Responses to Anonymous Reviewer 1 for paper: *Unraveling phenological to extreme drought and implications for water and carbon budgets*

The authors would like to thank the reviewer for taking the time to make thoughtful comments that will improve the manuscript. The questions, insights, and suggestions will help us to clarify, shape, and focus the readers on the main point of the article: that land-atmosphere interactions undergo rapid changes during flash drought not observed during drought or non-drought periods.

We respond to each of the Reviewer's comments below, which are in ***bold and italics***. Author responses are in blue with proposed manuscript changes in **bold**. New figures that we propose to include have labels N# throughout this document. Revised figures have the same figure number as in the original manuscript or point to the figure from the manuscript they will replace. Figures used to support claims in this document that are not going into the revised manuscript do not have Figure numbers but we have included figure captions.

Given the updated analysis and comparison of flash drought to non-flash drought periods, the authors use the following language throughout this document and will update accordingly in the revised manuscript. We use "flash drought" when referring to 2012 and "drought" or "drought years" for all other droughts. We use "non-drought" for all years that are not drought or flash drought.

The authors noticed the following major themes from the Reviewer comments.

1. A need to clarify differences between flash drought and drought
    a. Hypotheses need to better address what distinguishes flash drought from drought (i.e., timing, magnitude of effects, etc.)
    b. Strengthen analyses by using all available years from model outputs when comparing results from the flash drought year against 'non-flash' drought years
2. Some claims could be better supported with figures using model outputs that were not previously used, specifically:
    a. Including results from infiltration
    b. Exploring the relationship between phenology, GPP, stomatal conductance, and transpiration
3. Differences between model predictions, flux tower measurements, and MODIS observations need more discussion.
    a. Model values of GPP and ET were much lower than AmeriFlux and MODIS GPP during flash drought
    b. Discussion of observed differences should explore plant processes (e.g., reallocating carbon, tapping into deep groundwater, stomatal regulation, etc.) that the DCHM may be missing.

Response to Major Comment 1a:

We updated our hypotheses to distinguish flash drought from other droughts. They include language that allows us to answer more specifically how land-atmosphere interactions differ in a flash drought from other times. The hypotheses now read:

H1 During flash drought, there is an increase in days between precipitation events leading to larger reductions in total precipitation and infiltration as compared to non-flash drought events.

H2 Lower total infiltration observed during flash drought reduces soil water available for root water uptake, subsequently leading to reduced rates of transpiration, carbon uptake, and water use efficiency as compared to non-flash drought within a subseasonal time frame.

H3 In response to decreased water availability during flash drought, vegetation phenological states will be diminished as compared to non-flash drought years exacerbating the reduction of transpiration and carbon uptake.

Response to Major Comment 1b:

We now present our results using all years of available model output and framing the analyses in terms of flash drought vs “non-flash” drought conditions. We used the United States Drought Monitor (USDM, Svboda et al., 2002) to determine “drought” and “non-drought” years in Kansas. The Central and East Central Kansas climate regions contain the three study sites (see figure from USDM below). From the USDM time series of the climate regions, drought years were determined if an entire year was in drought or if parts of the region reached D2 (Severe Drought) or higher. The years 2006, 2011, 2013, 2014, and 2018 are labeled as drought years for analysis. The years 2007-2010, 2015-2017, and 2019 are labeled as non-drought years. The flash drought year 2012 is kept separate from other drought years in the analyses. This change will be updated in the Methods section as well (see response to specific comment below).

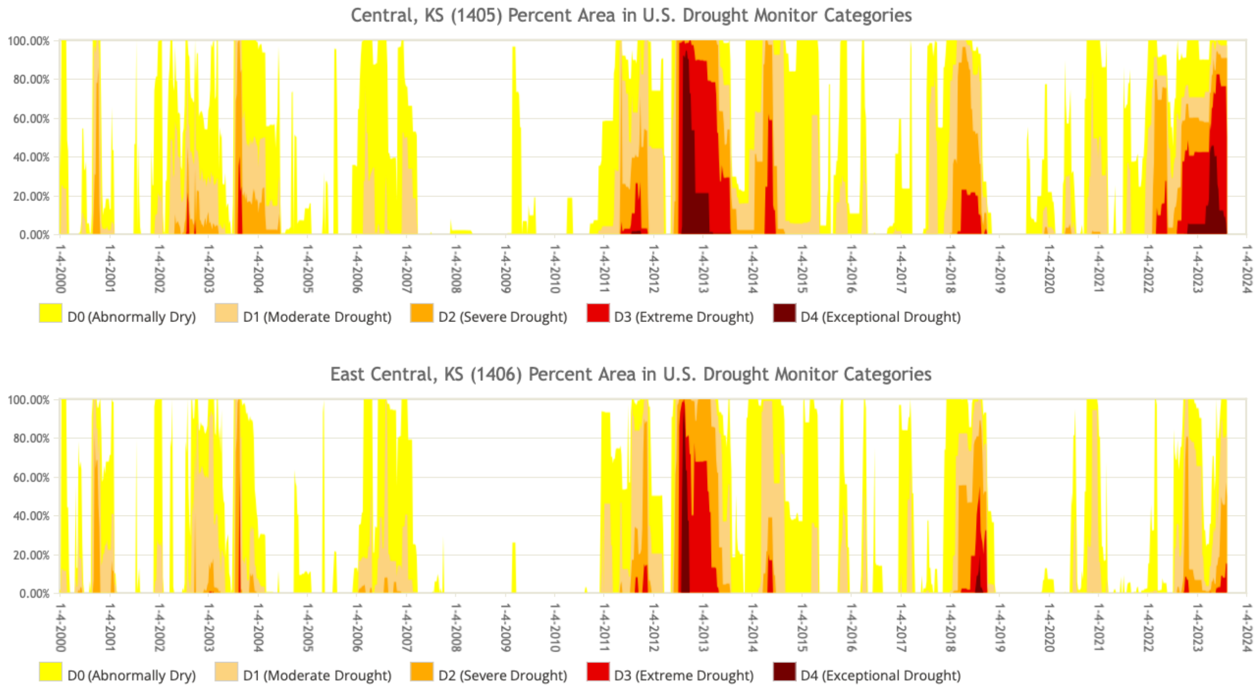

Figure. Percent land area in the U.S. Drought Monitor Categories for two Kansas climate regions that contain our study sites US-KFS, US-KLS, US-Kon.

Response to Major Comment 2a:

Throughout the paper, including the hypotheses, we made claims about infiltration and its relation to evaporation and root water uptake without showing any infiltration results. Hypothesis 1 has been updated (see above) to incorporate how timing of rainfall events impact total infiltration. To investigate this hypothesis we include plots of monthly accumulations of infiltration, evaporation, and root uptake, and figures relating infiltration to the amount of time between precipitation events to build a better understanding of how infiltration affects available water for plant use. We also address this in the discussion and conclusion. See below for new analyses related to specific comments.

Response to Major Comment 2b:

We take a closer look at sub-daily GPP and stomatal conductance during selected weeks throughout the growing season for a flash drought, drought, and non-drought year in order to evaluate how stomatal conductance reduces with increased vapor pressure deficit (VPD) and leads to subsequent declines in transpiration and GPP. We also investigate whether reductions in stomatal conductance or leaf area index (LAI) have a larger impact on GPP in order to evaluate if there exists a phenological dependence during flash drought. See below for new analyses related to specific comments.

In light of new analyses during the process of responding to the Reviewer comments, we propose a new title: ***Unraveling phenological and stomatal responses to flash drought and implications for water and carbon budgets***

Response to Major Comment 3a:

One reason for the discrepancies between modeled output and flux tower data was that plots of daily average rates of GPP and ET had to do with how we were calculating daily averages for the figure. While it made sense to average these variables over the entire 24-hour period for the flux tower data, the model shuts off GPP and evaporation when there is no incoming solar radiation leading to zeroes during half of the day. Thus, we only average GPP and ET over the active period within the model to avoid including the unrealistic zeros. This is how the DCHM model results were previously presented in Lowman and Barros (2016, 2018) and Lowman et al. (2018). Presenting the results differently was a mistake that has now been fixed. Additionally, we were able to find available gap-filled time series for GPP for US-KFS and US-KLS from the AmeriFlux FLUXNET database (Pastorello et al., 2020), allowing us to make comparisons where previous data was missing in the analysis.

Response to Major Comment 3b:

We propose to add the following text to the Discussion section of the manuscript to address this major comment:

In a recent paper, Giardina et al. (2023) argue that observed plant responses to water stress indicate the ability of plants to access deep groundwater and other stores of water that land surface models (LSMs) are not accounting for. The DCHM has similar soil moisture profiles to Noah-LSM and Hosseini et al., 2022, who used Noah-LM configurations, for both the 2012 flash drought and the 2018 drought. The DCHM also follows trends similar to AmeriFlux in 2012, but AmeriFlux top layer soil moisture values are much smaller from May to October of 2018, often under 0.1 m^3 m^{-3} during that time (Figure A1). Despite extremely low top layer soil moisture in 2018, AmeriFlux GPP reaches levels above 10 gC m^{-2} d{-1} coinciding with a brief recharge in soil moisture at the end of June. The DCHM estimates of GPP are often less than 50% of AmeriFlux GPP in 2012 and 2018. The model results from the Noah-LSM similarly underestimate GPP and overestimate soil moisture during these drought periods (Hosseini et al. 2022), suggesting that access to deep water reserves are responsible for these differences (Giardina et al. 2023).

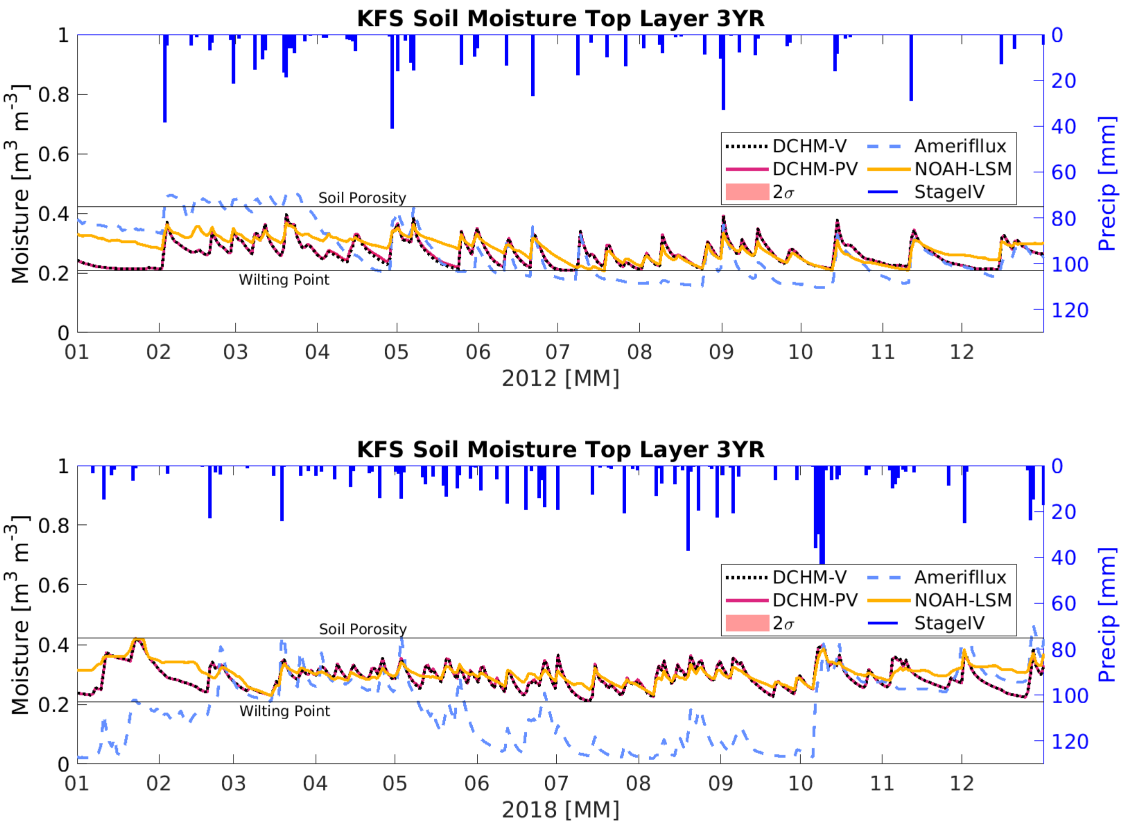

Figure A1 (2018 to be added to manuscript.) Top layer soil moisture at US-KFS from the DCHM compared to AmeriFlux, Noah-LSM (from NLDAS-2). Daily precipitation totals are indicated on the top axis.

Using predictive phenology with NOAH-LM, which can account for carbon reallocation to leaves, stems, roots, and soils, Hosseini et al. (2022), compared predicted estimates to flux tower measurements of GPP. Even while accounting for carbon movement, they found that during June, July, and August they underestimated tower data by 100 gC m^{-2} at US-Kon while overestimating by the same amount at US-KFS in April, May, and June (averaged across an 11-year study period encompassing wet and dry periods). The DCHM-PV, which does not account for carbon reallocation, performs similarity to the Hosseini et al. (2022) results, suggesting the accounting for carbon allocation cannot explain underestimating GPP in 2012 and 2013. The DCHM-PV compares more favorably against AmeriFlux data during 2012, the flash drought year, at US-KFS and US-KLS (Figure 8, A11) compared to 2018, a drought year. This suggests that there are missing processes in both the DCHM and the NOAH-LM that cannot capture plant water use during drought, and cannot be attributed to carbon allocation.

During drought and flash drought, DCHM-PV values also compare favorably with MODIS and tend to be slightly higher than MODIS during a non-drought year like 2019. During drought and flash drought, the DCHM-V and DCHM-PV tend to follow similar trajectories

but in response to little water stress, the predictive phenology model predicts increased carbon uptake compared to the DCHM-V results which align more with MODIS in 2019. AmeriFlux estimates of GPP during June and early July of 2012 and 2018 are also above estimates from MODIS. This suggests that during drought and flash drought, plants are able to maintain higher levels of GPP than what can be recreated in land surface models and satellite remote sensing. Differences in DCHM-PV and AmeriFlux GPP cannot be attributed to carbon reallocation since the NOAH-LM model accounts for carbon reallocation and similarly underestimated GPP compared to flux tower data (Hosseini et al. 2022). The working hypothesis is that plants have access to deeper water stores than can be accounted for in land surface models, as suggested by Giardina et al. (2023).

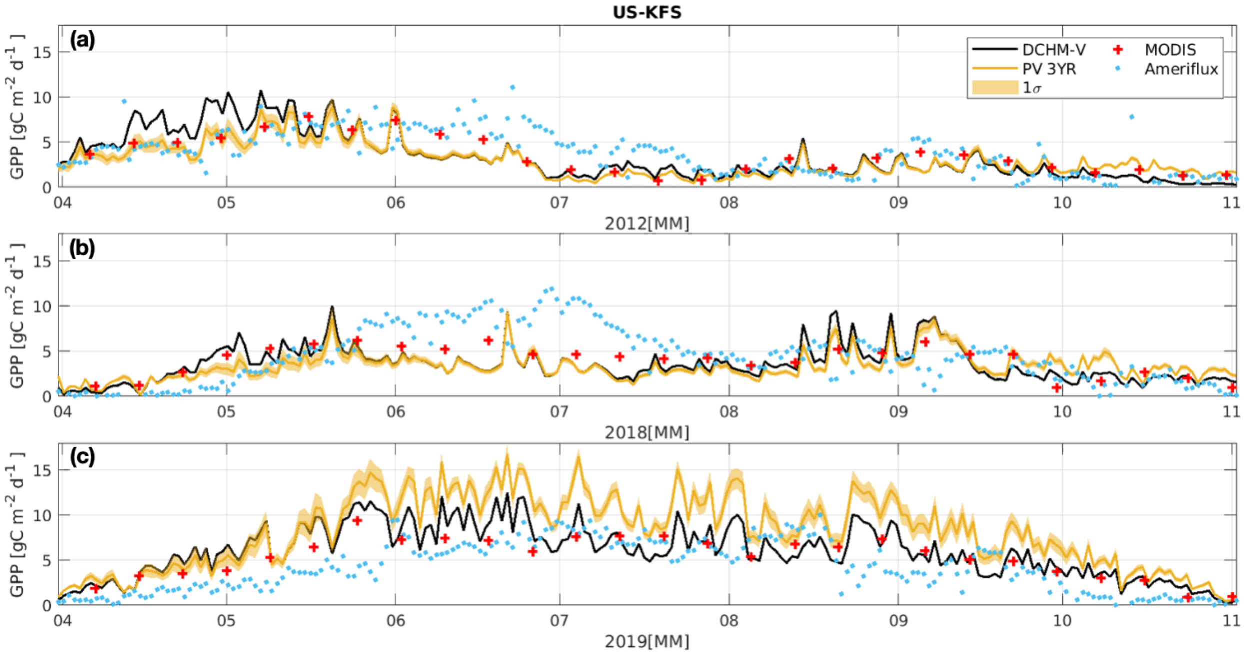

Figure 8 (updated). Time series of gross primary productivity, GPP, at US-KFS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. One standard deviation is shown for the DCHM-PV simulations. MODIS GPP are shown as red crosses and AmeriFlux GPP as small dots.

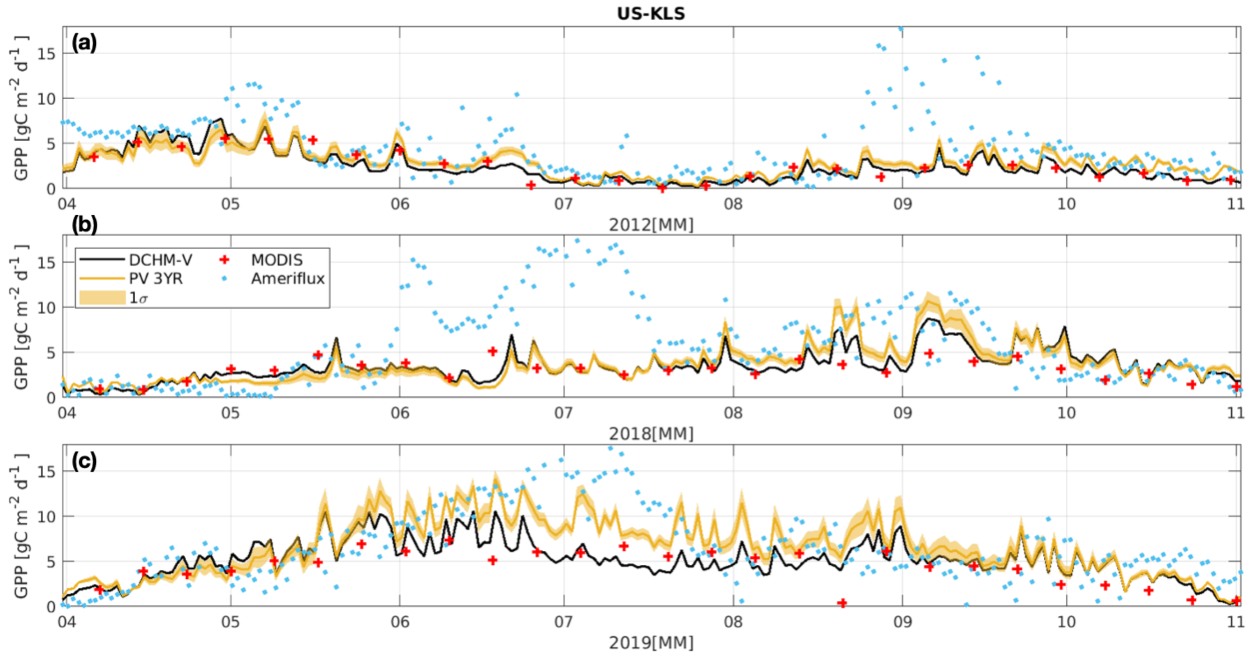

Figure A11 (updated). Time series of gross primary productivity, GPP, at US-KLS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. One standard deviation is shown for the DCHM-PV simulations. MODIS GPP are shown as red crosses and AmeriFlux GPP as small dots.

This papers offers a detailed analysis of drought responses for vegetation in the Midwestern US using an ecohydrologic model and assimilation of MODIS FPAR and LAI data along with flux tower data. The paper addresses some important ecohydrologic questions about how rapidly transpiration and carbon assimilation decline during drought and how changing phenology, specifically above-ground photosynthetic capacity, accelerates GPP losses. The paper utilizes advanced modeling techniques and builds on previous work that has established useful ways to assimilated remote sensing data into the DCHM model.

While that paper has significant potential, it did not really make strategic use of the model and observations to address some of the questions posed in the introduction.

For example, the hypothesis posed do not really address the issues of ‘flash’ drought.

In H2 The idea that drought causes both carbon uptake and transpiration to decline is something that is quite well understood - and there is ample evidence that this occurs- we know plant shut down when they run out of water. There are elements of the timing of this that are perhaps less well understood - and questions about how water use efficiency changes during a drought- and indeed the authors get at this to some extent in the paper - The hypothesis should reflect this. There would also be ways to frame the study (and hypothesis) to look at the relative impact of “flash” drought versus other

types of drought that could be interesting. Some additional thinking about how to use the model to test more nuanced (and informative) hypothesis would be strengthen this paper

The Reviewer's comments about H2 and other hypotheses are well taken. We have edited the Hypotheses in the paper to clearly articulate changes we expect to see in land-atmosphere interactions during flash drought that differ from drought (see response to Major Comment 1a above). The following changes were made to H1, H2, and H3.

H1 now addresses differences in timing of precipitation events during flash drought, drought, and non-drought years. H1 now reads as follows.

H1 During flash drought, there is an increase in days between precipitation events leading to larger reductions in total precipitation and infiltration as compared to non-flash drought events.

H2 now has language that broadens the types of vegetation responses we see due to decreased infiltration. We also have language to compare vegetation responses across flash, regular, and non-drought periods.

H2 Lower total infiltration observed during flash drought reduces soil water available for root water uptake, subsequently leading to reduced rates of transpiration, carbon uptake, and water use efficiency as compared to non-flash drought within a subseasonal time frame.

H3 was updated to draw comparisons of flash drought to non-flash drought years and to include more specific language other than "plant-atmospheric interactions".

H3 In response to decreased water availability during flash drought, vegetation phenological states will be diminished as compared to non-flash drought years exacerbating the reduction of transpiration and carbon uptake.

We also updated Figure 1 to incorporate the edited hypotheses.

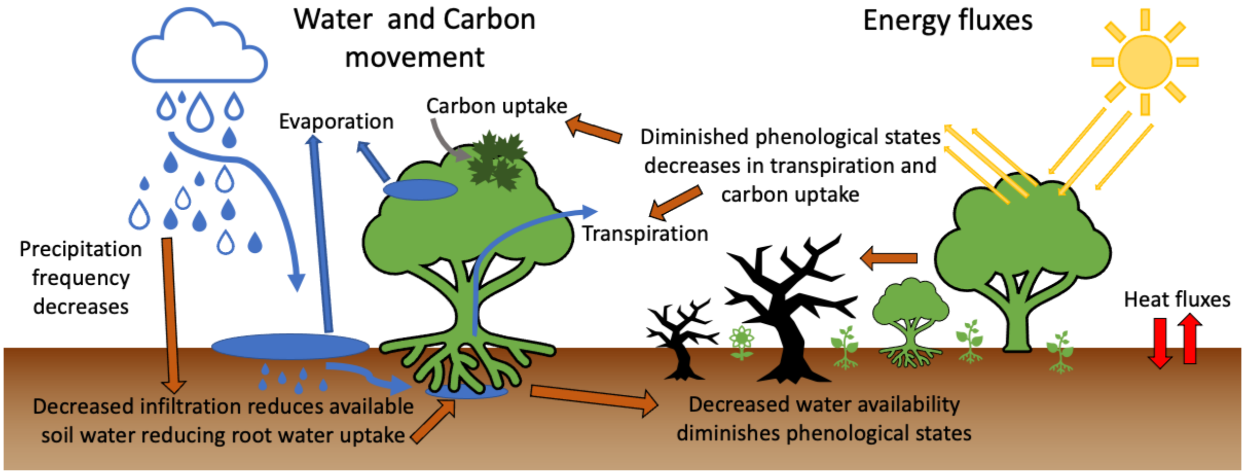

Figure 1 (updated). Schematic of water, carbon, and energy fluxes with hypotheses about ecological response to flash drought indicated with orange arrows. Decreased frequency of precipitation events leads to decreased infiltration and less water available for plant use during flash drought as compared to non-flash drought periods. The cascading effects of the decreased water availability, exacerbated by the decreased phenological states, include rapid reductions in transpiration and atmospheric carbon uptake to levels below drought.

Throughout the paper, there are statements made that are not well supported by graphs or analysis - for example - that evaporation exceeds infiltration (this is indirectly shown but it would be much convincing to show this directly - and model results could do this). In another perhaps more salient example for the paper, the authors state phenology declines reduce carbon and water exchanges (H3) - one could argue that because plants have already shut down stomates at that point in the season, phenological declines do not further reduce transpiration - I'm not suggesting this is true but graphs presented do not clearly rule this out in the testing of H3.

The Reviewer's comments are well taken. We have combed through the paper to make sure that any claims are fully substantiated and demonstrated with figures or data from our results. We have taken a critical look at unsubstantiated claims throughout the paper as referenced by this review. In doing so, we also noticed that we could better support the updated hypotheses by incorporating additional results and analyses. For changes to hypotheses, see response to Major Comment 1a above.

Regarding the first comment on evaporation exceeding infiltration, we acknowledge that we previously did not show any results for infiltration. In addition to updating the wording of our hypotheses, we will add a subsection to our results section of the manuscript that discusses infiltration.

We propose to add a section in the results **Section 3.X Plant Available Water** with subsections for infiltration and root water uptake.

Section 3.X.1 - Infiltration

During non-drought years, infiltration is over 100 mm per month, on average, from April to July with the highest rates in May (Figure N1). Conversely, during drought years, infiltration in April-July is less than non-drought years, and mean infiltration is often one standard deviation below infiltration in non-drought years. Furthermore, monthly accumulated infiltration is less for the flash drought year compared to both drought and non-drought years, indicating there is less available water for plant use. This is the case at US-KFS from April-October of 2012, with September totals similar to drought years. A drastic decline in May infiltration at US-KLS and US-Kon indicated infiltration rates that are 1-2 standard deviations below average drought conditions. All sites had infiltration rates below 100 mm for all months during 2012 with the exception of August at US-KLS.

Low infiltration totals during the flash drought year can be attributed to lower precipitation accumulations coupled with an increase in the number of days between precipitation events (Figure N2) and an increase in atmospheric demand for water (Figure N4). During the drought, and non-drought years, the average number of days between rainfall events within a month ranges from 1 to 7 days, while the lower end for the flash drought year is at 2.5 days. Additionally, during drought and non-drought years, monthly infiltration exceeds 150 mm, but remains at or below 75 mm for all sites aside from August 2012 at KLS where monthly infiltration is ~110 mm. In 2012, all three sites averaged over four days between rainfall events during May, June, and July with US-KFS averaging over six days between rainfall events during both May and June and more than five days in July (Figure N2a). Across all three sites from April-October 2012, there were more than four days between precipitation events 80% percent of the time compared to just 20% of the time in non-flash drought years.

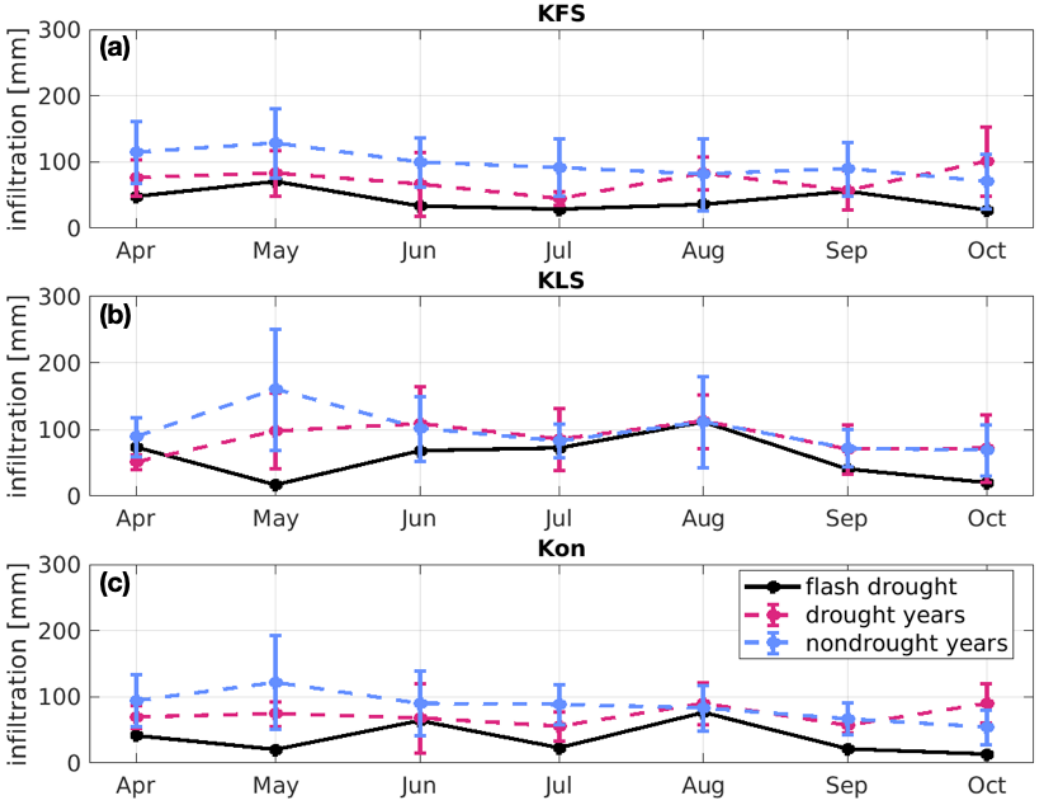

Figure N1. DCHM-PV 3 YR ensemble means of monthly infiltration accumulations for drought (red dashed line) and non-drought (blue dashed line) years compared to 2012 (black solid line) for all three study sites. Monthly sums are computed from the ensemble means of the 2000 Monte Carlo simulations then averaged across drought or non-drought years. Error bars represent one standard deviation across drought and non-drought years, respectively. Drought years are 2006, 2011, 2013, 2014, 2018 and non-drought years are 2007-2010, 2015-2017, 2019.

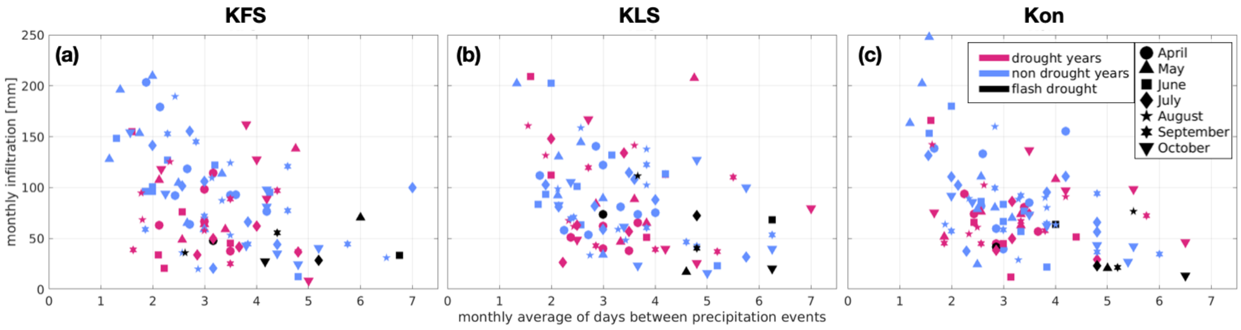

Figure N2. Monthly infiltration accumulation plotted against the monthly average of days between precipitation events for (a) US-KFS, (b) US-KLS, and (c) US-KON. Each shape indicates one month over which the averaging occurred and colors distinguish flash drought (black) from drought (red) and non-drought (blue) years.

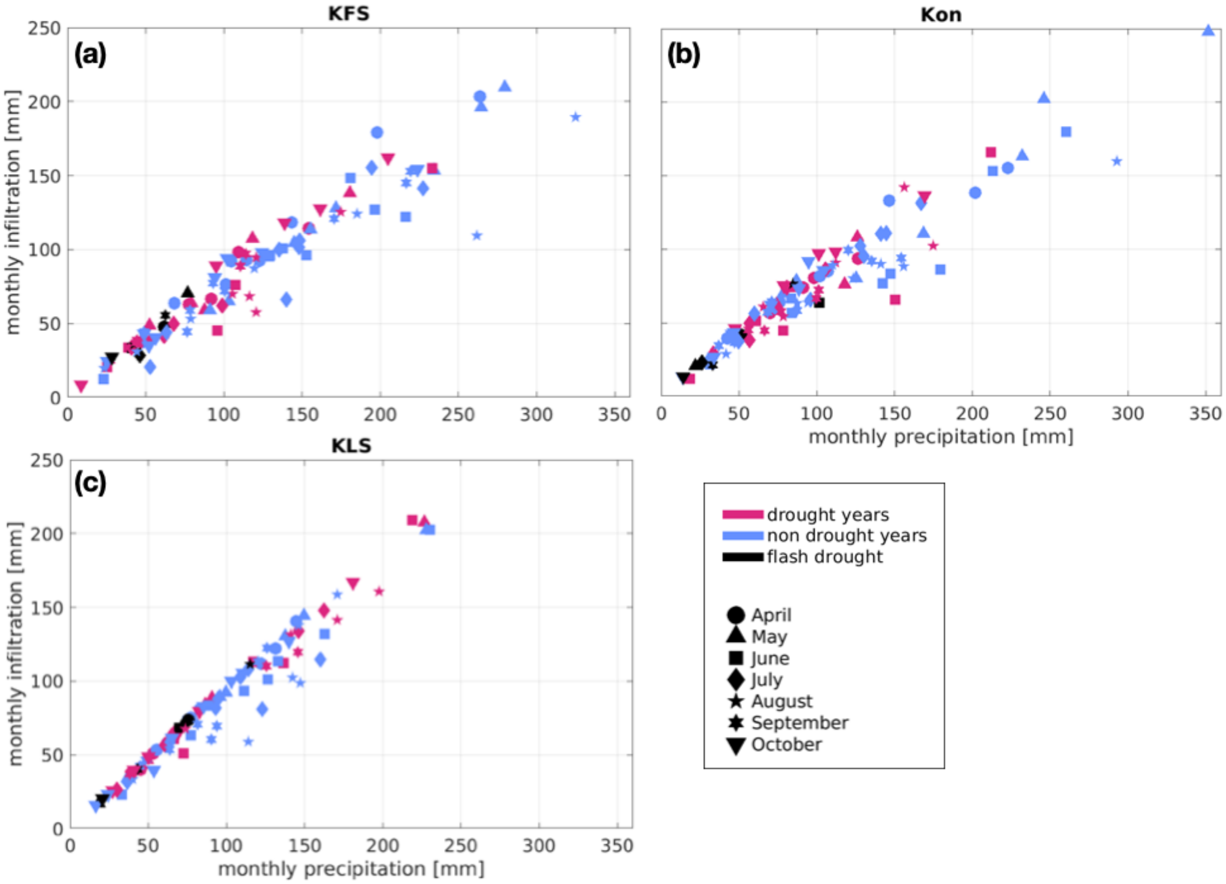

Figure N3 (appendix addition). Monthly accumulation of infiltration versus precipitation. Each shape indicates one month over which the averaging occurred and colors distinguish flash drought (black) from drought (red) and non-drought years (blue).

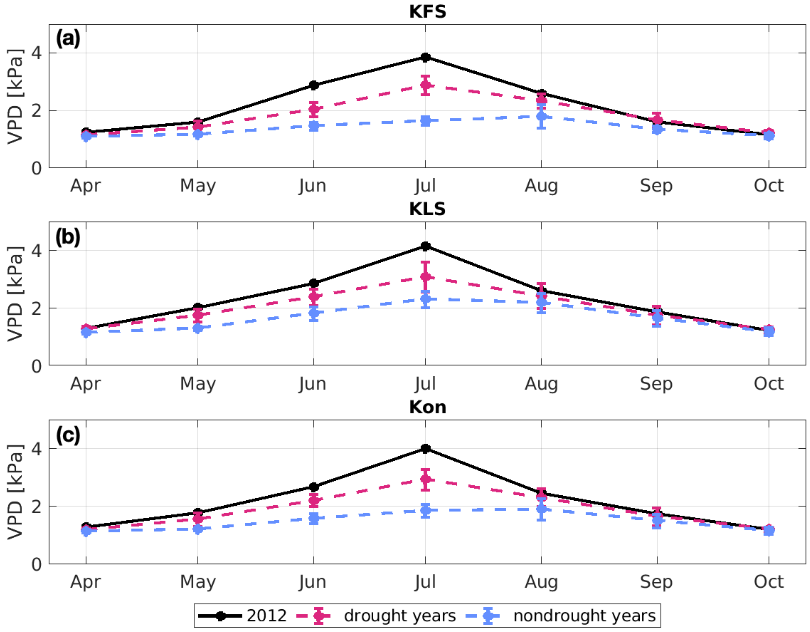

Figure N4 (appendix addition). Monthly average vapor pressure deficit [kPa] for the three AmeriFlux sites from April - October for the flash drought year 2012 (black), drought years (red), and non-drought years (blue). The error bar represents one standard deviation across drought and non-drought years.

In addition to a new section on infiltration, we will also incorporate root water uptake to this new section as a subsection.

Section 3.X.2 Root Water Uptake

Root water uptake is lower in flash drought than non-flash drought years for the period June to September (Figure A7). The middle soil layer is responsible for up to four times more root water uptake than the other layers so a major decline in root water uptake through the middle layer demonstrates how root water uptake is altered during drought (Figure A7). While root water uptake starts out in 2012 at levels above average non-drought years, it falls to more than one standard deviation below drought averages by July. This drastic shift is likely due to lower infiltration (Figure N1) and drives down rates of transpiration within the model over the same period.

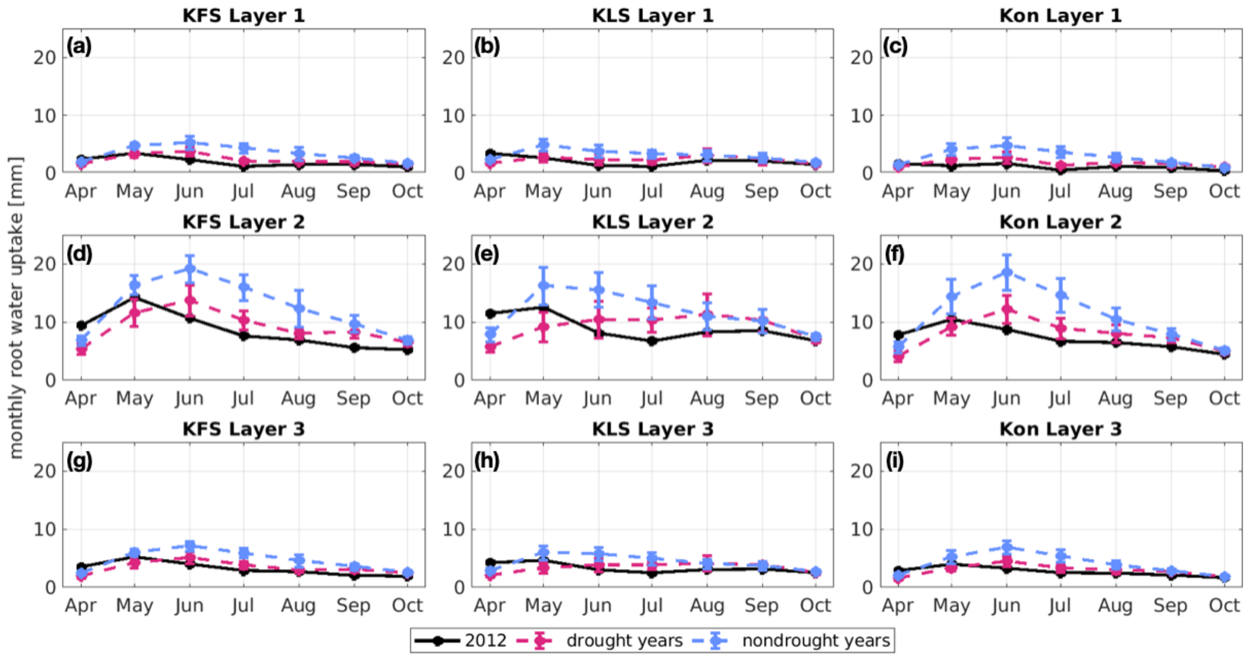

Figure A7 (Replacing A7 and A8). DCHM-PV 3YR monthly root water uptake totals for drought (red) and non-drought (blue) years compared to 2012 (black) across three soil layers for our three study sites. Monthly sums are computed from the ensemble means of the 2000 Monte Carlo simulations then averaged across drought or non-drought years. Error bars represent one standard deviation across drought and non-drought years, respectively. Drought years are 2006, 2011, 2013, 2014, 2018 and non-drought years are 2007-2010, 2015-2017, 2019.

Response to second part of comment:

The authors appreciate the reviewer bringing awareness about a potential counterpoint to our claim that ***"phenology declines reduce carbon and water exchanges (H3) - one could argue that because plants have already shut down stomates at that point in the season, phenological declines do not further reduce transpiration - I'm not suggesting this is true but graphs presented do not clearly rule this out in the testing of H3."***

We acknowledge that there may be more at play with phenology, stomatal conductance and plant-atmosphere interactions which has led to new ways of interpreting the model output. Below, is a new proposed section that evaluates the relative contributions of changes in stomatal conductance vs LAI during the flash drought. We also address this with the proposed change in title.

Section 3.XX Stomatal Conductance

An objective of this work is to evaluate whether changes in phenology vs changes in stomatal regulation have a stronger control on carbon uptake and transpiration during flash drought. We consider how GPP co-varies during flash drought, drought, and

non-drought years with sub-seasonal changes in LAI and stomatal conductance. During flash drought, both stomatal conductance and LAI remain low throughout the growing season and GPP falls below 10 gC m^{-2} (Figure N7).

As expected during non-drought years, there exists a wider range of values of stomatal conductance, LAI, and GPP throughout the peak growing season (Figure N6). In non-drought years, there is a clear seasonal cycle in the clockwise movement through the stomatal conductance-LAI parameter space. Stomatal conductances increase faster than LAI in the early season before reaching maximums around June-July. After LAI peaks, there is first a reduction in stomatal conductance and GPP at higher LAI before LAI decreases through August-September.

Once stomatal conductance values surpass 4 mm s^{-1}, we observe increases in GPP. LAI seems to have some effect on GPP, but not the same magnitude as stomatal conductance. At peak season values of LAI above 3 m^2 m^{-2}, we observe smaller fluctuations in GPP than observed with stomatal conductance. (Figures N6 and Appendix). When LAI decreases in the latter part of the growing season, stomatal conductances can still remain at relatively constant levels. It appears that stomatal conductance exerts a stronger control on daily GPP than LAI.

In drought years, stomatal conductance from April-October averages 1.4 mm s^{-1} across all sites (Figure N5) compared to an 2.3 mm s^{-1} in non drought years (Figure N6) and 1.1 mm s^{-1} in flash drought (Figure N7) . While both 2012 and 2018 have low values of stomatal conductance, LAI, and GPP, an important difference is the near-zero stomatal conductance during June and July for a range of LAI values (1-2 m^2 m^{-2}, Figure N7) that is not observed in 2018 and other drought years aside from US-KLS in 2011 (Figure N5l).

In 2012, stomatal conductances in the first week of May were as high or higher than in 2019, a non-drought year (Figure N9) at US-KFS. By July, due to increased VPD, stomatal conductances diminished (Figure N9b). Compared to a drought year, e.g. 2018 where stomatal conductance rates are similar in all three months, 2012 undergoes a major shift. This is likely due to plants regulating their stomata under dry atmospheric conditions and less likely attributed to the phenological changes as FPAR and LAI were similar in 2012 and 2018. In 2019, a non-drought year, some August stomatal conductance and GPP rates are slightly higher than in July at US-KFS despite decreased phenology (Figure N8c) suggesting that drought-induced changes in phenological state does not have as strong as a control on carbon and water exchanges as we have hypothesized.

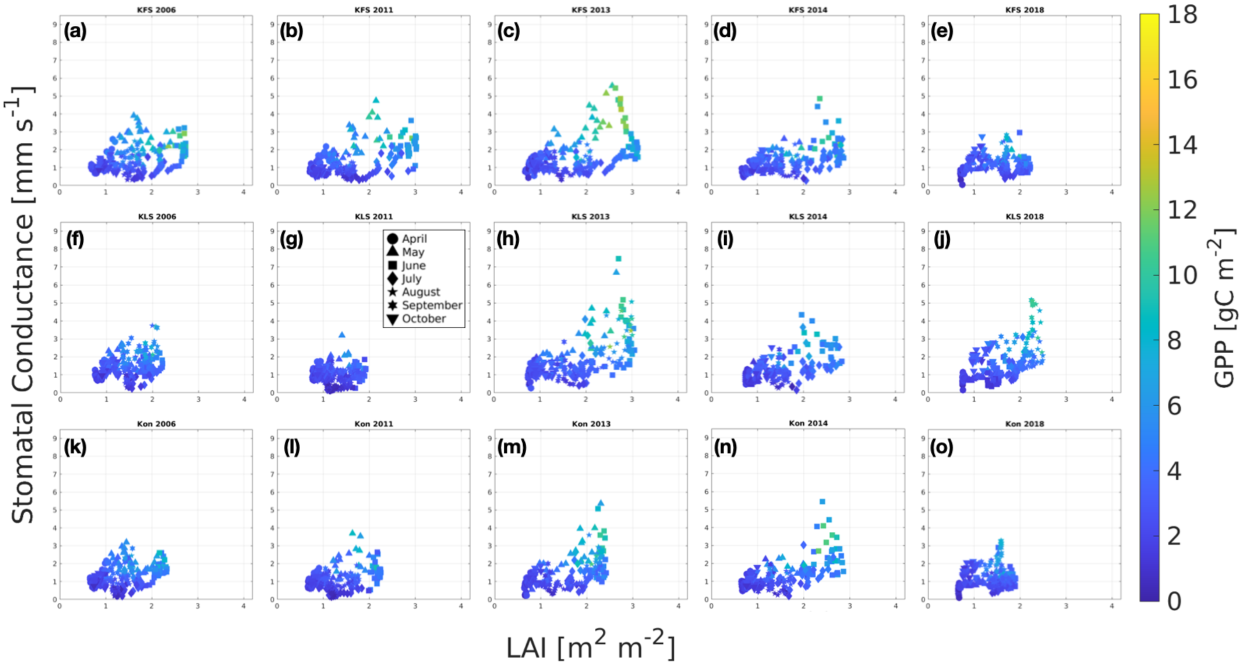

Figure (N5 - appendix addition). Daily stomatal conductance [mm s^{-1}] vs leaf area index, LAI [m^2m^{-2}] for all three sites and selected drought years. Marker shapes indicate an individual day from April 1 - October 31 from the selected drought year. Each month is given a unique shape and daily accumulations of gross primary productivity [gC m^{-2}] are indicated by color.

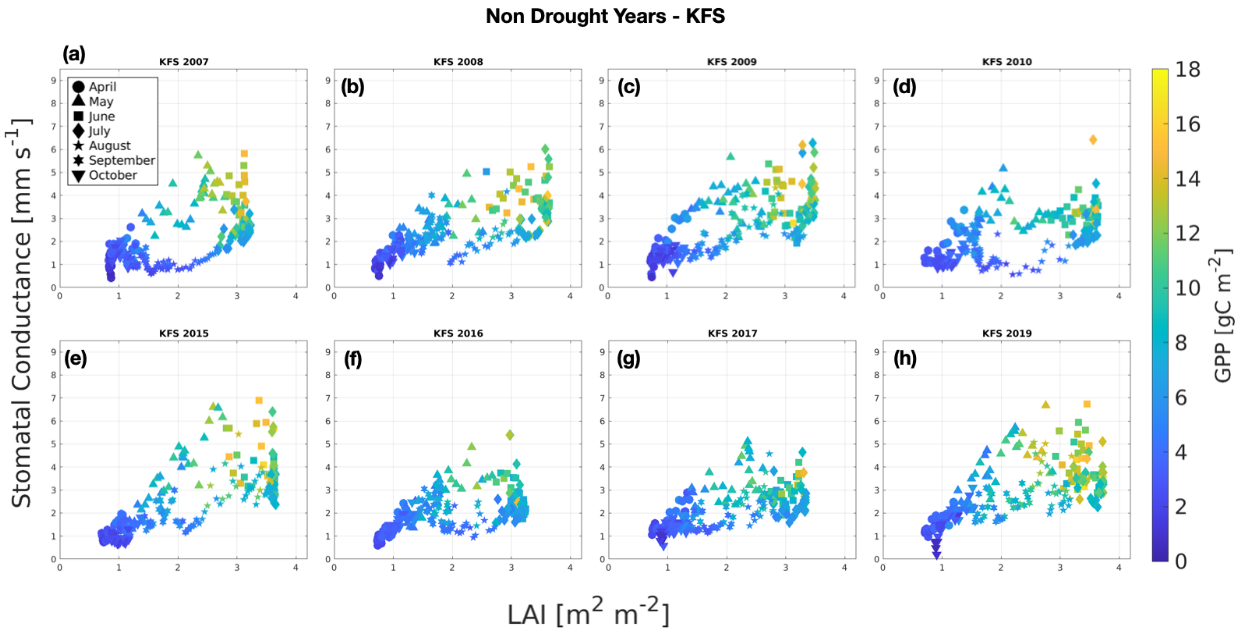

Figure (N6 will be added to the appendix. Similar figures for US-KLS and US-Kon will also be included). Daily stomatal conductance [mm s^{-1}] vs leaf area index, LAI [m^2m^{-2}] for US-KFS for selected non-drought years. Marker shapes indicate individual days from

April 1 - October 31 from the selected drought year. Each month is given a unique shape and daily totals of gross primary productivity [gC m^{-2}] are indicated by color.

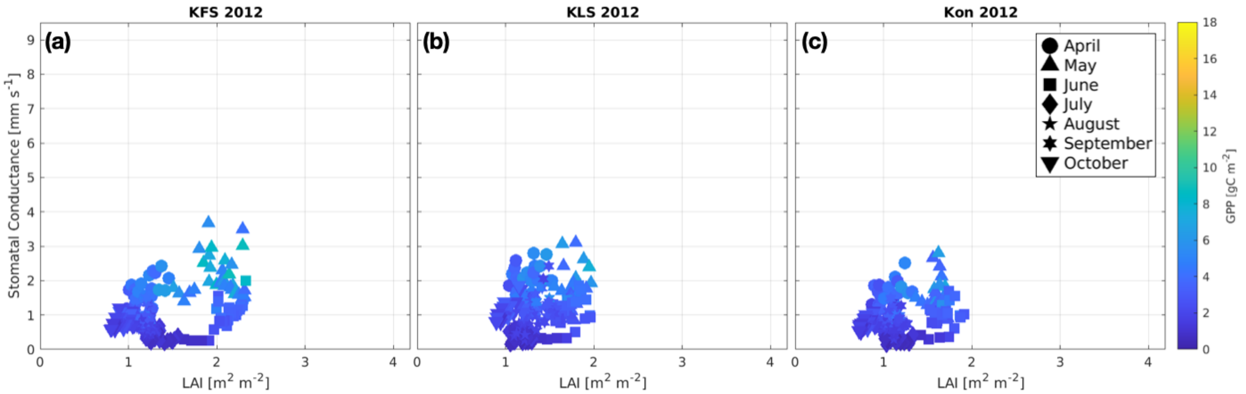

Figure (N7 appendix addition). Daily stomatal conductance [mm s^{-1}] vs leaf area index, LAI [m^2m^{-2}] for all three sites during the flash drought of 2012. Marker shapes indicate individual days from April 1 - October 31 from the selected year. Each month is given a unique shape and daily totals of gross primary productivity [gC m^{-2}] are indicated by color.

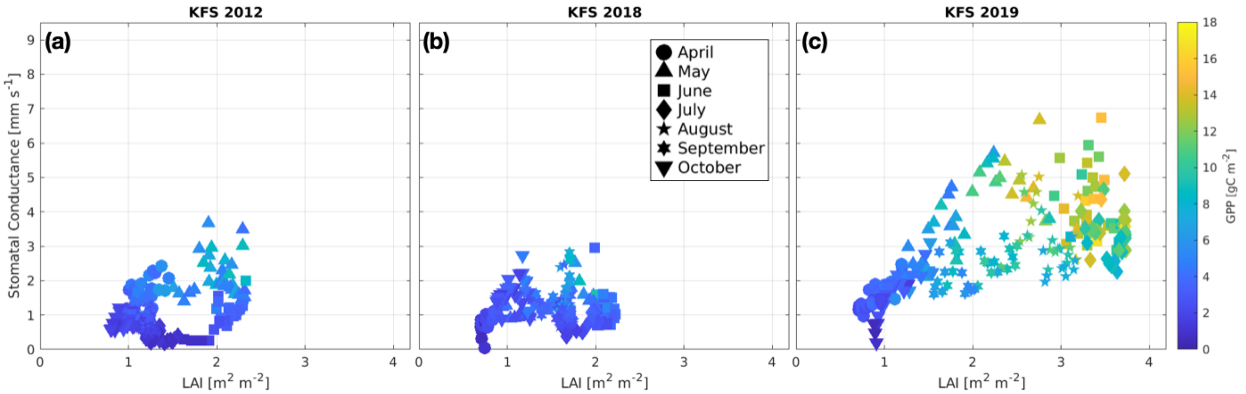

Figure (N8). Stomatal conductance [mm s^{-1}] vs leaf area index, LAI [m^2m^{-2}] for US-KFS for 2012, 2018, and 2019. Marker shapes indicate individual days between April 1 - October 31 from the year. Each month is given a unique shape and daily accumulations of gross primary productivity [gC m^{-2}] are indicated by color.

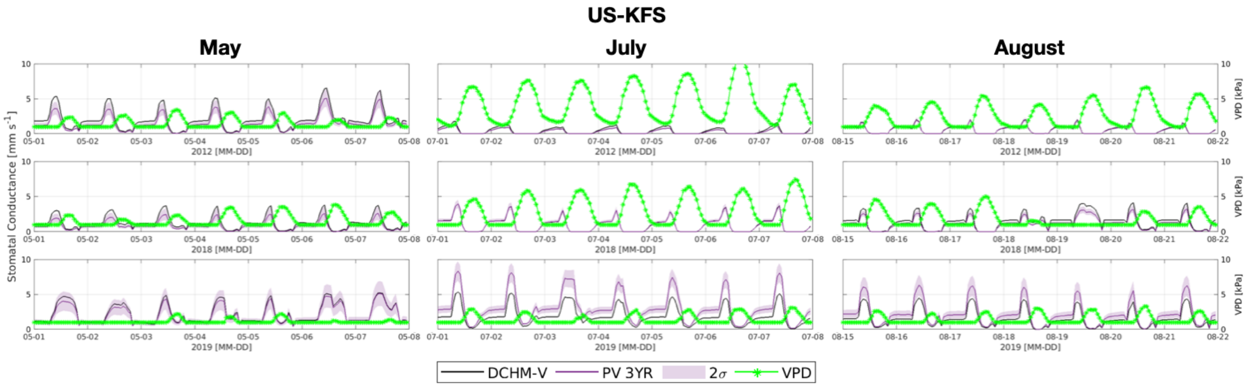

Figure (N9). Hourly stomatal conductances [mm s^{-1}] for one week in May, July, and August of 2012, 2018, and 2019 compared with vapor pressure deficit (VPD, kPa) for US-KFS.

The author also make statements about flash droughts but do not really distinguish flash drought from other types of drought in their analysis - They have multiple years that they do not really make use of.

The authors thank the review for this comment. This is one of the major comments we address in Major Comment 1 above and help drive most of the new analysis that better distinguishes flash drought from other drought and non-drought conditions.

As shown above, we are now making direct links to a drought year (2018) as well as averaging across drought and non-drought years. We determined the drought and non-drought years from the USDM (Svoboda et al., 2002). We will update the methods Section 2.7 beginning at line 243 with the change in bold.

Line 243: "…We highlight results from the three AmeriFluxes sites for 2012 (flash drought), **2018 (drought),** and 2019 (**non-drought)** to draw conclusions about plant response during flash drought **and how they differ from drought and non-drought years.** We also evaluate **model outputs** from 2006-2019 to assess the **differences** between the DCHM-V and DCHM-PV model configurations **during drought and non-drought years compared to a flash drought year. During this time period, we identified drought years as 2006, 2011, 2013, 2014, 2018 and non-drought years as 2007-2010, 2015-2017, 2019 using the USDM for the Central and East Central Kansas climate regions (Svoboda et al., 2002). Drought years were determined by whether parts of the region reached the D2 "Severe Drought" classification or higher. When computing drought and non-drought averages, we use the years listed here. Transpiration is calculated from total root water uptake through the three soil layers and total evaporation is computed from summing evaporation from ground and canopy surfaces allowing us to partition ET into evaporation and transpiration."**

Also to address this comment, we enhance our analysis to highlight that flash drought should be considered as the time period leading up to drought and the intensification rate (Otkin et al., 2018). As such, we will put more emphasis on the change leading from non-drought to drought (May to July of 2012) in results and discussion.

Finally there are also important differences between modeled and flux tower data that may be critical in the understanding flash drought responses. These differences need to be more rigorously explored (see detailed comments below)

The reviewer's comment is appreciated. As noted above in response to Major Comment 3, the reason for the discrepancies was an error made when creating the figures that has now been fixed. Please see the updates in Figures 8 (above in response to Major Comment 3) and Figure 9 below.

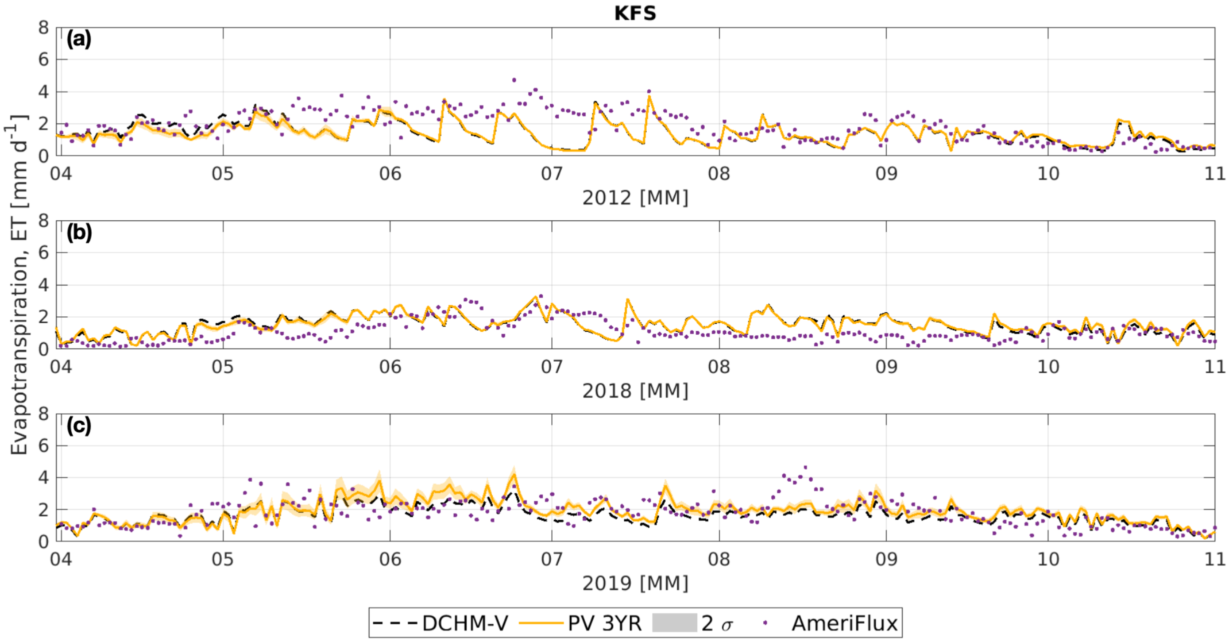

Figure 9 (updated). Time series of evapotranspiration, ET, at US-KFS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. Two standard deviations are shown for the DCHM-PV simulations. AmeriFlux ET is derived from latent heat measurements and shown as small dots.

One possible explanation for why flux tower data differs from model output is that the flux tower estimates incorporate a variety of vegetation types within the fetch contributing to the vertical fluxes, rather than the single vegetation type used within the model. Additionally, the size and orientation of the contributing fetch varies in time depending on measurement height and turbulent fluxes (Chu et al., 2021). Another difference could be that models may not be able to fully represent how vegetation can maintain ET by accessing groundwater or deep soil moisture, ultimately biasing models towards more severe effects of drought on vegetation (Giardina et al., 2023).

Further, the flux towers exist within a 4 km by 4 km region defined by the StageIV spatial grid cell used in the DCHM. Flux tower spatial extents range from a couple hundred meters to a few kilometers (Baldocci, 2003, Schmid, 1994) making the 4 km grid cell near the maximum range. Subgrid scale heterogeneity can lead to considerable discrepancies between parameterized and actual fluxes (Schmid, 1994). Since the DCHM treats the entire grid cell as a single vegetation type, our results hold some uncertainty as we cannot account for the heterogeneous mix of vegetation and land-use present on the ground (see Figure below). Inside of this grid is deciduous forest that could influence tower readings, that the DCHM does not account for.

Figure (non included in manuscript). US-KFS AmeriFlux tower site at the center of a 4 km by 4 km grid representing vegetation heterogeneity of the surrounding region.

We propose to update Section 4.5 Limitations:

Line 465: Capturing phenological responses **and subsequent changes to carbon and water fluxes** within a physically based model is not without its limitations.

We propose to remove lines 470-480, beginning with “For example…”

Line 470: …temporal and spatial scales. **The flux towers exist within a 4 km by 4 km region defined by the StageIV spatial grid cell used in the DCHM. Flux tower spatial extents range from a couple hundred meters to a few kilometers (Baldocci, 2003, Schmid, 1994) making the 4 km grid cell near the maximum range. Subgrid scale heterogeneity can lead to considerable discrepancies between parameterized and actual fluxes (Schmid, 1994). One explanation for why flux tower data differs from model output is that the flux tower estimates incorporate a variety of vegetation types within the fetch contributing to the vertical fluxes, rather than the single vegetation type used within the model. Additionally,**

the size and orientation of the contributing fetch varies in time depending on measurement height and turbulent fluxes (Chu et al., 2021).

Another difference between modeled and flux tower data could be that models may not be able to fully represent how vegetation can maintain ET by accessing groundwater or deep soil moisture, ultimately biasing models towards more severe effects of drought on vegetation (Giardina et al., 2023). Using predictive phenology with NOAM-LM, which can account for carbon reallocation to leaves, stems, roots, and soils, Hoessini et al. (2022), compared predicted estimates to flux tower measurements of GPP. Even while accounting for carbon movement, they found that during June, July, and August they underestimated tower data by 100 gC m^{-2} at US-Kon while overestimating by the same amount at US-KFS in April, May, and June (averaged across an 11-year study period encompassing wet and dry periods). The DCHM-PV, which does not account for carbon reallocation, responds to drought and flash drought differently than what is observed at flux tower sites. It matches better with AmeriFlux data during 2012, the flash drought year, at US-KFS and US-KLS (Figure 8, A11) compared to 2018, a drought year.

During drought and flash drought, DCHM-PV values also agree favorably with MODIS and tend to be slightly larger than MODIS during a non-drought year like 2019. During drought and flash drought, the DCHM-V and DCHM-PV tend to follow similar trajectories but in response to little water stress, the predictive phenology model predicts increased carbon uptake compared to the DCHM-V results which align more with MODIS in 2019. Drought levels of AmeriFlux observed GPP during June are above observed non-drought levels. Even during flash drought, GPP tended to be slightly higher than non-drought June levels. This suggests that during drought and flash drought, plants are able to maintain higher levels of GPP. Differences in DCHM-PV and AmeriFlux GPP less likely to be attributed to carbon reallocation since another predictive model that accounts for carbon reallocation underestimated AmeriFlux values, and more likely due the suggestion by Giardina et al., (2023) that vegetation is likely accessing water stores in ways that are different from how land surface models are currently representing it.

Some detailed comments.

H1 is actually two hypothesis - it would be useful to separate them

See response to Major Comment 1 above regarding our updated hypothesis H1.

The data sets and modeling proposed here tend to focus on relatively shallow surface soil - including citations of expected rooting depths for the PPTs would be helpful support for the implementation (especially given that flux tower observations of ET tend to be higher than the model in dry years)

The review comments are well taken. We propose two changes to the paper beginning at line 189.

1. For clarity, we remove parenthetical depths in line 189-190 since they are stated more clearly in the following sentence.
2. We add a reference to expected rooting depths using a combination of soil and PFT.

Line 189: **"We maintain the use of 8 cm for the top soil layer for model stability, but use 35 in for root zone depth and 72 in for the depth to the impermeable layer (Soil Survey Staff). This yields the three soil layers: top (0-8 cm), middle (8-89 cm) and bottom (89-183 cm) consistent with Kansas soil surveys. Rooting depths in the DCHM are determined using exponential root distribution functions whose input parameters are determined by PFT (Zeng, 2001 and Jackson et al., 1996). Soil layer and rooting depths align with the different combinations of soil textures and PFTs found in Thornthwaite and Mather (1957)."**

line 225 - Some additional (just one or two sentences) information about of how ensembles of phenology parameters are established is needed here (e.g what is done for each of the 3 periods to select the 2000 parameter distributions shown in Figure 4) - There needs to be a bit more context so that reader understands Figure 4 and what controls the variation in parameter sets.

We propose to address this comment by adding the following text on lines 215-221, where we put more emphasis on the data assimilation procedure.

Line 217: Using **soil water potential and VPD** outputs from the DCHM-V and updating…

Line 219: ...and mixed conditions). **We implement a dual state-parameter EnKF that simultaneously predicts FPAR and LAI and parameters that determine the growing season index. The parameter estimation procedure consists of initializing ensembles of parameters by sampling them from Gaussian distributions (Lowman and Barros, 2018), and predicting phenological state variables from the sampled parameters at every timestep (hourly). When MODIS LAI and FPAR are available (every 8 days), we update ensembles of the parameters in the DCBP before updating predictions of phenological state variables (this is the data assimilation step) to ensure that predictions of FPAR and LAI do not stray far from observations. The data assimilation period determines whether the generated parameters represent wet, dry, or mixed precipitation regimes.**

line 235 - The simulation period is relatively short - and isohydric-anisohydric differences may or not be distinguishable within the 3 years - thus you cannot really state that the vegetation model parameters trained on dry conditions will represent isohydroic vegetation?. Especially given that parameter values seem to change depending on period (Figure 4) but vegetation PFT does not.

The authors appreciate the reviewer's comments.

Lowman and Barros (2018) showed that assimilation period can determine the water stress adaptations for the modeled vegetation state. We propose the following edits beginning at line 235.

Line 235: …affects the development of flash drought. **It has been shown under varied climatological conditions plants can be highly adaptable, transitioning from isohydric to anisohydric in a single season (Guo et al., 2020). Lowman and Barros (2018) showed that assimilation period can determine the water stress adaptations for the modeled vegetation state.** Broadly speaking vegetation model parameters trained using **data from years with minimal rainfall represent plants that are accustomed to drier conditions and therefore exhibit more regulation in their water use tendencies (Lowman and Barros, 2018)**.

line 246 - That transpiration is calculated from root water uptake makes sense but it doesn't follow that this allows you to "to partition ET"…you would need to have a separate calculation of total ET to do that. Clarify

The reviewer's comment is well taken. The model computes total ET (Figure 9) from totaling surface evaporation from soil and canopy and from computing transpiration as root water uptake through the three soils layers. We address the partitioning in the addition to the Methods Section 2.7 above and copied here:

Line 243: "…We highlight results from the three AmeriFluxes sites for 2012 (flash drought), **2018 (drought),** and 2019 (**non-drought)** to draw conclusions about plant response during flash drought **and how they differ from drought and non-drought years.** We also evaluate **model outputs** from 2006-2019 to assess the **differences** between the DCHM-V and DCHM-PV model configurations **during drought and non-drought years compared to a flash drought year. During this time period, we identified drought years as 2006, 2011, 2013, 2014, 2018 and non-drought years as 2007-2010, 2015-2017, 2019 using the USDM for the Central and East Central Kansas climate regions (Svoboda et al., 2002). Drought years were determined by whether parts of the region reached the D2 "Severe Drought" classification or higher. When computing drought and non-drought averages, we use the years listed here. Transpiration is calculated from total root water uptake through the three soil layers and total evaporation is computed from summing evaporation from ground and canopy surfaces allowing us to partition ET into evaporation and transpiration."**

Line 259 - For clarity it would be helpful to be consistent in the naming conventions- e.g gamma or growth parameter not both

We now use the gamma once it is defined throughout the remainder of the manuscript.

Line 258- in what way is this in agreement with Lowman and Barros (e.g the choice of longer period for reducing uncertainty) - in a way that's not so surprising - more information usually reduces uncertainty?

The authors appreciate the need for clarification. Line 258 now reads: **"... in agreement with Lowman and Barros (2018) who found that using assimilation periods with both wet and dry conditions has the effect of capturing adaptive plant water use strategies."**

Line 259 That gamma values vary by site could be do to differences in climate (note that game values vary across wet and dry years) - so it is not a given that it varies by plant functional type - rather this is an assumption (e.g I think that you are assigning plant functional type parameters based on this analysis)- The wording of this paragraph could make that point more clearly

The reviewer's comments are well taken. The gamma parameter may vary based on differences in both climate and land cover type. We propose the following change:

Line 259: The values of γ vary by site due to **a combination of local climate and plant functional type (PFT)**. **US-KFS, modeled as a savanna, has the lowest value of gamma and standard deviation.** The smaller magnitudes…

Line 265 - "slower" relative to what?

We appreciate this comment and have incorporated language making more clear comparisons like the one needed here throughout other parts of the analysis. In particular, we make sure to note comparisons between "flash drought versus drought or non-drought years". This language also helps to distinguish between the model results from drought and non-drought periods from WET and DRY assimilation periods.

Line 264: ...slower senescence and reduced variance when using the 3YR assimilation parameters as **compared to the WET and DRY parameters** during…

The rationale for the continued focus (beyond Figure 4) of differences due to parameters sets based on wet, dry or both years is unclear - Given that using both wet and dry years clearly reduces uncertainty, I'm not sure why there is a need to compare estimates of FPAR, LAI, The authors may have a reason for this but if so it needs to be emphasized in the text. Removing this would allow the focus to be on DCHM-PV performance and the "actual" phenological mediated vegetation responses.

The reviewer's comments are well taken.

Figure 4 helps to establish the 3YR assimilation period as the inference period with minimal uncertainty. However, we still have the goal to capture differences in model outputs across the three different assimilation periods. This is done in Figures 5,6,7, and 9 which show FPAR, LAI, yearly sums of GPP and ET, and daily ET, respectively. Figures 5 and 6 support the conclusion that the 3YR inference period does indeed show slower change in phenological state compared to WET and DRY. We agree that further comparison with the WET and DRY results is no longer needed in the body of the manuscript. We can include WET and DRY results in the appendix. We replaced Figure 7 with monthly accumulations of GPP and ET from the DCHM-PV 3YR. We removed the DCHM-PV WET and DRY from Figure 9.

The WET inference period predicts parameters using model outputs of soil water potential and VPD during a year that received ample rain, and therefore plenty of water for plant use, and less atmospheric demand for water. We are assuming that this inference period represents vegetation that is not accustomed to water stress and therefore is less conservative in water use strategies (Lowman and Barros, 2018). This means that when plants run low on available water (e.g., in 2012), they will consume water normally and exacerbate dry down before abruptly shutting down functions dictating water and carbon cycling. This is in contrast to parameters produced during the DRY inference period. The assumption there is that vegetation will be more conservative in water use strategies, shutting down at the first sign of stress. In both cases, mean γ values are higher than the 3YR inference period meaning that vegetation during the 3YR inference period is more likely to make steady changes and adaptations to water stress and less likely to make abrupt changes as seen with both the WET and DRY simulations. However, the difference between 3YR, WET, and DRY is generally minimal in terms of the magnitude of water and carbon exchanges and detailed discussion of these results is no longer included in the analysis.

Section 3.12 and Section 3.1.3 - If I understand the methods correctly - Figure 6 and 7 show results using parameters conditions on prior information (e.g MODIS assimilation during the calibration period) and MODIS results for 2012 and 2019 - Given that, additional discussion about fit with MODIS would be helpful - How well does DCHM-PV do. Overall it captures patterns fairly well but there are some notable exceptions (e.g loss of FPAR in July and August at US-KFS in MOIDS that is not tracked by model) - It would be useful to have some presentation of model performance here

We appreciate the reviewer's comments. We think these comments refer to Figures 5 and 6 regarding FPAR and LAI. We updated Figures 5 and 6 to include 2018 (see below). We propose to provide additional text comparing MODIS to the DCHM-PV FPAR and LAI estimates.

Before section 3.2 Vegetation responses:

Line 309: **Generally, the predictive phenology model compares favorably with the seasonal changes observed in MODIS FPAR and LAI (Figures 5 and 6). In the summer, at US-KFS and US-KLS during 2019, the model tends to predict FPAR and LAI values higher than MODIS. In 2019, at US-KFS, MODIS observed a steady decline in FPAR from 0.8 to 0.6 throughout July with an increase back to 0.8 over an 8-day period at the beginning of August (Figure 5c). The DCHM-PV results do not show the same decline. Similarly for LAI at US-KFS (Figure 6c), MODIS observes a drop and then abrupt increase in LAI with the model estimates higher than MODIS. Yet, in June 2019 at US-Kon, the model estimates are lower than MODIS LAI.**

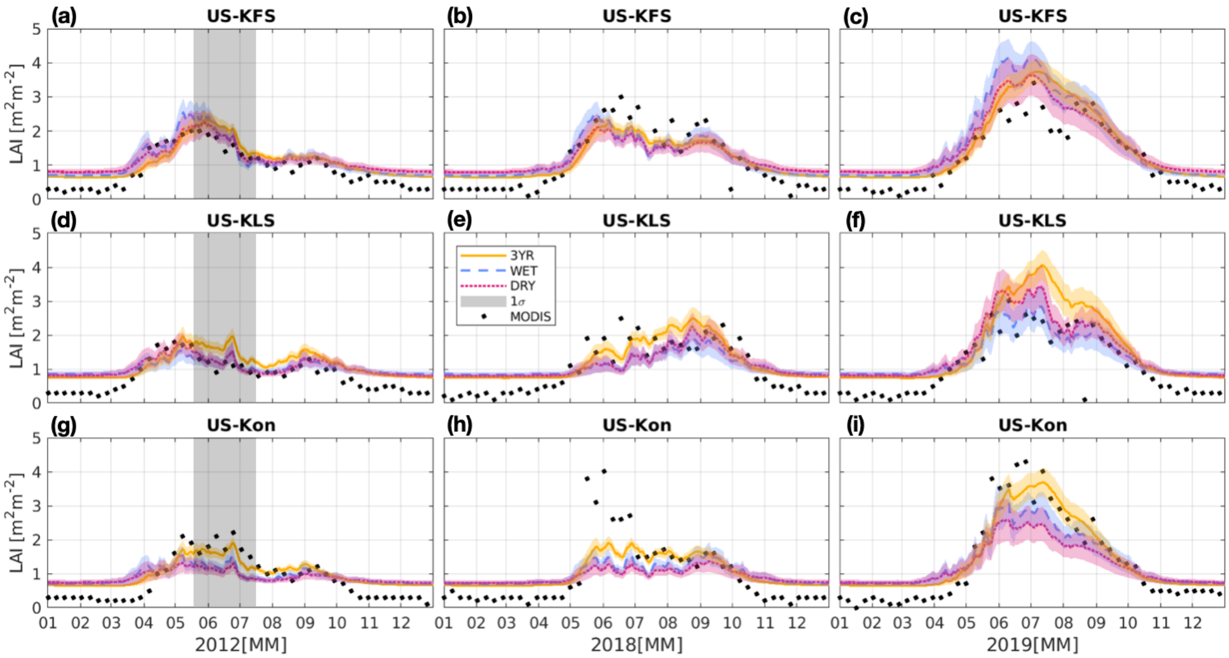

Figure 6. Time series of leaf area index (LAI) predicted from DCHM-PV for the flash drought year (2012), a drought year (2018), and a non-drought year (2019). Colors indicate the different data assimilation periods (yellow - 3YR (2003-2005), blue - WET (2005), red - DRY (2003)), with corresponding shaded regions representing one standard deviation of model outputs from the 2000 ensemble simulations. The 8-day MODIS MOD15A2H LAI is shown in black markers. The gray shaded region highlights the June to July decrease in FPAR during the 2012 flash drought.

To be added to discussion:

The higher model estimates of FPAR and LAI during summer 2019 could be due to the model accounting for excess water availability and other meteorological conditions (temperature, VPD, etc.) resulting in an increase in FPAR and LAI. MODIS estimates FPAR and LAI are based on radiative transfer models using bidirectional reflectance of incoming radiation from the red and near infrared bands (Myneni et al., 2015; Yan et al., 2016). MODIS GPP is directly dependent on observations of FPAR (Running & Zhao, 2015). This difference is apparent in DCHM-PV estimates of GPP exceeding estimates from the DCHM-V and MODIS GPP during the same period where the DCHM-PV predicts larger values of FPAR and LAI during 2019 (Figure 8). Our model performance against MODIS is similar to that found in the Hosseni et al. (2022), who also used a predictive phenology model coupled with Noah-LM. Across all 11 years in that study, their dynamic vegetation models tended to underestimate June and July LAI at US-Kon and slightly overestimate at US-KFS.

The DCHM-PV tends to compare favorably against MODIS during flash drought and non-drought at US-KFS and US-KLS (Figure 6 a,c,d,f) and underestimate those sites during drought (Figure 6 b,e). At US-Kon, MODIS LAI during May, June, and July tends to

be above DCHM-PV estimates. Differences between our model at 4 km and MODIS at 500 m spatial scales could depend on vegetation heterogeneity. Differences in land cover classification could be another reason for discrepancies between modeled and observed FPAR and LAI. Though we use MODIS to determine the land cover type, we take the most frequent land cover type at 500 m to determine the value of the 4 km grid cell. After upscale, the pixel at US-KFS is labeled as a Savanna but the 500 m MODIS grid cell containing US-KFS is classified as grassland. Regardless of the classification differences, the spectral reflectance method used by MODIS is inherently different from the predictive phenology routine used in the DCHM-PV, specifically in that it cannot account for how soil water availability influences vegetation growth (Lowman and Barros, 2018).

Line 315 - which are water stress years (e.g 2012). Also can you note which method (or averaged across all methods) does the 1kgCm2 reduction come from

The authors appreciate this comment from the reviewer. We are replacing Figure 7 and this paragraph has been rewritten so lines 310-319 will be removed. Instead, we compare DCHM-PV 3YR results across drought and non-drought years as listed above instead of showing yearly sums of GPP and ET.

3.2.1 GPP

Monthly averages of GPP accumulations from DCHM-PV ensemble means throughout the water year (April - October) indicate that carbon uptake falls below drought averages from May to June during the flash drought year of 2012 (Figure 7 a,c,e). Flash drought carbon assimilation amounts remain below drought levels before converging to average drought/non-drought levels by the end of October. GPP levels are consistently up to 50% lower in drought years compared to non-drought years. During the flash drought, GPP monthly totals in June to August 2012 are at least one standard deviation lower than drought years over the 2006-2019 simulation period. June GPP in 2012 is 40-50% of drought and less than 30% of non-drought years. An even greater discrepancy comes in July with carbon assimilation amounts being less than 30% of drought levels and 15% of non-drought levels.

Line 321: …for the flash drought year (2012), **a drought year (2018), and a non-drought year (2019). GPP rates are lower in drought than non-drought, but the decline in rates of GPP from at or above average in May to near zero by July in 2012 (Figure 8, A11, A12). During drought, rates of GPP are low but steady throughout the growing season. The rapid decline in GPP over the two month period is what distinguishes the 2012 flash drought as a period of time when rates of transpiration and carbon uptake are rapidly changing. Our results match closely with MODIS GPP, but the DCHM-PV overestimates MODIS in a drought year, which corresponds with the overestimation of MODIS phenology at the same time (Figure 6). Simulated GPP tends to underestimate flux tower GPP during June and July. From June to July in 2012, carbon uptake decreased from about more than 5.0**

(depending on site) to less than 1.0 gC m^{-2}day^{-1}. These types of declines do not happen in a drought year (e.g. 2018).

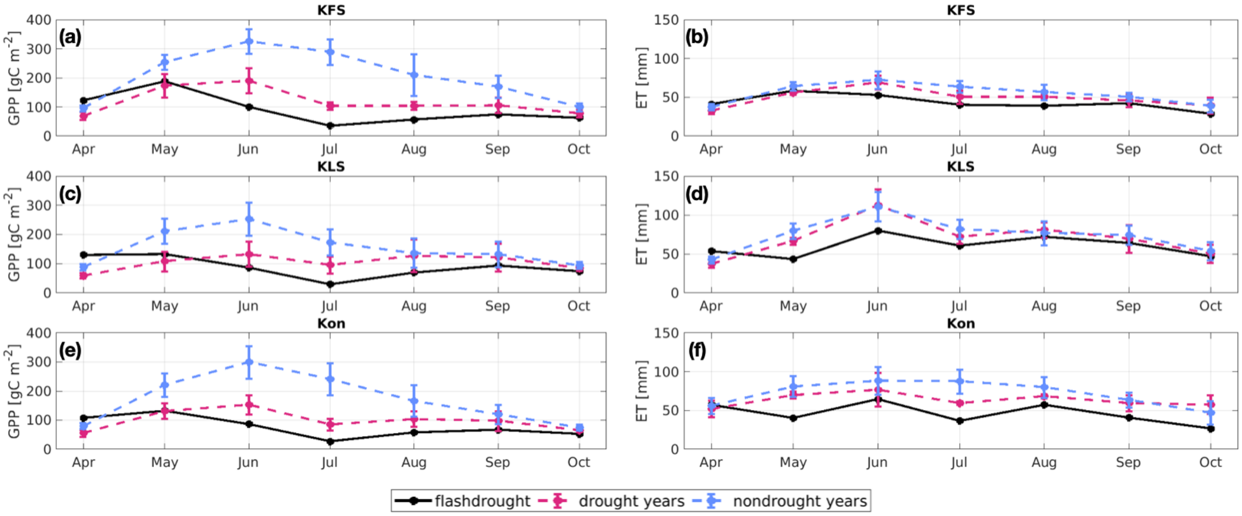

Figure 7 (replacement). DCHM-PV 3YR monthly totals of GPP (a,c,e) and ET (b,d,f) for drought (red) and non-drought (blue) years compared to 2012 (black) for US-KFS, US-KLS, and US-Kon AmeriFlux sites. Monthly sums are computed from the ensemble means of the 2000 Monte Carlo simulations then averaged across drought or non-drought years. Error bars represent one standard deviation across drought and non-drought years, respectively.

We replace yearly totals in Figure 7 with monthly averages of accumulated GPP and ET for drought and non-drought years. We also include monthly averages of ET separated as evaporation and transpiration in a new figure.

Line 328: **Monthly accumulations of ET for the flash drought years are shown against accumulations averaged across regular and non-drought years for the three study sites (Figure 7 b,d,f). Accumulations are from DCHM-PV 3YR simulations and averaged across ensemble members. ET accumulations are lower in the flash drought year, particularly at US-KLS and US-Kon. ET amounts during drought periods are slightly lower but generally similar to non-drought at US-KFS and US-KLS, indicating that total ET may not be a strong indicator of drought. However, looking at the same accumulations with ET parsed into evaporation and transpiration offer a different perspective.**

Simulated transpiration accumulations follow similar trajectories as GPP during flash drought (Figure N10 a,c,e). Indeed, transpiration rates are higher during flash drought than non-drought years in April but quickly level off before declining from May to July. July transpiration totals in 2012 fall below one standard deviation of drought years. At US-KFS, evaporation rates are not too different from normal drought and only slightly higher than non-drought years throughout the entire growing season (Figure N10b).

Evaporation accumulations follow different trajectories at all three study sites but have transpiration patterns that link transpiration and GPP.

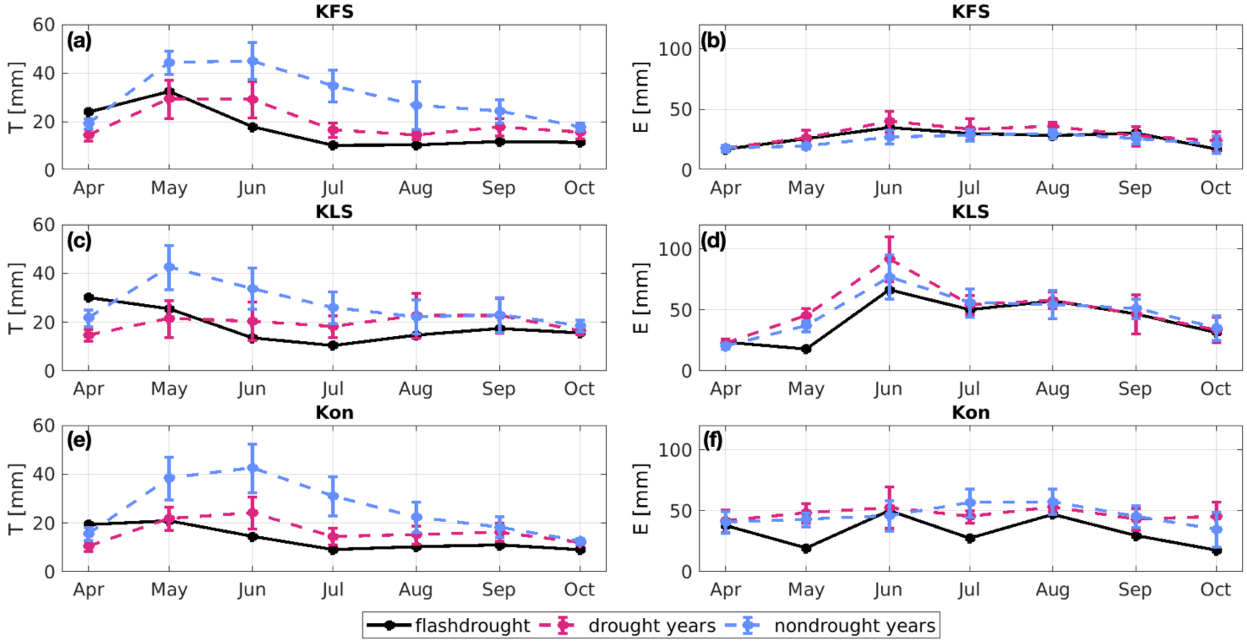

Figure N10. Simulated evapotranspiration partitioned into evaporation and transpiration from the DCHM-PV 3YR. Transpiration accumulations are computed from root water uptake through the three soil layers. Monthly averages are computed from the ensemble means of the 2000 Monte Carlo simulations then averaged across drought (red) or non-drought (blue) years. The flash drought year 2012 is shown in black. Error bars represent one standard deviation across drought and non-drought years. Drought years are 2006, 2011, 2013, 2014, 2018 and non-drought years are 2007-2010, 2015-2017, 2019.

lLine 380 - The arguments in the first 3 sentences of this paragraph need a bit more detail. Just because there are fluctuations in evaporation this doesn't not necessarily mean that "all" water evaporated before it had a chance to infiltrate.

We believe this quote comes from line 340. The authors appreciate this comment and we agree. The new infiltration section described in Major Comment 2a supports the claim that reduced infiltration drives down plant water availability.

The timing of infiltration and evaporation are different. While infiltration and evaporation occur in response to precipitation events, the timescales associated with each process are different. Water infiltrates the soils following precipitation but may not be available for plant use in instances when increased VPD (Figure N4) leads to enhanced evaporation of soil water. It was therefore prudent to account for accumulations over monthly timescales to better compare infiltration and evaporation totals.

Infiltration does occur, but it is drastically reduced during flash drought. Water availability is reduced due to both a lack of precipitation and increased days between precipitation events (Figures N1, N3) leading to reductions in root uptake (i.e. transpiration, Figures A7, N10). The large fluctuations in ET are therefore more associated with increased evaporative demand, reduced infiltration limiting root water uptake,and overall lack of water availability.

Line 339: …are a result of evaporation in response to precipitation (Figure 10a). This suggests **that following precipitation events during flash drought onset, ET is dominated by evaporation. Reduced available water infiltrating the soils limits water available for root water uptake.** Since…

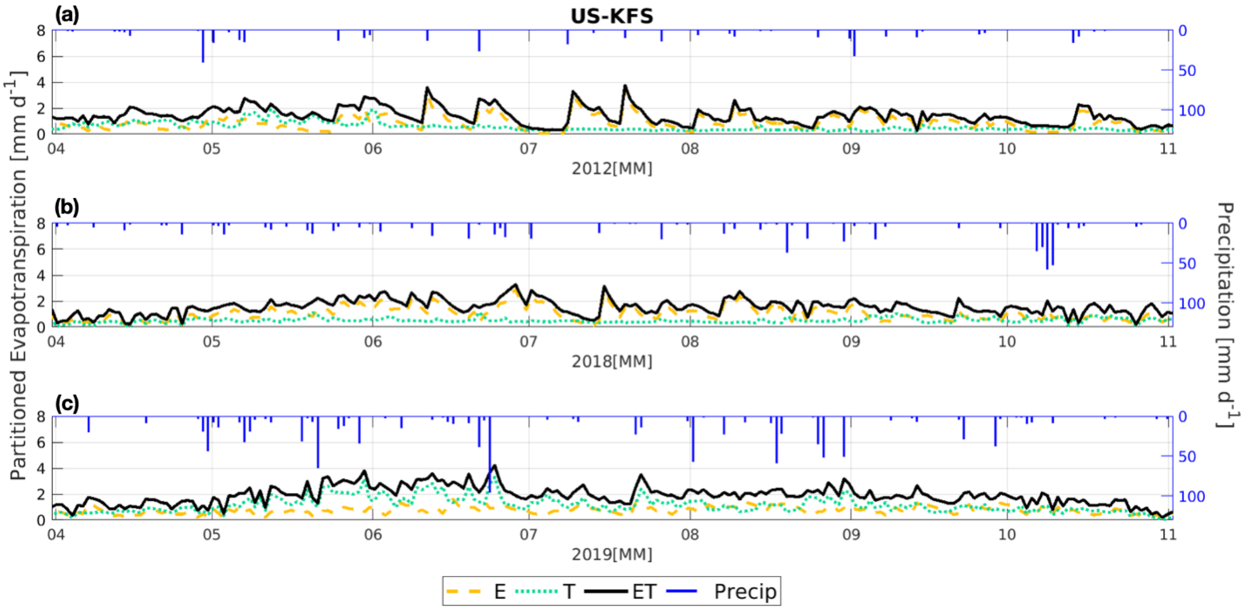

Figure 10 (updated). Time series of evapotranspiration, ET, at US-KFS partitioned into evaporation and transpiration for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. The curves represent ensemble means from the DCHM-PV 3YR.

Note the substantial underestimation of ET by the models relative to Ameriflux in 2012 should be noted here as well along with some discussion of why

This comment and suggestion from the review is well taken. Please see response to Major Comment 3a above.

As noted above, modeled ET should only consider the daytime values when computing daily averages because the model shuts down evaporation at nighttime when there is no incoming solar radiation. However, in the original manuscript we computed daily averages over a 24 hour period. With the update to Figure 9, we see that the modeled results match well against AmeriFlux. Differences still occur in 2012, with modeled ET agreeing with AmeriFlux starting in April through mid-May. Once the flash drought onsets (late May through July), modeled results tend to be lower than AmeriFlux. Once the flash drought ends in August, modeled ET once again agrees with AmeriFlux. In a drought and non-drought year, modeled ET appears to match better throughout most of the season (Figure 9). One explanation could be that water use by vegetation during flash drought is considerably different, and the model is not able to recreate this change in survival strategy.

In the results section which will be updated to be section 3.2.2 ET, we intend to add some of the above explanation to the first paragraph. We will also remove some of the comparisons of results that involve the DCHM-PV WET and DRY since we have agreed that analysis should focus on 3YR. We intend to cut the current line 332-337 and replace it with the following.

Line 332: …beginning of the growing season (April). **Modeled ET results match well against AmeriFlux at US-KFS during flash-drought, and non-flash drought periods (Figure 9). Differences still occur in 2012 at US-KFS, with modeled ET agreeing with AmeriFlux starting in April through mid-May. Once the flash drought onsets (late May through July), modeled results tend to be lower than AmeriFlux. Once the flash drought ends in August, modeled ET once again agrees with AmeriFlux. In a drought and non-drought year, modeled ET appears to match better throughout most of the season (Figure 9). While model estimates of ET are higher than flux tower measurements in non-drought at US-KLS, they match favorably in drought and flash drought (Figure A13). In contrast to model and flux tower comparisons at US-KFS and US-KLS, at US-Kon modeled ET (Figure A14) performs best in 2019 (non-drought) compared to 2012 (flash drought) and 2018 (drought). One explanation for the varied differences between model and tower ET data could be that water use by vegetation during flash drought is considerably different, and the model is not able to recreate this change in survival strategy. Depending on plant function type and soil characteristics, it may be difficult for the DCHM, or other models, to account for plant access to deep water stores (Giardina et al, 2023).**

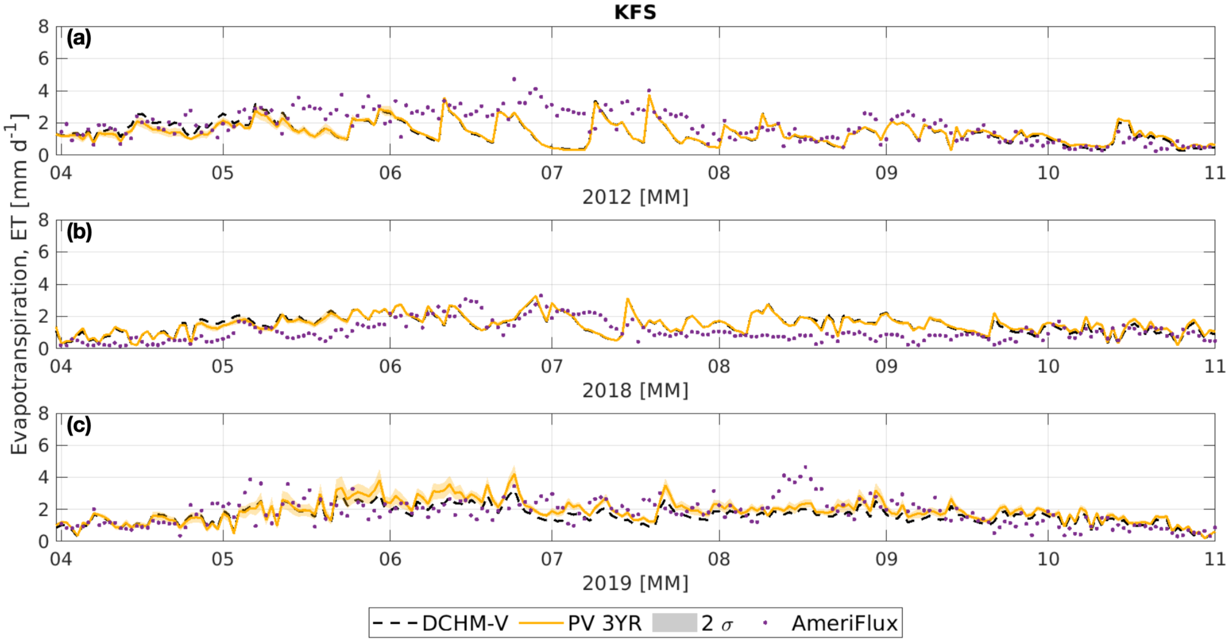

Figure 9 (updated). Time series of evapotranspiration, ET, at US-KFS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. Two standard deviations are shown for the DCHM-PV simulations. AmeriFlux ET is derived from latent heat measurements and shown as small dots.

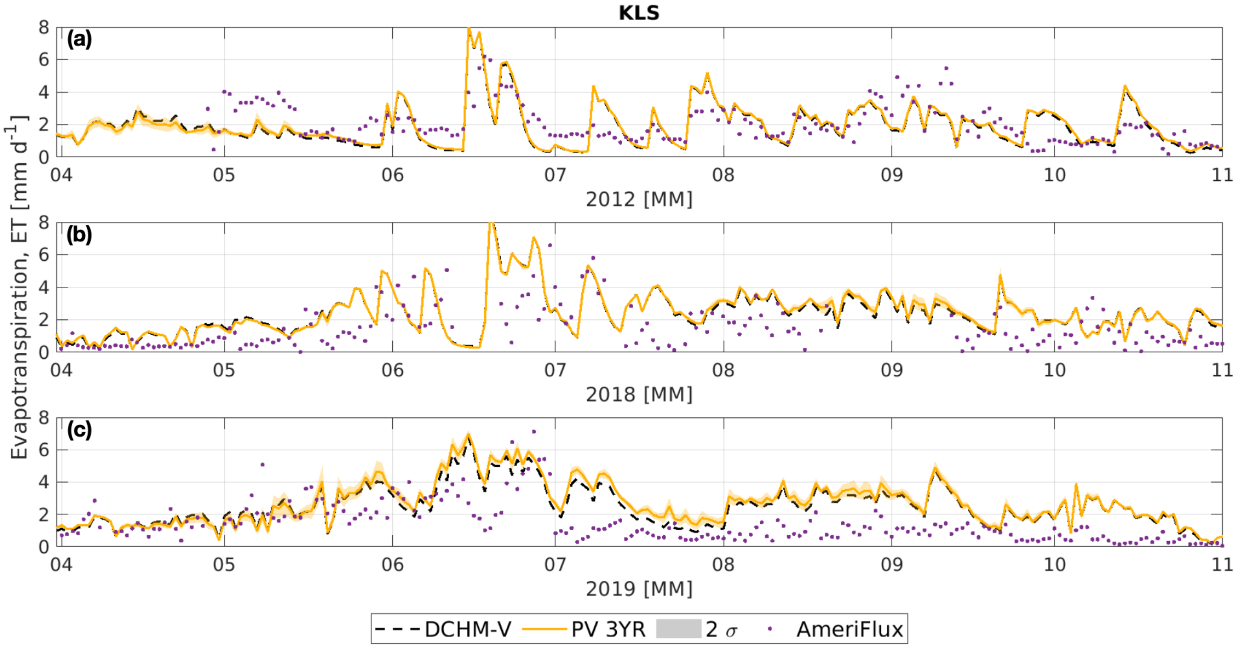

Figure A13 (updated). Time series of evapotranspiration, ET, at US-KLS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. Two standard deviations are shown for the DCHM-PV simulations. AmeriFlux ET is derived from latent heat measurements and shown as small dots.

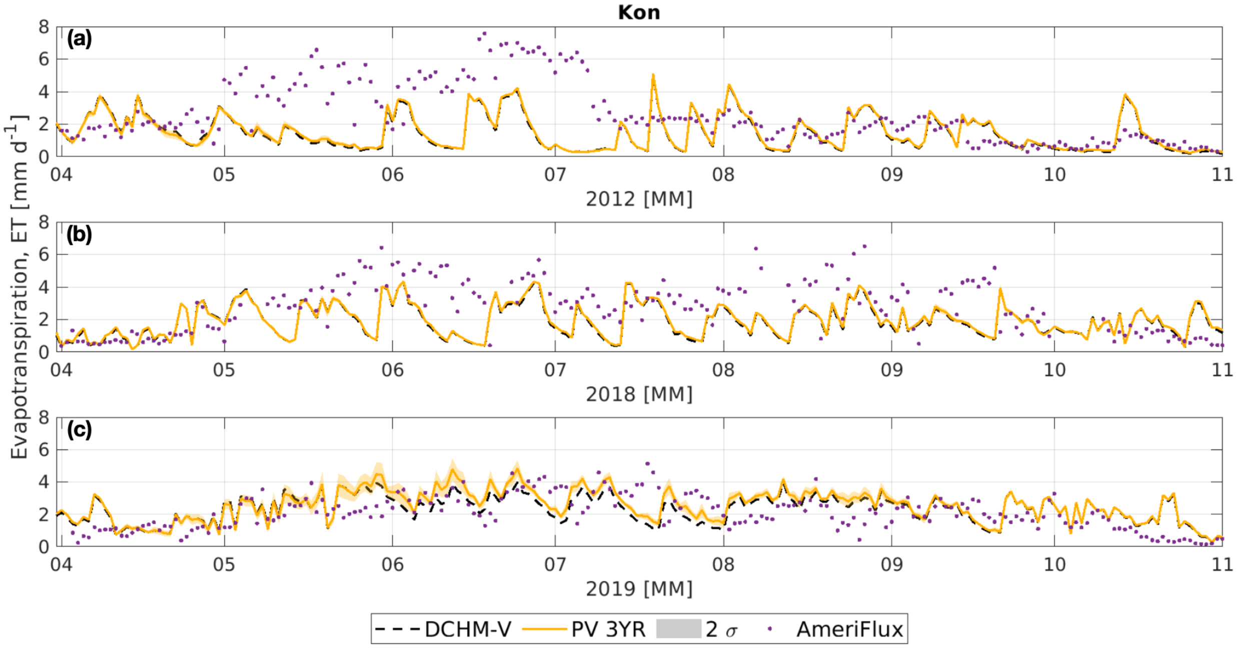

Figure A14 (updated). Time series of evapotranspiration, ET, at US-Kon for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. Two standard deviations are shown for the DCHM-PV simulations. AmeriFlux ET is derived from latent heat measurements and shown as small dots.

line 360 - Its worth noting that the drought response can be more complicated than simply shutting stomata - importantly grasses can shift their allocation of carbon - and this will be reflected in above ground biomass (the GPP measured by MODIS) but also in below ground stores and fluxes - For example see Ingrisch, Johannes, Stefan Karlowsky, Roland Hasibeder, Gerd Gleixner, and Michael Bahn. "Drought and recovery effects on belowground respiration dynamics and the partitioning of recent carbon in managed and abandoned grassland." Global Change Biology 26, no. 8 (2020): 4366-4378.

This helpful comment from the reviewer is well taken. Stomata closure implies less gas exchange. Ultimately, this drives the decreases in modeled losses to ET and GPP, but could explain the higher GPP observed in the AmeriFlux data. Vegetation, at least grasses, can reallocate already processed carbon to their roots when under drought stress mitigating GPP losses (Ingrisch, et al., 2020). This means that MODIS may see a reduction in phenological states (LAI and FPAR) but maybe GPP is less affected.

Line 360: …ET (Chen et al., 2019). **In some cases, vegetation can reallocate already processed carbon to their roots when under drought stress mitigating GPP losses (Ingrisch, et al., 2020). However, modeled GPP losses are likely a result of modeled stomatal behavior, as the model does not account for reallocation of carbon stores within the plants. This limitation of the model could explain why AmeriFlux GPP tends to be higher than the modeled GPP. Sub-daily scale stomatal conductance reduces to zero in response to increased VPD (Figure N9) leading to similar reductions in modeled GPP (Figure N11).**

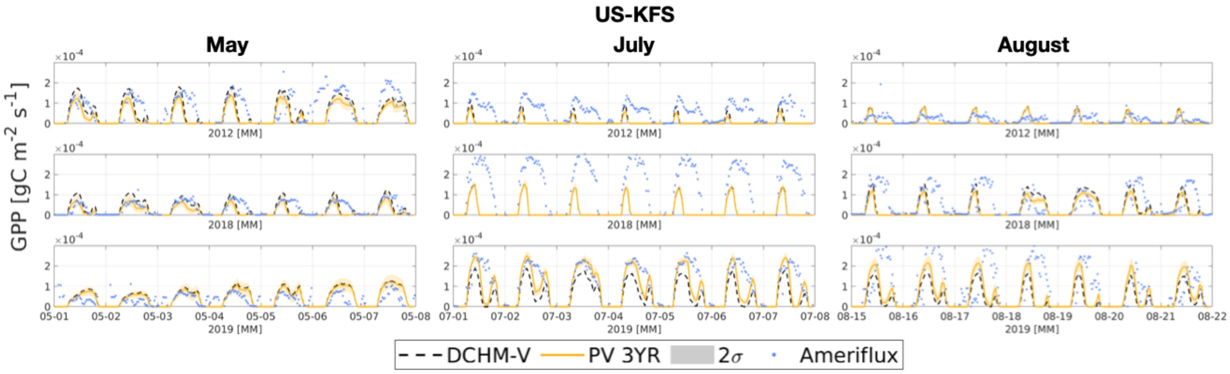

Figure N11. Hourly gross primary productivity [g C m^{-2} s^{-1}] from the DCHM-V and DCHM-PV shown against AmeriFlux 30-minute estimates for one week in May, July, and August of 2012, 2018, and 2019 atUS-KFS.

Similarly on line 380 - there is ample evidence of changing root allocation (and root respiration) for grasslands that would be worth citing here.

The reviewer's comment is well-taken. We now discuss changing root allocation and respiration with relevant citations. See edited text below.

Line 380: …leading to differences in hydraulic tendencies under variable water regimes **and atmospheric conditions which distinguish vegetation that is more likely to survive or recover from drought (McDowell et al., 2008; Martinez-Vilalta et al., 2002).**

Line 384 …role in root water uptake. **Moreover, models that can account for different vegetation behavior such as the reallocation of carbon storage and below ground respiration during drought may provide a better understanding of mechanisms driving drought resiliency and changes to carbon uptake during drought (Ingrisch et al., 2020; Sanaullah et al., 2012). These types of mechanisms could explain why results of studies like Wolf et al., (2016) who found that a warm and wet spring mitigated the effects of the 2012 flash drought on GPP losses.**

line 369 - Note that authors don't really show that evaporation after rain effects uses all available water - so it doesn't infiltrate (as stated in the hypothesis) - but they could do this at least with the model since both daily evaporation and precipitation is available.

We appreciate this comment and it is part of a larger theme addressed in the response to Major Comment 2. Note from Figure N3, precipitation exceeds infiltration. Excess amounts of water not contributing to infiltration can be attributed to runoff.

Line 365: layer…**Monthly infiltration (Figure N1) and evaporation accumulations (Figure 7 b,d,f) for June and July during the flash drought at all three sites are similar. However, in drought and non-drought periods, infiltration far exceeds evaporation. This implies that even as water is infiltrating the soils, comparable amounts are evaporating contributing to the soil drying down, leaving less new water available for plant use. In the time series of GPP during June and July,** at the onset of flash drought…

Line 367 …evaporation shuts down. **Despite evaporation tapering to zero during June and July of 2012 (Figure 10), pulses of rainfall lead to temporary rapid increases in rates of evaporation before falling back down. In May of 2012 at US-KFS there was 70 mm of water infiltrating the soils (Figure N1) with 35 mm of evaporation (Figure N10b). But in June and July total infiltration was 61 mm with 65 mm of evaporation over the two months. Similar comparisons can be found at US-KLS and US-Kon (Figures N1 and N10). In contrast, at US-KFS, during non-drought years, June averages of infiltration are in excess of 100 mm with 41 mm of evaporation. Average drought years have 66 mm of infiltration with 47 mm of evaporation (Figure N1a). Infiltration usually exceeds evaporation in the growing season, but infiltration accumulations comparable to evaporation totals appears to be an indication of flash drought.**

line 375 - “since GPP is decreasing” as the authors themselves note - declines in GPP do not always reflect what’s happening with transpiration so this statement needs some caveats

The point is well taken, especially with grasses as noted above, reallocating their stores of carbon which can be another reason why modeled GPP isn’t solely tied to changes in transpiration. However, the DCHM strongly links transpiration to GPP and does not account for reallocating carbon storages in other ways. We will use a new paragraph at Line 367-370 that leads into the current paragraph at line 371.

Line 367: **Despite major reductions in infiltration and fluctuations in top layer soil moisture during the peak of flash drought onset, modeled root water uptake indicated that plants were still pulling small amounts of water in through their roots, preventing plants from completely shutting down. With the ability to tap into water stores from deeper layers (Giardina et al., 2023) and small rates of transpiration still occurring, modeled carbon uptake is still maintained (Figure 8a, and A11a, A12a). Although it drastically slows, it does not stop.**

In the discussion, one of the challenges here is that observations and models suggest differences in plant ability to pull water from deeper layers - The first paragraph blurs these distinctions - so for example the line “we did find that even during the peak..plants were still” - is this based on flux towers, or model?. I also note that model and observed in Figure 11 suggest different stories about water use efficiency during the drought. The observed data suggests plants maintain higher water use efficiency, longer, during the drought than the model - this is very interesting - its suggests plants are doing something that the model misses which is informative - but needs to be much more of a focus in the paper.

The reviewer’s comments are well taken.

We now include language in the paragraph above (previous response, Line 367) to indicate that we were referring to model outputs regarding root water uptake. In the larger context, we acknowledge that there may be discrepancies between models and flux tower measurements due to below ground processes that models are missing (see response to Reviewer Comment 3b above).

Figure 11 has been updated to average over daytime hours for water use efficiency. Corresponding figures for US-KLS and US-Kon are included in the appendix. Updates to figure 11 include the results from averaging WUE over the daytime and the addition of 2018. We also updated the legend and y-axis label as well as an extension of the viewing window to include all of October. We now align with tower values better in 2012. However, in 2018 the DCHM underestimation of tower WUE can be attributed to the differences in the 2018 modeled and flux tower measurements of GPP since modeled ET at US-KFS matches well against tower measurements. See the updated Figures 8 and 9 above for reference to GPP and ET, respectively.

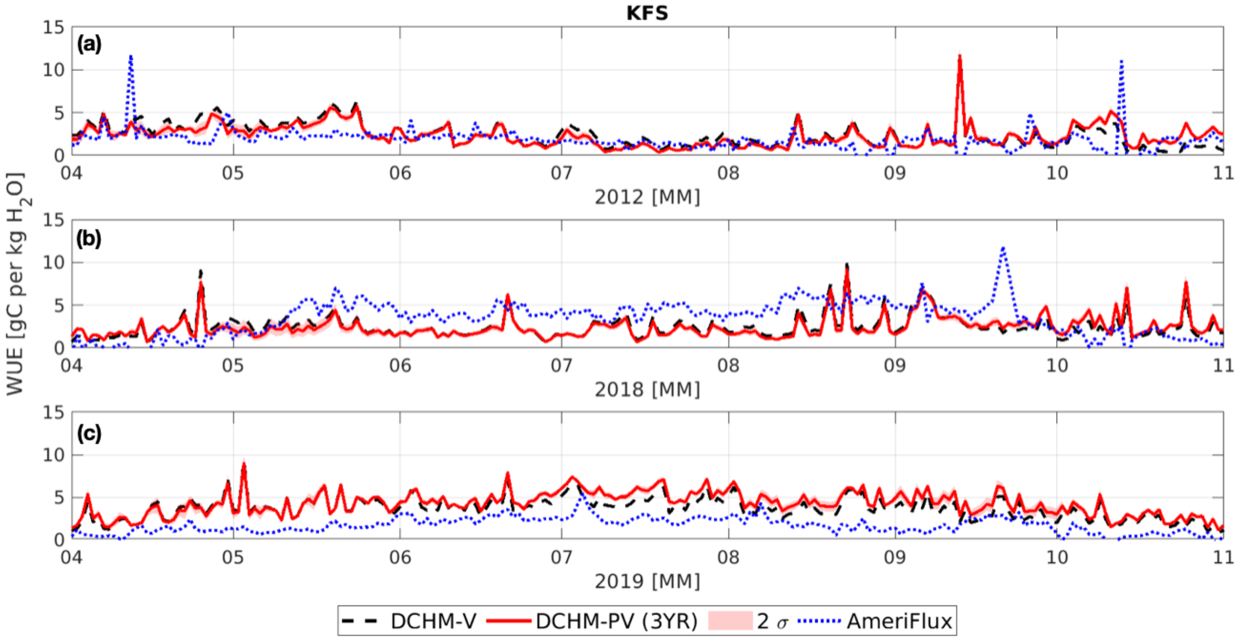

Figure 11. Computed growing season water use efficiency (WUE=GPP/ET) from the DCHM-V, DCHM-PV 3YR, and AmeriFlux for (a) 2012 (b) 2018, drought (c) 2019, non-drought at US-KFS. AmeriFlux WUE is computed by converting latent heat into eT by dividing by the coefficient of vaporization.

Figure 12 - its not so easy to see from this figure that when plant are transpiring more they are more efficient in their water use - Simply graphing transpiration vs WUE would show this much more clearly.

The reviewer's comments are well taken and appreciated. In fact, our results agree with the point that plants transpiring more are not necessarily more efficient in their water use.

Figure 12 has been updated below. The main point of the new Figure 12 is to show that transpiration as a percentage of overall ET and WUE follow the same transition from normal or above average levels to drought levels from May to July. This suggests that both follow the same pattern and can be used to mark the rapid transition to drought state.

Given the new analysis and figures of transpiration accumulation and average water use efficiency in drought and non-drought years compared with non-drought years we can better observe the trends associated with WUE and transpiration. The new Figure 12 suggests that plants are more efficient in non-drought years. There is an association with decreased T/ET and decreases WUE. However, that does not substantiate the claim that “plants that are transpiring more are more efficient.” We no longer like to make this claim.

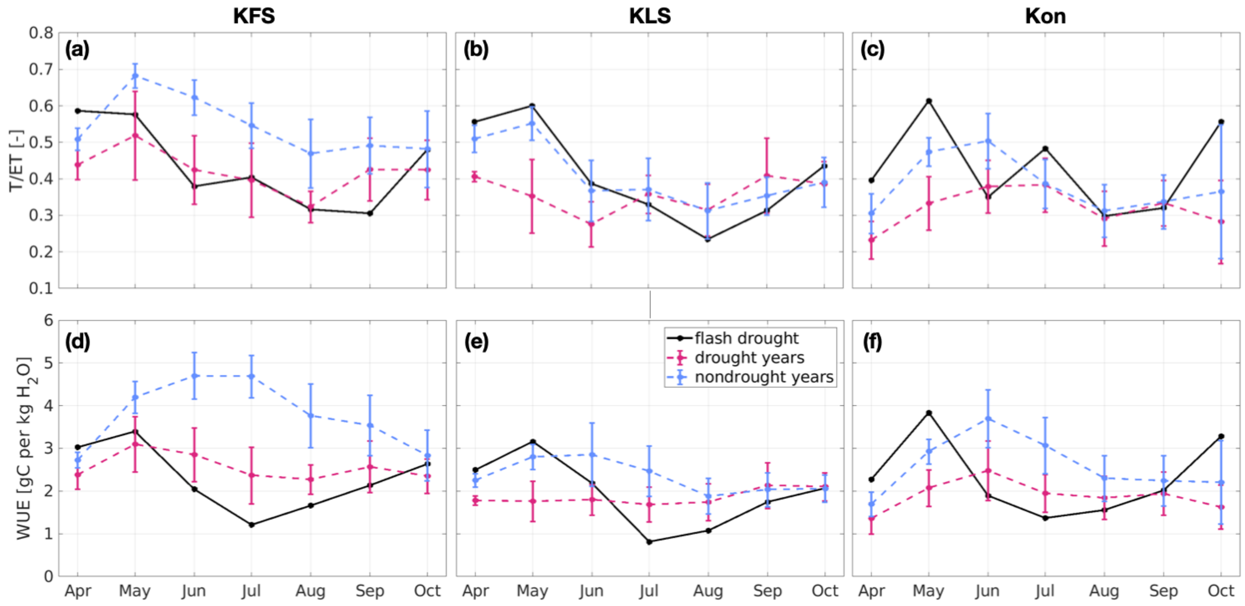

Figure 12 (replacement). Modeled growing season monthly averages of transpiration as a fraction of ET (a-c) and water use efficiency (WUE, d-f) for drought (red) and non-drought (blue) years compared with the flash drought year of 2012 (black) for US-KFS, US-KLS, and US-Kon AmeriFlux sites. Monthly averages are computed from the ensemble means of the 2000 Monte Carlo simulations then averaged across drought or non-drought years. Error bars represent one standard deviation across drought and non-drought years. Drought years are 2006, 2011, 2013, 2014, 2018 and non-drought years are 2007-2010, 2015-2017, 2019.

Figures 11 and 12 were updated. Figure 11 now includes a daily average time series at US-KFS for 2018 in addition to 2012 and 2019. Corresponding figures for US-KLS and US-Kon will be included in the appendix figures replacing A17 and A18. Figure 12 was updated to compare flash drought T/ET ration and WUE to aggregate monthly averages from drought and non-drought years. We also added October to the analysis period.

New comments for Figure 12.

A more apt comment using Figure 12 in the discussion is
Line 378: **Plants are more efficient during non-drought periods, and are especially inefficient during flash drought onset (Figure 12). Accumulated monthly averages of transpiration as a fraction of evapotranspiration show a transition from at or above non-drought levels to at or below drought levels. At US-KFS drought years show a lower fraction of transpiration throughout the growing season whereas drought and non-drought values are similar from July-October at US-KLS and US-Kon. US-Kon experiences larger fluctuations in the fraction of transpiration through the early and middle parts of the growing season (April - July). It is possible that US-Kon, modeled as a grassland, is showing an adaptation to the water stresses. Another possible explanation**

for fluctuations in T/ET is that at the same time E/ET is fluctuation in response to precipitation and increased atmospheric demand for water.

Water use efficiency is generally higher in non-drought years compared to drought, though WUE is similar across all non-flash drought years from August-October at US-KLS. WUE at all sites started off in 2012 with above average non-drought levels and an increase from April to May. However, from May-July WUE at all sites falls from above average to more than one standard deviation below drought years. With GPP differences being more substantial than ET between flash drought and non-flash drought periods (Figure 7), subseasonal reductions in WUE (WUE=GPP/ET) can be attributed to the losses in GPP. These reductions in WUE from above average drought conditions to below drought conditions (e.g., the 60%-70% reduction from May to July in 2012, Figure 12 d,e,f), could be used as an indicator of flash drought onset.

From Figure N12, we can see that changes in transpiration explain a small component, less than 10%, of the variability observed in WUE. Generally when plants have more available water (e.g., 2019), they transpire more, but higher values of WUE can be seen in flash drought and drought years (e.g. 2012, 2018) despite having lower rates of transpiration.

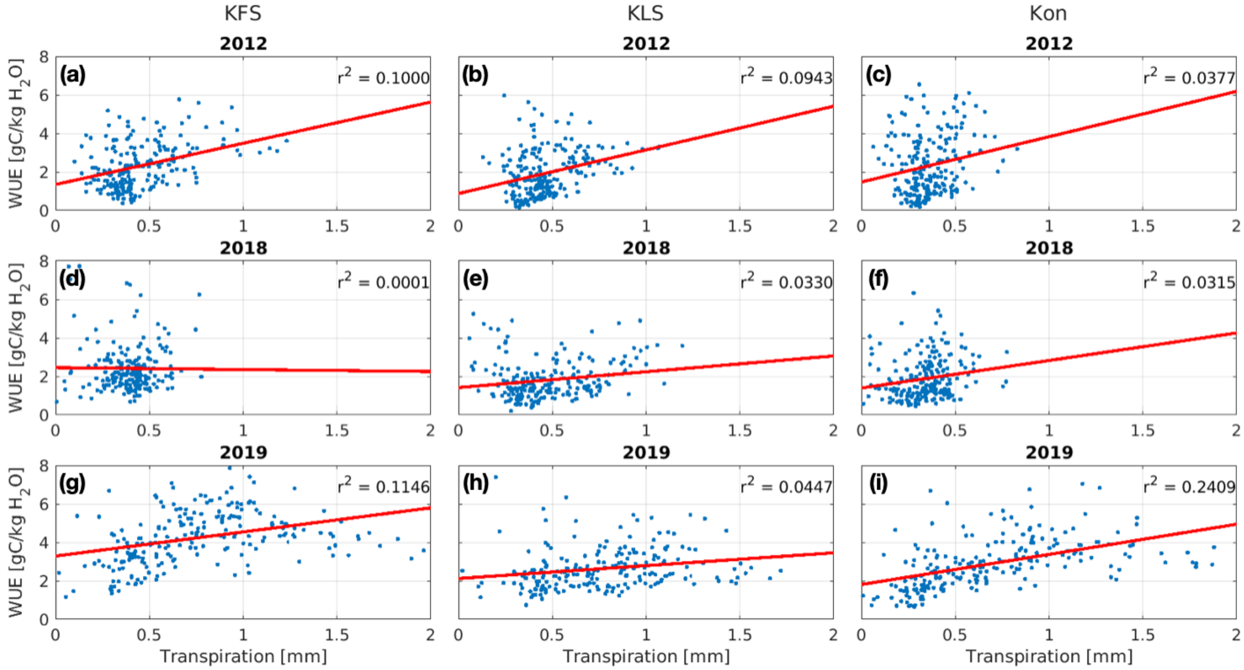

Figure N12 (appendix). Daily averages of water use efficiency versus transpiration for 2012, 2018, and 2019.

Also in figure 12 - some strategic use of color to differentiate wet versus dry years would be helpful here

We appreciate this comment.

Upon initial submission of the manuscript, the processing editor suggested making some changes and updates to make plots and figures more readable. Changing to figures include updated color schemes, font sizes, line weights, line styles, and markers. Color schemes were generated using a tool Coloring for Colorblindness https://davidmathlogic.com/colorblind/#%23D81B60-%231E88E5-%23FFC107-%23004D40 and figures were tested on Coblis - Color Blind Simulator https://www.color-blindness.com/coblis-color-blindness-simulator/.

Line 403 - The idea of flash drought responses is intriguing but I think the paper could do much more to support these ideas (e.g that the rapidness of the change is indicative of a flash drought) - Some more strategic comparison of the declines across different type of drought - flash versus "non-flash" (Of course this would require clearly distinguish what a flash drought is from other types of drought - but that seems to be part of the paper's motivation)

The authors thank the reviewer for this thoughtful comment. A major goal of the authors' responses to the reviewer comments throughout this document has been to refocus analysis to compare results (e.g., GPP, ET, WUE, etc.) from a flash drought year to drought and non-drought years (see response to Major Comment 1). It is our intent that the rapid changes seen from May to July in 2012 model results of vegetation-atmosphere interactions along with the decreased precipitation and increased atmospheric demand for water highlight the transitional period of flash drought intensification. Furthermore, we hope this emphasizes flash drought as being the rapid development of drought that can be observed through land-atmosphere interactions.

Concluding Remarks

The authors would like to thank the reviewer again for their comments, questions, and suggestions that help improve this manuscript. The authors acknowledge that the updated analysis necessitates updated language throughout the manuscript reflecting the major changes presented here.

References

1. Baldocchi, D. D. (2003). Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global change biology*, *9*(4), 479-492.
2. Chen, L. G., Gottschalck, J., Hartman, A., Miskus, D., Tinker, R., & Artusa, A. (2019). Flash drought characteristics based on US drought monitor. *Atmosphere*, *10*(9), 498.

3. Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., ... & Zona, D. (2021). Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. *Agricultural and Forest Meteorology*, *301*, 108350.
4. Giardina, F., Gentine, P., Konings, A. G., Seneviratne, S. I., & Stocker, B. D. (2023). Diagnosing evapotranspiration responses to water deficit across biomes using deep learning. *New Phytologist*.
5. Guo, J. S., Hultine, K. R., Koch, G. W., Kropp, H., & Ogle, K. (2020). Temporal shifts in iso/anisohydry revealed from daily observations of plant water potential in a dominant desert shrub. *New Phytologist*, *225*(2), 713-726.
6. Hosseini, A., Mocko, D. M., Brunsell, N. A., Kumar, S. V., Mahanama, S., Arsenault, K., & Roundy, J. K. (2022). Understanding the impact of vegetation dynamics on the water cycle in the Noah-MP model. *Frontiers in Water*, *4*, 925852.
7. Ingrisch, J., Karlowsky, S., Hasibeder, R., Gleixner, G., & Bahn, M. (2020). Drought and recovery effects on belowground respiration dynamics and the partitioning of recent carbon in managed and abandoned grassland. *Global Change Biology*, *26*(8), 4366-4378.
8. Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., & Schulze, E. D. (1996). A global analysis of root distributions for terrestrial biomes. *Oecologia*, *108*, 389-411.
9. Lowman, L. E., & Barros, A. P. (2016). Interplay of drought and tropical cyclone activity in SE US gross primary productivity. *Journal of Geophysical Research: Biogeosciences*, *121*(6), 1540-1567.
10. Lowman, L. E., & Barros, A. P. (2018). Predicting canopy biophysical properties and sensitivity of plant carbon uptake to water limitations with a coupled eco-hydrological framework. *Ecological Modelling*, *372*, 33-52.
11. Lowman, L. E., Wei, T. M., & Barros, A. P. (2018). Rainfall variability, wetland persistence, and water–carbon cycle coupling in the Upper Zambezi river basin in Southern Africa. *Remote Sensing*, *10*(5), 692.
12. Martínez-Vilalta, J., Piñol, J., & Beven, K. (2002). A hydraulic model to predict drought-induced mortality in woody plants: an application to climate change in the Mediterranean. *Ecological Modelling*, *155*(2-3), 127-147
13. McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., ... & Yepez, E. A. (2008). Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought?. *New phytologist*, *178*(4), 719-739.
14. Myneni, R., Knyazikhin, Y., Park, T. (2015). MOD15A2H MODIS Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC.
15. Otkin, J. A., Svoboda, M., Hunt, E. D., Ford, T. W., Anderson, M. C., Hain, C., & Basara, J. B. (2018). Flash droughts: A review and assessment of the challenges imposed by rapid-onset droughts in the United States. *Bulletin of the American Meteorological Society*, *99*(5), 911-919.
16. Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., ... & Law, B. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific data*, *7*(1), 1-27.

17. Running, S. W., & Zhao, M. (2015). Daily GPP and annual NPP (MOD17A2/A3) products NASA Earth Observing System MODIS land algorithm. *MOD17 User's Guide*, *2015*, 1-28.
18. Sanaullah, M., Chabbi, A., Rumpel, C., & Kuzyakov, Y. (2012). Carbon allocation in grassland communities under drought stress followed by 14C pulse labeling. *Soil Biology and Biochemistry*, *55*, 132-139.
19. Schmid, H. P. (1994). Source areas for scalars and scalar fluxes. *Boundary-Layer Meteorology*, *67*(3), 293-318.
20. Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., ... & Stephens, S. (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181-1190.
21. Thornthwaite, C. W., Mather J.R. (1957). Instructions and tables for computing potential evapotranspiration and the water balance. *Publications on Climatology*, *10*, 185-310.
22. Wolf, S., Keenan, T. F., Fisher, J. B., Baldocchi, D. D., Desai, A. R., Richardson, A. D., ... & Van Der Laan-Luijkx, I. T. (2016). Warm spring reduced carbon cycle impact of the 2012 US summer drought. *Proceedings of the National Academy of Sciences*, *113*(21), 5880-5885.
23. Yan, K., Park, T., Yan, G., Chen, C., Yang, B., Liu, Z., ... & Myneni, R. B. (2016). Evaluation of MODIS LAI/FPAR product collection 6. Part 1: Consistency and improvements. *Remote Sensing*, *8*(5), 359.
24. Zeng, X. (2001). Global vegetation root distribution for land modeling. *Journal of Hydrometeorology*, *2*(5), 525-530