Referee # 2

Specific comments:

1. -Title. The leaf-retention and shape strategies are only implied not studied in the manuscript. I suggest changing to a more relevant and accurate title.
   Title changed to: Response of carbon and water fluxes to meteorological and phenological variability in two Eastern North American forests of similar age but contrasting species composition – a multiyear comparison

1. -Line 16-24. The influences of drought and temperature on NEP and ET are entangled together here, which is a bit unclear. Also, some sentences seem to be repetitive. I suggest rewriting this part of the abstract to make it clearer. REVISED
   Summer meteorology greatly impacted the carbon and water fluxes in both stands, however the degree of response varied among the two stands. In general, warm temperatures caused higher ecosystem respiration (RE), resulting in reduced annual NEP values – an impact that was more pronounced at the deciduous broadleaf forest compared to the evergreen needleleaf forest. However, during warm and dry years, the evergreen forest had largely reduced annual NEP values compared to the deciduous forest.

3. - Abstract. Clarify and quantify (if possible) "greatly controlled", "greatly reduced", and "greatly impact". Updated the abstract so most uses of greatly were removed

4. -Line 55-57. Can you add a sentence or two summarizing the previous studies contrasting fluxes coniferous and deciduous forests? Ultimately reduced the focus on the previous studies in the revised introduction. Mentioned a few differences in past sentences.

5. -Methods. I noticed the distances of EC relative to the canopy top are different for the two sites. Would the heights of the EC affect the fluxes due to flux divergence or convergence? Following the assumptions that we are above the canopy roughness layer in each forest, and we're footprint-filtering appropriately, we don't think there is an effect.

6. -Line 157. Is friction velocity a good metric for filtering intermittent turbulence? Previous studies show intermittent turbulence is frequently observed during evening hours at forested sites. No, it's not. It should be paired with stationarity tests, to make it more appropriate. We also calculate the storage change as a means of capturing significant changes in carbon storage in the volume.

7. -Section 2.3. Have the data been filtered for stationarity? Yes. Stationarity test is done.

8. -Section 2.3. The threshold u* seem to be large (0.2 or 0.3 m/s are pretty standard)? Any explanations associated with the sites? I don't think our sites have particularly denser canopies than other sites. May hint at advection processes playing a role?

9. -Section 2.3. Add one or two sentences explaining how you processed/averaged the meteorological data. Meteorological variables were sampled at 5 second intervals and averaged at a half-hourly scale. A two-step cleaning process was used to remove outliers

in half-hourly meteorological data: coarse upper and lower thresholds were applied to half-hourly values to remove obvious outliers, and additional erroneous half-hourly data were removed from time series when instruments were known to be malfunctioning or visual inspection by multiple reviewers resulted in certain agreement that an outlier was present. Missing meteorological data of all lengths were gapfilled using extant data for the same half hours from either (in order of preference) a second sensor at the site, or an equivalent sensor from a nearby (1-3 km away) station in the network (sites described in Peichl et al., 2010).

10. -Section 2.4. Can you describe the uncertainties associated with the approach estimating phenological seasons? Uncertainties would be similar to gap-filling processes. While the estimation of the phenological seasons used 'non-gapfilled' GEP, this still includes the modeled RE and non-gapfilled NEE. A closing comment in Gonsamo et al. (2013) was that studies should also look into detailed uncertainty analysis with representative study sites from global distributions of plant functional types, as it was not previously done.

11. -Line 257 and Line 349. Clarify "responded similarly". REVISED – behaved similarly

12. -Line 255-262. Can you show the standard deviations of the annual mean Ta in Fig.1? Not entirely sure what was being asked, if it is a standard deviation of daily/annual temperature data or a comparison with the climate normals (deviations from mean). Added 30-year mean standard deviation in methods (8.0 ± 1.6°C).

13. -Line 265. Better explanation for the discrepancies is needed here. The discrepancies are over 300 umol m-2 day-1 in spring. Is it in the range of the measurement uncertainty? I'd suggest check the downward PAR to tease out the influences from the canopies and to evaluate the meteorological differences. This section was heavily edited. A paragraph was added in the methods section to highlight the reason for the discrepancies and how they were fixed. Once fixed, this sentence was edited accordingly.

14. -Line 267. Clarify "APAR was similar throughout the year". What are the values (mean and standard deviations) of the FPAR mentioned? At TP39, APAR exhibited a similar parabolic curve each year due to the seasonal amplitude in PARdn and the continuous presence of an apparently dense coniferous canopy promoting a nearly constant fraction (fPAR) of PARdn being absorbed (Fig 2a). Mean fPAR at TP39 was 0.9375 ± 0.05.

15. -Line 281. "Ts(5cm) at TP39 exceeded that of TPD" seems to suggest that the PARgroud at TPD is less, which implies that the APAR at TPD should be higher in summer and autumn. Please explain. However, during the summer and autumn of each year, Ts_{5cm} at TP39 exceeded that of at TPD due to differences in canopy cover. Also, a higher seasonal fPAR at TPD due to the presence of a dense deciduous canopy.

16. -Line 296. Can you explain why 6-year mean day of season growth was used instead of the days of individual years? The 6-year mean was used as it produced a better fit, but also helped explain a more long-term trend of growing season start dates.

17. -Line 327. Could you also add a sentence or two at the beginning of this paragraph to explain the physical meaning of the cumulative (seasonal and annual) fluxes, especially its differences from daily fluxes? Seasonal and total fluxes provide insight on each stands ability to sequester carbon and release water over interannually comparable timescales.

18. -Line 336. "spring was the only season when daily GEP was similar between the forests". As shown in Table 3, the seasonal GEP in spring show larger differences between sites, which I think to some extent contradicts with your statement in Line 336. Please reconcile. Also, when you compare the daily GEP for phenological seasons, how did you address the different lengths of the seasons (i.e. different number of data points)? The second part of this question answers the first part. They were similar in terms of daily rates of GEP not the total seasonal sum, which was impacted by the total length.

19. -There are a few places where I have similar comments as the previous one. I suggest adding some explanations for the statistical techniques (ANOVA and MANOVA) you used, which would shed some light on the discrepancies. -Line 338. The cumulative GEP in autumn (and 2012, 2014, 2015 summer) is higher at TP39 (except for 2012). Does it contradict the argument in Line 338? -Line 352. "RE was higher at TPD". But the cumulative RE were lower at TPD in spring and autumn. -Line 384. Seasonal ET is more different in spring not autumn. Also, "daily ET" or "seasonal ET"? The other reviewer suggested to remove the statistical techniques from the previous section. A sentence was added at Line 325 to briefly highlight the t-tests used. I revised the majority (if not all) the instances where I mentioned comparisons. I added time scales and key words to highlight the comparison of rates or averages in different periods.

20. -Line 339. "the 2016 summer was the only period . . .". Clarify "sufficiently". Also, it seems a false statement to me because summer GEP in 2013 and 2017 are also greater at TPD. REMOVED

21. -Line 353. Any figure or data to support this statement? Daily rates but REMOVED

22. -Line 399. How the low WUE in winter is reflected in Figure 6c? Did you only use data from spring to autumn? If so, clarify in the manuscript. All months were plotted

23. -Line 405. Can you clarify "similar results"? The LUE at TPD is 30% higher than that at TP39. Fixed the figure to implement corrected APAR data

24. -Line 406. Is the annual and seasonal LUE shown in the manuscript? If not, clarify it in the manuscript by adding "(data not shown)". Also, as shown in Table 3, TPD has lower annual GEP, which contradicts with the "greater GEP" referred here. Reconcile. Similarly, TPD had higher annual (data not shown) and summer LUE ($p < 0.01$), although spring and autumn LUE was similar at both sites.

25. -Line 435. Do you mean "deciduous forest" instead of "conifer"? If not, add the correlation of annual NEP and summer RE for the conifer forest to Table 4. If the answer

is yes, I'd suggest delete this sentence because it conveys the same meaning as the following two sentences. Meant conifer, but only included the key linear relationships

26. -Line 434-435. Can you add a brief explanation for the relationship of RE and spring Ta. Could be because of the fact that there's only 6 data points, but the warmest spring/year (2012) had the lowest annual RE, which the highest annual RE (2017) saw the coldest spring. Similarly, the coolest year in our record (2014) had a very warm spring.

27. -Line 439-448. The annual GEP has no significant relationships with meteorological variables as stated in Line 425. But this paragraph talks about GEP and meteo controls. Is it only summer GEP discussed in this paragraph? Yes, only looking at summer fluxes

28. -Line 439. What does "flux parameterizations" mean here? Is it explained in the methodology section? If not, I suggest adding it to the methods section. Yes. Added a new section to the methods: 2.4 Estimating effects of meteorological variables on carbon component fluxes

29. -Line 578. Is the assumption of similar carbon assimilation valid here given the different NEP? Changed to: Assuming similar daily rates of carbon assimilation (GEP)

30. -Table 3. Why the GEP sum for Jan 1 to SOS is missing? They seem to be available in Fig. 3. The assumption is that leaves aren't present so GEP remains zero until the SOS

31.-Table 4. Can you change this table to a figure similar to Fig. 4? The reasons are (i) you'd be able to show the standard deviations; (ii) the positive/negative correlation would be easier to tell. Ultimately chose not to, but it could be added to an appendix if needed

32.-Table 5. What model did you use for this calculation? Highlighted in methods (2.4)

33.-I notice the uncertainty analysis for measurements and calculations is missing. Can you add a brief subsection to Methods section (or wherever you find appropriate) dedicated to uncertainties? Added a paragraph on the uncertainty analysis in Section 2.3

Minor comments:

34. -I suggest changing all "warm temperatures/Ta" to "high temperatures/Ta" in the manuscript. REVISED

35. -Line 78. Clarify "controls". Environmental/meteorological controls? REVISED determine the impact of meteorological controls on overall forest productivities

36. -Line 88-91. Are percentages available for the tree species? Not that we know of for the specific study area. Could probably be done by students in the future.

37. -Line 119. Did you use the momentum and heat fluxes in this study? If not, there's no need to mention them. We do not. REMOVED

38. -Line 258. What is the value of “record Ta”? Also, “record high Ta”.
Record high Ta conditions (exceeding 30-year mean daily maximum values)

39. -Line 315. Are the “days 230 to 290” 6-year mean? Explain.
At both sites, the cumulative CDD from DOY 230 to 290 (mid-August to mid-October; loosely based on the range of dates in Oishi et al. [2018]). They used DOY 210 to 290.

40. -Line 325-326. This statement is not clear. Clarify or delete. DELETED

41. -Line 347. Define “outlier”. RE within the deciduous forest was greatly reduced, leading to an apparent outlier (exceeding mean and standard deviation) in annual RE

42. -Line 398. Clarify “the ratio of monthly ET”. Then modify the figure caption accordingly. … linear relationships of the monthly total ET and GEP (calculating WUE)

43. -Line 354-355. Confusing sentence. How do “comparable” results shape the “differences”? Rephrase. REMOVED

44. -Line 363. “for either site”? It’s hard to tell that the monthly NEP is negative at TPD in Figure 5b. Rephrase. Figure inset highlights the negative TP39 NEP during the summer

45. -Line 416. P value for being “significant”? “linear relationships of monthly Ta and monthly VPD”? Linear relationships of the 6-year monthly mean Ta and VPD ($p<0.01$).

46. -Be concise. See examples below. -Line 325. “at first glance” is not necessary. -Line 341-342. “significant daily minimums and maximums” seems to be repetitive as “highly variable”. -Line 417. Delete “,”. -Line 409-410. Delete “and”. Also, make the sentence clearer. -Line 372. “the highest” ——-> “highest”. REVISED ALL

47. -Given the different time scales used here, I suggest be more mindful about the uses of “daily, season, annual” when talking about fluxes. -Line 261. In “Ta at both sites”, do you mean “daily mean Ta”? daily mean Ta -Line 360. Change “The NEP in the conifer...” to “The annual NEP” or “The cumulative NEP”. annual NEP -Line 352. “spring and autumn RE was higher . . .”. Do you mean “daily RE in spring and summer”? Sentence removed -Line 410. Delete “When first considering . . .”. DELETED Change “ET”—-> “Annual ET”. -Line 325. Should “daily patterns” be “seasonal patterns”? Seasonal Also, substitute “expanded upon in Table3” with “the cumulative fluxes in Table 3”, just to be clear and accurate. REVISED

48. -I noticed quite a few miscitation or misspelling or inaccurate statements. See some examples below. -Line 270. “daily reductions in PAR (shouldn’t it be APAR?)”. -Line 401. 4.7 ——> 3.82 gC kg-1 H2O. -Line 406. R2 = 0.96 ——> R2 = 0.86. -Line 535. “increases” ——-> “decreases”? -Line 538. “most years” ——-> “half of the years”? -Line 553. “during drought years” is not accurate. It’s really just 2016. REVISED ALL

49. -I have a few minor comments regarding the tables and figures. See below. -Table 3. Can you highlight the highest and lowest annual fluxes with colored boxes? -Be more clear with figure captions, especially for words like “daily, monthly, seasonal, and annual”. For example, “A daily time series” in Fig. 2 is a bit confusing. -Figure 3. Green-red combination is not color-blind friendly. Also, can you annotate SOS, EOG, SOB, and EDS on the top panels? -Figure 4 caption. Two “and”. REVISED.