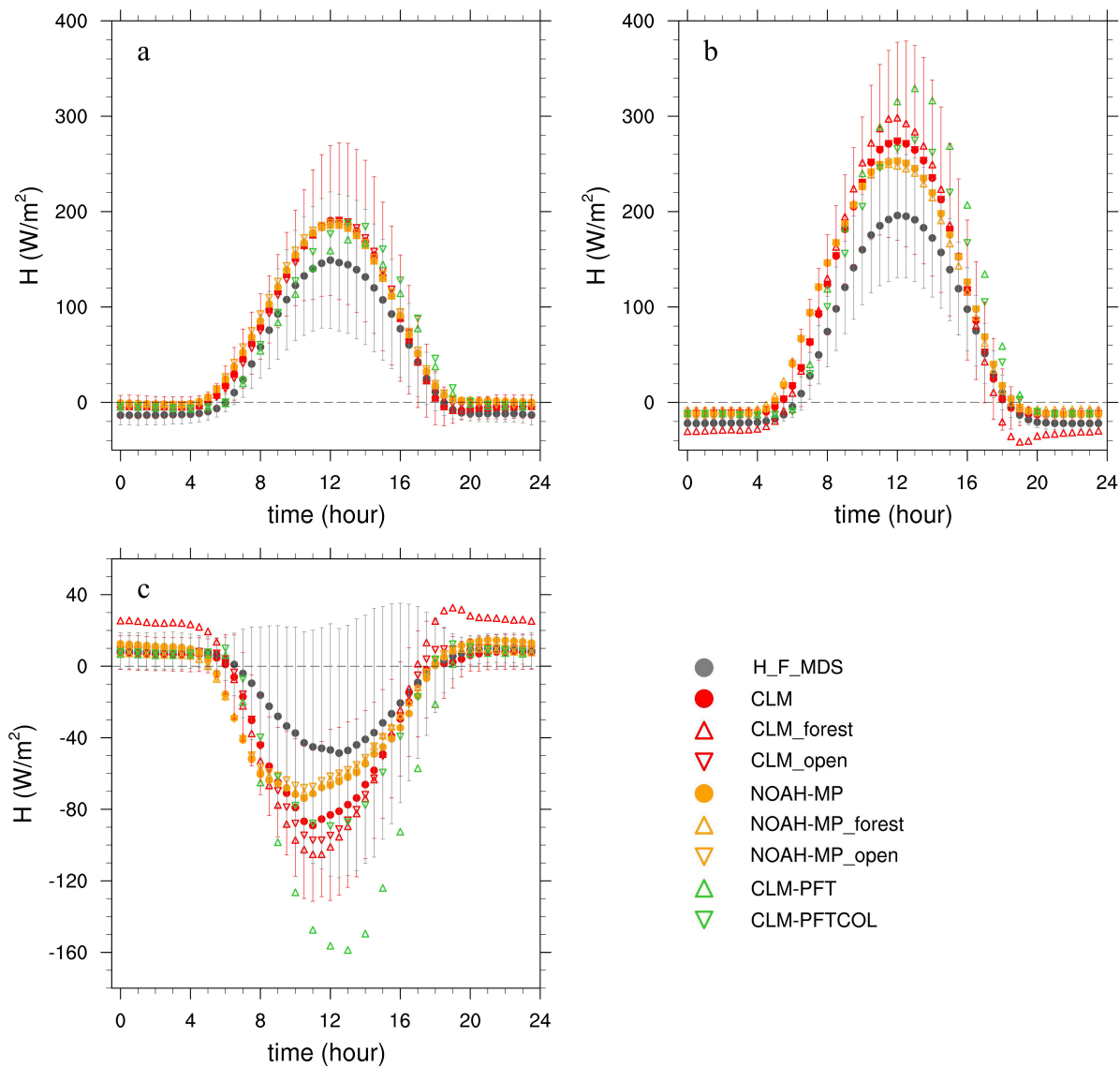



Figure 5. The diurnal cycle of H (W/m^2) averaged over all the open sites (a) and forest sites (b)

and their difference (open – forest, c) during the summer. The gray error bars indicate the

standard deviation of the observed H among the sites; the red error bars are for the simulated H

in the CLM case.

720

721

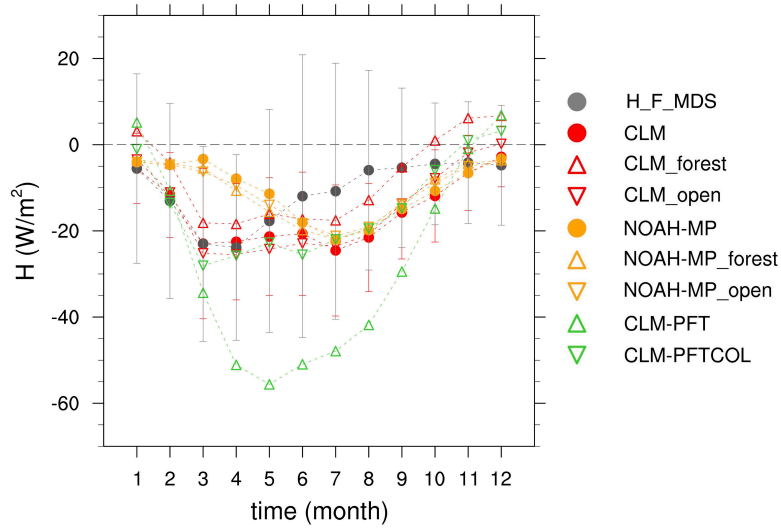



Figure 6. Change in the seasonal cycle of H (W/m^2) due to LULCC from forest to open land

(open − forest).

724

725

726

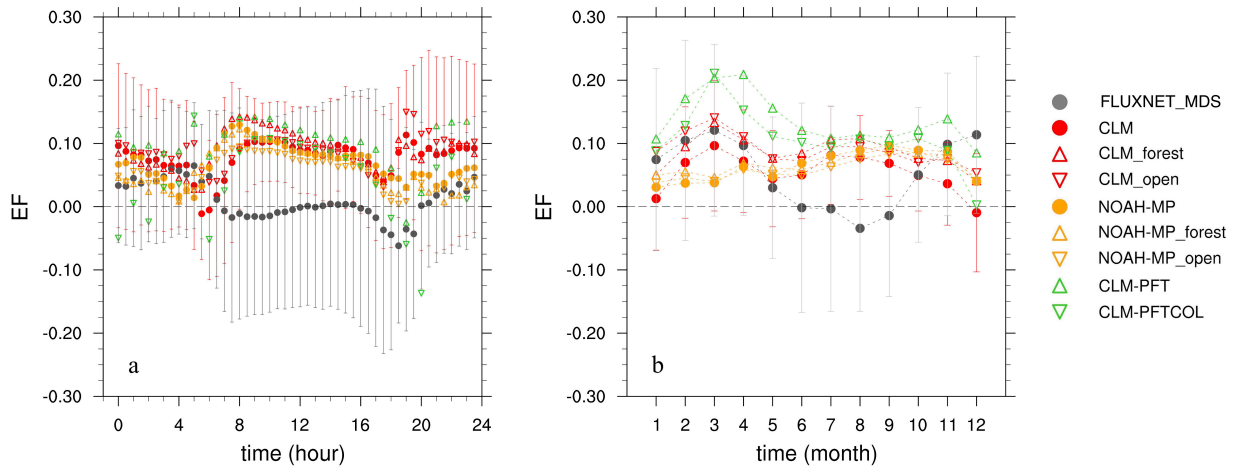



Figure 7. Change in the summer diurnal (a) and seasonal (b) cycle of EF (unitless) due to

LULCC from forest to open land (open − forest). The observed EF (FLUXNET_MDS) is

calculated based on the MDS gapfilled LE (LE_F_MDS) and H (H_F_MDS).

730

731

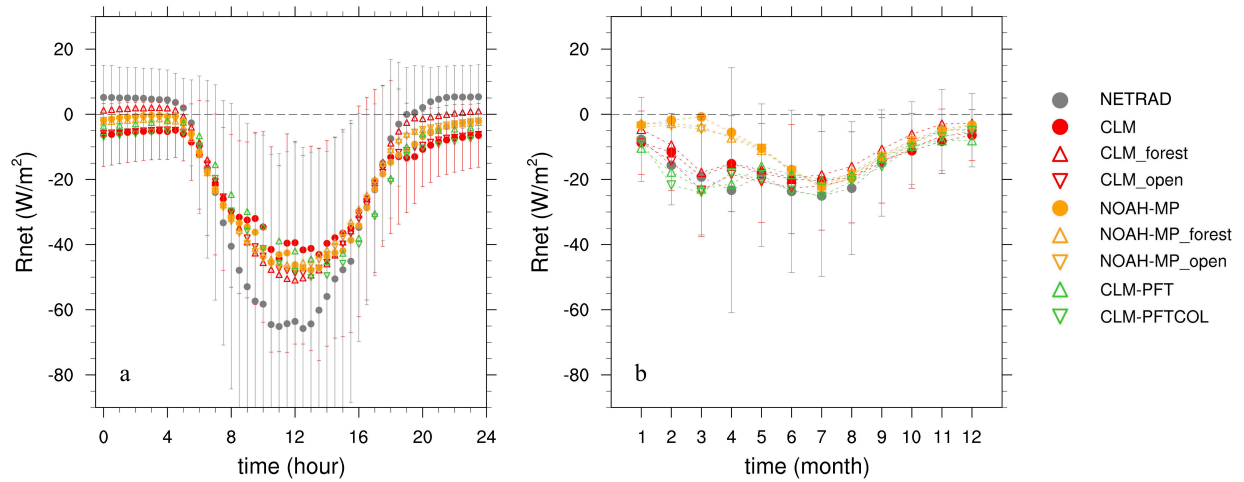



Figure 8. Change in the summer diurnal (a) and seasonal (b) cycle of *G* (W/m²) due to LULCC

from forest to open land (open − forest). It should be noted that the changes in the CLM-PFT

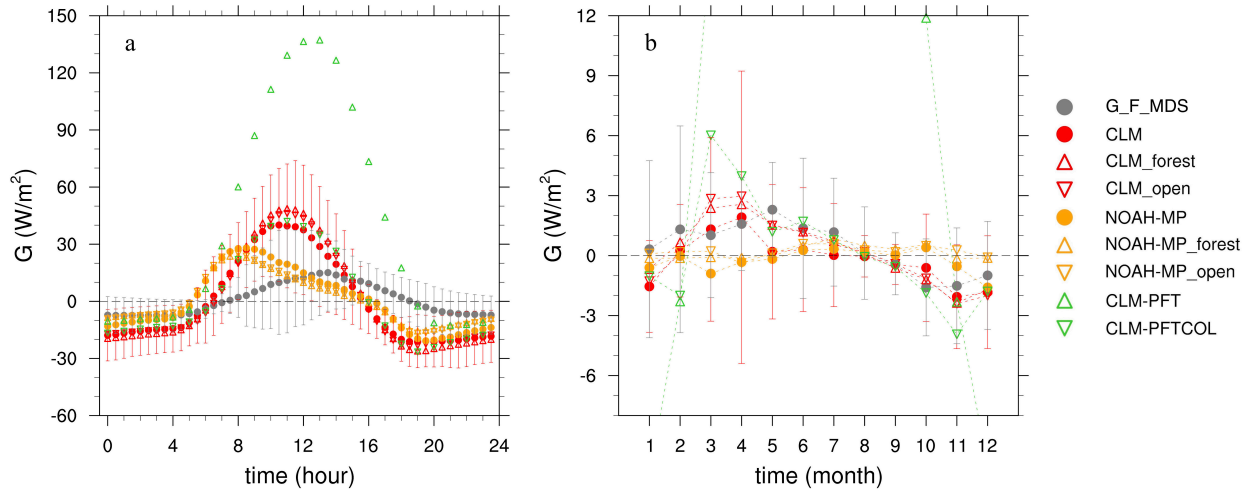
simulation are much further from the observations than the other simulations. Some of its values

are beyond the limit of the figure (b). The smallest value is -11.2 W/m² in January; while the

largest value is 52.9 W/m² in May.

737

738

739

740

741 **Figure 9**. Change in the summer diurnal (a) and seasonal (b) cycle of R_{net} (W/m^2) due to LULCC

742 from forest to open land (open – forest).

743

744

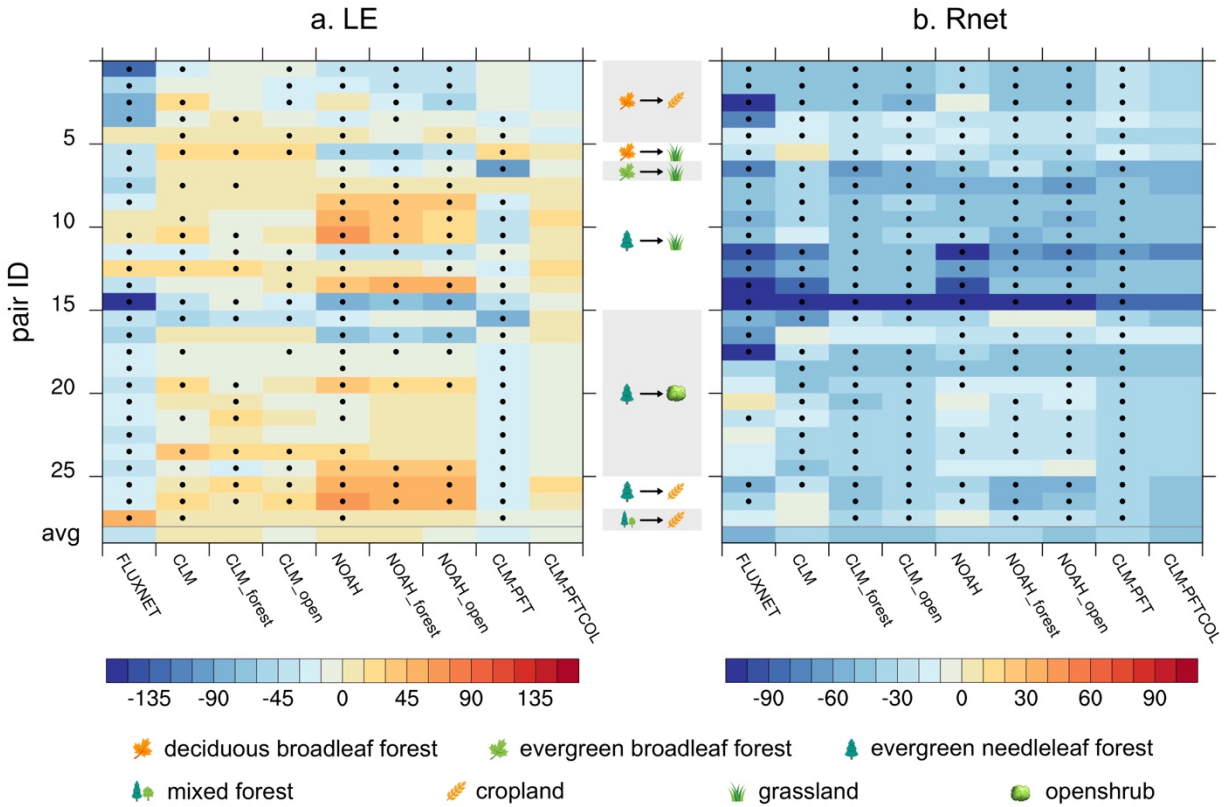



Figure 10. Change (open – forest) in observed and simulated *LE* (a) and R_{net} (b) during summer

daytime (averaged during the period 08:00 ~ 16:00) over individual pairs and their averages. The

vertical labels show the pair ID from 1 to 28 based on Table S1. The pairs are grouped based on

the type of LULCC (shown as the icons in the middle). The bottom row is the average over all

pairs. The Student's t-test is performed on the daily (daytime average) time series for each pair.

Dots indicate statistically significant changes at the 95% confidence level. No significant test is

carried out for the CLM-PFTCOL simulation (the last column), because we only have long-term

averaged hourly output for each month.

756

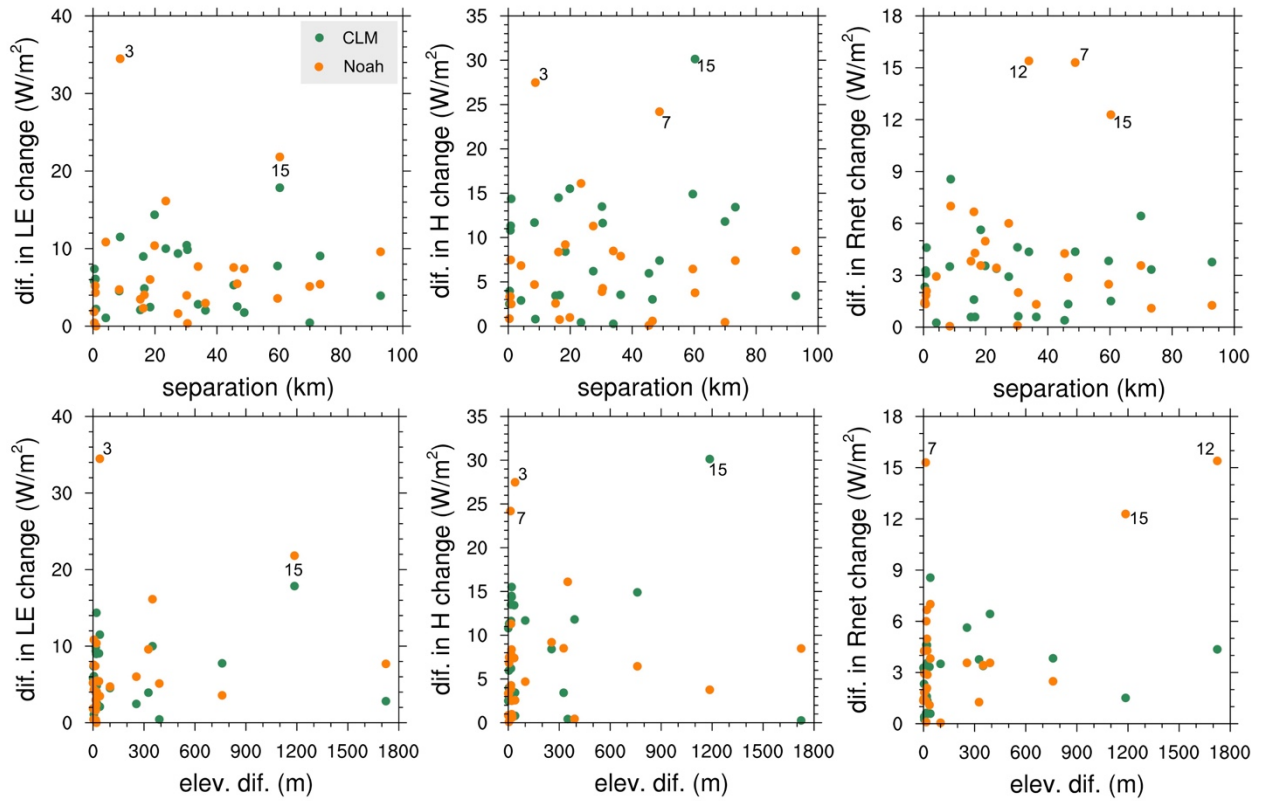



Figure 11. Sensitivity of differences in simulated surface energy flux changes (left column: *LE*,

middle: *H*, and right: *R~net~*) between "forest forcing" and "open forcing" simulations to site

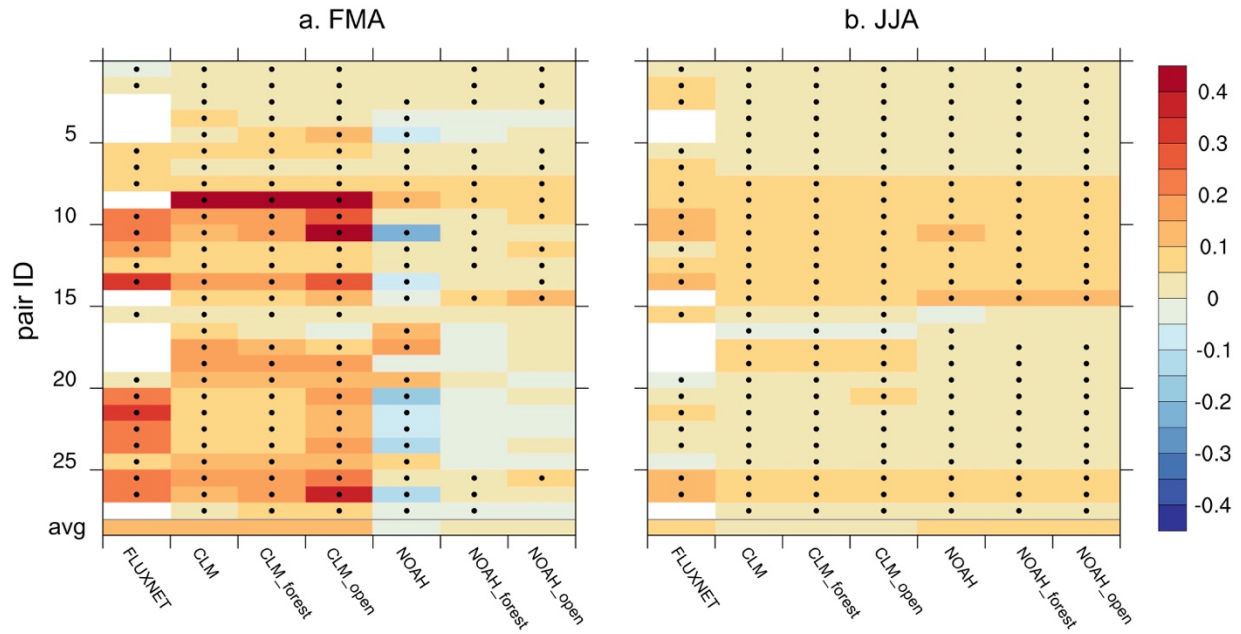
separation (top) and elevation difference (bottom) between the forest and open sites in individual

pairs. The pairs No. 3, 7, 12 and 15 are labeled because of the greatest differences in surface

fluxes changes.

762

763

764

765

766   **Figure 12**. Change (open – forest) in observed and simulated daytime albedo during late

767   winter/early spring (FMA, a) and summer (JJA, b). White areas indicate missing observations.

768