Overall response to both reviewers

We greatly thank both reviewers for the thorough and very helpful reviews. Synthesizing the two reviews indicated that the manuscript had a lot of interesting information but was too dense to effectively communicate the key ideas. In response, we have simplified the analysis so that it has fewer moving parts. Our reanalysis also represents improvements to the data assimilation approach that have occurred since the manuscript was first submitted.

We simplified and modified the analysis as follows:

1) We removed the need for the 2-stage data assimilation. Now there are two chains that assimilate all sites simultaneously: one that includes site-specific parameters for only the Duke site and one that does not include the site-specific parameters. This modification allows the analysis to focus on why the parameters are different rather than focusing on the need to weight the Duke site differently. Since we did not actually weigh the Duke site differently in the original analysis nor include a synthetic experiment that explores the influence of site weighing on parameter inference, we feel that the simplified, updated approach is more sound and easier to understand.

   The previous text on the two-stage vs. one-stage assimilation is now condensed to the following:

   In Methods:

   During preliminary analysis, we found that the Base assimilation predicted lower stem biomass than observed in the elevated CO2 plots in the Duke FACE study. Further analysis investigating the cause of the bias in the CO2 plots showed that three parameters (wSx1000, ThinPower, and pCRS) were required to be unique to the Duke FACE study in order to reduce the bias. Therefore, the Base assimilation included unique parameters for wSx1000, ThinPower, and pCRS parameters in all plots in the Duke FACE and US-DK3 studies. To highlight the need for the site-specific parameters, we repeated the Base assimilation approach without the three additional parameters for the Duke studies (NoDkPars assimilation).

   In Results:

   The plots at the Duke Forest study had a higher carrying capacity of stem biomass before self-thinning (WSx1000), smaller self-thinning parameter (ThinPower), and lower allocation to coarse root (pCRS) than values optimized from the other plots across the region (Table 6). The DA approach without these three study specific parameters (NoDkPars) predicted significantly lower accumulation of stem biomass in response to elevated CO2 than observed (df = 4, p = 0.002; Figure 5). The NoDKPars assimilation optimized the CO2 fertilization parameter (fCalpha700) to a value that predicted 45% less light-use efficiency at 700 ppm (1.13 in NoDKPar vs. 1.33 in Base; Table 6) than the Base assimilation.

In Conclusions:

Constraining the sensitivity to atmospheric CO_2 *differs from constraining the sensitivity to ASW because, unlike the multiple constraints on water sensitivity (drought, irrigation, and gradient studies), environmental conditions created by the few elevated* CO_2 *plots provided unique constraint on parameters. Our finding demonstrated that DA efforts should test for bias in unique ecosystem experiments before finalizing a set of model parameters used in optimization. In particular, we found that the parameter governing the photosynthetic response to elevated* CO_2 *(fCalpha700) was substantially lower when all parameters were assumed to be shared across all plots than when the* CO_2 *fertilization experiment was allowed to have unique parameters. The need for the three unique parameters at the Duke FACE study parameters can be explained by the constraint provided by multiple data streams and multiple plots. An assumption of the model was that an increase in stem biomass caused a decrease stem density through self-thinning, unless the average tree stem biomass was below a parameterized threshold (WSx1000). Therefore, an increase in photosynthesis and stem biomass through* CO_2 *fertilization could cause a decrease in stem density. For a single study, it is straightforward to simultaneously fit the* CO_2 *fertilization and self-thinning parameters to fit stem biomass and stem density observations for the site. However, regional DA presents a challenge because the self-thinning parameters are well constrained by the stem biomass and stem density observations across the region but the* CO_2 *fertilization parameters are not. As a result of the regional DA, the self-thinning parameters caused a stronger decrease in stem density than observed in the Duke FACE study. Therefore, the optimization favored a solution where there was a lower response to* CO_2*, thus a smaller decrease in stem density. Allowing the Duke FACE study to have unique self-thinning parameters that resulted in lower rates of self-thinning and allowed for simulated stem biomass to respond to* CO_2 *in a way that matched the observations without penalizing the optimization by degrading the fit to the stem density.*

Our finding that the Duke FACE study required unique self-thinning parameters to reduce bias in the simulated stem biomass suggests that when using DA to optimize parameters that are shared across plots, careful examination of prediction bias in key sites that provide unique constraint on certain parameters (like the Duke FACE) is critical. Based on this example, we suggest that DA efforts using multiple studies and multiple experiment types identify whether particular experiments at limited number of sites have the potential to uniquely constrain specific parameters. In this case, additional weight or site-specific parameters may be needed to avoid having the signal of the unique experiment overwhelmed by the large amount of data from the other sites and experiments. Additionally, the finding suggests that multi-site DA should consider using hierarchical approaches to predicting mortality, particularly because mortality is often not simulated as mechanistically as growth. A hierarchical approach, where each plot has a set of mortality parameters that are drawn from a regional distribution, could avoid having unexplained variation in mortality rates lead to bias in the parameterization of growth related processes (i.e., growth responses to CO_2*, drought, nutrient fertilization, etc.). The hierarchical approach to mortality could also highlight patterns in mortality* *rates across a region and allow for additional investigations in the mechanisms driving the patterns.*

2) We replaced the assimilations that separately removed the water and nutrient experiments with a single assimilation that removes all experiments (water, nutrient, and CO_2). We feel this is a better approach because the analysis included multi-factor experiments. For example, in the previous analysis, the removal of nutrient experiments also removed CO_2 and drought treatments. Now we present two sets of optimized parameters: with and without experiments. This allows us to more clearly address the question "how do the parameter distributions depend on the inclusion of ecosystem experiments in the data assimilation". Some of the figures were simplified in the process of this revision.

   The previous text because the assimilation with and without experiment is now condensed to the following:

   Methods:

   We also evaluated how parameter distributions and the associated environmental sensitivity of model predictions depended on the inclusion of ecosystem experiments in data assimilation. First, we repeated the Base assimilation, this time excluding the plots that included the manipulated treatments (NoExp). We removed all manipulation types at once, rather than individual experimental types, because all experimental types were involved multi-factor studies. The NoExp assimilation had the same number of data streams as the Base assimilation because it included the control treatments from the experimental studies. The NoExp assimilation represented the situation where only observations across environmental gradients were available. Second, we compared the parameterization of the ASW, soil fertility, and atmospheric CO2 environmental modifiers from the Base to the NoExp assimilation. The modifiers equations are described in Supplemental Material Section 1.2 and 1.3. Third, we repeated the same independent validation exercise for the 160 FMRC plots as described above for the Base assimilation. Fourth, we predicted the treatment plots in the irrigated, drought, nutrient addition (only plots where FR was assumed to be 1), and elevated CO2 plots. As for the Base assimilation, we used a t-test to compare the experimental response between the NoExp assimilation and observed and between the NoExp and Base assimilations. Since the experimental treatments were not used in the optimization, this was an independent evaluation of predictive capacity.

   Results:

   Excluding the experimental treatments from the data assimilation did not strongly influence the predictive capacity of the model. The RMSE validation plots in NoExp assimilation decreased slightly compared to Base assimilation (21.8 to 18.0 Mg ha-1) while the bias slightly increased (-3.7 to -4.1%)(Figure 4b). Excluding the experimental treatments resulted in a significantly lower response of stem biomass to elevated CO2 than observed ($df = 4$, $p < 0.001$; Figure 5). Furthermore, there was a slight negative response of stem biomass to CO2 in the NoExp assimilation because the parameter *governing the change in foliage allocation at elevated CO2 (fCpFS700) was unconstrained by observations (Table 6). This led to convergence on the lower bound of the prior distribution (0.5) where foliage allocation decreased with increased atmospheric CO2. The predictions of irrigation, drought, and nutrient addition experiments were not significantly different between the Base and NoExp assimilations (Figure 5).*

The parameters and associated response functions in the 3-PG for nutrients, ASW, and atmospheric CO2 differed between the Base and NoExp assimilations (Figure 6). First, the parameterization of the soil fertility rating (FR) showed a stronger dependence on SI in the NoExp assimilation than in the Base assimilation (Figure 6a). For a given SI there was a lower FR, thus stronger nutrient limitation, when experimental treatments were excluded from assimilation. Second, the parameterization of the function relating photosynthesis and canopy conductance to ASW resulted in lower photosynthesis and maximum conductance when soil available water was less than 50% in the NoExp than Base assimilations (Figure 6b). Finally, the response of photosynthesis to atmospheric CO2 was functionally zero in the NoExp assimilation, thus highlighting the importance of the elevated CO2 treatments in the Duke FACE study for constraining the parameterization of the CO2 response function (Figure 6c).

Discussion:

The most important experimental manipulation for constraining model parameters was the Duke FACE CO2 fertilization study because the CO2 fertilization parameters (fCalpha700 and fCpFS700) converged on the lower bounds of their prior distributions when the experiments were excluded from the assimilation. In contrast, excluding the nutrient fertilization, drought, and irrigation studies did not substantially alter the predictive capacity of the model. This finding suggests that data assimilation using plots across environmental gradients alone can constrain parameters associated with water and nutrient sensitivity. However, regardless of whether the experiments were included in the assimilation, the optimized model predicted higher sensitivity to drought than observed, highlighting that future studies should focus on improving the sensitivity to drought.

The 3-PG model included a highly-simplified representation of interactions between the water and carbon cycles that resulted in parameterizations that may contain assumptions that require additional investigation. First, transpiration was modeled as a function of a potential canopy transpiration that occurred if leaf area was not limiting transpiration. The LAI at which leaf area was no longer limiting was a parameter that was optimized (LAIgcx in Table 5), resulting in a value of 2.2. Interestingly, this optimized value is consistent with the scant literature on this topic. In their analysis of multi-year measurements of transpiration in loblolly pine, Phillips and Oren (2001) observed that transpiration per unit leaf area was relatively insensitive to increases in leaf area above LAI of approximately 2.5. Iritz and Lindroth (1996) reviewed transpiration data from a range of crop species and found only small increases in transpiration above LAI of 3-4. These authors suggest that the threshold-type responses observed were related to the *range of LAI at which self-shading increases most rapidly, therefore limiting increases in transpiration. The resulting model behavior of "flat" transpiration above 2.2 LAI, with gradually decreasing photosynthesis above that value, results in increasing water use efficiency at higher LAI values. Second, the relationship between relative ASW and the modifier of photosynthesis and transpiration predicted a modifier value greater than zero when the relative ASW was zero. This resulted in positive values from photosynthesis and transpiration when the average ASW during the month was zero. In practice, the monthly ASW was rarely zero during simulations, which presents a challenge constraining the shape of the ASW modifier. The priors for the two ASW modifiers (SWconst and SWpower) had ranges that permitted the modifier to be zero. Therefore, additional data are likely needed during very dry conditions to develop a more physically based parameterization. Alternatively, the parameterization of a non-zero soil moisture modifier at zero ASW may be due to trees having access to water at soil depths deeper than the top 1.5 m of soil represented by the bucket in 3-PG. Overall, it is important to view the parameterization presented here as a phenomenological relationship that is consistent with observations from drought and irrigation experiments as well as observations across regional gradients in precipitation.*

3) In response to Reviewer #2, we evaluated how well the model predicts the different experimental types. We now have a figure showing the observed and modeled experimental treatment responses for the data assimilation approaches. In the case of the data assimilation approach that did not include the experimental treatments, the comparison to the observed treatment responses are an independent validation of the model. We found that the data assimilation approach without the experiments predicts the experimental responses reasonably well, except for the CO_2 experiment.

   Reflecting this comment, the results section has been modified to the following:

   In Methods Section 2.4

   Finally, we compared the predicted responses to experimental manipulation to the observed responses. We focused the comparison on the percentage difference in stem biomass between the control and treatment plots. We used a paired t-test to test for differences between the predicted and observed responses within an experimental type (irrigated, drought, nutrient addition, and elevated CO2). We combined the single and multi-factor treatments for analysis. For the analysis of the nutrient addition studies we only used plots where FR was assumed to be 1 so that we were able to simulate the treatments without requiring the optimization of a site-specific FR parameter.

   In Methods Section 2.6

   Fourth, we predicted the treatment plots in the irrigated, drought, nutrient addition (only plots where FR was assumed to be 1), and elevated CO2 plots. As for the Base assimilation, we used a t-test to compare the experimental response between the NoExp assimilation and observed and between the NoExp and Base assimilations. Since the *experimental treatments were not used in the optimization, this was an independent evaluation of predictive capacity.*

In Result section 3.1

Furthermore, the response of stem biomass to irrigation (df = 7, p = 0.18), nutrient addition (df = 26, p = 0.29), and elevated CO2 (df = 4, p = 0.43) was not significantly different between the observed and the Base assimilation (Figure 5). The Base assimilation was significantly more sensitive to drought than observed (n = 31, p < 0.001; Figure 5).

In Results Section 3.2

Excluding the experimental treatments from the data assimilation did not strongly influence the predictive capacity of the model. The RMSE validation plots in NoExp assimilation decreased slightly compared to Base assimilation (21.8 to 18.0 Mg ha-1) while the bias slightly increased (-3.7 to -4.1%)(Figure 4b). Excluding the experimental treatments resulted in a significantly lower response of stem biomass to elevated CO2 than observed (df = 4, p < 0.001; Figure 5). Furthermore, there was a slight negative response of stem biomass to CO2 in the NoExp assimilation because the parameter governing the change in foliage allocation at elevated CO2 (fCpFS700) was unconstrained by observations (Table 6). This led to convergence on the lower bound of the prior distribution (0.5) where foliage allocation decreased with increased atmospheric CO2. The predictions of irrigation, drought, and nutrient addition experiments were not significantly different between the Base and NoExp assimilations (Figure 5).

4) In response to Anthony Walker's helpful suggestion, we added an additional focus on regional predictions by simulating the regional response to nutrient addition, elevated CO_2, and drought. Our new analysis goes beyond the previous analysis by propagating the parameter uncertainty for all HUC12 units in the Southeastern U.S.

   Reflecting this comment, the results section has been modified to the following:

   In Methods Section 2.6

   To demonstrate the capacity of the data assimilation system to create regional predictions with uncertainty, we simulated the regional response to a decrease in precipitation, an increase in nutrient availability, and an increase in atmospheric CO2 concentration, each as a single factor change from a 1985-2011 baseline. Each prediction included uncertainty by integrating across the parameter posterior distributions using a Monte-Carlo sample of the parameter chains. Our region corresponded to the native range of loblolly pine and used the HUC12 (USGS 12-digit Hydrological Unit Code) watershed as the scale of simulation. For each HUC12 in the region we used the mean SI, 30-year mean annual temperature, ASW aggregated to the HUC12 level, and monthly meteorology from Abatzoglou (2013) as inputs (Figure 3).

The SI of each HUC12 was estimated from biophysical variables in the HUC12 using the method described in Sabatia and Burkhart (2014). This SI corresponded to an estimated SI for stands without intensive silvicultural treatments or advanced genetics of planted stock.

To sample parameter uncertainty, we randomly drew 500 samples from the Base assimilation MCMC chain and simulated forest development from a 1985 planting to age 25 in 2011 in each HUC. We choose age 25 as the final age because it is a typical age of harvest in the region. For each sample, we repeated the regional simulation with 1) a 30% reduction in precipitation, 2) FR set to 1, and 3) atmospheric CO2 increased by 200 ppm. Within a parameter sample, we calculated the percentage change in stem biomass at age 25 between control simulation and three simulations with the environmental changes. We focused our regional analysis on the distribution of the percent change in stem biomass.

In Results Section 3.3

Regionally (i.e., the native range of loblolly pines), stem biomass at age 25 ranged from 52 Mg ha-1 to 292 Mg ha-1 with the most productive areas located in the coastal plains and the interior of Mississippi and Alabama (Figure 7a). The least productive locations were the western and northern extents of native range. The width of the 95% quantile interval for each HUC12 unit ranged from 6.2 to 29.8 Mg ha-1 with largest uncertainty located in most the productive HUC12 units and in the far western extent of the region (Figure 7b).

The predicted change in stem biomass at age 25 associated with an additional 200 ppm of atmospheric CO2 over the 1985-2011 levels was similar to the change associated with a removal of nutrient limitation (by setting FR = 1) (Figure 8a,c). The median change associated with elevated CO2 for a given HUC12 unit ranged from 19.2 to 55.7% with a regional median of 21.7% (Figure 8a). The change associated the removal of nutrient limitation ranged from 6.9 to 303.7% for a given HUC12 unit, with regional median of 24.1% (Figure 8b). The response to elevated CO2 was more consistent across space than the response to nutrient addition. The largest potential gains in productivity from nutrient addition were predicted in central Georgia, (Figure 3), the northern extent of the region, and the western extents, areas with the lowest SI (Figure 3).

Stem biomass was considerably less responsive to a 30% decrease in precipitation. The median change in stem biomass when precipitation was reduced from the 1985-2011 levels ranged from -11.6 to – 0.1% for a given HUC12 unit with a regional median of -5.1% (Figure 8c). Central Georgia was the most responsive to precipitation reduction reflecting the relatively low annual precipitation and warm temperatures (Figure 3).

For a given location, the predicted response to elevated CO2 had larger uncertainty than the predicted response to precipitation reduction and nutrient limitation removal (Figure 8c,d,f). The uncertainty, defined as the width of the 95% quantile interval, was consistent across the region for the response to elevated CO2 (Figure 8b). The uncertainty in the *response to precipitation reduction and nutrient limitation removal was largest in the regions with the largest predicted change (Figure 8df).*

5) Our discussion section is re-worked to reflect the simplified analysis described above.

   The discussion has the following paragraphs
   - *An overall of the findings (same paragraph as reviewed draft)*
   - *A paragraph about the hierarchical Bayesian approach (this paragraph is largely from the old methods section, as recommended by Walker)*
   - *A paragraph discussing that the CO2 response depended most strongly on the inclusion of ecosystem experiments in the data-assimilation*
   - *A paragraph discussing how the model predicted stronger sensitivity to drought than observed and what could be issues with the model*
   - *A paragraph discussing why unique parameters were needed for the Duke forest studies to get the CO2 fertilization response correct. (this is a more enriching discussion than the discussion about the 1-stage vs. 2-stage data-assimilation in the previous version*
   - *A paragraph about caveats associated with the regional simulations.*

6) In response to comments by both reviewers to justify the set of parameters that were fit, we added six more parameters to the assimilation. We also removed the confusing reference to a sensitivity study of model parameters (the methods describing it were buried in the footnote of a table)

   See table 3 for the parameters

7) Sub-sections were added throughout to improve clarity.

   The sub-sections follow the three objectives:

   1) to present and evaluate a new DA approach that integrates diverse data from multiple locations and experimental treatments with an ecosystem model to estimate the probability distribution of model parameters, 2) to examine how the predictive capacity and optimized parameters differ between an assimilation approach that only uses environmental gradients and an assimilation approach that uses both environmental gradients and ecosystem manipulations, and 3) to demonstrate the capacity of the DA approach to predict, with uncertainty, regional forest dynamics by simulating how forest productivity responds to drought, nutrient fertilization, and elevated atmospheric CO2 across the Southeastern U.S.

8) We fixed some minor issues with the model structure as follows
   a. The density independent mortality now removes all the biomass of an average individual rather than a proportion of an average individual. This was

accomplished by not using the parameter mS (the proportion of an average individual that is lost through turnover) in the density independent mortality calculation. Since density independent mortality represents random mortality it is more reasonable to not use mS in the calculation.

   b. The model now simulates throughfall experiments directly rather than just reducing rain. Now rain is intercepted by the canopy in the full amount but the rain that enters the soil is reduced when simulating the throughfall experiment. This is a small change that makes the comparison cleaner.
   c. FR is set to 1 in the fertilization studies that added nutrients at regular intervals. Many of these experiments were designed to fertilize to optimal nutrition so the assumption is well grounded and helps reduce the number of site level FR parameters that need to be optimized.
   d. The process error terms are allowed to be a linear function of the prediction. This allows for the uncertainty to increase with the magnitude of the prediction. This linear function is applied to stem biomass, GEP, and ET. It allows for more confidence in predictions of lower values (like winter GEP and ET).

9) There were improvements to the data assimilation algorithm under the hood that allowed for faster run times and convergence. The cost function did not change (though we have described the cost function more clearly in the text).

Overall, the updated manuscript is more streamlined (though with more explanation in the methods section) and represents the state-of-the-art for the DAPPER algorithm.

Specific responses Reviewer #1 (Walker) below
Our responses are in italics

Thomas et al present a data-assimilation (DA) study using constraints from multiple data streams from multiple sites and experiments to optimise parameters in the monthly timestep PG-3 model of loblolly pine production. The study has three specific objec- tives. Stated on lns 170-171, 1) a new regional and hierarchical data assimilation sys- tem with the capacity to assimilate multiple data streams from multiple experiments; stated on ln 179-180, 2) the consequences for parameter estimation and prediction of including or not including ecosystem manipulation experiments (this could be more broadly stated as evaluation of the DA); and stated on ln 181 3) model predictions with the optimised parameter set of forest biomass changes in response to changes in nutrient addition of precipitation. This study is well thought out and implemented, presents a useful advance to the use of DA in ecosystem modelling and forecasting, and will likely be of interest to many readers of Biogeosciences.

My main criticism is that the distinction between the three areas of this study is often not made explicitly throughout the manuscript and consequently the manuscript is not as readable or as clear as it could be.

The majority of my comments are an attempt to help improve the organisation and presentation of the manuscript with the goal that this study will be as widely read and cited as possible.

- With that in mind, I suggest organising the manuscript as much as possible by the three stated objectives. I suggest combining the sentence on lns 179-180 with the sentence on lns 170-171 and explicitly listing the three objectives together. The results and discussion section would benefit from organisation along the lines of the three stated objectives. I suggest breaking each into three subsections, each dealing with one of the objectives. Again the conclusions section should specifically address each objective.

  We have modified the structure of the manuscript so that the methods, results, and conclusion now have sections that address each of the three objectives.

Abstract

- It would be good to be specific about who the target audience is for this research. The research straddles a technical field that develops DA but the technique produces a tool at a level of maturity that could be used by foresters. These ultimate end users could be more explicitly targeted.

  The following text has been added:

  "Overall, we 1) demonstrated how three decades of research in southeastern U.S. planted pine forests can be used to develop data assimilation techniques that use multiple locations, multiple data streams, and multiple ecosystem experiment types to optimize parameters and 2) developed a tool for creating future predictions of forest productivity

for natural resource managers that are consistent with a rich history of ecosystem research across a region."

Introduction

- Is a bit long and could a page or so could be cut without loss of content. Paragraphs on lns 82-105 could be combined and reduced in length. The main point is that ecosystem experiments can help to reduce the problem of equifinality in DA.

  The paragraphs between 82 and 105 were shortened. There is now a single paragraph that is the following:

  "Using DA to parameterize ecosystem models with observations from multiple locations that leverage ecosystem manipulation experiments and environmental gradients will allow for predictions to be consistent with the rich history of global change research in forest ecosystems. Ecosystem manipulation experiments provide a controlled environment in which data collected can be used to describe how forests acclimate and operate under altered environmental conditions (Medlyn et al., 2015) and can potentially allow for the optimization of model parameters associated with the altered environmental factor in the experiment. Furthermore, the assimilation of data from ecosystem manipulation experiments may increase parameter identifiability (reducing equifinality (Luo et al., 2009)), where two parameters have compensating controls on the same processes, by isolating the response to a manipulated driver. Observations that span environmental gradients include measures of forests ecosystem stocks and fluxes across a range of climatic conditions, nutrient availabilities, and soil water dynamics. These studies leverage time and space to quantify the sensitivity of forest dynamics to environmental variation. However, covariation of environmental variation can pose challenges separating the responses to individual environmental factors. Overall, assimilating observations from a region that includes environmental gradients and manipulation experiments is a useful extension of prior DA research focused on DA at a single site with multiple types of observations (Keenan et al., 2012; Richardson et al., 2010; Weng and Luo, 2011).

- The paragraph on lns 108-141 makes some nice points but could be substantially shortened without loss of content. Much of the paragraph is methods like.

  The paragraph in the comment has been combined with the prior paragraph which is provided above.

- Weight to rare experiments (mentioned on ln 125) could also apply to rare data types. Later in the paragraph (ln 135-136) the authors state that data of different frequency is a problem in biasing the cost function toward high frequency data, but offer no solution other than a monthly timestep model. Rare data, or low frequency data, could also be given higher weights. Also high frequency data could be summarised at lower frequency.

  The discussion of the data weighting was removed.

Methods

Again long and could probably be made more concise. Also the organisation is tough to follow.

- I suggest leading with the observations, the various sites, and measurement campaigns/projects.Many of these are not properly introduced. This will provide a comprehensive introduction to the system and what measurements actually go into this DA system. Observation sites and projects are mentioned on ln 409-410, but these are not introduced and need to be described in the observations section of the methods.

  We moved the section on the observations to the beginning of the methods section. We structured the paragraph so that it gives an overview of all the measurement campaigns. The observations section (Section 2.1 is as follows)

  We used thirteen different data streams from 294 plots at 187 unique locations spread across the native range of loblolly pine trees to constrain model parameters (Table 1; Figure 1). The data streams covered the period between 1981 to 2015. The Forest Modeling Research Cooperative (FMRC) Thinning Study provides the largest number of plots that span the region (Burkhart et al., 1985). In this study, we only used the control plots that were not thinned. The Forest Productivity (FPC) Cooperative Region-wide 18 (RW18) study included control and nutrient fertilization addition plots that span the region (134.4 kg ha-1 N + 13.44 kg ha-1 P biannually) (Albaugh et al., 2015). The PINEMAP study included four locations dispersed across the region that included a replicated factorial experiment with control, nutrient fertilization (224 kg ha-1 N + 27 kg ha-1 P + micronutrients once at project initiation), a throughfall reduction (30% reduction), and fertilization by throughfall treatments (Will et al., 2015). The SETRES study was located at a single location and included replicated control, irrigation (~650 mm of added water per year), nutrient fertilization (~100 kg N ha-1 + 17 kg P ha-1 with micronutrients applied annually with absolute amount depending on foliar nutrient ratios), and fertilization by irrigation treatments (Albaugh et al., 2004). The Waycross study was a single site with a non-replicated fertilization treatment. The annual application of fertilization focused on satisfying the nutrient demand by the trees was one of the most productive stands in the region (Bryars et al., 2013). These five studies included data streams of stand stem biomass (defined as the sum of stemwood, stembark and branches) and live stem density. Waycross and SETRES included LAI measurements from litterfall traps (Waycross) or estimates from LICOR LAI-2000 (SETRES). SETRES also included fine root and coarse root measurements. In the PINEMAP, SETRES, and RW18 studies we only used foliage biomass estimates from the control plots. We excluded the foliage biomass estimates from the treatment plots because they were derived from allometric models that may not have captured changes in allometry due to the experimental treatment. We did use LAI measurements from both control and treatment plots where available (SETRES).

  We also included observations from the Duke FACE study where the atmospheric CO2 was increased by 200 ppm above ambient concentrations. Based on the data presented in McCarthy et al. (2010) the study included six control plots, four CO2 fumigated rings *(including the unfertilized half of the prototype), two nitrogen fertilization treatments (115 kg N ha-1 yr-1 applied annually) , and one CO2 by nitrogen addition treatment (fertilized half of prototype). The Duke FACE study included observations of stem biomass (loblolly pine and hardwood), coarse root biomass (loblolly pine and hardwood), fine root biomass (combined loblolly pine and hardwood), stem density (loblolly pine only), leaf turnover (combined loblolly pine and hardwood), fine root production (combined loblolly pine and hardwood), and monthly LAI (loblolly pine and hardwood).*

Finally, we included two Ameriflux sites with eddy-covariance towers in loblolly pine stands. The US-DK3 site was located in the same forest as the Duke FACE site described above (Novick et al., 2015). The US-NC2 site was located in coastal North Carolina (Noormets et al., 2010). We used monthly gross ecosystem production (GEP; modeled gross primary productivity from net ecosystem exchange measured at an eddy-covariance tower) and evapotranspiration (ET) estimates from the sites. The monthly GET and ET were gap-filled by the site PI. The GEP was a flux partitioned product created by the site PI. The biometric data from the US-DK3 site was assumed to be the same as the first control ring. The biometric data from the US-NC2 site included of stem biomass (loblolly pine and hardwood), coarse root biomass (loblolly pine and hardwood), fine root biomass (combined loblolly pine and hardwood), stem density (loblolly pine only), leaf turnover (combined loblolly pine and hardwood), and fine root production (combined loblolly pine and hardwood).

- I found section 2.3 very difficult to follow. I'm not expert on DA mathematical meth- ods but I have a reasonable conceptual handle on DA, and yet I was lost in the first paragraph. I also ran this section by a colleague who is expert in the mathematics underpinning DA and they agreed that this sections needs to be clearer. Their key criticism was that they could not see the derivation of Eq 7, perhaps the authors could add the derivation to an appendix. And that it is not clear how the MCMC was used to sample Eq 7. A clear description of the details of the MCMC procedure is necessary, along with the presentation of the cost function. Also the first term on the righthand side of Eq 7 is not the same as the righthand side of Eq 1, is this deliberate? And E is never defined.

  I strongly suggest reworking section 2.3 of the methods to be extremely clear about the DA process and how it was implemented. Start with a clear description of the goals of the DA – state estimation and estimation of parameter distributions. Then describe all the various sources of uncertainty and how the method accounts for them. Then take the reader step by step through the method. Perhaps a diagram would be useful. The following comments are an attempt to provide examples of where confusion arises but they are in no way comprehensive. The sentence on lns 281-283 is more or less stating the the same thing as the sentence on lns 284-285. I suggest fusing these together. Is the reference to a "latent model" really necessary, it is confusing with the mathematical model. Would "true" system states and fluxes convey the same meaning? Do not try to justify the method in comparison with previous methods (e.g. lns 286-291), in the methods this just confuses the description and this can be argued in the discussion. On lns 291-293, this is state estimation right? That's fine but is it really the focus of your method? None of the three stated objectives are for state estimation. How exactly was estimation of the latent state or flux the first step in the process when it includes the optimised parameters etc as described on lns 296-298? Seems like the statement on ln 306-308 should come before the previous paragraph.

We cleaned up the description of the cost function per the reviewer recommendation. (see Supplement to the review)

We used a hierarchal Bayesian framework to estimate the posterior distributions of parameters, latent states of stocks and fluxes, and process uncertainty parameters. The latent states represented a value of the stock or flux before uncertainty was added through measurement. The approach was as follows.

Consider a stock or flux (m) for a single plot (p) at time t ($q_{p,m,t}$). $q_{p,m,t}$ is influenced by the processes represented in the 3-PG model and a normally distributed model process error term,

$$q_{p,m,t} \sim N(f(\boldsymbol{\theta}, FR_p), \sigma_m) \qquad \textit{Equation 1}$$

where $\boldsymbol{\theta}$ is a vector of parameters that are optimized, FR_p is the site fertility, and σ_m is the model process error. Not shown are the vector of parameters that were not optimized (Supplemental Material Table 1), the plot ASW, an array climate inputs, and the initial conditions because these are assumed known and not estimated in the hierarchical model. The process error assumed that the error linearly scales with the magnitude of the prediction:

$$\sigma_m^2 = \gamma_m + \rho_m f(\boldsymbol{\theta}, FR_p) \qquad \textit{Equation 2}$$

While the structure of the Bayesian model allowed for all data streams to have process uncertainty that scales with the prediction, in this application we only allowed stem biomass, GEP, and ET process uncertainty to scale because they had large variation across space (stem biomass) and through time (i.e., there should be lower process uncertainty in the winter when GEP is lower). For the other data streams, the linear scaling term was removed by fixing ρ_m at 0.

FR_p did not have an explicit probability distribution. Rather the probability density evaluated to 1 if the plot was not fertilized, thus causing FR_p to be estimated from SI and MAT (Supplemental Material Equation 15), or if it was a fertilized plot and has an FR_p equal or higher than that of its non-fertilized control plot. The probability density evaluated to 0 if the estimated FR_p in a fertilized plot was less than the FR_p in the control plot or FR_p awas not contained in the interval between 0 and 1.

$$FR_p \sim \begin{cases} 1 \text{ if non-fertilized, } FR_p \geq 0, \text{ and } FR_p \leq 1 \\ 1 \text{ if } FR_p = 1 \text{ and fertilization levels are assumed to remove nutrient deficiencies} \\ 0 \text{ if } FR_p < 1 \text{ and fertilization levels are assumed to remove nutrient deficiencies} \\ 1 \text{ if fertilized but levels are not assumed to remove deficiencies and } FR_p \geq FR \text{ of control plot} \\ 0 \text{ if fertilized but levels are not assumed to remove deficiencies and } FR_p < FR \text{ of control plot} \\ 0 \;\; if\; FR_p < 0 \text{ or } FR_p > 1 \end{cases}$$

Equation 3

Our model included the effect of observational errors for measurements of stocks and fluxes. For a single stocks or flux for a plot at time t there is an observation ($y_{p,m,t}$). The normally distributed observation error model was:

$$y_{p,m,t} \sim N(q_{p,m,t}, \tau^2_{p,m,t}) \qquad \text{Equation 4}$$

where $\tau^2_{p,m,t}$ represented the measurement error of the observed state or flux. By including the observational error model, $q_{p,m,t}$ represented the latent, or unobserved, stock or flux. The variance was unique to each observation because it was represented as a proportion of the observed value. The $\tau^2_{p,m,t}$ was assumed known (see Table 2) and not estimated in the hierarchical model (Table 2).

The hierarchical model required prior distributions for all optimized parameters, including the parameters for the 3-PG model (θ), FR_p, and the process error parameters. The prior distributions for $\boldsymbol{\theta}$ are specified in Table 3. Some parameters were informed by previous research in loblolly pine ecosystems while other parameters were 'non-informative' with flat distributions (termed 'vague' in Table 3). The prior distributions *for the process error parameters were non-informative and had a uniform distribution with upper and lower bounds that spanned the range of reasonable error terms.*

$\gamma_m \sim U(0.001,100)$ *Equation 5*

$\rho_m \sim U(0,10)$ *Equation 6*

By combining the data, process, and prior models, our joint posterior that includes all thirteen data streams, plots, months with observations, and fitted parameters was

$$p(\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{q}|\boldsymbol{y},\boldsymbol{\tau},priors) \propto$$

$$\prod_{p=1}^{P}\prod_{m=1}^{M}\prod_{t=1}^{T} N(q_{p,m,t}|f(\boldsymbol{\theta},FR_p),\gamma_m + \rho_m f(\boldsymbol{\theta},FR_p))$$

$$\prod_{p=1}^{P}\prod_{m=1}^{M}\prod_{t=1}^{T} N(y_{p,m,t}|q_{p,m,t},\tau_{p,m,t}^{2})$$

$$\prod_{p=1}^{P} p(FR_p)\prod_{f=1}^{F} p(\theta_f)\prod_{m=1}^{M} p(\gamma_m)\prod_{m=1}^{M} p(\rho_m)$$

Equation 7

where bolded components represent vectors, P is the total number of plots, M is the total number of data streams, T is the total months with observations, and F is the total number of 3-PG parameters that are optimized.

We numerically estimated the joint posterior distribution using the Monte-Carlo Markov Chain – Metropolis Hasting (MCMC-MH) algorithm (Zobitz et al., 2011). This approach has been widely used to approximate parameter distributions in ecosystem DA research (Fox et al., 2009; Trudinger et al., 2007; Williams et al., 2005; Zobitz et al., 2011). Briefly, the algorithm proposes new values for the model parameters, uncertainty parameters, latent states, and FR. The proposed values were generated using a random draw from a normal distribution with a mean equal to the previously accepted value for that parameter and standard deviation equal to the parameter-specific jumping size. The ratio of proposed calculation of Equation 7 to the previously accepted calculation of Equation 7 was used to determine if the proposed parameters are accepted. If the ratio was greater than or equal to 1 the proposed values were always accepted. If the ratio was less than 1, a random number between 0 and 1 was drawn and the proposed values *are accepted if the ratio was greater than the random number. This allowed less probable parameter sets to be accepted, thus sampling the posterior distribution. We adapted the size of the jump size for each parameter to ensure the acceptance rate of the parameter set was between 22% and 43% (Ziehn et al., 2012) by adjusting the jump size if the acceptance rate for a parameter is outside the 22 – 43% range. All MCMC-MH chains were run for 30 million iterations with the first 15 million iterations discarded as the burn-in. Four chains were run and tested for convergence using the Gelman–Rubin convergence criterion, where a value for the criterion less than 1.1 indicated an acceptable level of convergence. We sampled every 1000th parameter in the final 15 million iterations of the MCMC-MH chain and used this thinned chain in the analysis described below. The 3-PG model and MCMC-MH algorithm were programed in FORTRAN 90 and used OpenMP to parallelize the simulation of each plot within an iteration of the MCMC-MH algorithm.*

- Section 2.4 jumps around between objectives. Some text would fit better in section 2.3, for example lns 408-428. Text on lns 454-461 would be better organised if it were to follow the text on 430-444, then the regional simulations can be presented afterwards.

  We reorganized as suggested by the reviewer. Section 2.1 is the observations, Section 2.2 is the Ecosystem Mode, Section 2.3 is the data assimilation method, Section 2.4 is the data assimilation evaluation, Section 2.5 is the Sensitivity to the inclusion of ecosystem experiments, Section 2.6 is the Regional predictions with uncertainty.

- I suggest defining sections 2.3, 2.4, and an additional 2.5 to be organised by the three stated objectives.

  We reorganized as suggested by the reviewer. Section 2.1 is the observations, Section 2.2 is the Ecosystem Mode, Section 2.3 is the data assimilation method, Section 2.4 is the data assimilation evaluation, Section 2.5 is the Sensitivity to the inclusion of ecosystem experiments, Section 2.6 is the Regional predictions with uncertainty.

- Also, while commonly used by the modelling community, I do not agree that you can run “experiments” with models. Models make predictions from a specific set of mathematical hypotheses and defined scenarios. An experiment is designed to test predictions and discriminate among hypotheses.

  We removed the ‘experiments’ language

Results

- Why were only 31 parameters optimised, can you describe why this set were chosen from the total 46?

In the revised manuscript, we included more parameters that were optimized (six more). The eight parameters that were not optimized did not have specific data to use as a constraint (leaf boundary layer, conductance, canopy light extinction coefficient, etc).

- Technically the parameters are not "sensitive" (ln 480), it is the model output that is sensitive to the parameter. "Influential" would be a better adjective to describe the parameters.

  To simplify the analysis and reduce the density of the manuscript we removed the sensitivity study and the reference to it in the text.

- Lns 486 & 488 variability is described as being reduced but no data are provided. Can you quantify these statements. There are many statements like this throughout the results and they ought to be quantified (e.g. lns 502, 508). Also on 508, is mean correct, isn't this the median of the parameter distribution?

  We added a column to the table that is the ratio of the size of the posterior 99% credible interval to the size of the prior 99% confidence interval. This ratio illustrates how the uncertainty is reduced by the data assimilation.

- Some kind of visual representation of the data in table 5 would be useful.

  Supplemental Material Figure 1 shows the PDF of the prior and posterior

- Ln 492 what do you mean strong priors? Well defined from measurements and litera- ture with low variance? Could you quantify this?

  We removed this language from the manuscript to reduce confusion

- Lns 494 the process uncertainty parameters are mentioned here and in the methods, but results are barely presented (only in the supplement) and are not discussed, or not that I noticed. This is a very interesting concept and I would like to see these data pre- sented a little more and at least a little discussed. What kind of impact does including these parameters have on the optimised parameter distributions? I understand you are already presenting a lot, but this is fairly novel as far as I'm aware and is of interest.

  We added a small discussion of the process error parameters to the discussion section

- Figure 10 and 11 would be more in keeping with your stated goal of forecasting on lns 65-68 if you removed the b panels in both plots. If you think that the parameter estimates when including the data from the manipulations gives a better estimate of those parameters then the data in panels b are not particularly useful for forecasts. In my view, and as stated on line 67 & 68 "provide information on both the expected future state of the forest and the probability distribution of those future states", the final figures would be much stronger if the probability distribution of the future states shown on the a panels were represented on the b panels.

We combined the Figure 10 and 11 into a single figure that has the median prediction on the left side and the uncertainty on the right side. This allows the figures to represent the forecasting capacities of the data assimilation approach. The paragraph is as follows:

Our hierarchal approach (Equation 7) was designed to partition uncertainty that is attributable to uncertainty in parameters, model process, and measurements (Hobbs and Hooten, 2015). Previous forest ecosystem DA efforts have either focused on parameter uncertainty, by using measurement uncertainty as the variance term in a Gaussian cost function (Bloom and Williams, 2015; Keenan et al., 2012; Richardson et al., 2010) or on total uncertainty by directly estimating the Gaussian variance term (Ricciuto et al., 2008). The latter combines measurement uncertainty and process uncertainty into the same parameter and is unable to be used for developing prediction intervals, as prediction intervals only include parameter and process errors (Dietze et al., 2013; Hobbs and Hooten, 2015). Our approach allows the estimation of the probability distribution of forest biomass before uncertainty is added through measurement. Considering that the method of assimilation can potentially have a large influences on posterior parameter distributions (Trudinger et al., 2007), future research should focus on comparing the hierarchal approach presented here to other approaches by using the same data constraints with alternative cost functions.

- While it is interesting to show the consequences for prediction of inclusion of manipulations or not, and the opposite sign of the change in predictions when water and nutrient manipulations are included, you already show this in Figures 6 & 7. If you want to keep the b panels in 10 and 11 I suggest you add them as extra panels to figures 6 & 7, showing the absolute delta (or similar) from the simulations that include the manipulation delta. This will allow you to address the question: what are the consequences of not including data from manipulations? Without confounding the predictions from the most appropriate DA product for the scenarios tested. Also, the scale ought to be the same for the data presented in Figs 10 and 11.

  We cut panel b from these figure.

- Was CO2 change included in the above projections of removal of nutrient limitation and precipitation reduction? Furthermore, it seems you have included data from water manipulation experiments, nutrient manipulation experiments, and CO2 manipulation experiments. But you have only made projections for nutrient and precipitation change. Why not CO2 change? CO2 projections would complete the study.

  We added a +200 ppm simulation to the set of regional predictions. The predicted regional changes are for +200 ppm, -30% precipitation, and removal of nutrient limitation. The uncertainty for each prediction is shown. This changed the description of the regional results to be the following:

Regionally (i.e., the native range of loblolly pines), stem biomass at age 25 ranged from 52 Mg ha-1 to 292 Mg ha-1 with the most productive areas located in the coastal plains and the interior of Mississippi and Alabama (Figure 7a). The least productive locations were the western and northern extents of native range. The width of the 95% quantile interval for each HUC12 unit ranged from 6.2 to 29.8 Mg ha-1 with largest uncertainty located in most the productive HUC12 units and in the far western extent of the region (Figure 7b).

The predicted change in stem biomass at age 25 associated with an additional 200 ppm of atmospheric CO2 over the 1985-2011 levels was similar to the change associated with a removal of nutrient limitation (by setting FR = 1) (Figure 8a,c). The median change associated with elevated CO2 for a given HUC12 unit ranged from 19.2 to 55.7% with a regional median of 21.7% (Figure 8a). The change associated the removal of nutrient limitation ranged from 6.9 to 303.7% for a given HUC12 unit, with regional median of 24.1% (Figure 8b). The response to elevated CO2 was more consistent across space than the response to nutrient addition. The largest potential gains in productivity from nutrient addition were predicted in central Georgia, (Figure 3), the northern extent of the region, and the western extents, areas with the lowest SI (Figure 3).

Stem biomass was considerably less responsive to a 30% decrease in precipitation. The median change in stem biomass when precipitation was reduced from the 1985-2011 levels ranged from -11.6 to – 0.1% for a given HUC12 unit with a regional median of -5.1% (Figure 8c). Central Georgia was the most responsive to precipitation reduction reflecting the relatively low annual precipitation and warm temperatures (Figure 3).

For a given location, the predicted response to elevated CO2 had larger uncertainty than the predicted response to precipitation reduction and nutrient limitation removal (Figure 8c,d,f). The uncertainty, defined as the width of the 95% quantile interval, was consistent across the region for the response to elevated CO2 (Figure 8b). The uncertainty in the response to precipitation reduction and nutrient limitation removal was largest in the regions with the largest predicted change (Figure 8df).

Additional points

- I think the title would benefit from the addition of "Loblolly Pine".

  Added to title

- Ln 50 Duke FACE experiment had 4 replicate plots, so where does the 5 come from on this line. An additional plot from the unreplicated prototype?

We removed the language from the abstract and later in the text we clarified that the replicated prototype was used (per the data reported in McCarthy et al. 2010)

- Ln 48 – 50 the sentence on this line would help flow if it were before the preceding sentence.

  Revised

- Ln 65 I don't think I would classify the three areas mentioned in the previous sentence as tools. They are more than tools, they are also knowledge.

  Removed the word 'tools' so that the sentence references the previous sentence terminology ('sources of information')

- Ln 67 What do you mean by "based on" here. Can probably delete. Also while I think your methods could be used for "forecasting" you don't really use the method in that sense.

  Removed the clause that contained 'based on'

- Ln 73 insert "can" in between "that generate"

  Fixed in text

- Ln 85 86 "carbon allocation and turnover" This is worded a little awkwardly

  Removed awkward language from text

- Ln 97-99 awkward way to start a paragraph.

  Paragraph was removed during the shortening of the introduction

- Ln 111 suggest replacing "important" with "useful" or something more descriptive

  Changed to 'useful'

- Ln 155-157 suggest replacing "nutrients" with "nutrient addition". Also suggest removing hyphens.

  Changed in text

- Ln 162-163 Awkward

  Removed 'available' to make less awkward

- Ln 171 Again I think you need to call out loblolly pine here

Changed in text when revision the statement of objectives

- Ln 175 The authors chosen acronym, in my view, somewhat undersells what they are doing. The DA method is hierarchical and considers data from multiple sites and of multiple different types. The acronym gives not indication of this and suggests that the DA method is only suitable for Pine Plantations. Of course it is the authors' choice though.

  Thank you for the suggestion to broaden the acronym. We kept the same acronym but changed the words to "Data Assimilation to Predict Productivity for Ecosystems and Regions" to emphasize the multi-site aspect of the DA.

- Ln 307 insert "considered" between "was a"

  Sentence was modified during revisions

- Ln 446 replace "regional" with "region"

  Changed in text

- Ln 522-524 I'm not sure what you mean here, could you clarify?

  Sentence removed during the revisions

- Ln 528 delete "a"

  Done

- Ln 576 replace "detangling" with "disentangling"

  Done

- Ln 582 I think "synthesised" would be a better word to use than "organised"

  Done

- Ln 591-591 I take your point about equifinality but can you really say this if predictions were not improved in some way? Just a thought. Is there a way that you can be sure that the mechanisms were correctly distinguished?

  We removed this sentence during revisions

- Ln 633-634 Agreed, but did your method strictly weight the data? Wasn't it more that the hierarchical method gave priority to the CO2 manipulation data?

  We removed this sentence during revisions

- Ln 646 replace “than” with “that”

  Done

- Ln 656 quantify this statement

  We removed this sentence during revisions

- Ln 662-663 this was news to me when I read this sentence. I think this would become clearer once the methods can be clarified as suggested above.

  We clarified in the method section. The method section more completely describes assumptions of the site index estimation. The following text was added to Section 2.6 in the methods:

  The SI of each HUC12 was estimated from biophysical variables in the HUC12 using the method described in Sabatia and Burkhart (2014). This SI corresponded to an estimated SI for stands without intensive silvicultural treatments or advanced genetics of planted stock.

- Ln 668 suggest changing “prior” to “previous”, just to maintain the meaning of prior in the Bayesian sense.

  Done

- Ln 673 you do not show any data on covariation of parameters.

  We removed this language

- Ln 676-680 I like this statement, makes a lot of sense. But is it most appropriate here? This point should be made clearly in the methods.

  Moved to methods

- Ln 685 suggest deleting “Multivariate Constructed Analogs (MACA)” it is not needed.

  Deleted

- Ln 692-697 This is a good point but I’m curious why the change in biomass in response to precipitation reduction was small given the large change in parameter values when water manipulations were included in the DA. Can you try to explain this based on the process hypotheses embedded in the model.

  We cut this sentence during revisions

- Ln 698 replace “reduced” with “reduction” Ln 707 insert “as a function of”

Done

- Ln 719 insert space in "fromadditional"

  Done

- Ln 760 While I'm sure the methods and tools developed by this study could be used for ecological forecasting, strictly speaking this study is not ecological forecasting. The third objective, which concerns optimised model predictions, is a scenario analysis rather than a forecast.

  We removed the term 'ecological forecast' from the sentence and changed to:

  DA is increasingly used for developing predictions from ecosystem models that include uncertainty estimation, due to its ability represent prior knowledge, integrate observations into the parameterization, and estimate multiple components of uncertainty, including observation, parameter, and process representation uncertainty (Dietze et al., 2013; Luo et al., 2011b; Niu et al., 2014).

- Ln 769 no need to cite Medlyn et al 2015 here

  Removed citation

Specific responses Reviewer #2
Our responses are in italics

Quinn Thomas et al. present a model-data fusion, or data assimilation, study that gathers 35 years of carbon cycle-related observations and manipulation experiments taken in Loblolly Pine ecosystems in the Southeastern US to optimize parameters of the 3-PG model within their new framework DAPPER. The authors examine the ability of the observations to constrain model parameters using a number of approaches for assimilating the different types of data, and they further examine the differences in model behavior/sensitivity and change in biomass stocks across the southeastern US as a result of the different experiments.

The authors have carried out an impressive and exhaustive collection of data for con- straining the 3-PG model in this study. This, and their investigation into different approaches for assimilating different types of data, in particular manipulation experimental data, make this study a noteworthy contribution to model–data assimilation literature in forested ecosystems, and therefore I would recommend publication in Biogeosciences. However, as it stands the manuscript is quite long and dense, which is understandable given the amount of detail that is required to present such a wide array of data and experiments. This being said, I recommend that the authors try to edit the article following some of the suggestions below (and their own views) to improve the clarity and readability of the text before this article is published.

- Overall, the objectives and key points of this study can get lost in the text. I think a few more sub-sections in the main text and supplementary, references and links between sections would help the reader to better follow and absorb the necessary amount of detail presented in the manuscript. I would also find it useful if the authors posed a few key scientific questions to help them highlight the main messages of the study.

  We clarified the last paragraph of the introduction to directly state the three objectives of the study. We also added section to the Methods, Results, and Discussion that parallel the objectives

  The objectives paragraph is as follows:

  Using loblolly pine plantations across the southeastern U.S as a focal application, our objectives are to 1) present and evaluate a new DA approach that integrates diverse data from multiple locations and experimental treatments with an ecosystem model to estimate the probability distribution of model parameters, 2) examine how the predictive capacity and optimized parameters differ between an assimilation approach that only uses environmental gradients and an assimilation approach that uses both environmental gradients and ecosystem manipulations, and 3) demonstrate the capacity of the DA approach to predict, with uncertainty, regional forest dynamics by simulating how forest productivity responds to drought, nutrient fertilization, and elevated atmospheric CO2 across the Southeastern U.S.

- Some sections in the methods could do with more explanation for why certain approaches were used (see comments below) or better links to the supplementary material, as I have just mentioned.

  See comments below for response

- The introduction and discussion are quite long and this can prevent some of the key points from being highlighted. I suggest the authors try to cut down the text where they see fit, including some sentences that essentially are repetitions of earlier statements.

  We cut the introduction and removed paragraphs

- The paragraphs in the results section could be separate sections with sub-headings in order to guide the reader, while at the same time the results could benefit from stronger links between each section, especially before line 522, in particular comparing the between the 1st and 2nd stages, or the different 2-stage approaches with the 1- stage approach. At the moment, the results section before line 522 is a bit fragmented, making it harder to weave together a coherent story that brings out the key points.

  We added sub-sections to the results section

- Reading this manuscript I found myself asking: What do you expect from each experiment/approach? What will you gain/lose? Which approach is the right approach, going forward? These questions were largely answered in the discussion, and therefore I have made a suggestion below that perhaps some of the results and discussion could be merged within the sub-sections suggested above. This is a personal style issue however.

  We hope that updated analysis and discussion section helps answer these questions more clearly. There are now sub-sections in the results and discussion that help provide continuity between the sections.

- Finally, the authors may consider cutting other sections of the discussion that are not fully pertinent to the results as the paper is already quite full of detail. I would like to stress that despite this suggestion I did find the discussion to be interesting and comprehensive, but I would like to see the key messages highlighted more and am concerned the length of the paper may overwhelm the reader.

  We have cut out the paragraphs that aren't directly related to the results. These include the paragraph about the connections to the Community Land Model and the paragraph about the connections to sap-flux measurements

Introduction

- Line 97: "relative contribution of each environmental control should be separated in order to correctly parameterize the sensitivity to changes in the environment". I agree to some extent but this is very hard to do and should we be separating each environmental control, as the interaction between different environmental changes may produce different outcomes than if each were treated separately? I would be interested to hear the authors thoughts on this and what they think the impact of assimilating manipulation experiments data separately has on their results.

Per reviewer #1 comment to shorten this paragraph, this sentence is now removed from the manuscript.

- Line 124-128: See previous studies Wutzler and Carvalhais (2014) and Section 2 of MacBean et al. (2016) for further discussion on debate of how to deal with the issue of weighting to account for the number of observations and/or using a multi-stage assimilation approach to address challenges of assimilating a diverse set of observations. Both issues are the subject of debate in the literature. On the issue of weighting by the number of observations, from a mathematical standpoint there would be no need if the error covariance matrix is properly characterized; however, this is difficult to achieve in practice. Similarly, a joint or simultaneous assimilation, in which all observations are assimilated together, is mathematically more rigorous as the error covariance between the observations can be properly taken into account. I appreciate that you have dis- cussed the benefit of weighting by the type of data in the discussion, but this debate in the literature (for and against weighting, due to the abovementioned reasons) should perhaps be referred to more clearly in this study.

  Per reviewer #1 comment to shorten this paragraph and review #2 comment that the discussion lacks of the data weighting lacks precision, we cut this discussion.

- Line 129: It is true of course that to constrain changes in biomass monthly time-scale models are sufficient, but note that monthly time-scale models are not the only way to overcome computational challenges associated with inverting a complex ecosystem model. There are sophisticated yet simple algorithms that dramatically improve the sampling of parameter space in a limited number of iterations. See the work of Jasper Vrugt: https://scholar.google.co.uk/citations?user=zkNXecUAAAAJ&hl=en&oi=ao

  We cut the discussion about monthly time-step models while shortening the paragraph but will definitely look more closely into the work by Vrught. Thanks for highlighting!

Methods

- Section 2.1 It would be good if you could refer to references and/or relevant sections in the Supplement in Section2.1 to depict between standard characteristics of the 3PG model specific additions or alternative choices you made and (and to explain why you made those choices). For example:

  Added subsections to the Supplemental Material and added the references to Supplemental Material to the main text

- Line 201-202: Was this additional function based on a published study?

The function was developed as part of this study

- Line 209: Is the site-index a new addition to the model that you developed? If so, from where?

  The text now reads:

  For unfertilized plots, we used site index (SI), a measure of the height of a stand at a specified age (25 years), to estimate FR. This approach is in keeping with previous efforts (Gonzalez-Benecke et al., 2016; Subedi et al., 2015)

- Lines 218-220: Why did you remove the dependence of total root allocation on FR for the DA study?

  We removed the dependence of total root allocation on FR because we separated root allocation into the coarse and fine roots. Therefore, the previous function was not applicable. Future studies should investigate how best to build this function back in and ask whether we currently have the observational constraints to parameterize it.

- Line 229-231: A reference for or further explanation of this modification would be good here. –

  Added text

- Line 245: “implicit irrigation in very dry conditions.” Is this a realistic feature of these sites? How does this affect the results? Especially for the water availability manipulation experiments.

  We added text explain how this assumption could influence the results. “This assumption may cause the model to be less sensitive to low soil availability but the optimized parameterization may compensate. “

- Line 250: do you mean to say “mean monthly GPP”?

  GPP was a sum for each month so ‘monthly GPP’ is correct. Mean monthly GPP might imply that multiple months are averaged.

- Line 251-252: How did you select the 31 parameters to be optimized?

  In the revised manuscript, we included more parameters that were optimized (six more). The eight parameters that were not optimized did not have specific data to use as a constraint (leaf boundary layer, conductance, canopy light extinction coefficient, etc).

  The paragraph in the results section now reads as follows:

Our multi-site, multi-experiment, multi-data stream DA approach (Base assimilation) increased confidence in the model parameters (Table 5). Averaged across parameters, the posterior 99% quantile range from the Base assimilation was 60% less than the prior range. The largest reduction in parameter uncertainty was for the parameters associated with light-use efficiency (alpha) and the conversion of GPP to NPP (y), which on average had ranges that were 85% lower in the posterior than the prior. Parameters associated with allocation and allometry had a 63% reduction in the range while parameters associated with mortality processes had 70% reduction in the range. Parameters associated with environmental modifiers had the least reduction in the range with a 40% decrease. In addition to the parameters associated with the 3-PG model, the model process error parameters for each data stream were well constrained with large reductions in the range (> 99% decrease; Supplemental Material Table 2)

- Table 1: Please can you give the equation for how the sensitivity is calculated? Also, please could you explain why there is both a number and "vague" given for the uncer tainty of some parameters? If "vague", please can you detail how you defined the prior uncertainty/ranges in the text?

  We cut out the sensitivity analysis and added more parameters to the optimization.

- Finally, I appreciate you have a lot of information to con- vey and the tables are large, but it might be good to have all optimized parameters here and just indicated which ones are referred to in the discussion.

  We expanded the table to include all optimized parameters

- As a general comment, it is hard to find some of the information you refer to in the Supplement (e.g. the other optimized parameters you refer to in the caption of Table 1). Please could you split the Supplement into numbered/indexed sections and then refer specifically to the relevant section to help the reader?

  We added section divisions to the supplemental material

- Line 255-265: How did you initiate the biomass pools? Based on site-level data for the start of the simulation period? Please detail with references. If no site data were available, how sensitive were your DA experiments on the method used to initiate the biomass pools? Later note: I see you have addressed this in Section 2.4. It might be useful to refer to that section here so the reader is not questioning this in this section.

  We moved the text on the initialization described to the section on the model description

- Section 2.2 Table 2: Last column – Table 3 instead of Table 4. Also, please could uou explain, or give references, for why the SD for observations sometimes varied between 10% and 2.5% of the observation.

To reduce confusion, we used 10% for LAI observations. Future applications of the method can focus more on the influence of data uncertainty on parameter estimates.

- Section 2.3 Equation 4: Please explain why you picked a uniform distribution between 0.001 and 100?

  We added text to state that the bounds of 0.001 to 100 were designed allow the priors to be vague. The bounds include reasonable ranges of standard deviation parameters.

- Lines 348-349: Please explain why (only) 3 MCMC chains were run? Was a convergence metric such as R-hat used?

  We re-ran our optimization with the updates described at the top of the response. We ran 4 chains and used the Gelman R criteria to test for convergence. The methods section now includes the following text:

  Four chains were run and tested for convergence using the Gelman–Rubin convergence criterion, where a value for the criterion less than 1.1 indicated an acceptable level of convergence.

- Section 2.4 Lines 398-399: Although I understand the reasoning that these sites are close together and the most data rich, I don't understand why you lump the Duke CO2 enrichment site with DK3 and NC2 in the 1st stage when you stated that you wanted to test the influence of the CO2 fertilization – why not just test the Duke CO2 enrichment site by itself in the 1st stage and the remaining sites/plots in the 2nd stage to answer this question?

  Addressing this comment was the one of the primary reasons that we re-ran and simplified our analysis. Our updated analysis removed the need for a 2-step analysis. (see beginning of this response for more info).

- Further to the above point, I appreciate the extra experiments to understand the influence of the CO2 fertilization on the posterior parameters, and the further experiments to determine the influence of the water treatments and nutrient addition. But how dependent are your results on which type of observation and/or treatment is assimilated in the 1st stage vs 2nd stage? Would the results different if you reverse the stages you have in your current set-up? Again, see Wutzler and Carvalhais (2014) and/or MacBean et al. (2016) who discuss these issues (as well as the issue of the weight of different types of data, as you discuss below. A pseudo-test with synthetic observations would have been useful prior to assimilating real data to determine whether the exact set-up of a 2-stage assimilation is sensitive to the order of observation assimilation as well as to confirm if the assimilation system is able to constrain the parameters to their correct values.

  Our updated analysis removed the need for a 2-step analysis. (see beginning of this response for more info)

- Lines 430-465: While the tests and approaches put forward here are interesting, the text is dense. Any efforts the authors could make to simplify the description of the experiments and simulations performed (perhaps with the use of a table and simula- tion/experiment code names?) would likely help the reader.

  We reorganized and clarified this text in response to this comment and comments from Reviewer 1. We have a Base (all plots, three unique parameters for the Duke site), NoExp (no experimental treatments, three unique parameters for the Duke site) and NoDkPars (all plots, no unique parameters for the Duke site)

- Lines 467-475: The cross-validation exercise presented here is a useful one. Was a similar test used to assess the validity of the posterior distributions of the manipulation experiments, even though there are fewer sites?

  We added a cross validation of the experiments treatments. We now include optimized parameter set that did not include the experimentally treated plots in the assimilation. This parameter set is now used to predict the experimental treatments.

Results

- Line 480-484: Description of the sensitivity analysis and choice of parameters should be in the methods. Was this a one-at-a-time sensitivity analysis or a full global method? What is the justification for using this approach versus an existing global sensitivity analysis that accounts for correlations between parameters and explores the whole parameter space (unless I have misunderstood what was done)?

  We cut the reference to the sensitivity analysis

- Why did you fix the light extinction coefficient as opposed to the quantum yield parameter?

  We fixed the light extinction coefficient because it was more known than the canopy quantum yield.

- Supplemental Table 3 and Table 5: As mentioned above I would suggest having all the optimized parameters in one table. I would also suggest putting the prior min/max in Table 5 even though it might mean having an extra line/column per parameter and taking this information out of table 1 so it is easier to see how well the optimization has constrained the parameters.

  We moved all parameters to the table in the main text and added the range uncertainty in the priors to the same table

- Finally, I would suggest splitting up the parameter tables into the sections you refer to in the text, e.g. “temperature sensitivity of quantum yield” or “physiological parameters” etc. This will make it easier for the reader to refer to the tables when reading the text.

Done

- Which experiment do the supplemental figures correspond to? The "ALL" experiment? This should be detailed.

  The assimilation approaches have been renamed and clarified in the supplemental figures.

- Are you talking about the 1st stage experiment in the first paragraph of the results? If so, it would be good to specify this, and I would further suggest splitting the results into sections to more easily guide the reader.

  We clarified by using the names of the data assimilation approaches. Our results section is better organized in response to review 2.

- Do you discuss DK+NC2-fert in the results, or have I missed it? Perhaps more needed on the 1-stage versus 1st and 2nd stages before you discuss the experiments with and without nutrient and water addition (i.e. before line 522)?

  Our updated analysis did not require the 2-stage approach so we no longer need to report the DK+NC2-fert results. This helps simplify the description of the results.

- Figure 5 comes before Figure 4 in the text – switch around?

  Fixed in text

- Lines 507-515: I am a bit confused by the sentence "The two-stage assimilation was critical for constraining the CO2 quantum yield enhancement parameter (Calpha700)" as you then go on to say (and show, in Figure 5) that the 1 stage resulted in a narrower uncertainty interval? I guess you mean that despite the higher 95% confidence interval, the 2-stage approach results in a more realistic parameter value but I am not at all sure on that? Please could you clarify this in the text?

  Paragraph was modified in the revisions

- Line 517: I would suggest putting the names of the soil fertility parameters in brackets to aid the reader, or again put sub-headings in the parameter tables.

  Paragraph was removed during revisions

- As you did not have a strong difference in predictive capability between experiments with and without nutrient or water addition, even though you had different parameters, that presumably means you have a certain amount of model equifinality? You discuss and show the difference in model behavior as a result of the different approaches in Figures 5 – 7, but you do not discuss which one you think leads to the right behav- ior? Do you have an idea? Perhaps a synthetic experiment with pseudo-observations taken from the model simulations might help with this (a so-called "observing system simulation experiment", or OSSE)?

This was a very insightful comment. Our response reflects the updated analysis described above that has two assimilation approaches: with and without ecosystem experiments. Our new Figure 5 (the bar graph with the experimental responses from the observations and model predictions) helps support the following:

- *Including experiments in the assimiliation substantially increases the predictive capacity of the model in the* CO_2 *experiments.*
- *The predictive capacity of drought, irrigation, and nutrient fertilization experiments did not substantially change whether experiments where included or not.*

We think that an OSSE would be a great follow on study that more specifically explores of the issues that are brought up in this analysis. An OSSE could explore how locations of plots within a region and the different types of individual experiments influence the ability to retrieve known parameters. Such a study would build on the description of the cost function and general approach presented in this manuscript. Since we do not include an OSSE, we now try to avoid making general statements in the discussion that would require an OSSE to quantitatively support.

- Lines 522 onwards show very interesting results. However, I would suggest that the patterns detailed in last two paragraphs (Lines 553-572) would benefit from explanations linking back a bit more (not just referring to figures) to the different model behavior/mechanisms identified and discussed in the RW-fert and RW-water sections just above.

  In response to Reviewer 1, we cut the results of the regional simulations from the RW-fert and RW-water simulations

Discussion

- First paragraph is more of a summary than a discussion and could be cut or added to conclusions.

  We prefer to provide a summary at the beginning of discussions to remind the reviewer of key points.

- Although perhaps a little too long, this is a useful discussion that ties the results to- gether and answers some of the questions I raised in my comments on the results. Perhaps it would be useful to combine some of the summary points raised in the dis- cussion with relevant sections in the results with separate sub-headings as I mentioned above.

  We added subheadings to the discussion

- Lines 650-652: Interesting point and in addition, as I have mentioned above, I think a synthetic experiment would also be very helpful in this regard.

  We agree that a synthetic experiment would be an excellent next study. The synthetic experiment could create 'fake' region with different environmental gradients and explore the types of gradients that allow for the retrieval of parameters from the OSSE study.

Minor comments

- Line 87: Do you mean the "assimilation of manipulation experimental data", rather than the "assimilation of experiments"?

  Yes. Fixed

- Line 88: two or more

  Fixed

MacBean, N., Peylin, P., Chevallier, F., Scholze, M., and Schürmann, G.: Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, Geosci. Model Dev., 9, 3569-3588, doi:10.5194/gmd-9-3569-2016, 2016

Leveraging 35 years of *Pinus taeda* research in the southeastern U.S. to constrain forest carbon cycle predictions: regional data assimilation using ecosystem experiments

Style Definition: Normal (Web)

Deleted: forest

R. Quinn Thomas[1*], Evan Brooks[1], Annika Jersild[1], Eric Ward[2], Randolph Wynne[1], Timothy J. Albaugh[1], Heather Dinon Aldridge[3], Harold E. Burkhart[1], Jean-Christophe Domec[4,5], Thomas R. Fox[1], Carlos A. Gonzalez-Benecke[6], Timothy A. Martin[7], Asko Noormets[8], David A. Sampson[9], Robert O. Teskey[10]

[1]Department of Forest Resources and Environmental Conservation, Virginia Tech, USA,
[2]Climate Change Science Institute and Environmental Sciences Division, Oak Ridge National Laboratory, USA,
[3]State Climate Office of North Carolina, North Carolina State University, USA,
[4]Bordeaux Sciences Agro, UMR 1391 INRA-ISPA, 33175 Gradignan Cedex, France,
[5]Nicholas School of the Environment, Box 90328, Duke University, Durham, NC 27708, USA,
[6]Department of Forest Engineering, Resources and Management, Oregon State University, USA,
[7]School of Forest Resources and Conservation, University of Florida, USA,
[8]Department of Forestry and Environmental Resources, North Carolina State University, USA,
[9]Decision Center for a Desert City, Arizona State University, USA,
[10]Warnell School of Forestry and Natural Resources, University of Georgia, USA,

*Corresponding author: R. Quinn Thomas (rqthomas@vt.edu)

Formatted: Font color: Auto

Formatted: Widow/Orphan control

Abstract

Predicting how forest carbon cycling will change in response to climate change and management depends on the collective knowledge from measurements across environmental gradients, ecosystem manipulations of global change factors, and mathematical models. Formally integrating these sources of knowledge through data assimilation, or model-data fusion, allows the use of past observations to constrain model parameters and estimate prediction uncertainty. Data assimilation (DA) focused on the regional scale has the opportunity to integrate data from both environmental gradients and experimental studies to constrain model parameters. Here, we introduce a hierarchical Bayesian DA approach (Data Assimilation to Predict Productivity for Ecosystems and Regions, DAPPER) that uses observations of carbon stocks, carbon fluxes, water fluxes, and vegetation dynamics from loblolly pine plantation ecosystems across the Southeastern U.S. to constrain parameters in a modified version of the 3-PG forest growth model. The observations included major experiments that manipulated atmospheric carbon dioxide (CO_2) concentration, water, and nutrients, along with non-experimental surveys that spanned environmental gradients across an 8.6 x 10^5 km^2 region. We optimized regionally representative posterior distributions for model parameters, which dependably predicted data from plots withheld from the data assimilation. While the mean bias in predictions of N fertilization experiments, irrigation experiments, and CO_2 enrichment experiments was low, future work needs to focus modifications to model structure that decrease the bias in predictions of drought experiments. Predictions of how growth responded to elevated CO_2 strongly depended on whether ecosystem experiments were assimilated and whether the assimilated field plots in the CO_2 study were allowed to have different mortality parameters than the other field plots in the region. We present predictions of stem biomass productivity under elevated CO_2,

Formatted: Line spacing: double, Don't keep with next

Deleted:

Deleted: However,

Deleted: influence of different

Deleted: treatments on those predictions depends on the exact methods and techniques used for data assimilation.

Deleted: of Pine Plantation Ecosystem Research

Deleted: studies

Deleted: the most sensitive

Formatted: Font:Times New Roman

Deleted: The posterior distributions

Deleted: parameters associated with ecosystem responses to CO_2, precipitation, and nutrient addition, along with the corresponding regional changes in production associated with nutrient

Deleted: and drought, depended on how the experimental data were assimilated. In particular, assimilating nutrient addition

Deleted: reduced the predicted sensitivity to nutrient fertilization while assimilated water manipulation

Deleted: increased the sensitivity to drought. Further, it was necessary to assimilate data from the CO_2 experimental

Deleted: site before other studies to constrain the parameters associated with the influence

Deleted: CO_2 on canopy photosynthesis. The ambient CO_2 plots were numerous and had a large contribution to the cost function compared to the low number

Deleted: (289 ambient vs. 5

Deleted: plots).

decreased precipitation, and increased nutrient availability that include estimates of uncertainty for the Southeastern U.S. Overall, we: 1) demonstrated how three decades of research in southeastern U.S. planted pine forests can be used to develop DA techniques that use multiple locations, multiple data streams, and multiple ecosystem experiment types to optimize parameters, and 2) developed a tool for the development of future predictions of forest productivity for natural resource managers that leverage a rich dataset of integrated ecosystem observations across a region.

Deleted: data assimilation

Deleted: . This approach allows

Deleted: to be consistent with

Deleted: history

Deleted: research

Deleted:

Deleted:

Formatted: Widow/Orphan control

Formatted: Font:Not Bold

1 Introduction

Forest ecosystems absorb and store a large fraction of anthropogenic carbon dioxide (CO_2) emissions (Le Quere et al., 2015; Pan et al., 2011) and supply wood products to a growing human population (Shvidenko et al., 2005). Therefore, predicting future carbon sequestration and timber supply is critical for adapting forest management practices to future environmental conditions and for using forests to assist with the reduction of atmospheric CO_2 concentrations. The key sources of information for developing these predictions are results from global change ecosystem manipulation experiments, observations of forest dynamics across environmental gradients, and process-based ecosystem models. The challenge is integrating these three sources into a common framework for creating probabilistic predictions that provide information on both the expected future state of the forest and the probability distribution of those future states.

Data assimilation (DA), or data-model fusion, is an increasingly used framework for integrating ecosystem observations into ecosystem models (Luo et al., 2011; Niu et al., 2014; Williams et al., 2005). DA integrates observations with ecosystem models through statistical, often Bayesian, methods that can generate probability distributions for ecosystem model parameters and initial states. DA allows for the explicit accounting of observational uncertainty (Keenan et al., 2011), the incorporation of multiple types of observations with different time scales of collection (MacBean et al., 2016; Richardson et al., 2010), and the representation of prior knowledge through informed parameter prior distributions or specific relationships among parameters (Bloom and Williams, 2015).

Using DA to parameterize ecosystem models with observations from multiple locations that

Formatted: Font:Bold

Formatted: Abstract

Deleted: Forest ecosystems absorb and store a large fraction of anthropogenic carbon dioxide (CO_2) emissions (Le Quere et al., 2015; Pan et al., 2011) and supply wood products to a growing human population (Shvidenko et al., 2005). Therefore, predicting future carbon sequestration and timber supply is critical for adapting forest management practices to future environmental conditions and for using forests to assist with reduction of atmospheric CO_2 concentrations. The key sources of information for developing these predictions are results from global change ecosystem manipulation experiments, observations of forest dynamics across environmental gradients, and process-based ecosystem models. The challenge is integrating these three tools into a common framework for creating probabilistic predictions, or forecasts (based on (Luo et al., 2011a)), that provide information on both the expected future state of the forest and the probability distribution of those future s ... [1]

leverage ecosystem manipulation experiments and environmental gradients will allow for predictions to be consistent with the rich history of global change research in forest ecosystems. Ecosystem manipulation experiments provide a controlled environment in which data collected can be used to describe how forests acclimate and operate under altered environmental conditions (Medlyn et al., 2015) and can potentially allow for the optimization of model parameters associated with the altered environmental factor in the experiment. Furthermore, the assimilation of data from ecosystem manipulation experiments may increase parameter identifiability (reducing equifinality (Luo et al., 2009)), where two parameters have compensating controls on the same processes, by isolating the response to a manipulated driver. Observations that span environmental gradients include measures of forests ecosystem stocks and fluxes across a range of climatic conditions, nutrient availabilities, and soil water dynamics. These studies leverage time and space to quantify the sensitivity of forest dynamics to environmental variation. However, covariation of environmental variation can pose challenges separating the responses to individual environmental factors. Overall, assimilating observations from a region that includes environmental gradients and manipulation experiments is a useful extension of prior DA research focused on DA at a single site with multiple types of observations (Keenan et al., 2012; Richardson et al., 2010; Weng and Luo, 2011).

Southeastern U.S. planted pine forests are ideal ecosystems for exploring the application of DA to carbon cycle and forest production predictions. These ecosystems are dominated by loblolly pine (*Pinus taeda* L.), thus allowing for a single parameter set to be applicable to a large region containing many soil types and climatic gradients. Loblolly pine represents more than one half of the standing pine volume in the southern United States (11.7 million ha) and is by far the single most commercially important forest tree species for the region, with more than 1 billion seedlings planted annually (Fox et al., 2007; McKeand et al., 2003). There is also a rich history of experimental research located across the region focused on global change factors that have included nutrient addition (Albaugh et al., 2016; Carlson et al., 2014; Raymond et al., 2016), water exclusion (Bartkowiak et al., 2015; Tang et al., 2004; Ward et al., 2015; Will et al., 2015), and water addition experiments (Albaugh et al., 2004; Allen et al., 2005; Samuelson et al., 2008). The region also includes a multi-year ecosystem CO_2 enrichment study (McCarthy et al., 2010). Furthermore, many of these experiments are multi-factor with water exclusion by nutrient addition (Will et al., 2015), water addition by nutrient addition (Albaugh et al., 2004; Allen et al., 2005; Samuelson et al., 2008), and CO_2 by nutrients addition treatments (McCarthy et al., 2010; Oren et al., 2001). Beyond experimental treatments, Southeastern U.S. loblolly pine ecosystems include at least two eddy-covariance sites with high frequency measurements of C and water fluxes along with biometric observations over many years (Noormets et al., 2010; Novick et al., 2015), and sites with multi-year sap flow data (Ewers et al., 2001; Gonzalez-Benecke and Martin, 2010; Phillips and Oren, 2001). Finally, there are studies that include plots that span the regional environmental gradients and extend back to the 1980s (Burkhart et al., 1985). Overall, the multi-decadal availability of observations of C stocks (or biomass), leaf area index (LAI), C fluxes, water fluxes, and vegetation dynamics in plots with experimental manipulation and plots across environmental gradients, is well suited to potentially constrain model parameters and predictions of how carbon cycling responds to environmental change.

Using loblolly pine plantations across the southeastern U.S as a focal application, our objectives were to 1) develop and evaluate a new DA approach that integrates diverse data from multiple locations and experimental treatments with an ecosystem model to estimate the probability distribution of model parameters, 2) examine how the predictive capacity and optimized parameters differ between an assimilation approach that only uses environmental gradients and an assimilation approach that uses both environmental gradients and ecosystem manipulations, and 3) demonstrate the capacity of the DA approach to predict, with uncertainty, regional forest dynamics by simulating how forest productivity responds to drought, nutrient fertilization, and elevated atmospheric CO$_2$ across the Southeastern U.S.

2 Methods

2.1 Observations

We used thirteen different data streams from 294 plots at 187 unique locations spread across the native range of loblolly pine trees to constrain model parameters (Table 1; Figure 1). The data streams covered the period between 1981 to 2015. The Forest Modeling Research Cooperative (FMRC) Thinning Study provides the largest number of plots that span the region (Burkhart et al., 1985). In this study, we only used the control plots that were not thinned. The Forest Productivity Cooperative (FPC) Region-wide 18 (RW18) study included control and nutrient fertilization addition plots that span the region (134.4 kg ha^{-1} N + 13.44 kg ha^{-1} P biannually) (Albaugh et al., 2015). The PINEMAP study included four locations dispersed across the region that included a replicated factorial experiment with control, nutrient fertilization (224 kg ha^{-1} N + 27 kg ha^{-1} P + micronutrients once at project initiation), throughfall reduction (30% reduction), and fertilization by throughfall treatments (Will et al., 2015). The SETRES study was located at a single location and included replicated control, irrigation (~650 mm of added water per year),

Formatted: Don't keep with next

Formatted: Font:Bold

Moved (insertion) [1]

Formatted: Font:Bold

nutrient fertilization (~100 kg N ha^{-1} + 17 kg P ha^{-1} with micronutrients applied annually with absolute amount depending on foliar nutrient ratios), and fertilization by irrigation treatments (Albaugh et al., 2004). The Waycross study was a single site with a non-replicated fertilization treatment. The annual application of nutrient fertilization was focused on satisfying the nutrient demand by the trees and resulted in one most productive stands in the region (Bryars et al., 2013). These five studies included data streams of stand stem biomass (defined as the sum of stemwood, stembark and branches) and live stem density. Waycross and SETRES included LAI measurements from litterfall traps (Waycross) or estimates from LICOR LAI-2200 (SETRES). SETRES also included fine root and coarse root measurements. In the PINEMAP, SETRES, and RW18 studies we only used foliage biomass estimates from the control plots. We excluded the foliage biomass estimates from the treatment plots because they were derived from allometric models that may not have captured changes in allometry due to the experimental treatment. We did use LAI measurements from both control and treatment plots where available (SETRES).

We also included observations the Duke FACE study where the atmospheric CO_2 was increased by 200 ppm above ambient concentrations. Based on the data presented in McCarthy et al. (2010) the study included six control plots, four CO_2 fumigated rings (including the unfertilized half of the prototype), two nitrogen fertilization treatments (115 kg N ha^{-1} yr^{-1} applied annually) , and one CO_2 by nitrogen addition treatment (fertilized half of prototype). The Duke FACE study included observations of stem biomass (loblolly pine and hardwood), coarse root biomass (loblolly pine and hardwood), fine root biomass (combined loblolly pine and hardwood), stem density (loblolly pine only), leaf turnover (combined loblolly pine and hardwood), fine root production (combined loblolly pine and hardwood), and monthly LAI (loblolly pine and hardwood).

Finally, we included two Ameriflux sites with eddy-covariance towers in loblolly pine stands. The US-DK3 site was located in the same forest as the Duke FACE site described above (Novick et al., 2015). The US-NC2 site was located in coastal North Carolina (Noormets et al., 2010). We used monthly gross ecosystem production (GEP; modeled gross primary productivity from net ecosystem exchange measured at an eddy-covariance tower) and evapotranspiration (ET) estimates from the sites. The monthly GET and ET were gap-filled by the site PI. The GEP was a flux partitioned product created by the site PI. The biometric data from the US-DK3 site was assumed to be the same as the first control ring. The biometric data from the US-NC2 site included observations of stem biomass (loblolly pine and hardwood), coarse root biomass (loblolly pine and hardwood), fine root biomass (combined loblolly pine and hardwood), stem density (loblolly pine only), leaf turnover (combined loblolly pine and hardwood), and fine root production (combined loblolly pine and hardwood).

2.2 Ecosystem Model

We used a modified version of the Physiological Principles Predicting Growth (3-PG) model to simulate vegetation dynamics in loblolly pine stands (Bryars et al., 2013; Gonzalez-Benecke et al., 2016; Landsberg and Waring, 1997). 3-PG is a stand-level vegetation model that runs at the monthly time-step and includes vegetation carbon dynamics and a simple soil water bucket model (Figure 2). While a complete description of the 3-PG model and our modifications can be found in the Supplemental Material Section 1, the key concept for interpreting the results is that gross primary productivity (GPP) was simulated using a light-use efficiency approach where the

Formatted: Font:Bold

Formatted: Don't keep with next

Deleted: Model

Deleted: (Bryars et al., 2013; Gonzalez-Benecke et al., 2016; Landsberg and Waring, 1997)

Deleted: 1

absorbed photosynthetically active radiation (APAR) was converted to carbon based on a quantum yield (Supplemental Material Section 1.1). Quantum yield was simulated using a parameterized maximum quantum yield (alpha) that was modified by environmental conditions including atmospheric CO_2, available soil water (ASW) and soil fertility (Supplemental Material Section 1.2-1.3). The ASW and soil fertility modifiers were values between 0 and 1, while the atmospheric CO_2 modifier had a value of 1 at 350 ppm (thus values greater than 1 at higher CO_2 concentrations).

Elevated CO_2 modified tree physiology by increasing quantum yield, based on an increasing but saturating relationship with atmospheric CO_2 (Supplemental Material Section 1.2). Based on initial results from the data assimilation, we also added a function where the allocation to foliage relative to stem biomass decreased as atmospheric CO_2 increased (Supplemental Material Section 1.2). ASW and quantum yield were positively related through a logistic relationship between relative ASW and the quantum yield modifier, where relative ASW was the ratio of simulated ASW to a plot-level maximum ASW. Soil fertility and quantum yield were proportionally related, where quantum yield was scaled by an estimate of relative stand-level fertility (a value of 1 was the maximum fertility). The fertility modifier (FR) was constant throughout a simulation of a plot and was either based on site characteristics or directly optimized as a stand-level parameter (Supplemental Material Section 1.3). For plots with nutrient fertilization, FR was a directly optimized parameter or set to 1, depending on the level of fertilization (see below). For unfertilized plots, we used site index (SI), a measure of the height of a stand at a specified age (25 years), to estimate FR. This approach is in keeping with previous efforts (Gonzalez-Benecke et al., 2016; Subedi et al., 2015); however, SI does not solely represent nutrient availability of an

Deleted: .

Deleted: α

Deleted: air temperature,

Deleted: .

Deleted: available soil water

Deleted: and

Deleted: .

Deleted: . We

Deleted: . Available soil water

Deleted: available soil water

Deleted: available soil water

Deleted: available soil water

Deleted: available soil water.

Deleted: where

Deleted: .

Deleted: . Here we

Deleted: -

Deleted: ,

Deleted: and the 35-year mean annual temperature as site characteristics to predict FR.

ecosystem. For a given climate SI captures differences in soil fertility, where a lower SI corresponded to a site with lower fertility, but regional variation in SI also included the influence of climate on growth rates that were already accounted for in the other environmental modifiers in the 3-PG model. When a climate term is not used in the empirical FR model, FR is relative to the highest SI in the region, which does not occur in the northern extent of the region even in fertilized plots due to climatic constraints. Thus, we also included the historical (1970-2011) 35-year mean annual temperature (MAT) as an additional predictor, resulting in an empirical relationship that predicted FR as an increasing, but saturating, function of SI within areas of similar long-term temperature. For our application of the 3-PG model using DA, we removed the previously simulated dependence of total root allocation on FR (Bryars et al., 2013; Gonzalez-Benecke et al., 2016) because we separated coarse and fine roots. Other environmental conditions influenced GPP, including temperature, frost days, and vapor pressure deficit (VPD). A description of these modifiers can be found in Supplemental Material Section 1.2.

Each month, net primary production (a parameterized and constant proportion of GPP) was allocated to foliage, stem (stemwood, stembark, and branches), coarse roots, and fine roots (Supplemental Material Section 1.4). Differing from previous applications of 3-PG to loblolly pine ecosystems, we modified the model to simulate fine roots and coarse roots separately. 3-PG also simulated simple population dynamics by including stem density as a state variable. Stem density and stem biomass pools were reduced by both density-dependent mortality, based on the concept of self-thinning (Landsberg and Waring, 1997), and density-independent mortality, a new modification where a constant proportion of individuals die each month (Supplemental Material Section 1.5). Finally, we added a simple model of hardwood understory vegetation to

Deleted: , site index

Deleted: site index

Deleted: . However,

Deleted: site index

Deleted: To account for the climatic influence on site index, a long-term climate variable (35-year mean annual temperature) was included in the empirical

Deleted: site index. For plots with nutrient fertilization, FR was a directly optimized parameter.

Deleted: (Bryars et al., 2013; Gonzalez-Benecke et al., 2016). Therefore, plots with lower FR could be interpreted to have lower quantum yield.

Deleted: frosts

Deleted: with a

Deleted: the

Deleted: .

Deleted: .

Deleted: and density-independent mortality (a new modification), with the former based on the concept of self-thinning. Finally, we added a simple model of hardwood understory vegetation to enable the use of estimates of gross primary productivity and evapotranspiration from eddy-covariance tower studies with significant understories. Details of the model can be found in the Supplemental Material.

enable the assimilation GEP and ET observations from eddy-covariance tower studies with significant understories (Supplemental Material Section 1.7).

The water cycle was a simple bucket model with transpiration predicted using a Penman-Monteith approach (Bryars et al., 2013; Gonzalez-Benecke et al., 2016; Landsberg and Waring, 1997)(Supplemental Material Section 1.6). The canopy conductance used in the Penman-Monteith subroutine was modified by environmental conditions. The modifiers included the same ASW and VPD modifier as used in the GPP calculation. Maximum canopy conductance occurred when simulated LAI exceeded a parameterized value of LAI (LAIgcx). Evaporation was equal to the precipitation intercepted by the canopy. Runoff occurred when the ASW exceeded a plot-specific maximum ASW. As in prior applications of 3-PG, ASW was not allowed take a value below a minimum ASW, resulting in an implicit irrigation in very dry conditions. This assumption may cause the model to be less sensitive to low ASW but the optimized parameterization may compensate.

The 3-PG model used in this study simulated the monthly change in eleven state variables per plot: four stocks for loblolly pines, five stocks for understory hardwoods, loblolly pine stem density (stems ha^{-1}), and ASW. The key fluxes that were used for DA included monthly GEP, monthly ET, annual root turnover, and annual foliage turnover. In total, 46 parameters were required by 3-PG. The model required mean daily maximum temperature, mean daily minimum temperature, mean daily PAR, total frost days per month, total rain per month, annual atmospheric CO$_2$, and latitude. Each plot also required maximum ASW, SI, MAT, and the initial condition of the eleven state variables as model inputs (Figure 3).

Deleted: The water cycle was a simple bucket model with transpiration predicted using a Penman-Monteith approach (Bryars et al., 2013; Gonzalez-Benecke et al., 2016; Landsberg and Waring, 1997). The canopy conductance used in the Penman-Monteith subroutine was modified by environmental conditions. The modifiers include the same available soil water and vapor pressure deficit modifier as used in the GPP calculation. Maximum canopy conductance occurred when simulated LAI exceeded a parameterized value of leaf area index (LAI). Evaporation was equal to the precipitation intercepted by the canopy. Runoff occurred when the available soil water exceeded a plot-specific maximum available soil water. As in prior applications of 3-PG, available soil water was not allowed take a value below a minimum available soil water, resulting in an implicit irrigation in very dry conditions.

... [2]

Formatted: Don't keep with next

Deleted: available soil water.

Deleted: GPP

Deleted: evapotranspiration (

Deleted:),

Deleted: with 31 of the parameters optimized using DA (Table 1, Supplemental Table 1, SI Table 2).

Deleted: at the monthly time scale, monthly

Deleted: available soil water, site index, mean annual temperature

Deleted: 2

We used the first observation at the plot as the initial conditions for the loblolly pine vegetation states (foliage biomass, stem biomass, coarse root biomass, fine root biomass, and stem number). When observations of coarse biomass and fine root biomass were not available, these stocks were initialized as a mean region-wide proportion of the observed stem biomass. However, the value of initial root biomass in plots without observations was not important because root biomass did not influence any other functions in the model. The hardwood understory stocks at US-DK3 and US-NC2 were also initialized using the first set of observations. Initial fine root and coarse biomass were distributed between loblolly pine and hardwoods based on their relative contribution of total initial foliage biomass. The initialized ASW was assumed to be equal to the maximum ASW because most plots were initialized in winter months when plant demand for water was minimal. The maximum ASW in each plot was extracted from the SSURGO soils dataset (Staff, 2016). The value we used corresponded to the maximum ASW for the top 1.5 m of the soil. We assumed that the minimum ASW was zero. Because we focused on a region-wide optimization, we used region-wide 4-km estimates of observed monthly meteorology as inputs and to calculate the 35-year MAT for each plot (Abatzoglou, 2013). SI was based on height measurements at age 25 in each plot or calculated by combining observations of height at younger ages with an empirical model (Dieguez-Aranda et al., 2006).

We simulated ecosystem manipulation experiments in the 3-PG model by altering the environmental modifiers or by modifying the environmental inputs. Nutrient addition experiments were simulated by setting FR equal to 1 for the studies that applied nutrients at regular interval to remove nutrient deficiencies (RW18, SETRES, Waycross). FR was directly estimated for fertilized plots in two of the studies either because nutrients were only added once

Moved (insertion) [2]

Formatted: Don't keep with next

Moved up [1]: The data streams covered the period between 1981 to 2015.

Deleted: We used a hierarchal Bayesian framework to approximate the posterior probability distributions of model parameters in Table 1, the model process uncertainty parameters, and the latent model states and fluxes. The latent model states represented the 'true' stock or flux before measurement uncertainty was included in the observation. Our hierarchal approach was designed to partition uncertainty that is attributable to uncertainty in param ... [5]

Moved down [4]: Equation 2

Deleted: **2.2 Observations** ... [3]

Deleted: ... [6]

Formatted: Font:Bold

Deleted: All data streams were not available in all p ... [4]

Moved down [3]: **Data assimilation method**

Moved down [5]: ... [7]

Deleted: posterior distributions using the Monte-Ca ... [8]

Moved down [6]: All MCMC-MH chains were ru ... [9]

Deleted: Three chains were run and compared for ... [10]

Moved down [7]: The 3-PG model and MCMC- ... [11]

Formatted: Font:Bold

Deleted: **2.4 Model simulations** ... [12]

Moved up [2]: We used the first observation at th ... [13]

Deleted: the plots without

Deleted: observations

Deleted: contribute to the root cost function and ro ... [14]

Deleted: In the two plots with flux observations (U ... [15]

Deleted: was

Deleted: was

Deleted: available soil water

Deleted: available soil water

Deleted: is

Deleted: available soil water

Deleted: (Staff, 2016). We assumed that the minin ... [16]

Deleted: available soil water

Deleted: collect

Deleted: mean annual temperature for each plot ... [17]

Deleted: Site index

Deleted: (Dieguez-Aranda et al., 2006)

Deleted: the

Deleted: directly estimating FR, rather than calcul ... [18]

Deleted: or greater than the FR in the control plots.

at the beginning of the study (PINEMAP), thus potentially not removing nutrient limitation, or nitrogen was the only element added (Duke FACE), thus allowing the potential for nutrient limitation by other elements. For these plots, we also assumed that the FR of the fertilized plot was equal to or larger than the control plot. Throughfall exclusion experiments were simulated by decreasing the throughfall by 30% in the treatment plots. The SETRES irrigation experiments were simulated by adding 650 mm to ASW between April and October. CO$_2$ enrichment experiments were simulated by setting the atmospheric CO$_2$ input equal to the treatment mean from the elevated CO$_2$ rings (570 ppm). One plot (US-NC2) included a thinning treatment during the period of observation. We simulated the thinning by specifying a decrease in the stem count that matched the proportion removed at the site, with the biomass of each tree equivalent to the average of trees in the plot.

2.3 Data assimilation method

We used a hierarchal Bayesian framework to estimate the posterior distributions of parameters, latent states of stocks and fluxes, and process uncertainty parameters. The latent states represented a value of the stock or flux before uncertainty was added through measurement. The approach was as follows.

Consider a stock or flux (m) for a single plot (p) at time t ($q_{p,m,t}$). $q_{p,m,t}$ is influenced by the processes represented in the 3-PG model and a normally distributed model process error term,

$$q_{p,m,t} \sim \mathrm{N}(\mathrm{f}(\boldsymbol{\theta}, \mathrm{FR_p}), \sigma_m) \qquad \text{Equation 1}$$

Deleted: rain inputs by 30% in the treatment plots. This assumed that the fractional reduction in precipitation and

Deleted: were equal.

Deleted: Irrigation

Deleted: precipitation

Moved (insertion) [8]

Moved (insertion) [9]

Moved up [8]: CO$_2$ enrichment experiments were simulated by setting the atmospheric CO$_2$ input equal to the treatment mean from the elevated CO$_2$ rings (570 ppm).

Formatted: Don't keep with next

Deleted: While not an experiment, one

Moved up [9]: plot (US-NC2) included a thinning treatment during the period of observation. We simulated the thinning by specifying a decrease in the stem count that matched the proportion removed at the site, with the biomass of each tree equivalent to the average of trees in the ... [19]

Moved (insertion) [3]

where $\boldsymbol{\theta}$ is a vector of parameters that are optimized, $\mathrm{FR_p}$ is the site fertility, and σ_m is the model process error. Not shown are the vector of parameters that were not optimized (Supplemental Material Table 1), the plot ASW, an array climate inputs, and the initial conditions because these were assumed known and not estimated in the hierarchical model. The process error assumed that the error linearly scales with the magnitude of the prediction:

$$\sigma_m^2 = \gamma_m + \rho_m \mathrm{f}\left(\boldsymbol{\theta}, \mathrm{FR_p}\right) \qquad \text{Equation 2}$$

Formatted: Centered

Moved (insertion) [4]

While the structure of the Bayesian model allowed for all data streams to have process uncertainty that scales with the prediction, in this application we only allowed stem biomass, GEP, and ET process uncertainty to scale because they had large variation across space (stem biomass) and through time (i.e., there should be lower process uncertainty in the winter when GEP is lower). For the other data streams, the linear scaling term was removed by fixing ρ_m at 0.

$\mathrm{FR_p}$ did not have an explicit probability distribution. Rather the probability density evaluated to 1 if the plot was not fertilized, thus causing $\mathrm{FR_p}$ to be estimated from SI and MAT (Supplemental Material Equation 15), or if it was a fertilized plot and has an $\mathrm{FR_p}$ equal or higher than that of its non-fertilized control plot. The probability density evaluated to 0 if the estimated $\mathrm{FR_p}$ in a fertilized plot was less than the $\mathrm{FR_p}$ in the control plot or $\mathrm{FR_p}$ was not contained in the interval between 0 and 1.

$$\mathrm{FR_p} \sim \begin{cases} 1 \text{ if non-fertilized, } \mathrm{FR_p} \geq 0, \text{ and } \mathrm{FR_p} \leq 1 \\ 1 \text{ if } \mathrm{FR_p} = 1 \text{ and fertilization levels are assumed to remove nutrient deficiencies} \\ 0 \text{ if } \mathrm{FR_p} < 1 \text{ and fertilization levels are assumed to remove nutrient deficiencies} \\ 1 \text{ if fertilized but levels are not assumed to remove deficiencies and } \mathrm{FR_p} \geq \text{FR of control plot} \\ 0 \text{ if fertilized but levels are not assumed to remove deficiencies and } \mathrm{FR_p} < \text{FR of control plot} \\ 0 \text{ if } \mathrm{FR_p} < 0 \text{ or } \mathrm{FR_p} > 1 \end{cases}$$

Equation 3

Our model included the effect of observational errors for measurements of stocks and fluxes. For a single stocks or flux for a plot at time t there was an observation ($y_{p,m,t}$). The normally distributed observation error model was:

$$y_{p,m,t} \sim \mathrm{N}(q_{p,m,t}, \tau^2_{p,m,t})$$

Equation 4

where $\tau^2_{p,m,t}$ represented the measurement error of the observed state or flux. By including the observational error model, $q_{p,m,t}$ represented the latent, or unobserved, stock or flux. The variance was unique to each observation because it was represented as a proportion of the observed value. The $\tau^2_{p,m,t}$ was assumed known (Table 1) and not estimated in the hierarchical model.

The hierarchical model required prior distributions for all optimized parameters, including the parameters for the 3-PG model (θ), $\mathrm{FR_p}$, and the process error parameters. The prior distributions for $\boldsymbol{\theta}$ are specified in Table 3. Some parameters were informed by previous research in loblolly pine ecosystems while other parameters were ‘non-informative’ with flat distributions (termed ‘vague’ in Table 3). The prior distributions for the process error parameters were non-informative and had a uniform distribution with upper and lower bounds that spanned the range of reasonable error terms.

$$\gamma_m \sim U(0.001,100) \quad \text{Equation 5}$$

$$\rho_m \sim U(0,10) \quad \text{Equation 6}$$

By combining the data, process, and prior models, our joint posterior that includes all thirteen data streams, plots, months with observations, and fitted parameters was

$$p(\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{q}|\boldsymbol{y},\boldsymbol{\tau},priors) \propto$$

$$\prod_{p=1}^{P}\prod_{m=1}^{M}\prod_{t=1}^{T}\mathrm{N}\left(q_{p,m,t}|\mathrm{f}\left(\boldsymbol{\theta},\mathrm{FR_p}\right),\gamma_m+\rho_m\mathrm{f}\left(\boldsymbol{\theta},\mathrm{FR_p}\right)\right)$$

$$\prod_{p=1}^{P}\prod_{m=1}^{M}\prod_{t=1}^{T}\mathrm{N}\left(y_{p,m,t}|q_{p,m,t},\tau_{p,m,t}^{2}\right)$$

$$\prod_{p=1}^{P}p(FR_p)\prod_{f=1}^{F}p(\theta_f)\prod_{m=1}^{M}p(\gamma_m)\prod_{m=1}^{M}p(\rho_m)$$

Equation 7

where bolded components represent vectors, P is the total number of plots, M is the total number of data streams, T is the total months with observations, and F is the total number of 3-PG parameters that are optimized.

Moved (insertion) [5]

We numerically estimated the joint posterior distribution using the Monte-Carlo Markov Chain – Metropolis Hasting (MCMC-MH) algorithm (Zobitz et al., 2011). This approach has been widely used to approximate parameter distributions in ecosystem DA research (Fox et al., 2009; Trudinger et al., 2007; Williams et al., 2005; Zobitz et al., 2011). Briefly, the algorithm proposed new values for the model parameters, uncertainty parameters, latent states, and FR. The proposed values were generated using a random draw from a normal distribution with a mean equal to the previously accepted value for that parameter and standard deviation equal to the parameter-specific jumping size. The ratio of the proposed calculation of Equation 7 to the previously accepted calculation of Equation 7 was used to determine if the proposed parameter was accepted. If the ratio was greater than or equal to 1 the proposed value was always accepted. If the ratio was less than 1, a random number between 0 and 1 was drawn and the proposed value was accepted if the ratio was greater than the random number. This allowed less probable parameter sets to be accepted, thus sampling the posterior distribution. We adapted the size of the jump size for each parameter to ensure the acceptance rate of the parameter set was between 22% and 43% (Ziehn et al., 2012) by adjusting the jump size if the acceptance rate for a parameter was outside the 22 – 43% range. All MCMC-MH chains were run for 30 million iterations with the first 15 million iterations discarded as the burn-in. Four chains were run and tested for convergence using the Gelman–Rubin convergence criterion, where a value for the criterion less than 1.1 indicated an acceptable level of convergence. We sampled every 1000th parameter in the final 15 million iterations of the MCMC-MH chain and used this thinned chain in the analysis described below. The 3-PG model and MCMC-MH algorithm were programed in FORTRAN 90 and used OpenMP to parallelize the simulation of each plot within an iteration of the MCMC-MH algorithm.

2.4 Data assimilation evaluation

Using the observations, model, and hierarchical Bayesian method described above, we assimilated both the non-manipulated and manipulated plots (Base assimilation; Table 4). We

Moved (insertion) [6]

Moved (insertion) [7]

Formatted: Font:12 pt

Formatted: Normal

Deleted: <#>Model experiments and analysis ... [20]

assessed model performance first by calculating the RMSE and bias of stem biomass predictions (the most common data stream). In the evaluation, we only used the most recent observed values to increase the time length between initialization and validation. Second, we assessed the predictive capacity by comparing model predictions to data not used in the parameter optimization in a cross-validation study. In this evaluation, we repeated the Base assimilation without 160 FMRC thinning study plots (Table 2), predicted the 160 plots using the median parameter values, and calculated the RMSE and bias stem biomass of the independent set of plots. Rather than holding out all 160 plots from a single assimilation and not generating a converged chain, we divided the 160 plots into four unique sets of 40 plot and repeated the assimilation for each set. Finally, we compared the predicted responses to experimental manipulation to the observed responses. We focused the comparison on the percentage difference in stem biomass between the control and treatment plots. We used a paired t-test to test for differences between the predicted and observed responses within an experimental type (irrigated, drought, nutrient addition, and elevated CO_2). We combined the single and multi-factor treatments for analysis. For the analysis of the nutrient addition studies we only used plots where FR was assumed to be 1 so that we were able to simulate the treatments without requiring the optimization of a site-specific FR parameter.

During preliminary analysis, we found that the Base assimilation predicted lower stem biomass than observed in the elevated CO_2 plots in the Duke FACE study. Further analysis investigating the cause of the bias in the CO_2 plots showed that three parameters (wSx1000, ThinPower, and pCRS) were required to be unique to the Duke FACE study in order to reduce the bias. Therefore, the Base assimilation included unique parameters for wSx1000, ThinPower, and pCRS parameters in all plots in the Duke FACE and US-DK3 studies. To highlight the need for the site-specific parameters, we repeated the Base assimilation approach without the three additional parameters for the Duke studies (NoDkPars assimilation).

2.5 Sensitivity to inclusion of ecosystem experiments

We also evaluated how parameter distributions and the associated environmental sensitivity of model predictions depended on the inclusion of ecosystem experiments in data assimilation. First, we repeated the Base assimilation, this time excluding the plots that included the manipulated treatments (NoExp). We removed all manipulation types at once, rather than individual experimental types, because all experimental types involved multi-factor studies. The NoExp assimilation had the same number of data streams as the Base assimilation because it included the control treatments from the experimental studies. The NoExp assimilation represented the situation where only observations across environmental gradients were available. Second, we compared the parameterization of the ASW, soil fertility, and atmospheric CO_2 environmental modifiers from the Base to the NoExp assimilation. The modifiers equations are described in Supplemental Material Section 1.2 and 1.3. Third, we repeated the same independent validation exercise for the 160 FMRC plots as described above for the Base assimilation. Fourth, we predicted the treatment plots in the irrigated, drought, nutrient addition (only plots where FR was assumed to be 1), and elevated CO_2 plots. As for the Base assimilation, we used a t-test to compare the experimental response between the NoExp assimilation and observed and between the NoExp and Base assimilations. Since the experimental treatments were not used in the optimization, this was an independent evaluation of predictive capacity.

2.6 Regional predictions with uncertainty

To demonstrate the capacity of the data assimilation system to create regional predictions with uncertainty, we simulated the regional response to a decrease in precipitation, an increase in nutrient availability, and an increase in atmospheric CO_2 concentration, each as a single factor change from a 1985-2011 baseline. Each prediction included uncertainty by integrating across the parameter posterior distributions using a Monte-Carlo sample of the parameter chains. Our region corresponded to the native range of loblolly pine and used the HUC12 (USGS 12-digit Hydrological Unit Code) watershed as the scale of simulation. For each HUC12 in the region we used the mean SI, 30-year mean annual temperature, ASW aggregated to the HUC12 level, and monthly meteorology from Abatzoglou (2013) as inputs (Figure 3). The SI of each HUC12 was estimated from biophysical variables in the HUC12 using the method described in Sabatia and Burkhart (2014). This SI corresponded to an estimated SI for stands without intensive silvicultural treatments or advanced genetics of planted stock.

To sample parameter uncertainty, we randomly drew 500 samples from the Base assimilation MCMC chain and simulated forest development from a 1985 planting to age 25 in 2011 in each HUC. We chose age 25 as the final age because it is a typical age of harvest in the region. For each sample, we repeated the regional simulation with 1) a 30% reduction in precipitation, 2) FR set to 1, and 3) atmospheric CO_2 increased by 200 ppm. Within a parameter sample, we calculated the percent change in stem biomass at age 25 between control simulation and the three simulations with the environmental changes. We focused our regional analysis on the distribution of the percent change in stem biomass.

Formatted: Don't keep with next

Moved down [10]: Table 1.

Deleted: These sites were grouped together because they were the most data rich, had the high frequency data streams (monthly GEP, ET, and LAI), and were relatively close in geography. FR was directly estimated for all plots in the first stage, with the FR of a fertilized plot required to be equal to or higher than its control plot. The FR of the CO_2 experiment was equal to the corresponding control plot estimated FR. The FR of the control plot was required to be greater than 0 and, if associated with a nutrient fertilization plot, less than the FR of the fertilized plot.
To examine how the exclusion of the nutrient addition experiment influenced parameter inference and prediction, we focused on the difference in maximum quantum yield parameter (α) and the relationship between site index and soil fertility modifier (FR) with and without assimilating the nutrient experiments. Additionally, we simulated how stem biomass at age 25 ($STEM_{25}$) responded to a complete removal of nutrient limitation (FR = 1) for the focal site in Georgia. As in the precipitation sensitivity described above, we represented the percentage change in $STEM_{25}$ between simulations with estimated FR and FR =1 as a distribution by integrating across parameter uncertainty. We predicted the regional response to nutrient fertilization by setting the FR at all HUC12 units (see previous paragraph) equal to 1 using the median posterior parameter values from data assimilation where nutrient addition experiments were either included or not. We focused on the regional pattern in the percentage change in stem biomass with the predicted FR (current level fertility) and FR = 1 (nutrient limitation removed). ... [21]

Deleted: manipulation experiments in the region.
To examine how the exclusion of the nutrient addition experiment influenced parameter inference and prediction, we focused on the difference in maximum quantum yield parameter (α) and the relationship between site inde ... [22]

Deleted: and a 30% decrease in annual precipitation with and without assimilating the water experiments. A 30% percent decrease in precipitation mirrors the magnitude of reduction in the experimental throughfall reduction ... [23]

Deleted: uncertainty by repeating simulations using 1000 random draws from the converged chain of the posteriors. Finally, we predicted the regional response to a reduction in precipitation from historical using the median

Deleted: values from the data assimilation with and without the water experiments included.

Deleted: regional

Deleted: site index

Deleted: available soil water

Deleted: as inputs (Figure 2). We simulated forest development from 1989 to 2014 using actual precipitation and again with a 30% reduction in precipitation. We focused our analysis on the percent change in $STEM_{25}$ betwe ... [24]

Deleted: To examine how the exclusion of the nutrient addition experiment influenced parameter inference and prediction, we focused on the difference in maximum quantum yield parameter (α) and the relationship be ... [25]

3 Results

3.1 Data assimilation evaluation

Our multi-site, multi-experiment, multi-data stream DA approach (Base assimilation) increased confidence in the model parameters (Table 5). Averaged across parameters, the posterior 99% quantile range from the Base assimilation was 60% less than the prior range. The largest reduction in parameter uncertainty was for the parameters associated with light-use efficiency (alpha) and the conversion of GPP to NPP (y), which on average had ranges that were 85% lower in the posterior than the prior. Parameters associated with allocation and allometry had a 63% reduction in the range while parameters associated with mortality processes had 70% reduction in the range. Parameters associated with environmental modifiers had the least reduction in the range with a 40% decrease. In addition to the parameters associated with the 3-PG model, the model process error parameters for each data stream were well constrained with large reductions in the range (> 99% decrease; Supplemental Material Table 2)

The Base assimilation reliably predicted data from the regionally distributed non-manipulated plots that were not used in the optimization. The mean bias in stem biomass of the cross-validation was -3.7 % and the RMSE was 21.8 Mg ha^{-1} (Figure 4a). Furthermore, the response of stem biomass to irrigation (df = 7, p = 0.18), nutrient addition (df = 26, p = 0.29), and elevated CO$_2$ (df = 4, p = 0.43) was not significantly different between the observed and the Base assimilation (Figure 5). The Base assimilation was significantly more sensitive to drought than observed (n = 31, p < 0.001; Figure 5).

Formatted: Don't keep with next

Formatted: Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 2 + Alignment: Left + Aligned at: 0" + Indent at: 0.25", Don't keep with next

Deleted: Our multi-site, multi-experiment, multi-data stream DA approach was able to constrain most parameters in the 3-PG model (31 of 46 parameters were optimized; Table 6; Supplemental Table 3; Supplemental Figure 1-3). The 31 optimized parameters were the most sensitive parameters in the 3-PG model, defined by the change in total biomass at age 25 for the focal site in Georgia to a 10% change in the parameter (Table 1; Supplemental Table 1). One exception was the light extinction coefficient (k), which showed high sensitivity but was assumed to be fixed because it strongly co-varied with the quantum yield parameter (α). Parameters associated with biomass allocation had priors with large variance but DA was able to provide posteriors with relatively low variance (pFS2, pFS20, pR, and pCRS; Supplemental Figure 1; Supplemental Table 3). The DA process also produced posterior distributions that had less variability than the prior distribution for the important parameters associated with light-use efficiency (α, y, FR1, and FR2; Table 5). DA did not change the parameter distributions, i.e., the posterior and prior distributions were similar, for the parameters that governed the temperature sensitivity of quantum yield, the VPD sensitivity of quantum yield, and the maximum canopy conductance (Supplemental Figure 1-2; Supplemental Table 3). These parameters had strong priors supported by previous research on loblolly pine physiology. Finally, the DA approach was able to estimate the distributions of the process uncertainty parameters (Supplemental Figure 3; Supplement Table 4). ... [26]

The plots at the Duke Forest study had a higher carrying capacity of stem biomass before self-thinning (WSx1000), smaller self-thinning parameter (ThinPower), and lower allocation to coarse root (pCRS) than values optimized from the other plots across the region (Table 6). The DA approach without these three study specific parameters (NoDkPars) predicted significantly lower accumulation of stem biomass in response to elevated CO_2 than observed (df = 4, p = 0.002; Figure 5). The NoDKPars assimilation optimized the CO_2 fertilization parameter (fCalpha700) to a value that predicted 45% less light-use efficiency at 700 ppm (1.13 in NoDKPar vs. 1.33 in Base; Table 6) than the Base assimilation.

3.2 **Sensitivity to inclusion of ecosystem experiments**

Excluding the experimental treatments from the data assimilation did not strongly influence the predictive capacity of the model. The RMSE validation plots in NoExp assimilation decreased slightly compared to Base assimilation (21.8 to 18.0 Mg ha^{-1}) while the bias slightly increased (-3.7 to -4.1%)(Figure 4b). Excluding the experimental treatments resulted in a significantly lower response of stem biomass to elevated CO_2 than observed (df = 4, p < 0.001; Figure 5). Furthermore, there was a slight negative response of stem biomass to CO_2 in the NoExp assimilation because the parameter governing the change in foliage allocation at elevated CO_2 (fCpFS700) was unconstrained by observations (Table 6). This led to convergence on the lower bound of the prior distribution (0.5) where foliage allocation decreased with increased atmospheric CO_2. The predictions of irrigation, drought, and nutrient addition experiments were not significantly different between the Base and NoExp assimilations (Figure 5).

The parameters and associated response functions in the 3-PG for nutrients, ASW, and atmospheric CO_2 differed between the Base and NoExp assimilations (Figure 6). First, the parameterization of the soil fertility rating (FR) showed a stronger dependence on SI in the NoExp assimilation than in the Base assimilation (Figure 6a). For a given SI there was a lower FR, thus stronger nutrient limitation, when experimental treatments were excluded from assimilation. Second, the parameterization of the function relating photosynthesis and canopy conductance to ASW resulted in lower photosynthesis and maximum conductance when soil available water was less than 50% in the NoExp than Base assimilations (Figure 6b). Finally, the response of photosynthesis to atmospheric CO_2 was functionally zero in the NoExp assimilation, thus highlighting the importance of the elevated CO_2 treatments in the Duke FACE study for constraining the parameterization of the CO_2 response function (Figure 6c).

3.3 Regional predictions with uncertainty

Regionally (i.e., the native range of loblolly pines), stem biomass at age 25 ranged from 52 Mg ha^{-1} to 292 Mg ha^{-1} with the most productive areas located in the coastal plains and the interior of Mississippi and Alabama (Figure 7a). The least productive locations were the western and northern extents of the native range. The width of the 95% quantile interval for each HUC12 unit ranged from 6.2 to 29.8 Mg ha^{-1} with largest uncertainty located in most the productive HUC12 units and in the far western extent of the region (Figure 7b).

Deleted: using the two-stage approach (RW),

Deleted: were

Deleted: 8). These patterns were largely driven by patterns in the soil fertility factor (FR; Figure 9), reflecting the sensitivity of the 3-PG model to the FR parameters (Table 1). The area weighted mean $STEM_{25}$ response to fertilization (represented by setting FR = 1) across the region was 28% with the highest response occurring in the far west of the region, the Piedmont of Georgia, the interior of the gulf coast, and the northern reach of the region (Figure 10a). These were all areas with the lowest soil fertility parameter. The least responsive region to nutrient addition was in Florida (Figure 10a). Excluding the nutrient addition experiments from the DA increased the sensitivity to nutrient addition

The predicted change in stem biomass at age 25 from an additional 200 ppm of atmospheric CO_2 (over the 1985-2011 concentrations) was similar to the change associated with a removal of nutrient limitation (by setting FR = 1) (Figure 8a,c). The median change associated with elevated CO_2 for a given HUC12 unit ranged from 19.2 to 55.7% with a regional median of 21.7% (Figure 8a). The change associated the removal of nutrient limitation ranged from 6.9 to 303.7% for a given HUC12 unit, with regional median of 24.1% (Figure 8b). The response to elevated CO_2 was more consistent across space than the response to nutrient addition. The largest potential gains in productivity from nutrient addition were predicted in central Georgia, the northern extent of the region, and the western extents, areas with the lowest SI (Figure 3).

Stem biomass was considerably less responsive to a 30% decrease in precipitation, than to nutrient addition and an increase in atmospheric CO_2. The median change in stem biomass when precipitation was reduced from the 1985-2011 levels ranged from -11.6 to – 0.1% for a given HUC12 unit with a regional median of -5.1% (Figure 8c). Central Georgia was the most responsive to precipitation reduction reflecting the relatively low annual precipitation and warm temperatures (Figure 3).

For a given location, the predicted response to elevated CO_2 had larger uncertainty than the predicted response to precipitation reduction and nutrient limitation removal (Figure 8c,d,f). The uncertainty, defined as the width of the 95% quantile interval, was consistent across the region for the response to elevated CO_2 (Figure 8b). The uncertainty in the response to precipitation reduction and nutrient limitation removal was largest in the regions with the largest predicted change (Figure 8df).

4 Discussion

Using DA to parameterize models for predicting ecosystem change requires disentangling the

Formatted: Normal

Formatted: Font:12 pt, Kern at 14 pt

Deleted: 10b), as shown for the focal Georgia site (Figure 6b), but did not change the spatial patterns of the response.

Formatted: Font:12 pt, Kern at 14 pt

Deleted: ... [27]

Formatted: Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, ... + Start at: 3 + Alignment: Left + Aligned at: 0" + Indent at: 0.25", Don't keep with next

Deleted:

Formatted: Don't keep with next

Deleted: applied to forecasting

Deleted: detangling

vegetation responses to temperature, precipitation, nutrients, and elevated CO_2. To address this challenge, we introduced a regional-scale hierarchical Bayesian approach (DAPPER) that assimilated data across environmental gradients and ecosystem manipulation experiments into a modified version of the 3-PG model. Furthermore, we synthesized observations of carbon stocks, carbon fluxes, water fluxes, vegetation structure, and vegetation dynamics that spanned 35 years of forest research in a region (Table 1, Figure 1) with large and dynamic carbon fluxes (Lu et al., 2015). By combining the DAPPER system with the regional set of observations, we were able to estimate parameters in a model with high predictive capacity (Figure 4) and with quantified uncertainty on parameters (Table 5) and regional simulations (Figures 7 and 8).

Our hierarchical approach (Equation 7) was designed to partition uncertainty among parameters, model process, and measurements (Hobbs and Hooten, 2015). Separating the parameter and process uncertainty is required to estimate prediction intervals, as prediction intervals only include parameter and process errors (Dietze et al., 2013; Hobbs and Hooten, 2015). Previous forest ecosystem DA efforts have either focused on parameter uncertainty, by using measurement uncertainty as the variance term in a Gaussian cost function (Bloom and Williams, 2015; Keenan et al., 2012; Richardson et al., 2010) or on total uncertainty by directly estimating the Gaussian variance term (Ricciuto et al., 2008). Our approach allowed the estimation of the probability distribution of forest biomass before uncertainty is added through measurement. Considering that the method of DA can potentially have a large influence on posterior parameter distributions (Trudinger et al., 2007), future research should focus on comparing the hierarchical approach presented here to other approaches by using the same data constraints with alternative cost functions.

Deleted: to estimate parameters and generate uncertainty estimates on predictions of carbon and water cycling across the whole native range of loblolly pine. Furthermore, we organized

Deleted: (Figure 3; Table 3) in a region with large and dynamic carbon fluxes (Lu et al., 2015).

Deleted:). We also found that the predictions of forest productivity response to rising CO_2, altered precipitation, and altered nutrient availability were highly sensitive to the types of experiments used in DA as well as the methodological approach applied.

Deleted: We found that including nutrient and water manipulation experiments aided in distinguishing the mechanisms driving patterns in biomass across the region. Including these experiments in the data-assimilation did not improve the predictive capacity of the independent validation set of non-manipulation plots. However, including nutrient and water manipulation did change the underlying mechanisms explaining the patterns in stem biomass. Without the nutrient and water manipulation experiments, the same biomass predictions were attributable to a higher level of nutrient limitation and a lower level of water limitation. This resulted in differing sensitivities to changes in nutrient or water availability. Overall, this finding highlights a key challenge when parameterizing ecosystem models that will be used for global change predictions, that different combinations of environmental drivers can produce similar predictions of current observations. Ecosystem manipulation experiments are an important tool for addressing this challenge.

... [28]

4.1 Sensitivity to inclusion of ecosystem experiments

The most important experimental manipulation for constraining model parameters was the Duke FACE CO_2 fertilization study because the CO_2 fertilization parameters (fCalpha700 and fCpFS700) converged on the lower bounds of their prior distributions when the experiments were excluded from the assimilation. In contrast, excluding the nutrient fertilization, drought, and irrigation studies did not substantially alter the predictive capacity of the model. This finding suggests that data assimilation using plots across environmental gradients alone can constrain parameters associated with water and nutrient sensitivity. However, regardless of whether the experiments were included in the assimilation, the optimized model predicted higher sensitivity to drought than observed, highlighting that future studies should focus on improving the sensitivity to drought.

The 3-PG model included a highly-simplified representation of interactions between the water and carbon cycles that resulted in parameterizations that may contain assumptions that require additional investigation. First, transpiration was modeled as a function of a potential canopy transpiration that occurred if leaf area was not limiting transpiration. The LAI at which leaf area was no longer limiting was a parameter that was optimized (LAIgcx in Table 5), resulting in a value of 2.2. Interestingly, this optimized value is consistent with the scant literature on this topic. In their analysis of multi-year measurements of transpiration in loblolly pine, Phillips and Oren (2001) observed that transpiration per unit leaf area was relatively insensitive to increases in leaf area above LAI of approximately 2.5. Iritz and Lindroth (1996) reviewed transpiration data from a range of crop species and found only small increases in transpiration above LAI of 3-

Deleted:

Formatted: Don't keep with next

Deleted: , while consistent with observations,

Deleted: For example

Deleted: is

Deleted: SI

Deleted: 3

Moved (insertion) [11]

Deleted: 3.

Moved up [11]: Interestingly, this optimized value is consistent with the scant literature on this topic.

Deleted: (2001)

Moved (insertion) [12]

4. These authors suggest that the threshold-type responses observed were related to the range of LAI at which self-shading increases most rapidly, therefore limiting increases in transpiration. The resulting model behavior of "flat" transpiration above 2.2 LAI, with gradually decreasing photosynthesis above that value, results in increasing water use efficiency at higher LAI values. Second, the relationship between relative ASW and the modifier of photosynthesis and transpiration predicted a modifier value greater than zero when the relative ASW was zero. This resulted in positive values from photosynthesis and transpiration when the average ASW during the month was zero. In practice, the monthly ASW was rarely zero during simulations, which presents a challenge constraining the shape of the ASW modifier. The priors for the two ASW modifiers (SWconst and SWpower) had ranges that permitted the modifier to be zero. Therefore, additional data are likely needed during very dry conditions to develop a more physically based parameterization. Alternatively, the parameterization of a non-zero soil moisture modifier at zero ASW may be due to trees having access to water at soil depths deeper than the top 1.5 m of soil represented by the bucket in 3-PG. Overall, it is important to view the parameterization presented here as a phenomenological relationship that is consistent with observations from drought and irrigation experiments as well as observations across regional gradients in precipitation.

Constraining the sensitivity to atmospheric CO_2 differs from constraining the sensitivity to ASW because, unlike the multiple constraints on water sensitivity (drought, irrigation, and gradient studies), environmental conditions created by the few elevated CO_2 plots provided unique constraint on parameters. Our finding demonstrated that DA efforts should test for bias in unique ecosystem experiments before finalizing a set of model parameters used in optimization. In particular, we found that the parameter governing the photosynthetic response to elevated CO_2

Moved up [12]: reviewed transpiration data from a range of crop species and found only small increases in transpiration above LAI of 3-4. These authors suggest that the threshold-type responses observed were related to the range of LAI at which self-shading increases most rapidly, therefore limiting increases in transpiration.

Deleted: (1996)

Deleted: 3

Deleted: The parameterization of the relationships between transpiration and photosynthesis in 3-PG would likely benefit fromadditional data beyond the two eddy-covariance studies with ET observations used here. For example, canopy conductance estimates, and their associated uncertainty have been derived from assimilating observations from sap-flow measurements into a model that scales from the sensor measurements to canopy transpiration using LAI observations (Bell et al., 2015). This sap-flow to canopy conductance scaling approach (the State Space Canopy Conductance (StaCC) model (Bell et al., 2015)) produces a probability distribution of monthly canopy conductance that could be integrated into the DAPPER system by treating the posterior estimates of StaCC as the distribution of the data in equation 2. Second, the optimized parameters that described the relationship between relative available soil water

Deleted: available soil water

Deleted: available soil water

Deleted: available soil water

Deleted: available soil water

Deleted: available soil moisture

Deleted: SW1

Deleted: SW2

Deleted: is

Deleted: available soil water

Deleted: throughfall exclusion

Deleted: Beyond the specifics of the 3-PG modeling efforts, the DA of regional observations into a monthly, computationally tractable ecosystem model can potentially inform Earth system modeling efforts. While the details of physiology differ between 3-PG and global land-surface models, the concepts governing NPP allocation are similar. Therefore, DA using the 3-PG model can be used to parameterize the allocation patterns of similar plant types in a global model. One land-surface model, the Community Land Model (CLM), includes parameters that govern the ratio of stem to leaf allocation, ratio of coarse root to stem allocation, and the ratio of leaf to fine root allocation, parameters that are also optimized in DAPPER. As an example, the ratio of fine root to leaf allocation in CLM 4.0 and 4.5 for temperate pine plant function type is set to 1, resulting in equal annual allocation of carbon to foliage and fine roots (Oleson et al., 2013). In contrast, we found that the median ratio of fine root to foliage allocation was substantially lower at 0.13 (Table 6). Therefore, simulations in the CLM with the lower value of root allocation would have higher allocation to aboveground tissues if the loblolly pine parameters from our analysis were used. This w ... [29]

(fCalpha700) was substantially lower when all parameters were assumed to be shared across all plots than when the CO_2 fertilization experiment was allowed to have unique parameters. The need for the three unique parameters at the Duke FACE study parameters can be explained by the constraint provided by multiple data streams and multiple plots. An assumption of the model was that an increase in stem biomass caused a decrease stem density through self-thinning, unless the average tree stem biomass was below a parameterized threshold (WSx1000). Therefore, an increase in photosynthesis and stem biomass through CO_2 fertilization could cause a decrease in stem density. For a single study, it is straightforward to simultaneously fit the CO_2 fertilization and self-thinning parameters to fit stem biomass and stem density observations for the site. However, regional DA presents a challenge because the self-thinning parameters are well constrained by the stem biomass and stem density observations across the region but the CO_2 fertilization parameters are not. As a result of the regional DA, the self-thinning parameters caused a stronger decrease in stem density than observed in the Duke FACE study. Therefore, the optimization favored a solution where there was a lower response to CO_2, thus a smaller decrease in stem density. Allowing the Duke FACE study to have unique self-thinning parameters that resulted in lower rates of self-thinning and allowed for simulated stem biomass to respond to CO_2 in a way that matched the observations without penalizing the optimization by degrading the fit to the stem density.

Our finding that the Duke FACE study required unique self-thinning parameters to reduce bias in the simulated stem biomass suggests that when using DA to optimize parameters that are shared across plots, careful examination of prediction bias in key sites that provide unique constraint on certain parameters (like the Duke FACE) is critical. Based on this example, we suggest that DA efforts using multiple studies and multiple experiment types identify whether particular experiments at limited number of sites have the potential to uniquely constrain specific parameters. In this case, additional weight or site-specific parameters may be needed to avoid having the signal of the unique experiment overwhelmed by the large amount of data from the other sites and experiments. Additionally, the finding suggests that multi-site DA should consider using hierarchical approaches to predicting mortality, particularly because mortality is often not simulated as mechanistically as growth. A hierarchical approach, where each plot has a set of mortality parameters that are drawn from a regional distribution, could avoid having unexplained variation in mortality rates lead to bias in the parameterization of growth related processes (i.e., growth responses to CO_2, drought, nutrient fertilization, etc.). The hierarchical approach to mortality could also highlight patterns in mortality rates across a region and allow for additional investigations in the mechanisms driving the patterns.

4.2 Regional predictions with uncertainty

Our predictions of how stem biomass responses to elevated CO_2, nutrient addition, and drought were designed to illustrate the capacity of the DAPPER approach to simulate the uncertainty in future predictions. By using DA, our regional predictions and the uncertainty are consistent with observations but are associated with key caveats. First, only parameter uncertainty was presented in the regional simulations. There is additional uncertainty associated with model process error. We showed the parameter uncertainty because it isolated the capacity to parameterize the individual environmental response functions in the model. Second, the response to drought may be too strong because of the bias in the model predictions of the drought studies. However, there is potential that the drought studies underestimated the sensitivity to ASW since they are relatively short term (< 5 years) and manipulate local ASW without manipulating large scale ASW (i.e., regional water tables). Third, the large responses to N fertilization at the western and northern extents of the study region may be too high. The large responses are attributed to the low SI and the low predicted site fertility index (FR_p). The low SI may be attributable to water limitation and temperature limitation that is not fully accounted for in the parameterization. Additional nutrient addition experiments in the northern and western extent along with further development of the representation of nutrient availability in the 3-PG model may allow for a more robust representation of soil fertility. Finally, the baseline fertility used in our regional analysis was derived from an empirical model of SI that was developed using field plots with minimal management (Sabatia and Burkhart, 2014). Subsequently our estimate of baseline fertility is likely on the low end of forest stands currently in production and the response to nutrient addition may be higher than a typical stand under active management.

5 Conclusions

DA is increasingly used for developing predictions from ecosystem models that include uncertainty estimation, due to its ability represent prior knowledge, integrate observations into the parameterization, and estimate multiple components of uncertainty, including observation, parameter, and process representation uncertainty (Dietze et al., 2013; Luo et al., 2011; Niu et al., 2014). Our application of DA to loblolly pine plantations of the southeastern U.S demonstrated that these ecosystems are well suited as a test-bed for the development of DA techniques, particularly techniques for assimilating ecosystem experiments. We found that assimilating observations across environmental gradients can provide substantial constraint on many model parameters but that ecosystem manipulative experiments, particularly elevated CO_2

Formatted: Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 3 + Alignment: Left + Aligned at: 0" + Indent at: 0.25", Don't keep with next

Deleted: DA is increasingly used for ecological forecasting due to its ability represent prior knowledge, integrate observations into the parameterization, and estimate multiple components of uncertainty, including observation, parameter, and process representation uncertainty (Dietze et al., 2013; Luo et al., 2011b; Niu et al., 2014). Our application of DA to loblolly pine plantations of the southeastern U.S demonstrated that these ecosystems are well suited as a test-bed for the development of DA techniques, particularly techniques for assimilating ecosystem experiments. Further, we found that assimilating ecosystem manipulative experiments into a simple ecosystem model changed predictions quantifying how forest productivity responds to environmental change, highlighting the importance of networks of ecosystem manipulation experiments for helping to parameterize and evaluate ecosystems models (Medlyn et al., 2015). ... [30]

studies, were critical for constraining parameters associated forest productivity in a more CO_2 enriched atmosphere. This highlights the importance of whole-ecosystem manipulation CO_2 experiments for helping to parameterize and evaluate ecosystem models. Finally, we present an approach for the development of future predictions of forest productivity for natural resource managers that leverage a rich dataset of integrated ecosystem observations across a region.

6 Data availability

Observations used in the DA can be found in the following: Duke FACE study can be found in McCarthy et al. (McCarthy et al., 2010), the PINEMAP studies are available through the TerraC database (http://terrac.ifas.ufl.edu), the US-DK3 eddy-flux tower data are available through the Ameriflux database (http://ameriflux-data.lbl.gov), the Waycross data can be found in Bryars et al. (2013), the US-NC2 data are available upon request from Asko Noormets, the FMRC and FPC are available through membership with the cooperatives. The parameter chains and 3-PG model code are available upon request from R. Quinn Thomas.

Deleted:) ,

Deleted: 2003

Deleted: with

Acknowledgments

Funding support came from USDA-NIFA Project 2015-67003-23485 and the Pine Integrated Network: Education, Mitigation, and Adaptation project (PINEMAP), a Coordinated Agricultural Project funded by the USDA National Institute of Food and Agriculture, Award #2011-68002-30185. Additional funding support came from USDA-NIFA McIntire-Stennis Program. The Virginia Space Grant Consortium Graduate STEM Research Fellowship Program provided partial support for A. Jersild. Computational support was provided by Virginia Tech Advanced Research Computing. This research was also supported by grants from the French

Formatted: Widow/Orphan control

Research Agency (MACACC ANR-13-AGRO-0005 and MARIS ANR-14-CE03-0007). We thank Luke Smallman and Mat Williams for helpful discussions about data assimilation. We thank the corporate and government agency members of the FPC and FMRC research cooperatives for supporting the extensive long-term experimental and observational plots in those datasets. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract number DE-AC05-00OR22725.

References

Abatzoglou, J. T.: Development of gridded surface meteorological data for ecological applications and modelling, International Journal of Climatology, 33(1), 121–131, doi:10.1002/joc.3413, 2013.

Albaugh, T. J., Albaugh, J. M., Fox, T. R., Allen, H. L., Rubilar, R. A., Trichet, P., Loustau, D. and Linder, S.: Tamm Review: Light use efficiency and carbon storage in nutrient and water experiments on major forest plantation species, Forest Ecol Manag, 376, 333–342, doi:10.1016/j.foreco.2016.05.031, 2016.

Albaugh, T. J., Allen, H. L. and Kress, L. W.: Root and stem partitioning of *Pinus taeda*, Trees, 20(2), 176–185, doi:10.1007/s00468-005-0024-4, 2005.

Albaugh, T. J., Lee Allen, H., Dougherty, P. M. and Johnsen, K. H.: Long term growth responses of loblolly pine to optimal nutrient and water resource availability, Forest Ecol Manag, 192(1), 3–19, doi:10.1016/j.foreco.2004.01.002, 2004.

Albaugh, T., Fox, T., Allen, H. and Rubilar, R.: Juvenile southern pine response to fertilization is influenced by soil drainage and texture, Forests, 6(8), 2799–2819, doi:10.3390/f6082799, 2015.

Allen, C. B., Will, R. E. and Jacobson, M. A.: Production efficiency and radiation use efficiency of four tree species receiving irrigation and fertilization, Forest Science, 51(6), 556–569, 2005.

Bartkowiak, S. M., Samuelson, L. J., McGuire, M. A. and Teskey, R. O.: Fertilization increases sensitivity of canopy stomatal conductance and transpiration to throughfall reduction in an 8-year-old loblolly pine plantation, Forest Ecol Manag, 354, 87–96, doi:10.1016/j.foreco.2015.06.033, 2015.

Bloom, A. A. and Williams, M.: Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological “common sense” in a model–data fusion framework,

Formatted: Don't keep with next

Deleted: Abatzoglou, J. T. and Brown, T. J.: A comparison of statistical downscaling methods suited for wildfire applications, International Journal of Climatology, 32(5), 772–780, doi:10.1002/joc.2312, 2012.

Deleted: Bell, D. M., Ward, E. J., Oishi, A. C., Oren, R., Flikkema, P. G. and Clark, J. S.: A state-space modeling approach to estimating canopy conductance and associated uncertainties from sap flux density data, edited by D. Whitehead,, 35(7), 792–802, doi:10.1093/treephys/tpv041, 2015.

Biogeosciences, 12(5), 1299–1315, doi:10.5194/bg-12-1299-2015, 2015.

Bryars, C., Maier, C., Zhao, D., Kane, M., Borders, B., Will, R. and Teskey, R.: Fixed physiological parameters in the 3-PG model produced accurate estimates of loblolly pine growth on sites in different geographic regions, Forest Ecol Manag, 289, 501–514, doi:10.1016/j.foreco.2012.09.031, 2013.

Burkhart, H. E., Cloeren, D. C. and Amateis, R. L.: Yield relationships in unthinned loblolly pine plantations on cutover, site-prepared lands, Southern Journal of Applied Forestry, 9(2), 84–91, 1985.

Carlson, C. A., Fox, T. R., Allen, H. L., Albaugh, T. J., Rubilar, R. A. and Stape, J. L.: Growth responses of loblolly pine in the Southeast United States to midrotation applications of nitrogen, phosphorus, potassium, and micronutrients, Forest Science, 60(1), 157–169, doi:10.5849/forsci.12-158, 2014.

DeLucia, E. H., Drake, J. E., Thomas, R. B. and Gonzalez-Meler, M.: Forest carbon use efficiency: is respiration a constant fraction of gross primary production? Global Change Biology, 13(6), 1157–1167, doi:10.1111/j.1365-2486.2007.01365.x, 2007.

Dieguez-Aranda, U., Burkhart, H. E. and Amateis, R. L.: Dynamic site model for loblolly pine (Pinus taeda L.) plantations in the United States, Forest Science, 52(3), 262–272, 2006.

Dietze, M. C., LeBauer, D. S. and Kooper, R.: On improving the communication between models and data, Plant Cell Environ, 36(9), 1575–1585, doi:10.1111/pce.12043, 2013.

Ewers, B. E., Oren, R., Phillips, N., Stromgren, M. and Linder, S.: Mean canopy stomatal conductance responses to water and nutrient availabilities in Picea abies and Pinus taeda, Tree Physiology, 21(12-13), 841–850, 2001.

Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., Ricciuto, D., Reichstein, M., Tomelleri, E., Trudinger, C. M. and Van Wijk, M. T.: The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data, Agr Forest Meteorol, 149(10), 1597–1615, doi:10.1016/j.agrformet.2009.05.002, 2009.

Fox, T. R., Jokela, E. J. and Allen, H. L.: The Development of Pine Plantation Silviculture in the Southern United States, Journal of Forestry, 105(7), 337–347, 2007.

Gonzalez-Benecke, C. A. and Martin, T. A.: Water availability and genetic effects on water relations of loblolly pine (Pinus taeda) stands, Tree Physiology, 30(3), 376–392, doi:10.1093/treephys/tpp118, 2010.

Gonzalez-Benecke, C. A., Gezan, S. A., Albaugh, T. J., Allen, H. L., Burkhart, H. E., Fox, T. R., Jokela, E. J., Maier, C. A., Martin, T. A., Rubilar, R. A. and Samuelson, L. J.: Local and general above-stump biomass functions for loblolly pine and slash pine trees, Forest Ecol Manag, 334, 254–276, doi:10.1016/j.foreco.2014.09.002, 2014.

Deleted: Gao, C., Wang, H., Weng, E., Lakshmivarahan, S., Zhang, Y. and Luo, Y.: Assimilation of multiple data sets with the ensemble Kalman filter to improve forecasts of forest carbon dynamics, Ecological Applications, 21(5), 1461–1473, doi:10.1890/09-1234.1, 2011.

Gonzalez-Benecke, C. A., Teskey, R. O., Martin, T. A., Jokela, E. J., Fox, T. R., Kane, M. B. and Noormets, A.: Regional validation and improved parameterization of the 3-PG model for Pinus taeda stands, Forest Ecol Manag, 361, 237–256, doi:10.1016/j.foreco.2015.11.025, 2016.

Hobbs, N. T. and Hooten, M. B.: Bayesian Models: A Statistical Primer for Ecologists, Princeton University Press, Princeton. 2015.

Iritz, Z. and Lindroth, A.: Energy partitioning in relation to leaf area development of short-rotation willow coppice, Agr Forest Meteorol, 81(1-2), 119–130, doi:10.1016/0168-1923(95)02306-2, 1996.

Keenan, T. F., Carbone, M. S., Reichstein, M. and Richardson, A. D.: The model–data fusion pitfall: assuming certainty in an uncertain world, Oecologia, 167(3), 587–597, doi:10.1007/s00442-011-2106-x, 2011.

Keenan, T. F., Davidson, E., Moffat, A. M., Munger, W. and Richardson, A. D.: Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling, Global Change Biology, 18(8), 2555–2569, doi:10.1111/j.1365-2486.2012.02684.x, 2012.

Landsberg, J. and Waring, R.: A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning, Forest Ecol Manag, 95(3), 209–228, doi:10.1016/S0378-1127(97)00026-1, 1997.

Le Quere, C., Moriarty, R., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Friedlingstein, P., Peters, G. P., Andres, R. J., Boden, T. A., Houghton, R. A., House, J. I., Keeling, R. F., Tans, P., Arneth, A., Bakker, D. C. E., Barbero, L., Bopp, L., Chang, J., Chevallier, F., Chini, L. P., Ciais, P., Fader, M., Feely, R. A., Gkritzalis, T., Harris, I., Hauck, J., Ilyina, T., Jain, A. K., Kato, E., Kitidis, V., Klein Goldewijk, K., Koven, C., Landsch utzer, P., Lauvset, S. K., Lef evre, N., Lenton, A., Lima, I. D., Metzl, N., Millero, F., Munro, D. R., Murata, A., Nabel, J. E. M. S., Nakaoka, S., Nojiri, Y., O'Brien, K., Olsen, A., Ono, T., P 'erez, F. F., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., R odenbeck, C., Saito, S., Schuster, U., Schwinger, J., S 'ef 'erian, R., Steinhoff, T., Stocker, B. D., Sutton, A. J., TAKAHASHI, T., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Vandemark, D., Viovy, N., Wiltshire, A., Zaehle, S. and Zeng, N.: Global Carbon Budget 2015, Earth Syst. Sci. Data, 7(2), 349–396, doi:10.5194/essd-7-349-2015, 2015.

LeBauer, D. S., Dietze, M., Long, S., Mulrooney, P., Rohde, G. S., Wang, D. and Kooper, R.: *Biofuel Ecophysiological Traits and Yields Database (BETYdb)*,, doi:doi:10.13012/J8H41PB9, 2010.

Lu, X., Lu, X., Kicklighter, D. W., Kicklighter, D., Melillo, J. M., Melillo, J. M., Reilly, J. M., Reilly, J. M., Xu, L. and Wu, L.: Land carbon sequestration within the conterminous United States: Regional- and state-level analyses, J Geophys Res-Biogeo, 120(2), 379–398, doi:10.1002/2014JG002818, 2015.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S. and Schimel, D. S.: Ecological forecasting and data assimilation in a data-rich era, Ecological Applications, 21(5),

Deleted: . Q., Luo, Y., Ogle, K.,

Deleted: Tucker, C., Fei, S.,

Deleted: Gao, C.,

Deleted: LaDeau, S., Clark, J. S.,

1429–1442, doi:10.1890/09-1275.1, 2011.

Luo, Y., Weng, E., Wu, X., Gao, C., Zhou, X. and Zhang, L.: Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models, Ecological Applications, 19(3), 571–574, doi:10.1890/08-0561.1, 2009.

MacBean, N., Peylin, P., Chevallier, F., Scholze, M. and Schürmann, G.: Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, Geosci. Model Dev., 9(10), 3569–3588, doi:10.5194/gmd-9-3569-2016, 2016.

Matamala, R., Gonzàlez-Meler, M. A., Jastrow, J. D., Norby, R. J. and Schlesinger, W. H.: Impacts of fine root turnover on forest NPP and soil C sequestration potential, Science, 302(5649), 1385–1387, doi:10.1126/science.1089543, 2003.

McCarthy, H. R., Oren, R., Johnsen, K. H., Gallet-Budynek, A., Pritchard, S. G., Cook, C. W., LaDeau, S. L., Jackson, R. B. and Finzi, A. C.: Re-assessment of plant carbon dynamics at the Duke free-air CO2 enrichment site: interactions of atmospheric [CO2] with nitrogen and water availability over stand development, New Phytol, 185(2), 514–528, doi:10.1111/j.1469-8137.2009.03078.x, 2010.

McKeand, S., Mullin, T., Byram, T. and White, T.: Deployment of genetically improved loblolly and slash pines in the south, Journal of Forestry, 101(3), 32–37, 2003.

Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., Hickler, T., Jain, A. K., Luo, Y., Parton, W., Prentice, I. C., Thornton, P. E., Wang, S., Wang, Y.-P., Weng, E., Iversen, C. M., McCarthy, H. R., Warren, J. M., Oren, R. and Norby, R. J.: Using ecosystem experiments to improve vegetation models, Nature Climate Change, 5(6), 528–534, doi:10.1038/nclimate2621, 2015.

Niu, S., Luo, Y., Dietze, M. C., Keenan, T. F., Shi, Z., Li, J. and III, F. S. C.: The role of data assimilation in predictive ecology, Ecosphere, 5(5), art65–16, doi:10.1890/ES13-00273.1, 2014.

Noormets, A., Gavazzi, M. J., McNulty, S. G., Domec, J.-C., Sun, G., King, J. S. and Chen, J.: Response of carbon fluxes to drought in a coastal plain loblolly pine forest, Glob Change Biol, 16(1), 272–287, doi:10.1111/j.1365-2486.2009.01928.x, 2010.

Novick, K. A., Oishi, A. C., Ward, E. J., Siqueira, M. B. S., Juang, J.-Y. and Stoy, P. C.: On the difference in the net ecosystem exchange of CO2 between deciduous and evergreen forests in the southeastern United States, Glob Change Biol, 21(2), 827–842, doi:10.1111/gcb.12723, 2015.

Oren, R., Ellsworth, D., Johnsen, K., Phillips, N., Ewers, B., Maier, C., Schafer, K., McCarthy, H., Hendrey, G., McNulty, S. G. and Katul, G.: Soil fertility limits carbon sequestration by forest ecosystems in a CO_2-enriched atmosphere, Nature, 411(6836), 469–472, doi:10.1038/35078064, 2001.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S. L., Rautiainen, A., Sitch, S. and Hayes, D.: A large and persistent carbon sink in

Deleted: 2011a. ... [31]

Deleted: 2011b

Deleted: Oleson, K., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, J. F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D. M., Sacks, W., Sun, Y., Tang, J. and Yang, Z. L.: Technical Description of version 4.5 of the Community Land Model (CLM), NCAR Technical Note NCAR, Boulder. 2013.

the world's forests, Science, 333(6045), 988–993, doi:10.1126/science.1201609, 2011.

Phillips, N. and Oren, R.: Intra- and inter-annual variation in transpiration of a pine forest, Ecological Applications, 11(2), 385–396, 2001.

Raymond, J. E., Fox, T. R., Strahm, B. D. and Zerpa, J.: Differences in the recovery of four different nitrogen containing fertilizers after two application seasons in pine plantations across the southeastern United States, Forest Ecol Manag, 380(C), 161–171, doi:10.1016/j.foreco.2016.08.044, 2016.

Ricciuto, D. M., Davis, K. J. and Keller, K.: A Bayesian calibration of a simple carbon cycle model: The role of observations in estimating and reducing uncertainty, Global Biogeochem. Cycles, 22(2), n/a–n/a, doi:10.1029/2006GB002908, 2008.

Richardson, A. D., Williams, M., Hollinger, D. Y., Moore, D. J. P., Dail, D. B., Davidson, E. A., Scott, N. A., Evans, R. S., Hughes, H., Lee, J. T., Rodrigues, C. and Savage, K.: Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints, Oecologia, 164, 25–40, doi:10.1007/s00442-010-1628-y, 2010.

Sabatia, C. O. and Burkhart, H. E.: Predicting site index of plantation loblolly pine from biophysical variables, Forest Ecol Manag, 326, 142–156, doi:10.1016/j.foreco.2014.04.019, 2014.

Samuelson, L. J., Butnor, J., Maier, C., Stokes, T. A., Johnsen, K. and Kane, M.: Growth and physiology of loblolly pine in response to long-term resource management: defining growth potential in the southern United States, Can. J. For. Res., 38(4), 721–732, doi:10.1139/X07-191, 2008.

Shvidenko, A., Barber, C. V. and Persson, R.: Forest and Woodland Systems, in Ecosystems and Human Well-being Current State and Trends, Volume, edited by R. Hassan, R. Scholes, and N. Ash, pp. 585–621, Island Press, Washington. 2005.

Deleted: Samuelson, L. J., Pell, C. J., Stokes, T. A., Bartkowiak, S. M., Akers, M. K., Kane, M., Markewitz, D., McGuire, M. A. and Teskey, R. O.: Two-year throughfall and fertilization effects on leaf physiology and growth of loblolly pine in the Georgia Piedmont, Forest Ecol Manag, 330, 29–37, doi:10.1016/j.foreco.2014.06.030, 2014.

Staff, S. S.: Soil Survey Geographic (SSURGO) Database for Southeastern U.S,, 2016.

Subedi, S., Fox, T. and Wynne, R.: Determination of fertility rating (FR) in the 3-PG model for loblolly pine plantations in the Southeastern United States based on site index, Forests, 6(9), 3002–3027, doi:10.3390/f6093002, 2015.

Tang, Z., Sayer, M. A. S., Chambers, J. L. and Barnett, J. P.: Interactive effects of fertilization and throughfall exclusion on the physiological responses and whole-tree carbon uptake of mature loblolly pine, Canadian Journal of Botany, 82(6), 850–861, doi:10.1139/b04-064, 2004.

Trudinger, C. M., Raupach, M. R., Rayner, P. J., Kattge, J., Liu, Q., Pak, B., Reichstein, M., Renzullo, L., Richardson, A. D., Roxburgh, S. H., Styles, J., Wang, Y.-P., Briggs, P., Barrett, D. and Nikolova, S.: OptIC project: An intercomparison of optimization techniques for parameter estimation in terrestrial biogeochemical models, J. Geophys. Res., 112(G2), G02027–17, doi:10.1029/2006JG000367, 2007.

Deleted: Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Ward, E. J., Domec, J.-C., Laviner, M. A., Fox, T. R., Sun, G., McNulty, S., King, J. and Noormets, A.: Fertilization intensifies drought stress: Water use and stomatal conductance of Pinus taeda in a midrotation fertilization and throughfall reduction experiment, Forest Ecol Manag, 355, 72–82, doi:10.1016/j.foreco.2015.04.009, 2015.

Weng, E. and Luo, Y.: Relative information contributions of model vs. data to short- and long-term forecasts of forest carbon dynamics, Ecological Applications, 21(5), 1490–1505, doi:10.1890/09-1394.1, 2011.

Will, R., Fox, T., Akers, M., Domec, J.-C., González-Benecke, C., Jokela, E., Kane, M., Laviner, M., Lokuta, G., Markewitz, D., McGuire, M., Meek, C., Noormets, A., Samuelson, L., Seiler, J., Strahm, B., Teskey, R., Vogel, J., WARD, E., West, J., Wilson, D. and Martin, T.: A range-wide experiment to investigate nutrient and soil moisture interactions in loblolly pine plantations, Forests, 6(6), 2014–2028, doi:10.3390/f6062014, 2015.

Williams, M., Schwarz, P., Law, B. E., Irvine, J. and Kurpius, M.: An improved analysis of forest carbon dynamics using data assimilation, Global Change Biology, 11(1), 89–105, doi:10.1111/j.1365-2486.2004.00891.x, 2005.

Zobitz, J. M., Desai, A. R., Moore, D. J. P. and Chadwick, M. A.: A primer for data assimilation with ecological models using Markov Chain Monte Carlo (MCMC), Oecologia, 167(3), 599–611, doi:10.1007/s00442-011-2107-9, 2011.

Deleted: Wightman, M., Martin, T., González-Benecke, C., Jokela, E., Cropper, W., Jr. and WARD, E.: Loblolly Pine Productivity and Water Relations in Response to Throughfall Reduction and Fertilizer Application on a Poorly Drained Site in Northern Florida, Forests, 7(10), 214–19, doi:10.3390/f7100214, 2016.

Deleted: Wutzler, T. and Carvalhais, N.: Balancing multiple constraints in model-data integration: Weights and the parameter block approach, J Geophys Res-Biogeo, 119(11), 2112–2129, doi:10.1002/2014JG002650, 2014.

Formatted: Font:12 pt

Formatted: Widow/Orphan control

Table 1. Regional observational data streams used in data assimilation.

Data stream	Measurement frequency	Measurement or estimation technique	Uncertainty	Stream ID for Table 3
Foliage biomass (Pine)	Annual or less	Allometric relationship	Based on propagating the allometric model uncertainty in Gonzalez-Benecke et al. 2014. Varied by observation.	1
Foliage biomass (hardwood)	Annual or less	Allometric relationship	Assumed zero	2
Stem biomass (pine)	Annual or less	Allometric relationship	Based on propagating the allometric model uncertainty in Gonzalez-Benecke et al. 2014. Varied by observation.	3
Stem biomass (hardwood)	Annual or less	Allometric relationship	Assumed zero	4
Coarse root biomass (combined)	Annual or less	Allometric relationship	Assumed zero*	5
Fine root biomass (combined)	Annual or less	Allometric relationship	SD = 10% of observation	6
Foliage biomass production (combined)	Annual	Litterfall traps	SD = 10% of observation	7
Fine root biomass production (combined)	Annual	Mini-rhizotrons	SD = 10% of observation	8
Pine stem density	Annual or less	Counting individuals	1% (assumed small)	9
Leaf area index (pine)	Monthly to annual	Litter traps or LI 2000	SD = 10% of observation	10
Leaf area index (hardwood)	Monthly to annual	Litter traps or LI 2000	SD = 10% of observation	11
Leaf area index (combined)	Only used if not separated into pine and hardwood	Litter traps or LI 2000	SD = 10% of observation	12
Gross Ecosystem Production	Monthly	Modeled from flux eddy-covariance net ecosystem exchange	SD = 10% of observation	13
Evapotranspiration	Monthly	Eddy-covariance	SD = 10% of observation	14

*the relatively low number of observations prevented convergence when using the observational uncertainty model so observational uncertainty was assumed to be zero to allow convergence.

Moved (insertion) [10]

Formatted Table

Deleted: Table 1. A subset of parameters optimized using data assimilation, prior distributions, and the sensitivity of total biomass at age 25 to the parameter. These are the parameters referred to in the results and discussion, other optimized model parameter can be found in the supplemental material.

Deleted: Parameter

Deleted: Parameter description

Deleted: Units

Deleted: Sensitivity*

Deleted: Prior distribution

Deleted Cells

Deleted: Prior parameters

Deleted Cells

Deleted: Reference

Deleted: prior

Deleted: α ... [32]

Deleted: Standard deviation (SD) = 10% of observation

Deleted: 2.5

Deleted: turnover

Deleted: turnover

Deleted: count

Deleted: If litter trap method: SD = 2.5% of obser ... [33]

Deleted: If litter trap method: SD = 2.5% of obser ... [34]

Deleted: If litter trap method: SD = 2.5% of obser ... [35]

Deleted:

Formatted: Widow/Orphan control

Table 2. Descriptions of the studies used in data assimilation.

Study name	Number of locations	Number of plots per site	Experimental treatments (plots)	Data streams (Table 2)	Measurement Years	Measurement Stand Ages (years)	Reference
FMRC[1] Thinning Study	163	1	None	1, 3,9	1981 - 2003	8 - 30	(Burkhart et al., 1985)
FPC[2] Region-wide 18	18	2	Nutrient addition	1, 3,9	2011-2014	12-21	(Albaugh et al., 2015)
PINEMAP[3]	4	16	Nutrient addition, 30% throughfall, Nutrient x throughfall	1, 3,9	2011-2015	3 – 13	(Will et al., 2015)
Waycross	1	2	Nutrient addition	3,9,10	1991-2010	4-23	(Bryars et al., 2013)
SETRES[4]	1	16	Nutrient addition, irrigation, nutrient x irrigation	1,3,5,6,9, 10	1991-2006	8 - 23	(Albaugh et al., 2004)
Duke FACE[5] and US-DK3 Flux	1	12	CO_2, nutrient addition, CO_2 x nutrient addition	2,3,4,5,6, 7,8,9,10, 11,13,14	1996-2004	13-22	(McCarthy et al., 2010; Novick et al., 2015)
NC2 Flux	1	1	None	2,3,4,5,6, 7,9,10,11 ,12,13,14	2005-2014	12-22	(Noormets et al., 2010)
Total	187	294			1981 - 2014	4 - 30	

[1] Forest Modeling Research Cooperative; [2] Forest Productivity Cooperative; [3] Pine Integrated Network: Education, Mitigation, and Adaptation project (PINEMAP); [4] Southeast Tree Research and Education Site; [5] Free Air Carbon Enrichment

Deleted: 3

Deleted: (Burkhart et al., 1985)

Deleted: (Albaugh et al., 2015)

Deleted: (Will et al., 2015)

Deleted: (Bryars et al., 2013)

Deleted: (Albaugh et al., 2004)

Deleted: (McCarthy et al., 2010; Novick et al., 2015)

Deleted: flux

Deleted: (Noormets et al., 2010)

Deleted:

Table 3. The prior distributions of all 3-PG model parameters optimized using data assimilation

Parameter	Parameter description	Units	Prior distribution	Prior parameters	Reference for prior
Allocation and structure					
pFS2	Ratio of foliage to stem allocation at stem diameter = 2 cm	-	uniform	min = 0.08 max = 1.00	vague
pFS20	Ratio of foliage to stem allocation at stem diameter = 20 cm	-	uniform	min = 0.10 max =1.00	vague
pRF	Ratio of fine roots to foliage allocation	-	uniform	min = 0.05 max = 2.00	vague
pCRS	Ratio of coarse roots to stem allocation	-	uniform	min = 0.15 max = 0.35	1
SLA0	Specific leaf area at stand age 0	$m^2 kg^{-1}$		mean =5.53 sd = 0.44	2
SLA1	Specific leaf area for mature aged stands	$m^2 kg^{-1}$	normal	mean = 3.58 sd = 0.11	2
tSLA	Age at which specific leaf area = ½(SLA0 + SLA1)	Years	normal	mean = 5.97 sd = 2.15	2
fCpFS700	Proportional decrease in allocation to foliage between 350 and 700 ppm CO_2	-	uniform	min = 0.50 max = 1.00	vague
StemConst	Constant in stem mass vs. diameter relationship	-	normal	mean = 0.022 sd = 0.005	3
StemPower	Power in stem mass vs. diameter relationship	-	normal	mean = 2.77 sd = 0.2	3
Canopy photosynthesis, autotrophic respiration, and transpiration					
alpha	Canopy quantum efficiency (pines)	mol C mol PAR^{-1}	uniform	min = 0.02 max = 0.06	vague
y	Ratio NPP/GPP	-	uniform	min = 0.30 max= 0.65	4
MaxCond	Maximum canopy conductance	$m s^{-1}$	uniform	min = 0.005 max = 0.03	2
LAIgcx	Canopy LAI for maximum canopy conductance	-	uniform	min = 2 max = 5	2,5,6
Environmental modifiers of photosynthesis and transpiration					
kF	Reduction rate of production per degree Celsius below zero	-	normal	mean = 0.18 sd = 0.016	2
Tmin	Minimum monthly mean temperature for growth	°C	normal	mean = 4.0 sd = 2.0	2,5,6
Topt	Optimum monthly mean temperature for growth	°C	normal	mean = 25.0 sd = 2.0	2,5,6
Tmax	Maximum monthly mean temperature for growth	°C	normal	mean = 38.0 sd = 2.0	2,5,6
SWconst	Moisture ratio deficit when downregulation is 0.5	-	uniform	min = 0.01 max = 1.8	vague
SWpower	Power of moisture ratio deficit	-	uniform	min = 1 max= 13	vague
CoeffCond	Defines stomatal response to VPD	$mbar^{-1}$	normal	mean = 0.041 sd = 0.003	2

fCalpha700	Proportional increase in canopy quantum efficiency between 350 and 700 ppm CO$_2$	-	uniform	min = 1.00 max = 1.8	vague
MaxAge	Maximum stand age used to compute relative age	Years	uniform	min = 16 max =200	vague
nAge	Power of relative age in fage	-	uniform	min = 0.2 max = 4.0	vague
rAge	Relative age to where fage = 0.5	-	uniform	min = 0.01 max = 3.00	vague
FR1	Fertility rating parameter 1 (mean annual temperature coefficient)	-	uniform	min = 0.0 max = 1.0	vague
FR2	Fertility rating parameter 2 (site index age 25 coefficient)	-	uniform	min = 0.0 max = 1.0	vague
Mortality					
wSx1000	Maximum stem mass per tree at 1000 trees/ha	kg tree^{-1}	normal	mean = 235 sd = 25	2,5,6
ThinPower	Power in self thinning law	-	uniform	min = 1.0 max = 2.5	2,5,6
ms	Fraction of mean stem biomass per tree on dying trees	-	uniform	min = 0.1 max = 1.0	vague
Rttover	Average monthly root turnover rate	Month^{-1}	uniform	min = 0.017 max = 0.042	7
MortRate	Density independent mortality rate (pines)	Month^{-1}	uniform	min = 0.0002 max = 0.004	vague
Understory hardwoods					
alpha_h	Canopy quantum efficiency (understory hardwoods)	mol C mol PAR^{-1}	uniform	min = 0.005 max = 0.07	vague
pFS_h	Ratio of foliage to stem partitioning (understory hardwoods)	-	uniform	min = 0.2 max = 3.0	vague
pR_h	Ratio of foliage to fine roots (understory hardwoods)	-	uniform	min = 0.05 max = 2	vague
SLA_h	Specific leaf area (understory hardwoods)	m^2 kg^{-1}	normal	mean = 16 sd = 3.8	8
fCalpha700_h	Proportional increase in canopy quantum efficiency between 350 and 700 ppm CO$_2$ (understory hardwood)	-	uniform	min = 1.00 max = 2.5	vague

[1](Albaugh et al., 2005); [2](Gonzalez-Benecke et al., 2016); [3](Gonzalez-Benecke et al., 2014) [4](DeLucia et al., 2007);[5](Bryars et al., 2013);[6](Subedi et al., 2015);[7](Matamala et al., 2003);[8](LeBauer et al., 2010)

Table 4. Description of the different data assimilation approaches used.

Simulation Name	Treatments included in assimilation	Number of plots
Base	All plots and experiments in the region were used simultaneously. Includes unique pCRS, wSx1000, and ThinPower parameters for plots in the Duke FACE study	294
NoExp	Same as Base assimilation but excluding all plots with experimental manipulations. Includes control plots that are part of experimental studies.	208
NoDkPars	Same as Base assimilation but without pCRS, wSx1000, and ThinPower parameter for plots in the Duke FACE and US-DK3 studies	294

Deleted: All

Deleted: 1-stage data assimilation.

Deleted: DK+NC2 ... [36]

Formatted Table

Deleted: 1st stage of 2-stage

Deleted: . Same as DK+NC2

Deleted: 10

Deleted: without nutrient fertilization

Deleted: RW

Deleted: 2nd stage of 2-stage assimilation. Region-wide assimilation of FRMC, FPC, PINEMAP, Waycross, and SETRES sites. Uses the posteriors of the DK+NC2 simulation as priors. Includes nutrient addition and water manipulation experiments. This simulation is repeated four times for four different out-of-sample validation plots.

Deleted: 281

Deleted: RW-fert ... [37]

Table 5. The optimized medians, range of the 99% quantile intervals of the posterior distributions and the 99% quantile range for priors with normally distributed priors or the range of the upper and lower bounds for priors with uniform distributions.

Parameter	Posterior median	Posterior 99% C.I. range	Prior range	Posterior/Prior Range
Allocation and structure				Parameter group mean = 0.38
pFS2	0.58	0.55 - 0.61	0.08 – 1.00	0.06
pFS20	0.57	0.55 - 0.59	0.10 – 1.00	0.05
pR	0.11	0.07 - 0.15	0.05 – 2.00	0.04
pCRS	0.26	0.25 - 0.27	0.15 - 0.35	0.11
pCRS (Duke)	0.21	0.18 - 0.23	0.15 - 0.35	0.20
SLA0	8.44	7.67 - 9.25	4.4 - 6.66	0.70
SLA1	2.84	2.72 - 2.96	3.59 - 4.16	0.43
tSLA	4.13	3.88 - 4.41	0.43 - 11.51	0.05
fCpFS700	0.74	0.60 - 0.90	0.50 – 1.00	0.60
StemConst	0.022	0.009 - 0.035	0.009 - 0.035	1.00
StemPower	2.78	2.29 - 3.27	2.25 - 3.29	0.95
Canopy photosynthesis, autotrophic respiration, and transpiration				Parameter group mean = 0.14
alpha	0.029	0.026 - 0.031	0.02 - 0.06	0.14
y	0.50	0.47 - 0.53	0.30 - 0.65	0.15
MaxCond	0.011	0.01 - 0.012	0.005 - 0.03	0.09
LAIgcx	2.2	2.0 - 2.48	2.0 - 5 .0	0.16
Environmental modifiers of photosynthesis and transpiration				Parameter group mean = 0.61
kF	0.16	0.12 - 0.2	0.14 - 0.22	1.04
Tmin	-5.56	-8.88 - -2.69	-1.15 - 9.15	0.60
Topt	23.42	21.1 - 26.31	19.85 - 30.15	0.51
Tmax	39.56	34.71 - 44.39	32.85 - 43.15	0.94
SWconst	1.09	0.91 - 1.56	0.01 - 1.8	0.36
SWpower	8.86	3.39 - 12.98	1.00 – 13.00	0.80
CoeffCond	0.036	0.029 - 0.043	0.034 - 0.048	0.91
fCalpha700	1.33	1.18 - 1.52	1.0 - 1.80	0.43
MaxAge	151.5	54.4 - 199.6	16.0 - 200 .0	0.79

Deleted: Table 5. Posterior means and 95% credib ... [38]
Formatted Table ... [39]
Deleted: RW
Deleted: RW-fert
Deleted: RW-water
Deleted: DK+NC2
Deleted: All
Deleted Cells ... [40]
Deleted: α
Formatted Table ... [41]
Deleted: 037 ... [42]
Deleted: 040 ... [43]
Deleted: 037 ... [44]
Deleted: 035 ... [45]
Deleted: 0.032 ... [46]
Deleted Cells ... [47]
Deleted: y
Formatted Table ... [48]
Moved (insertion) [13] ... [49]
Inserted Cells ... [50]
Deleted: 48 ... [51]
Deleted: 48 ... [52]
Deleted: 48 ... [53]
Deleted: 0.48 ... [54]
Deleted Cells ... [55]
Deleted: 0.52 ... [56]
Deleted Cells ... [57]
Moved down [15]: 0.50
Deleted: – 0.54)
Moved down [14]: 0.51
Deleted:)
Moved (insertion) [15] ... [58]
Moved (insertion) [14] ... [59]
Moved (insertion) [16] ... [60]
Formatted Table ... [61]
Deleted: 31 ... [62]
Deleted: 31 ... [63]
Deleted: 31 ... [64]
Deleted: 1.32 ... [65]
Deleted: 1.11 ... [66]
Deleted Cells ... [67]
Deleted: fCpFS700
Deleted: 0.84 ... [68]
Deleted: 0.83 ... [69]
Deleted: .84 ... [70]
Deleted: 84 ... [71]
Moved down [17]: 0.76
Deleted: -0.93)
Deleted: 0.99 ... [72]

nAge	3.35	1.77 - 3.99	1.00 – 4.00	0.74
rAge	2.25	0.81 - 2.99	0.01 – 3.00	0.73
FR1	0.073	0.061 - 0.086	0.00 – 1.00	0.03
FR2	0.17	0.15 - 0.19	0.0 – 1.0	0.04
Mortality				Parameter group mean = 0.37
wSx1000	176.9	169.6 - 184.4	165.6 - 294.4	0.15
wSx1000 (Duke)	243.3	196.89 - 305.02	165.6 - 294.4	0.76
ThinPower	1.68	1.60 - 1.78	1.00 - 2.5	0.12
ThinPower v(Duke)	1.26	1.00 - 1.85	1.00 - 2.5	0.56
mS	0.52	0.37 - 0.71	0.10 – 1.00	0.38
Rttover	0.023	0.017 - 0.031	0.017 - 0.042	0.55
MortRate	0.001	9e-04 - 0.0011	2e-04 - 0.004	0.06
Understory hardwoods				Parameter group mean = 0.28
alpha_h	0.02	0.02 - 0.02	0.005 - 0.07	0.01
pFS_h	1.78	1.54 - 2.06	0.2 – 3.0	0.19
pR_h	0.21	0.06 - 0.43	0.05 – 2.00	0.19
SLA_h	16.3	14.1 – 19.0	6.2 - 25.8	0.25
fCalpha700_h	1.84	1.58 - 2.17	1.0 – 2.50	0.74

Deleted: SWconst

Deleted: 31 ... [75]

Inserted Cells ... [73]

Deleted: 48 ... [74]

Deleted: 1.8 ... [76]

Deleted: 1.30 ... [77]

Deleted: 1.57 ... [78]

Deleted Cells ... [79]

Moved up [16]: 1.09

Deleted: SWpower

Deleted: 1.61 ... [80]

Deleted: 1.29 ... [81]

Deleted Cells ... [82]

Deleted: 93 ... [83]

Deleted: 2.20 ... [84]

Deleted: 1.47 ... [85]

Deleted: – 2.26)

Inserted Cells ... [86]

Deleted: 094 ... [87]

Deleted: 096 ... [88]

Deleted: 118 ... [89]

Deleted: not fit

Deleted Cells ... [90]

Deleted: 0.094 ... [91]

Deleted Cells ... [92]

Deleted: 144 ... [93]

Deleted: 124 ... [94]

Deleted: 179 ... [95]

Deleted: not fit

Deleted: 0.153 ... [96]

Deleted: ... [98]

Deleted: 180 ... [99]

Deleted: 180 ... [100]

Deleted: 258 ... [101]

Deleted: 181 ... [102]

Deleted Cells ... [103]

Formatted Table ... [97]

Moved (insertion) [17] ... [104]

Deleted: thinPower

Formatted Table ... [105]

Deleted: 67 ... [106]

Deleted: 70 ... [107]

Deleted: 71 ... [108]

Deleted: 1.28 ... [109]

Deleted: 1.61 ... [110]

Deleted Cells ... [111]

Moved (insertion) [18] ... [112]

Deleted: pCRS

Formatted Table ... [113]

Moved up [13]: 0.26

Deleted: ... [114]

Inserted Cells ... [115]

Deleted: 0.24 ... [116]

... [117]

... [118]

Deleted Cells ... [119]

... [120]

Deleted Cells ... [121]

... [122]

Table 6. Median and range of the 99% quantile intervals of the posterior distributions for the parameters in the NoExp and NoDkPars assimilations

Parameter	NoExp median	NoExp 99% range	NoDkPars median	NoDkPar 99%
Allocation and structure				
pFS2	0.63	0.61 - 0.68	0.57	0.55 - 0.60
pFS20	0.63	0.60 - 0.65	0.57	0.55 - 0.59
pR	0.11	0.06 - 0.16	0.11	0.08 - 0.15
pCRS	0.29	0.27 - 0.30	0.26	0.25 - 0.27
pCRS (Duke)	0.25	0.23 - 0.28	N/A	N/A
SLA0	7.47	6.57 - 8.41	8.56	7.73 - 9.32
SLA1	3.00	2.88 - 3.12	2.89	2.79 - 2.99
tSLA	4.75	4.30 - 5.26	4.12	3.90 - 4.38
fCpFS700	0.50	0.50 - 0.53	0.94	0.83 – 1.00
StemConst	0.022	0.01 - 0.04	0.02	0.01 - 0.04
StemPower	2.79	2.27 - 3.26	2.77	2.28 - 3.30
Canopy photosynthesis, autotrophic respiration, and transpiration				
alpha	0.030	0.028 - 0.033	0.029	0.026 - 0.031
y	0.48	0.45 - 0.51	0.49	0.46 - 0.52
MaxCond	0.017	0.015 - 0.021	0.011	0.011 - 0.012
LAIgcx	4.4	3.9 – 5.0	2.1	2.0 - 2.5
Environmental modifiers of photosynthesis and transpiration				
kF	0.15	0.11 - 0.20	0.16	0.11 - 0.20
Tmin	-7.8	-10.97 - -4.95	-6.04	-9.06 - -3.03
Topt	21.55	19.15 - 24.39	22.71	20.54 - 25.42
Tmax	40.56	36.51 - 45.62	39.82	35.62 - 44.56
SWconst	0.93	0.8 - 1.1	1.14	0.91 - 1.62
SWpower	6.27	2.98 - 11.49	7.99	3.29 - 12.95
CoeffCond	0.041	0.034 - 0.047	0.036	0.030 - 0.042
fCalpha700	1.01	1.0 0- 1.06	1.15	1.10 - 1.25
MaxAge	152.84	54.18 - 199.5	152.0	49.2 - 199.3
nAge	3.36	1.93 - 3.99	3.36	1.89 - 3.99
rAge	2.26	0.80 - 2.99	2.24	0.83 - 2.99
FR1	0.12	0.09 - 0.14	0.08	0.07 - 0.09
FR2	0.20	0.16 - 0.24	0.17	0.15 - 0.19
Mortality				
wSx1000	191.6	180.2 - 210.2	181.32	173.26 - 196.32
wSx1000 (Duke)	235.1	175.0 - 297.5	N/A	N/A
ThinPower	1.76	1.61 - 1.92	1.59	1.46 - 1.72
ThinPower (Duke)	1.42	1.01 - 2.02	N/A	N/A
mS	0.54	0.33 - 0.80	0.5	0.25 - 0.71
Rttover	0.019	0.02 - 0.03	0.022	0.017 - 0.030
MortRate	0.0013	0.0011 - 0.0014	0.0011	9e-04 - 0.0013
Understory hardwoods				
alpha _h	0.031	0.025 - 0.040	0.02	0.017 - 0.023
pFS_h	2.39	1.86 - 2.96	1.79	1.59 - 2.09
pR_h	0.25	0.05 - 0.67	0.21	0.06 - 0.41

SLA_h	12.37	9.96 - 15.07	16.42	14.37 - 18.55
fCalpha700_h	1.08	1.00 - 1.83	1.83	1.56 - 2.15

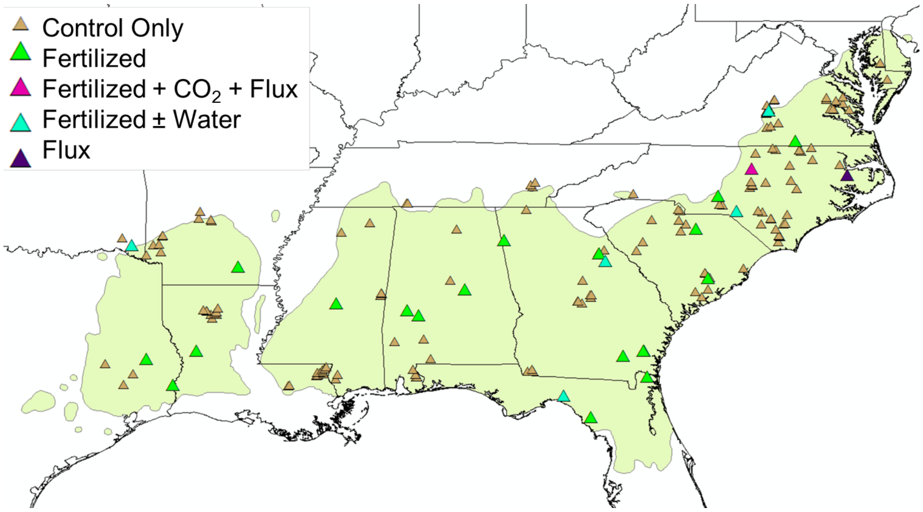


Figure 1. Map of loblolly pine distribution, plot locations used in data assimilation, and the experiment type associated with each plot. The control-only treatments were plots without any associated experimental treatment or flux measurements. Fertilized were plots with nutrient additions. CO$_2$ were plots with free-air concentration enrichment treatments. The flux treatments were plots with eddy-covariance measurements of ecosystem-scale carbon and water exchange. The water treatments included throughfall exclusion and irrigation experiments.

Moved (insertion) [19]

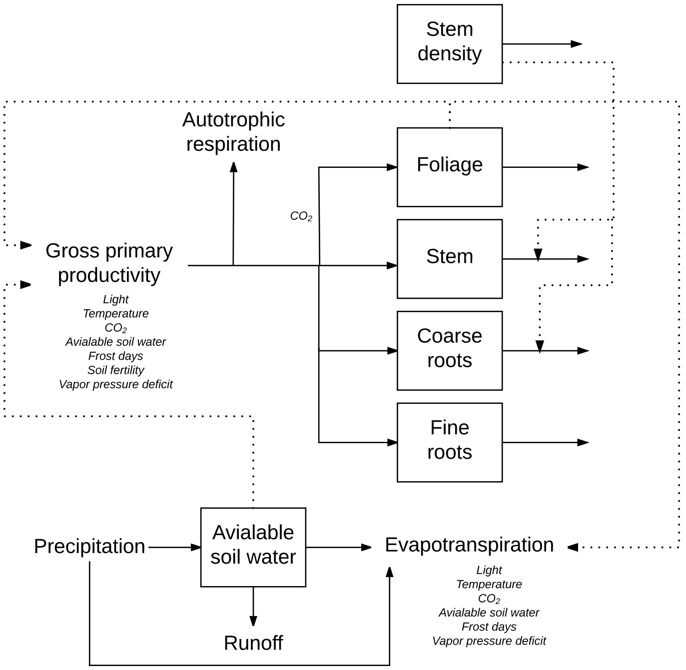


Figure 2. A diagram of the monthly time-step 3-PG model used in this study. The stocks are represented by the boxes and the fluxes by the arrows. An influence of a stock on a flux that is not directly related to that stock is represented by the dotted lines. The environmental influences on a flux is described using italics. A description of the model can be found in the supplemental information.

Formatted: Font:12 pt

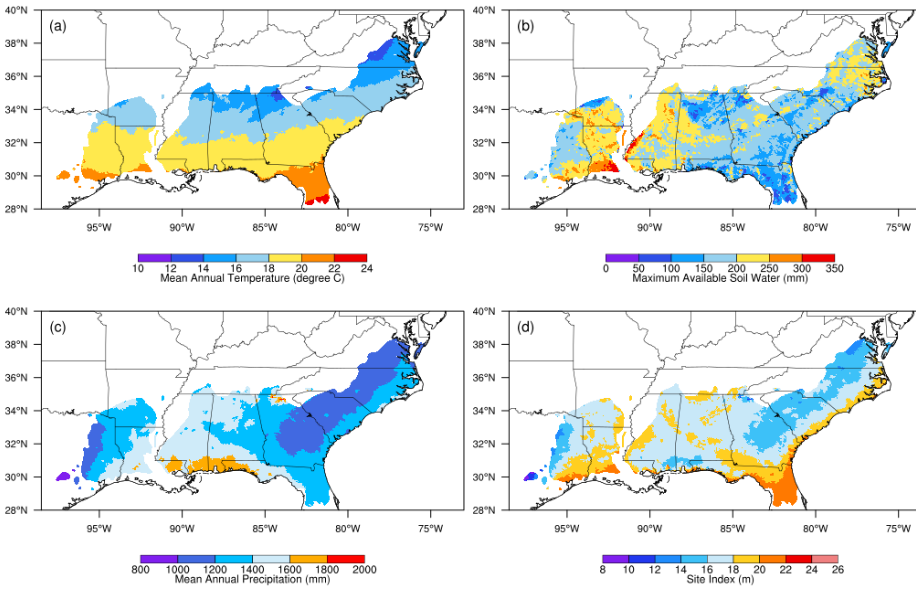


Figure 3. Key climatic and stand characteristic inputs to the regional 3-PG simulations: (a) Mean annual temperature (1979-2011) as a summary of the gradient in monthly temperature inputs used in simulations, (b) maximum available soil water for the top 1.5 meters of soil from SSURGO, (c) mean annual precipitation (1979-2011) as a summary of the gradient in monthly precipitation inputs used in simulations, and (d) site index. The area shown is the natural range of loblolly pine (*Pinus taeda L.*).

Deleted: ... [123]

Deleted: 2

Deleted: The focal site in Georgia highlighted in Figures 5c and 6b is represented by the circle containing the dot.

Deleted: *<sp><sp>* ... [124]

Moved up [19]: Map of loblolly pine distribution, plot locations used in data assimilation, and the experiment type associated with each plot. The control-only treatments were plots without any associated experimental treatment or flux measurements. Fertilized were plots with nutrient additions. CO_2 were plots with free-air concentration enrichment treatments. The flux treatments were plots with eddy-covariance measurements of ecosystem-scale carbon and water exchange. The water treatments included throughfall exclusion and irrigation experiments.

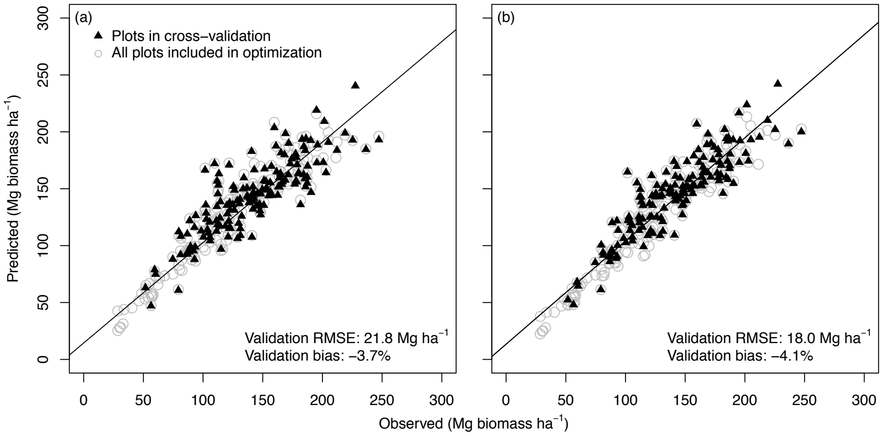
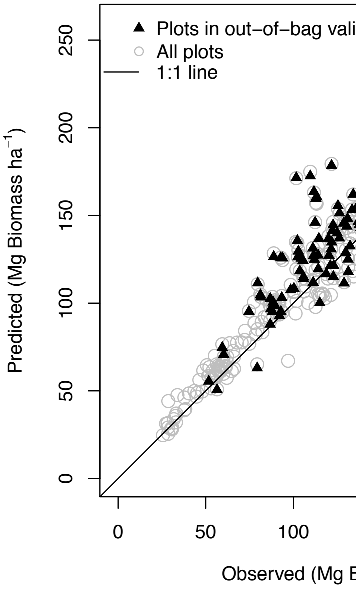


Figure 4. Model evaluation of stem biomass when assimilating (a) observations across environmental gradients and ecosystem manipulation experiments (Base; Table 4), and (b) assimilation only observations across environmental gradients (NoExp; Table 4). The gray circles correspond to predictions where all plots were used in data assimilation. The black triangles correspond to predictions where 160 plots were not included in data assimilation and represent an independent evaluation of model predictions (out-of-bag validation). For each plot, we used the measurement with the longest interval between initialization and measurement for evaluation.

Deleted:



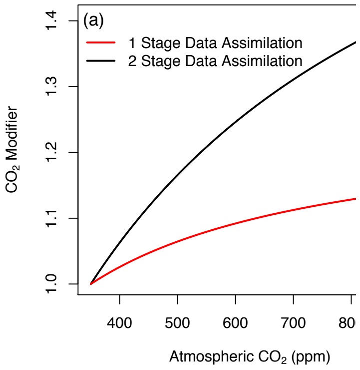
Deleted: using the RW simulation described in

Deleted: 5.

Deleted: 120

Deleted: Page Break

Deleted:



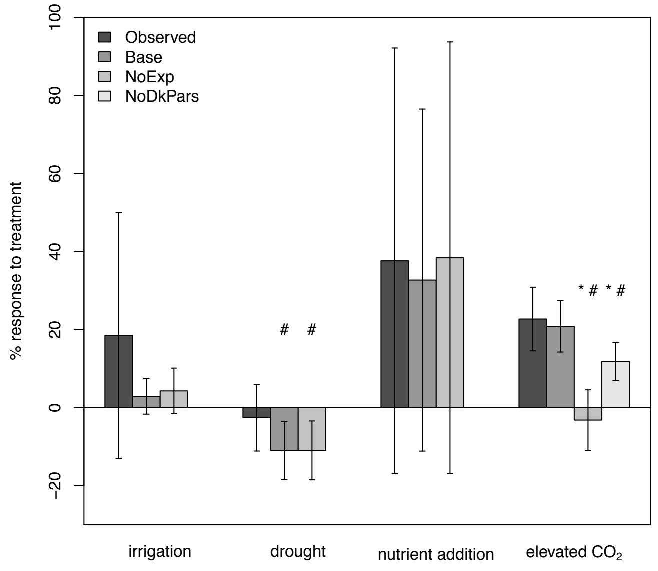


Figure 5. The mean response, expressed as a percentage change in stem biomass from the control treatment, for irrigation, drought (as a reduction in throughfall), nutrient addition, and elevated CO_2 experiments. The observed response and the response simulated by the Base, NoExp, and NoDkPars assimilation approaches are shown. # signifies that value below was significantly different from the observed ($p < 0.05$). * signifies that value was significantly from the Base assimilation ($p<0.05$). Error bars are ±1 standard deviation.

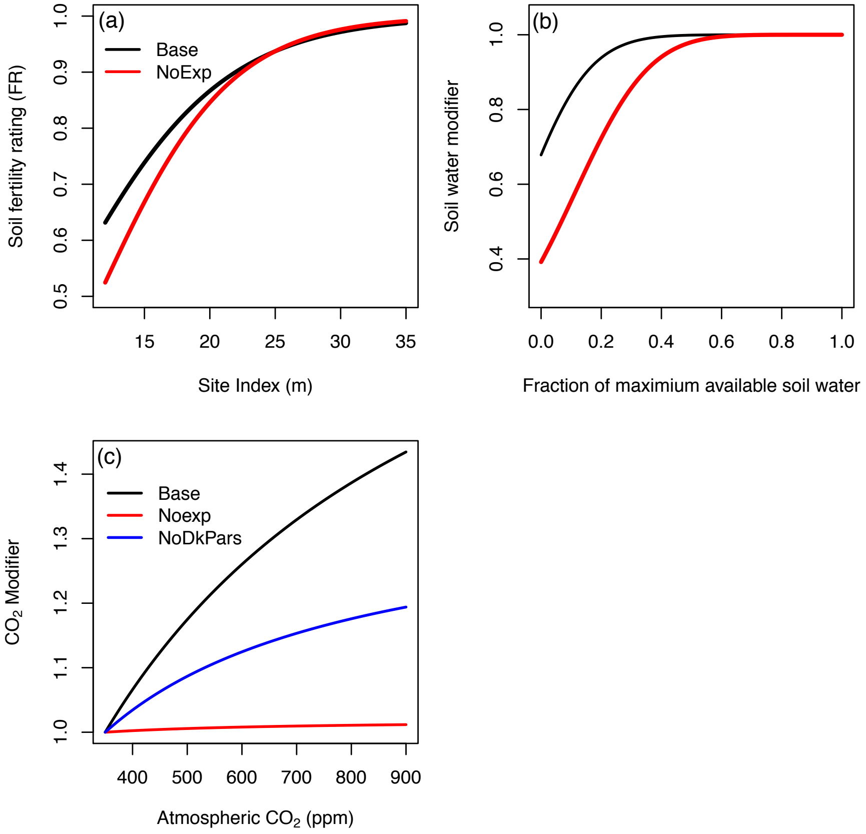


Figure 6. Optimized environmental response functions in the 3-PG model for the (a) soil fertility influence on photosynthesis), (b) available soil water influence on photosynthesis and conductance, and (c) atmospheric CO_2 influence on photosynthesis. The function shapes were derived from the parameters in the Base, NoExp, and NoDkPars assimilations (Table 4).

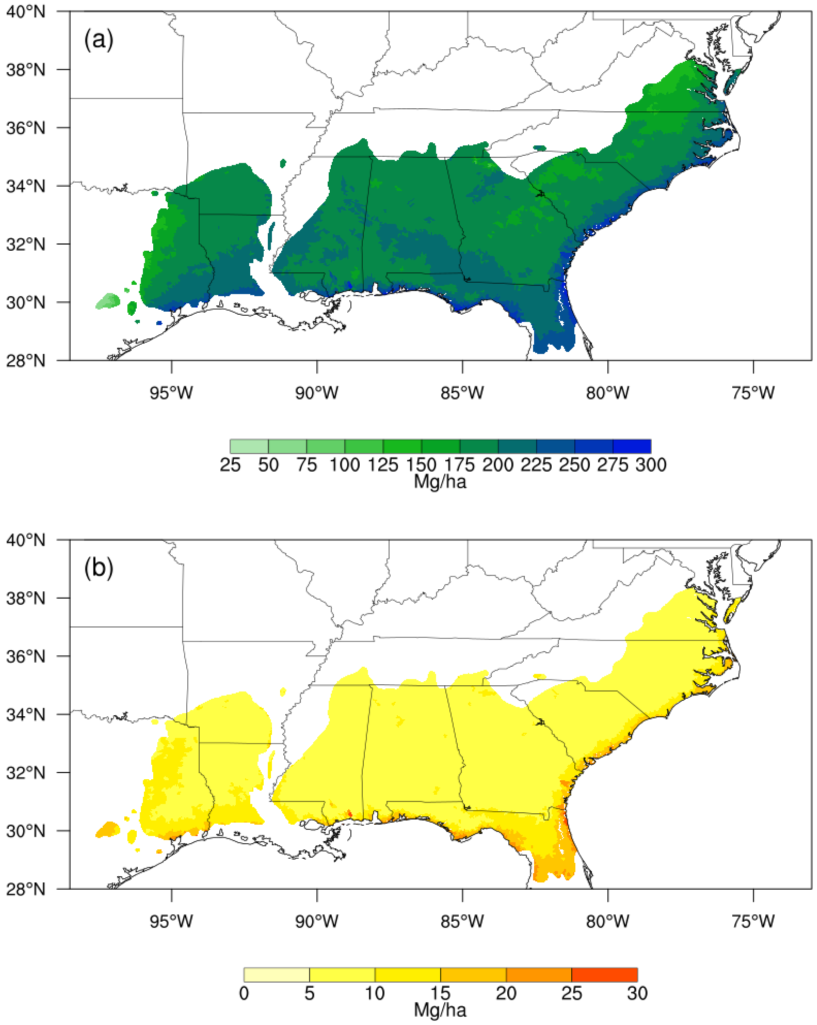


Figure 7. (a) Regional predictions of stem biomass stocks for a 25-year-old stand planted in 1985. Parameters used in the predictions were from the Base assimilation approach described in Table 5. (b) The width of the 95% quantile interval associated with uncertainty in model parameters.

Moved down [20]: Figure 8.

Deleted: 1979

Deleted: the predictions were from the RW data assimilation approach described in Table 5. The focal site in Georgia highlighted in Figures 5c and 6b is represented by the circle containing the dot.

Deleted: RW data

Deleted: The focal site in Georgia highlighted in Figures 5c and 6b is represented by the circle containing the dot.

Page Break

Deleted:

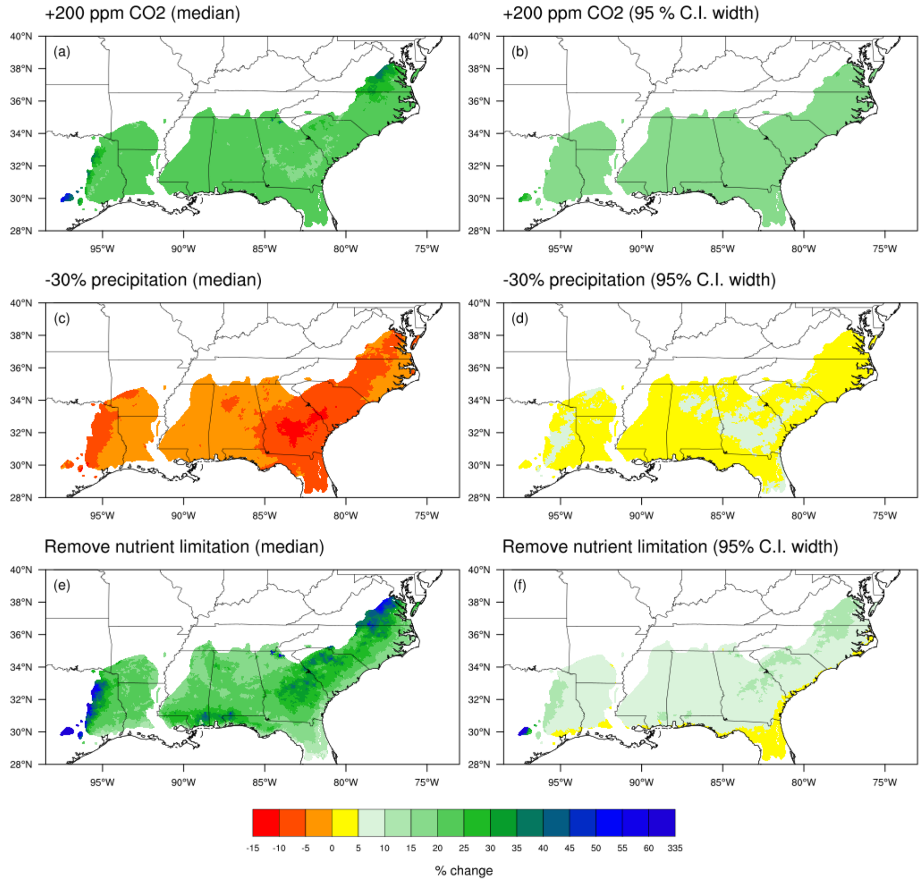


Figure 8. Predictions of the percentage change in stem biomass at age 25 in response to (a,b) a 200 ppm increase in atmospheric CO_2 over 1985-2011 concentrations, (c,d) a 30% reduction in precipitation from 1985-2011 levels, and (e,f) a removal of nutrient limitation by setting the soil fertility rating in the model equal to 1. The left column is the median prediction and the right column is the width of the 95% quantile interval associated with parameter uncertainty. The predictions used the Base assimilation.

Moved (insertion) [20]

Deleted: change in stem biomass of a 25-year stand when nutrient limitation is completely removed through nutrient addition (simulated by setting FR = 1). Predictions from data assimilation that included nutrient addition experiments are shown in (a) and prediction data assimilation that did not include nutrient addition experiments are shown in (b). The focal site in Georgia highlighted in Figures 5c and 6b is represented by the circle containing the dot. ... [125]

Forest ecosystems absorb and store a large fraction of anthropogenic carbon dioxide (CO_2) emissions (Le Quere et al., 2015; Pan et al., 2011) and supply wood products to a growing human population (Shvidenko et al., 2005). Therefore, predicting future carbon sequestration and timber supply is critical for adapting forest management practices to future environmental conditions and for using forests to assist with reduction of atmospheric CO_2 concentrations. The key sources of information for developing these predictions are results from global change ecosystem manipulation experiments, observations of forest dynamics across environmental gradients, and process-based ecosystem models. The challenge is integrating these three tools into a common framework for creating probabilistic predictions, or forecasts (based on (Luo et al., 2011a)), that provide information on both the expected future state of the forest and the probability distribution of those future states.

Data assimilation (DA), or data-model fusion, is an increasingly used framework for integrating ecosystem observations into ecosystem models (Luo et al., 2011a; Niu et al., 2014; Williams et al., 2005). DA integrates observations with ecosystem models through statistical, often Bayesian, methods that generate probability distributions for ecosystem model parameters and initial states. DA allows for the explicit accounting of observational uncertainty (Keenan et al., 2011), the incorporation of multiple types of observations with different time scales of collection (Richardson et al., 2010), and the representation of prior knowledge through informed parameter prior distributions or specific relationships among parameters (Bloom and Williams, 2015). Using DA to parameterize ecosystem models with observations from multiple locations that leverage environmental gradients and from ecosystem manipulation experiments will allow for forecasts to be consistent with the rich history of global change research in forest ecosystems.

Ecosystem manipulation experiments provide a controlled environment in which data collected can be used to describe how forests acclimate and operate under altered environmental conditions (Medlyn et al., 2015). These data may be used to constrain model parameters that are associated with specific physiological functions associated with, for example, carbon allocation and turnover as related to the controlled manipulation. Furthermore, the assimilation of experiments may increase parameter identifiability (reducing equifinality (Luo et al., 2009)), where two parameters have compensating controls on the same processes, by isolating the response to a manipulated driver. For example, carbon assimilation and primary productivity can be modeled as a light and temperature controlled process that is adjusted by nutrients, water, and atmospheric CO_2 concentration. In this case, the productivity may mathematically be equal between a parameterization that has high potential conversion of light to photosynthesis (high quantum yield) but low relative nutrient availability and a parameterization with low quantum yield but high relative nutrient availability. Therefore, the challenge is that the same rate of production can emerge from different contributions of environmental controls.

For future predictions with changing environmental conditions, the relative contribution of each environmental control should be separated in order to correctly parameterize the sensitivity to changes in the environment. Key examples of existing and past ecosystem experiments that have the potential to isolate specific parameters in DA include CO_2 enrichment, water manipulation, nutrient addition, and elevated soil temperature experiments. Many of these experiments are common, particularly when including nutrient addition experiments in managed forests. Other types of experiments are less common, but the few sites with the experiments, such as whole-ecosystem CO_2 enrichment, include intensive measurements of numerous carbon pools and fluxes required for model optimization.

Developing optimized parameters that apply to a region requires assimilating observations that span environmental gradients to support the application of model predictions to a range of climatic conditions, nutrient availabilities, and soil water dynamics. Therefore, the DA of multiple research sites across a region is an important extension of prior DA research focused on DA at a single site with multiple types of observations (Keenan et al., 2012; Richardson et al., 2010; Weng and Luo, 2011). Incorporating multiple locations that include global change experiments in DA is associated with numerous challenges. First, prior research has demonstrated that high frequency observations (i.e., daily, or more frequent, net ecosystem exchange observations) can overwhelm the contribution of low frequency observations (i.e., annual tree diameter measurements) to the cost-function used for optimization (Richardson et al., 2010), resulting in a parameter set that predominately represents the high-frequency dynamics. DA of ecosystem experiments and regional observations can present similar issues because key contrasts isolated in an ecosystem experiment with relatively few plots may be overwhelmed by the contribution of more numerous regional observations from non-manipulated plots. For example, whole ecosystem CO_2 enrichment experiments are uncommon but are the only observations representing ecosystem dynamics in an environment with over 550 ppm atmospheric CO_2 (McCarthy et al., 2010). Therefore, DA techniques may be required that assign additional weight to unique, but rare, experiments in the DA approach. As an example, a multi-stage Bayesian approach could be used where the observations from the unique experiment are assimilated first and the posteriors from that assimilation are used as priors for the assimilation of the remaining observations. Second, DA requires using highly simplified ecosystem models because many DA methods use millions of iterations to explore parameter distributions and these iterations have to be applied to both control and manipulated treatments. However, in tension with the need for simple models in DA, more complex models that simulate carbon, water, and nutrient dynamics are also needed to fully leverage the diversity of ecosystem manipulation experiments. Monthly time-scale models of ecosystem processes may be well suited to overcome these challenges for application to predicting changes in biomass over decades in response to global change. First, the contribution of monthly flux and annual biomass measurements to the optimized cost function is more similar in monthly than daily models (12:1 vs. 365:1). Second, they are computationally more efficient than daily models commonly used in DA, allowing data spanning hundreds of plots and multiple decades to be assimilated. Finally, DA is able to calibrate parameters associated with carbon, nitrogen, and water cycles so that they are appropriate for an aggregated monthly time step, helping prevent potential issues associated when applying daily parameterizations to coarser temporal time-steps.

Southeastern U.S. planted pine forests are ideal ecosystems for exploring the application of DA to carbon cycle and forest production predictions. These ecosystems are dominated by loblolly pine (*Pinus taeda* L.), thus allowing for a single parameter set to be applicable to a large region containing many soil types and climatic gradients. Loblolly pine represents more than one half of the standing pine volume in the southern United States (11.7 million ha) and is by far the single most commercially important forest tree species for the region, with more than 1 billion seedlings planted annually (Fox et al., 2007; McKeand et al., 2003). There is also a rich history of experimental research focused on global change factors including region-wide nutrient addition (Albaugh et al., 2016; Carlson et al., 2014; Raymond et al., 2016), water exclusion (Bartkowiak et al., 2015; Tang et al., 2004; Ward et al., 2015; Will et al., 2015), and water addition experiments (Albaugh et al., 2004; Allen et al., 2005; Samuelson et al., 2008). The region also includes a long-term ecosystem CO_2 enrichment study (McCarthy et al., 2010). Furthermore, many of these experiments are multi-factor with water exclusion-by-nutrients (Will et al., 2015), water addition-by-nutrients (Albaugh et al., 2004; Allen et al., 2005; Samuelson et al., 2008), and CO_2-by-nutrients treatments (McCarthy et al., 2010; Oren et al., 2001). Beyond experimental treatments, Southeastern U.S. loblolly pine ecosystems include at least two eddy-covariance sites with high frequency measurements of carbon and water fluxes along with biometric observations over many years (Noormets et al., 2010; Novick et al., 2015), and sites with multi-year sap flow data (Ewers et al., 2001; Gonzalez-Benecke and Martin, 2010; Phillips and Oren, 2001). Finally, there are available studies that include plots that span the regional environmental gradients and extend back to the 1980s (Burkhart et al., 1985). Overall, the high availability of observations of biomass stocks, leaf area index (LAI), carbon fluxes, water fluxes, and vegetation dynamics that span the past 35 years in loblolly pine ecosystems, including plots with experimental manipulation and plots across environmental gradients, is well suited to potentially constrain model parameters and predictions of how carbon cycling responds to environmental change.

Our objective was to develop a DA approach that integrated diverse data from multiple locations, including ecosystem experiments, for predicting how forest productivity may respond to global change. We applied DA techniques to optimize a monthly-time step, simple forest productivity model using southeastern U.S.-wide experimental (nutrient addition, CO_2 enrichment, and water manipulations) and non-experimental data from 35 years of loblolly pine plantation research in the region. Our DA approach, DAPPER (Data Assimilation of Pine Plantation Ecosystem Research), is unique in its focus on simultaneously assimilating observations from multiple locations, experimental types, and data streams into a simple ecosystem model that includes carbon, water, and (implicitly) nutrients using a hierarchal Bayesian technique to develop parameter distributions. We used the DAPPER system to evaluate the sensitivity of biomass predictions and parameter distributions to the inclusion of ecosystem experiments in DA and to predict the regional sensitivity of forest production to nutrient fertilization and drought.

Page 12: [2] Deleted Revisions 5/22/17 1:33:00 PM

The water cycle was a simple bucket model with transpiration predicted using a Penman-Monteith approach (Bryars et al., 2013; Gonzalez-Benecke et al., 2016; Landsberg and Waring, 1997). The canopy conductance used in the Penman-Monteith subroutine was modified by environmental conditions. The modifiers include the same available soil water and vapor pressure deficit modifier as used in the GPP calculation. Maximum canopy conductance occurred when simulated LAI exceeded a parameterized value of leaf area index (LAI). Evaporation was equal to the precipitation intercepted by the canopy. Runoff occurred when the available soil water exceeded a plot-specific maximum available soil water. As in prior applications of 3-PG, available soil water was not allowed take a value below a minimum available soil water, resulting in an implicit irrigation in very dry conditions.

Page 13: [3] Deleted Revisions 5/22/17 1:33:00 PM

2.2 Observations

We used thirteen different data streams from 294 plots at 187 unique locations spread across the region to constrain model parameters (Table 2; Figure 3).

Page 13: [4] Deleted **Revisions** **5/22/17 1:33:00 PM**

All data streams were not available in all plots (Table 2; Table 3). The most common set of data streams were annual or less frequent observations of stand stem biomass (defined as the sum of stemwood, stembark and branches), winter foliage biomass, and living tree counts. The stem and foliage biomass were optimized using regional allometric models based on measurements of tree diameter, height, and plot level-stem size distributions (Gonzalez-Benecke et al., 2014). The most comprehensive set of data streams was from Duke Forest where annual measurements and allometric-based estimates were made of stem biomass (loblolly pine and hardwood), coarse root biomass (loblolly pine and hardwood), fine root biomass (combined loblolly pine and hardwood), stem count (loblolly pine only), leaf turnover (combined loblolly pine and hardwood), and fine root production (combined loblolly pine and hardwood). The Duke Forest dataset (DK3 combined with the Duke FACE CO_2 fertilization study) also included monthly observations of LAI, gross ecosystem production (GEP; modeled gross primary productivity from net ecosystem exchange measured at an eddy-covariance tower), and ET. The set of data streams associated with a particular site and experimental design is shown in Table 3. The measurement uncertainty associated with each data stream is listed in Table 2. Since the model used a monthly time-step, and plots with only biomass and stem density observations were more common than plots with monthly flux estimates, the data used in the optimization cost function were not dominated by high frequency data streams (GEP and ET).

Page 13: [5] Deleted **Revisions** **5/22/17 1:33:00 PM**

We used a hierarchal Bayesian framework to approximate the posterior probability distributions of model parameters in Table 1, the model process uncertainty parameters, and the latent model states and fluxes. The latent model states represented the 'true' stock or flux before measurement uncertainty was included in the observation. Our hierarchal approach was designed to partition uncertainty that is attributable to uncertainty in parameters, model process, and measurements (Hobbs and Hooten, 2015). Previous forest ecosystem DA efforts have either focused on parameter uncertainty, by using measurement uncertainty as the variance term in a Gaussian cost function, or on total uncertainty by directly estimating the Gaussian variance term. The latter combines measurement uncertainty and process uncertainty into the same parameter and is unable to be used for developing prediction intervals, as prediction intervals only include parameter and process errors (Dietze et al., 2013; Hobbs and Hooten, 2015). Here, our focus was on estimating the probability distribution of forest biomass before uncertainty is added through measurement.

First, we estimated the probability of a latent state or flux ($z_{i,m,p}$) for each data point (i) from each data stream (m) in a plot (p) using the 3-PG model with the plot FR . This included the optimized parameters (θ_F), fixed parameters (θ_C), soil characteristic inputs (S), climate inputs (C), site index (SI), fertility (FR_p), and initial conditions (I) required by the 3-PG to simulate each plot, $f(\theta_F, \theta_c, C, S, I, FR_p)$. The latent state ($z_{i,m,p}$) was assumed to be normally distributed with the mean from the 3-PG simulation and an optimized, data stream-specific, process variance $\sigma^2_{m,(process)}$

$$p(\text{process}|\text{process parameters})=$$

$$P\left(z_{i,m,p}\middle|f\left(\theta_F,\theta_C,C,S,I,FR_p\right),\sigma^2_{m\ (process)}\right)$$

$$\sim\text{Normal}\left(z_{i,m,p}\middle|f\left(\theta_F,\theta_C,C,S,I,FR_p\right),\sigma^2_{m\ (process)}\right) \qquad \text{Equation 1}$$

The unobserved true state related to the observed state through a data observation model. In the sampling model, the measured state ($y_{i,m,p}$) was a random sample from a normal distribution with a mean of the true state and a data point-specific standard deviation ($\sigma^2_{i,m,p}$).

p(data|process,data parameters)=

$$\mathbf{P}\left(\mathbf{y}_{\mathbf{i,m,p}}\middle|\mathbf{z}_{\mathbf{i,m,p}},\boldsymbol{\sigma}^{\mathbf{2}}_{\mathbf{i,m,p}}\right)\boldsymbol{\sim}\mathbf{Normal}\left(\mathbf{y}_{\mathbf{i,m,p}}\middle|\mathbf{z}_{\mathbf{i,m,p}},\boldsymbol{\sigma}^{\mathbf{2}}_{\mathbf{i,m,p}}\right)$$

Page 13: [6] Deleted **Revisions** **5/22/17 1:33:00 PM**

This standard deviation ($\sigma^2_{i,m,p}$) represented measurement uncertainty and was similar to the denominator in least-squares approach that is commonly used in DA (Bloom and Williams, 2015; Keenan et al., 2011).

Each parameter (θ_F) that was optimized using the Bayesian method had a prior probability that is specified in Table 1. The prior distribution for the standard deviation $\sigma^2_{m,(process)}$ parameters were uniformly distributed:

$$p(\text{process parameters}|\text{priors})\times p(\text{priors}) = P(\sigma^2_m)\times P(\theta_F) \qquad \text{Equation 3}$$

where

$$P(\sigma^2_m)\sim \text{unif}(0.001,100) \qquad \text{Equation 4}$$

and

$$P(\theta_F) \sim \text{See Supplemental Table 1} \qquad \text{Equation 5}$$

Finally, following the description of the plot specific FR_p described above, the probability for fertilized treatments was based on a comparison to the control treatment FR.

$$P(FR_p|\theta_F,E)=\begin{cases} 1 \text{ if non-fertilized} \\ 1 \text{ if fertilized and } FR_p \geq FR \text{ of control plot} \\ 0 \text{ if fertilized and } FR_p < FR \text{ of control plot} \end{cases} \qquad \text{Equation 6}$$

Our complete Bayesian model for estimating the posterior distributions for the parameters (θ_F), process uncertainty ($\sigma^2_{m,(process)}$), and unobserved true states ($z_{i,m,p}$) was:

$$P\left(\theta_F,\sigma^2_m,z_{i,m,p}\middle|y_{i,m,p},\sigma^2_{i,m,p},\theta_C,S,C,SI,I\right) \propto$$

$$P(z_{i,m,p}|f(\theta_F,\theta_C,FR_p,E),\sigma^2_m)P(y_{i,m,p}|z_{i,m,p},\sigma^2_{i,m,p})P(FR_p|\theta_F,E)P(\theta_F)P(\sigma^2_m) \qquad \text{Equation 7}$$

Page 13: [7] Moved to page 17 (Move #5) Revisions 5/22/17 1:33:00 PM

We numerically estimated the

Page 13: [8] Deleted Revisions 5/22/17 1:33:00 PM

posterior distributions using the Monte-Carlo Markov Chain – Metropolis Hasting (MCMC-MH) algorithm (Zobitz et al., 2011). This approach has been widely used to approximate parameter distributions in ecosystem DA research (Fox et al., 2009; Trudinger et al., 2007; Williams et al., 2005; Zobitz et al., 2011). We adapted the size of the jump for each parameter (i.e., how far a proposed new value can potentially be from the

current value) to ensure the acceptance rate of the parameter set is between 22% and 43% (Ziehn et al., 2012).

Page 13: [9] Moved to page 18 (Move #6) Revisions 5/22/17 1:33:00 PM

All MCMC-MH chains were run for 30 million iterations with the first 15 million iterations discarded as the burn-in.

Page 13: [10] Deleted Revisions 5/22/17 1:33:00 PM

Three chains were run and compared for convergence and we sampled every 1000th parameter in the final 15 million iterations of the MCMC-MH chain. This thinned chain was used in the analysis described below.

Page 13: [11] Moved to page 18 (Move #7) Revisions 5/22/17 1:33:00 PM

The 3-PG model and MCMC-MH algorithm were programed in FORTRAN 90 and used OpenMP to parallelize the simulation of each plot within an iteration of the MCMC-MH algorithm.

Page 13: [12] Deleted Revisions 5/22/17 1:33:00 PM

2.4 Model simulations

Each plot simulated required initial conditions for each model state, climate inputs, soil characteristic inputs, and site index.

Page 13: [13] Moved to page 13 (Move #2) Revisions 5/22/17 1:33:00 PM

We used the first observation at the plot as the initial conditions for the loblolly pine vegetation states (foliage biomass, stem biomass, coarse root biomass, fine root biomass, and stem number). When observations of coarse biomass and fine root biomass were not available, these stocks were initialized as a mean region-wide proportion of the observed stem biomass.

Page 13: [14] Deleted Revisions 5/22/17 1:33:00 PM

contribute to the root cost function and root biomass does not

Page 13: [15] Deleted **Revisions** **5/22/17 1:33:00 PM**

In the two plots with flux observations (US-Dk3 and US-NC2),

Page 13: [16] Deleted **Revisions** **5/22/17 1:33:00 PM**

(Staff, 2016). We assumed that the minimum available soil water was zero.

Page 13: [17] Deleted **Revisions** **5/22/17 1:33:00 PM**

mean annual temperature for each plot (Abatzoglou, 2013)

Page 13: [18] Deleted **Revisions** **5/22/17 1:33:00 PM**

directly estimating FR, rather than calculating from Equation 2, and by requiring the optimized FR in the fertilized plot to be

Page 14: [19] Moved to page 14 (Move #9) **Revisions** **5/22/17 1:33:00 PM**

plot (US-NC2) included a thinning treatment during the period of observation. We simulated the thinning by specifying a decrease in the stem count that matched the proportion removed at the site, with the biomass of each tree equivalent to the average of trees in the plot.

Page 18: [20] Deleted **Revisions** **5/22/17 1:33:00 PM**

Model experiments and analysis

Our analysis focused on comparing parameter distributions and predictions among simulations that used different experimental treatments to estimate the posterior distributions (Table 4). To examine the influence of the Duke FACE CO$_2$ fertilization, we compared a one stage vs. a two-stage data assimilation process. The one stage process assimilated all observations in all plots and experiments simultaneously. In this approach, the elevated CO$_2$ plots only represented 5 of the 294 plots across the region and thus a relatively minor contribution to the likelihood (cost-function) calculation. The two-stage

process used the observations from Duke FACE, US-Dk3 flux site, the other flux site in North Carolina (US-NC2) to estimate parameter posteriors using the priors in Table 1 and SI

2.3

Page 21: [21] Deleted Revisions 5/22/17 1:33:00 PM

These sites were grouped together because they were the most data rich, had the high frequency data streams (monthly GEP, ET, and LAI), and were relatively close in geography. FR was directly estimated for all plots in the first stage, with the FR of a fertilized plot required to be equal to or higher than its control plot. The FR of the CO_2 experiment was equal to the corresponding control plot estimated FR. The FR of the control plot was required to be greater than 0 and, if associated with a nutrient fertilization plot, less than the FR of the fertilized plot.

For the second DA stage, the posterior distributions from the first stage were used as priors for the assimilation of the region-wide observations from the PINEMAP, FPC RW 18, FMRC Thinning, SETRES, and Waycross studies (Table 4). We compared the CO_2 quantum yield enhancement parameter (Calpha700) between the one and two stage approaches to evaluate how the estimation of CO_2 fertilization of plant growth depended on how the Duke FACE data are used in data assimilation. We also estimated the distribution of the percentage increase in net primary productivity (NPP) associated with the elevated CO_2 treatment using the one and two stage data assimilation approaches. The distribution of the percentage increase in NPP was calculated by randomly selecting 1000 parameter sets, with replacement, from the 1-stage converged MCMC chains. This calculation was repeated using the 2-stage approach.

Based on the results from comparing the one and two stage approaches (see results below), we proceeded using the two-stage approach to examine the influence of the water manipulation and nutrient fertilization experiments on posterior distributions and predictions. To evaluate the influence of water manipulation experiments, we repeated the second stage of the data assimilation without the plots where water was added or subtracted. To evaluate the influence of the nutrient manipulation experiments, we first repeated the first stage of data assimilation without the nutrient addition plots in the Duke FACE experiment and used those posteriors as priors to the second stage. This ensured that the priors to the second stage of data assimilation did not include information from nutrient addition experiments. The second stage then excluded the other

Page 21: [22] Deleted Revisions 5/22/17 1:33:00 PM

manipulation experiments in the region.

To examine how the exclusion of the water manipulation experiments influenced parameter inference and predictions, we first examined how the parameter distributions changed from initial priors through the two assimilation stages. With respect to the water manipulation experiments, we focused on the shape of the relationship between available soil water and the quantum yield and stomatal conductance modifier (governed by parameters SW1 and SW2) with and without assimilating the water manipulation experiments. To illustrate the capacity to estimate the probability distribution of predictions using the posterior uncertainty in parameters, we analyzed a focal site in Georgia, near the center of the loblolly pine range (circle in Figure 2). At the focal site, we predicted the sensitivity of stem biomass at age 25 (hereby referred to as STEM_{25}) to a 30%

Page 21: [23] Deleted Revisions 5/22/17 1:33:00 PM

and a 30% decrease in annual precipitation with and without assimilating the water experiments. A 30% percent decrease in precipitation mirrors the magnitude of reduction in the experimental throughfall reduction studies used in DA (Table 3 and Figure 3). Our prediction distributions were calculated

Page 21: [24] Deleted **Revisions** **5/22/17 1:33:00 PM**

as inputs (Figure 2). We simulated forest development from 1989 to 2014 using actual precipitation and again with a 30% reduction in precipitation. We focused our analysis on the percent change in STEM_{25} between the two simulations

Page 21: [25] Deleted **Revisions** **5/22/17 1:33:00 PM**

To examine how the exclusion of the nutrient addition experiment influenced parameter inference and prediction, we focused on the difference in maximum quantum yield parameter (α) and the relationship between site index and soil fertility modifier (FR) with and without assimilating the nutrient experiments. Additionally, we simulated how stem biomass at age 25 (STEM_{25}) responded to a complete removal of nutrient limitation (FR = 1) for the focal site in Georgia. As in the precipitation sensitivity described above, we represented the percentage change in STEM_{25} between simulations with estimated FR and FR =1 as a distribution by integrating across parameter uncertainty. We predicted the regional response to nutrient fertilization by setting the FR at all HUC12 units (see previous paragraph) equal to 1 using the median posterior parameter values from data assimilation where nutrient addition experiments were either included or not. We focused on the regional pattern in the percentage change in stem biomass with the predicted FR (current level fertility) and FR = 1 (nutrient limitation removed).

Finally, we assessed overall model performance of the 2-stage approach for data assimilation with all experimental types included in DA, excluding the nutrient addition experiments, and excluding the nutrient addition experiments using an out-of-sample approach. The approach held 40 random FMRC thinning study plots (Table 3) out from the assimilation, predicted the 40 plots using the median parameter values, and compared the predicted stem biomass to the observed stem biomass. These were plots without any manipulations of nutrients or water, were located throughout the region, and had measurement ages up to 30 years old. For each plot, we only used the most recent observed values to increase the time length between initialization and validation. We repeated the validation for four unique sets of 40 FMRC thinning study plots.

Page 22: [26] Deleted Revisions 5/22/17 1:33:00 PM

Our multi-site, multi-experiment, multi-data stream DA approach was able to constrain most parameters in the 3-PG model (31 of 46 parameters were optimized; Table 6; Supplemental Table 3; Supplemental Figure 1-3). The 31 optimized parameters were the most sensitive parameters in the 3-PG model, defined by the change in total biomass at age 25 for the focal site in Georgia to a 10% change in the parameter (Table 1; Supplemental Table 1). One exception was the light extinction coefficient (k), which showed high sensitivity but was assumed to be fixed because it strongly co-varied with the quantum yield parameter (α). Parameters associated with biomass allocation had priors with large variance but DA was able to provide posteriors with relatively low variance (pFS2, pFS20, pR, and pCRS; Supplemental Figure 1; Supplemental Table 3). The DA process also produced posterior distributions that had less variability than the prior distribution for the important parameters associated with light-use efficiency (α, y, FR1, and FR2; Table 5). DA did not change the parameter distributions, i.e., the posterior and prior distributions were similar, for the parameters that governed the temperature sensitivity of quantum yield, the VPD sensitivity of quantum yield, and the maximum canopy conductance (Supplemental Figure 1-2; Supplemental Table 3). These parameters had strong priors supported by previous research on loblolly pine physiology. Finally, the DA approach was able to estimate the distributions of the process uncertainty parameters (Supplemental Figure 3; Supplement

Table 4).

The addition of the second stage of assimilation that used region-wide observations and posteriors from the DK+NC2 assimilation modified the distributions of the parameters that related to allocation and mortality but did not provide additional constraint on the physiological parameters (Table 5). In particular, the parameters associated with the self-thinning curve and allocation of coarse roots had non-overlapping 95% credible intervals between the DK+NC2 and RW assimilation. The larger estimate for Wsx1000 and lower value for thinPower in the DK+NC2 indicated self-thinning was lower at the sites in the DK+NC2 assimilation than the average of the other sites in the region. The lower value for the pCRS parameter indicated that less NPP was allocated to coarse roots in the DK+NC2 assimilation than the RW assimilation.

The two-stage assimilation was critical for constraining the CO_2 quantum yield enhancement parameter (Calpha700). Both the mean of the posterior distribution and the range of the 95% credible interval were smaller for fCalpha700 when all observations were assimilated simultaneously (1-stage approach) than the distribution estimated using the 2-stage approach (Duke and NC2 assimilated before the region-wide assimilation) (Figure 5a; Table 5). Despite the same data used in both approaches, the differences in fCalpha700 led to a predicted lower enhancement of NPP associated with elevated CO_2 in the experiment. The 1-stage assimilation approach had a median increase in NPP between the control and elevated CO_2 treatments of 15% compared to a 27% in the two-stage approach (Figure 5b).

The RW assimilation constrained the soil fertility parameters that were necessary to enable regional simulations. Our regional model using the 2-stage approach performed well compared to stem biomass data not used in the assimilation. The mean bias in stem biomass of the four out-of-sample validation sets was -6.7 % and the RMSE was 21.2 Mg ha^{-1} (Figure 4).

Excluding the nutrient addition experiments from the DA increased the simulated level of nutrient limitation but did not change the predictive capacity of the independent non-manipulated validation set. DA without nutrient fertilization experiments had a greater and more uncertain value for the maximum quantum yield parameter (α; Figure 6a; Table 5). This parameter was shared across all plots and modified by the environmental conditions at each plot. To compensate for the higher α parameter when nutrient fertilization experiments were excluded from DA, the two soil fertility parameters (FR1 and FR2) combined to predict a 10% lower FR values for a given site index and mean annual temperature (Figure 6b). Subsequently, the prediction for the percentage change in STEM$_{25}$ associated with maximum fertilization (i.e., setting FR = 1) at the focal site in Georgia was 7% higher and had greater uncertainty when nutrient fertilization experiments were excluded from the DA (Figure 6c). The RMSE and mean bias of the non-manipulated validation set was 20.4 Mg ha^{-1} and -4.8 %, respectively (SI Figure 1a)

Excluding the water manipulation experiments from the DA reduced the sensitivity to available soil water but, similar to the inclusion of the nutrient addition experiments, did not change the predictive capacity of the independent non-manipulated validation set. The combined differences in the SW1 and SW2 parameters between the DA with and without the water manipulation experiments decreased the sensitivity of quantum yield and canopy conductance to a reduction in available soil water (Figure 7a). For example, at an available soil water to maximum available soil water ratio of 0.50, the quantum yield and canopy conductance modifier decreased from 0.95 without water experiments to 0.8 with water experiments (Figure 7a). At the focal site in Georgia, the sensitivity of $STEM_{25}$ to a reduction in annual precipitation (Figure 7b) was larger when the water experiments were included in the DA (-8.5% median change in $STEM_{25}$ for a 30% reduction in precipitation) than when the experiments were excluded (-4.1% median change in $STEM_{25}$ for a 30% reduction in precipitation). Similarly, the predictions of $STEM_{25}$ change associated with a 30% increase in precipitation (median: 3.8%) were higher when water experiments were included than when not included (median: 1.1%). The magnitude of uncertainty in the predictions did not differ substantially between forecasts with and without water experiments (Figure 7b). The RMSE and mean bias of the non-manipulated validation set was 19.3 Mg ha^{-1} and -5.8 %, respectively (SI Figure 1b)

Page 25: [27] Deleted **Revisions** **5/22/17 1:33:00 PM**

The sensitivity of forest production to a 30% reduction in precipitation varied across the region. The most sensitive areas, the Piedmont of Georgia and the western edge of the region, predicted up to a 13.1% decline in $STEM_{25}$ (Figure 11a). These were warm areas with relatively low precipitation before the 30% reduction (Figure 2c). The least sensitive area was the interior of the gulf coast (<1% decline; Figure 11a), the area with the highest precipitation in the region (Figure 2c). The regional mean reduction in $STEM_{25}$ associated with a 30% decrease in precipitation was 5.7% (Figure 11a). Excluding the water manipulation experiments from DA reduced the regional mean sensitivity to 1.7% (Figure 11b).

Page 26: [28] Deleted **Revisions** **5/22/17 1:33:00 PM**

We found that including nutrient and water manipulation experiments aided in distinguishing the mechanisms driving patterns in biomass across the region. Including these experiments in the data-assimilation did not improve the predictive capacity of the independent validation set of non-manipulation plots. However, including nutrient and water manipulation did change the underlying mechanisms explaining the patterns in stem biomass. Without the nutrient and water manipulation experiments, the same biomass predictions were attributable to a higher level of nutrient limitation and a lower level of water limitation. This resulted in differing sensitivities to changes in nutrient or water availability. Overall, this finding highlights a key challenge when parameterizing ecosystem models that will be used for global change predictions, that different combinations of environmental drivers can produce similar predictions of current observations. Ecosystem manipulation experiments are an important tool for addressing this challenge.

Parameter and process identifiability, or equifinality, presents a challenge when parameterizing ecosystem models using DA (Luo et al., 2009). One important source of equifinality is the tradeoff between parameters governing the potential productivity of the vegetation and the downregulation of productivity due to nutrient limitation. When using observational data at a single site, a single parameter is often optimized to set a photosynthetic rate per absorbed light, i.e., a quantum yield. This single parameter combines the potential photosynthesis set by climate and the influence of nutrient limitation on photosynthesis into a single parameter. However, separating these two processes into two or more parameters is challenging because a high potential quantum yield parameter (α) and high nutrient limitation (FR) can mathematically yield the same photosynthetic rate as low potential quantum yield and low nutrient limitation. The former implies a larger potential response to nutrient addition than the latter. We found that including nutrient addition experiments in DA helped overcome this challenge. In the case of the

3-PG model used in this study, the maximum quantum parameter (α) and soil fertility parameters (FR1 and FR2) were more constrained and inferred lower levels of nutrient limitation across the region when nutrient fertilization experiments were included in the DA. This finding likely extends to other models that include the concept of potential productivity and productivity downregulated by nutrient limitation. For example, the applications of the Data Assimilation Link Ecosystem Carbon (DALEC) model (Williams et al., 2005) to DA often assumed nine of the ten parameters associated with photosynthesis were fixed, thus using a single parameter to represent both the quantum yield (defined as nitrogen use efficiency in DALEC) and the magnitude of nitrogen limitation of a site (Fox et al., 2009). The use of a single parameter, rather than using nutrient addition experiments to separate into multiple parameters, is appropriate when assuming nutrient availability is static. Applications of DA to predictions of ecosystems with changing nutrient availability, either through management, elevated CO_2, or nitrogen addition, would benefit from using nutrient addition studies to quantify the magnitude of nutrient limitation. Studies of known nutrient gradients could be used in lieu of nutrient addition studies, but effort must be made to account for confounding abiotic factors, such as available soil water or climatic conditions, that may co-vary with nutrient availability.

Another challenge in DA is deciding how to weigh different types of data used in model fitting (Gao et al., 2011; Wutzler and Carvalhais, 2014). Here we demonstrate that DA efforts should also consider how to weigh different types of ecosystem experiments. In our analysis, we included three types of experiments: nutrient addition, water manipulation, and CO_2 fertilization. The nutrient addition and water manipulation experiments were represented by multiple sites across the region while the CO_2 fertilization only occurred at a single location (Figure 3). We found that the parameter that represents the increase in maximum quantum yield under elevated CO_2 was substantially lower when all observations, sites, and experiments were assimilated simultaneously than when the CO_2 fertilization experiment was given greater weight. The greater weight was applied by first assimilating the CO_2 fertilization experiment and using the posteriors as priors for assimilating the remaining observations. Providing additional weight on the single site with unique environmental conditions (i.e., atmospheric CO_2 at 570 ppm) using a two-stage data-assimilation, we were able to more accurately represent the observed differences in NPP between the ambient and elevated CO_2 treatments at the Duke site (McCarthy et al., 2010). Given than only a few of the parameters were significantly different between the Duke site and the other studies across the region, it may be possible to optimize one parameter for the Duke site and another parameter for the other studies in a 1-stage approach that combines all the plots into a single assimilation. However, the 2-stage approach was required to identify which parameters were different between the Duke site and the other studies. Overall, we suggest that DA efforts using multiple studies and multiple experiment types identify whether particular experiments at limited number of sites have the potential to uniquely constrain specific parameters. In this case, additional weight may be needed to avoid having the signal of the unique experiment overwhelmed by the large amount of data from the other sites and experiments.

Our analysis highlights that nutrient limitation of productivity was widespread across the region. The largest potential gains in productivity from nutrient addition were predicted in central Georgia, an area with warm annual temperatures but poor soils, as expressed in the low site index. The baseline fertility used in our regional analysis was derived from an empirical model of site index that was developed using field plots with minimal management (Sabatia and

Burkhart, 2014). Subsequently our estimate of baseline fertility is likely on the low end of forest stands currently in production. Further, we recognize that the site index model had uncertainty that could be formally incorporated into the hierarchal Bayesian approach in future applications.

The soil fertility modifier has commonly been used to calibrate the 3-PG for applications to a single site, with recent work focused on developing an approach to predicting the soil fertility modifier from environmental conditions (Gonzalez-Benecke et al., 2016; Subedi et al., 2015). We have extended prior efforts to develop a simple predictive model of FR in two ways. First, we simultaneously calibrated the parameters in the empirical FR model alongside the other parameters in the 3-PG model. Prior studies have assumed fixed values for the 3-PG model parameters, fitted FR for plots with observations, and developed a relationship between FR and site index. Our Bayesian approach to simultaneously calibrating the 3-PG parameters and the FR model allowed for the estimation of uncertainty and covariation among parameters in the 3-PG and FR models. Second, we included a climate term (mean annual temperature) in the relationship between site index and FR. This resulted in a lower FR for a given site index in warmer locations. By including the climate term, FR can be interpreted as relative to the climate at a given location and the potential productivity of a plot can be optimized by setting FR equal to 1. When a climate term is not used in the empirical FR model, FR is relative to the greatest site index in the region, which does not occur in the northern extent of the region even in fertilized plots due to climatic constraints.

Our simulations show that loblolly pine productivity was not strongly sensitive to changes in precipitation at present day temperatures and atmospheric CO_2. We simulated a 30% reduction in annual precipitation and found a maximum of a 13.1% reduction in productivity. A 30% reduction in precipitation is plausible but is more extreme than most Multivariate Adaptive Constructed Analogs (MACA) downscaled climate model projections for the Representative Concentration Pathway (RCP) 8.5 scenario from the CMIP5 Project (comparing the 1971-2000 period to the 2070-2099) (Abatzoglou and Brown, 2012; Taylor et al., 2012). Central Georgia was the most responsive to precipitation reduction, paralleling the spatial patterns in the response to nutrient addition, suggesting that the region is able to support high productivity but is sensitive to nutrient and precipitation levels. The simulated sensitivity was likely due to poor soils (low site index) and low baseline precipitation relative to the warm climate. Our predictions of low sensitivity to precipitation reduction or addition were derived from assimilating observations from throughfall exclusion and irrigation experiments across the region. Prior publications from the studies used in DA also reported low sensitivities to water manipulations, indicating that our predictions are likely not biased (Albaugh et al., 2004; Samuelson et al., 2014; Ward et al., 2015; Wightman et al., 2016). For example, the throughfall exclusion experiment at the focal site in Georgia, reported a 13% reduction in stem production during a dry year but a 0% reduced during a wet year, resulting in a 7% reduction of productivity over a 2-year period in response to a 30% reduction in throughfall (Samuelson et al., 2014). Our predicted 8.5% reduction to a 30% reduction in precipitation compares well to the observed change, noting that our sensitivity integrated over a 25-year rotation and included a mix of relatively wet and dry years.

Page 28: [29] Deleted Revisions 5/22/17 1:33:00 PM

Beyond the specifics of the 3-PG modeling efforts, the DA of regional observations into a monthly, computationally tractable ecosystem model can potentially inform Earth system modeling efforts. While the details of physiology differ between 3-PG and global land-surface models, the concepts governing NPP allocation are similar. Therefore, DA using the 3-PG model can be used to parameterize the allocation patterns of similar plant types in a global model. One land-surface model, the Community Land Model (CLM), includes parameters that govern the ratio of stem to leaf allocation, ratio of coarse root to stem allocation, and the ratio of leaf to fine root allocation, parameters that are also optimized in DAPPER. As an example, the ratio of fine root to leaf allocation in CLM 4.0 and 4.5 for temperate pine plant function type is set to 1, resulting in equal annual allocation of carbon to foliage and fine roots (Oleson et al., 2013). In contrast, we found that the median ratio of fine root to foliage allocation was substantially lower at 0.13 (Table 6). Therefore, simulations in the CLM with the lower value of root allocation would have higher allocation to aboveground tissues if the loblolly pine parameters from our analysis were used. This would increase carbon accumulation in woody tissues and could alter predictions of nutrient limitation because stems have higher C:N ratios. Other parameters, including the stem to coarse root ratio, are closer to the values used in the CLM.

5

Page 31: [30] Deleted **Revisions** **5/22/17 1:33:00 PM**

DA is increasingly used for ecological forecasting due to its ability represent prior knowledge, integrate observations into the parameterization, and estimate multiple components of uncertainty, including observation, parameter, and process representation uncertainty (Dietze et al., 2013; Luo et al., 2011b; Niu et al., 2014). Our application of DA to loblolly pine plantations of the southeastern U.S demonstrated that these ecosystems are well suited as a test-bed for the development of DA techniques, particularly techniques for assimilating ecosystem experiments. Further, we found that assimilating ecosystem manipulative experiments into a simple ecosystem model changed predictions quantifying how forest productivity responds to environmental change, highlighting the importance of networks of ecosystem manipulation experiments for helping to parameterize and evaluate ecosystems models (Medlyn et al., 2015).

6 Data availability

Observations used in the DA can be found in the following: Duke FACE study can be found in McCarthy et al. (McCarthy et al., 2010)

Page 36: [31] Deleted Revisions 5/22/17 1:33:00 PM

2011a.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S. and Schimel, D. S.: Ecological forecasting and data assimilation in a data-rich era, Ecological Applications, 21(5), 1429–1442,

Page 39: [32] Deleted Revisions 5/22/17 1:33:00 PM

α	Canopy quantum efficiency (pines)	mol C mol PAR^{-1}	0.84	Uniform	Min = 0.02 Max = 0.1	Vague
y	Ratio NPP/GPP	-	0.84	Uniform	Max= 0.66 Min = 0.30	1
fCalpha700	Proportional increase in canopy quantum efficiency between 350 and 700 ppm CO2	-	0.08	Uniform	Min = 1.05 Max = 2.0	Vague
fCpFS700	Proportional decrease in allocation to foliage between 350 and 700 ppm CO2	-	0.00[#]	Uniform	Min = 0.50 Max = 1.00	Vague
SWconst	Moisture ratio deficit when downregulation is 0.5	-	0.06	Uniform	Min = 0.6 Max = 1.8	2, Vague
SWpower	Power of moisture ratio deficit	-	0.06	Uniform	Min = 1 Max= 13	2, Vague
FR1	Fertility rating parameter 1 (mean annual temperature coefficient)	-	0.23	Uniform	Min = 0.0 Max = 1.0	Vague
FR2	Fertility rating parameter 2 (site index age 25 coefficient)	-	0.39	Uniform	Min = 0.0 Max = 1.0	Vague
wSx1000	Maximum stem mass per tree at 1000 trees/ha	kg tree^{-1}	0.43	Normal	Mean = 235 Sd = 25	3,4
thinPower	Power in self thinning law	-	0.25	Uniform	Min = 1.1	3,4

					Max = 1.80	
pCRS	Ratio of coarse roots to stem allocation	-	0.08	Uniform	Min = 0.15 Max = 0.35	5

1(DeLucia et al., 2007);[2](Landsberg and Waring, 1997), [3](Bryars et al., 2013),[4](Gonzalez-Benecke et al., 2016), 5(Albaugh et al., 2005)
* Sensitivity is 1 when a 10% increase in the parameter results in a 10% change in total biomass. [#]Sensitivity is 0 when a 10% increase in the parameters does not change total biomass by a value greater than 0.01%.

Table 2. Regional observational data streams used in data assimilation.

Data stream	Measurement frequency	Measurement or estimation technique	Uncertainty	Stream ID for Table 4

Page 39: [33] Deleted Revisions 5/22/17 1:33:00 PM

If litter trap method: SD = 2.5% of observation
If LI-2000 method:

Page 39: [34] Deleted Revisions 5/22/17 1:33:00 PM

If litter trap method: SD = 2.5% of observation
If LI-2000 method:

Page 39: [35] Deleted Revisions 5/22/17 1:33:00 PM

If litter trap method: SD = 2.5% of observation
If LI-2000 method:

Page 43: [36] Deleted Revisions 5/22/17 1:33:00 PM

DK+NC2	1st stage of 2-stage assimilation. All plots at the Duke eddy flux (DK3), Duke Free Air CO_2 Enrichment Study, and NC2 eddy flux site; includes CO_2 enrichment and nutrient addition experiments at the Duke site	13
DK+NC2-fertNoExp	1st stage of 2-stageSame as Base assimilation. Same as DK+NC2 but without nutrient fertilizationexcluding all plots with experimental manipulations. Includes control plots that are part of experimental studies.	10208

Page 43: [37] Deleted Revisions 5/22/17 1:33:00 PM

RW-fert	2nd stage of 2-stage assimilation. Same as RW but without nutrient addition experiments; uses the posteriors of the DK+NC2-fert simulation as priors	222
RW-water	2nd stage of 2-stage assimilation. Same as RW but without water manipulation experiments	241

Page 44: [38] Deleted Revisions 5/22/17 1:33:00 PM

Table 5. Posterior means and 95% credible intervals for parameters listed in Table 1 using the data assimilation approaches listed in Table 4.

Page 44: [39] Formatted Table Revisions 5/22/17 1:33:00 PM

Formatted Table

Page 44: [40] Deleted Cells Revisions 5/22/17 1:33:00 PM

Deleted Cells

Page 44: [41] Formatted Table Revisions 5/22/17 1:33:00 PM

Formatted Table

Page 44: [42] Deleted Revisions 5/22/17 1:33:00 PM

037
(0.034 – 0.040)

Page 44: [43] Deleted Revisions 5/22/17 1:33:00 PM

040
(

Page 44: [43] Deleted Revisions 5/22/17 1:33:00 PM

040

(

**Page 44: [44] Deleted	Revisions	5/22/17 1:33:00 PM**

037
(0.035 – 0.040)

**Page 44: [45] Deleted	Revisions	5/22/17 1:33:00 PM**

035
(0.030 – 0.042)

**Page 44: [46] Deleted	Revisions	5/22/17 1:33:00 PM**

0.032
(0.030 – 0.035)

**Page 44: [47] Deleted Cells	Revisions	5/22/17 1:33:00 PM**

Deleted Cells

**Page 44: [48] Formatted Table	Revisions	5/22/17 1:33:00 PM**

Formatted Table

**Page 44: [49] Moved from page 0 (Move #13) Revisions	5/22/17 1:33:00 PM**

0.26

**Page 44: [50] Inserted Cells	Revisions	5/22/17 1:33:00 PM**

Inserted Cells

**Page 44: [51] Deleted	Revisions	5/22/17 1:33:00 PM**

48
(

**Page 44: [51] Deleted	Revisions	5/22/17 1:33:00 PM**

48
(

**Page 44: [52] Deleted	Revisions	5/22/17 1:33:00 PM**

48
(

**Page 44: [52] Deleted	Revisions	5/22/17 1:33:00 PM**

48
(

**Page 44: [53] Deleted	Revisions	5/22/17 1:33:00 PM**

48
(0.46 – 0.51)

**Page 44: [54] Deleted	Revisions	5/22/17 1:33:00 PM**

0.48
(0.45 – 0.51)

**Page 44: [55] Deleted Cells	Revisions	5/22/17 1:33:00 PM**

Deleted Cells

**Page 44: [56] Deleted	Revisions	5/22/17 1:33:00 PM**

0.52
(

**Page 44: [57] Deleted Cells	Revisions	5/22/17 1:33:00 PM**

Deleted Cells

Page 44: [58] Moved from page 44 (Move #15)Revisions 5/22/17 1:33:00 PM

0.50

Page 44: [59] Moved from page 44 (Move #14)Revisions 5/22/17 1:33:00 PM

0.51

Page 44: [60] Moved from page 0 (Move #16) Revisions 5/22/17 1:33:00 PM

1.09

Page 44: [61] Formatted Table Revisions 5/22/17 1:33:00 PM

Formatted Table

Page 44: [62] Deleted Revisions 5/22/17 1:33:00 PM

31
(1.22 – 1.40)

Page 44: [63] Deleted Revisions 5/22/17 1:33:00 PM

31
(

Page 44: [63] Deleted Revisions 5/22/17 1:33:00 PM

31
(

Page 44: [64] Deleted Revisions 5/22/17 1:33:00 PM

31
(

Page 44: [64] Deleted Revisions 5/22/17 1:33:00 PM

31
(

Page 44: [65] Deleted Revisions 5/22/17 1:33:00 PM

1.32
(1.23 – 1.41)

Page 44: [66] Deleted Revisions 5/22/17 1:33:00 PM

1.11
(1.08 – 1.15)

Page 44: [67] Deleted Cells Revisions 5/22/17 1:33:00 PM

Deleted Cells

Page 44: [68] Deleted Revisions 5/22/17 1:33:00 PM

0.84
(0.75 – 0.93)

Page 44: [69] Deleted Revisions 5/22/17 1:33:00 PM

0.83
(0.75 – 0.93)

Page 44: [70] Deleted Revisions 5/22/17 1:33:00 PM

.84
(

Page 44: [70] Deleted Revisions 5/22/17 1:33:00 PM

.84
(

Page 44: [71] Deleted **Revisions** **5/22/17 1:33:00 PM**

84
(

Page 44: [72] Deleted **Revisions** **5/22/17 1:33:00 PM**

0.99
(0.95 – 1.0)

Page 45: [73] Inserted Cells **Revisions** **5/22/17 1:33:00 PM**

Inserted Cells

Page 45: [74] Deleted **Revisions** **5/22/17 1:33:00 PM**

48
(1.09 – 1.85)

Page 45: [75] Deleted **Revisions** **5/22/17 1:33:00 PM**

31
(0.95 – 1.70)

Page 45: [76] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.8
(1.47 – 2.15)

Page 45: [77] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.30
(0.89 – 1.76)

Page 45: [78] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.57
(1.08 – 1.79)

Page 45: [79] Deleted Cells **Revisions** **5/22/17 1:33:00 PM**

Deleted Cells

Page 45: [80] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.61
(0.90 –

Page 45: [80] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.61
(0.90 –

Page 45: [81] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.29
(0.78 – 1.98)

Page 45: [82] Deleted Cells **Revisions** **5/22/17 1:33:00 PM**

Deleted Cells

Page 45: [83] Deleted **Revisions** **5/22/17 1:33:00 PM**

93
(1.48 – 3.82)

Page 45: [84] Deleted **Revisions** **5/22/17 1:33:00 PM**

2.20
(1.47

Page 45: [84] Deleted **Revisions** **5/22/17 1:33:00 PM**

2.20
(1.47

Page 45: [85] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.47
(

Page 45: [86] Inserted Cells **Revisions** **5/22/17 1:33:00 PM**

Inserted Cells

Page 45: [87] Deleted **Revisions** **5/22/17 1:33:00 PM**

094
(

Page 45: [87] Deleted **Revisions** **5/22/17 1:33:00 PM**

094
(

Page 45: [88] Deleted **Revisions** **5/22/17 1:33:00 PM**

096
(0.088 – 0.103)

Page 45: [89] Deleted **Revisions** **5/22/17 1:33:00 PM**

118
(0.110 – 0.128)

Page 45: [90] Deleted Cells **Revisions** **5/22/17 1:33:00 PM**

Deleted Cells

Page 45: [91] Deleted **Revisions** **5/22/17 1:33:00 PM**

0.094
(0.087 – 0.102)

Page 45: [92] Deleted Cells **Revisions** **5/22/17 1:33:00 PM**

Deleted Cells

Page 45: [93] Deleted **Revisions** **5/22/17 1:33:00 PM**

144
(0.133 – 0.154)

Page 45: [94] Deleted **Revisions** **5/22/17 1:33:00 PM**

124
(

Page 45: [94] Deleted **Revisions** **5/22/17 1:33:00 PM**

124
(

Page 45: [95] Deleted **Revisions** **5/22/17 1:33:00 PM**

179
(

Page 45: [95] Deleted **Revisions** **5/22/17 1:33:00 PM**

179
(

Page 45: [95] Deleted **Revisions** **5/22/17 1:33:00 PM**

179
(

Page 45: [96] Deleted **Revisions** **5/22/17 1:33:00 PM**

0.153
(0.140 – 0.168)

Page 45: [97] Formatted Table **Revisions** **5/22/17 1:33:00 PM**

Formatted Table

Page 45: [98] Deleted **Revisions** **5/22/17 1:33:00 PM**

(171 – 181)

Page 45: [99] Deleted **Revisions** **5/22/17 1:33:00 PM**

180
(174 – 186)

Page 45: [100] Deleted **Revisions** **5/22/17 1:33:00 PM**

180
(176 – 186)

Page 45: [101] Deleted **Revisions** **5/22/17 1:33:00 PM**

258
(228 – 295)

Page 45: [102] Deleted **Revisions** **5/22/17 1:33:00 PM**

181
(174 0 187)

Page 45: [103] Deleted Cells **Revisions** **5/22/17 1:33:00 PM**

Deleted Cells

Page 45: [104] Moved from page 44 (Move #17)Revisions **5/22/17 1:33:00 PM**

0.76

Page 45: [105] Formatted Table **Revisions** **5/22/17 1:33:00 PM**

Formatted Table

Page 45: [106] Deleted **Revisions** **5/22/17 1:33:00 PM**

67
(1.60 – 1.74)

Page 45: [107] Deleted **Revisions** **5/22/17 1:33:00 PM**

70
(1.63 –

Page 45: [107] Deleted **Revisions** **5/22/17 1:33:00 PM**

70
(1.63 –

Page 45: [108] Deleted **Revisions** **5/22/17 1:33:00 PM**

71
(1.65 – 1.78)

Page 45: [109] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.28
(1

Page 45: [109] Deleted **Revisions** **5/22/17 1:33:00 PM**

1.28

(1

Page 45: [110] Deleted Revisions 5/22/17 1:33:00 PM

1.61
(1.51 – 1.69)

Page 45: [111] Deleted Cells Revisions 5/22/17 1:33:00 PM

Deleted Cells

Page 45: [112] Moved from page 45 (Move #18)Revisions 5/22/17 1:33:00 PM

0.19

Page 45: [113] Formatted Table Revisions 5/22/17 1:33:00 PM

Formatted Table

Page 45: [114] Deleted Revisions 5/22/17 1:33:00 PM

(0.25 – 0.27)

Page 45: [115] Inserted Cells Revisions 5/22/17 1:33:00 PM

Inserted Cells

Page 45: [116] Deleted Revisions 5/22/17 1:33:00 PM

0.24
(0.23 – 0.

Page 45: [116] Deleted Revisions 5/22/17 1:33:00 PM

0.24
(0.23 – 0.

Page 45: [117] Deleted Revisions 5/22/17 1:33:00 PM

(0.24 – 0.26)

Page 45: [118] Deleted Revisions 5/22/17 1:33:00 PM

0.17
(0.16 –

Page 45: [119] Deleted Cells Revisions 5/22/17 1:33:00 PM

Deleted Cells

Page 45: [120] Deleted Revisions 5/22/17 1:33:00 PM

0.28
(0.27 – 0.29)

Page 45: [121] Deleted Cells Revisions 5/22/17 1:33:00 PM

Deleted Cells

Page 45: [122] Deleted Revisions 5/22/17 1:33:00 PM

Figures

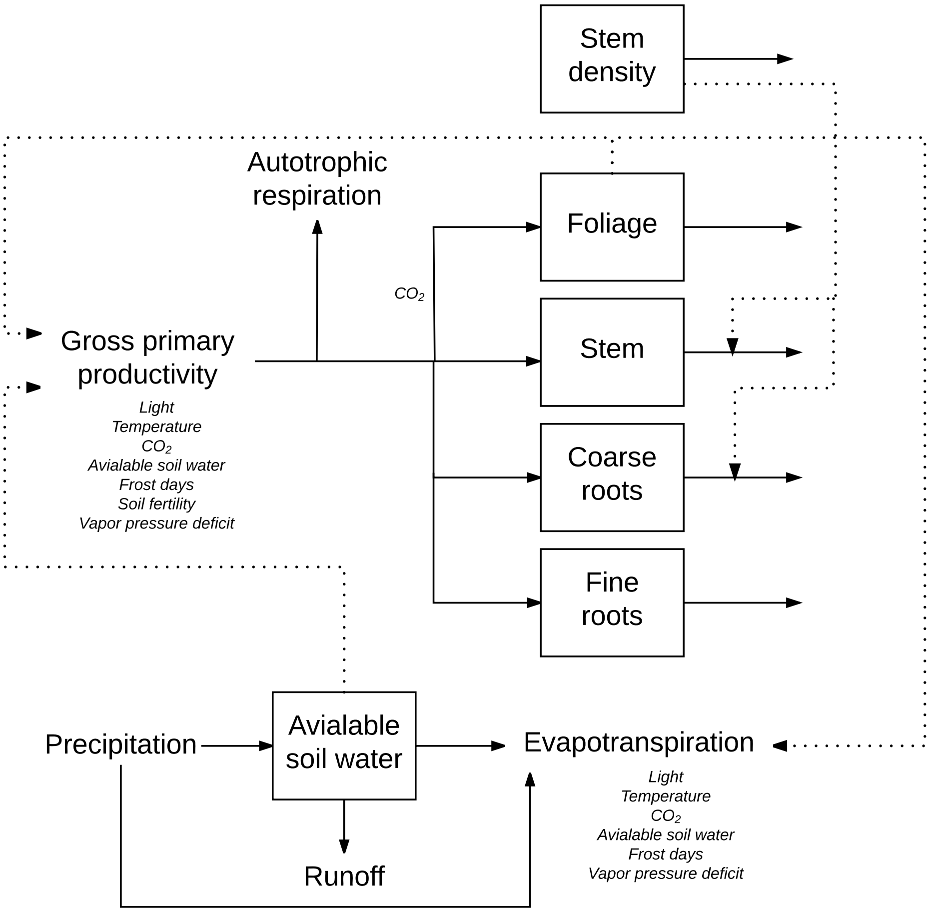
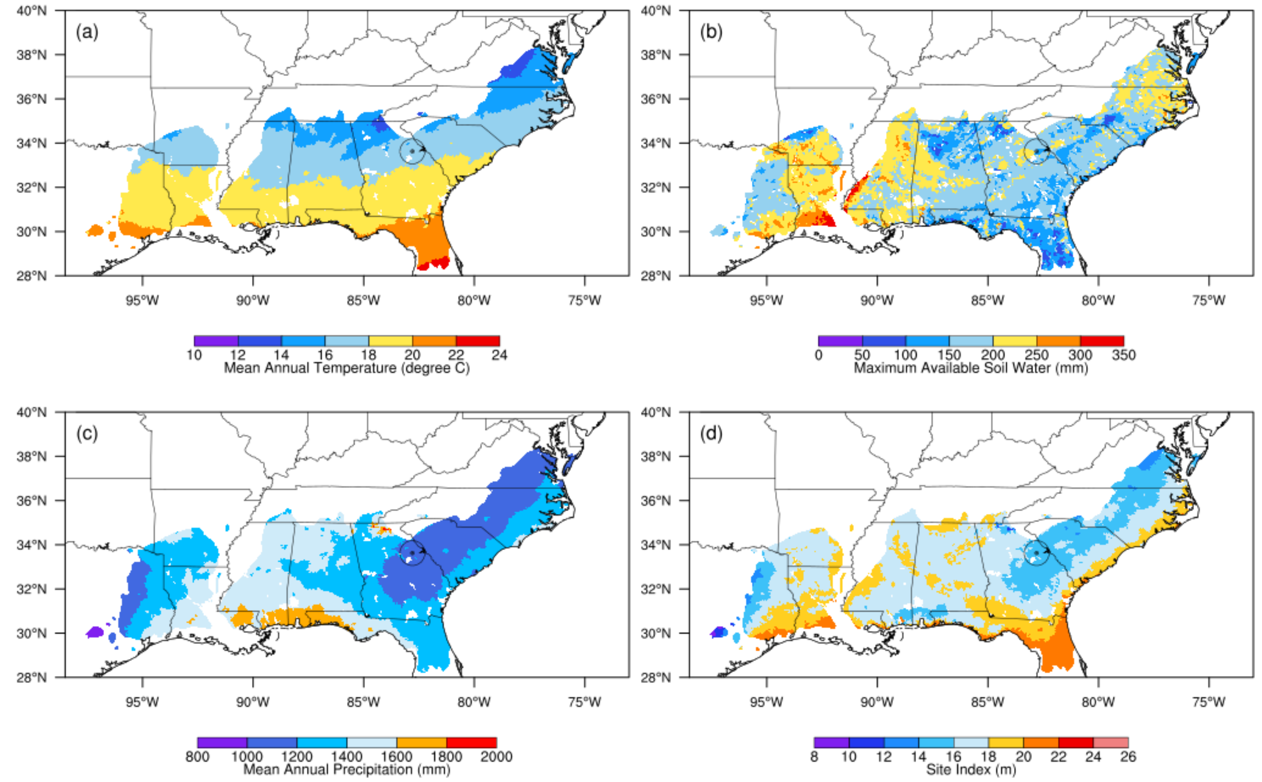


Figure 1.

Page 50: [123] Deleted **Revisions** **5/22/17 1:33:00 PM**



Page 50: [124] Deleted **Revisions** **5/22/17 1:33:00 PM**

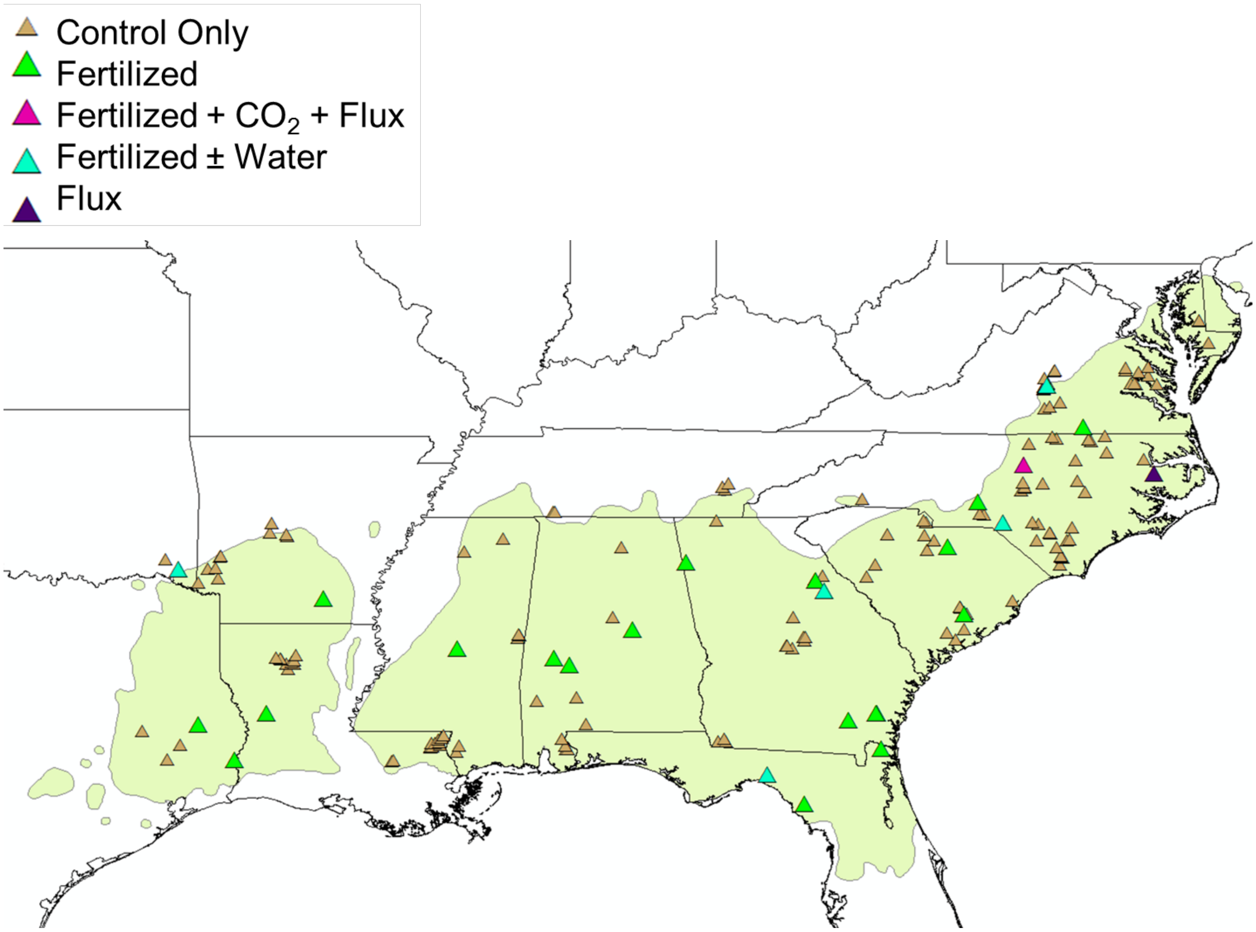


Figure 3.

Page 55: [125] Deleted **Revisions** **5/22/17 1:33:00 PM**

change in stem biomass of a 25-year stand when nutrient limitation is completely removed through nutrient addition (simulated by setting FR = 1). Predictions from data assimilation that included nutrient addition experiments are shown in (a) and prediction data assimilation that did not include nutrient addition experiments are shown in (b). The focal site in Georgia highlighted in Figures 5c and 6b is represented by the circle containing the dot.

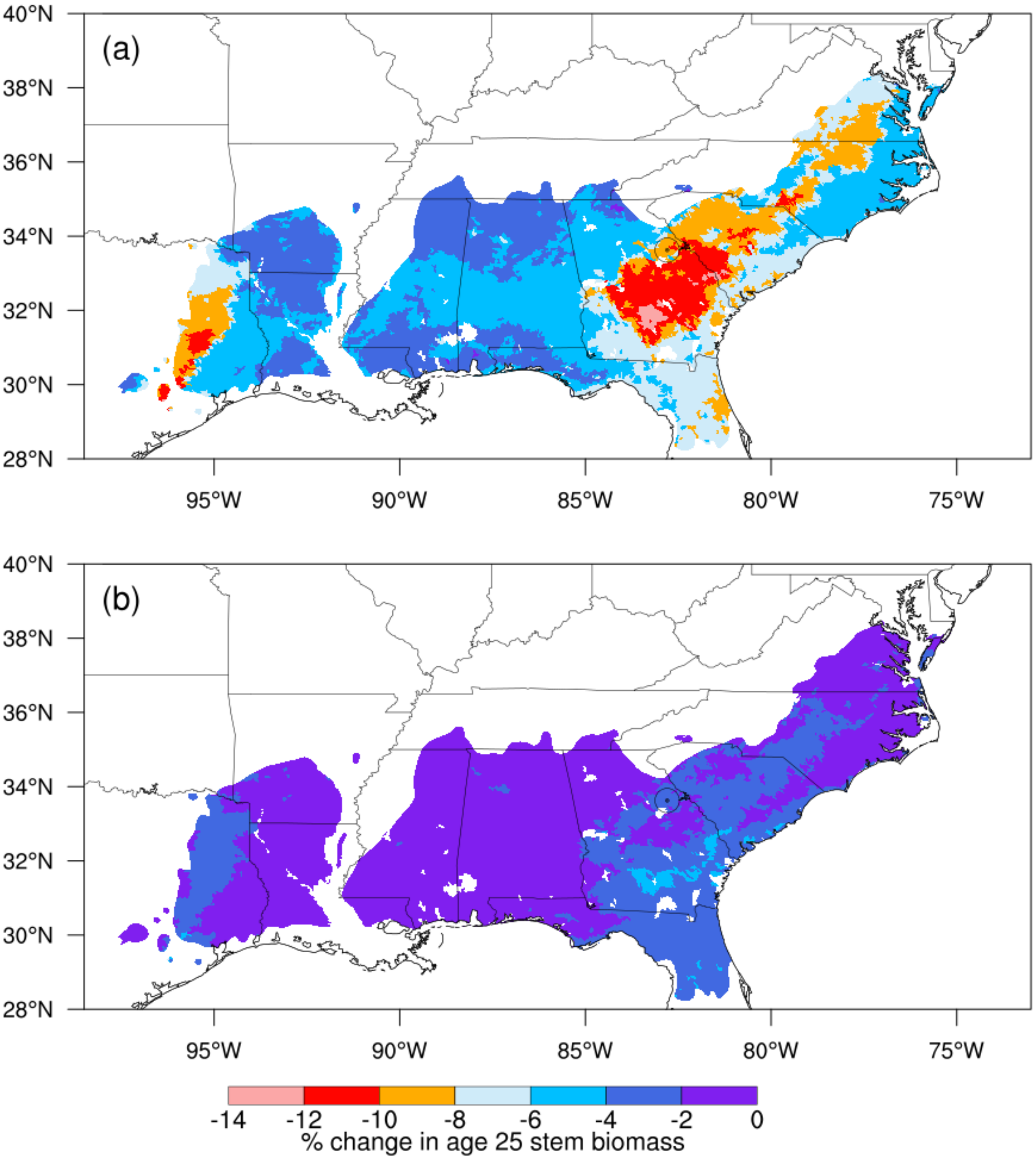


Figure 11. Regional predictions of the change in stem biomass of a 25-year stand when annual precipitation is reduced by 30%. Predictions from data assimilation that included water manipulation experiments are shown in (a) and prediction data assimilation that did not include water manipulation experiments are shown in (b). The focal site in Georgia highlighted in Figures 6c and 7b is represented by the circle containing the dot.