We thank the reviewer for the thoughtful and detailed comments and are pleased to note that the reviewer finds the study significant. Our responses to the comments are given below. We believe that the revisions we made in response to the reviewers comments have improved the manuscript significantly.

[MaC1]: *The discussion is too general, and a more in-depth discussion would make this paper more interesting. For example, why do most remote sensing proxies and models fail to represent spatial and inter-annual variations of GPP for DBF areas? Could this be due to uncertainty in the remote sensing data or model representation of spring phenology during the early growing season? It would be more interesting if the authors could explain why the different RS proxies/models perform differently for biome types (e.g. MOD17 and VPRM model for GRS in Fig. 9). This would provide the reader with more specific information on what processes might be missing in a certain model. Also, most models/proxies do not do well in representing inter-annual variability in forests. It would be useful if the authors could provide a more detailed explanation of the reasons for this behavior.*

We agree that more specific discussion would be helpful. The challenge in doing this is that it's difficult to diagnose specific mechanisms and reasons for poor model and proxy performance with the data we have available. While it is certainly reasonable to discuss possible explanations, we prefer not to be speculative, and therefore provided a somewhat general discussion in our original manuscript. However, we also recognize that the reviewer makes a valid point. To address the reviewer's concern, we have added the following new text to the discussion that attempts to clarify these issues:

"Relatively weak performance of models in DBF at annual scale has also been noted in the context of dynamic ecosystem models (e.g., Schwalm et al., 2007). Most of the DBF sites included in this study are located in temperate regions where phenology co-varies with temperature and PAR, and exerts significant control on annual GPP across sites (Richardson et al., 2012). Across sites, one day of error in spring and fall phenology can lead to errors of 12 gCm-2 in January to June GPP and 6 gCm^{-2} in July to December GPP estimates respectively (Richardson et al., 2010). Phenology is also an important factor that controls interannual variations in GPP and one day of phonological anomaly in spring and fall can lead to differences of 7gCm-2 in January to June and 8 gCm-2 in July to December GPP respectively (Richardson et al., 2010). In LUE-based models such as those included in this analysis, interaction between PAR and phenology is primarily captured through APAR. Thus, the ability of LUE-based models to capture variations in GPP is closely tied with the accuracy of remotely sensed estimates (or surrogate) for FPAR. For a variety of reasons it appears that 8-day MODIS data (FPAR in MOD17, and EVI in VPRM and TG) do not consistently capture rapid phenological changes occurring over relatively short time scales in spring and fall, thereby introducing error to remote sensing-based estimates of annual GPP. Recent studies have also suggested that lagged effects can significantly affect annual GPP, especially in ecosystems that are not strongly dependent on a single climatic factor (Gough et al. 2008, Marcolla et al., 2011). For example, Zielis et al. (2013) used long term eddy covariance data collected at a spruce forest site to show that inclusion of meteorological data from the previous year significantly improved estimates of net ecosystem exchange, suggesting that next generation models need to include lagged effects and functional responses to climate forcing in previous year. “

[MaC2(a)]: *There may be some inconsistency in this paragraph. The previous discussion on the “Proxy+Met” model indicated that the spatial variation in annual terrestrial GPP over large areas might reflect an equilibrium response to climate. But then the authors state that “the influence of environmental variables on GPP becomes progressively weaker as the temporal scale increases.*

We thank the reviewer for pointing this out. We agree that in its current form there appears to be a discrepancy in the paragraph. In the context of the ‘Proxy+Met’ model our claim is that annual productivity reflects equilibrium response to *mean annual climate*. Here, we are suggesting that *fluctuations* in temperature or vapor pressure deficit (VPD) at daily and 8-day time scale do not significantly influence estimates of annual GPP. The two statements are thus consistent with each other. We will replace the current sentence with the following sentence:

“Other studies have shown that fluctuations in environmental variables have weak effect on GPP at daily to 8-day time scale. “

[MaC2(b)]: *I would think a lack of understanding of how leaf level processes scale to daily and longer time scales might largely explain why the LUE model fails to account for the spatial variability of annual GPP for certain biome types. This might also apply to the results on modeled inter-annual anomalies.*

We agree. We highlight this in discussing our analysis of geographic variation in annual GPP (L 2-3, page 11642). We also agree that this might apply to the results in temporal domain. However, as we discussed above and in the manuscript, processes operating at the level of individual trees (e.g., functional responses and lagged effect of weather) appear to be also important for explaining inter-annual variability.

[MaC2(c)]: *The declaration that “the LUE-based remote sensing approaches need to incorporate processes occurring at sub-diurnal time scales” does not seem to be supported by the results, as this paper focuses on annual time scales.*

We agree. We will delete the sentence in the revised manuscript.

[MaC3]: *Interannual anomalies in mean growing season greenness (EVI, NDVI) and annual GPP were highly correlated in EBF”: this seems odd to me. Vegetation indices (especially NDVI) tend to saturate in dense vegetation such as EBF, and their seasonal cycle is likely partially obscured by cloud/aerosol effects. Besides, the failure of remote* *sensing models for EBF is mostly likely due to a lack of understanding of the processes controlling seasonality or inter-annual variability of EBF photosynthesis (not just due to increased model complexity). Further explanation is needed here.*

The EBF sites included in our study span a large range of annual GPP from 500 gCm-2 to 2500 gCm-2, and a number of these sites are not densely vegetated sites. Moreover, we used quality flags and removed contaminated MODIS data. We therefore do not believe that saturation or contamination is likely to be an issue at most sites, although it cannot be ruled out entirely in NDVI. We agree with the later part of the comment that a lack of understanding about controls on photosynthesis is one of the main reasons why models working at finer time scales cannot capture inter-annual variations in EBF. To further explain our results, we will add the following paragraph:

"The relationship between remotely sensed VIs, environmental change, and GPP at seasonal time scale is particularly complex in EBF (Huete et al., 2006). Moreover, our current understanding of controls on seasonal scale photosynthesis and its effect on annual productivity in EBF is limited. Despite these challenges, our results suggest that year-to-year variations in GPP may be partly driven by the corresponding variations in LAI. A number EBF sites included in this study are located in the sub-tropical Mediterranean climate where annual productivity is partly controlled by precipitation and large annual precipitation anomalies cause corresponding anomalies in VIs."

[MiC1]: (*Page 11629, Line 17*) *"variability GPP" should be "variability in GPP"*

Done.

[MiC2]: (*Page 11631, Line 25*) *"charactering" should be "characterizing"*

Done.

[MiC3]: (*Page 11634, Line 19*) *The most recent (Collection 5) MOD17 product (Zhao & Running, Science, 2010) uses 11-biome specific parameters.*

MOD17 includes gross- (GPP) and net primary productivity (NPP) estimates. It uses 11 biome-specific parameters to estimate NPP. However, only five of these are used in modeling GPP.

[MiC3]: (*Page 11634, Line 26-27*) *"Following the same procedure that is used by the operational MOD17 algorithm. . .": it would be better to include the reference (Zhao et al. RSE, 2005)*

We have included the suggested reference in the revised manuscript.

[MiC4]: (*Page 11637, Line 10*) *"jackknifed": what does this mean?*

We have removed the word since rest of the sentence conveys the meaning.

[MiC5]: *The number of figures could be reduced. For example, Figs 4, 6, 8 and 10 could be removed and summarized in the text. Most of the critical information can be found in Figs 3, 5, 7, and 9.*

We agree with this. We will remove the four figures and make appropriate changes in the revised manuscript.

[MiC6]: *It would be useful to distinguish different biomes to give readers a better idea of the spatial representation of flux towers for different biomes.*

Done

[MiC7]: *Fig. 2: What does the red line/cross represent?*

Red line marks median and cross indicates outliers. We have added this in figure description.

[MiC8]: *Give the definition of RMSE and MBE in Fig. 5.*

Done.

REFERENCES

Huete, A. R., et al., Amazon rainforests green-up with sunlight in dry season, Geophys. Res. Lett., 33, L06405, doi:10.1029/ 2005GL025583, 2006.

Marcolla, B., et al., Climatic controls and ecosystem responses drive the inter-annual variability of the net ecosystem exchange of an alpine meadow, Agr. Forest Meteorol., vol. 151 (9), pp. 1233-1243, 2011.

Richardson, A. D., et al., Climate change, phenology, and phenological control of vegetation feedbacks to the climate system, Agr. Forest Meteorol., vol. 169, pp. 156-173, 2013.

Schwalm, C., et al., A model-data intercomparison of CO 2exchange across North America: Results from the North American Carbon Program site synthesis, J. Geophys. Res., vol. 115, pp. G00H05, 2010.

Zielis, S., et al., Forest NEP is significantly driven by previous year's weather, Biogeosciences Discuss., vol. 10 (10), pp. 15587-15611, 2013.