Comment on bg-2021-49
Anonymous Referee #2

Turner and others use SIF observations to estimate GPP across the U.S. and note that its interannual variability is driven in large part by extreme events. There are many interesting elements to the manuscript, but many aspects were difficult to follow and/or incompletely described which made it difficult to be confident in the results.

Referencing could be improved in many places (e.g. lines 19 & 22. On p. 2. Consecutive line numbering is so helpful, please don’t re-start the line numbers on every page as a courtesy to reviewers.)

Figure 1: I have questions. Some of the sites shown are meant to study lakes or are in mountainous terrain, but it is hard to tell. Is there no list of sites used? Such a table would be extremely useful in the Supplement, and also to credit the data providers for their efforts in making the data available. At least one, if not more, appears to be a NEON site (anything starting with 'x'), which is fine but NEON should be credited. Ah, I see now that the sites are listed in the Acknowledgements. This is nice but a table would be more useful to the reader. (And US-Men is meant to study a lake. A table with ecosystem type and latitude/longitude would be helpful all around.)
Why were no eddy covariance sites in three of the four areas denoted as important for interannual variability (Texas, South Dakota, and Illinois/Indiana/Ohio) used? I understand that data are only intermittently available, but the Nebraska Mead sites may be a reasonable stand-in and the Indiana sites may be helpful.

How was GPP estimated? Was this consistently using the nighttime (Reichstein et al., 2005 or similar) approach? If so was it consistently with the Reichstein algorithm? This uses a unique u* threshold every few months and respiration model parameters that change as a function of time. If these are interspersed with GPP estimates that use a different approach there will be differences in seasonality of GPP estimates due to methodological differences alone.

Figure 2: I’m not entirely convinced that the ‘dominant cluster’ approach is useful or even necessary given that the peaks of the pdfs vary between about 15 (mixed forest) and 17 (evergreen needleleaf forest) and that estimating these land cover types using MODIS is rather well-established.

In Table 1, is the value behind the +/- sign the standard error of the mean or the standard deviation? (I’m assuming its not the variance). The table legend was not described in very much detail.

Per the previous comment, the structure of the manuscript was a bit bewildering. It was difficult to determine how the methodology took place because methods were interspersed throughout the manuscript starting on Page 1 with the equations. For example, I appreciate that uncertainty is (mostly) noted about carbon flux estimates but how is this uncertainty determined? Is the bootstrapping approach used to determine the uncertainty of the SIF-GPP relationship which is then propagated? Is uncertainty owing from the 16-day moving window used?
Per the previous comment again, extreme events often happen quickly by definition. Is a 16-day moving window sufficient to fully describe how flooding of flash drought for example impacts ecosystems? I appreciate that the proposed precipitation mechanisms are described as ‘hypothesized’ and the notion that precipitation is the culprit makes great sense, but the devil might be in the details, which were not described again in sufficient detail.

30: GPP does not scale linearly with PAR; the GPP/PAR relationship will usually be saturated at 1:30 pm during the growing season. How does this impact equation 4?

Please note the typo in the legend of Figure 3.