We are grateful to the referees for the positive and helpful reviews that have improved the manuscript. Detailed point-by-point responses to the reviewers' comments are provided below. The reviewers' comments are shown in italics with our responses embedded using regular font text. The page and line numbers below refer to the revised manuscript submission version.

Response to Referee #1: Benjamin Felzer, bsf208@lehigh.edu

1. I find the authors have responded to all my concerns, with one minor exception. The point I was trying to make in my comment below was not just the effect of including N-limitation, but also that including N deposition would more than likely offset the negative effects of ozone on GPP. All this requires is one additional sentence or so on their part to acknowledge this. I can provide one of my own references, though I am happy if they choose any reference (does not have be mine) that makes the point.

Felzer, B. S., Cronin, T., Reilly, J. M., Melillo, J. M. and Wang, X. 2007. Impacts of ozone on trees and crops. Comptes rendus Geosience. 339/11-12: 784-798.

We emphasize the possible offsetting effects of N deposition on Page 19, Line 575: "However, other studies have suggested that the O_3 damage to GPP may be offset by the benefits of co-located nitrogen deposition (Ollinger et al., 2002; Felzer et al., 2007), or even limited by carbon-nitrogen interactions (Kvalevag and Myhre, 2013)."

Response to Referee #2

The authors have made considerable improvements to the paper and have addressed many of the reviewers' concerns. The text is much clearer, with methodology and results more understandable and accurate. The analyses and results are discussed within the context of observations and other studies and seem reasonable. Here are a few additional comments, and these and other minor comments are annotated within the attached text:

1) What is a "distributed" simulation? Do you mean regional simulation? This terminology comes up throughout the text, and is not clear what is meant. If the simulations are regional, please change "distributed" to "regional", or else include an explanation of what makes a simulation "distributed".

We have deleted the word "distributed" and used "regional" throughout the manuscript.

2) You have included a great number of sensitivity analyses, but none test the sensitivity of the particular ozone parameterization used for each PFT. The Sitch (2007) approach is laden with potential problems, including that the parameterization is based on only a few species of plants, and was developed based on decreases in biomass rather than photosynthesis (which are not equivalent). Is it possible that results would show greater differences if different PFT parameters were used? This is key to your results and currently completely overlooked, so at least some discussion of this weakness should be included.

We do agree with the reviewer. Despite the known shortcomings in the Sitch et al. algorithm, our

analyses indicate that the Sitch model performs reasonably well for different parameters against an extensive meta-analysis dataset of field and laboratory measurements (Figure 6). In addition, we purposefully include the low and high O_3 sensitivity results in our study to account for uncertainties in the parameterization.

In the revised manuscript, we have emphasized uncertainties in the Sitch et al. (2007) algorithm (Page 21, Line 611-613): "In addition, the O_3 damage algorithm parameters were calibrated using limited measurements for a few plant species, and were based on biomass yield not photosynthetic rate (Sitch et al., 2007). Future work will exploit recent extensive meta-data analyses (Lombardozzi et al., 2013; Wittig et al., 2007) to refine the ozone damage parameterization in YIBs including the decoupled modification of photosynthesis and stomatal conductance."

3) The figures are much cleaner, but 40 panels is still too many for the main body of the text. If you don't want to average by pft, can you choose one representative panel for each pft (e.g. 5 panels total) and then include the others as supplemental? Also, for the bar charts, can you plot histograms of the X2 instead, and have parts of each bar/bin color-coded for the PFT type?

We have moved original Figs. 1-3 to supplementary materials and added new Figs. 1-2 as suggested. The new Fig. 1 compares monthly average simulated and observed GPP by PFT type. Fig. 2 shows the histograms of χ^2 and $\Delta\chi^2$ with colored stacks indicating PFTs.

Detailed comments:

1. Lines 49-50: Because of fossil fuel emissions?

We clarified it as follows: "Currently, North America is acting as a net source of CO_2 to the atmosphere (King et al., 2012), mainly due to fossil-fuel combustion in the U.S. and the deforestation in Mexico." (Lines 49-51).

2. Lines 54-56: Will the carbon sink decrease as forests mature?

This is a fascinating and controversial question beyond the scope of our current study. For example, it is generally assumed that young trees may be more effective carbon sinks than old trees. However, recent studies indicate that old trees are still important carbon sinks (Luyssaert et al. Nature, 2008).

3. Line 57: More sensitive to climate forcing: It's not clear exactly what this means or how it relates to the CO2 sink. Please make the connection clear.

The original sentence referred to surface climate response to radiative forcing over the North American land surface. We agree with the reviewer and have deleted this sentence in the revised manuscript.

4. Section 1, paragraph 2: This paragraph is only related to ozone through the first sentence. Can you add an estimate from previous work about how much ozone might be decreasing this

sink? I now see it in a paragraph below. Still, the opening sentence in this paragraph needs to connect ozone to the change in C sink in some way -- a mechanism, or referencing amount of decrease found by previous study.

We added a sentence to the end of this paragraph to make the connection to ozone damage (Page 2, Line 57): "However, O_3 damage may in part dampen the level of carbon sequestration by North American ecosystems (Felzer et al., 2004; 2005)."

5. Line 100: I'm not sure what regionally distributed configuration means. Please you clarify here since you use "distributed" throughout the rest of the paper. Do you mean regional simulations? If so, you can easily remove "distributed" and just call the simulations "regional" and then explain how you chose vegetation type & climate data for the regional simulations.

Confirming that we do mean "regional" simulations. We have removed the word "distributed" and used "regional" throughout the manuscript. We explain how we chose vegetation & climate data for regional simulations in section 2.2.2.

6. Line 275: Was there a reason that you chose this forcing for the ozone simulations?

We have added an explanatory sentence (Page 10, Line 274): "Later analyses show that METsite_LAImerra has the lowest biases relative to other O_3-free simulations."

7. Line 285: Standard deviations?

We have clarified (Page 10, Line 286): "ε_i are the observational uncertainties resulting from turbulence, gap-filling, flux partitioning, and u* threshold determination (Barr et al., 2013)." (Lines 286-288)."

8. Section 2.2.2: You do a great job of describing the vegetation cover & meteorological variables, and conduct many sensitivity tests on different aspects of the simulations. However, you do not consider how ozone damage might vary by PFT, and only include a high and a low sensitivity. The Sitch parameters are not robust (for example, the damage calculations are based on decreasing biomass, not photosynthesis, which is the parameter adjusted in the model), so it would be prudent to include either additional sensitivities or at least some sort of discussion about this. It is striking that you are testing the sensitivity of the model to many more parameters than ozone, even though ozone is of primary interest to this paper.

Please see response to the main comment #2 above.

9. Title of section 3: This should be called "Results and Discussion", and the next section should be "Conclusions".

Changed as suggested.

10. Lines 348-349: It's not clear whether you mean for all NACP sites or for just the deciduous broadleaf and shrubland NACP sites. Please clarify what vegetation type you mean and where

the sites are located.

We have clarified (Page 12, Line 348): "The largest model overestimate (factor of 3-8) occurs at two ENF sites, CA-SJ1 and CA-SJ2 in North Central Canada."

11. Line 349: Are these the only two grassland sites? If not, do others capture the seasonality better?

There are 5 grassland sites in total and simulations are generally weaker at these sites. We have clarified (Page 12, Line 349): "For the grassland sites, the model-observation correlation is low because the seasonality is not well captured, especially for US-ARM (in Great Plains) and US-Var (in California), where the modeled maximum GPP occurs in summer (July), 2-3 months later than in the measurements (April)."

12. Line 359: over all simulated sites?

We have clarified (Page 12, Line 361): "The annual GPP averaged over all 40 sites is …".

13. 22 simulations, correct? Are these for ozone-free GPP, as in the previous paragraph? Modis or MERRA? Please be more specific.

We have clarified (Page 13, Line 364): "Among the 40 NACP sites, 22 have reasonable performance with $\chi^2 < 4$ for the simulation METsite_LAImerra".

14. Line 379: species, not speciation. Speciation refers to an evolutionary process.

We have changed the word "speciation" to "species".

15. Lines 411-413: Please remind us whether these are monthly (or seasonal) averages, or hourly data. The figure 5 legend says that the values plotted are for diurnal mean and daily max measurements. However, you only have a few data points in Fig. 5b, which (according the to legend) are the diurnal mean values. Shouldn't this mean that you have points for every day that you're comparing? Please clarify both in the text and in the figure legend.

We have clarified in the Figure caption (now Fig. 4): "Validation of simulated June-July-August (JJA) summertime average surface (a, b) diurnal mean and (c, d) daily maximum 8-hour average O_3…" and added to the end of the caption: "Each point in (b) and (d) represents the mean value for JJA at one specific site."

16. Line 504: How are you defining eastern and western US?

We have added (Page 17, Line 504): "…in the eastern U.S. (east of 95°W) and 3.9 g C m^{-2} day^{-1} in the western U.S. (west of 95°W) …"

17. Lines 565-566: Is there any way to quantify the uncertainty associated with these values?

Measurement years are not consistent across all of the sites, which complicates the uncertainty quantification. Instead, we have quantified the uncertainty in the regional model carbon uptake estimate based on interannual variability (Page 20, Line 582): "The total carbon uptake was estimated to be 4.43 ± 0.18 Pg C for the summer, accounting for 57-60% of the annual value of 7.59 ± 0.25 Pg C".

18. Lines 581-583: This stuck out in the above discussion as well. It seems a little confusing to separate out summer and annual GPP, especially when you only have annual estimates to compare to.

We emphasize summer results because the O_3 vegetation damage effect is the strongest during this season.

19. Figs. 1-2: 40 panels is just too much for a figure within the main body of the paper. This can go in the supplementary material, but should not go in the main text. You have 5 different PFTs here, so maybe choose 1 representative figure for each PFT and then include the rest in supplemental. Same for Figure 2.

We have moved the original Figs. 1-2 to supplementary material. We added a new Fig. 1 to compare monthly average GPP between simulation and observations by PFT type.

20. Fig.3: Similar to the 40-panel plots above, these figures have too many bars and it is difficult to interpret. Consider instead making a histogram of X2 values, with stacked bars that are color-coded by pft type.

We have moved the original Fig. 3 to supplementary material. We plotted stacked histograms of χ^2 in Fig. 2 as suggested.

21. Fig. 6: This figure seems unnecessary and is confusing. Much of the information conveyed here is also present in the maps that overlay the NACP sites (Fig. 8) and is summarized in Fig. 7. Both Figs. 7 and 8 are much more clear, so this figure can be removed.

The original Fig. 6 (now Fig. 5) is one of the most important results for this study. The figure presents the quantified ozone vegetation damage and its connections with observed GPP and $[O_3]$ at each individual site. Fig. 7 (now Fig. 6) present results with fixed level of $[O_3]$, which is not from observations. Fig. 8 (now Fig. 7) has no PFT-information for those sites. As a result, we retained this figure in the revised manuscript.