Author Response to Anonymous Reviewer #2
The authors appreciate the thoughtful review and critique of the manuscript, with many constructive suggestions. Here we provide a general response, addressing points in the order they were presented in the review. Review comments are in bold face.

The paper titled "Unravelling the impacts of precipitation, temperature, and land-cover change for extreme drought over the North American High Plains" uses the ParFlow-CLM model and a set of synthetic climate scenarios to investigate the role of different extreme climate conditions and their interactions in the severity of drought. The novelty in the paper is in its objective to quantify how comprehensively integrating subsurface flow into the modeling of the land surface acts to enhance or reduce drought conditions and the role that temperature, precipitation, and land cover play in those emergent conditions. The paper is appropriate for HESS and I believe it will eventually be suitable for publication. However, the scenarios need to be revisited and rerun before publication.

Main issues: The meteorological scenarios used in this study are physically inconsistent. One cannot just add or subtract from the temperature without accounting for how that impacts the rest of the meteorological states or fluxes. For example, downward longwave radiation is very strongly correlated to temperature and specific humidity is strongly connected to temperature. As such, if you are modifying one you have to modify the others. If not, the meteorology with which you force the model is mostly fantasy and not realistic. There are methods out there to deal with this. This needs to be addressed before publication by first making physically consistent meteorological scenarios and then rerunning the simulations.

- We agree the meteorological changes are an approximation, but we believe they are suitable for the proposed research questions in this paper. For the goals of this study, we are not simulating a projected climate change (with, for example, a Global Climate Model prediction) nor are we predicting the actual drought. This is a difference between the present study and many published studies, as outlined in the literature review. Instead, we are using one-factor perturbations to the model (forcing, land cover) in a systematic way to run numerical experiments.
- By making a single change at a time, we can attribute any differences between the baseline and the perturbed run to the single variable that was perturbed. This is an advantage of modeling studies over real world observations, as we can assess process interaction with much greater precision and detail. If we changed humidity and pressure with associated temperature changes we would be modifying three things at once, which limit the linearity arguments and the strength of the experiment as such.
- It is worth noting that the model runs are computationally expensive. For this paper, each year used about a week of wall-clock time and over 20,000 processor hours. This expense means that understanding individual impacts is

valuable from a modeling perspective, to lay a foundation before proceeding to more complex simulations.

- Other studies have also used the approach of changing individual variables or small groups of variables. For example, Kollet and Maxwell perturbed temperature without changing other meteorological variables for a study of drought in the Little Washita watershed. Rasmussen et al (2011) also employed a simplified approach they called “pseudo-global warming”, adding estimated climate perturbations in temperature, vapor mixing ratio, boundary layer height and windspeed to a forcing dataset. Markovich et al (2016) perturbed temperature alone in a similar study of climate change in California, and Pribulick et al (2016) perturbed temperature and land cover in studying the impacts of vegetation change under global warming in a Colorado watershed. While the reviewer is correct that changing one meteorological variable does not fully represent the physical system, it is a documented simplification used in multiple published papers.
- We will explain these points more clearly and add more details and justification on this modeling approach and perturbations to the methods section. We will also note that model runs with changes in all meteorological forcing is a good possibility for future study.

Other comments:
Abstract, Line 10 - physically based equations? A lot of land surface and hydrologic models have physically-based equations. I really don’t understand the need to use this terminology here. ParFlow is indeed an advanced groundwater model that uses our best understanding of how to model the flow of subsurface water over regional scales.
However, that does not mean that other models do not use physically based equations. As long as models are not just complete black boxes, they will always have physically based equations. Those equations might be over-simplistic but they are still physically- based. In my opinion, the better analogy of ParFlow is the “superparameterization” term used in climate models. Anyway, I don’t say this to lessen the value of ParFlow but using this terminology distracts the reader from the real strengths of ParFlow.
We agree with the reviewer that most models have some physical basis for their equations. We used the term “physically based” to distinguish from statistically based models. While we believe this is useful shorthand for the abstract, we appreciate the suggestion from the reviewer and we don’t want this terminology to distract from our message. We will expand our discussion in the methods to make it clear that we understand ParFlow is not the only physically based model.

Abstract, Line 13 - I don’t think there was a need to run such a complex model to prove that decreases in precipitation will cause the largest negative anomalies in the evapotranspiration, soil moisture, streamflow, and water table levels. The only flux that would compete with precipitation would be shortwave radiation. However, fortunately for us, the sun is not going anywhere anytime soon.

Indeed, the first research question has been addressed in other literature; reviewing impacts of different factors with this model shows that our results are consistent with other literature. Additionally, in this exercise we are able to quantify the impacts in detail, rather than making qualitative arguments. Characterizing the impacts in general terms is a preliminary to more detailed analysis of nonlinearity and scaling of these impacts in research questions 2 and 3.

Page 1, Line 22 - There are studies over the past few years that show that drought recurrence has changed very little over the past half century. Droughts have devastating impacts regardless of if we consider climate change or not. The value of studying the impact of subsurface flow on droughts is critical regardless of the impact of climate change.
This is a good point; we will expand the discussion on importance to include it.

Page 3, Line 6 - It is true that the different meteorological factors are strongly interconnected. However, there is a reason for that. The states and fluxes are inherently interconnected. You can't think about precipitation without thinking about temperature. And you especially can't think about specific humidity, radiation fluxes, among others without thinking about temperature. As such, although separating them is a nice thought experiment it is not any realistic representation of what is going on in the observed physical system.
We agree with the reviewer's point, but perhaps not their conclusion. In addition to referring back to our response to this point, above, we would like to note that our model goal is not to represent all the complexity of meteorology, but to test narrowly defined specific scenarios for the purpose of a numerical experiment. The intent of experiments like these is not to represent the observed physical system in all ways, but rather to lay the groundwork for more complex experiments and for interpreting observations in the future. We have attempted to be transparent in these goals and the simplifications we intentionally made to the system throughout the manuscript. We appreciate the reviewer's comments and we will make sure that the limitations of this simplification and our reasons for making it are more clear in the revised manuscript.

Page 3, Line 27 - Could we move away from the term "fully integrated hydrologic model"? I understand that you are reemphasizing the coupled groundwater, land surface model, surface flow advantages of the model. However, "fully integrated" has a very comprehensive meaning that leads to expecting too much from ParFlow-CLM and leads to stronger criticisms for the model than there should be. Fully integrated would mean a complete characterization of urban hydrology, water management, dynamic coupling with the ecosystem processes (i.e., ecohydrology), complete representation of flooding dynamics, among many many other processes. Again, ParFlow-CLM has significant advantages that should not be discounted or disregarded over more simplistic approaches and those should be emphasized. But you can do that in a more direct way instead of broad terminology such as "fully integrated".

We will replace “fully integrated” with “integrated hydrologic model” for greater clarity, as the reviewer suggests, and add a brief discussion of terminology to the methods.

Page 4, Line 23 - Van Genuchten functions should be Van Genuchten water retention curve. Also, it would be a good idea to cite the corresponding VG paper.
We will revise to make it clear that Van Genuchten functions refers to the Van Genuchten pressure-saturation and pressure-relative permeability relationship, and add a citation to the Van Genuchten 1980 paper.

Page 5, Line 2 - It is certainly true that VIC does not model the lateral flow between macroscale grid cells and this is a feature that should have been addressed a long time ago. Thankfully most land surface models are finally moving in the right direction and starting to include groundwater flow between macroscale cells. In any case, VIC does account for sub-grid flow of water even though it only does so implicitly in the variable infiltration curve. Stating that there is no lateral flow at all is not entirely accurate.
We will make the statement more precise to state that VIC “does not model lateral flow between macroscale grid cells”, as suggested.

Page 5, Line 7 - I agree with this statement to a certain extent. Although it is fairly comprehensive down to the 1km scale grid scale (if we disregard the role of faults and karst terrain in the regional flow), it still completely disregards finer scales. There are plenty of hydrologic processes that play a fundamental role in ecosystem function and structure that are far below the spatial resolution that you capture in ParFlow. As a result, there is still an underlying need to parameterize meter-scale coupling between the water, energy, and biogeochemical cycles. CLM does that somewhat but it is certainly still far from the objective.
We agree this is an important point and will add some discussion on fine-scale processes that are not captured in the current representation.

Page 6, Line 6 - This should say sink-filled instead of processed. However, why is this necessary? The sinks in elevation maps can be real features. For example depressions in the landscape certainly exist. This is one of the large advantages in my opinion of ParFlow vs more simplified models since it allows you to directly model those depressions through subsurface redistribution or through accumulation over the surface and redistribution. Maybe I am misunderstanding something here. Please clarify.
This is correct; the terrain was processed with a sink-filling algorithm and we will be more explicit about this in the revised manuscript. We agree with the reviewer that natural depressions in the landscape exist; however, literature suggests that it is difficult to distinguish these natural depressions and sinks from noise within the processed DEM (eg Kenny et al 2008). For the simulations presented here we use the kinematic wave approximation of the overland flow equations and therefore we require a domain with a connected drainage network. Additionally, we would like to note that this approach is

consistent with the previous work modeling the continental United States by Maxwell and Condon (2016) that our domain was built from.

Page 6, Line 17 - This is problematic. You can't assume pre-industrial conditions in water management but then use land use change from the past 30 years. They are fairly inconsistent. The agriculture over the plains is inherently dependent on irrigation. If you are going to have one you shouldn't throw out the other especially given the argument of a fully integrated approach. As you revisit the scenarios I would recommend also trying to address this inconsistency.
We agree that the agriculture over the plains depends on irrigation, and that this water use is very different from pre-industrial water management. Ideally, we agree that water use and land use would be consistent within the model; unfortunately, there is limited pre-industrial data on land use. In the case of this project, we are not reconstructing any specific historical drought, so it is less important in addressing the research questions to match all forcings and settings to one period of time.

Page 6, Line 25 - Recursively? Do you mean repeatedly? Recursively has a very specific meaning in computing that doesn't make much sense here.
We will change recursively to repeatedly for clarity.

Table 2 - It is important to realize that the moment that you start varying the climatic conditions you will start also varying the response of the biogeochemical cycles. Certain plant species will die while others will flourish. As a result, the land use scenarios that you use are inconsistent with what would actually occur with your prescribed forcing. I don't expect you to account for this because this is actually a problem in climate models. However, it might be a good topic to address this explicitly in the discussion section.
This is an important point and we will add discussion of this point to the paper.
We would also note that the historical Dustbowl was marked by widespread crop failures and massive topsoil loss. While we are not reconstructing the Dustbowl, this does imply that our land use scenario of replacing "crops" with "bare soil" is not unreasonable in the face of extreme drought.

Page 13, Line 23 - I would say not strongly correlated. "Independent" is too strong of a word here.
We will make the suggested terminology change.

Page 22, Line 16 - I would stay km-scale instead of small scale. That term is fairly nebulous and means different things to different communities. For example, small scale in watershed hydrology is cm to meter scale.
We agree that km-scale is more precise and will revise to clarify.

References Cited

Maxwell, R. M., & Condon, L. E.: Connections between groundwater flow and transpiration partitioning, *Science*, 353(6297), 377-380, 2016.

Maxwell, R. M., & Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under climate change. *Nature Geoscience*,1(10), 665-669, 2008

Kenny, F., Matthews, B., & Todd, K.: Routing overland flow through sinks and flats in interpolated raster terrain surfaces. *Computers & Geosciences* 34(11), 1417-1430, 2008

Markovich, K.H., Maxwell, R.M. and Fogg, G.E., Hydrogeological response to climate change in alpine hillslopes. *Hydrological Processes*, 30(18), 3126-3138, 2016

Pribulick, C.E., Foster, L.M., Bearup, L.A., Navarre-Sitchler, A.K., Williams, K.H., Carroll, R.W. and Maxwell, R.M. Contrasting the hydrologic response due to land cover and climate change in a mountain headwaters system. *Ecohydrology*, 9(8), 1431-1438, 2016

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., Tewari, M., Barlage, M., Dudhia, J., Yu, W. and Miller, K. High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *Journal of Climate*, 24(12), 3015-3048, 2011