Interactive comment on "Quantifying the Impacts of Compound Extremes on Agriculture and Irrigation Water Demand" by Iman Haqiqi et al.

Comments and Responses to Anonymous Referee #1 *(Reviewer comments in italics)*

Comment: The paper provides a novel approach to quantify the compounding effects of soil moisture and heat stressors on crop yield in the US over a historical time period. The study investigates multiple statistical representations to try to tease out the importance of the interactions between heat stress and soil moisture conditions on crop yield, and takes advantage of a large scale hydrologic model to extract the necessary soil moisture data to build the various models. The paper looks technically sound, and the paper is a great contribution to the literature.

We would like to thank the referee for his/her helpful comments that helped to improve the manuscript. We have revised the paper accordingly and provided overall and specific answers below. Also, many thanks for the positive feedback on the technical details and the significance of the paper.

As the majority of the comments are around the organization of the paper, we have revised the flow of the paper and transitions within the sections. We have dropped the sections identified less relevant by the referees. This has resulted in a substantial re-ordering of the material presented, and these changes have substantially shortened the paper as requested by the reviewer. Now, the paper is focused on the main messages. The manuscript introduces the problem by stating the research gap as "current statistical models of crop yield prediction ignore the compound extreme". And we establish the discussion around the main finding that "statistical models ignoring compound hydroclimatic extremes will significantly underestimate the yield response to water in hot days while they will significantly overestimate the yield response to water in moderate days". The referee's comments also helped us identify the unclear terms and less critical ideas. They helped us to improve the cohesion of the writings by providing clarifying definitions for unfamiliar terms and by removing the ideas not critical for the argument. The background information has been moved to the Supplementary Materials. We have also clarified the methods, moved some parts of the appendix to the text, and moved some parts of the Methods section to the Supplementary. These are major changes:

Introduction: We have included some of the text from the section "Empirical concerns" to provide adequate background on the models and metrics of individual and compound hydroclimatic extremes for predicting corn yields. We limited the text on the state of the art in the statistical prediction of corn yields to highlight current shortcomings. We kept the text on the description of the objectives to give a clear view of the originality of the research. We have removed the sentences more relevant to the Results and Conclusion.

Empirical concerns: A shortened version of this section has been merged into "Methods" and "Introduction" sections as follows. The sentences regarding the Schlenker and Roberts (2009) model are moved to the Methods section making the base for our model with individual extremes. The sentences regarding spatial aggregation are removed, we only kept our method for spatial aggregation in the Methods section. The sentences regarding average versus extreme metrics of water availability are moved to the introduction as they show the shortcomings in the current literature and how we are going to address them in the paper. The sentences regarding "interaction of soil moisture and heat" are shortened, rephrased, and moved to the introduction as they are base for our arguments about compound extreme. We have also clarified the meaning of the statistical term "interaction" when it first

appeared in the manuscript. Finally, the sentences regarding measurement errors and endogeneity concerns are moved to Supplementary.

Methods: This section has some minor changes. We re-order the sub-sections introducing the data before the models. Also, technical terms are described including the "panel fixed effect" method, "daily interaction of heat and soil moisture", and "conditional marginal impact". Figures 1-3 are improved to support definitions and methods.

Results: The results from Model 1 (individual extremes) and Model 2 (compound extremes) have not changed. However, we added a couple of sentences to provide a comparison with previous studies. We added two critical subsections here. A new sub-section on "Model comparison" compares the performance of each model in predicting yields and to illustrate why we have estimated different models with different assumptions and different water metrics. It clearly shows the advantages of using a model with compound extremes. Also, a new sub-section on "Robustness checks" describes why we do these checks and what we learn. Figures 4-6 are moved to the Results section with more details.

Discussion: This section is substantially shortened. We dropped contents about methods and results. The section on "implications for climate studies" and the related text is dropped. The section on "implications for irrigation water demand" and the related text is dropped. Based on our findings we argue that "As we find that the coefficient on extreme heat is significantly different when considering soil moisture, it is possible that previous statistical studies have over- or under-estimated the yield impacts". The revised Discussion section is provided below.

In the following sections, we offer detailed responses to each comment.

Comment: However, the paper could almost be cut in half to get the key messages of the paper, and much of the text can be either moved to supplementary materials or completely omitted. For example, I would suggest moving the first 5 figures to supplementary materials. I struggled with the flow of the ideas and text, and there is quite a bit of redundancy, and unnecessarily verbose. I would recommend major revisions, with most of the efforts on rearranging and streamlining the flow of the paper.

Overall response: Thank you for these excellent suggestions. These comments helped us to improve the organization of the paper. To minimize redundancies and maximize audience engagement, we re-organized the manuscript. We omitted the less relevant parts in order to focus on the main message. This has resulted in a substantial re-ordering of the material presented, and substantially shortened the paper.

Comment: "the paper could almost be cut in half… much of the text can be either moved to supplementary materials or completely omitted"

Regarding the length of the paper, we have shortened the paper substantially from 52 pages (around 19,000 words) to 29 pages (around 10,000 words).

Comment: "there is quite a bit of redundancy, and unnecessarily verbose"

Thanks for this comment that helped us to improve the flow of the paper and the cohesion of the writings. We have revised the organization of the paper. The flow of the Introduction section has been

revised as you will see from the following responses. We have omitted the contents related to the conclusion, discussion, and summary from the Introduction. The Discussion section has been revised substantially as you will see below. We have omitted the equations, methods, and results type of content from it. In the revised version, we have focused on the main message. We have revised the flow of the paper focusing on the significance of compound extreme metrics and their advantage over the individual extreme metrics.

Comment: "I would suggest moving the first 5 figures to supplementary materials"

Regarding the figures, we have dropped panel b from figure 1. We also moved figure 3 to the Supplementary. However, Figure 2 is revised to illustrate the critical concepts and definitions necessary for this study. Figures 2 is central to our proposed method for constructing the soil moisture variable. We wanted to show the importance of introducing the metrics based on deviation from normal in figure 2; and the heterogeneity of mean soil moisture across space in figure 4. Figure 5 is illustrating critical results in rejection of the hypothesis that precipitation and soil moisture are the same metric for statistical studies. Below we illustrate the revised figures.

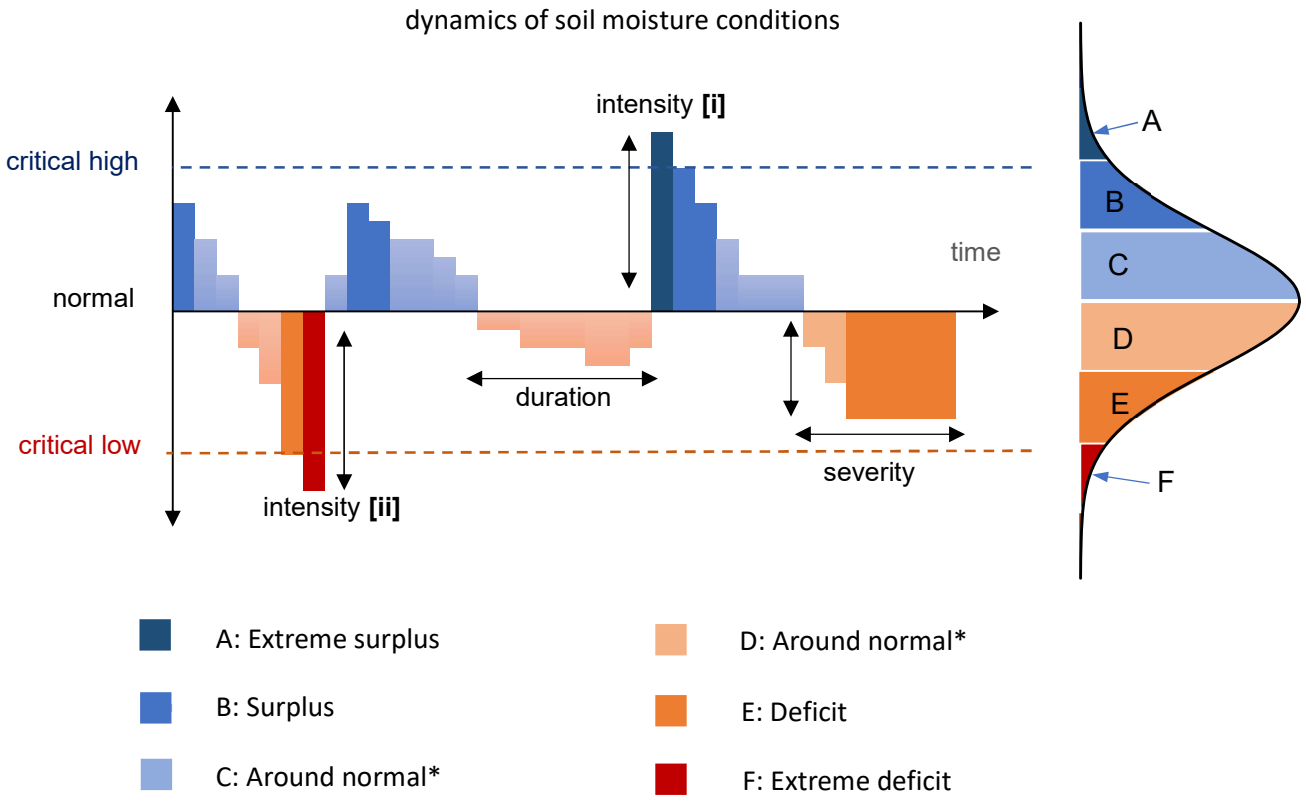

Figure 1. Soil moisture dynamics within a typical growing season. Some soil moisture conditions can be harmful to crops including excess wetness [i], moisture stress intensity[ii], duration of moisture stress [iii], and severity of soil moisture stress [iv]. Normal level of soil moisture is defined as the historical average of volumetric soil moisture within the growing season.

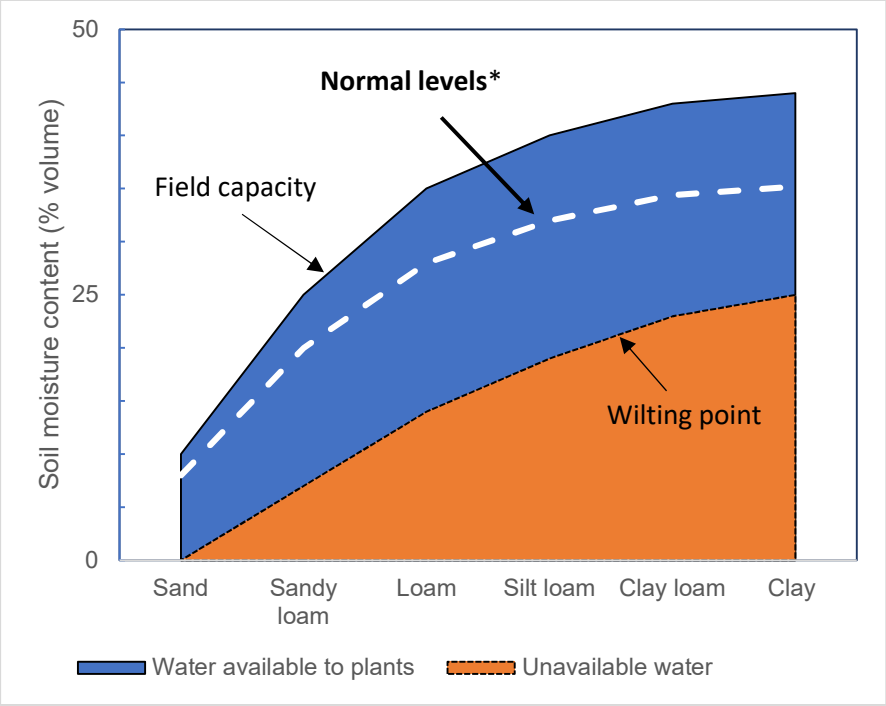

Figure 2. Soil texture affects normal moisture levels. The sandy soil has the lowest normal level while the clay has the highest normal levels.

Specific comments:

Lines 1-2: I would suggest changing the title to something like "Quantifying the compounding effects of soil moisture and heat on crop yield" The paper does not talk about the impacts on irrigation water demand, and none of the figures show results looking at water demands.

We agree that this section may disrupt the flow of the paper. To improve the flow of the paper and to focus on the main message, we decided to follow the reviewer's suggestion and eliminate the "irrigation demand" section.

Line 12: Are high-resolution and fine-scale intended to mean different things?

No. We did not mean different things by using these two separate terms. In the revised version, we only use "fine-scale" throughout the paper to avoid confusion.

Introduction: The introduction section needs better arrangement for better flow of the ideas. I would recommend focusing on the importance of this work, what is the current state of the knowledge in this space, how this differs or builds on previous efforts, the key novelties it adds to the field, and the specific science/research questions it is trying to tackle. All of this is pretty much there, but it needs to flow better, and certainly results should be omitted from the introduction section to avoid redundancies.

Many thanks for highlighting the relevance of the work. This section has been shortened and re-written according to these recommendations. We have also removed the results and summary contents from the Introduction section. We have included some of the text from the section "Empirical concerns" to provide adequate background on the models and metrics of individual and compound hydroclimatic extremes for predicting corn yields. We limit the text on the state of the art in the statistical prediction of corn yields to highlight current shortcomings. We also kept the text on the description of the objectives to give a clear view of the originality of the research.

Lines 22-32: this section is redundant with some of the content of the abstract and talks about the approach and key messages before even articulating the importance of the work. I would recommend omitting.

We have omitted these lines.

Lines 57-62: "We show that the coefficient . . ." Avoid throwing results in the middle of the introduction section to avoid redundancy. I would suggest omitting.

We have omitted these lines.

Line 79: it is a bit weird to talk about concerns before even talking about the approach.

This section has largely been moved, with key items moved to the Methods section and some are moved to the Introduction.

Line 82: spell out Sec for consistency sake.

According to the Manuscript Preparation Guideline, "the abbreviation 'Sect.' should be used when it appears in running text and should be followed by a number unless it comes at the beginning of a sentence". However, we make them consistent by putting them all at the beginning of the sentences here. We have also removed many of the Section references, as they are not needed.

Line 83: Do you mean "background: Key factors impacting yield" or something along that line?

This section has been moved to the Supplementary Materials and other relevant sections. This type of section is standard in the econometric literature from which the methods are mainly derived, but we agree it does not fit in the flow of the paper here.

Line 85: "before starting our discussion" please rephrase.

This section has been revised and moved to the Methods section.

Line 94: "we will briefly talk about" please rephrase

This section has been removed.

Lines 97-101: I would suggest omitting this paragraph. Water is discussed in section 2.3, and here the focus is on spatial aggregation.

These lines have been omitted, with key details succinctly described in the methods section.

Line 100: a sample of what? The sentence is somewhat vague.

This paragraph is now eliminated.

Lines 105-110: "we construct our. . ." this reads like a methodology section and should be part of section 3.

We removed this part in the revised version. The data construction is explained in the Methds.

Line 111: "another empirical challenge" This reads like you are talking about a different challenge than what was discussed in the above two paragraphs under 2.1. I would suggest separating this section as '2.2 Degree of temporal aggregation' and keeping the previous sub-section on the spatial aspect only.

Omitted, with key details succinctly described in the methods section.

Line 130: if you are going to end of each challenge with how this study tackles this challenge or differs from previous efforts, then I would suggest that this is done here as well, and at the end of each of the other challenges discussed in section 2.

Omitted, with key details succinctly described in the methods section.

Line 144: "To undertake. . ." please omit sentence. It does not add much.

This sentence is omitted in the revised version.

Lines 145-147: Omit. I would suggest not throwing results at this stage. Plus, the reader does not know anything about WBM yet.

These lines are omitted in the revised version.

Line 166: "a fixed effect panel regression" I am not sure what that means. Please explain. Also, in the following sentence, what coefficients are you referring to? Please be specific.

Thanks for raising the need for clarification about this method. We added a brief description of the fixed-effect panel regression. Also, we removed this term in any text before the methods section. The following sentence is added in the description of variables in the Model (1)

> "The fixed effect variable (also termed the unobserved individual effect) allows us to control for other biophysical or economic characteristics of each location which are not varying over time and can potentially explain the yield differences between counties."

Also, we added the following in the estimation strategy section:

> "A panel fixed-effect approach is a statistical method for analyzing two-dimensional (e.g. time and location) panel data. This method is helpful for analyzing data that is collected repeatedly for the same locations over time with a relatively short time span (Wooldridge, 2016). As our data set contains information for counties over time, a panel data analysis is appropriate. In addition, a fixed-effect model is appropriate as there are unique biophysical and economic attributes of counties that can explain yield differences across counties and are not changing over time. When we conduct a statistical test (Hausman test), it rejects the random effects model in favor of the fixed effect models we use."

References:

Wooldridge, Jeffrey M. *Introductory Econometrics: A modern approach*. Nelson Education, 2016.

Line 182: rephrase "as we prefer to take care of. . ."

This sentence is omitted, reworded in the methods.

Line 189: this section needs a concluding sentence to connect the dots.

The section has been dropped from the revised manuscript.

Line 217: having read through section 2, it leaves the reader wondering what all of this has to do with compounding extremes. I wonder if section 2 can be shrunk or moved to a later section after describing the method section of the paper to improve the flow of the paper.

Thanks for your suggestion. To improve the flow of the paper, we have shortened the content of this section and moved much of the material to the Supplementary, Methods, or other relevant sections. Here are some of the major changes:

Line 84-92: shortened and moved to the Methods.

Line 93-116: omitted.

Line 117-126: shortened and moved to the Introduction.

Line 127-130: moved to the Methods.

Line 131-147: shortened and moved to the Introduction.

Line 148-163: shortened and moved to the Introduction.

Line 164-171: shortened and moved to the Introduction.

Line 172-177: shortened and moved to the Methods.

Line 178-189: shortened and moved to the Introduction.

Line 190-217: omitted.

Line 219: "we introduce two models." What kind of models? Please specify.

Specified now as "statistical models"

Lines 219-221: why design the two models in this manner? Explain the logic.

A brief introduction to why these two models are used has been added to the beginning of the Methods section. Here is the related text:

> "Model 1 assumes the impacts of heat and water on corn yields are separable. This model considers metrics of individual extremes (heat stress and water availability). … Within this framework, we investigate which indicator of individual extremes is a better predictor of corn yields. Relaxing the separability assumption, model 2 assumes the yield impacts of heat and water are mutually interdependent. Model 2 considers indicators of compound extremes."

Line 226-227: "in summary," omit. You just started talking about the model here.

Omitted.

Line 229: "as reported by WBM" omit since the reader has not read about WBM yet unless you go with my recommendation to have section 3.3 moved to 3.1 as explained later.

The sentence is omitted. We have also moved section 3.3 to 3.1 following your suggestion.

Lines 227-233: these equations (1a-d) need to be shown here. They are core to the whole paper and deserve more attention in the paper.

We have discussed the models in the estimation strategy. In the revised version, the following is added:

"Considering the exposure to each temperature interval to capture the marginal impact of heat and water on crop yields, we estimate the following for model (1-a):

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \delta_a P_{it} + \delta'_a P_{it}^2 + \lambda_s t + \lambda'_s t^2 + c_i + \varepsilon_{it} \quad (1)$$

where *i* is an index for counties, *t* is the index of time, *s* is the index for states, y_{it} is the log corn yields, D_{it} represents growing degree day variables, *P* shows cumulative precipitation over the growing season, *t* shows the time trend variable (*t* = year – 1950), c_i is a time-invariant county fixed effect, ε is the residual, and *α, β, δ, λ* are the regression parameters showing the marginal impacts. The subscript *a* is used to show the water coefficients (*δ*) are related to metrics in Model (1-a).

To evaluate the importance of soil moisture metrics in Model (1-b), we estimate the following:

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \delta_b M_{it} + \delta'_b M_{it}^2 + \lambda_s t + \lambda'_s t^2 + c_i + \varepsilon_{it} \quad (2)$$

where the variables are defined as Model (1-a) except for the water availability metric. Here *M* shows the seasonal mean soil moisture index calculated as average daily root zone soil moisture from the first day of April to the end of September. The subscript *b* is used for *δ* to distinguish the water coefficients in Model (1-b).

For Model (1-c) we estimate the following model:

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \delta_c N_{it}^{def} + \delta'_c N_{it}^{sur} + \lambda_s t + \lambda'_s t^2 + c_i + \varepsilon_{it} \quad (3)$$

where we replace seasonal mean or cumulative metrics with two new metrics to control the impacts of water extremes on corn yields. Here, N^{def} is the number of days that soil moisture is under 25 mm below normal levels (deficit); and N^{sur} is the number of days that soil moisture is higher than 25 mm above normal levels. The rest of the variables are defined as Model (1-a). The subscript *c* shows δ_c is specific to Model (1-c).

Finally, we estimate the following equation for Model (1-d):

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \delta_d M_{it}^{pos} + \delta'_d M_{it}^{neg} + \lambda_s t + \lambda'_s t^2 + c_i + \varepsilon_{it} \quad (4)$$

where M^{pos} is a cumulative measure of positive soil moisture deviations compared to the normal levels (equivalent to A+B+C in Figure 1). And M^{neg} is the cumulative measure of negative soil moisture deviations compared to the normal levels (equivalent to D+E+F in Figure 1). The subscript *d* distinguished estimated *δ* from previous models."

Lines 225-234: are metrics, indicators, and water variables the name thing here?

In the revised paper we only use "water metric" when writing specifically about methods used in this paper.

Line 243: I would suggest making this sub-section (3.3 Data) as the first sub-section in the methods section for better flow. Sub-section 3.4 builds nicely on what's covered under the first two subsections, and the data comes in the middle and breaks the flow.

We have moved Data to Section 2.1 (first methods section)

Line 250: "Daily interaction" how is this defined or calculated? Is this a term in equation 2? If so, then please state so.

We added more details on the daily interaction. Here is the text in the estimation strategy discussion:

> "For Model (2), we consider the daily interaction of heat and soil moisture as the compound metric. The interaction term is defined when the marginal impact of an explanatory variable depends on the magnitude of yet another explanatory variable (Wooldridge, 2016). Here, the marginal impact of heat on yield depends on water availability; also, the marginal impact of water on yield depends on heat. This is called conditional marginal impact."

References:

Wooldridge, Jeffrey M. *Introductory Econometrics: A modern approach*. Nelson Education, 2016.

Lines 261-265: omit this paragraph. It is identical to lines 249-255.

Omitted.

Line 231: Was there any validation work done on the soil moisture data? I am not necessarily asking for that to be shown here, and rather some citations of the previous validation work using WBM should suffice.

Previous work that uses WBM in an agricultural context is provided in the following references (Grogan, 2016; Grogan et al., 2017; Wisser et al., 2008, 2010):

Grogan, D.: Global and regional assessments of unsustainable groundwater use in irrigated agriculture, Doctoral Dissertations [online] Available from: https://scholars.unh.edu/dissertation/2, 2016.

Grogan, D. S., Wisser, D., Prusevich, A., Lammers, R. B. and Frolking, S.: The use and re-use of unsustainable groundwater for irrigation: a global budget, Environ. Res. Lett., 12(3), 034017, doi:10.1088/1748-9326/aa5fb2, 2017.

Wisser, D., Frolking, S., Douglas, E. M., Fekete, B. M., Vörösmarty, C. J. and Schumann, A. H.: Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets, Geophysical Research Letters, 35(24), doi:10.1029/2008GL035296, 2008.

Wisser, D., Fekete, B. M., Vörösmarty, C. J. and Schumann, A. H.: Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network-Hydrology (GTN-H), Hydrology and Earth System Sciences, 14(1), 1–24, 2010.

Line 304: are you using a single scan, or are you capturing the evolution of the cropland over the historical time period?

The cropland data product, the Crop Data Layer (CDL, USDA NASS, 2017), is an annual time series of cropland area. This captures the evolution over time. We have revised the sentence to the following:

> "We employed the Crop Data Layer from the US Department of Agriculture to exclude grid cells with no cropland and to aggregate the grid cell information to the county level (Boryan et al., 2012; USDA-NASS, 2017)."

Boryan, C., Yang, Z. and Di, L.: Deriving 2011 cultivated land cover data sets using usda national agricultural statistics service historic cropland data layers, in 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 6297–6300, IEEE., 2012.

USDA-NASS: USDA-National Agricultural Statistics Service, Cropland Data Layer, United States Department of Agriculture, National Agricultural Statistics Service, Marketing and Information Services Office, Washington, DC [online] Available from: http//nassgeodata. gmu. edu/Crop-Scape, 2017.

Line 319: I would suggest moving the materials here to be merged with subsections 3.1 and 3.2. For example, I would suggest moving lines 320 to 331 to appear in line 234. This would mean deleting the sentence "the estimation strategy is described in section 3.4." Similarly, I would move lines 232 through 361 to line 242.

The methods section has been substantially reorganized, including changing the order of how the data, model equations, and estimation strategy are described. We have also omitted cross-references to sections.

Line 363: "This section provides estimation results for different representations of Model (1)" well the authors discuss results from Model (2) as well (starting around line 389).

This sentence has been removed. The results section is re-organized to focus on the main results and to improve the flow of the paper. Here is the new order:

> 3.1. Model (1): predicting yield responses to individual extremes
>
> 3.2 Model (2): predicting yield responses to compound extremes
>
> 3.3 Model comparison
>
> 3.4 Decomposing the variation in US corn yields
>
> 3.5 Robustness checks

Line 384: so what does all of this mean? Which is the 'best' model formulation, on what basis, and how does this compare with previous findings?

We have added a section on the model comparison, and section titles have been added for clarity. The performances of the models are compared based on AIC and BIC. While R-squared is not necessarily the best measure for model comparison, we have reported it for interested readers.

"A comparison of model performance metrics is given in Table 5, along with a description of the water metric and the extreme metric used in each model. We find that for Models 1b-d and Models 2a-d the coefficients on the soil moisture metrics are significant and with expected signs. Comparing the models' performance suggests that Model (1-b), with mean soil moisture, performs better than the Model (1-a), with cumulative precipitation. Also, Model (1-d), with the extreme soil moisture metrics, outperforms all previous models (with cumulative precipitation or with mean soil moisture). The best corn yield predictor is from Models (2-a) and (2-b), considering compound extremes through the daily interaction of heat and soil moisture."

Line 387: "the deficit and by 2300" – delete 'and'

Deleted.

Line 397: "The figure shows that Model (1) would. . ." It was not clear that the intent was to compare the two models (1 and 2) to get at the compounding aspect. Some articulation of that upfront would help the reader follow through.

Many thanks for your comment which helped us focus on the central finding of the paper. As this is a significant point, we have talked about it at the beginning of the Results section. We have added the following:

"Here we describe the regression results from each individual model, and compare their performance to identify which metrics are important to include in the statistical estimate of corn yields. The central finding is that measures of soil moisture extremes are statistically significant, and models including intensity, duration, and severity metrics (as illustrated in Fig. 1) better capture both mean and year-to-year variation in U.S. corn yields. This point is illustrated in Figure 7, which compares Model 1 (a-d range) to Model 2a: each model estimates the percentage change in corn yields assuming an additional 10 degree-days above 29˚C and no change in mean soil moisture. The figure shows that Model (1) would significantly underestimate the damage for conditions with extreme water surplus or extreme water deficit."

Line 44: "from previous models" which models are you referring to? 1a,b,c,d, 2a?

Thanks for pointing to this issue. We have clarified the sentence as:

"This is not significantly different from previous models (1-a, 1-b, 1-c, 1-d, and 2-a)."

Lines 409-422: I would suggest moving this to be part of the results section. And to change the title for section 5 to be simply "Discussion" and then to jump to 5.1 directly.

Thanks for your suggestion. This section is shortened and moved to the "Model comparison" subsection of the Results.

Lines 416-420: please expand on this section to explain what you found out from these additional analyses that are in the appendix. Currently, they come across as throw away sentences.

We have moved these sentences to the "Robustness check" as a subsection of the Results.

> "The Supplementary Materials provide several robustness checks. The goal is to investigate whether different assumptions can improve the predictive power of Model (1) such that it outperforms Model (2). We answer three questions. First, are the estimation results from Model (1) different from those using alternative water metrics from WBM output? Second, are the estimates in Model (1) different from those obtained using a model considering growth stages? And third, do the main findings change if we alter the geographical scope of the study?
>
> For the first robustness question, alternative water metrics, we re-estimate Model (1) using daily evapotranspiration (which is related to the water requirements of plants) and soil moisture fraction (soil moisture content divided by field capacity). Overall, the findings remain robust to alternative soil moisture metrics from WBM including the mean of soil moisture fraction, the seasonal mean of evapotranspiration as well as within season standard deviation of them. We also look at the results using an alternative interpolation of WBM data to PRISM resolution (nearest neighbor versus bilinear interpolations). We reject the null hypothesis that the coefficient on yield response to heat is different between these two metrics. Also, we reject the null hypothesis that the prediction power across these models is higher than Model (2).
>
> To test the second robustness question, time separability, we re-estimate Model (1-b) for two-month intervals (Apr-May, Jun-Jul, Aug-Sep), and the findings remain robust. We find that considering bi-monthly variables does not change the yield response to heat. Although this alternative formulation does improve the predictive power of Model (1-b) a little bit, the performance is not better than the original Models (2-a) and (2-b) with compound extremes.
>
> To test the sensitivity of our findings to the geographical area, we re-estimate the models for the Eastern US and the Western US. We find that the estimated coefficients of Models (1-a) and (1-b) are not robust to the geographical choice, while those of Model (2) remain robust."

Lines 420-422: "Finally, we have provided . . ." I would omit these two sentences.

The sentence is omitted, and the whole section is re-organized.

Line 432: "we recommend the use of soil . . ." I can't tell if this recommendation is based on the findings in this study, or simply an opinion based on past efforts/studies. Please clarify.

We have shortened the discussion section and focused on the main messages and central findings. This is the revised Discussion:

> In this paper, we have identified new water availability metrics that improve the predictive power of statistical corn yield models. While predictive power is an important outcome of this

analysis, the insights gained from incrementally adding higher temporal-resolution metrics of water extremes to the models are also valuable for understanding the drivers of corn yield variability, and for revealing the resolution of water availability data required to capture future extremes under climate change scenarios. Statistical crop models have been used to both elucidate drivers of crop yield trends and variability, and to evaluate potential climate change impacts on crop production in the future (e.g., Lobell and Burke, 2010; Diffenbaugh et al. 2012). However, these models typically use seasonally averaged water availability metrics (e.g., total growing season precipitation), and utilize precipitation more often than soil moisture. Generally, if the location of the study does not expect a significant change in the within-season distribution of the soil moisture, a mean soil moisture index will work. However, if there is an expected change in this distribution, using the mean variable will create biased yield projections. Because climate models project significant changes in the frequency and intensity of both extreme precipitation and temperature (Myhre et al., 2019; Zscheischler et al., 2018; Manning et al., 2019; Bevacqua et al., 2019; Poschlod et al., 2020; Potopová et al., 2020; Wehner, 2019), the results presented here show that the mean metrics of water availability – especially mean precipitation - are not sufficient to capture the impacts on yields. It is necessary to consider the metrics of extreme events as illustrated in Figure 1. As we find that the coefficient on extreme heat is significantly different when considering soil moisture, it is possible that previous climate impact studies have over- or under-estimated the yield impacts. Further, farm management practices can alter soil moisture – and therefore yields – independent of precipitation. Supplemental irrigation, as well as no-till farming, cover cropping, and soil conservation, can increase soil moisture. These adaptations may occur in places predicted to face higher mean precipitation coupled with more extreme water events. The results of these management practices cannot be captured by statistical models looking at precipitation metrics alone. Such precipitation-based studies could potentially lead to over-estimation of yield damages under future climate extremes by not accounting for human adaptations designed to conserve soil moisture.

Line 454: "Model (2)" 2a or 2b?

This part is omitted in the revised version.

Line 454: "while Model (3) predicts ..." do you mean model 1 here?

This part is omitted in the revised version.

Lines 484-499: Subsection 5.4 comes as a surprise to the reader. It also reads more like a methods section. I would suggest dropping this subsection.

Thanks for your comment. To improve the flow of the paper, we followed your suggestion. This part is dropped in the revised version.

Lines 500-509: The first sentence is redundant. The subsection is relatively shallow as compared to previous subsections. Also, it is not clear if there is any conclusion that can be drawn from Figure 10. I wonder if this would fit better if moved to the results section instead of being a discussion subsection.

The redundant section has been removed, and the remainder has been moved to the results. We have also clarified the methodology.

> "To show the significance of weather variation for crop yields, we estimated the historical impacts of heat and water using Model (2-a). The trend is estimated assuming no variation in heat and water availability. Then, we predicted the impact of heat on yields considering observed variation in heat and assuming normal soil moisture. Finally, we predicted the yield considering observed variation in heat and simulated variation in soil moisture. The residual is not reported."

Lines 510-536: Subsection 5.6 is another big surprise to the reader. I was not expecting this as this was never baked in the framing of the paper in the initial sections. All the previous sections including the data subsection focused on the US. Though this extends the work globally, which begs the question of how the extrapolation was done. I would suggest omitting this from this paper and keeping it for a follow-on paper.

Thanks for your suggestion. This section is omitted in the revised section.

Figure 6: what are the units for the y-axis and for the color bar on the far right?

Thanks for pointing to the missing units. It is the ratio of soil moisture to normal soil moisture. We have corrected this in the revised version.