# Emergent climate and CO2 sensitivities of net primary productivity in ecosystem models do not agree with empirical data in temperate forests of eastern North America

Article

Accepted Version

www.reading.ac.uk/centaur

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

1   **Title:** Emergent climate and $CO_2$ sensitivities of net primary productivity in ecosystem models

2   do not agree with empirical data in temperate forests of eastern North America

3

4   **Running Head:** Model and data NPP sensitivity to climate

5

6   **Authors**

7   **Rollinson, Christine R.[1,2]\*,** Yao Liu[3], Ann Raiho[4], David J.P. Moore[3], Jason McLachlan[4],

8   Daniel A. Bishop[5], Alex Dye[6], Jaclyn Hatala Matthes[7], Amy Hessl[6], Thomas Hickler[8,9], Neil

9   Pederson[5], Benjamin Poulter[10], Tristan Quaife[11], Kevin Schaefer[12], Jörg Steinkamp[8], Michael C.

10  Dietze[1]

11

12  **Author Affiliations**

13      1.  Department of Earth & Environment, Boston University, Boston, MA USA

14      2.  Morton Arboretum, Lisle, IL, USA

15      3.  School of Natural Resources and the Environment, University of Arizona, AZ, USA

16      4.  Department of Biological Sciences, University of Notre Dame, South Bend, IN, USA

17      5.  Harvard Forest, Harvard University, Petersham, MA, USA

18      6.  Department of Geology and Geography, West Virginia University, Morgantown, WV,

19          USA

20      7.  Department of Biological Sciences, Wellesley College, Wellesley, MA USA

21      8.  Senckenberg Biodiversity and Climate Research Centre (BiK-F), Frankfurt/Main,

22          Germany

23   9.  Department of Physical Geography, Geosciences, Goethe University, Frankfurt/Main,

24      Germany

25   10. Institute on Ecosystems and Department of Ecology, Montana State University,

26      Bozeman, MT 59717, USA

27   11. Department of Meteorology, University of Reading, Reading, BERKS, UK.

28   12. National Snow and Ice Data Center, Cooperative Institute for Research in Environmental

29      Sciences, University of Colorado at Boulder, Boulder, CO, USA

30

31   **\* Corresponding Author:** crollinson@mortonarb.org, +1 630-719-2422

32

33   **Keywords:** climate sensitivity, emergent response, climate change, net primary productivity,

34   ecosystem model, model-data comparison, paleoecology, tree rings

35

36   **Paper Type:** Primary research article

37

38      **Abstract**

39          Ecosystem models show divergent responses of the terrestrial carbon cycle to global

40      change over the next century. Individual model evaluation and multi-model comparisons with

41      data have largely focused on individual processes at sub-annual to decadal scales. Thus far, data-

42      based evaluations of emergent ecosystem responses to climate and $CO_2$ at multi-decadal and

43      centennial time scales have been rare. We compared the sensitivity of net primary productivity

44      (NPP) to temperature, precipitation, and $CO_2$ in ten ecosystem models with the sensitivities

45      found in tree-ring reconstructions of NPP and raw ring-width series at six temperate forest sites.

46      These model-data comparisons were evaluated at three temporal extents to determine whether

47      the rapid, directional changes in temperature and $CO_2$ in the recent past skew our observed

48      responses to multiple drivers of change.  All models tested here were more sensitive to low

49      growing season precipitation than tree-ring NPP and ring widths in the past 30 years, although

50      some model precipitation responses were more consistent with tree rings when evaluated over a

51      full century. Similarly, all models had negative or no response to warm growing season

52      temperatures while tree-ring data showed consistently positive effects of temperature.  Although

53      precipitation responses were least consistent among models, differences among models to $CO_2$

54      drive divergence and ensemble uncertainty in relative change in NPP over the past century.

55      Changes in forest composition within models had no effect on climate or $CO_2$ sensitivity.  Fire in

56      model simulations reduced model sensitivity to climate and $CO_2$, but only over the course of

57      multiple centuries. Formal evaluation of emergent model behavior at multi-decadal and multi-

58      centennial time scales is essential to reconciling model projections with observed ecosystem

59      responses to past climate change. Future evaluation should focus on improved representation of

60 Model and data NPP sensitivity to climate 4
disturbance and biomass change as well as the feedbacks with moisture balance and $CO_2$ in

61 individual models.

62

63    **Introduction**

64            Changes in temperature and precipitation regimes over the past millennium have been

65    associated with shifts in ecosystem composition and structure in the paleoecological record

66    (Prentice *et al.*, 1991; Davis & Shaw, 2001; Shuman *et al.*, 2002; Clifford & Booth, 2015).

67    Modern empirical data from experiment- and observation-based studies provide evidence that

68    rapidly increasing temperatures, altered precipitation regimes, and rising atmospheric $CO_2$

69    concentrations are causing changes in ecosystem dynamics today (Boisvenue & Running, 2006;

70    Morin *et al.*, 2009; Fisichelli *et al.*, 2013; Peñuelas *et al.*, 2013). However, the effects of climate

71    change in long-lived ecosystems such as the temperate forests of the eastern and midwestern

72    United States remain contested because changes in forest composition and structure take decades

73    to centuries to occur (Renwick & Rocca, 2014). These observation-based studies capture

74    emergent, long-term ecosystem responses to climate and $CO_2$ variability that are the product of

75    feedbacks and interactions among physiological and biogeochemical processes.

76            Terrestrial ecosystem models are used to make mechanistic, process-based projections of

77    ecosystem response to changing climate and $CO_2$ in the past, present, and future.  However, the

78    mechanistic, bottom-up approach used to build and evaluate ecosystem models is mismatched in

79    spatial and temporal scale from observations of ecosystem response to climate change.

80    Differences in model parameterization or structural representation of physiological process can

81    cause major divergences in the resulting ecosystem dynamics through time (De Kauwe *et al.*,

82    2013; Walker *et al.*, 2015).  Model intercomparisons of ecosystem response to climate change

83    over the next century typically show increased divergence of ecosystem dynamics at the end of

84    the century (e.g. Friedlingstein *et al.*, 2014).  Although most policy and management decisions

85    occur at sub-centennial temporal scales, life spans of common temperate tree species in the

86  northeastern United States can range from 120 years for *Betula papyrifera* to over 800 years for

87  *Tsuga canadensis* (Burns & Honkala, 1990). This means that in the absence of widespread

88  disturbance or management, even the model simulations of ecosystem response to climate

89  change are far shorter than the multi-generational scales at which changes in forest composition

90  or structure in response to shifts in climate occur.  These slow, gradual shifts in forest

91  composition or structure may mediate ecosystem sensitivity to climate and $CO_2$ in both models

92  and reality. For example, mortality and recruitment are inherent processes underlying the types

93  of ecosystem-scale responses to climate change observed in the past, and occur at temporal and

94  spatial scales beyond that which we can observe with current available ecological data.  This can

95  then cause the observed responses of ecosystems to past climate change over multiple centuries

96  to be different from that observed in shorter studies focused on a few decades.

97          Tree rings provide annually resolved records of individual and forest response to

98  environmental variation over the temporal scales of decades and centuries across temperate

99  forests.  Tree rings have been used to understand forest responses to climate variability at

100  individual sites and entire continents (Williams *et al.*, 2013; Charney *et al.*, 2016; D'Orangeville

101  *et al.*, 2016), but these approaches have been largely disconnected from assessments of how

102  climate change impacts ecosystems in models.  Tree rings are increasingly used to quantify and

103  constrain components of the terrestrial carbon cycle in forests as well as individual- and forest-

104  level responses to climate (Graumlich *et al.*, 1989; Davis *et al.*, 2009; Babst *et al.*, 2013; Dye *et*

105  *al.*, 2016).  These efforts are critical for understanding long-term forest responses to climate

106  variability as multiple field experiments have indicated that long-term individual and ecosystem

107  responses to warming and elevated $CO_2$ diverge from initial responses found in the first few

108  years of manipulation (Melillo *et al.*, 2002; Hollister *et al.*, 2005; Matesanz *et al.*, 2009; Norby *et*

109    *al.*, 2010).  This long-term evaluation of climate impacts on forest ecosystems is particularly

110    important for disentangling the simultaneous directional shifts in temperature, precipitation, $CO_2$,

111    and disturbance that complicate much of modern ecological research (Foster *et al.*, 1998; Turner

112    *et al.*, 2003; Gómez-Aparicio *et al.*, 2011).

113         This paper compares emergent sensitivity of net primary productivity (NPP) to climate

114    and $CO_2$ in ecosystem models to those found in tree-ring data at multiple temporal scales.  The

115    goals of this paper are: 1) determine whether annual ecosystem NPP sensitivity to climate and

116    $CO_2$ in ecosystem models matches those found in individual tree- and forest-level tree-ring data;

117    2) compare the climate and $CO_2$ sensitivities in models and data from short temporal extents

118    where patterns are dominated by inter-annual climate variability with those from centennial-scale

119    records that contain low-frequency climatic shifts; and 3) quantify the contribution of differences

120    in model responses to climate and $CO_2$ to model ensemble uncertainty in NPP dynamics through

121    time.  To achieve these goals, we analyze the temporal trends of NPP in ten ecosystem models

122    that have been run from 850 to 2010 A.D. at six temperate forest sites in the Upper Midwestern

123    and Northeastern United States.  The influences of growing season temperature, precipitation,

124    and $CO_2$ on change in model NPP are then compared to those found in plot-level NPP

125    reconstructions from tree rings at two sites as well as raw ring widths from plot-based sampling

126    and the International Tree Ring Databank (ITRDB) at four sites.  Climate and $CO_2$ responses are

127    analyzed for three temporal extents: 1) the scale of past paleoecological responses to climate

128    change (850-2010 A.D.); 2) the period of historical temperature and precipitation records (1901-

129    2010 A.D.); and 3) the scale of robust modern ecological data (1980-2010 A.D.).  We use these

130    results to explain model-data discrepancies in regional drivers of NPP through time as well as

131    quantify the drivers of uncertainty within the model ensemble over the past millennium.

132

133     **Materials and Methods**

134     *Model experimental overview*

135          We modeled ecosystem carbon flux and composition change in response to climate

136     variation from 850-2010 A.D. at six sites in the Northeastern and Upper Midwestern United

137     States. These sites are located on an east-west gradient that represent present-day differences in

138     temperature, precipitation, and forest types (Table 1, S1). Models were presented with common,

139     continuous 6-hourly meteorological climate and monthly $CO_2$ forcing data.  $CO_2$ was taken from

140     the Law Dome time series (850-2000) and the NOAA Mauna Loa record (2001-2010) (Keeling

141     *et al.*, 2005; Wei *et al.*, 2014). Climate drivers included air temperature, water-equivalent

142     precipitation rate, total incoming shortwave radiation, incoming longwave radiation, surface air

143     pressure, specific humidity, and wind speed.  The continuous, 6-hourly 850-2010 meteorological

144     driver set was developed by using an artificial neural network to spatially and temporally

145     downscale output from CCSM4 output from the Paleoclimate Modeling Intercomparison Project,

146     Phase III (PMIP3) past millennium simulations (850-1849, Crucifix *et al.*, 2012) and Coupled

147     Model Intercomparison Project, Phase 5 (CMIP5) simulations (1850-1900, Taylor *et al.*, 2012)

148     using 6-hourly, 0.5-degree CRUNCEP data (1901-2010, Wei *et al.*, 2014) according to Kumar *et*

149     *al.*, (2012).  Due to mismatches at transitional periods between CCSM4 and CRUNCEP

150     products, all variables except wind were bias-corrected to avoid sharp jumps in climatology in

151     the time series.  Temperature was corrected as an additive bias whereas short- and long-wave

152     radiation, and precipitation were corrected using a ratio bias to conserve the hydrological

153     sensitivity among meteorological variables (Hempel *et al.*, 2013).

154          Ten models representing variants of five independent ecosystem models completed

155    simulations of ecosystem dynamics at all six sites from 850-2010 A.D. and provided monthly- or

156    annual-resolution output on composition and the carbon cycle (Table 2).  Full model protocol

157    can be found in Supporting Information 2. Dynamic vegetation models allow plant communities

158    to self-assemble and change through time while static vegetation models prescribe plant

159    functional types (PFTs) by fractional area based on potential vegetation distribution from

160    Ramankutty and Foley (1999).  All models were spun-up to steady-state at 850 A.D. by cycling

161    the first 20 years of forcing data (850-869 A.D.) with a constant $CO_2$ concentration of 277

162    ppm.   Not all models used all meteorological variables as drivers.  LINKAGES was the only

163    model to not include $CO_2$ as a driver and only includes temperature and precipitation drivers

164    whereas all other models included at a minimum, temperature, precipitation, shortwave radiation,

165    and $CO_2$.  Even though LINKAGES does not include $CO_2$ as a driver of ecosystem dynamics, it

166    was included in our analyses to help isolate potential confounding effects of increasing $CO_2$ with

167    other drivers of ecosystem change.

168          Our analyses focus on two model characteristics (dynamic vegetation and fire) and the

169    dynamics of two ecosystem properties (composition, aboveground biomass) that are linked to

170    slow processes associated with ecosystem responses to climate and $CO_2$ change (Table 2).

171    However, because PFTs in all models except for SiBCASA, a biome-based model, respond to

172    climate independently, they may still experience shifts in the relative community composition if

173    defined by fluctuations in PFT biomass.  Composition variability through time was quantified as

174    the standard deviation of percent evergreen biomass through time within each site for each

175    model.  Although fire was possible in ED, CLM, and LPJ model variants, the conditions for fire

176    were not triggered in ED2 and ED2-LU, so the effects of fire were only present in four models

177    (CLM-BGC, CLM-CN, LPJ-WSL, LPJ-GUESS).  Fire in ecosystem models primarily affects

178    ecosystems by reducing biomass, but mortality or changes in growth can also cause similar

179    biomass fluctuations in all models. To calculate biomass variability, biomass was first

180    normalized to the site mean so that biomass through time was expressed as a percent of mean.

181    Biomass variability was then calculated as the standard deviation of this normalized biomass

182    through time at each site for each model.  All models except JULES-STATIC used aboveground

183    biomass (AGB) as the measure of biomass variability.  Leaf area index was used as a proxy for

184    biomass in JULES-STATIC because there is no biomass or vegetation dynamics beyond leaf

185    area simulated in JULES without coupling to the TRIFFID model.  In other models LAI had a

186    mean correlation with AGB of $0.73 \pm 0.18$ and ranged from 0.55 in CLM-BGC to 0.99 in

187    LINKAGES.  In JULES-TRIFFID, the correlation between LAI and AGB was 0.93.

188

189

190    *Empirical Data*

191           We used two types of tree-ring data to provide an empirical estimate of climate

192    sensitivity for trees and forests from our study region.  First, stand-level aboveground NPP was

193    reconstructed from sampling mapped fixed-area plots according to Dye *et al.*, (2016), which

194    enable aboveground NPP to be calculated on a per area basis that is comparable to ecosystem

195    model output (Table 1).  NPP was reconstructed from five plots in two separate stands at

196    Harvard and three plots in a single stand at Howland. A nested sampling scheme was used at

197    both locations where two to three increment cores were taken from all trees greater than 10 cm

198    diameter at breast height (DBH) in a 13-m radius plot and trees greater than 20 cm DBH were

199    cored in a 20-m radius plot.  Growth measurements from all cores for each were averaged and

200   used to reconstruct DBH through time that was then converted to aboveground biomass and

201   annual aboveground biomass increment (proportional to NPP) for each tree and plot using

202   species-specific equations that can be found in the appendix of Dye *et al.*, (2016).  Due to the

203   potential decline in sample replication, and the potential for unmeasured mortality, analyses of

204   tree-ring NPP sensitivity to climate were restricted to 1980-2010 (Clark *et al.*, 2001; Foster *et al.*,

205   2014).

206         In addition to tree-ring estimates of NPP, we also analyzed the sensitivity of raw ring

207   width (RW) to climate from the trees used to generate the NPP estimates as well as from nine

208   datasets from the International Tree Ring Databank (ITRDB) (Grissino-Mayer & Fritts, 1997).

209   Records from the ITRDB were found in close proximity to the Demming Lake, Howland, and

210   UNDERC sites, allowing more robust comparisons of climate sensitivity across the modeled

211   study sites (S3).  In cases where multiple cores existed for a tree, the mean of ring width for each

212   year was used. Although we include supplemental sensitivity analyses using *a priori* detrended

213   ring width index (RWI, S4), we chose raw ring widths as our response variable rather than RWI

214   or basal area increment (BAI) for two reasons.  First, we chose not to detrend rings widths *a*

215   *priori* for our primary analyses in order to preserve potential low-frequency effects of changing

216   climate or $CO_2$ in the modern era that might be confounded with increasing stem size. Second,

217   BAI was not used because measurements of actual diameters were not available for ITRDB

218   trees. The CRUNCEP climate records used to drive the models were matched to the tree-ring

219   data for each site.

220

221   *Sensitivity Analysis*

222          Model and tree-ring sensitivities to climate were analyzed using generalized additive

223   models (GAMs) where NPP or RW is the sum of non-linear temperature, precipitation and $CO_2$

224   effects. In these analyses, we used site-level annual NPP as the response variable for ecosystem

225   models, plot-level annual aboveground woody increment for tree-ring NPP, and mean raw ring

226   width for each individual in the RW analysis.  Temperature, precipitation, and $CO_2$ effects were

227   estimated with three-knot thin plate regression splines while an additional size effect was fit with

228   three-knot cubic smoothing spline using the *gam* function of the *mgcv* package in R 3.2.3 (Wood,

229   2012).  Even though LINKAGES lacks $CO_2$ as a model driver, this effect was included in our

230   statistical analyses as a test of the statistical model's ability to attribute variation in NPP to

231   climate and $CO_2$. The size term was included to account for effects of biomass (models, tree-ring

232   NPP) or individual size (raw ring width) that could alter the maximum potential NPP or ring

233   width independent of climate.  Total site or plot aboveground biomass was used for the size

234   effect in models and tree-ring NPP GAMs, while reconstructed DBH was used for size in the

235   RW GAMs.  DBH was reconstructed by subtracting ring widths from the DBH at the time of

236   sampling at Harvard and Howland, but was reconstructed by summing ring widths from the

237   ITRDB records, where no DBH information was available. Because ring width sums may not

238   accurately reflect the true DBH of ITRDB trees, in RW models the size effect was fit separately

239   for each ITRDB core. This effectively individually detrends ITRDB cores by removing effects of

240   increasing size through time while simultaneously assessing the climate and $CO_2$ sensitivity of

241   both ITRDB and plot-based samples.  We also performed supplemental analyses with *a priori*

242   detrended RWI using a two-thirds spline that removes long-term trends in tree-ring series (S5).

243   Because *a priori* detrending removes age- and size-based trends in the data, the size factor was

244   removed from our climate and $CO_2$ GAM in RWI analyses.  We also include an analysis with all

245   datasets restricted to Harvard and Howland sites where tree-ring NPP records are available (S6).

246         We normalized the sensitivities of model NPP and tree-ring width to climate and $CO_2$ for

247   all analyses to facilitate comparison of model agreement of change in NPP through time and

248   relative influences of temperature, precipitation, and $CO_2$ as drivers of that change. In all cases,

249   the response variable of NPP or RW was converted to a percent deviation from the model or

250   tree-ring dataset mean.  We assessed model agreement of both change in NPP and drivers of

251   change as the standard deviation around the ensemble mean before and after 1901, where $CO_2$

252   begins to sharply increase. To prescribe the effects of individual drivers on change in NPP or

253   RW through time, we quantified the relative effects of temperature, precipitation, and $CO_2$ on

254   NPP or RW in each year. The $CO_2$ effect was offset from the GAM-estimated effect so that the

255   beginning of simulation $CO_2$ concentration in 850-855 A.D. received a weight of zero.

256         Models were analyzed at three temporal scales to determine whether observations from

257   short periods in the modern era were able to capture the relationships between climate and $CO_2$

258   change and ecosystem responses over periods with low-frequency, directional shifts in climate.

259   These three temporal extents are: 1) scale of robust, modern empirical records (1980-2010), 2)

260   the scale of observational climate data (1901-2010), and 3) the full model simulation extent

261   (850-2010).   This model-centric analysis was complemented with comparisons to the shifts in

262   sensitivity seen in the RW models at the two shorter scales.  Because GAMs center the spline-

263   based effects on the means of the given data (i.e. the mean of observed temperature,

264   precipitation, and $CO_2$), all three temporal scales were re-centered on their respective 1980-2010

265   means to facilitate cross-scale comparisons. Differences in climate and $CO_2$ sensitivities within

266   and across temporal scales both within individual models and at the ensemble level were

267    assessed by comparing the mean slopes and 95% confidence intervals around each effect.  We

268    used an ANOVA to compare variability in the model ensemble across temporal scales.  In this

269    ANOVA we compared the absolute deviation of mean climate and $CO_2$ sensitivity for each

270    model to that of the ensemble mean at each scale.

271          We also quantified the effects of slow ecosystem processes among models on NPP

272    sensitivity to climate and $CO_2$ using linear regression.  Here, we considered two categorical

273    characteristics of models as well as two related continuous variables to identify trends in model

274    dynamics and sensitivity to climate and $CO_2$. Categorical model characteristics included

275    vegetation scheme (static or dynamic) and the presence of fire (yes or no) while composition

276    (fraction evergreen) and biomass variability described similar model dynamics as continuous

277    effects. Each characteristic of slow ecosystem processes was correlated with the mean slopes of

278    model sensitivities to temperature, precipitation and $CO_2$.  We performed this correlation

279    analysis at all three temporal scales to determine if the effects of these slow processes were

280    significant at particular, characteristic scales.

281

282    **Results**

283    *Spatial and temporal patterns of NPP in ecosystem models*

284          Terrestrial ecosystem models disagreed about both the spatial patterns of NPP in the

285    Northeastern and Upper Midwestern United States as well as the change in NPP through time

286    (Fig. 1).  Midwestern sites tended towards lower NPP than those further east with ensemble-

287    mean NPP ranging from $3.60 \pm 2.14$ MgC m$^{-2}$ yr$^{-1}$ (mean $\pm$ SD) at Billy's Lake to $6.0 \pm 2.5$ MgC

288    m$^{-2}$ yr$^{-1}$ at Harvard Forest.  ED-LU had the highest mean NPP of the ensemble ($7.12 \pm 2.80$ MgC

289    m$^{-2}$ yr$^{-1}$) and JULES-STATIC had the lowest ($1.51 \pm 1.17$ MgC m$^{-2}$ yr$^{-1}$), but this pattern varied

290    across individual sites and times. Tree-ring estimates of NPP (NPP$_{TR}$) at both Harvard and

291    Howland were much lower than predicted by models (Fig. 1).  From 1980-2010, NPP$_{TR}$ was 1.73

292    $\pm$ 0.27 MgC ha$^{-1}$ yr$^{-1}$ at Harvard and 1.17 $\pm$ 0.13 MgC ha$^{-1}$ yr$^{-1}$ at Howland.  The mean NPP of

293    the model ensemble from 1980-2010 was 7.47 $\pm$ 3.73 MgC ha$^{-1}$ yr$^{-1}$ at Harvard and 7.07 $\pm$ 3.93

294    MgC ha$^{-1}$ yr$^{-1}$ at Howland.  NPP$_{TR}$ was lower than the lowest model NPP at both sites (2.90 $\pm$

295    0.38 MgC ha$^{-1}$ yr$^{-1}$ at Harvard, 3.48 $\pm$ 0.30 at Howland MgC ha$^{-1}$ yr$^{-1}$).

296            To facilitate comparisons of the impacts of climate and $CO_2$ on NPP, through time, we

297    analyzed percent change in NPP relative to each model's mean.  Once relativized, models and

298    tree rings displayed similar levels of variability in NPP and ring width: model NPP variability

299    was 16 $\pm$ 8%, NPP$_{TR}$ was 12%, and RW was 20%.  Even though ED2-LU had the highest NPP

300    variability in absolute terms, with a standard deviation through time of 1.83 MgC m$^{-2}$ yr$^{-1}$, but

301    JULES-STATIC displayed higher temporal NPP variability relative to its mean (29%).  In

302    contrast, SiBCASA showed the least absolute and relative NPP variability through time: 0.28

303    MgC m$^{-2}$ yr$^{-1}$, 5% of its mean.  The higher relative NPP variability of JULES-STATIC was

304    accompanied by the most stable biomass (temporal standard deviations of 3%) while JULES-

305    TRIFFID showed the least temporal variability of composition with only a mean 2% variability

306    through time (Table 2).  LPJ-GUESS, whose mean NPP and NPP variability were close to the

307    ensemble mean, had the highest variability of both composition and biomass (35% and 24%,

308    respectively).

309

310    *NPP and RW sensitivity to climate and CO$_2$*

311            Comparisons between models and tree-ring data at sub-centennial temporal scales reveal

312    large disparities between the effects of climate and $CO_2$ on NPP and tree growth (Fig. 2). At the

313    1980-2010 temporal extent, the ensemble of models and tree-ring data agreed on a positive

314    relationship between NPP or growth and precipitation, but mean model sensitivity (0.09 ± 0.05%

315    mm$^{-1}$ yr$^{-1}$) was nine times higher than both tree-ring NPP and ring widths (both 0.01% mm$^{-1}$ yr$^{-1}$

316    $^{-1}$).  The overall signs of sensitivity were opposite between models and tree-ring data for

317    temperature (model ensemble = -0.99 ± 4.03% $^{\circ}$C$^{-1}$, RW = 11.0% $^{\circ}$C$^{-1}$, NPP$_{TR}$ = 14.2% $^{\circ}$C$^{-1}$) and

318    $CO_2$ (model ensemble = 0.20 ± 0.16% ppm$^{-1}$, RW = -0.7% ppm$^{-1}$, NPP$_{TR}$ = 0.01% ppm$^{-1}$).

319    However, the 95% confidence interval for NPP$_{TR}$ does encompass 0 and is consistent with

320    relatively $CO_2$-insensitive models such as LPJ-WSL. *A priori* detrending of ring width to ring

321    width index dramatically reduced temperature, precipitation, and $CO_2$ sensitivity in the

322    individual-based tree-ring analysis, causing temperature and $CO_2$ sensitivities to be more

323    consistent with model sensitivities, but less consistent with NPP$_{TR}$ (S5).  Trends within and

324    among the sensitivities of models and tree-ring data were similar between analyses including all

325    sites and when analyses were restricted to just Harvard and Howland (S6).  Increasing temporal

326    scale had relatively minor effects on climate sensitivity in model NPP and increased sensitivity

327    in RW (Fig. 2, S7).  This resulted in greater consistency among precipitation responses in RW

328    and models, but continued model-data disagreement in temperature and $CO_2$ at the 1901-2010

329    scale.  Temperature sensitivity of NPP in LINKAGES was a notable exception to this trend, and

330    showed similar positive effects of lower growing season temperatures as the tree-ring datasets,

331    but then diverged and displayed negative effects of warm temperatures.

332         Spatial and temporal variability in NPP among models corresponded to differences

333    among models in emergent ecosystem sensitivities to temperature, precipitation, and $CO_2$. At the

334    full temporal extent of our model simulations (850-2010), models showed consistent positive

335    effects of precipitation and $CO_2$ on NPP, although with varying magnitudes of the effect (Fig. 2).

336    For precipitation, the ensemble of models had an average $0.13 \pm 0.11\%$ increase in NPP per

337    millimeter precipitation, but individual models showed sensitivities ranging from $0.02 \pm 0.03\%$

338    $mm^{-1}$ $yr^{-1}$ in SiBCASA to $0.32 \pm 0.16\%$ $mm^{-1}$ $yr^{-1}$ in JULES-STATIC.   Ensemble NPP

339    sensitivity to $CO_2$ was $0.36 \pm 0.34\%$ $ppm^{-1}$ with individual effects in models with $CO_2$ ranging

340    from $0.06 \pm 0.04\%$ $ppm^{-1}$ in CLM-BGC to $1.03 \pm 0.20\%$ $ppm^{-1}$ in JULES-STATIC.

341    LINKAGES, which lacks $CO_2$ effects in its model structure, supports the ability of the GAM

342    approach to correctly attribute impacts of climate and $CO_2$ on NPP by displaying a small NPP

343    response of $-0.01 \pm 0.00\%$ $ppm^{-1}$.  Temperature was the only effect included in models to show

344    differences in the direction of NPP sensitivity.  Most models had a generally negative

345    temperature effect and the ensemble mean effect was a decrease of $2.37 \pm 9.14\%$ in NPP per ˚C.

346    JULES-STATIC showed the strongest effect with a $23.48 \pm 0.89\%$ decrease per ˚C.  Overall,

347    LINKAGES had a positive temperature effect ($14.77 \pm 19.43\%$ per ˚C), although this effect

348    became negative at higher temperatures.  Over the full course of model simulations, the presence

349    of fire was correlated with decreased sensitivity to temperature (t=-2.3, p=0.03) and $CO_2$ (t=-2.8,

350    p<0.01).  Decreased model NPP sensitivity to temperature was also associated with increased

351    biomass variability through time (t=-2.7, p=0.01).  Models with more variable composition over

352    multiple centuries also tended to be less sensitive to $CO_2$, but this trend was not significant (t=-

353    2.0, p=0.06).

354        An analysis of variance indicated there was greater agreement among model sensitivities

355    to precipitation and $CO_2$ at the two shorter temporal scales than at the multi-centennial (850-

356    2010) extent (Fig. 2; precipitation: t=4.6, p<0.01, $CO_2$: t=4.0, p<0.01), but temperature showed

357    similar ensemble variability across temporal scales (p>0.05).  The presence of fire in ecosystem

358    models correlated with reduced sensitivity to temperature, precipitation, and $CO_2$, but only at the

359    multi-centennial temporal scale (Table 3).  Slow ecosystem processes had weakened effects on

360    model NPP sensitivity to climate and $CO_2$ at the shorter temporal scales of empirical data and

361    were not significantly associated with patterns in model sensitivities at these multi-centennial

362    scales (Table 3).  Similarly, models with high temporal variability of biomass were also less

363    sensitive to temperature, but only over the course of multiple centuries.  Factors associated with

364    changes in composition including composition stability and whether a model had dynamic or

365    static vegetation had no correlation with climate sensitivity in our model ensemble.

366

367    *Drivers of model ensemble patterns agreement through time*

368            Differences in model sensitivity to increasing $CO_2$ explained increasing model ensemble

369    uncertainty in change in NPP since 1900 (Fig. 3). Model ensemble variability of normalized NPP

370    prior to 1900 was $11.0 \pm 5.3\%$ (Fig. 3b).  After 1900, this variability more than doubled to $20.6 \pm$

371    $7.4\%$.  This shift in model agreement occurred as NPP in most models shifted from being

372    primarily influenced by precipitation to showing varying degrees of $CO_2$ enhancement (Fig. 3b,

373    c).  In the 1980-2010 extent, the effect of $CO_2$ on NPP in models that include dynamic $CO_2$

374    ranged from +2% to +110% with model variability in $CO_2$ response proportionate to that of

375    inter-model NPP deviation during this time period ($CO_2$ variability = $29 \pm 3\%$; NPP variability =

376    $28 \pm 5\%$, Fig. 3c).  LINKAGES was the only model to not include a $CO_2$ driver and our

377    statistical method misattributed a slight 0.6% decrease in NPP to increasing $CO_2$.

378

379    **Discussion**

380            Analysis of emergent responses to temperature, precipitation, and $CO_2$ at multi-decadal

381    and multi-centennial time scales reveals many inconsistencies among models and data. The ten

382    terrestrial ecosystem models included in our study showed variability in both mean NPP across

383    space and relative changes in NPP through time (Fig. 1).  Similar discrepancies have been widely

384    observed in other multi-model comparisons, which have attributed differences among models to

385    numerous causes including ecosystem feedbacks and uncertainties in model processes and

386    parameterization (Piao *et al.*, 2013; Friedlingstein *et al.*, 2014; Walker *et al.*, 2014).  We found

387    widely varying patterns of climate and $CO_2$ effects among models and data across multiple

388    temporal scales (Fig. 2).  However, the ensemble of models displayed less variation in climate

389    and $CO_2$ responses at short temporal scales.  Small differences in model structure and

390    parameterization compound over time and drive model divergence, skewing model projections at

391    ecology- and policy-relevant timescales.  Fire and changes in biomass are only associated with

392    differences in precipitation and $CO_2$ sensitivity among models at the multi-centennial scale,

393    indicating that feedbacks involving disturbance and biomass only have discernible impacts on

394    climate and $CO_2$ over multiple forest generations (Table 3).  Although model responses to

395    precipitation are the most variable, responses to $CO_2$ are the greatest source of ensemble

396    divergence of the past 100 years.

397         Models and empirical datasets have widely varying estimates of baseline NPP, so

398    standardization is important to compare the impacts of climate and $CO_2$ through time across

399    datasets. To be consistent across all models in the ensemble, we analyzed total ecosystem NPP,

400    which includes the total carbon allocated to both above- and belowground tissues.  Most models

401    in our ensemble simulated NPP values within the range of what has been observed for forests

402    common in the Northeastern and Upper Midwestern United states.  Representative field-based

403    estimates of NPP for our study region range from 1.3 MgC ha$^{-1}$ yr$^{-1}$ in mature boreal forests to as

404    high as 8.7 MgC ha$^{-1}$ yr$^{-1}$ in the oak-hickory forests that dominate further south when a 50%

405    carbon content of biomass is assumed (Jenkins *et al.*, 2001; Luyssaert *et al.*, 2007; Goulden *et*

406    *al.*, 2011). Even if the models were perfectly accurate, their values would still be greater than

407    the aboveground-only NPP quantified through tree rings in our study, which is estimated to be

408    between 45 and 65% of the total carbon. Furthermore, our tree-ring productivity estimates only

409    include the aboveground biomass increment from one year to the next, which will not capture

410    any biomass that is produced and turns over within a year. A recent comparison of tree-ring

411    estimates of NPP at Harvard Forest indicated good agreement with repeat forest censuses (Eisen

412    & Plotkin, 2015; Dye *et al.*, 2016), but are substantially lower than values estimated through

413    remote sensing for coniferous and deciduous forests at the same site (5.5 and 6.8 MgC ha$^{-1}$ yr$^{-1}$,

414    respectively; Turner *et al.*, 2005). In order to reconcile total and aboveground NPP estimates,

415    better quantification of amounts and controls of allocation and turnover among tissues is

416    essential (Jenkins *et al.*, 2001; Litton *et al.*, 2007; Luyssaert *et al.*, 2007). Some empirical studies

417    suggest that the fraction of NPP allocated to aboveground wood could vary from year to year due

418    to tree maturation, increasing $CO_2$, or interannual climate variability (DeLucia *et al.*, 2005;

419    Norby *et al.*, 2005; Doughty *et al.*, 2014). Although some ecosystem models allocate carbon in

420    response to changing resource limitations (De Kauwe *et al.*, 2014), most do not and the models

421    used in this analysis are based on simple allometric rules. Despite these inconsistencies in NPP

422    values among models and empirical datasets, it is informative to compare the emergent

423    sensitivities of productivity to climate and $CO_2$ across datasets.

424            Emergent model NPP sensitivities to temperature and precipitation are not supported by

425    the responses observed in plot or individual tree-ring records. Specifically, most models were

426    over-sensitive to precipitation relative to tree rings and had predominantly negative or no

427    response to temperature whereas both tree-ring NPP and raw rings widths showed consistently

428   positive responses (Fig. 2).  The strongly negative effect of temperature coupled with a strongly

429   positive effect of precipitation in the models suggests that the ecosystem models in our ensemble

430   may be over-sensitive to both temperature- and precipitation-driven droughts despite rapid post-

431   drought recovery seen in many ecosystem models (Anderegg *et al.*, 2015).  Although the tree

432   rings in our study show positive correlations with temperature in both raw ring width and *a*

433   *priori* detrended ring width index, other studies provide support for negative effects of high

434   temperatures on tree growth in the eastern United States (e.g. Rollinson *et al.*, 2016).

435   Nonetheless, models appear to be lacking moderating feedbacks that can cause positive

436   temperature effects at low temperatures or reduce precipitation sensitivity.  Indeed, the model in

437   our ensemble that most closely matches the empirical relationship between temperature and NPP

438   is LINKAGES, which calculates NPP directly from growing degree days rather than from

439   separate GPP and autotrophic respiration functions (Post & Pastor, 1996).

440          Models were also over-sensitive to precipitation relative to tree rings.  Traditional tree-

441   ring analyses have focused on growth responses to drought indices such as the Palmer Drought

442   Severity Index that combines temperature, precipitation, and soil moisture into a single drought

443   metric (e.g. Speer *et al.*, 2009; McEwan *et al.*, 2011).  However, process-based ecosystem

444   models drive productivity from independent forcing of temperature and precipitation and

445   temperature- versus precipitation-based droughts may impact ecosystems differently (Anderegg

446   *et al.*, 2013).  Furthermore, models can vary in their representation of soil and hydrology so that

447   the same temperature and precipitation drivers may result in different droughts both among

448   models and compared to empirical system (Cook *et al.*, 2015).  Tree-ring studies that have used

449   mixed-modeling frameworks to look at the simultaneous influences of temperature and

450   precipitation on growth have revealed stronger influences of temperature on growth than

451    precipitation (Rollinson *et al.*, 2016).  In order to improve the ecosystem-scale model responses

452    to both temperature and precipitation, further empirical research is needed to separate

453    temperature and precipitation effects on tree growth.

454          Model responses to precipitation were more variable than temperature or CO2 and

455    consequently there was little consensus in our model ensemble about when and where ecosystem

456    productivity was driven by temperature versus precipitation (Figs. 2 & 3). Despite sensitivity

457    disparities with tree-ring data, the strong control of precipitation in many models at the multi-

458    centennial scale supports hypotheses that drought may have been instrumental in driving

459    ecosystem dynamics and species shifts in New England 500-600 years before present (Clifford &

460    Booth, 2015). However, the relative insensitivity of productivity to temperature is difficult to

461    reconcile with apparent temperature-driven millennial-scale compositional shifts in the

462    paleoecological record (Prentice *et al.*, 1991; Blois *et al.*, 2013).  In the model ensemble, fire and

463    changes in biomass variability reduce model sensitivity to temperature rather than changes in

464    relative composition, which suggests this lack of temperature sensitivity is not due to stabilizing

465    shifts in plant functional types.  Challenges of recreating spatial and temporal patterns of

466    observed composition is a pervasive problem in ecosystem modeling and is known to have

467    cascading impacts that bias other aspects of ecosystem dynamics including NPP and

468    transpiration (Matthes *et al.*, 2016).  Improved plant functional type parameterization within and

469    across models through careful data collection will undoubtedly help address some of the

470    discrepancies in magnitude and relative importance of temperature and precipitation seen

471    between models and data in our study.

472          Differences in model responses to increasing $CO_2$ drive model ensemble uncertainty in

473    NPP since 1900 (Fig. 3).  The discrepancy of $CO_2$ enhancement effects in the model ensemble

474    reflects the long-running debate in empirical studies about whether or not increased $CO_2$ has

475    sustained effects on tree growth and ecosystem productivity. One synthesis from Free-Air $CO_2$

476    Enrichment (FACE) studies, where conditions were raised to ~170 ppm above modern, increased

477    productivity by 23%, or 13% $ppm^{-1}$ (Norby *et al.*, 2005). This is slightly less than half of the

478    43% increase in NPP attributed to $CO_2$ in our model ensemble over a similar increase in $CO_2$

479    over the past 100 years (Fig. 3).  However, a 30-year study that found 12% increased stem width

480    from increased $CO_2$ attributed most of this difference to early growth increases that diminished

481    as the trees matured (Hättenschwiler *et al.*, 1997).  The effect of increasing competition for

482    resources such as nitrogen and light during forest maturation has been used to explain a lack of

483    $CO_2$ enhancement on tree growth in older forests quantified through tree rings (van der Sleen *et*

484    *al.*, 2014; Fernández-de-Uña *et al.*, 2016).  In tree rings, rising $CO_2$ has been associated with

485    increased water use efficiency, but there has been little evidence that this translates into

486    increased stem growth (Gedalof & Berg, 2010; Andreu-Hayles *et al.*, 2011; Peñuelas *et al.*,

487    2011; van der Sleen *et al.*, 2014; Frank *et al.*, 2015).

488           Many of the explanations for a lack of $CO_2$-induced growth increases in tree-ring widths

489    center around the challenges of separating out the effects of increasing $CO_2$ from increasing tree

490    size and stand-level forest dynamics (Jacoby & D'Arrigo, 1997; Andreu-Hayles *et al.*, 2011). In

491    our results, stand-level NPP from tree rings showed no $CO_2$ effect, but individual ring widths

492    displayed a negative effect that arises from the aforementioned challenges even though we

493    explicitly accounted for increasing individual size in our analyses (Fig. 2, S5). This model- and

494    data-based uncertainty in the long-term effect of increasing $CO_2$ on ecosystems is the greatest

495    source of uncertainty in current and future forest carbon cycle modeling (Sitch *et al.*, 2008;

496    IPCC, 2013).  Both fire and changes in biomass are correlated with $CO_2$ sensitivity in our

497     ensemble of models across temporal scales (Table 3), indicating that disturbance and recovery

498     processes may have strong influence on how $CO_2$ impacts ecosystem NPP.  Consequently,

499     accurate representation of causes of forest disturbances and process of recovery at local scales is

500     necessary to reconcile models with empirical data and to make improved predictions of future

501     forest dynamics under continued global change (Seidl *et al.*, 2011; Dietze *et al.*, 2014; Steinkamp

502     & Hickler, 2015).

503

504     **Conclusions**

505         Models and data display conflicting responses of NPP to climate and $CO_2$ and these

506     differences are more pronounced at the multi-centennial time scales of past climate-driven

507     ecosystem change.  Models are generally over-sensitive to growing season precipitation relative

508     to tree-ring datasets and display negative responses to temperature while tree rings show

509     consistent positive effects.  Further research on independent temperature and precipitation effects

510     is necessary to diagnose systematic weaknesses in ecosystem models. Similarly, differences in

511     model responses to $CO_2$ drive divergence of model ensemble NPP over the past century as the

512     effects compound through time while $CO_2$ concentration consistently increases.  Despite

513     empirical evidence for positive $CO_2$ effects on individual processes included in ecosystem

514     models, the net effects on tree growth and forest-level NPP remain unclear and poorly

515     constrained.  Differences in disturbance as measured through the presence of fire in model

516     simulations only impacts the sensitivity of productivity to climate and $CO_2$ at multi-centennial

517     times scales that capture multiple tree generations. The paths toward model improvement and

518     reconciling discrepancies with data will be model-specific as each model has unique structural

519     and parameterization requirements.  However, accurate representation of fire and other

520  disturbance processes and ecosystem recovery in models will require improved synthesis of

521  short-term ecophysiological processes with gradual shifts in forest composition and structure

522  seen in historical and paleoecological records.  We advocate that this process of refinement

523  include consideration and formal evaluation of emergent, ecosystem-level behavior at the multi-

524  decadal and multi-centennial temporal scales at which changes in forest composition and

525  biomass occur.  This approach would better align model evaluation with the organizational and

526  temporal scales of forest dynamics and policy decisions.

527

541

542  **References**

543  Anderegg LDL, Anderegg WRL, Berry JA (2013) Not all droughts are created equal: Translating

544      meteorological drought into woody plant mortality. *Tree Physiology*, **33**, 701–712.

545  Anderegg WRL, Schwalm C, Biondi F et al. (2015) Pervasive drought legacies in forest

546      ecosystems and their implications for carbon cycle models. *Science*, **349**, 528–532.

547  Andreu-Hayles L, Planells O, Gutiérrez E, Muntan E, Helle G, Anchukaitis KJ, Schleser GH

548      (2011) Long tree-ring chronologies reveal 20th century increases in water-use efficiency but

549      no enhancement of tree growth at five Iberian pine forests. *Global Change Biology*, **17**,

550      2095–2112.

551  Babst F, Poulter B, Trouet V et al. (2013) Site- and species-specific responses of forest growth to

552      climate across the European continent. *Global Ecology and Biogeography*, **22**, 706–717.

553  Best MJ, Pryor M, Clark DB et al. (2011) The Joint UK Land Environment Simulator (JULES),

554      model description. Part 1: Energy and water fluxes. *Geoscientific Model Development*, **4**,

555      677–699.

556  Blois JL, Williams JW, Fitzpatrick MC et al. (2013) Modeling the climatic drivers of spatial

557      patterns in vegetation composition since the Last Glacial Maximum. *Ecography*, **36**, 460–

558      473.

559  Boisvenue C, Running SW (2006) Impacts of climate change on natural forest productivity -

560      evidence since the middle of the 20th century. *Global Change Biology*, **12**, 862–882.

561  Burns RM, Honkala BH (1990) Silvics of North America: 1. Conifers; 2. Hardwoods.

562      *Agricultural Handbook*, **654**.

563  Charney ND, Babst F, Poulter B et al. (2016) Observed forest sensitivity to climate implies large

564      changes in 21st century North American forest growth. *Ecology Letters*, **19**, 1119–1128.

565  Clark D a, Brown S, Kicklighter DW, Chambers JQ, Thomlinson JR, Ni J (2001) Measuring net

566      primary production in forest concepts and field methods. *Ecological Applications*, **11**, 356–

567      370.

568   Clark DB, Mercado LM, Sitch S et al. (2011) The Joint UK Land Environment Simulator

569      (JULES), model description. Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific*

570      *Model Development*, **4**, 701–722.

571   Clifford MJ, Booth RK (2015) Late-Holocene drought and fire drove a widespread change in

572      forest community composition in eastern North America. *The Holocene*, **25**, 1102–1110.

573   Cook BI, Ault TR, Smerdon JE (2015) Unprecedented 21st century drought risk in the American

574      Southwest and Central Plains. *Science Advances*, **1**, 1–7.

575   Crucifix M, Harrison S, Brierley C (2012) Recent and deep pasts in paleoclimate model

576      intercomparison project. *Eos, Transactions American Geophysical Union*, **93**, 539–539.

577   D'Orangeville L, Duchesne L, Houle D, Kneeshaw D, Côté B, Pederson N (2016) Northeastern

578      North America as a potential refugium for boreal forests in a warming climate. *Science*,

579      **352**, 1452–1455.

580   Davis MB, Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change.

581      *Science (New York, N.Y.)*, **292**, 673–9.

582   Davis SC, Hessl AE, Scott CJ, Adams MB, Thomas RB (2009) Forest carbon sequestration

583      changes in response to timber harvest. *Forest Ecology and Management*, **258**, 2101–2109.

584   DeLucia EH, Moore DJ, Norby RJ (2005) Contrasting responses of forest ecosystems to rising

585      atmospheric CO2: Implications for the global C cycle. *Global Biogeochemical Cycles*, **19**,

586      1–9.

587   Dietze M, Serbin S, Davidson C et al. (2014) A quantitative assessment of a terrestrial biosphere

588      model's data needs across North American biomes. *Journal of Geophysical Research:*

589          *Biogeosciences*, **119**, 286–300.

590     Doughty CE, Malhi Y, Araujo-Murakami A et al. (2014) Allocation trade-offs dominate the

591          response of tropical forest growth to seasonal and interannual drought. *Ecology*, **95**, 2192–

592          2201.

593     Dye A, Barker-Plotkin A, Bishop D, Pederson N, Poulter B, Hessl A (2016) Comparing tree-ring

594          and permanent plot estimates of aboveground net primary production in three Eastern U.S.

595          forests. *Ecosphere*, in press.

596     Eisen K, Plotkin AB (2015) Forty years of forest measurements support steadily increasing

597          aboveground biomass in a maturing , Quercus -dominant northeastern forest. *Journal of the*

598          *Torrey Botanical Society*, **142**, 97–112.

599     Fernández-de-Uña L, McDowell NG, Cañellas I, Gea-Izquierdo G (2016) Disentangling the

600          effect of competition, CO2 and climate on intrinsic water-use efficiency and tree growth.

601          *Journal of Ecology*, n/a-n/a.

602     Fisichelli NA, Frelich LE, Reich PB (2013) Temperate tree expansion into adjacent boreal forest

603          patches facilitated by warmer temperatures. *Ecography*, **36**, 001–010.

604     Foster DR, Motzkin G, Slater B (1998) Land-Use History as Long-Term Broad-Scale

605          Disturbance: Regional Forest Dynamics in Central New England. *Ecosystems*, **1**, 96–119.

606     Foster JR, D'Amato AW, Bradford JB (2014) Looking for age-related growth decline in natural

607          forests: Unexpected biomass patterns from tree rings and simulated mortality. *Oecologia*,

608          **175**, 363–374.

609     Frank DC, Poulter B, Saurer M et al. (2015) Water-use efficiency and transpiration across

610          European forests during the Anthropocene. *Nature Climate Change*, **5**, 579–583.

611     Friedlingstein P, Meinshausen M, Arora VK, Jones CD, Anav A, Liddicoat SK, Knutti R (2014)

612     Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of*

613     *Climate*, **27**, 511–526.

614  Gedalof Z, Berg AA (2010) Tree ring evidence for limited direct CO2 fertilization of forests over

615     the 20th century. *Global Biogeochemical Cycles*, **24**, 2–7.

616  Gerten D, Schaphoff S, Haberlandt U, Lucht W, Sitch S (2004) Terrestrial vegetation and water

617     balance - Hydrological evaluation of a dynamic global vegetation model. *Journal of*

618     *Hydrology*, **286**, 249–270.

619  Gómez-Aparicio L, García-Valdés R, Ruíz-Benito P, Zavala MA (2011) Disentangling the

620     relative importance of climate, size and competition on tree growth in Iberian forests:

621     implications for forest management under global change. *Global Change Biology*, **17**,

622     2400–2414.

623  Goulden ML, Mcmillan AMS, Winston GC, Rocha A V, Manies KL, Harden JW, Bond-

624     Lamberty BP (2011) Patterns of NPP, GPP, respiration, and NEP during boreal forest

625     succession. *Global Change Biology*, **17**, 855–871.

626  Graumlich LJ, Brubaker LB, Grier CC (1989) Long-term trends in forest net primary

627     productivity: Casade Mountains, Washington. *Ecology*, **70**, 405–410.

628  Grissino-Mayer HD, Fritts HC (1997) The International Tree-Ring Data Bank: an enhanced

629     global database serving the global scientific community. *The Holocene*, **7**, 235–238.

630  Hättenschwiler S, Miglietta F, Raschi A, Körner C (1997) Thirty years of in situ tree growth

631     under elevated CO2: a model for future forest responses? *Global Change Biology*, **3**, 463–

632     471.

633  Hempel S, Frieler K, Warszawski L, Schewe J, Piontek F (2013) A trend-preserving bias

634     correction &ndash; The ISI-MIP approach. *Earth System Dynamics*, **4**, 219–236.

635    Hollister RD, Webber PJ, Tweedie CE (2005) The response of Alaskan arctic tundra to

636        experimental warming: differences between short- and long-term responses. *Global Change*

637        *Biology*, **11**, 525–536.

638    IPCC (2013) Summary for Policy Makers. In: *Climate Change 2013: The Physical Science*

639        *Basis. Contribution of Working Group I to the Fifth Assessment Report of the*

640        *Intergovernmental Panel on Climate Change* (eds Stocker TF, Qin D, Plattner G-K, Tignor

641        M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM). Camberidge University

642        Press, Cambridge, United Kingdom and New York, NY, USA.

643    Jacoby GC, D'Arrigo RD (1997) Tree rings, carbon dioxide, and climatic change. *Proceedings*

644        *of the National Academy of Sciences of the United States of America*, **94**, 8350–3.

645    Jenkins JC, Birdsey RA, Pan Y (2001) Biomass and NPP estimation for the mid-Atlantic region

646        (USA) using plot-level forest inventory data. *Ecological Applications*, **11**, 1174–1193.

647    De Kauwe MG, Medlyn BE, Zaehle S et al. (2013) Forest water use and water use efficiency at

648        elevated CO2: a model-data intercomparison at two contrasting temperate forest FACE

649        sites. *Global Change Biology*, **19**, 1759–1779.

650    De Kauwe MG, Medlyn BE, Zaehle S et al. (2014) Where does the carbon go? A model-data

651        intercomparison of vegetation carbon allocation and turnover processes at two temperate

652        forest free-air CO2 enrichment sites. *New Phytologist*, **203**, 883–899.

653    Keeling CD, Whorf TP, (Carbon Dioxide Research Group) (2005) Atmospheric CO2

654        concentrations (ppmv) derived from in situ air samples collected at Mauna Loa

655        Observatory, Hawaii. *Scripps Institude of Oceanography*.

656    Kumar J, Brooks B-GJ, Thornton PE, Dietze MC (2012) Sub-daily Statistical Downscaling of

657        Meteorological Variables Using Neural Networks. *Procedia Computer Science*, **9**, 887–896.

658    Litton CM, Raich JW, Ryan MG (2007) Carbon allocation in forest ecosystems. *Global Change*

659          *Biology*, **13**, 2089–2109.

660    Luyssaert S, Inglima I, Jung M et al. (2007) CO2 balance of boreal, temperate, and tropical

661          forests derived from a global database. *Global Change Biology*, **13**, 2509–2537.

662    Matesanz S, Brooker RW, Valladares F, Klotz S (2009) Temporal dynamics of marginal steppic

663          vegetation over a 26-year period of substantial environmental change. *Journal of Vegetation*

664          *Science*, **20**, 299–310.

665    Matthes JH, Goring S, Williams JW, Dietze MC (2016) Benchmarking historical CMIP5 land-

666          climate feedbacks across the Upper Midwest and Northeastern United States. *Journal of*

667          *Geophysical Research: Biogeosciences*.

668    McEwan RW, Dyer JM, Pederson N (2011) Multiple interacting ecosystem drivers: toward an

669          encompassing hypothesis of oak forest dynamics across eastern North America. *Ecography*,

670          **34**, 244–256.

671    Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR (2009) Mechanistic scaling of

672          ecosystem function and dynamics in space and time: Ecosystem Demography model

673          version 2. *Journal of Geophysical Research*, **114**, 1–21.

674    Melillo JM, Steudler PA, Aber JD et al. (2002) Soil Warming Feedbacks the Carbon-Cycle

675          Climate. *Science*, **298**, 2173–2176.

676    Morin X, Lechowicz MJ, Augspurger C, O'Keefe J, Viner D, Chuine I (2009) Leaf phenology in

677          22 North American tree species during the 21st century. *Global Change Biology*, **15**, 961–

678          975.

679    Norby RJ, Delucia EH, Gielen B et al. (2005) Forest response to elevated CO2 is conserved

680    across a broad range of productivity. *Proceedings of the National Academy of Sciences of*

681    *the United States of America*, **102**, 18052–6.

682  Norby RJ, Warren JM, Iversen CM, Medlyn BE, McMurtrie RE (2010) CO2 enhancement of

683    forest productivity constrained by limited nitrogen availability. *Proceedings of the National*

684    *Academy of Sciences of the United States of America*, **107**, 19368–73.

685  Oleson KW, Lawrence DM, Gordon B et al. (2010) Technical description of version 4.0 of the

686    Community Land Model (CLM). *NCAR/TN-478+STR NCAR Technical Note*, 266.

687  Peñuelas J, Canadell JG, Ogaya R (2011) Increased water-use efficiency during the 20th century

688    did not translate into enhanced tree growth. *Global Ecology and Biogeography*, **20**, 597–

689    608.

690  Peñuelas J, Sardans J, Estiarte M et al. (2013) Evidence of current impact of climate change on

691    life: a walk from genes to the biosphere. *Global change biology*, **19**, 2303–38.

692  Piao S, Sitch S, Ciais P et al. (2013) Evaluation of terrestrial carbon cycle models for their

693    response to climate variability and to CO2 trends. *Global Change Biology*, **19**, 2117–2132.

694  Post WM, Pastor J (1996) Linkages - an individual-based forest ecosystem model. *Climatic*

695    *Change*, **34**, 253–261.

696  Prentice IC, Bartlein PJ, Webb T (1991) Vegetation and climate change in eastern North

697    America since the last glacial maximum. *Ecology*, **72**, 2038–2056.

698  Ramankutty N, Foley JA (1999) Estimating historical changes in global land cover: croplands

699    from 1700 to 1992. *Global Biogeochemical Cycles*, **13**, 997–1027.

700  Renwick KM, Rocca ME (2014) Temporal context affects the observed rate of climate-driven

701    range shifts in tree species. *Global Ecology and Biogeography*, **24**, 44–51.

702  Rollinson CR, Kaye MW, Canham CD (2016) Interspecific variation in growth responses to

703         climate and competition of five eastern tree species. *Ecology*, **97**, 1003–1011.

704    Schaefer K, Collatz GJ, Tans P et al. (2008) Combined Simple Biosphere/Carnegie-Ames-

705         Stanford Approach terrestrial carbon cycle model. *Journal of Geophysical Research*, **113**,

706         G03034.

707    Seidl R, Fernandes PM, Fonseca TF et al. (2011) Modelling natural disturbances in forest

708         ecosystems: a review. *Ecological Modelling*, **222**, 903–924.

709    Shuman B, Webb T, Bartlein P, Williams JW (2002) The anatomy of a climatic oscillation:

710         Vegetation change in eastern North America during the Younger Dryas chronozone.

711         *Quaternary Science Reviews*, **21**, 1777–1791.

712    Sitch S, Smith B, Prentice IC et al. (2003) Evaluation of ecosystem dynamics, plant geography

713         and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change*

714         *Biology*, **9**, 161–185.

715    Sitch S, Huntingford C, Gedney N et al. (2008) Evaluation of the terrestrial carbon cycle, future

716         plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation

717         Models (DGVMs). *Global Change Biology*, **14**, 2015–2039.

718    van der Sleen P, Groenendijk P, Vlam M et al. (2014) No growth stimulation of tropical trees by

719         150 years of $CO_2$ fertilization but water-use efficiency increased. *Nature Geoscience*, **8**,

720         24–28.

721    Smith B, Warlind D, Arneth A, Hickler T, Leadley P, Siltberg J, Zaehle S (2014) Implications of

722         incorporating N cycling and N limitations on primary production in an individual-based

723         dynamic vegetation model. *Biogeosciences*, **11**, 2027–2054.

724    Speer JH, Grissino-Mayer HD, Orvis KH, Greenberg CH (2009) Climate response of five oak

725         species in the eastern deciduous forest of the southern Appalachian Mountains, USA.

726      *Canadian Journal of Forest Research*, **39**, 507–518.

727   Steinkamp J, Hickler T (2015) Is drought-induced forest dieback globally increasing? *Journal of*

728      *Ecology*, **103**, 31–43.

729   Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design.

730      *Bulletin of the American Meteorological Society*, **93**, 485–498.

731   Turner MG, Collins SL, Lugo AL, Magnuson JJ, Rupp TS, Swanson FJ (2003) Disturbance

732      Dynamics and Ecological Response: The Contribution of Long-Term Ecological Research.

733      *BioScience*, **53**, 46–56.

734   Turner DP, Ritts WD, Cohen WB et al. (2005) Site-level evaluation of satellite-based global

735      terrestrial gross primary production and net primary production monitoring. *Global Change*

736      *Biology*, **11**, 666–684.

737   Walker AP, Hanson PJ, De Kauwe MG et al. (2014) Comprehensive ecosystem model-data

738      synthesis using multiple data sets at two temperate forest free-air CO2 enrichment

739      experiments: Model performance at ambient CO2 concentration. *Journal of Geophysical*

740      *Research: Biogeosciences*, **119**, 937–964.

741   Walker AP, Zaehle S, Medlyn BE et al. (2015) Predicting long-term carbon sequestration in

742      response to CO2 enrichment: How and why do current ecosystem models differ? *Global*

743      *Biogeochemical Cycles*, **29**, 476–495.

744   Wei Y, Liu S, Huntzinger DN et al. (2014) NACP MsTMIP: Global and North American Driver

745      Data for Multi-Model Intercomparison. *Data Set*.

746   Williams AP, Allen CD, Macalady AK et al. (2013) Temperature as a potent driver of regional

747      forest drought stress and tree mortality. *Nature Climate Change*, **3**, 292–297.

748    Wood S (2012) mgcv: GAMs with GCV/AIC/REML smoothness estimation and GAMMs by

749        PQL.

750

751

**Supporting Information Captions**

*Supporting Information 1:* Growing season temperature and precipitation time series for 850-2010 A.D. for all six sites that were used as model drivers and in for sensitivity analysis.

*Supporting Information 2:* Model simulation protocol.

*Supporting Information 3:* Location and record information for tree-ring width information from the International Tree Ring Databank (Grissino-Mayer & Fritts, 1997).

*Supporting Information 4:* Mean and 95% confidence interval of spine-detrended ring width index (RWI) for all tree ring records used in analyses.

*Supporting Information 5:* Sensitivity of relativized NPP and tree-ring width index (RWI) to growing season temperature, growing season precipitation, and $CO_2$ across three temporal scales.

*Supporting Information 6:* Sensitivity of relativized NPP and tree-ring width to growing season temperature, growing season precipitation, and $CO_2$ across three temporal scales restricted to Harvard and Howland sites.

*Supporting Information 7:* Relativized sensitivity of model NPP or ring width to growing season (May – September) temperature, growing season precipitation, and $CO_2$ by temporal extent.

**Table 1:** Location, modern forest type, and mean model driver growing season (May through September) temperature and precipitation for the six modeling locations and which data types were available at each site for sensitivity analysis. Temperature and precipitation values presented are mean ± standard deviation from 850-2010 A.D.  Data type codes are as follows: M = ecosystem models, RW = raw tree-ring widths, $NPP_{TR}$ = tree-ring NPP reconstruction.

| Name | Longitude | Latitude | Forest Type | Temp (˚C) | Precip (mm) | Data Types |
|---|---|---|---|---|---|---|
| Demming Lake | -95.17 | 47.17 | Mixed | 15.6 ± 1.0 | 375 ± 74 | M, RW |
| Billy's Lake | -94.58 | 46.28 | Mixed | 16.4 ± 1.0 | 398 ± 81 | M |
| UNDERC | -89.53 | 46.22 | Mixed | 14.3 ± 0.9 | 411 ± 74 | M, RW |
| Minden Bog | -82.83 | 43.61 | Evergreen | 16.4 ± 0.8 | 375 ± 65 | M |
| Harvard Forest | -72.18 | 42.54 | Deciduous | 15.6 ± 0.7 | 520 ± 86 | M, RW, $NPP_{TR}$ |
| Howland Forest | -68.73 | 45.25 | Mixed | 13.9 ± 0.8 | 492 ± 80 | M, RW, $NPP_{TR}$ |

**Table 2:** List of models and key model characteristics used in analyses. Static vegetation models were prescribed site composition based on modern forest type information (Table 1). However, because composition was assessed as fraction biomass or leaf area (LPJ & JULES models), relative composition could shift through time. Composition and biomass variability was quantified as the standard deviation of the relative fraction evergreen or biomass through time over the full modeling temporal extent. For fire occurrence, "No" indicates no fire occurred in the model simulations at any locations. Version numbers are not tracked in SiBCASA.

| Model | Vers. | Vegetation Scheme | Composition Variability (%) | Fire Occurrence | Biomass Variability (%) | Citation |
|---|---|---|---|---|---|---|
| CLM-BGC | 4.5 | Static | 32% | Yes | 22% | Oleson *et al.*, 2010 |
| CLM-CN | 4.5 | Static | 29% | Yes | 17% | Oleson *et al.*, 2010 |
| ED2 | 2.1 | Dynamic | 20% | No | 20% | Medvigy *et al.*, 2009 |
| ED2-LU | 2.1 | Dynamic | 21% | No | 19% | Medvigy *et al.*, 2009 |
| JULES-STATIC | 4.1 | Static | 25% | No | 3% | Best *et al.*, 2011; Clark *et al.*, 2011 |
| JULES-TRIFFID | 4.1 | Dynamic | 02% | No | 6% | Best *et al.*, 2011; Clark *et al.*, 2011 |
| LINKAGES | 1 .0 | Dynamic | 22% | No | 19% | Post & Pastor, 1996 |
| LPJ-GUESS | 3.1 | Dynamic | 35% | Yes | 24% | Sitch *et al.*, 2003; Gerten *et al.*, 2004; Smith *et al.*, 2014 |
| LPJ-WSL | 1.0 | Dynamic | 28% | Yes | 12% | Sitch *et al.*, 2003; Gerten *et al.*, 2004 |
| SiBCASA | | Static | 00% | No | 4% | Schaefer *et al.*, 2008 |

**Table 3:** ANOVA effect sizes of changes in climate and $CO_2$ sensitivity for four key characteristics of ecosystem models and ecosystem dynamics. Effects are expressed as mean change in normalized NPP (%) per unit climate effect ± standard error. Static vegetation and fire effects are relative to dynamic vegetation scheme and absence of fire, respectively. * and bold indicate significance at $p < 0.05$.

| Effect | Character | 1980-2010 | 1901-2010 | 850-2010 |
|---|---|---|---|---|
| Temperature | Static Vegetation | $3.8 \pm 3.68$ | $4.13 \pm 3.68$ | $4.17 \pm 3.68$ |
| | Composition Var. | $12.67 \pm 14.92$ | $7.67 \pm 17.3$ | $-8.16 \pm 20.47$ |
| | Fire Occurs | $0.99 \pm 3.31$ | $-1.4 \pm 3.31$ | **$-8.18 \pm 3.31$** * |
| | Biomass Var. | $-0.67 \pm 7.2$ | $-9.89 \pm 8.82$ | **$-20.03 \pm 9.15$** * |
| Precipitation | Static Vegetation | $-0.01 \pm 0.2$ | $0.15 \pm 0.2$ | $0.06 \pm 0.2$ |
| | Composition Var. | $0.03 \pm 0.2$ | $-0.06 \pm 0.23$ | $-0.25 \pm 0.27$ |
| | Fire Occurs | $-0.01 \pm 0.04$ | $-0.02 \pm 0.04$ | **$-0.11 \pm 0.04$** * |
| | Biomass Var. | $0.06 \pm 0.1$ | $0.01 \pm 0.12$ | $-0.2 \pm 0.13$ |
| $CO_2$ | Static Vegetation | $-0.01 \pm 0.2$ | $0.15 \pm 0.2$ | $0.06 \pm 0.2$ |
| | Composition Var. | $-0.34 +/- 0.7$ | $-0.33 +/- 0.81$ | $-1.81 \pm 0.96$ |
| | Fire Occurs | $-0.17 +/- 0.15$ | $-0.15 +/- 0.15$ | **$-0.45 \pm 0.15$** * |
| | Biomass Var. | $0.09 +/- 0.39$ | $-0.13 +/- 0.48$ | $-0.61 \pm 0.5$ |

**Figure Captions**

**Figure 1:** Net primary production (NPP) across all sites for the full modeling temporal extent (850-2010) for all ten ecosystem models. Sites are as follows: a) Demming Lake, b) Billy's Lake, c) UNDERC, d) Minden Bog, e) Harvard Forest, f) Howland Forest. Lines indicate the 10-year running means.  Black dots at Harvard and Howland indicate tree-ring estimates of NPP.

**Figure 2:** Normalized sensitivity of NPP and tree ring width to growing season temperature, growing season precipitation, and $CO_2$ across three temporal scales.  Climate and $CO_2$ effects have been normalized to the mean NPP or ring width for each model to facilitate comparison across models and data of different scales.  Colored lines indicate the mean estimated response and color shaded area indicate a 95% confidence interval around that response.  Gray shaded regions indicate the ranges of climate or $CO_2$ not observed in the models at that temporal scale and are presented only for aiding visualization and as such differences among models in these regions should not be interpreted.

**Figure 3:** *a)* 10-year running mean net primary productivity for ten terrestrial ecosystem models (lines) with 95% confidence intervals from among-site variability (shaded areas).  *b)* 10-year running mean of the ensemble mean of normalized model NPP (colored line) with shaded 95% confidence interval (from spread among models) showing relative change in NPP through time in the model ensemble.  Color of each line indicate the relative control of temperature (red), precipitation (blue) and $CO_2$ (green) on change in relative NPP, dashed line indicates the normalized mean of 100%.  *c)* Mean effect of temperature (red), precipitation (blue), and $CO_2$ (green) on normalized NPP in the model ensemble. Lines indicate the ensemble mean and the shaded area indicates the 95% confidence intervals.