Responses to Anonymous Reviewer 2 for paper: *Unraveling phenological to extreme drought and implications for water and carbon budgets*

The authors greatly appreciate the reviewer taking time to provide feedback on our manuscript. We think that the comments are helpful in shaping the new analysis and clarifying how important mechanisms driving plant controls of water and carbon movements may be affected during flash drought.

In light of new analysis described in our responses below, we propose a new title: ***Unraveling phenological and stomatal responses to flash drought and implications for water and carbon budgets***

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**. Responses in *italics* are used to reference exact quotes from the author's responses to Reviewer Comment 1 (RC1). New figures that we propose to include have labels N# throughout this document and the number corresponds to how the new figures were introduced starting with RC1. Revised figures have the same figure number as in the original manuscript. 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.

This study investigates vegetation phenological processes during drought and non-drought periods using a set of DCHM-V and DCHM-PV model simulations. The focus is on three sites in Kansas USA, as they experienced extreme drought and pluvial conditions in recent decades, and also have ground-based and satellite-based observations available to compare with model simulations and study observed processes. The modeling experiments are neatly designed, and the investigation is systematic to study sources of uncertainty in vegetation phenology. I however have a number of comments that need to be addressed – please see below.

Main comments

1)   This study uses 2012 and 2019 to exemplify the contrast of vegetation phenology between flash drought vs non-drought years. It would help to add a non-flash/conventional drought year to the study, as the vegetation phenology could differ considerably between flash drought and conventional drought. It would be interesting to see how the evolution of plant phenology, water use, and productivity may differ between the two drought cases.

The authors appreciate this comment and agree with the reviewer that we should highlight the differences between flash drought and "non-flash/conventional drought". To this end, we propose to add 2018 (conventional drought) as a specific case study to compare against 2012 (flash drought) and 2019 (non-drought). Additionally, we focus components of the analysis to consider average conditions across conventional "drought" and "non-drought" over the course of

the period 2006-2019. This major change is described in our response to Major Comment #1 from Reviewer #1.

From response to Review Comment 1 (RC1): *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.*
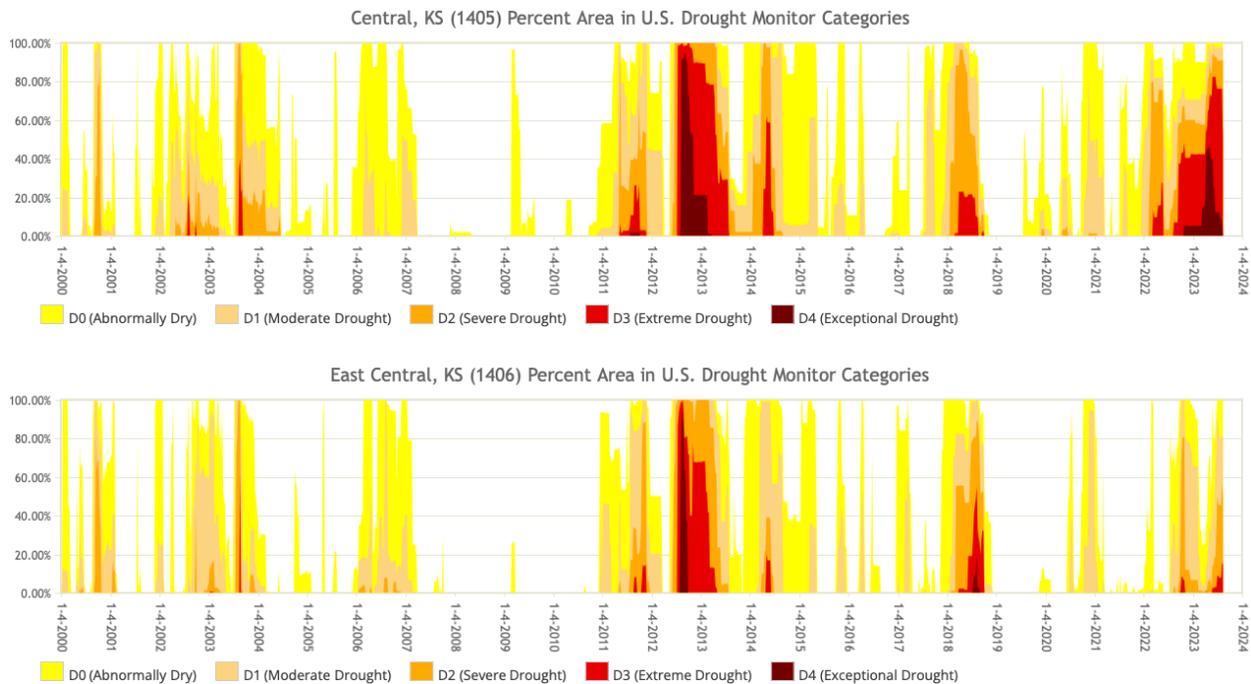

Figure Caption. Percent land area in U.S. Drought Monitor Categories for two Kansas climate regions that contain our study sites US-KFS, US-KLS, US-Kon.

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."

An example of how we compare flash drought to non-flash drought can be seen with updated Figure 7. 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.
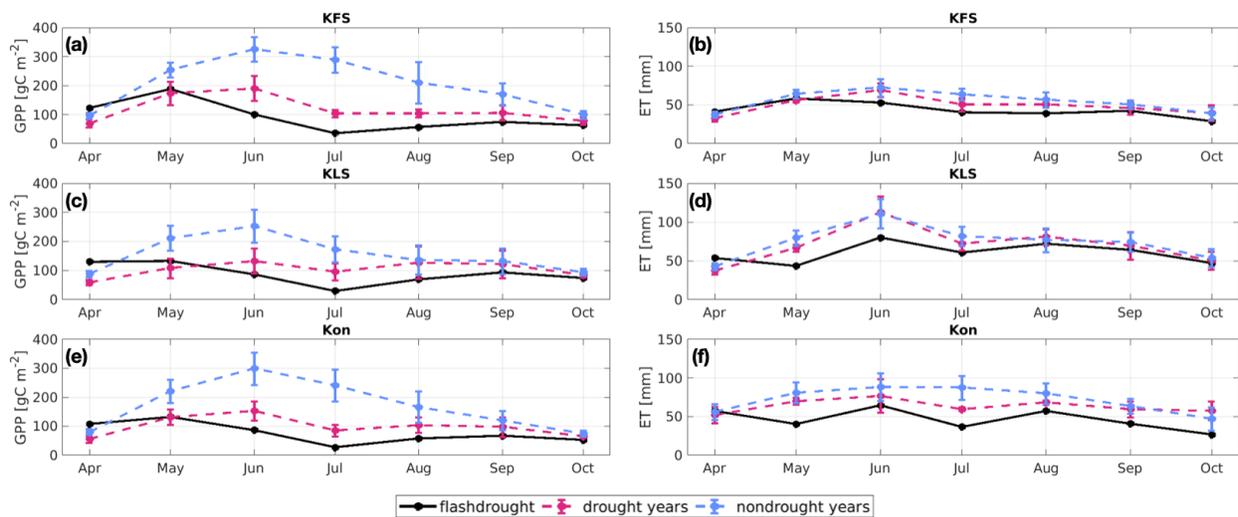

Figure 7 (replacement). DCHM-PV 3YR monthly totals of GPP (a,c,e) and ET (b,d,f) for drought and non-drought years compared to 2012 for US-KFS, US-KLS, and US-Kon AmeriFlux sites. Monthly sums are computed from the ensemble means of the 2000 Monte Carlo simulations then averaged across drought or non-drought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

An example of how we integrate 2018, a non-flash drought, to compare against 2012 and 2019 can be seen in the updated Figure 11.
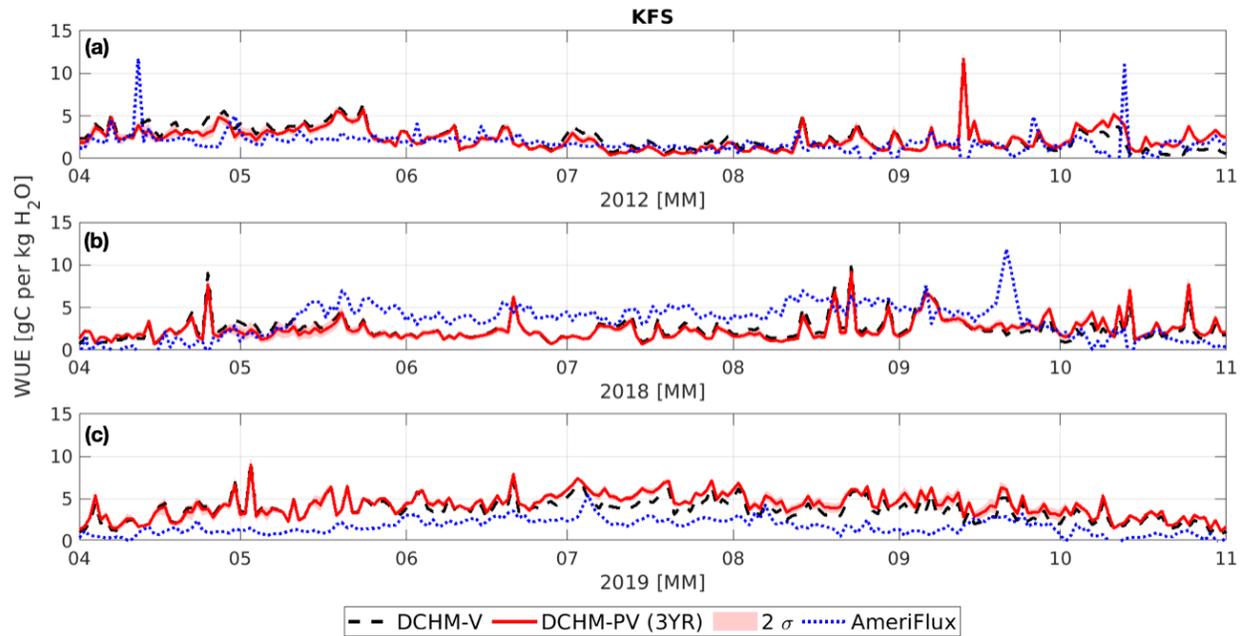
Figure 11. Computed growing season water use efficiency (WUE=GPP/ET) from the DCHM-V, DCHM-PV 3YR, and AmeriFlux for (a) 2012 (b) 2018, drought (c) 2019, non-drought at US-KFS. AmeriFlux WUE is computed by converting latent heat into eT by dividing by the coefficient of vaporization.

2)   Much of the findings are based on the DCHM-V and DCHM-PV simulations, and are thus subject to the performance of the DCHM and its predictive phenology in simulating observed land surface and vegetation processes. The comparison between the model results and independent observations (e.g., MODIS, AmeriFlux) however shows considerable differences: some of the models vs. observations differences are so substantial that they are much larger than the differences between different model experiments (e.g., Figs.5-9). While these differences could be in part due to the data comparison across different spatial and temporal scales (Section 4.5), they also make one wonder about the performance of the DCHM and its predictive phenology. I suggest the study provides more information/results on the fidelity of the DCHM and its predictive component is simulating basic land surface variables (e.g., soil moisture, evapotranspiration) and observed vegetation phonology related fields (e.g., LAI, FPAR), e.g., in terms of climatology, seasonal cycle and year-to-year variations, to make the model-based findings more convincing. Also see some of my detailed comments below.

The author's appreciate this reviewer's comment and we agree that more discussion is needed involving the comparison of the results from the DCHM and measurements and observations from AmeriFlux and MODIS.

Major Comment 3 from our responses to RC1, who shared similar feedback, addressed this point.

From our responses to RC1:

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.

We proposed the following text to be added to the Discussion section of the manuscript in RC1 to discuss the discrepancies.

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 the NLDAS-2 and Hosseini et al., 2022, who used Noah-MP 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}, which is below wilting point, 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 that LSMs cannot reproduce could explain these differences (Giardina et al. 2023).

Hosseini et al. (2022) compared predicted estimates of GPP to flux tower measurements at US-KFS and US-KON using predictive phenology with Noah-MP, which accounts for carbon reallocation to leaves, stems, roots, and soils. Even while accounting for carbon movement, they found that during June, July, and August they underestimated tower carbon uptake 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 model employed by Hosseini et al. (2022), suggesting the accounting for carbon allocation cannot explain underestimating GPP in 2012 and 2018. The DCHM-PV compares more favorably against AmeriFlux data during 2012, the flash drought year, at US-KFS and US-KLS as opposed to 2018, a drought year (Figure 8, A11). This suggests that there are missing processes in both the DCHM and the Noah-MP that

cannot capture plant water use during drought, and cannot be attributed to carbon allocation.

During drought and flash drought, the DCHM-V and DCHM-PV tend to follow similar trajectories. However, when not under water stress, the predictive phenology model predicts higher carbon uptake than the DCHM-V, which aligns more with MODIS in 2019. AmeriFlux estimates of GPP during June and early July of 2012 and 2018 are also above estimates from MODIS. GPP estimates from the flux tower are higher than the DCHM and MODIS, suggesting that plants are able to maintain higher levels of GPP than what can be recreated in land surface models and satellite remote sensing during drought and flash drought. Differences in DCHM-PV and AmeriFlux GPP cannot be attributed to carbon reallocation since the Noah-MP model accounts for carbon reallocation and similarly underestimated GPP compared to flux tower data (Hosseini et al. 2022). A likely 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).

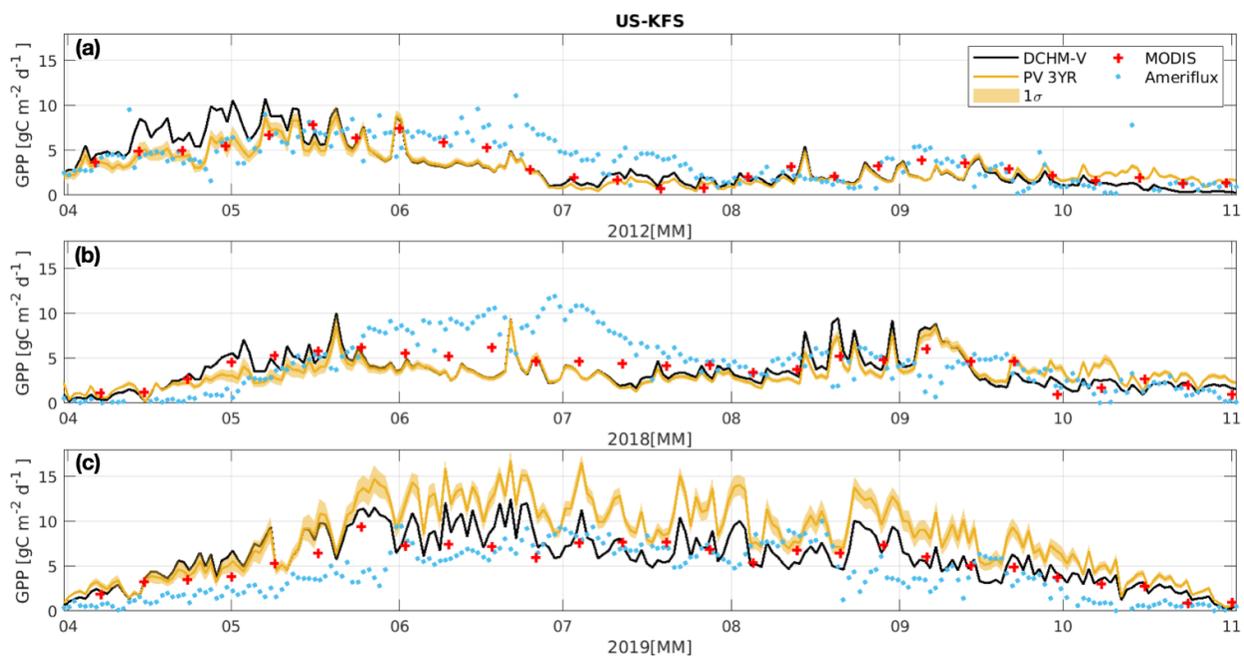Updated Figure 8 includes averaging DCHM only over daytime and the inclusion of 2018.

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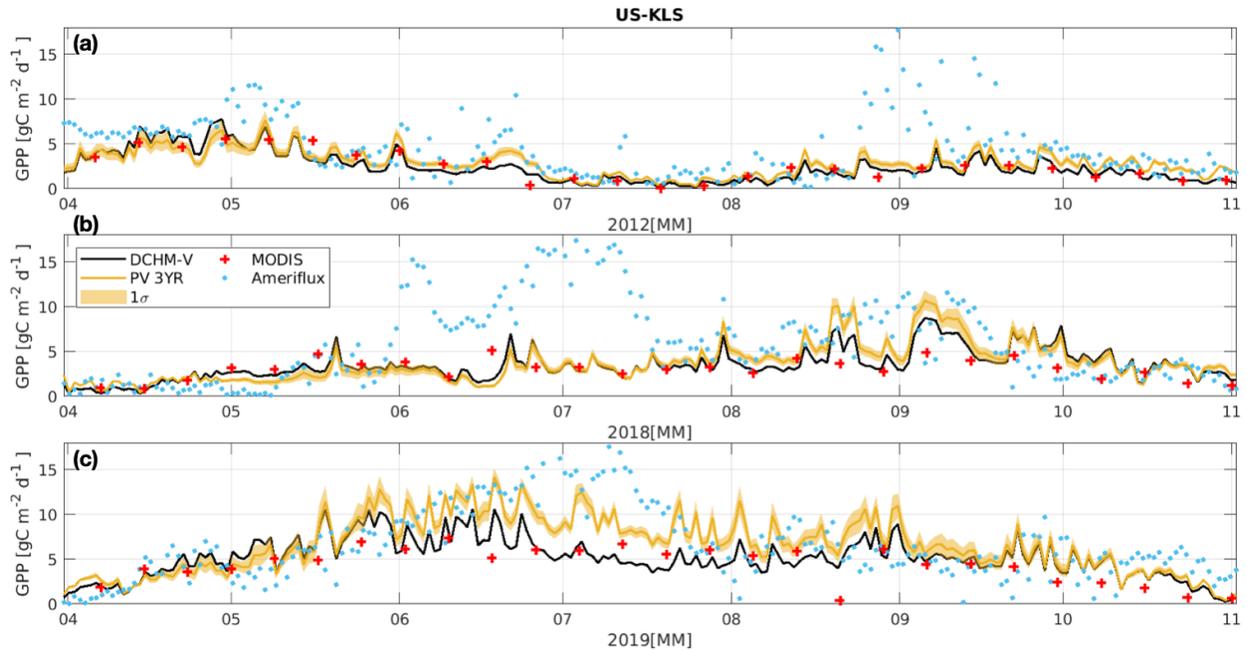
Figure A11 (updated). Time series of gross primary productivity, GPP, at US-KLS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year. One standard deviation is shown for the DCHM-PV simulations. MODIS GPP are shown as red crosses and AmeriFlux GPP as small dots.

As further explanation, 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 footprints cover areas with length dimensions ranging from a couple hundred meters to a few kilometers (Baldocchi, 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). In addition to savanna, there are deciduous forests within this gridcell that could influence tower readings, and that the DCHM does not account for.
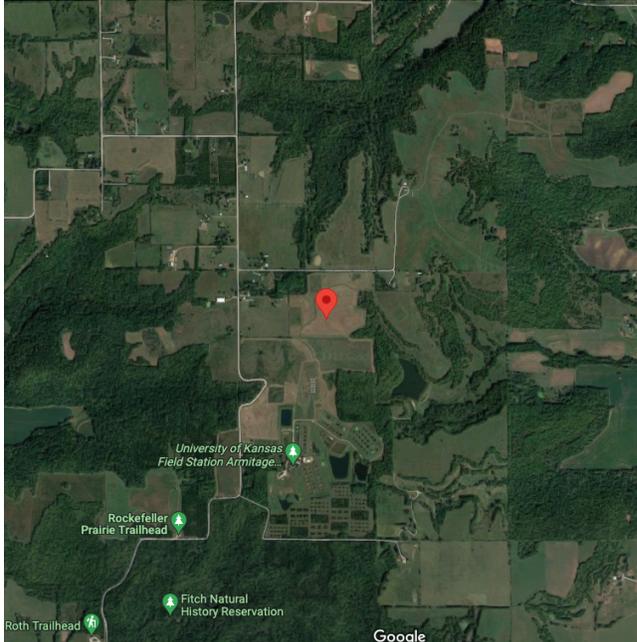
Figure (non included in manuscript). US-KFS AmeriFlux tower site at the center of a 4 km by 4 km grid representing vegetation heterogeneity of the surrounding region.

We will also add to the discussion 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 are less likely to be attributed to carbon reallocation since the model used by Hosseini et al. (2022) accounted for carbon reallocation and still underestimated AmeriFlux.

Detailed comments

1) It would help to briefly discuss the implications of the findings (e.g., based on WET vs DRY vs 3YR) to subseasonal prediction of vegetation.

Our original intention was to simulate different plant isohydric and anisohydric tendencies following Lowman and Barros (2018) who showed that the data assimilation period can be used to generate phenology model parameters that represent different water use strategies. Following this logic led us to test parameters using WET, DRY, or mixed conditions (3YR) to simulate anisohydric vs isohydric tendencies among the different plants. Our results show that the data assimilation period may not be the only factor to consider when trying to simulate water use strategies. The DCHM predicts stomatal conductance depending on vapor pressure deficit (VPD), light exposure, and soil moisture. High temperatures and low relative humidity lead to increases in VPD. In the model, high VPD leads to very low (or zero) stomatal conductance. (Figure N9). With little water available and high VPD, the DCHM-V and the DCHM-PV follow very closely. The DCHM-PV predicts higher stomatal conductance than the DCHM-V when ample water is available and there are lower values of VPD (Figure N9). This translates to higher GPP predictions in non-drought years (Figure 8).

We suggest the following updates to the discussion section 4.1 Vegetation Responses to Flash Drought.

While phenology is an important component to consider when computing changes to transpiration and carbon uptake (Lowman and Barros, 2018; Flack-Prain et al., 2019), our results indicate that stomatal conductance is also critical for accurately representing

these fluxes. Plants adaptively regulate their stomata during periods of water stress (Guo et al., 2020), and some have been demonstrated to maintain open stomata or even increase stomatal conductance under high VPD conditions (Urban et al., 2017). Stomatal conductance shuts down under high VPD in the DCHM (Figure N9), which does not account for the possibility of an adaptive stomatal regulation strategy. Since GPP is directly dependent on stomatal conductance (Farquhar and Sharkey, 1982), DCHM estimates of sub-daily GPP decrease in response to elevated VPD (Figure N11). Moreover, changes in phenological growth state (i.e. LAI) occur across longer (i.e. seasonal) time scales (Katul et al., 2001) than stomatal regulation, which controls carbon and water exchange at sub-daily timescales (Guo et al., 2020). The differences between modeled and observed GPP and ET suggest that there are mechanisms controlling plant responses to drought stress not accounted for within the DCHM. For example, the DCHM could be too strict in representing the sensitivity of stomatal closure to elevated VPD for the Kansas study sites. There could be plant or climate specific VPD dependence (Grossiord et al., 2020), plants could have access to stores of water not accounted for (Giardina et al., 2023), or both.

Guo et al. (2020) showed that isohydricity (i.e. stomatal regulation) exists on a spectrum and that some plants are able to move along that spectrum at sub-daily time-scales with varying environmental conditions, such as higher VPD. Given the high VPD in 2012 at our test sites (Figures N4, N13), we expect the DCHM to estimate low stomatal conductance, and thus low GPP relative to AmeriFlux observations when under atmospheric water stress. We also highlight that the VPD estimated by the DCHM using the NLDAS-2 Forcing File A atmospheric variables is higher during 2012 and 2018 and lower in 2019 compared to AmeriFlux (Figure N13), explaining in part the discrepancies between model and AmeriFlux GPP. As stomatal response to increasing VPD is more complex than how it is represented in LSMs, we agree with Grossiord et al. (2020) who suggest that future modeling studies should focus on how rising VPD drives stomatal closure across different plant functional types.

Daily GPP from the DCHM tends to match the magnitude of AmeriFlux daily GPP at US-KFS in 2012 (flash drought) throughout much of the growing season while greatly underestimating June and July observations in 2018 (drought). The larger discrepancies are also apparent in hourly estimates of GPP (Figure N13). The DCHM halts midday GPP in July 2018, but AmeriFlux values remain high. The differences are smaller in 2012, where AmeriFlux observed carbon assimilation rates of 1 gC m^{-2} s^{-1} throughout the daytime and the DCHM shut down carbon assimilation due to elevated VPD. This again points to the ability for vegetation to access water in ways that current LSMs cannot account for (Giardina et al. 2023). If plants have access to deeper water or are able to tap into stores of water not currently accounted for, they may be able continue (at least temporarily) exchanging water and carbon despite lower precipitation or increased VPD. As stomata control the movement of water and carbon, affecting GPP and water use efficiency (Lawson and Vialet-Chabrand, 2019), accounting for plant adaptations that adaptively regulate stomatal sensitivity to drought stress may improve model accuracy.
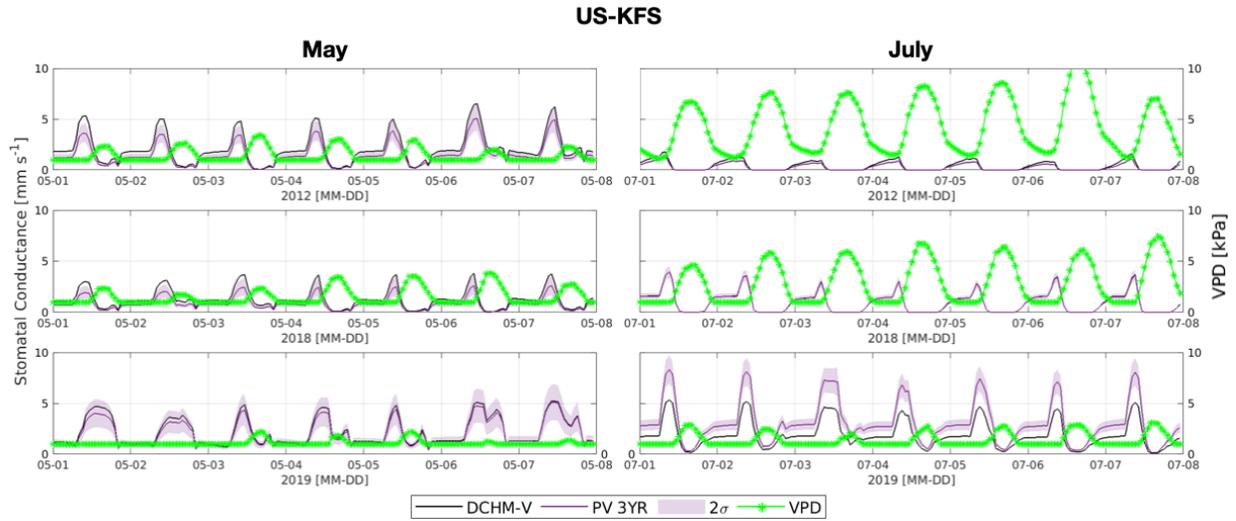
Figure N9 (Adapted). Hourly stomatal conductances [mm s^{-1}] for one week in May, and July of 2012, 2018, and 2019 compared with vapor pressure deficit (VPD, kPa) for US-KFS.
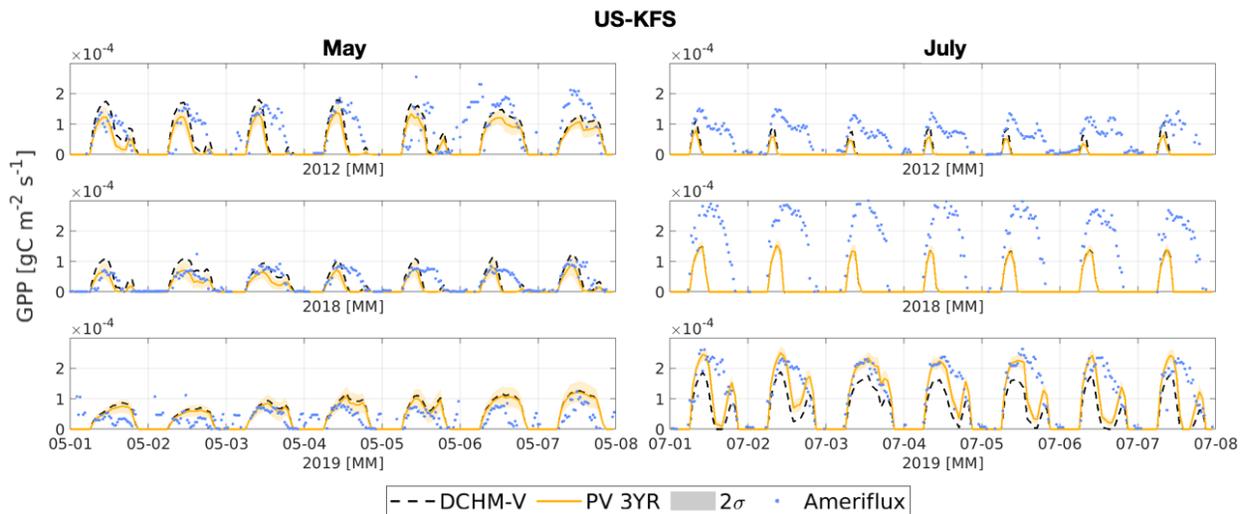

Figure N11 (Adapted). Hourly gross primary productivity [g C m^{-2} s^{-1}] from the DCHM-V and DCHM-PV shown against AmeriFlux 30-minute estimates for one week in May, July, and August of 2012, 2018, and 2019 atUS-KFS.
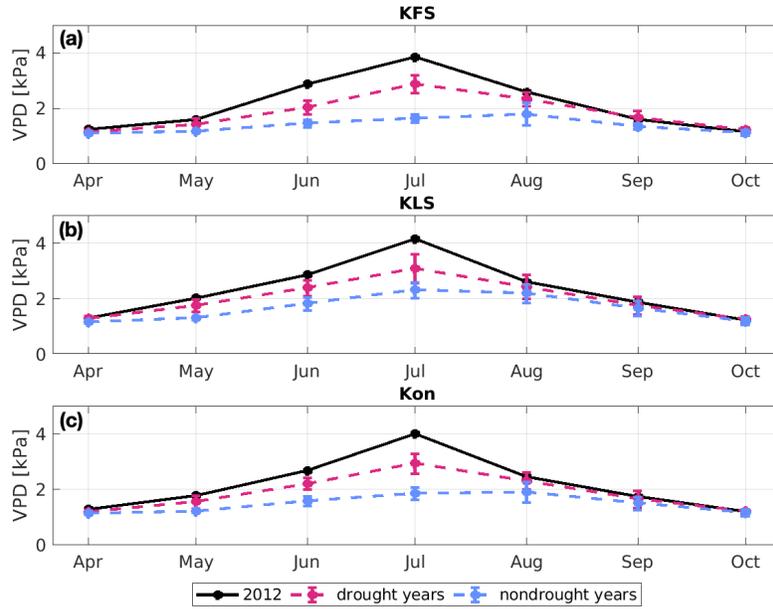
Figure N4 (appendix addition). Monthly average vapor pressure deficit [kPa] for the three AmeriFlux sites from April - October for the flash drought year 2012 (black), drought years (red), and non-drought years (blue). The error bar represents one standard deviation across drought and non-drought years.
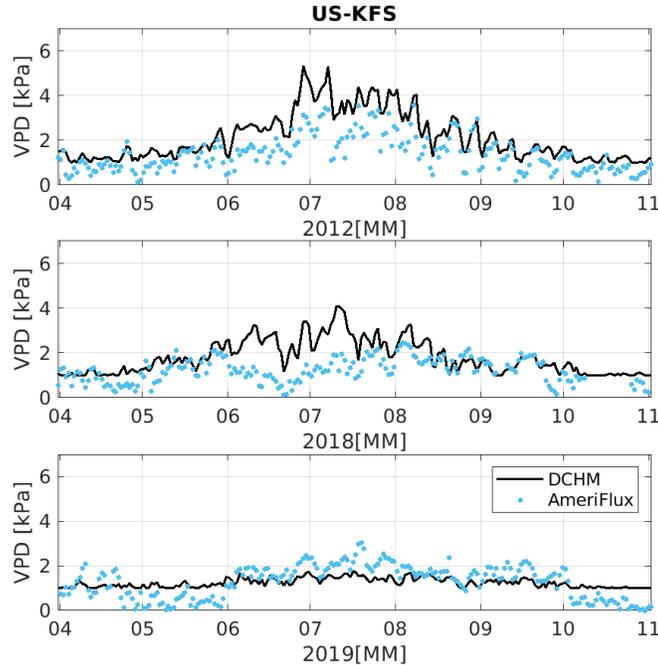

Figure N13 (Appendix) Daily vapor pressure deficit at US-KFS for (a) 2012 - flash drought, (b) 2018 - drought and (c) 2019 - non-drought. The DCHM computes VPD using air temperature and vapor pressure from NLDAS-2 Forcing File A.

2)   Noah-LSM: Noah LSM has multiple versions. If the Noah-LSM used in this study refers to the Noah in NLDAS-2, please specify.

The reviewer's point is well-taken. The Noah-LSM in this study does refer to the Noah model employed in NLDAS-2 (Xia et al., 2012). We will update the soil moisture figure captions and any references to NLDAS-2 soil moisture computed using Noah-LSM in the main body of the manuscript and in the Appendix. We also propose to combine figures from the Appendix so that Figure A1 and A2 become A1 a,b,c to represent the top layer soil moisture at US-KFS for 2012, 2018, and 2019.
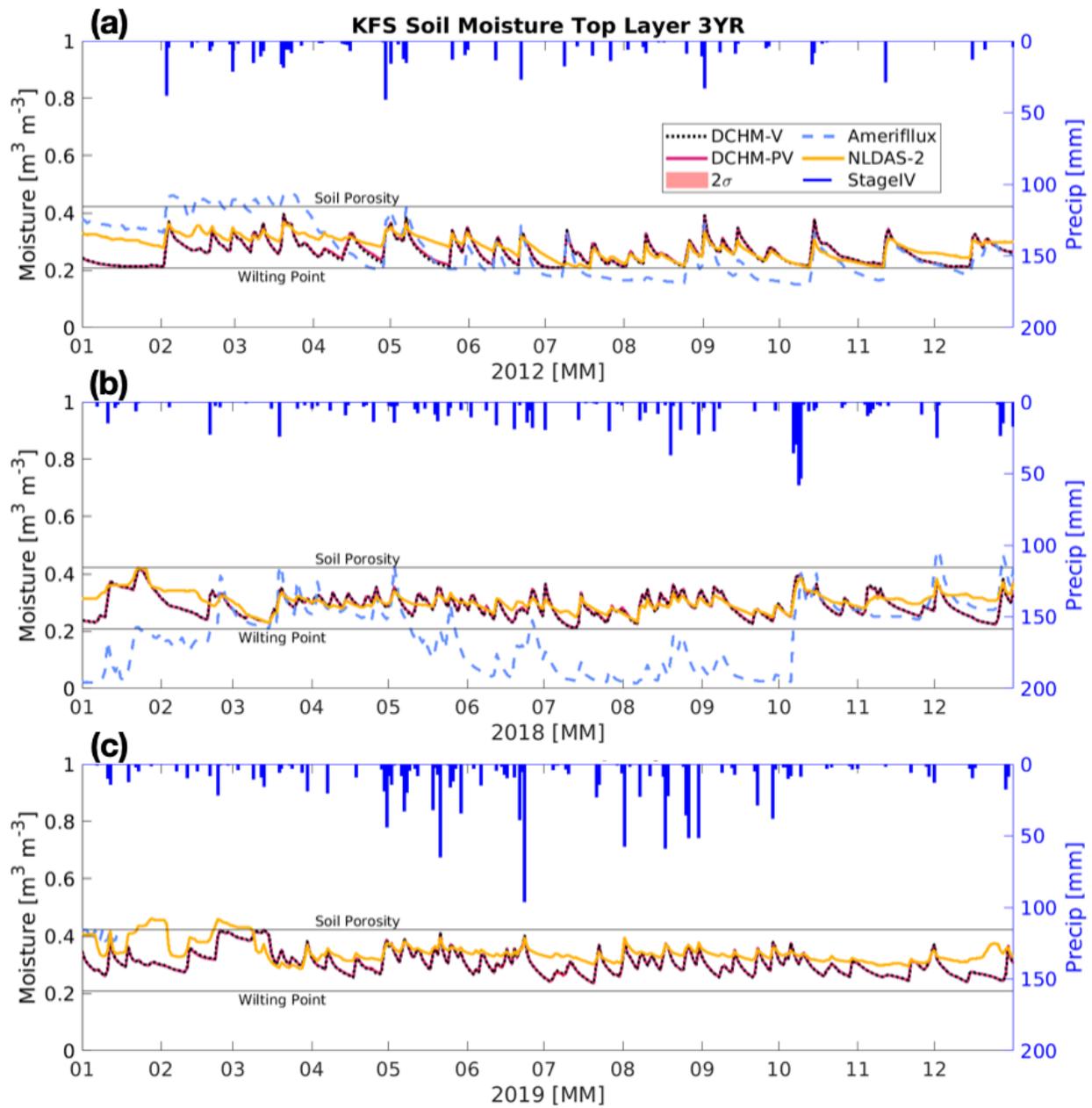
3)   line 259: of gamma => of the growth rate parameter

This comment is well taken. Following a similar comment from Review 1, we now use gamma once it is defined throughout the remainder of the manuscript rather than going back and forth between gamma and the growth rate parameter.

4)   Figure 12: May want to increase the thickness of curves for 2012 and 2019 to highlight the results for these two years

This comment is well taken. We have updated many figures to use thicker lines, varied color, dashed lines, and new marker shapes to help distinguish between simulations/years. Figure 12 has been completely reformatted so the flash drought can be compared to other drought and non-drought periods, as opposed to solely 2019.
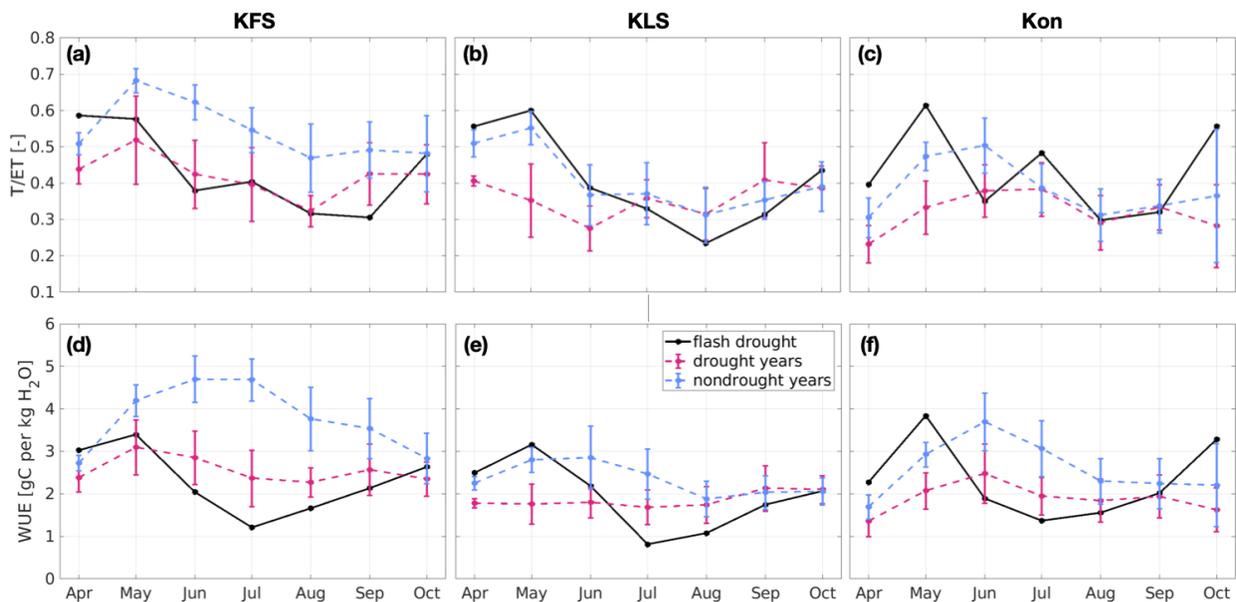

Figure 12 (replacement). Modeled growing season monthly averages of transpiration as a fraction of ET (a-c) and water use efficiency (WUE, d-f) for drought and non-drought years compared with the flash drought year of 2012 for US-KFS, US-KLS, and US-Kon AmeriFlux sites. Monthly averages are computed 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.

5) *Line 390: (Figure 10 => (Figure 10)*

This review comment is well taken, and we will implement this change.

6) *Figures A3, A5. Middle and deep layer soil moisture for the flash drought year 2012. How to explain the substantial differences between DCHM-V/DCHM-PV and Noah-LSM? Noah-LSM seems to make more sense as it shows a notable decline after June 2012. In contrast, the soil moisture in DCHM-V/DCHM-PV remains relatively steady throughout 2012 and does not seem to be responsive to the strong precipitation deficits during 2012, which looks odd; this is concerning as any issues in simulating soil moisture would adversely impact the simulation of vegetation and evapotranspiration processes etc. Please also see my second main comment.*

The reviewer's comment is well taken. First, see updates to Figure A3, which will now be A2 and combine 2012, 2018, and 2019 middle layer soil moisture for US-KFS. We respond below by (1) explaining why we see differences between the DCHM and NLDAS-2 soil moisture, and (2) by describing how these differences impact estimates of carbon uptake (GPP) and transpiration (T). We investigate soil moisture, GPP, and T by comparing our results to another modeling study who investigated US-KFS and US-Kon during 2012 and 2018 (Hosseini et al. 2022).

**NOTE: We cannot reproduce the figures referenced from Hosseini et al. (2022) here. Instead, we reference specific figures and panels from their paper for comparison. In reference to soil moisture, see the bottom four panels of Figure 6 in Hosseini et al. (2022). In reference to GPP, see bottom panels of Figure 3 in Hosseini et al. (2022). In reference to transpiration, see the third panel of Figure 5 in Hosseini et al. (2022). In reference to LAI, see the top panels of Figure 6 in Hosseini et al. (2022).

    (1) Why we see differences

**NOTE: In the following paragraphs we compare DCHM soil moisture from different layers to other products (SMERGE, NLDAS-2) and model outputs Noah-MP (Hosseini et al., 2022). Layer depths do not directly compare so for reference, we briefly state the various depths used.

The DCHM top layer soil moisture is an average over 0-8 cm, the middle layer is 8-89 cm, and the deep layer is 89-183). Depths were determined from the Kansas Soil Survey (Soil Survey Staff). In Hosseini et al. (2022) the top layer in Noah-MP is 0-10 cm and the deep layer average soil moisture they present comes from three layers with thicknesses of 30, 60, and 100 cm. Effectively, this is an average over 10-200 cm vs the DCHM which ranges from 8-183 cm . We average the DCHM middle and deep layers for comparison (see Figure below) and convert Noah-MP estimates into volumetric soil water content for comparison. The NLDAS-2 soil moisture depths used for comparison are 0-10 cm, 0-100 cm, 100-200 cm (Xia et al. 2012) to compare against the DCHM top, middle, and deep layers, respectively. In figures of the middle layer soil moisture, we include comparisons to SMERGE 0-40 cm, computed from "merging" NLDAS and the European Space Agency satellite soil moisture (Tobin et al. 2019).

A first explanation for why we see differences between DCHM modeled soil moisture and NLDAS-2 is that NLDAS-2 soil moisture was estimated from the Noah-LSM without predictive phenology (Xia et al., 2012). However, Hosseini et al. (2022) used various Noah-MP configurations (including with and without predictive phenology) to compute soil moisture, and the DCHM results match well with their soil moisture at US-Kon in 2012 and 2018 (see Figure 6 in Hosseini et al. 2022). Converting units from mm to $m^3 m^{-3}$, we see that Noah-MP predicts a drop in 2012 soil moisture at US-Kon from ~0.35 to 0.28 $m^3 m^{-3}$ from January to September while the DCHM sees a drop of about from ~0.36 to 0.30 $m^3 m^{-3}$. The Noah-MP model configuration that uses dynamic LAI and vegetation fraction (V3-LD-FD) predicts soil moisture decay from June-September that shows the least steep decline in soil moisture from late June to late August (Hosseini et al. 2022, Figure 6 bottom panel), aligning with results from the DCHM-V and -PV (Figure N14, N15 and additional figure below).
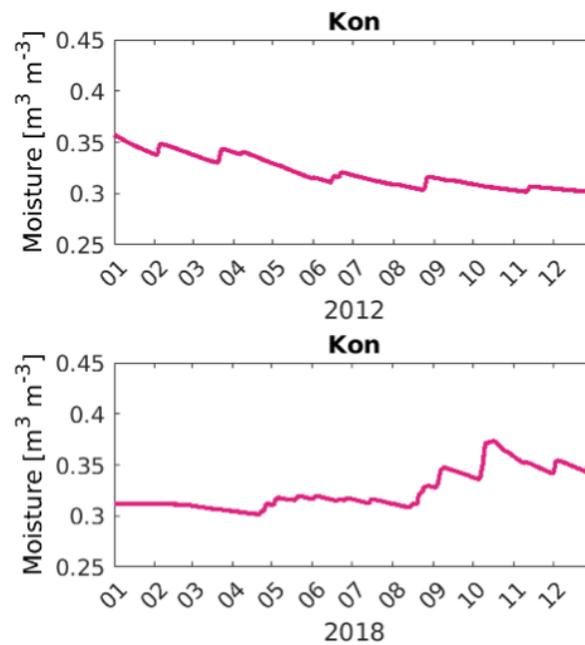

Figure (averaging outputs from N14 and N15). DCHM-PV 3YR volumetric soil moisture averaged across middle and deep layers for US-Kon in 2012, 2018.

It is also important to note that Hosseini et al. (2022) estimates of the top 10 cm of soil moisture match well the magnitude of flux tower soil moisture, fluctuating between ~0.15-0.3 $m^3 m^{-3}$, between May and July. These findings agree favorably with DCHM top layer soil moisture in 2012 (Figure N14). However, like the DCHM, all model configurations of Noah-MP in Hosseini et al. (2022) estimate lower soil moisture compared to field measurements in the top layer from mid-February to early May 2012 and higher soil moisture from early May through the rest of the year except for some spikes preceding larger rainfall events. Similarly for top layer soil moisture results from 2018, all of the Hosseini models and the DCHM overestimate soil moisture compared to field observations starting in late April and throughout the end of the year (Figure N14a). Thus, the DCHM model results for soil moisture in 2012 and 2018 at KON are in line with what has previously been estimated from different configurations of the Noah-MP that use

predictive phenology and differ similarly from the NLDAS-2 dataset and field observations of soil moisture.

A second explanation of the DCHM estimating higher soil moisture than NLDAS-2 might have to do with cascading effects high VPD has on stomatal conductance. In response to high VPD in the DCHM, stomatal conductance shuts down (Figure N9). Therefore plants are not transpiring. Reduced transpiration is directly tied to reduced root water uptake, resulting in the soils retaining comparatively higher levels of moisture. Figure A7 shows that modeled middle and deep layer root water uptake decreases ~50% from May to July 2012 at US-KFS. Within the DCHM, reduced root water uptake (Figure A7) is likely why estimates of soil moisture in the middle and deep layers remain higher compared to SMERGE and NLDAS-2 (using Noah-LSM) soil moisture (Figures A2 and A3) at US-Kon. However, the DCHM and SMERGE agree favorably in 2012 and 2018 throughout most of the growing season at US-KFS. Note that the DCHM matches well middle and deep layer estimates of soil moisture from NLDAS-2 and SMERGE in 2019 when there is ample water available for plant use within the DCHM.

   (2) How these differences impact estimates of GPP and transpiration

GPP

The DCHM estimates low GPP and stomatal conductance rates during the flash drought period in 2012, while eddy covariance data recorded elevated rates of GPP (e.g., Figure 8, N9, N11). The low estimates of GPP and stomatal conductance from the DCHM are directly related to high atmospheric aridity (or VPD) indicating that the DCHM slows carbon and water exchanges under atmospheric water stress, despite sufficient soil moisture to undergo photosynthesis.

Hosseini et al. (2022) report 11-year (2008-2018) averages of GPP for US-Kon and US-KFS using different Noah-MP configurations, MODIS and AmeriFlux data (Figure 3 in Hosseini et al. 2022). In the figure below, we show the same 11-year averages computed from the DCHM-PV. Noah-MP using predictive LAI configurations estimates higher GPP in April (~150-200 gC m^{-2}) and May (~300 gC m^{-2}), than the DCHM by ~100 gC m^{-2} for similar soil moisture during this time (see Figure 6 in Hosseini et al. 2022 and Figure A2 and N14 below). Both Noah-MP and the DCHM GPP peak in June and the Noah-MP results fall within one standard deviation of the DCHM in June and July at both sites. However, the DCHM 11-year averages of GPP match well the Apr-Oct averages from flux towers for KFS. Noah-MP includes routines for reallocating carbon to different parts of plants (i.e. stems, roots, etc.) that may account for the higher estimates of GPP compared to the DCHM, which does not include such processes.
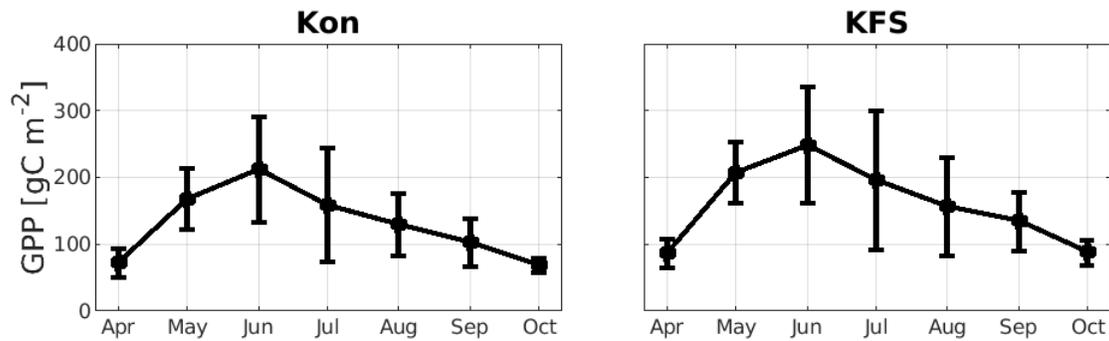
Figure. Monthly GPP averages across the same 11-year period (2008-2018) as Hosseini et al. (2022) using ensemble mean estimates from the DCHM-PV 3YR. Error bars represent one standard deviation from the 11-year average.

Transpiration

The maximum daily transpiration rate estimated from the DCHM, which computes transpiration from root water uptake, is 1.25 mm d^{-1} in 2012 and 2018 (Figure below), but the Noah-MP modeled transpiration reach over 2mm d^{-1} in May and June for both 2012 and 2018. July - September rates of transpiration for both the DCHM and Noah-MP (with dynamic LAI) fall to less than 0.5 mm d^{-1}. Peak transpiration in May and June of 2012 before a decrease to lower transpiration rates in July-October is observed in both the DCHM and Noah-MP (see the third panel of Figure 5 Hosseini et al., 2022) although there are differences in magnitude of transpiration, some of which can be attributed to the differences in computed LAI. Like Hosseini et al., (2022), the DCHM estimates two seasonal peaks of transpiration in June and September of 2018. The late season peak seems to align with large increases in late season precipitation.

Some of the discrepancies in transpiration may result from differences in estimated LAI from both models. The DCHM-PV estimates of LAI tend to agree favorably with the timing of green up and seasonal changes compared to MODIS (see RC1 for full Figure 6 showing LAI at all three sites from 2012, 2018, 2019). At US-Kon, the DCHM-PV 3YR shows April LAI less than 1 m^2 m^{-2} (see our Figure 6g below,), but Hosseini et al., (2022) estimates leaf out earlier and with April LAI at ~2.7 m^2 m^{-2} (see top panels of their Figure 6). The uptick in transpiration seen by Hoesseini in September 2012 might also be due to the increase in LAI from 0.2 to 2.0 m^2 m^{-2} that they found at the same time. Meanwhile, the uptick in LAI seen by the DCHM-PV was from 1.0 to 1.2 m^2 m^{-2}.

Overestimating LAI leads to overestimating latent heat fluxes, as transpiration is a component of latent heat. DCHM estimates of latent heat in May and June of 2012 are less than that of flux tower by ~100 W m^{-2} and match tower measurements well when during wet periods, like 2019 (Figure below). In Noah-MP (Niu et al, 2011; Ma et al., 2017, Li et al., 2021) and in the DCHM, transpiration is directly related to root water uptake which depends on canopy (and stomatal) conductance and both models compute canopy conductance using LAI. Soil moisture across the two models was similar, but LAI varied by over 1 m^2 m^{-2} during the growing

season. Thus, LAI and not differences in soil moisture are likely responsible for differences in modeled GPP and transpiration. .
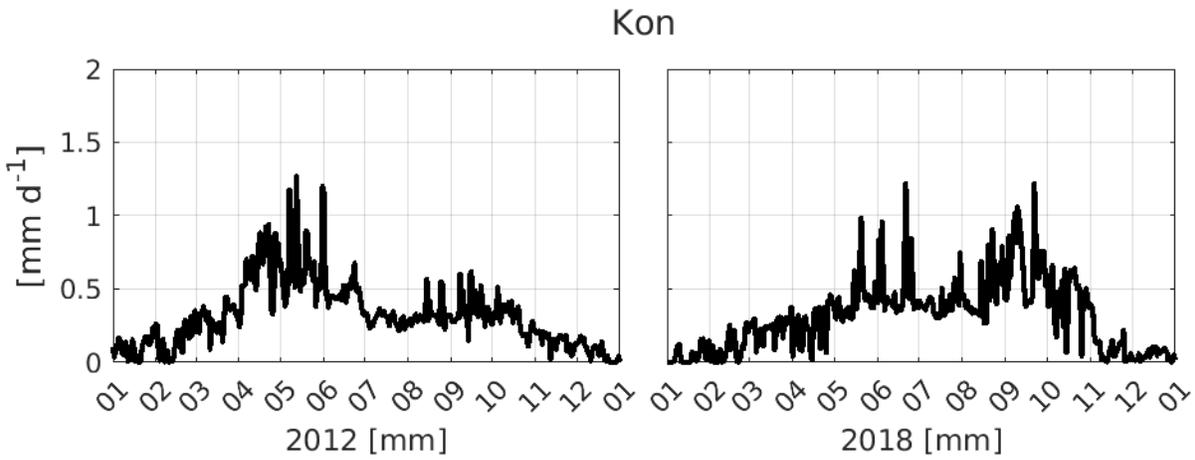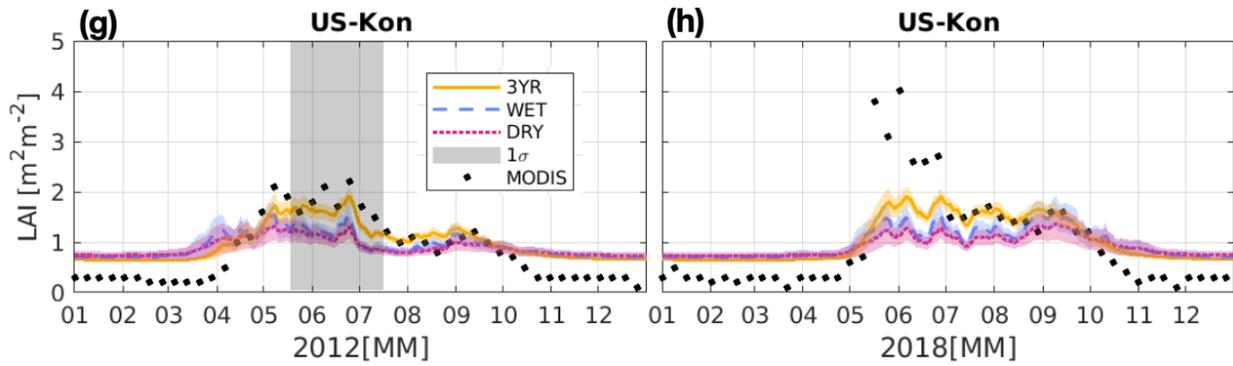

Figure. Daily transpiration averaged over daytime.

Cropped from 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.
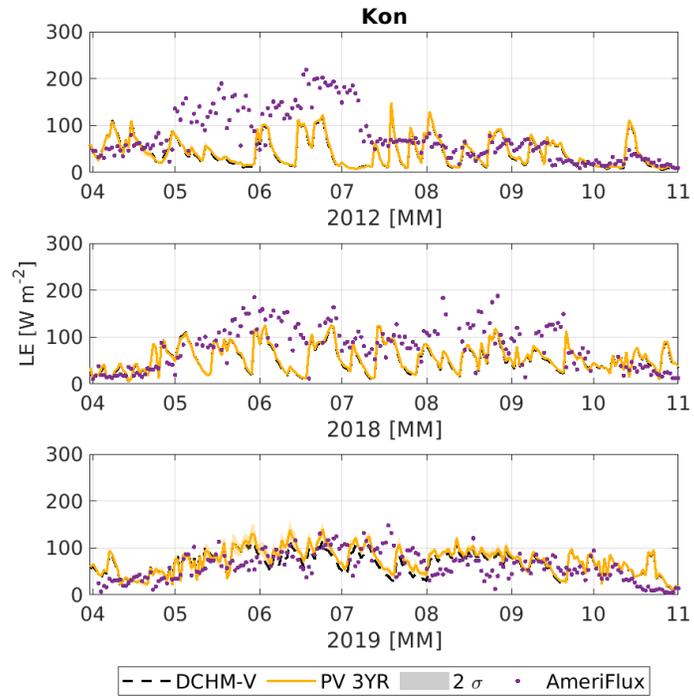
Figure. DCHM estimates of latent heat at US-Kon for 2012, 2018, 2019 compared with AmeriFlux.
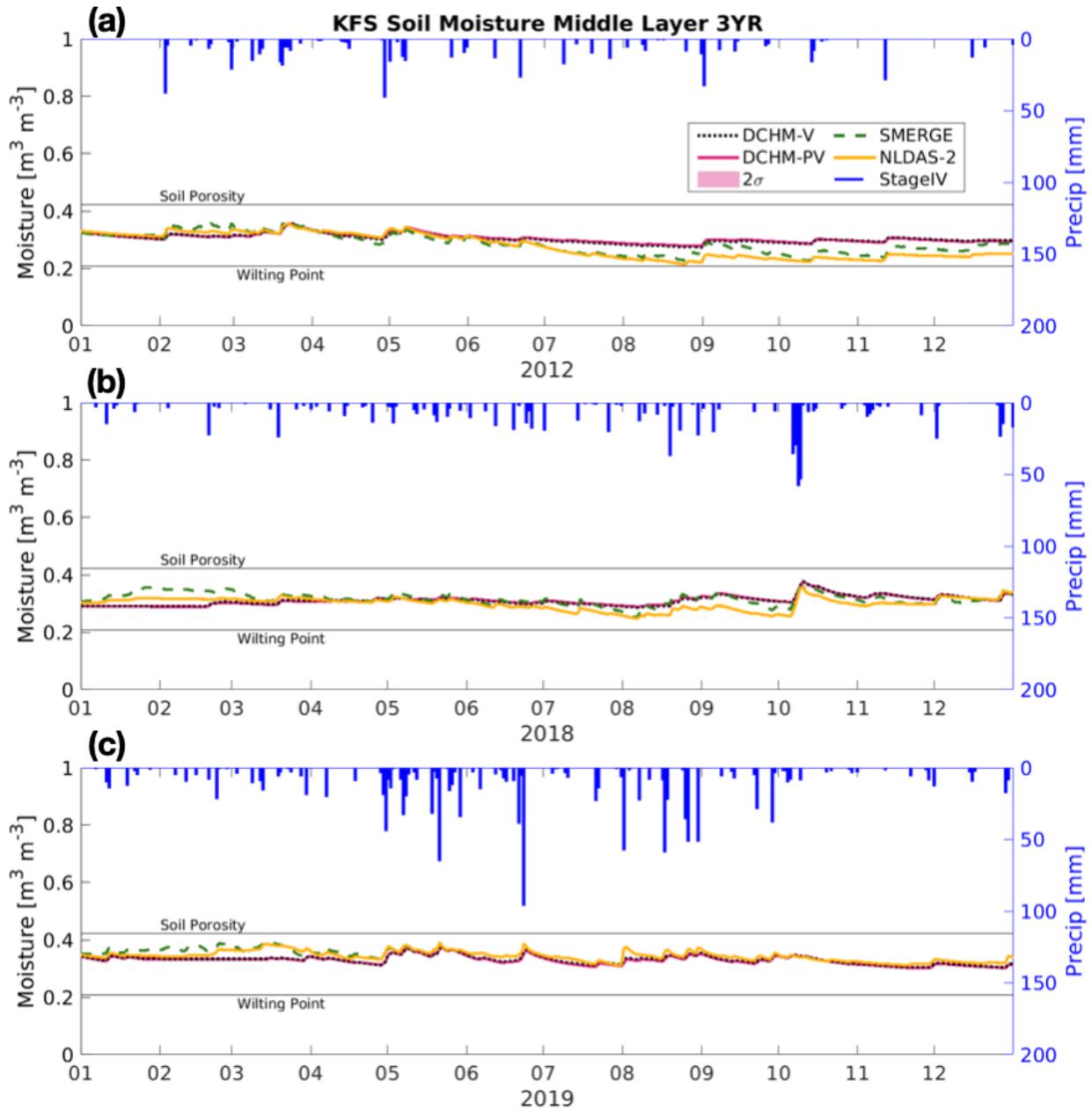
Figure A2 (newly created to combine A3 and A4 and adding 2018). **Middle layer soil moisture at US-KFS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year using the DCHM-V (black dotted line), the DCHM-PV with two standard deviations (red), SMERGE (green dashed line), NLDAS-2 derived from Noah-LSM (yellow) and Stage IV precipitation on the top and right axes (blue).**
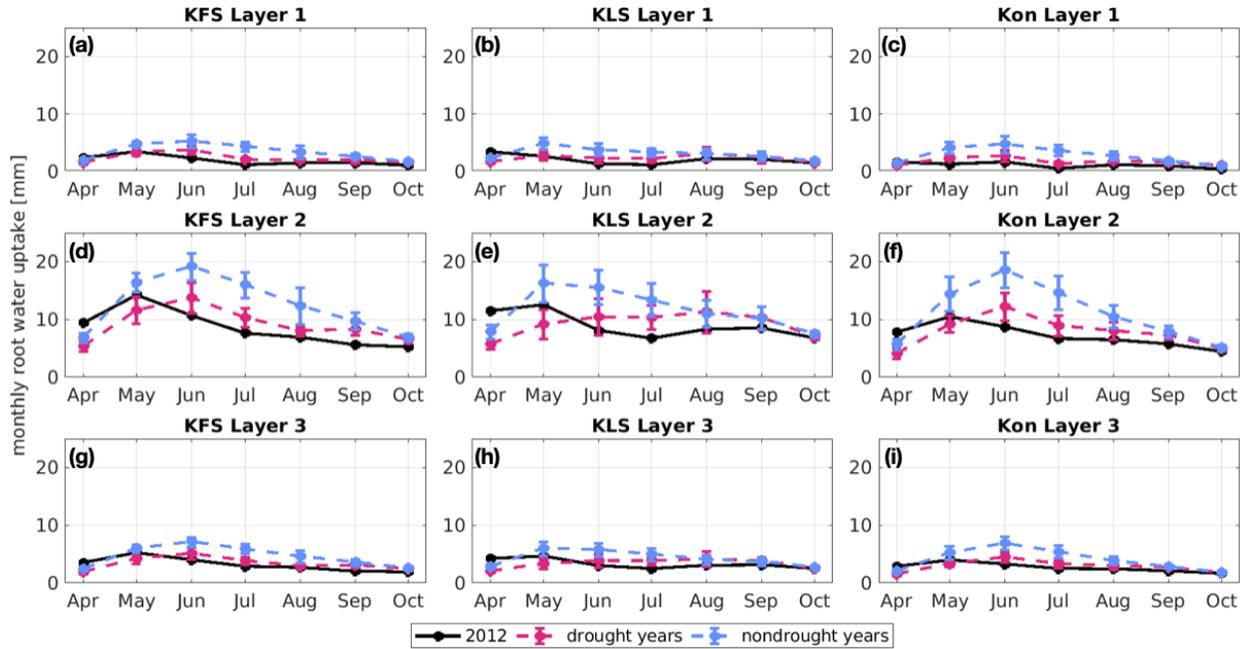
Figure A7 (Replacing A7 and A8). DCHM-PV 3YR monthly root water uptake totals for drought (red) and non-drought (blue) years compared to 2012 (black) across three soil layers for our three study sites. Monthly sums are computed from the ensemble means of the 2000 Monte Carlo simulations then averaged 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.
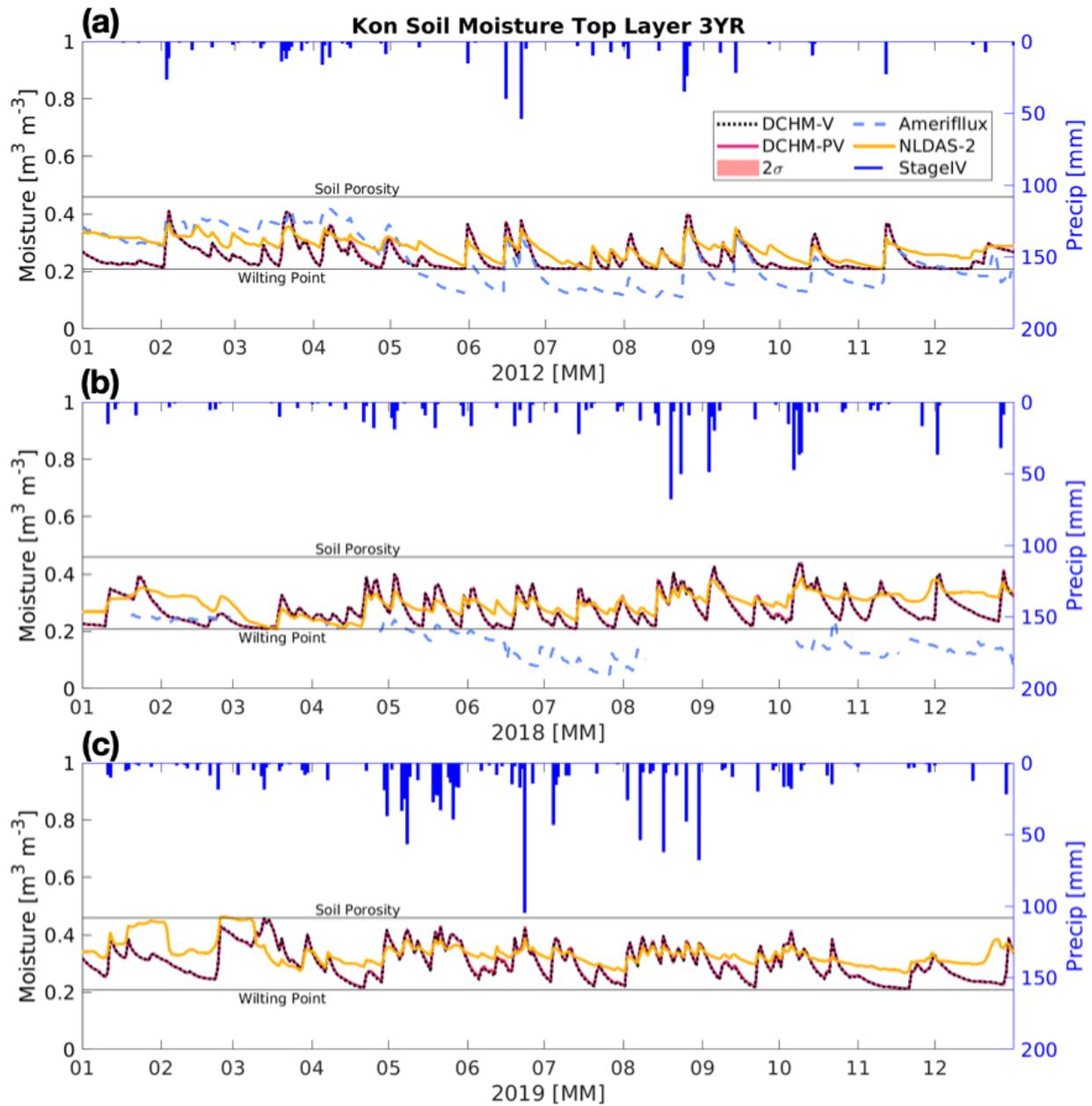
Figure (N14 - appendix). Top layer soil moisture at US-Kon for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year using the DCHM-V (black dotted line), the DCHM-PV with two standard deviations (red), AmeriFlux (blue dashed line), NLDAS-2 derived from Noah-LSM (yellow) and Stage IV precipitation on the top and right axes (blue).
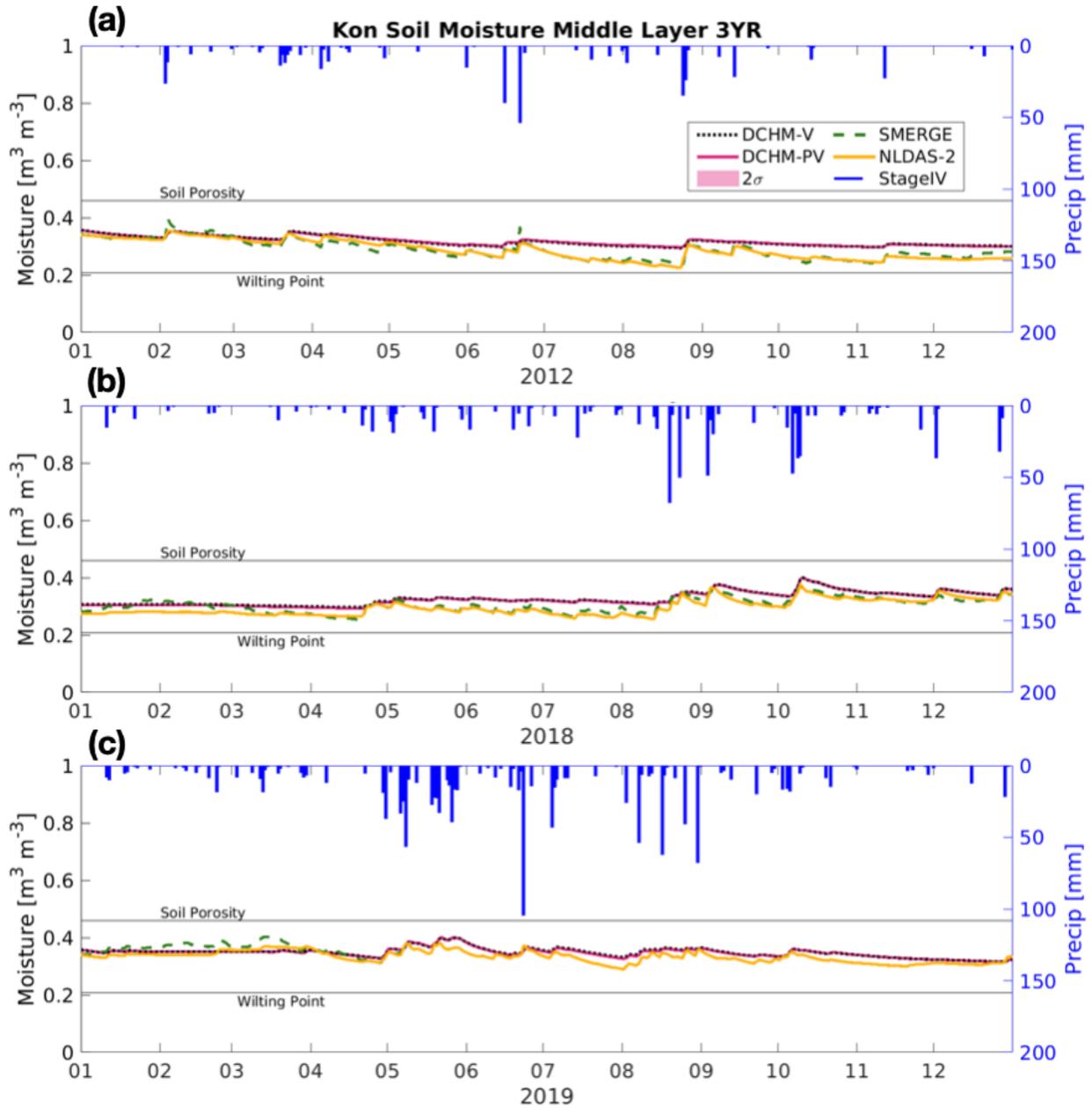
Figure N15 (appendix). **Middle layer soil moisture at US-Kon for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year using the DCHM-V (black dotted line), the DCHM-PV with two standard deviations (red), SMERGE (green-dashed line), NLDAS-2 derived from Noah-LSM (yellow) and Stage IV precipitation on the top and right axes (blue).**

7)   Figure A6 is identical to Figure A5 and appears to be incorrect. Please check if it plots the results for 2019.

The author's appreciate the reviewer pointing this out. We have fixed this mistake and combined into one figure while adding 2018. This mistake also happened with A1 and A2 (see combination above). We can make similar combinations of soil moisture plots for other sites and layers to add to the appendix.
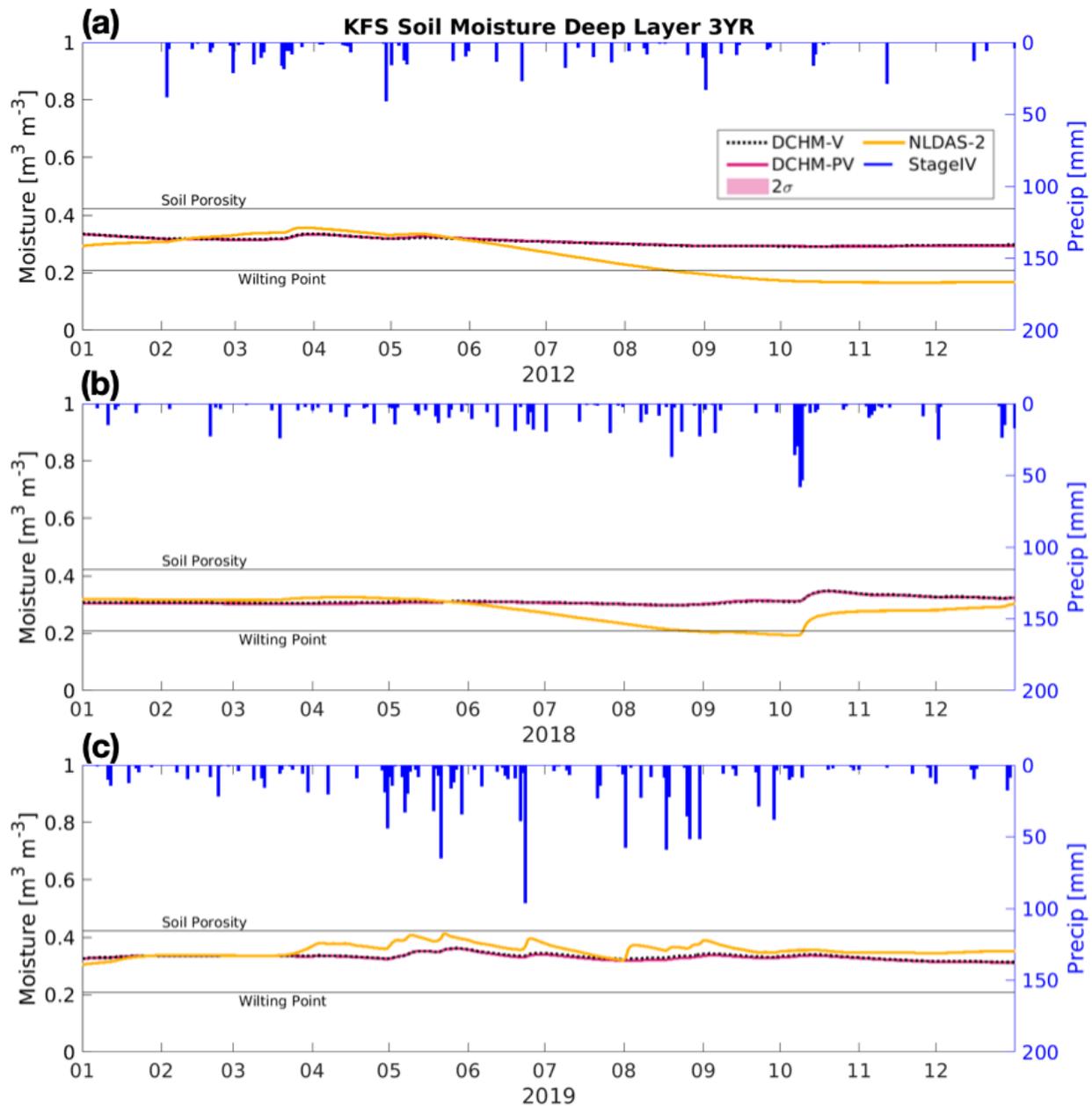

Figure A3 (new and result of combining A5 and A6 with results from 2018**). Deep layer soil moisture at US-KFS for (a) 2012, flash drought, (b) 2018 drought and (c) 2019 a non-drought year using the DCHM-V (black dotted line), the DCHM-PV with two standard deviations (red), NLDAS-2 derived from Noah-LSM (yellow) and Stage IV precipitation on the top and right axes (blue).**

8) *Figure A10: "during 2012"=>"during 2019'?*

The authors thank the reviewer for pointing out this error. We will update the figure caption accordingly. We also propose to provide updated figures with the DCHM averaged over only the daytime hours as mentioned above in response to Major Comment 2. We update the color scheme to be monochromatic grayscale to be more vision friendly. It should be noted that the WET and DRY were identical to the 3YR. This was a bug in the plotting code that we fixed.

An example of one of the new figures is below. With the addition of a new figure, A10 might not be the label in the revised manuscript.
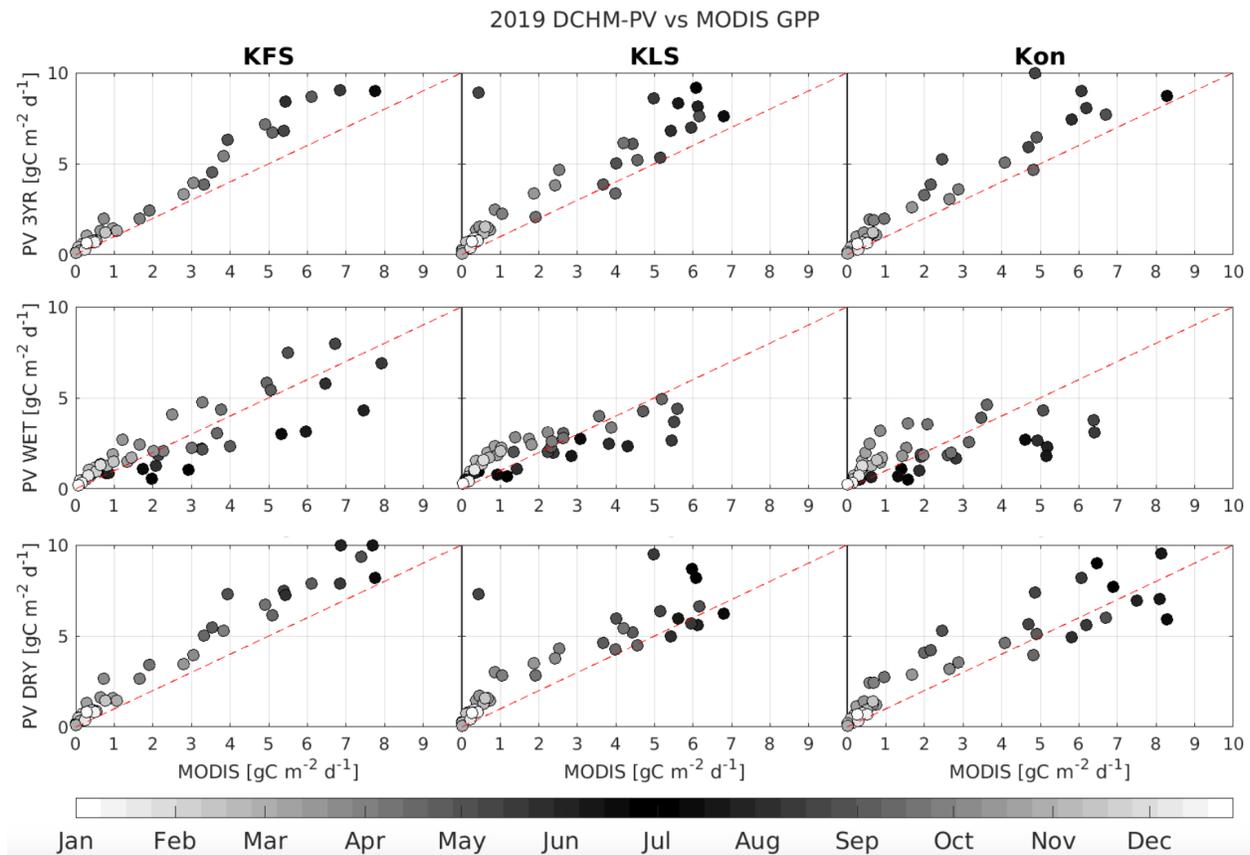

Figure A10 (replacement). MODIS (MOD17A2H) vs DCHM-PV 3YR, WET, and DRY for all three sites during 2019.

9) *Figure A11a: The difference between Ameriflux and model simulation is striking. The inclusion of Ameriflux appears to cause confusion rather than providing a truthful evaluation of the model results.*

The authors appreciate this comment from the reviewer. The data discrepancies were striking and were the result of an error made when plotting. See Response to Major Comment 3a from the responses to RC1 and Response to Major Comment 2 above.

With updates to how we compute daily averages from model GPP and the use of AmeriFlux FLUXNET, we see that model and AmeriFlux are in better alignment. There is still a striking difference in June and July of 2018 (newly added drought year) that suggests during drought there may be something plants are doing below ground to maintain higher rates of GPP that the DCHM is not capturing. We feel that the use of AmeriFlux FLUXNET in updated figures (including Figure A11 above) are now more useful in evaluating model performance.

Closing remarks

The authors would like to express our gratitude for the thoughtful comments and that our replies provide clearer and deeper analysis of evaluating the role of vegetation of the movement of water and carbon during flash drought. We understand that should this manuscript be accepted for publication, that there are several new passages and figures (both here and in our response to Reviewer Comment 1) that will need to be included (or removed) and that other changes to enhance cohesiveness of the manuscript in light of the new analysis will need to be incorporated.

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. 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.
3. Farquhar, G. D., & Sharkey, T. D. (1982). Stomatal conductance and photosynthesis. *Annual review of plant physiology*, *33*(1), 317-345.
4. Flack-Prain, S., Meir, P., Malhi, Y., Smallman, T. L., & Williams, M. (2019). The importance of physiological, structural and trait responses to drought stress in driving spatial and temporal variation in GPP across Amazon forests. *Biogeosciences*, *16*(22), 4463-4484.
5. Garcia-Quijano, J. F., & Barros, A. P. (2005). Incorporating canopy physiology into a hydrological model: photosynthesis, dynamic respiration, and stomatal sensitivity. *Ecological Modelling*, *185*(1), 29-49.
6. 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*.
7. Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R. T., ... & McDowell, N. G. (2020). Plant responses to rising vapor pressure deficit. *New Phytologist*, *226*(6), 1550-1566.

8. 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.

9. 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.

10. Katul, G., Lai, C. T., Schäfer, K., Vidakovic, B., Albertson, J., Ellsworth, D., & Oren, R. (2001). Multiscale analysis of vegetation surface fluxes: from seconds to years. *Advances in Water Resources*, *24*(9-10), 1119-1132.

11. Lawson, T., & Vialet‑Chabrand, S. (2019). Speedy stomata, photosynthesis and plant water use efficiency. *New Phytologist*, *221*(1), 93-98.

12. Li, L., Yang, Z. L., Matheny, A. M., Zheng, H., Swenson, S. C., Lawrence, D. M., ... & Leung, L. R. (2021). Representation of plant hydraulics in the Noah‑MP land surface model: Model development and multiscale evaluation. *Journal of Advances in Modeling Earth Systems*, *13*(4), e2020MS002214.

13. 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.

14. 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.

15. 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.

16. Ma, N., Niu, G. Y., Xia, Y., Cai, X., Zhang, Y., Ma, Y., & Fang, Y. (2017). A systematic evaluation of Noah‑MP in simulating land‑atmosphere energy, water, and carbon exchanges over the continental United States. *Journal of Geophysical Research: Atmospheres*, *122*(22), 12-245.

17. Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., ... & Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah‑MP): 1. Model description and evaluation with local‑scale measurements. *Journal of Geophysical Research: Atmospheres*, *116*(D12).

18. 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.

19. Schmid, H. P. (1994). Source areas for scalars and scalar fluxes. *Boundary-Layer Meteorology*, *67*(3), 293-318.

20. Soil Survey Staff, Natural Resources Conservation nService, USDA: Web Soil Survey. Available online. Accessed 07 December 2022, https://www.nrcs.usda.gove/resources/data-and-reports/web-soil-survey.

21. 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.

22. Tobin, K. J., Crow, W. T., Dong, J., & Bennett, M. E. (2019). Validation of a new root-zone soil moisture product: Soil MERGE. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(9), 3351-3365.

23. Urban, J., Ingwers, M., McGuire, M. A., & Teskey, R. O. (2017). Stomatal conductance increases with rising temperature. *Plant signaling & behavior*, *12*(8), e1356534.

24. Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., ... & Lohmann, D. (2012). Continental‐scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS‐2): 2. Validation of model‐simulated streamflow. *Journal of Geophysical Research: Atmospheres*, *117*(D3).