

1
2
3
4
5
6
7
8

DR. COLIN AVERILL (Orcid ID : 0000-0003-4035-7760)
PROF. MICHAEL DIETZE (Orcid ID : 0000-0002-2324-2518)

Article type : Primary Research Articles

Title: Continental-scale nitrogen pollution is shifting forest mycorrhizal associations and soil carbon stocks

Running head: Nitrogen pollution and forest mycorrhizas

Authors: Colin Averill^{1,2*}, Michael C. Dietze² and Jennifer M. Bhatnagar¹

Affiliation: ¹Department of Biology, Boston University, Boston, MA, U.S.A.; ²Department of Earth & Environment, Boston University, Boston, MA, U.S.A.

*correspondence should be sent to caverill@bu.edu, 5 Cummington Mall, Boston University, Boston, MA 02215, 203-444-8925

Keywords: mycorrhizal fungi, nitrogen deposition, soil carbon storage, forest ecology, microbial ecology, nutrient limitation.

Paper Type: Primary Research Article

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/gcb.14368](https://doi.org/10.1111/gcb.14368)

This article is protected by copyright. All rights reserved

9 **Abstract:** Most tree roots on Earth form a symbiosis with either ecto- or arbuscular mycorrhizal
10 fungi. Nitrogen fertilization is hypothesized to favor arbuscular mycorrhizal tree species at the
11 expense of ectomycorrhizal species due to differences in fungal nitrogen acquisition strategies,
12 and this may alter soil carbon balance, as differences in forest mycorrhizal associations are
13 linked to differences in soil carbon pools. Combining nitrogen deposition data with continental-
14 scale U.S. forest data, we show that nitrogen pollution is spatially associated with a decline in
15 ectomycorrhizal vs. arbuscular mycorrhizal trees. Furthermore, nitrogen deposition has
16 contrasting effects on arbuscular vs. ectomycorrhizal demographic processes, favoring
17 arbuscular mycorrhizal trees at the expense of ectomycorrhizal trees, and is spatially correlated

18 with reduced soil carbon stocks. This implies future changes in nitrogen deposition may alter the
19 capacity of forests to sequester carbon and offset climate change via interactions with the forest
20 microbiome.

21 **Introduction:** Forests represent 31% of global land surface area (Food and Agriculture
22 Organization, 2010) and currently offset ~2.4 Pg of CO₂ emissions every year by storing carbon
23 (C) in live plant biomass and soil (Pan *et al.*, 2011). A central component of this C storage
24 capacity may be the presence of different members of the forest microbiome. Two dominant
25 classes of mycorrhizal fungi, the arbuscular mycorrhizal fungi and ectomycorrhizal fungi (AM
26 and EM, respectively), form a symbiosis with the roots of most trees on Earth, enhancing access
27 to soil nutrients and water (Smith & Read, 2009; van der Heijden *et al.*, 2015). AM fungi
28 primarily rely on inorganic forms of N (Hodge & Fitter, 2010; Phillips *et al.*, 2013) and small
29 organic N compounds (Whiteside *et al.*, 2012). In contrast, EM fungi are thought to rely more
30 heavily on organic N sources (Phillips *et al.*, 2013), having a greater capacity to invest in N

31 degrading extracellular enzymes that access complex organic forms of N in soil, such as proteins
32 and chitin (Fernandez & Kennedy, 2016). EM fungi are associated with both slower
33 decomposition of soil organic matter by free-living microbial populations and increased soil C
34 storage (Gadgil & Gadgil, 1971; Averill *et al.*, 2014; Averill & Hawkes, 2016; Fernandez &
35 Kennedy, 2016; Kyaschenko *et al.*, 2017), potentially by competing with free-living soil
36 microbes for these organic N resources or differences in plant or fungal organic matter chemistry
37 (Averill & Hawkes, 2016; Fernandez & Kennedy, 2016). These distinctions between AM and
38 EM fungi lead to two important predictions: 1- that inorganic N inputs to ecosystems will favor
39 AM associated trees at the expense of EM associated trees, and 2- that inorganic N driven
40 declines in EM composition of forests will reduce the belowground C storage capacity of the
41 forest biome.

42 Continental scale inorganic N deposition provides an opportunity to test predictions of
43 how N, plant-microbial systems, and soil C interact at large spatial scales. N deposition in the
44 U.S. has increased 5-10 times over preindustrial levels, predominately due to fossil fuel burning
45 (Aber *et al.*, 2003). Variability in N emission source locations, topography, wind and rainfall
46 patterns produce a heterogeneous geographic distribution of N deposition in the U.S., allowing N
47 deposition patterns to be leveraged as an unintentional inorganic N addition experiment.
48 Previous species-level analyses suggest that both AM and EM tree growth rates can respond
49 positively to N deposition (Thomas *et al.*, 2010); however, it is unclear whether or not these
50 species-level responses translate to changes in the relative balance of EM vs. AM associated
51 trees at the community-level, which regulates landscape-level soil biogeochemical cycling.

52 To determine if N deposition is changing the total composition of these symbiont types at
53 a scale that could impact the C storage capacity of the biosphere, we analyzed the relative
54 abundance of AM and EM trees in 2,760 U.S. forest plots, comprised of 65,769 trees, coupled
55 with direct site-level observations of soil C stocks collected from the U.S. Forest Service's Forest
56 Inventory and Analysis database (Forest Inventory and Analysis, 2014, Fig. 1). We analyzed
57 how anthropogenic N deposition is spatially correlated with the relative abundance of AM and
58 EM trees, as well as current forest growth, tree recruitment, mortality rates, and total soil C
59 stocks, using a Bayesian statistical framework. Doing so enabled us to ask if N-fertilization via
60 anthropogenic N deposition favors AM at the expense of EM associated trees and if these
61 changes in forest mycorrhizal composition may translate to differences in soil C balance.

62

63 **Materials and Methods:**

64 **General Analysis Approach:** This analysis aimed to isolate the effect of N deposition on soil C
65 storage, the relative abundance of EM vs. AM trees, and EM vs. AM basal area growth,
66 recruitment and mortality rates across the continental U.S., while controlling for the potential
67 effects of mean annual temperature, mean annual precipitation, soil pH, soil C:N or total N
68 stock, and soil clay content. This is done using a Bayesian multiple regression framework,
69 where the relationship between all environmental predictors and individual EM or AM tree
70 dynamics are estimated simultaneously. In Bayesian analysis, a probability distribution is
71 estimated for each dependent variable (i.e. EM or AM relative abundance data, growth rates,
72 recruitment counts, mortality occurrences and soil C stocks) as it is regressed against the
73 independent variables in the multiple regression model (Dietze, 2017). A relationship between
74 a dependent and independent variable can be considered significant if the 95% credible interval
75 of the corresponding parameter estimate does not overlap zero. We implemented these models
76 in a Bayesian statistical framework because it allowed us to account for variable re-census
77 intervals in forest inventory data without transforming data, and in basal area growth,
78 recruitment and mortality models it allowed us to flexibly fit non-linear relationships and more
79 complex probability distributions. Furthermore, Bayesian models allow us to propagate
80 uncertainty by sampling from the distribution of each parameter estimate in our model when
81 estimating the predicted effect of N deposition on a dependent variable of interest (Dietze,
82 2017). For example, when we estimate the effect of N deposition on the relative abundance of
83 EM trees in a forest holding all other predictors constant at their means, the 95% credible
84 interval of the mean effect at a given level of N deposition reflects not only the uncertainty in
85 the N deposition effect, but also the uncertainty in temperature, precipitation, pH, C:N, clay
86 and intercept parameter estimates.

87 Only predictors with parameter estimates where 95% credible intervals did not overlap
88 zero are discussed as having a significant effect in the main text. When reporting or visualizing
89 the effect of a single predictor variable (say N deposition) on a dependent variable, we varied
90 the predictor over its entire range in the data set, holding all other predictors constant at their
91 mean values. In this way, we evaluated the effect size of a predictor in isolation, while

92 accounting for the effects and uncertainties of all other predictors in the model. Complete
93 description of data collection and analysis are detailed below.

94 **Data Collection:** This analysis used all plots within the U.S. forest inventory and analysis
95 (FIA) data set (Forest Inventory and Analysis, 2014) that have paired soil C observations. Soil
96 C profiles are aggregated down to 20 cm depth in the mineral soil (if soils are this deep), plus
97 any overlying organic horizon profiles. This approach causes the total soil profile depth to vary
98 across sites, because while most profiles contain 20cm of mineral soil, the depth of the forest
99 floor horizon will vary. Because our analysis is premised on the hypothesis that EM
100 ecosystems may slow decomposition due to an ecological interaction, it may be that variation
101 in organic horizon thickness is an outcome of this process. In contrast, we do not expect
102 mycorrhizal associations to modify the depth of the mineral soil horizon. Therefore, we
103 standardized soil element aggregation to a fixed depth in the mineral soil plus any overlying
104 organic material. This generated soils data (soil C, N, soil pH), paired with forest composition
105 data for each site. A subset of sites had been re-measured in time, allowing calculation of forest
106 basal area growth, recruitment and mortality rates. These two data sets were used for five down
107 stream analyses. We refer to the complete data set as the "abundance data set" and the second
108 that includes temporal remeasurement of forest properties as the "remeasurement data set". We
109 only used FIA plots that are forested and have no recent evidence of active management or
110 human harvesting. We note that most forests in the United States have some history of human
111 disturbance, and our analysis cannot account for time since last time stand replacing
112 disturbance. Full details of data filtering, soil aggregation, and other site selection procedures
113 are presented in the Supporting Information 1.

114 Within each plot, each tree was assigned a mycorrhizal type based on the
115 supplementary data file in Phillips et al. (Phillips *et al.*, 2013), which provides mycorrhizal
116 assignments for most tree species within the FIA database. This file was supplemented with
117 information on mycorrhizal status for western U.S. tree species (References provided in
118 Supporting Information 2). We calculated the relative abundance of EM trees as the basal area
119 of all live EM trees within a plot, divided by the total basal area of live trees within the plot.
120 We only analyzed sites where more than 90% of forest basal area consisted of AM or EM
121 mycorrhizal tree species (AM, EM, or a combination of the two types) at the time of soil
122 sampling (i.e. $(\text{Basal}_{\text{AM}} + \text{Basal}_{\text{EM}}) / \text{Basal}_{\text{total}} > 0.9$). Hence, a plot consisting of 50% EM and

123 50% AM trees by basal area would be retained in this analysis. We excluded plots with >10%
124 non-mycorrhizal, arbutoid, and eriocid mycorrhizal symbionts, as well as hosts that are known
125 to strongly associate with both AM and EM, but this only constituted 15% of the available
126 sites. The final abundance data set included 2,760 unique sites, encompassing 65,769 trees.
127 Past analyses along AM to EM gradients have demonstrated that the relative abundance of AM
128 vs. EM trees aboveground can be indicative of the abundance of these fungal community
129 members belowground (Nave *et al.*, 2013; Soudzilovskaia *et al.*, 2015; Cheeke *et al.*, 2016).

130 The second FIA database subset (“re-measurement dataset”) represents the first re-
131 measurement of all FIA forest plots after the initial soil sampling and was used for analyses of
132 basal area growth, recruitment and mortality. The majority of plots (1,912 of 2,760 plots) had
133 been re-measured since the time of the initial soil sampling. Initial visits were conducted
134 between 2000-2011. Re-measurement visits were conducted in 2004-2015. The average
135 duration of the re-measurement interval was 6.3 years. Because the soil properties used in this
136 analysis change slowly at this time scale, we are confident that the demographic patterns
137 observed in the re-measurement analysis are relatable to the initial soil conditions. We chose to
138 analyze the observation after the initial measurements, rather than the one prior, because many
139 of the soil sites included in the Phase 3 data are new to the FIA, and as a result, had no prior
140 observation of forest composition. Using the observations during and after the soil sampling
141 enabled more data to be included in the tree re-measurement analysis. The final data set used
142 for the analysis of tree mortality included 1,912 unique sites, encompassing 54,378 trees (Fig.
143 1).

144
145 *Climate, N deposition and soil clay data:* Climate data for each plot were assigned using
146 PRISM 30-year 1981-2010 climate normal data products at 800-meter resolution for mean
147 annual temperature (MAT) and mean annual precipitation (MAP) (PRISM Climate Group &
148 Oregon State University, 2017). Soil texture (% clay) data was extracted from the North
149 American Carbon Project Unified North American Soil Map at 0.25 degree resolution (Liu *et*
150 *al.*, 2014). N deposition rates were determined using the National Atmospheric Deposition
151 Program (NADP) 15-year mean wet and dry N deposition ($\text{NH}_4^+ + \text{NO}_3^-$) from the 2000-2014
152 interval at 0.25° resolution (National Atmospheric Deposition Program, 2015). While
153 deposition data are for the most recent time interval available, comparison of wet N deposition

154 data from the 15-year 2000-2014 interval is strongly correlated with the 30-year 1985-2014
155 interval ($r^2 = 0.96$). Therefore, we are confident that 2000-2014 data are representative of
156 historical patterns of N deposition loads in the United States. It is important to note that the FIA
157 does not report latitude and longitude data with exact precision; latitude and longitude
158 coordinates are limited in accuracy to ~800 meters. However, given the resolution of the data
159 products used here, and the fact that climate and N deposition variables are highly spatially
160 auto-correlated at the kilometer scale, the spatial accuracy of the climate and N deposition
161 products is on par with the spatial accuracy of the FIA site locations (Dietze & Moorcroft,
162 2011). In addition to not reporting exact spatial locations, the FIA randomly swaps 20% of
163 plots within a county. Because plots are swapped randomly, swapping should not bias our
164 parameter estimates, only increase uncertainty in our parameter estimates associated with
165 predictors derived from spatial products (MAT, MAP and N deposition). Furthermore,
166 variation in these predictors is much larger between counties than within, and therefore we can
167 still capture variation in forest and soil processes driven by these predictors, even if all sites
168 were randomly swapped within a county. Finally, the number and spatial scale of wet and dry
169 observation network sites is unlikely to capture small-scale point sources, which can generate
170 high levels of N deposition at very local scales. These effects will necessarily be missed by our
171 analysis. Despite this limitation, the >20x variation captured in the NADP N deposition data
172 ($1.2\text{-}27.0 \text{ kg N ha}^{-1} \text{ yr}^{-1}$) may still inform the underlying drivers of variation in mycorrhizal
173 functional types.

174 Spurious relationships between N deposition and AM vs. EM relative abundance or
175 forest demographic processes are possible, if the geographic distribution of N deposition is
176 spatially confounded with other environmental factors known to influence forest dynamics. For
177 example, the highest levels of N deposition are found in the Midwest, where there is some
178 evidence of a shift from EM to AM species due to changes in fire regime (McEwan *et al.*,
179 2011). To address this potential confounding among N deposition, forest mycorrhizal type, and
180 other environmental factors, we replicated our analyses, first excluding all states west of the
181 Mississippi River, and then excluding both Western states (comprised predominantly of EM
182 associated stands, and has, on average, the lowest levels of N deposition) and Midwestern
183 states.

184

Statistical Analysis

185 *Modeling the relative abundance of EM trees:* We modeled the relative abundance of EM trees
186 at the plot level as a continuous beta distribution, which is designed to deal with proportional
187 data on the interval (0,1), accounting for changes in variance as a function of the mean.
188 Because our data set included many 0's and 1's, we transformed relative abundance values
189 according to:

$$y = (y * (n - 1) + 0.5) / n$$

190 Where y is the relative abundance of EM basal area, bounded on the interval [0,1] and n is the
191 sample size (Smithson & Verkuilen, 2006). This had the effect of transforming 0 and 1 values
192 to 0.000210 and 0.99979. We also ran models with 0,1 transformation to 0.010,0.990 and
193 0.001,0.999. All models generated analogous results, with similar effect sizes. The parameters
194 of the beta distribution, p_i and q_i were described as:

$$p_i = m_i * t$$
$$q_i = (1 - m_i) * t$$

195 We modeled m_i as the logit transform of the linear combination of MAT, MAP, soil C:N, soil
196 pH, soil clay and N deposition. All parameters were assigned normally distributed,
197 uninformative priors, save for t , which was assigned a gamma distributed, uninformative prior.

198
199 *Modeling plot-scale basal area growth of surviving trees:* Basal area growth was estimated at
200 the plot-level, on plots visited one measurement period after the initial soil measurement (i.e.
201 “re-measurement dataset”). Because this analysis is primarily concerned with which plant
202 mycorrhizal type is most abundant, we chose to analyze growth as basal area increment, rather
203 than C increment, as this better reflects which trees are largest and growing the fastest within a
204 plot. Basal area increment was modeled as a log-normal distribution, where the log of the
205 survival basal area increment was modeled as a linear function of previous basal area, climate,
206 soil and N deposition, and an interaction between N deposition and the relative abundance of
207 EM associated trees within a plot at the end of the measurement interval. Once models were fit,
208 visual inspection of the relationship between basal area growth and N deposition suggested an
209 intermediate 'hump' in the relationship between N deposition and stand basal area growth in the
210 full data set. To model this, we added a Gaussian relationship between N deposition and basal
211 area growth, in addition to the linear relationships, which took the form of:

$$g_1 * e^{-\frac{(x - g_2)^2}{2g_3^2}}$$

212 Where g_1 , g_2 and g_3 are the parameters of the Gaussian distribution and x is N deposition. We
 213 placed a positive constraint on Gaussian parameter priors to aid in model convergence. Finally,
 214 models with Gaussian perturbations resulted in multiple alternative parameter combinations
 215 that were equally parsimonious, and therefore not all chains converged. To satisfy convergence
 216 criteria we subsetting model output to chains that converged on a similar area of parameter
 217 space, though we note there were several Gaussian parameter combinations that returned
 218 equivalent results.

219

220 *Modeling tree-level mortality:* Mortality probability (M_i) was estimated based on the logistic
 221 regression model for binary mortality data. Because the FIA census interval varies at the plot
 222 level, we related annual mortality probability (p_i) of tree i to the observed binary mortality data
 223 (M_i) using a Bernoulli likelihood (Dietze & Moorcroft, 2011),

$$M_i = \text{Bern}(1 - (1 - p_i)^{t_i})$$

224 Where t_i is the duration of the census interval for a given plot. p_i was modeled as the inverse
 225 logit transform of z_i , where z_i is a linear combination of MAT, MAP, soil C:N, soil pH, soil
 226 clay, mycorrhizal type, N deposition, and a linear interaction between N deposition and
 227 mycorrhizal association. In calculating z_i , we accounted for the non-linear relationship between
 228 tree diameter (D) and mortality probability by including two exponential terms. The first term
 229 accounts for the initial decline in tree mortality probability as trees increase in size. The second
 230 term accounts for the increasing mortality probability of large diameter, older trees. The
 231 summation of these two relationships will generate the commonly observed "J" shaped
 232 relationship between tree mortality probability and tree diameter (Dietze & Moorcroft, 2011).

233 The full expression of z_i was constructed as,

$$z_i = b_1 * MAT + b_2 * MAP + b_3 * soil_{C:N} + b_4 * soil_{pH} + b_5 * soil_{clay} + b_6$$

$$* myc_type + b_7 * N_{dep} + b_8 * myc_type * N_{dep} + e^{-b_9 * D} + e^{b_{10} * D}$$

234 Parameters b_1 - b_8 were assigned normally distributed, uninformative priors. Parameters b_9 and
 235 b_{10} priors were assigned exponentially distributed priors, constrained to values greater than or
 236 equal to 0 to avoid fitting redundant model terms.

237

238 *Modeling plot level recruitment:* Recruitment of trees was modeled using a zero-inflated
239 Poisson model, to account for the fact that the abundance of zeros in tree recruitment data is
240 often poorly fit by traditional Poisson models (Fortin & DeBlois, 2007). A multi-model
241 average was fit, modeling the occurrence of recruitment as a Bernoulli process, and then the
242 number of recruits as a Poisson process. The Poisson model was then weighted by the
243 probability of recruitment occurring estimated by the Bernoulli model in order to account for
244 zero inflation. Because the FIA census interval varies at the plot level, we related the annual
245 number of recruits (r_i) of plot i to the observed number of recruits over a given census interval
246 (R_i) using a Poisson likelihood,

$$R_i = Pois(r_i * t_i)$$

247 Where t_i is the duration of the census interval for a given plot. We fit separate models for AM
248 and EM recruitment. MAT, MAP, soil C:N, soil pH, soil clay and N deposition were used as
249 predictors for both Poisson and Bernoulli components of the model. Finally, the number of AM
250 or EM recruits within a plot will strongly depend on the relative abundance of AM or EM trees
251 already present within a plot (i.e. the apple does not fall far from the tree). To account for this,
252 we included the relative abundance of AM or EM trees as covariates within models of AM or
253 EM recruitment.

254
255 *Modeling soil C storage:* Soil carbon was modeled as a linear function of MAT, MAP, soil pH,
256 soil clay, soil N stock, the relative abundance of EM tree basal area, and an interaction between
257 the relative abundance of EM tree basal area and soil N stocks. Models relating soil C and N at
258 large spatial scales generally violate the assumption of normality of residuals (Averill *et al.*,
259 2014), as there are multiplicative errors associated with the measurement of soil C on an area
260 basis. Therefore, we modeled soil C with a lognormal distributed probability density, and log
261 transformed the soil N predictor. All parameters were assigned normally distributed,
262 uninformative priors. We considered accounting for residual spatial autocorrelation in this
263 model, yet initial semivariograms of soil C fit to an intercept only model exhibited little
264 residual spatial autocorrelation. Because of this, we proceeded to analyze all soil C
265 observations under the assumption they are spatially independent.

266

267 **Bayesian model diagnostics and convergence:** Forest basal area growth, mortality, recruitment
268 and soil C storage analyses were all analyzed in a Bayesian framework. Bayesian models were
269 fit using the runjags package for R statistical software (Denwood, in review; R Core Team,
270 2017). JAGS models were initialized with three chains, checked for convergence, and had
271 initial burn-in iterations removed before estimation of final parameter values. Parameter values
272 were estimated based on a minimum of 10,000 samples of 3 independent chains, after 2,000
273 adaptive iterations and 4,000 burn-in iterations. Parameter estimates were deemed converged
274 based on visual inspection of chain parameter estimates for the last 10,000 iterations, and
275 checking that Gelman diagnostic values for each parameter were below 1.1 (Gelman, 2014).
276 Additional samples were taken as necessary to achieve convergence.

277
278 **Reporting of Results:** We discuss results for predictors where the 95% credible interval did not
279 overlap zero, which is our criteria for statistical significance, and refer to predictors that do not
280 meet this condition as having no effect on the modeled dependent variable (Gelman, 2014).
281 Recruitment mean and 95% confidence interval is reported as the mean number of recruits based
282 on the fit of the Poisson model prior to discretization via the Poisson distribution, weighted by
283 the probability of recruitment occurring based on the binomial model.

284 We plotted binned means of dependent variables in response to N deposition, controlling
285 for the effects of other covariates in the model. To do this, we subtract the product of each
286 parameter-covariate combination of each observation from the dependent variable, and add back
287 in the mean value of the parameter-covariate combination from the data set. For example, if
288 forest basal area growth was modeled by the function,

$$basal\ area\ growth_i = 0.5 * temperature_i + 0.1 * precipitation_i$$

289 To isolate the effect of precipitation on basal area growth and remove the effect of temperature,
290 we would transform our observations of basal area growth to remove the variance associated
291 with temperature using the following transformation:

$$basal\ area\ growth_{T_i} = growth_i - 0.5 * temperature_i + 0.5 * temperature_{mean}$$

292 Where basal area $growth_{T_i}$ is the transformed value of an individual basal area growth
293 observation, controlling for the effect of temperature. Error bars represent the standard error of
294 the mean within a bin, except for mortality bins where error bars represent 95% confidence

295 intervals, calculated using the Pearson-Klopper method, implemented in the binom package for
296 R statistical software (Sundar, 2014).

297
298 **Results:** Across the continental U.S., N deposition was associated with a greater abundance of
299 AM compared to EM dominated forests (Fig. 2a). Furthermore, we observed contrasting
300 relationships between AM vs. EM basal area growth and recruitment rates with increasing N
301 deposition. AM trees grew faster with increasing levels of atmospheric N deposition compared to
302 EM trees (Fig. 2c,d,e). In addition, N deposition increased AM tree recruitment rates, while EM
303 recruitment rates declined (Fig. 2f, 2g,h). N deposition also increased the probability of tree
304 mortality, however we detected no difference between the mortality rates of AM vs. EM
305 associated trees (Fig. 2b). These relationships were robust to subsetting the data set to exclude
306 the western U.S. or both the western and midwestern U.S. states (Supporting Information Fig. 1-
307 2).

308 In an effort to quantify this potential influence of forest disturbance, we analyzed the
309 relationship between N deposition and stand age in the FIA data set. We found younger forest
310 sites tended to have higher levels of N deposition, yet this relationship was lost once western
311 U.S. states were excluded. To determine if correlation between forest age and N deposition may
312 be biasing our results we re-ran models of EM relative abundance including forest age as an
313 additional covariate. We found a negative relationship between forest age and EM relative
314 abundance, inconsistent with a positive correlation between N deposition and forest age driving
315 our findings. Furthermore, inclusion of the stand age covariate did not change the parameter
316 estimate of N deposition in our model of EM relative abundance.

317 Aboveground forest responses to N deposition were also associated with changes in the C
318 content of U.S. forest soils. At low levels of N deposition, EM forests stored more C in soils than
319 AM forests (Fig. 3a). However, N deposition interacted with mycorrhizal type; at the highest
320 levels of N deposition C-storage decreased and the difference in C-storage between AM and EM
321 forests disappeared (Fig. 3b). Subsetted analyses showed EM sites were still associated with
322 elevated soil C content, though the interactive effect of mycorrhizal type and N deposition on
323 soil C was lost (Fig. S3, S4).

324 We report full model output (parameter estimates, standard deviations, and significance
325 values) in Supporting Information 1 Tables 1-6. We also provide beta factors (the change in

326 dependent variable associated with 1 standard deviation variation in each independent variable)
327 in Supporting Information 1 Tables 7-12.

328

329 **Discussion:** AM and EM fungi are the dominant groups of forest mycorrhizal fungal symbionts
330 worldwide (Read & Perez-Moreno, 2003; Phillips *et al.*, 2013). These fungi have long been
331 thought to represent adaptations to differences in ecosystem N availability (Read, 1991), but
332 ecosystem-scale validation of this hypothesis has been challenging, as it takes decades to observe
333 changes in forest composition. Our analysis demonstrates that N deposition from anthropogenic
334 sources is spatially correlated with lower EM abundance and greater AM abundance in U.S.
335 forests, supporting this hypothesis (Figure 2a). Our analysis of demographic process rates also
336 shows higher levels of N deposition drive a suite of changes in basal area growth, recruitment
337 and mortality rates favoring AM trees at the expense of EM trees (Figure 2b-2h). This result
338 implies that N deposition is currently changing the balance of AM and EM associated trees in
339 U.S. forests. In addition, the interaction between mycorrhizal strategies, N deposition, and soil C
340 storage indicates that ongoing changes in forest composition – particularly of the soil
341 microbiome – with N inputs may have downstream consequences for belowground C
342 sequestration and global climate change, to the extent that a loss of EM forest community
343 members with increasing levels of N deposition translates to a loss of soil C from forest floor and
344 upper mineral soil layers.

345 Our analyses of forest composition and demographic process rates control for soil C:N
346 ratio and soil pH (key indicators of soil fertility), and we excluded sites with evidence of recent
347 human management. However, we cannot rule out the potential influence of correlations between
348 historical forest disturbance regimes and forest mycorrhizal composition, which may be
349 confounded with N deposition. In an effort to quantify this potential influence of forest
350 disturbance, we analyzed the relationship between N deposition and stand age in the FIA data
351 set. We did find that younger forest sites tended to have higher levels of N deposition, yet this
352 relationship was lost once western U.S. states were excluded. Furthermore, incorporating stand
353 age into our analysis showed a *negative* relationship between stand age and relative abundance of
354 EM trees- i.e. older stands had more AM trees, and did not change N deposition parameter
355 estimates. Based on these findings we believe it is unlikely that correlations between N
356 deposition and historical forest disturbance regimes are driving our results.

357 Our analysis supports the conclusion that N deposition is currently shaping the relative
358 abundance of AM and EM trees in U.S. forests and will continue to do so in the future. This is
359 because N deposition favors AM basal area growth and recruitment at the expense of EM basal
360 area growth and recruitment (Fig. 2c,h), but does not have a differential effect on mortality
361 through time. Furthermore, basal area growth and recruitment curves cross. This combination of
362 crossing temporal demographic responses to N deposition means that, given sufficient time, a
363 pressure of high N deposition rates will transition an EM forest to an AM forest, or the
364 suspension of high N deposition rates may revert an AM forest to an EM state.

365 We observed strong spatial relationships between N deposition, mycorrhizal type, and
366 soil C stocks after accounting for major environmental factors known to influence mycorrhizal
367 basal area growth in forests across space (climate, total N stock, and other soil properties),
368 suggesting that these patterns are foundational for predicting changes in soil C stocks with N
369 pollution. Our observations are also consistent with theoretical work showing that EM associated
370 trees can stabilize soil C by inducing N limitation of microbial decomposer communities (Read
371 & Perez-Moreno, 2003; Orwin *et al.*, 2011; Phillips *et al.*, 2013) and with experimental data
372 validating predictions from this theory (Averill & Hawkes, 2016). Elevated C storage in EM
373 ecosystems may also be driven by additional mechanisms, such as differences in input litter
374 chemistry between AM and EM trees or fungi (Cornelissen *et al.*, 2001; Fernandez & Koide,
375 2012; Phillips *et al.*, 2013). Regardless of the particular mechanism, our analysis suggests that
376 global change impacts on plant-microbial interactions in forest ecosystems may have large
377 consequences for C sequestration at continental scales.

378 The negative correlation we observed between N deposition and soil C stocks contrasts
379 other empirical work showing that N inputs broadly suppress the activity of microbial
380 decomposers in forest ecosystems (Janssens *et al.*, 2010). One reason for this underlying
381 discrepancy may be fundamental differences in how N is added via N deposition compared to N
382 fertilization experiments. While both additions may result in the same annual rate of N input, N
383 fertilizer intended to simulate anthropogenic N deposition is frequently added as several doses at
384 far higher concentration than is experienced in continuous low-concentration N deposition in
385 North American temperate forests (Vadeboncoeur, 2010). High concentration additions of
386 mineral N to soil may transiently overwhelm microbial N demand, driving elevated nitrification
387 and acidity (Gundersen & Rasmussen, 1990; Tian & Niu, 2015; Chen *et al.*, 2016) compared to

388 what would occur under more realistic anthropogenic N loading regimes. N fertilization
389 experiments frequently observe changes in pH (Ramirez *et al.*, 2010; Riggs *et al.*, 2015) and soil
390 water potential (Braddock *et al.*, 1997), suggesting that there may also be a direct effect of
391 elevated salinity and osmotic stress on microbial growth and activity when adding N fertilizers
392 (Averill and Waring 2018). Our analysis explicitly controls for soil pH, and this may explain the
393 discrepancy in results. In addition, we found that N deposition effects on soil C depend on the
394 relative abundance of EM community members (Fig. 3a). Such a phenomenon may result if N
395 deposition relieves N limitation of saprotrophic microbial populations in soils due to competition
396 with EM fungi for N resources, or if there are differences in organic matter chemistry between
397 AM and EM ecosystems. Either mechanism would result in N deposition increasing
398 decomposition activity and driving further losses of soil C.

399 The FIA soil sampling only extends to 20cm depth in the mineral soil. It is possible AM
400 and EM forests have differential effects at greater soil depths, and these patterns will be missed
401 by this analysis of shallower soil profiles. Furthermore, there is increasing recognition that
402 particulate vs. mineral associated organic matter may respond differentially and in contrasting
403 directions to changes in soil microbial activity (Schmidt *et al.*, 2011; Cotrufo *et al.*, 2013; Averill
404 & Waring, 2018). This analysis only considers bulk soil C, and cannot speak to differential
405 effects of mycorrhizal association on these finer scale soil C pools. Additional measurements of
406 soil mineral sorptive capacity, beyond what can be captured by basic soil texture measurements,
407 will likely improve statistical models designed to capture spatial variation in soil C storage.

408 Our analysis of the relative abundance of EM basal area in forests support the long-
409 standing paradigm of what controls the abundance of EM vs. AM trees within a forest (Read,
410 1991; Smith & Read, 2009). In temperate forests of the U.S., EM trees are most abundant in cool
411 and wet habitats, while AM trees are more abundant in warmer and drier habitats (Fig. S6). We
412 found that soil C:N and soil pH were also strong predictors of the relative abundance of EM vs.
413 AM trees. Given that EM trees often generate acidic litter inputs to ecosystems (Phillips *et al.*,
414 2013), it is difficult to say whether the soil pH is a cause or effect of the presence of EM trees.
415 Nevertheless, our finding that AM associations are more beneficial to trees under high soil N
416 availability compared to EM associations supports long-standing theories of differences among
417 mycorrhizal types (Read & Perez-Moreno, 2003; Phillips *et al.*, 2013). Overall, our analysis

418 provides critical empirical support for long hypothesized environmental controls over the relative
419 abundance of AM vs. EM trees within the temperate forest biome.

420 It is important to note that while this analysis assumes all EM trees behave the same,
421 there is substantial variability in ECM fungal communities even in forests dominated by a single
422 tree species (Clemmensen *et al.*, 2013; Averill & Hawkes, 2016). This variation in composition
423 likely drives changes in the functioning of these mycorrhizal communities, as different EM fungi
424 are known to be differentially sensitive to N deposition (Lilleskov *et al.*, 2002), and these
425 different N response sensitivities have been linked to different mycorrhizal exploration types and
426 resource preferences (Hobbie & Agerer, 2010; Hobbie *et al.*, 2014). Differences in EM fungal
427 communities within forest types that affect tree demographic processes and soil C stocks will be
428 missed by this analysis, and likely contribute to uncertainty in model fits. We see simultaneous
429 modeling of both plant and soil microbial communities through time, as well as their interaction,
430 as an urgent next step in predicting future states of temperate forest ecosystems.

431 This study suggests a potential for losses of belowground C in response to N deposition,
432 which may substantially mitigate previously reported N deposition induced increases in
433 aboveground C storage (Thomas *et al.*, 2010). Validating the relative contributions of N
434 deposition to C storage vs. loss will require repeated measurements of forest soil C profiles
435 across N deposition gradients in time. Our study provides a framework for how large-scale forest
436 inventory datasets could be used to predict the balance of ecosystem C stocks, specifically by
437 injecting critical aspects of soil microbial community composition into ecosystem models of the
438 terrestrial C cycle. AM and EM mycorrhizal associations have been linked to the ability of
439 ecosystems to sustain plant productivity in response to elevated CO₂ (Terrer *et al.*, 2016) and
440 soil C storage capacity (Averill *et al.*, 2014; Averill & Hawkes, 2016) via interactions with soil
441 N availability. The spatial extent of forest inventory data, coupled with the developing capacity
442 to remote sense forest mycorrhizal associations (Fisher *et al.*, 2016), provides information to
443 develop and calibrate new simulation models that explicitly represent these divergent categories
444 of forest microbial symbiont types (Brzostek *et al.*, 2014). This coupled approach promises to
445 advance a new class of ecosystem models that can address the role of the forest microbiome in
446 the C cycle at a scale necessary to understand ecosystem responses to global environmental
447 change.

448

449 **Acknowledgements:** Ryan Kelly provided code necessary to query the FIA database. Bonnie
450 Waring, Will Pearse and Edward Brzostek provided feedback on an earlier version of this
451 manuscript. The project was funded by the Peter Paul Professorship to J.M. Bhatnagar at Boston
452 University. Colin Averill was supported by the NOAA Climate and Global Change Postdoctoral
453 Fellowship Program, administered by Cooperative Programs for the Advancement of Earth
454 System Science (CPAESS), University Corporation for Atmospheric Research (UCAR),
455 Boulder, Colorado, USA. Michael Dietze was supported by NSF Macrosystems #1318164. All
456 forest and soils data used in this analysis are freely available via the U.S. Forest Service Forest
457 Inventory and Analysis DataMart (<https://apps.fs.usda.gov/fia/datamart>).

References

- Aber JD, Goodale CL, Ollinger SV et al. (2003) Is Nitrogen Deposition Altering the Nitrogen Status of Northeastern Forests? *BioScience*, **53**, 375.
- Averill C, Hawkes CV (2016) Ectomycorrhizal fungi slow soil carbon cycling. *Ecology Letters*, **19**, 937–947.
- Averill C, Waring B (2018) Nitrogen limitation of decomposition and decay: How can it occur? *Global Change Biology*, **24**, 1417–1427.
- Averill C, Turner BL, Finzi AC (2014) Mycorrhiza-mediated competition between plants and decomposers drives soil carbon storage. *Nature*, **505**, 543–545.
- Braddock JF, Ruth ML, Catterall PH, Walworth JL, McCarthy KA (1997) Enhancement and Inhibition of Microbial Activity in Hydrocarbon-Contaminated Arctic Soils: Implications for Nutrient-Amended Bioremediation. *Environmental Science & Technology*, **31**, 2078–2084.
- Brzostek ER, Fisher JB, Phillips RP (2014) Modeling the carbon cost of plant nitrogen acquisition: Mycorrhizal trade-offs and multipath resistance uptake improve predictions of retranslocation: Carbon cost of mycorrhizae. *Journal of Geophysical Research: Biogeosciences*, **119**, 1684–1697.
- Cheeke TE, Phillips RP, Brzostek ER, Rosling A, Bever JD, Fransson P (2016) Dominant mycorrhizal association of trees alters carbon and nutrient cycling by selecting for microbial groups with distinct enzyme function. *New Phytologist*.

- Chen D, Li J, Lan Z, Hu S, Bai Y (2016) Soil acidification exerts a greater control on soil respiration than soil nitrogen availability in grasslands subjected to long-term nitrogen enrichment (ed Niu S). *Functional Ecology*, **30**, 658–669.
- Clemmensen KE, Bahr A, Ovaskainen O et al. (2013) Roots and Associated Fungi Drive Long-Term Carbon Sequestration in Boreal Forest. *Science*, **339**, 1615–1618.
- Cornelissen J, Aerts R, Cerabolini B, Werger M, van der Heijden M (2001) Carbon cycling traits of plant species are linked with mycorrhizal strategy. *Oecologia*, **129**, 611–619.
- Cotrufo MF, Wallenstein MD, Boot CM, Deneff K, Paul E (2013) The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Global Change Biology*, **19**, 988–995.
- Denwood MJ (in review) runjags: An R package providing interface utilities, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of Statistical Software*.
- Dietze MC (2017) *Ecological Forecasting*. Princeton University Press.
- Dietze MC, Moorcroft PR (2011) Tree mortality in the eastern and central United States: patterns and drivers. *Global Change Biology*, **17**, 3312–3326.
- Fernandez CW, Kennedy PG (2016) Revisiting the “Gadgil effect”: do interguild fungal interactions control carbon cycling in forest soils? *New Phytologist*, **209**, 1382–1394.
- Fernandez CW, Koide RT (2012) The role of chitin in the decomposition of ectomycorrhizal fungal litter. *Ecology*, **93**, 24–28.
- Fisher JB, Sweeney S, Brzostek ER et al. (2016) Tree-mycorrhizal associations detected remotely from canopy spectral properties. *Global Change Biology*, **22**, 2596–2607.
- Food and Agriculture Organization (2010) *Global Forest Resources Assessment 2010*. Food and Agricultural Organization, Rome.
- Forest Inventory and Analysis (2014) *The Forest Inventory and Analysis Database: Database description and user guide version 6.0.1 for Phase 3*. U.S. Department of Agriculture, Forest Service. 182 pp.
- Fortin M, DeBlois J (2007) Modeling tree recruitment with zero-inflated models: the example of hardwood stands in southern Québec, Canada. *Forest Science*, **53**, 529–539.
- Gadgil RL, Gadgil PD (1971) Mycorrhiza and Litter Decomposition. *Nature*, **233**, 133–133.

- Gelman A (2014) *Bayesian data analysis*, Third edition edn. CRC Press, Boca Raton, 661 pp.
- Gundersen P, Rasmussen L (1990) Nitrification in Forest Soils: Effects from Nitrogen Deposition on Soil Acidification and Aluminum Release. In: *Reviews of Environmental Contamination and Toxicology*, Vol. 113 (ed Ware GW), pp. 1–45. Springer New York, New York, NY.
- van der Heijden MGA, Martin FM, Selosse M-A, Sanders IR (2015) Mycorrhizal ecology and evolution: the past, the present, and the future. *New Phytologist*, **205**, 1406–1423.
- Hobbie EA, Agerer R (2010) Nitrogen isotopes in ectomycorrhizal sporocarps correspond to belowground exploration types. *Plant and Soil*, **327**, 71–83.
- Hobbie EA, Hofmockel KS, van Diepen LTA, Lilleskov EA, Ouimette AP, Finzi AC (2014) Fungal carbon sources in a pine forest: evidence from a ¹³C-labeled global change experiment. *Fungal Ecology*, **10**, 91–100.
- Hodge A, Fitter AH (2010) Substantial nitrogen acquisition by arbuscular mycorrhizal fungi from organic material has implications for N cycling. *Proceedings of the National Academy of Sciences*, **107**, 13754–13759.
- Janssens IA, Dieleman W, Luysaert S et al. (2010) Reduction of forest soil respiration in response to nitrogen deposition. *Nature Geoscience*, **3**, 315–322.
- Kyaschenko J, Clemmensen KE, Hagenbo A, Karlton E, Lindahl BD (2017) Shift in fungal communities and associated enzyme activities along an age gradient of managed *Pinus sylvestris* stands. *The ISME Journal*.
- Lilleskov EA, Hobbie EA, Fahey TJ (2002) Ectomycorrhizal fungal taxa differing in response to nitrogen deposition also differ in pure culture organic nitrogen use and natural abundance of nitrogen isotopes. *New Phytologist*, **154**, 219–231.
- Liu D, Wang G, Mei R, Yu Z, Yu M (2014) Impact of initial soil moisture anomalies on climate mean and extremes over Asia: INITIAL SOIL MOISTURE IMPACT ON CLIMATE. *Journal of Geophysical Research: Atmospheres*, **119**, 529–545.
- McEwan RW, Dyer JM, Pederson N (2011) Multiple interacting ecosystem drivers: toward an encompassing hypothesis of oak forest dynamics across eastern North America. *Ecography*, **34**, 244–256.
- National Atmospheric Deposition Program (2015) *NRSP-3*. NADP Program Office, Illinois State Water Survey, University of Illinois, Champaign, IL, USA.

- Nave LE, Nadelhoffer KJ, Le Moine JM, van Diepen LTA, Cooch JK, Van Dyke NJ (2013) Nitrogen Uptake by Trees and Mycorrhizal Fungi in a Successional Northern Temperate Forest: Insights from Multiple Isotopic Methods. *Ecosystems*, **16**, 590–603.
- Orwin KH, Kirschbaum MUF, St John MG, Dickie IA (2011) Organic nutrient uptake by mycorrhizal fungi enhances ecosystem carbon storage: a model-based assessment. *Ecology Letters*, **14**, 493–502.
- Pan Y, Birdsey RA, Fang J et al. (2011) A Large and Persistent Carbon Sink in the World's Forests. *Science*, **333**, 988–993.
- Phillips RP, Brzostek E, Midgley MG (2013) The mycorrhizal-associated nutrient economy: a new framework for predicting carbon-nutrient couplings in temperate forests. *New Phytologist*, **199**, 41–51.
- PRISM Climate Group, Oregon State University (2017).
- R Core Team (2017) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramirez KS, Craine JM, Fierer N (2010) Nitrogen fertilization inhibits soil microbial respiration regardless of the form of nitrogen applied. *Soil Biology and Biochemistry*, **42**, 2336–2338.
- Read DJ (1991) Mycorrhizas in ecosystems. *Experientia*, **47**, 376–391.
- Read DJ, Perez-Moreno J (2003) Mycorrhizas and nutrient cycling in ecosystems—a journey towards relevance? *New Phytologist*, **157**, 475–492.
- Riggs CE, Hobbie SE, Bach EM, Hofmockel KS, Kazanski CE (2015) Nitrogen addition changes grassland soil organic matter decomposition. *Biogeochemistry*, **125**, 203–219.
- Schmidt MWI, Torn MS, Abiven S et al. (2011) Persistence of soil organic matter as an ecosystem property. *Nature*, **478**, 49–56.
- Smith SE, Read DJ (2009) *Mycorrhizal symbiosis*, 3. ed., Repr edn. Elsevier/Acad. Press, Amsterdam, 787 pp.
- Smithson M, Verkuilen J (2006) A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, **11**, 54–71.
- Soudzilovskaia NA, van der Heijden MGA, Cornelissen JHC et al. (2015) Quantitative assessment of the differential impacts of arbuscular and ectomycorrhiza on soil carbon cycling. *New Phytologist*, **208**, 280–293.

- Sundar D-R (2014) *binom: Binomial Confidence Intervals For Several Parameterizations*.
- Terrer C, Vicca S, Hungate BA, Phillips RP, Prentice IC (2016) Mycorrhizal association as a primary control of the CO₂ fertilization effect. *Science*, **353**, 72–74.
- Thomas RQ, Canham CD, Weathers KC, Goodale CL (2010) Increased tree carbon storage in response to nitrogen deposition in the US. *Nature Geoscience*, **3**, 13–17.
- Tian D, Niu S (2015) A global analysis of soil acidification caused by nitrogen addition. *Environmental Research Letters*, **10**, 24019.
- Vadeboncoeur MA (2010) Meta-analysis of fertilization experiments indicates multiple limiting nutrients in northeastern deciduous forests. *Canadian Journal of Forest Research*, **40**, 1766–1780.
- Whiteside MD, Digman MA, Gratton E, Treseder KK (2012) Organic nitrogen uptake by arbuscular mycorrhizal fungi in a boreal forest. *Soil Biology and Biochemistry*, **55**, 7–13.

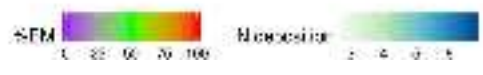
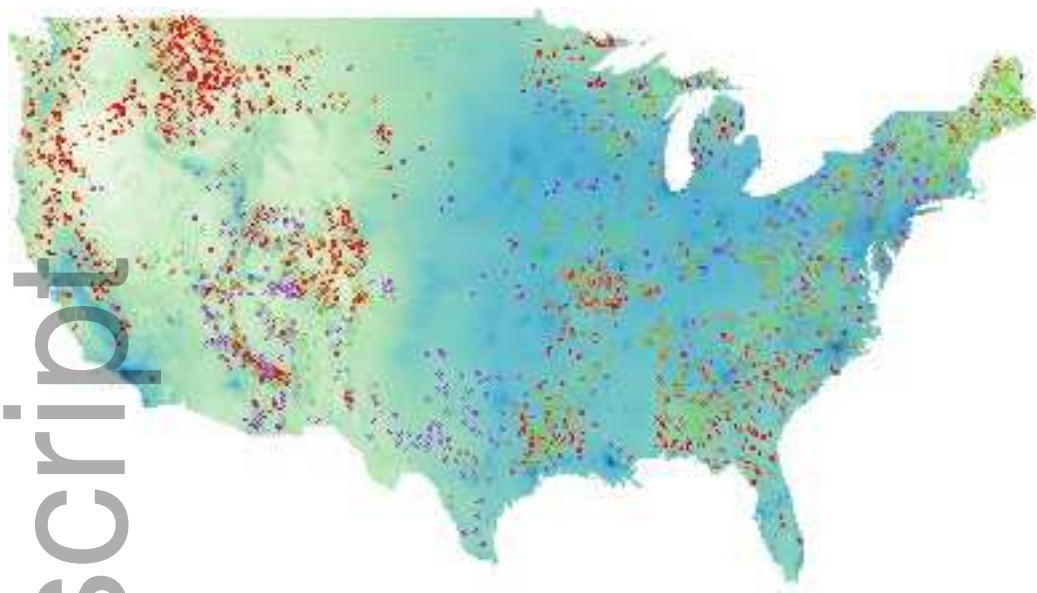
Figure Captions

Fig. 1. Distribution of sites used in all analyses, as well as the relative abundance of EM trees by basal area in each plot. All points were used in the analyses of the relative abundance of EM basal area and soils. Circles represent plots that were used in both the demographic analyses and the relative abundance analysis, while X's represent plots used in the relative abundance analysis only. N deposition data are based on the 2000-2014 mean wet + dry N deposition rates reported by the National Atmospheric Deposition Program. N deposition data are log transformed for visualization purposes only.

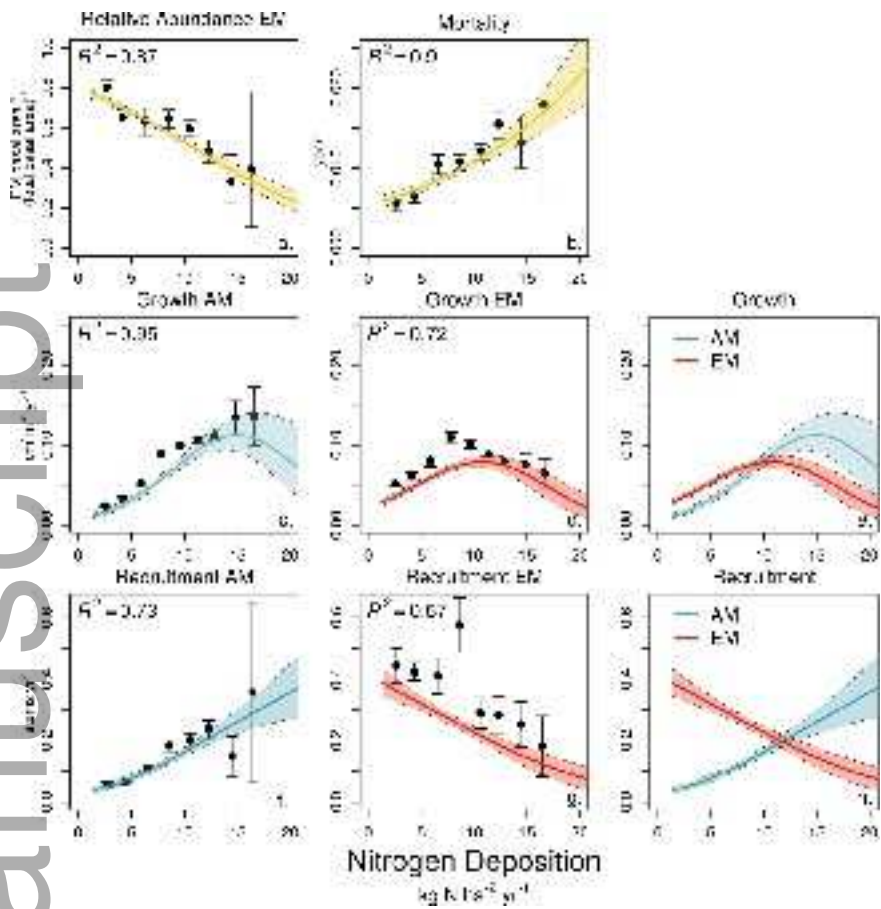
Fig. 2. Relationships between forest mycorrhizal abundance and mycorrhizal-specific demographic processes and N deposition. The relative abundance of EM trees is negatively correlated with N deposition (a.) N deposition had a positive effect on individual tree mortality rates, but there was no significant difference between AM and EM trees (b.) N deposition has a significantly more positive effect on arbuscular mycorrhizal (AM) basal area growth than EM basal area growth at the plot level (c., d.) such that these curves cross (e.) Nitrogen deposition had a positive effect on AM tree recruitment at the plot level but a negative effect on EM recruitment (f., g.) such that these curves cross (h.) Regression lines are based on the full

multiple regression output, holding all other predictors constant at their mean values. Shaded regions represent the 95% credible interval of the mean response. Points plotted are binned mean values and associated standard errors of soil C observations that have been detrended to remove variance associated with other predictors in the model. For additional detail see "Reporting of Results" in the Methods Section.

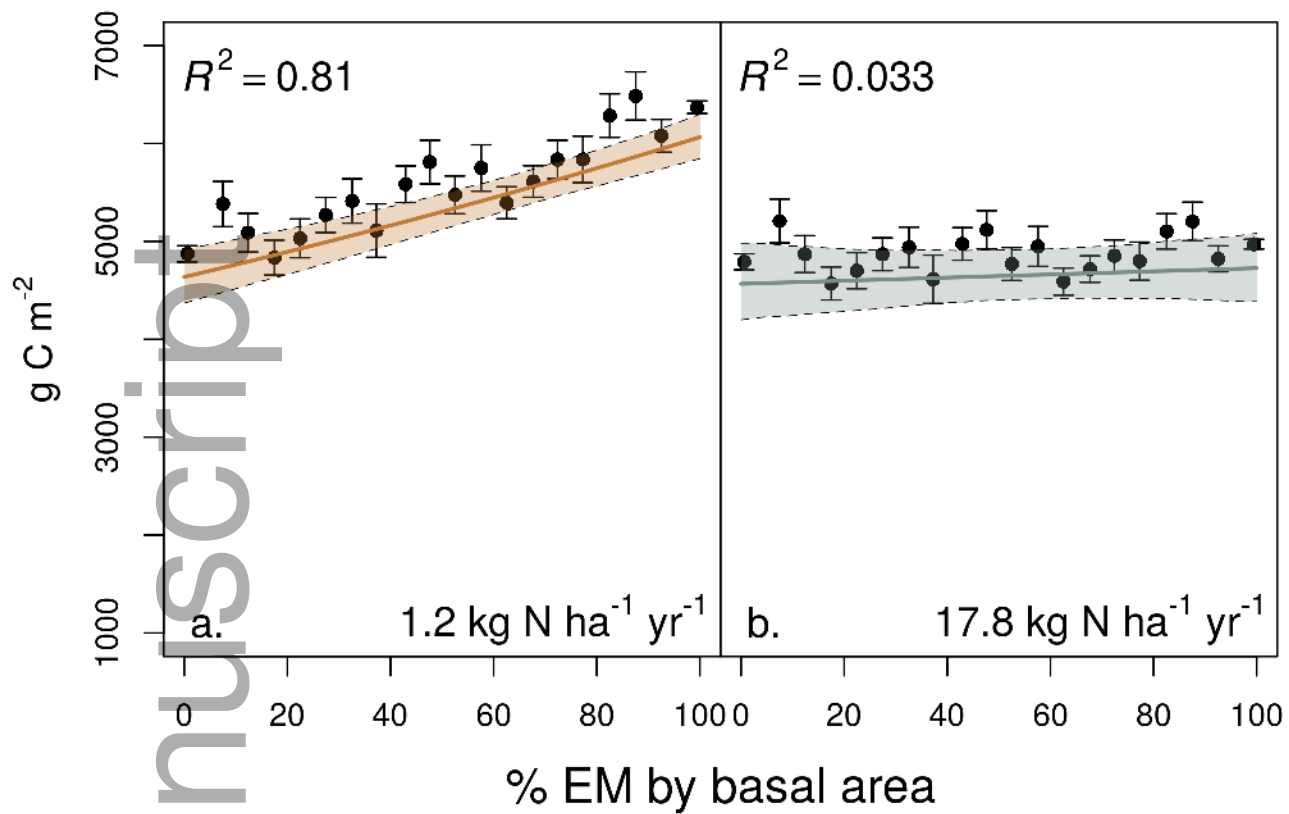
Fig. 3. Soil C storage as a function of EM relative abundance at low and high levels of N deposition. There is a positive relationship between soil C storage and the relative abundance of EM trees within a plot. However, there is an interaction between the relative abundance of EM trees and N deposition such that the effect is strong at low levels of N deposition (a.) but is completely absent at the highest levels of N deposition observed in the data set (b.) Regression lines are based on the full multiple regression output, holding all other predictors constant at their mean values. Shaded regions represent the 95% credible interval of the mean response. Points plotted are binned mean values and associated standard errors of soil C observations that have been detrended to remove variance associated with other predictors in the model. For additional detail see "Reporting of Results" in the Methods section.



gcb_14368_f1.png



gcb_14368_f2.png



gcb_14368_f3.png