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

Comments and Responses to the Editor:

Authors' comment:

We would like to thank the Editor for his helpful comments and considerations. We have revised the paper according to the referees' suggestions and provided overall and specific answers separately. Regarding Fig. 2, we agree that it provides standard textbook material. We have revised it to show the importance of introducing the metrics based on deviation from normal. The revised figure is moved to the Supplementary material. The followings are:

1- Comments and Responses to Anonymous Referee #1 (pages RC1-1 to RC1-15)
2- Comments and Responses to Anonymous Referee #2 (pages RC2-1 to RC2-17)
3- Track-change version of the manuscript (pages 1-66)

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 28 pages (around 10,000 words). Around 3,000 words are moved to the supplementary.

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 figures 3 to the Supplementary. Figure 2 is also moved to the supplementary with revisions to illustrate the critical concepts and definitions necessary for this study. Figures 4 and 5 are important. We wanted to show 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.

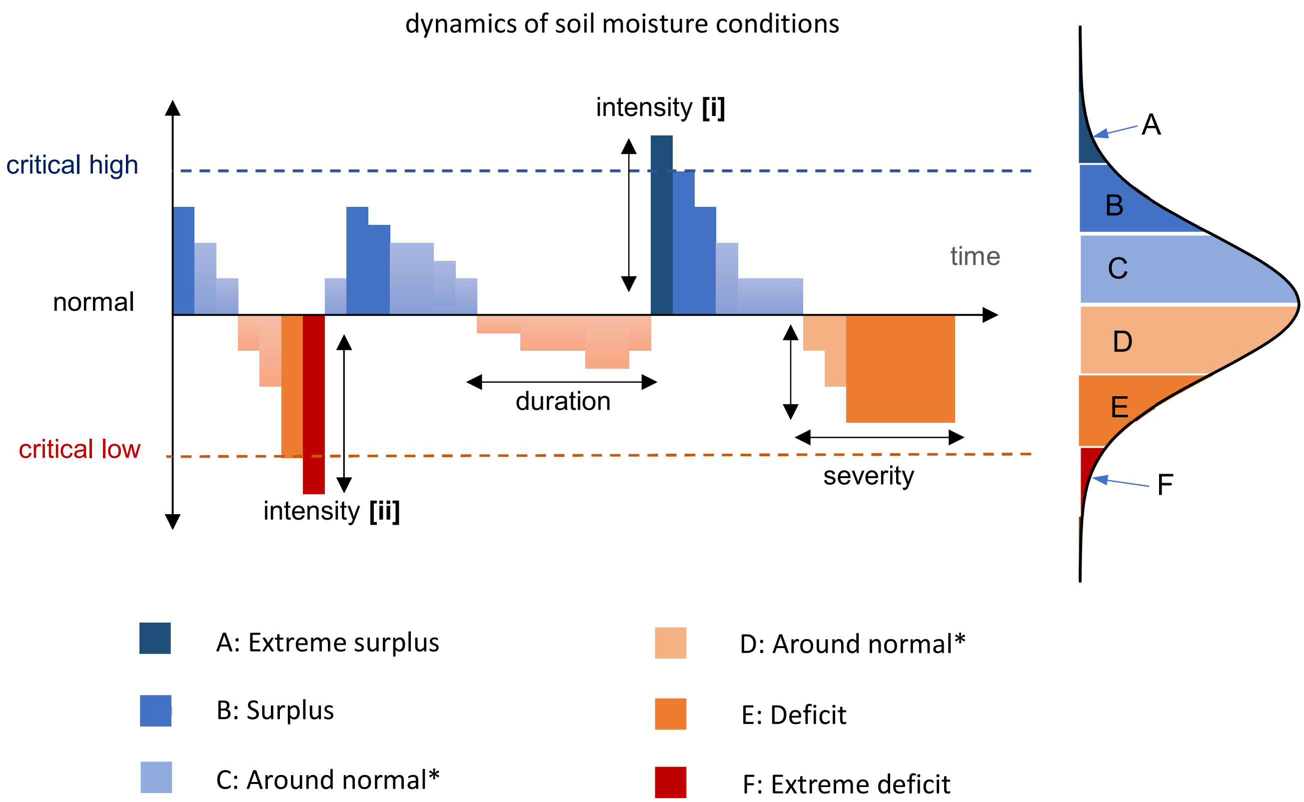


Figure 1. Soil moisture dynamics within a typical growing season. Some soil moisture conditions can be harmful to crops including excess wetness [i], moisture stress intensity[ii], duration of moisture stress [iii], and severity of soil moisture stress [iv]. Normal level of soil moisture is defined as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

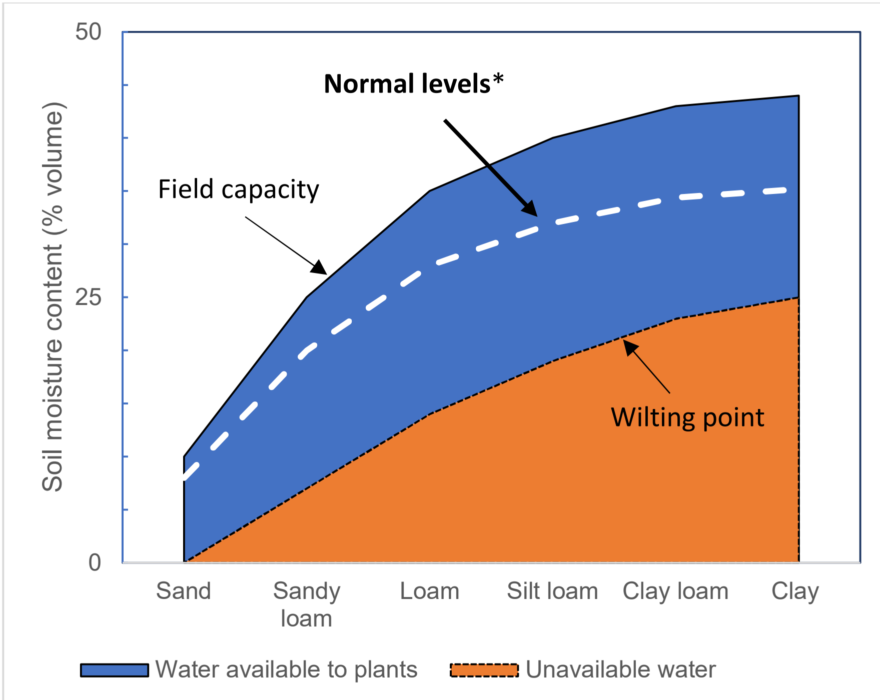


Figure 2. Soil texture affects normal moisture levels. The sandy soil has the lowest normal level while the clay has the highest normal levels.

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.

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 #2 *(Reviewer comments in italics)*

Overall: *The paper is about a well-designed study aiming to elaborate individual and compound extreme event impacts on corn yields in the USA using statistical approach. The significance of extreme events on yield anomalies were studied using various indicators of soil moisture (representing water stress) as well. The outcomes of the paper can be insightful for further studies of predicting crop yield anomalies and assessing impacts of extreme weather conditions to crop yields. Consequently, the paper is worth for publishing with some revisions.*

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.

My major comments on the paper are:

1- *The paper needs to be re-structured/re-written. First, it is too lengthy including textbook information (e.g. Figure 1b, and Figure 2) which are not necessary for the reader (peer knowledge). Second, its structure is chaotic: the introduction chapter includes results and discussions points etc; it is like a short summary of the whole paper; the discussion section includes equations, methods, results and data sources. The authors claim to include results/conclusions which are too diverse and out of scope of the analysis (e.g. irrigation, farm soil management, marginal value, decision making as specified in the abstract). The framework of analysis do not support to make conclusions about these topics. The authors should revise their goals and associated conclusions accordingly. The paper is about compound vs individual extreme events on crop yield and comparison of different soil moisture indicators. Other conclusions not taken from this analysis can be excluded. Furthermore, the empirical concerns are relevant however too lengthy for readers. It can be reduced and can be removed to SI.*

Overall response: Thanks for these excellent suggestions. These comments helped us to improve the organization of the paper. To minimize redundancies and maximize the 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: “it is too lengthy”

Response: 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: “including textbook information (e.g. Figure 1b, and Figure 2) which are not necessary for the reader (peer knowledge)”

Response: Regarding the textbook information, we have dropped panel b from figure 1. Figure 2 and 3 are revised to illustrate the critical concepts and definitions necessary for this study. We have moved figures 2 and 3 to the supplementary.

Comment: “*the introduction chapter includes results and discussions points etc; it is like a short summary of the whole paper”*

Response: The flow of the Introduction section has been revised as you will see from the following responses. We have omitted the contents related to conclusion, discussion and summary from the Introduction. The first paragraph and the last paragraph are omitted too.

Comment: “*the discussion section includes equations, methods, results and data sources”*

Response: 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.

Comment: “*The authors claim to include results/conclusions which are too diverse and out of scope of the analysis (e.g. irrigation, farm soil management, marginal value, decision making as specified in the abstract). The framework of analysis do not support to make conclusions about these topics. … Other conclusions not taken from this analysis can be excluded.”*

Response: We agree that some of the discussions required further details and their relevance to the main message were not well-defined. Hence, we have focused on the main message and omitted the discussions about marginal value, farm soil management, supplemental irrigation. Below we have included the shortened and revised Discussion section.

Comment: “*The authors should revise their goals and associated conclusions accordingly. The paper is about compound vs individual extreme events on crop yield and comparison of different soil moisture indicators.”*

Response: Thanks for this very helpful comment. 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: “*Furthermore, the empirical concerns are relevant however too lengthy for readers. It can be reduced and can be removed to SI”.*

Response: Thanks for highlighting the relevance of this material. The content of this section is shortened and moved to SI and other relevant sections. Below, we will describe the changes in more details.

2- The authors claim that "marginal value of water" will be calculated and utilized in the paper. There is nothing about it in the method and result section (only shown in the discussion section – a short paragraph without any substantial info). I think having this goal of economic analysis is not relevant and beyond the scope the. It is better to exclude this part of the analysis so that the paper is coherent and consistent with its framework.

It is true that the paper does not provide details on the implications for irrigation water demand. While the paper could potentially talk about economic and agronomic water demand, it only briefly discussed the economic demand. To improve the flow of the paper and to focus on the main message, we decided to cut the "irrigation demand" section.

3- Discussion sections were boldly written (e.g. like for climate change discussion and farmer management). I recommend drawing conclusions only if it is supported by the data and analysis.

Thanks for this comment that helped us focus on the critical findings. We omitted the climate change implications. We have omitted the contents are not critical to our main message. Also, we have revised the conclusion and discussion to only draw the conclusions supported by our analysis. 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 (Diffenbaugh et al., 2012; Lobell and Burke, 2010) . 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 (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."

References:

Bevacqua, E., Maraun, D., Vousdoukas, M. I., Voukouvalas, E., Vrac, M., Mentaschi, L. and Widmann, M.: Higher probability of compound flooding from precipitation and storm surge in Europe under anthropogenic climate change, Science Advances, 5(9), eaaw5531, doi:10.1126/sciadv.aaw5531, 2019.

Diffenbaugh, N. S., Hertel, T. W., Scherer, M. and Verma, M.: Response of corn markets to climate volatility under alternative energy futures, Nature Climate Change, 2(7), 514–518, doi:10.1038/nclimate1491, 2012.

Lobell, D. B. and Burke, M. B.: On the use of statistical models to predict crop yield responses to climate change, Agricultural and Forest Meteorology, 150(11), 1443–1452, doi:10.1016/j.agrformet.2010.07.008, 2010.

Manning, C., Widmann, M., Bevacqua, E., Loon, A. F. V., Maraun, D. and Vrac, M.: Increased probability of compound long-duration dry and hot events in Europe during summer (1950–2013), Environ. Res. Lett., 14(9), 094006, doi:10.1088/1748-9326/ab23bf, 2019.

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M. and Stohl, A.: Frequency of extreme precipitation increases extensively with event rareness under global warming, Scientific Reports, 9(1), 1–10, doi:10.1038/s41598-019-52277-4, 2019.

Poschlod, B., Zscheischler, J., Sillmann, J., Wood, R. R. and Ludwig, R.: Climate change effects on hydrometeorological compound events over southern Norway, Weather and Climate Extremes, 28, 100253, doi:10.1016/j.wace.2020.100253, 2020.

Potopová, V., Trnka, M., Hamouz, P., Soukup, J. and Castraveț, T.: Statistical modelling of drought-related yield losses using soil moisture-vegetation remote sensing and multiscalar indices in the south-eastern Europe, Agricultural Water Management, 236, 106168, doi:10.1016/j.agwat.2020.106168, 2020.

Wehner, M.: Estimating the probability of multi-variate extreme weather events, in Workshop on Correlated Extremes, Columbia University., 2019.

Zscheischler, J., Westra, S., Van Den Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch, D. N., Leonard, M. and Wahl, T.: Future climate risk from compound events, Nature Climate Change, 8(6), 469–477, 2018.

For more-detailed comments:

1) Abstract

- which crops were addressed in the article? Please specify. It is important to mention corn here.

The paper is focused on corn in the US, we have added this in the revised abstract.

- "the value of water experiences a four-fold increase on hot days": not clear, what do the authors refer to by "value of water"? Is this volume? Value of water is generally associated with significance, importance, true cost etc.

This sentence is omitted from the abstract. This term was used to refer to economic value, but the related section and discussions are removed from the revised paper.

- This paper also improves our understanding of the conditional marginal value (or damage)". Which way? And what is conditional marginal value? It is important to provide necessary descriptions in the text as well.

This sentence is related to a section which is omitted from the revised paper. However, the concept of conditional marginal value has been defined in the paper. This is added in the text:

> "Marginal impact and conditional marginal impact are two statistical concepts equivalent to partial derivatives in mathematics. When the partial derivative of one variable does not depend on other variables, we use the term "marginal impact". When it depends on other variables, we use "conditional marginal impact". A conditional marginal impact shows the impact of a compound extreme. A non-conditional marginal impact can show the impact of individual extremes."

2) Introduction

- The first paragraph was written like a conclusion section (after line 26). It includes a short summary, reminding "an abstract". This part needs revision or can be completely excluded (or moved to discussion/conclusion sections).

This paragraph is excluded in the revised paper.

- Ln 33: there can be other factors affecting crop yield significantly such as soil, management, nutrients etc.

This is completely right. The word "variation" was missing. We revised the sentence to the following:

> "In agricultural production, water and heat extremes are key determinants of yield variations".

- Ln 37-38: "Other metrics of extreme water conditions", please specify.

We revised the sentence as:

> "While soil moisture is a more appropriate measure of water availability for crops, extreme water indicators based on soil moisture have been only minimally explored".

- Ln 38-39: "Current statistical studies had limited success in statistically capturing the yield response to soil moisture metrics", please explain why.

We added the following explanation:

> "There are several potential reasons for the limited success of previous statistical studies in capturing yield response to soil moisture. Direct measures of soil water availability include complex biophysical and hydrological processes that are difficult to capture in a rather simple statistical model. On the other hand, seasonal mean soil moisture is highly correlated to seasonal precipitation. Thus, including an average of soil water content may not add value to a statistical model."

- Ln 43: "the impact of climate change on soil moisture". The paper is about individual extreme response of yield vs compound. It is not clear why the authors refer to CC studies.

This is omitted. The climate change section is dropped now, so this sentence is no longer relevant.

- Ln 46: "conditional marginal impact". Please explain what this means.

See explanation above.

- Ln 50: please explain "wet-heat stress"

Wet heat stress or moist heat stress are the terms have been used in different disciplines to talk about hot and humid or moist conditions (Buzan and Huber, 2020). Soil water can exacerbate the heat stress under conditions of high humidity. This is not a prevalent condition. However, it can arise in the context of complex meteorological, hydrological, and agronomic interactions. In the US Midwest, a combination of heatwave and corn sweat can create "moist heat stress" which is dangerous for people, animals, and plants.

Reference:

Buzan, J. R. and Huber, M.: Moist heat stress on a hotter Earth, Annual Review of Earth and Planetary Sciences, 48, 2020.

- Ln 55-60: this part is an outcome of the study. Please remove it to another section (e.g. discussion).

This part has been shortened and moved to Methods and Results.

- What are exactly marginal and conditional marginal impacts? It is better if definitions are given for readers.

See explanation above.

- Ln 64-79: this part is related to discussion/conclusion. I recommend deleting these parts or move to the other relevant sections.

In order to shorten the length of the paper, this part has been removed.

- Ln 77/78: the authors claim that they will show how the results can be used to economically quantify the marginal value of water, in the form of soil moisture, for corn production in the US under different hydroclimatic conditions. I couldn't see this in the rest of the paper. Please clarify.

This topic has now been omitted as it is tangential to the main theme of this paper.

3) Empirical concerns

- This section is mostly about discussion of the method and assumptions taken for the study. It can be presented as supplementary information, rather than in the main text. That can help reader to focus on the results of the paper and its wider implications. In its current form, it is too lengthy.

Thanks for your suggestion. To improve the flow of the paper, we have shortened the content of this section and moved them 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.

- Equation 1: please describe what exactly each letter in the equation refers to? For example please refer last variable in the equation as error and describe g(h) function?

Thanks for catching this. We have added the description for the missing variables. Here, g(h) is a general function showing the yield growth as function of heat.

- Ln 126: "measure the value of water". Not clear what the authors refer to as "value of water". Please clarify.

This part is omitted in the revised version.

- Figure 1: this is nothing new, a known information– like a textbook. Excluding this figure does not change anything about the paper. I recommend not to include it.

We have dropped panel b of the Figure 1. We believe that Figure 1-a illustrates the concepts that are central to the Methods. While illustration itself might look like textbook information, it helps us to define the metrics of soil moisture extremes. To distinguish this from a common-knowledge figure, we have modified it as follows:

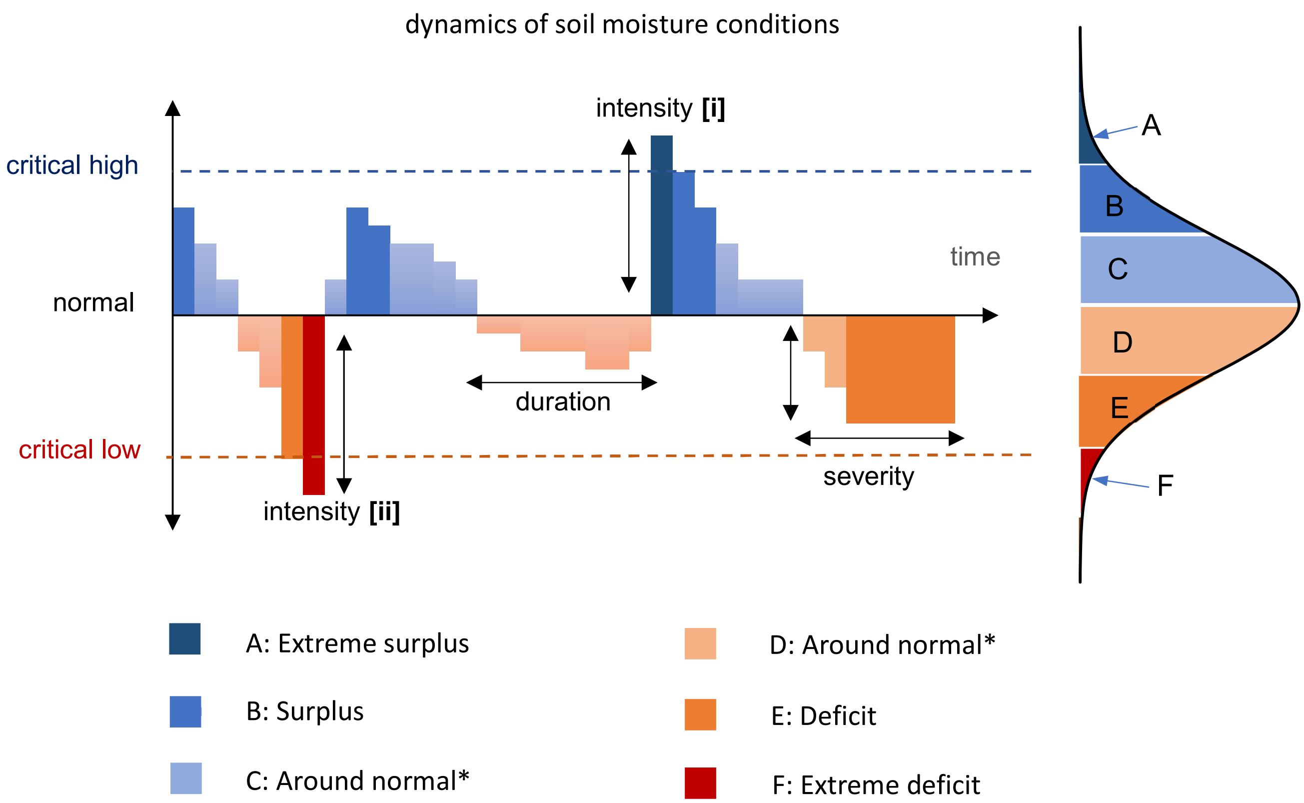


Figure 1. Soil moisture dynamics within a typical growing season. Some soil moisture conditions can be harmful to crops including excess wetness [i], moisture stress intensity[ii], duration of moisture stress [iii], and severity of soil moisture stress [iv]. Normal level of soil moisture is defined as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

- Ln 134: "Many researchers have acknowledged the need for soil moisture data to predict the response of crop yields to variations in water availability." Please provide references to those researchers.

This sentence has been omitted in the revised version.

- Ln 171: please provide references to those studies.

This sentence has been rephrased and moved to the introduction:

"It has become a standard practice either to focus on a limited geographical area (Rizzo et al., 2018; Wang et al., 2017) or to employ a proxy variable like precipitation, evapotranspiration, or vapor pressure deficit estimates (Comas et al., 2019; Roberts et al., 2013)."

References:

Comas, L. H., Trout, T. J., DeJonge, K. C., Zhang, H. and Gleason, S. M.: Water productivity under strategic growth stage-based deficit irrigation in maize, Agricultural Water Management, 212, 433–440, doi:10.1016/j.agwat.2018.07.015, 2019.

Rizzo, G., Edreira, J. I. R., Archontoulis, S. V., Yang, H. S. and Grassini, P.: Do shallow water tables contribute to high and stable maize yields in the US Corn Belt?, Global Food Security, 18, 27–34, doi:10.1016/j.gfs.2018.07.002, 2018.

Roberts, M. J., Schlenker, W. and Eyer, J.: Agronomic Weather Measures in Econometric Models of Crop Yield with Implications for Climate Change, Am J Agric Econ, 95(2), 236–243, doi:10.1093/ajae/aas047, 2013.

Wang, R., Bowling, L. C., Cherkauer, K. A., Cibin, R., Her, Y. and Chaubey, I.: Biophysical and hydrological effects of future climate change including trends in CO2, in the St. Joseph River watershed, Eastern Corn Belt, Agricultural Water Management, 180, 280–296, doi:10.1016/j.agwat.2016.09.017, 2017.

4) Method

- This section is too long. Please shorten it and provide detailed information in SI.

We have substantially revised the organization and transitions within the Methods section. The section is re-organized to focus on the critical parts of the methods and to improve the flow of the paper. Here is the new order:

2.1. Data

2.2 Model (1): individual extremes

2.3 Model (2): compound extremes

2.4 Estimation strategy

- Equation 2: please define each variable and function used in the equation.

Thanks for pointing to the missing definitions. We have corrected it.

"where y_{it} is the crop yield, $g(h, m)$ is the yield response function to each combination of soil moisture level, *m,* and temperature (heat), *h*; $\varphi(h, m)$ is the distribution of soil moisture and heat; $\overline{m}$ and $\underline{m}$ are maximum and minimum soil moisture; $\overline{h}$ and $\underline{h}$ are maximum and minimum temperature; and c_i is a time-invariant county fixed effect. Here, we do not separate the impact of heat from water. In other words, the marginal impact of heat depends on water; and the marginal impact of water depends on heat."

- Ln 230: "some indicators", please clarify which indicators.

We have clarified this as:

> "In Model (1-c), we consider the number of days that soil moisture is either too high or too low. The model with these metrics of soil moisture extremes further improves the fit, revealing a negative marginal relationship associated with the number of days with low/high soil moisture."

- Please provide numbers to the equations.

Thanks for your comment. We added equation numbers in the revised version.

- Ln 277: g(Ws), please define the parameter

This description has been added to the text: "and $g(W_s)$ is 1 for all crops, while it is an exponential function of soil moisture depth for non-crop soil areas."

- Ln 290-295: this is a result of the analysis, not related to data/method or assumptions.

Thanks for your comment. We have moved this figure to the Supplementary Material. This information is important to ensuring that soil moisture is a different metric than precipitation. This information is added, and the statement re-contextualized and rephrased.

> "In a statistical study, a natural first step is to look at the correlation between these variables. To show that mean soil moisture is a different metric than mean precipitation, we have plotted the annual mean soil moisture versus annual cumulative precipitation in Fig. S1. This figure is a scatter plot for US counties for the growing season from 1981 to 2015. The simple correlation coefficient between them is 0.44. This rejects the hypothesis that soil moisture is highly correlated with precipitation. As mean precipitation has a linear relationship with cumulative precipitation, the results show that mean soil moisture is a different metric than cumulative or mean precipitation."

- Figure 4,5 and 6 are outcomes of the model/analysis. They can be presented in the result section.

These figures have moved to the results section. We have also added more explanations about the figures and their messages.

> "The overall simulation results from WBM are illustrated in Fig. 4-6, showing the gridded historical mean for the cultivated continental US, average annual variations for the cultivated continental US, and bivariate distribution of soil moisture and heat for the corn growing grid cells. To illustrate the spatial heterogeneity, Fig. 4 shows the growing season mean soil moisture content (in mm in 1000 mm topsoil) as calculated based on daily root-zone soil moisture level from Apr-Sep for 1981-2015 at 2.5 x 2.5 arcmin grids excluding non-cultivated area. Average growing season soil moisture is

heterogeneous across the Continental US, with distinct regional patterns (see Fig. 4). For the corn belt, the soil moisture level is relatively high compared to other regions. The mean of volumetric soil moisture ranges from below 50 mm in southern California to above 250 mm in the Corn Belt and around Mississippi.

To compare the variation of simulated soil moisture and precipitation, Fig 4 illustrates the weighted average soil moisture and precipitation over the cultivated US for 1981-2015. In general, variation in soil moisture average is higher than in that of precipitation (Fig. 5), showing how this new water metric is different from previous approaches. One interesting finding is that for some years the mean precipitation and the mean soil moisture move in opposite directions. For example, in 1990 the mean precipitation declined by around 5% while mean soil moisture increased by around 13%.

To show the dynamics of soil moisture and heat, Fig. 6 shows their bivariate distribution by month based on daily information for all the cultivated grid cells in the US Corn Belt for 1981-2015. Heat and soil moisture combinations vary through the growing season (Fig. 6) The data shows significant month-to-month variation, with the second half of the season facing hotter and dryer days. Also, July has the highest variation in soil moisture deviation with high probability of compound extremes as the distribution moves toward the lower right. "

5) Results

- Ln 363/364: "We will discuss the implications of these results in Sect. 5." The authors use lots of cross references between the sections as seen in here. This is not necessary, since discussion section means discussion of the results by definition. Please through the entire text and remove unnecessary cross-section references.

Good point. By cutting the length of the manuscript an improved flow of the paper, there is no need to these references. Thus, the superfluous section cross-references have been removed.

- Table 2: note section is repetition of the previous sections, thus it is not necessary.

The table notes have been removed or shortened for all the Tables.

- Ln 404: "This indicates that water is up to four times more valuable in hot weather." The authors can consider revising the sentence and be more explicit, "value of water" may mean several things.

As we omitted the value of water section, we have revised this as follows:

"The estimated parameters show the yield response to changes in soil water content. Comparing the parameter values can show the difference in yield response to soil moisture in hot weather and moderate weather…. This indicates that the average yield response to water is up to four times higher in hot weather."

- Model (2-a) and Model (2-b) were mentioned here for the first time. Please describe the differences between these models in method/data section.

The Methods section is revised to consider this comment. We have introduced the models in the relevant subsections on the Methods section. 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

6) Discussion

- Ln 410/411: this is related to differences between model 1 & 2, right? Please clarify which model outcome supports (or all?) the statement.

These lines are omitted. The clarification has been added in subsection 3.2 "Model (2): predicting yield response to compound extremes".

- Performance: does this mean best correlation between indicators of extreme events yield anomalies? Please clarify.

For comparing the models, we have Looked at statistical criteria. We have added Table 5 to compare the performance metrics of the models.

Table 5: Performance metrics for Models 1(a-d) and 2(a-d).

Model	Water metric	Extreme metric	R-squared	AIC (Akaike's information criterion)	BIC (Bayesian information criterion)
1-a	Avg. precipitation	Precipitation sqr	0.469	-21,238	-21,201
1-b	Avg. soil moisture	Soil moisture sqr	0.471	-21,612	-21,576
1-c	Avg. soil moisture	Number of days with low/high soil moisture	0.480	-22,697	-22,660
1-d	Avg. soil moisture	Avg soil moisture deficit/surplus	0.491	-24,303	-24,267
2-a	Avg. soil moisture	T binned by extreme deficit/surplus	0.492	-24,402	-24,328
2-b	normal soil moisture x T	extreme deficit/surplus x T	0.501	-25,582	-25,509

- First paragraph: what about model 1-c ?

This section is omitted. We have presented the results from model 1-c in subsection 3.1 " Model (1): predicting yield response to individual extremes".

> "Regarding Model (1-c), the coefficient on the number of days with low moisture is also significant and negative. Our estimation sample shows on average 26 days of high soil moisture and 27 days of low soil moisture. The implication is that eliminating 25 days of high soil moisture and 25 days of low soil moisture can improve the corn yields by up to 12.6%."

- Model 2 a-b were not defined in the previous parts of the paper. Please check consistency.

The Methods section is revised to consider this issue. As mentioned above, we have introduced the models in the relevant subsections on the Methods section.

> "First, we construct a binning estimator based on daily interaction on heat and soil moisture in model (2-a). …. We estimate a coefficient for each combination of excess heat and soil moisture; i.e., we estimate a model with metrics of degree days while controlling for soil moisture. The model provides the conditional marginal impact of excess heat as:

$$y_{it} = \alpha D_{it}^{10-29} + \left\{ \sum_m \beta_m D_{mit}^{29} \right\} + \delta M_{it} + \delta' M_{it}^2 + \lambda_s t + \lambda_s' t^2 + c_i + \varepsilon_{it} \quad (11)$$

> where i is the county index, t is the time index, m is an index of soil moisture condition (high, low, normal), s is an index for states, y is average corn yields, D represents conditional growing degree day variables, M shows the seasonal mean soil moisture content, T stands for the time trend variable, c_i is a time-invariant county fixed effect. Here, β is indexed by m; i.e., the marginal impact of heat is conditional to soil moisture conditions. α, β, δ, λ are the regression parameters showing the marginal impacts.

> Second, we estimate a model with metrics of soil moisture while controlling for temperature in model (2-b). We define an index of soil moisture when the temperature is above the threshold and an index of soil moisture when the temperature is below the threshold. In this model, the soil moisture is separated by a temperature threshold *H**.

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \left\{ \sum_m \delta_m M_{mit}\big|_{H<H*} + \delta_m' M_{mit}\big|_{H>H*} \right\} + \lambda_s t + \lambda_s' t^2 + c_i + \varepsilon_{it} \quad (12)$$

> where i is the county index, t is the time index, m is an index of soil moisture condition, s is an index for states, y shows average corn yields, D represents growing degree day variables, M shows conditional seasonal mean soil moisture, t stands for the time trend variable, H is the average daily temperature, $H*$ is the temperature threshold, and c_i is a time-invariant county fixed effect. Here, we define δ and δ' to test whether the marginal impact of soil moisture depends on heat. The soil

moisture metrics are calculated from daily gridded data and aggregated to county and growing season. This includes the index of normal soil moisture (*SM* 0-25+ mm around normal) when $H > H^*$, the index of normal soil moisture when $H < H^*$, the index of moisture deficit (*SM* 25+ mm below normal) when $H > H^*$, index of moisture deficit when $H < H^*$, the index of moisture surplus (*SM* 25+ mm above normal) when $H > H^*$, and the index of moisture surplus when $H < H^*$. $\alpha, \beta, \delta, \lambda$ are the regression parameters showing the marginal impacts. "

- Ln 416-421: These are newly introduced topics. None of these research goals (including why to have them), methods and results were mentioned in the previous sections of the paper (e.g. new interaction model, why do you have that and this was never mentioned in the paper). It is like Appendix is another paper with its own results, methods and goals. Please revise the paper accordingly.

We have substantially shortened and revised the Discussions and Appendix sections. The paper has been revised to focus on the main contribution and major messages. Thus, we dropped Model 2-c and 2-d as well as the discussions on "Implications for irrigation water demand and subsurface drainage" and "Implications for climate studies."

- Ln 424-428: Is this an outcome supported by the results? If so, please indicate how. It is more like a general knowledge.

This paragraph is omitted. We have revised the discussion section around the advantages of using the metrics of individual and compound extremes.

- Ln 429-430: Please provide supporting data/result from the analysis.

We have removed this paragraph as it requires further investigations which are not related to the main message of the paper.

- Ln 433: what are the other metrics suggested in the literature?

This section is omitted in the revised version.

- Ln 434-438: Is this a conclusion related to compound vs individual extreme weather event analysis? Can we say the same if we use other metrics of water stress than soil moisture?

Thanks for your helpful question. This section is omitted in the revised version. However, we used your suggestion in revising the paper. We focused on comparing models with individual extremes and models with compound extremes. This has improved the flow of the paper and highlighted the significance of this study.

- Ln 465-469: I question that the authors' research is critical for climate change studies. First, their analysis was based on historical data and says nothing about counterfactual analysis. This is not the first time impacts of a compound event was researched and like other studies this paper shows stronger impact of a compound event. It does not bring anything to climate change impact studies.

We have omitted this subsection in the revised version and briefly talked about it in the revised manuscript. However, we believe that the findings are critical for climate impact studies for several reasons. First, the current literature follows methods like Schlenker and Roberts (2009) by modelling yield response functions looking only at average water conditions. They ignore individual and compound extremes related to water. As we find that the coefficient on heat stress variable is significantly different when considering soil moisture and compound extremes, it is possible that previous climate impact studies have over- or under-estimated the yield impacts of climate change. Second, we are introducing simple but operational metrics of individual and compound extremes that can be constructed using hydroclimatic models for the future. These metrics can improve the prediction of crop yields. We are not aware of any other study suggesting such a simple yet powerful prediction framework.

- Ln 472: please clarify benefit of this collaboration. In which way it helps to solve the challenge.

We believe that collaboration between hydrologists, climate scientists, and statisticians can improve data generating processes and leads to better models and metrics to help better decisions among people and policymakers. Here is the revised text:

> "Applying this framework to climate impact studies will face a key challenge —namely projecting the future compound extremes with the high temporal resolution of Model 2. It requires collaboration between hydrologists, climate scientists, and statisticians (Zscheischler et al., 2020). For future yield projections, we need reliable future projections of daily temperature (maximum and minimum) and soil moisture. Unfortunately, to the best of our knowledge, available data sets including predictions of future soil moisture have a relatively coarse spatial and temporal resolution, and rely on climate model projections with known difficulties representing daily temporal resolution events (Hempel et al., 2013). Further research is required to improve the ability of climate models and impact models in projecting the bivariate distribution of heat-moisture (Sarhadi et al., 2018)."

References:

Zscheischler, J., van den Hurk, B., Ward, P. J. and Westra, S.: Multivariate extremes and compound events, in Climate Extremes and Their Implications for Impact and Risk Assessment, pp. 59–76, Elsevier., 2020.

Sarhadi, A., Ausín, M. C., Wiper, M. P., Touma, D. and Diffenbaugh, N. S.: Multidimensional risk in a nonstationary climate: Joint probability of increasingly severe warm and dry conditions, Science Advances, 4(11), eaau3487, doi:10.1126/sciadv.aau3487, 2018.

- Ln 479-ln 483: this recommendation is not related to the sub-section heading. The authors stated a discussion point which is out of scope of their analysis and not supported with the overall goal of the paper. Recommendations can be given to farmers etc; however their model/research is not aimed for decision - support guidance. Please remove this section of revise it.

Thanks for your comment. We have omitted this part.

- Section 5.4: This section includes literature, method, data and equation related to an estimation. This is not a discussion section. Please previse it accordingly. This additional analysis doesn't bring anything to the value of the paper. I would recommend excluding this analysis from the paper in order to keep its coherence and consistency.

Thanks for your comments which helped to improve the flow of the paper. We have omitted this subsection.

- Ln 501: "We find that the average damage from excess heat has been up to four times more severe when combined with water stress" what is the damage, yield losses?

Thanks for your comment. Originally benefits and damages were considered from an economics point of view. In the revised version, we removed the economic analysis of the value of soil moisture. Now we have revised and clarified the sentence as:

> "Finally, the marginal impact of heat index on crop yields depends on the soil moisture level. We show the average yield damage from heat stress is up to four times more severe when combined with water stress; and therefore the value of water in maintaining crop yield is up to four times larger on hot days."

- Line 517-525: the CC knowledge and analysis were not included in previous parts (method, data, results) section of the paper. Please include info about this analysis in adequate sections.

To improve the flow of the paper and reduce the redundancy, the climate change material is omitted.

- Line 525- 535: There is almost no economic analysis thus the paper does not contribute to CC economics. No policy analysis or research were provided either; also paper does not say/bring anything to regional resilience of agroecosystems, global food security, and as well as future climate impacts. These two paragraphs have to be re-written. These claims are bold and cannot be taken from the research as described in the paper.

Thanks for your comment. As we have dropped the subsection, theses paragraphs are also omitted.

Track-change color guides:

Inserted text

~~Deleted text~~

~~Moved from~~

Moved to

~~Moved to and deleted~~

Quantifying the Impacts of Compound Extremes on Agriculture ~~and Irrigation Water Demand~~

Iman Haqiqi[1], Danielle S. Grogan[2], Thomas W. Hertel[1,3], and Wolfram Schlenker[4]

[1] Department of Agricultural Economics, Purdue University, West Lafayette, IN, USA.
[2] Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH, USA.
[3] Purdue Climate Change Research Center, Purdue University, West Lafayette, IN, USA.
[4] School of International and Public Affairs, Columbia University, New York City, NY, USA.

Correspondence to: Iman Haqiqi (ihaqiqi@purdue.edu)

Abstract. Agricultural production and food prices are affected by hydroclimatic extremes. There has been a large literature measuring the impacts of individual extreme events (heat stress or water stress) on agricultural and human systems. Yet, we lack a comprehensive understanding of the significance and the magnitude of the impacts of compound extremes. Here, we combine a ~~high-resolution~~ fine-scale weather product with ~~fine-scale~~ outputs of a hydrological model to construct functional ~~indicator~~metrics of individual and compound hydroclimatic extremes for agriculture. Then, we ~~measure the impacts of~~ estimate a yield response function with individual and compound ~~extremes~~ metrics ~~on crop yields~~ focusing on corn in the United States during the 1981-2015 period. Supported by statistical evidence, we show that metrics of compound hydroclimatic extremes are better predictors of corn yield variations than metrics of individual extremes. We also confirm that wet heat is more damaging than dry heat for corn. We show that the average yield damage from heat stress has been up to four times more severe when combined with water stress~~; and the value of water experiences a four-fold increase on hot days~~. ~~In a robust framework with only a few parameters of compound extremes, this paper also improves our understanding of the conditional marginal value (or damage) of water in crop production. This value is critically important for irrigation water demand and farmer decision making – particularly in the context of supplemental irrigation and sub-surface drainage.~~

Keywords. agriculture; climate impacts; water balance model; extreme heat; extreme drought.

1 Introduction

~~In this paper, we quantify the response of agricultural production to individual and compound extremes. This study employs high-resolution daily estimates of soil moisture from a hydrological model, in combination with fine-scale daily weather data and a county-level panel of crop yields across the United States (US). By considering metrics of individual and compound extremes in a statistical framework from the literature estimating climate impacts on agriculture, we explore the importance of compound hydroclimatic extremes versus individual extremes. We show that the compound-extremes approach provides a significantly better prediction model compared to the individual-extreme approach that is currently employed in the literature. The proposed~~

~~framework also allows us to estimate the conditional marginal value of water (economic water demand) in agricultural production – a key metric for producers making supplemental irrigation decisions. We show the implications of our findings for irrigation water demand and subsurface drainage. This paper adds important nuances to the climate-agriculture-water links and therefore helps to refine projections of future impacts of anthropogenic climate change.~~

We construct various ~~indicator~~metrics of individual and compound hydroclimatic extremes appropriate for agricultural studies. In agricultural production, water and heat extremes are key determinants of yield variations. They affect agricultural yields, farm revenues, and crop markets. The relationship between extreme heat~~, cumulative seasonal precipitation,~~ and crop yields has been well-documented, particularly across the United States (US) and particularly for corn (Schlenker and Roberts, 2009; Urban et al., 2012; Diffenbaugh et al., 2012; Roberts et al., 2013; Lobell et al., 2013; Urban et al., 2015; Wing et al., 2015; Burke and Emerick, 2016). However, the precipitation-based ~~indicator~~metrics of water conditions used previously are either mean or cumulative measures calculated over the growing season or stages of crop growth. ~~However,~~These cumulative indices, monthly mean, or seasonal average metrics do not capture extreme events during the season (e.g. early-season floods and late-season droughts can cancel out when taking the average). The mean variable can be misleading as the plants respond to day to day variability. Furthermore, the mean water index may not represent hydrological extremes (D'Odorico and Porporato, 2004; Lobell and Burke, 2010; Schaffer et al., 2015; Werner and Cannon, 2016). While the average conditions are important, exposure to extreme water stress can cause permanent unrecoverable damage to the plant (Denmead and Shaw, 1960). In addition, too much water can cause flooding, waterlogging, or may wash out soil nutrients and fertilizers (Kaur et al., 2018; Schmidt et al., 2011; Urban et al., 2015). Therefore, it is necessary to introduce ~~indicator~~metrics of extreme soil moisture stress. This will be even more important in the future, as climate ~~scientist~~projections are predicting more extreme drought and precipitation events (Myhre et al., 2019). In other words, mean variables can create biases in future climate impact analysis by ignoring the extreme events. It is important to ~~introduce different~~ evaluate new metrics of daily water availability to ~~to measure the value of water at the time which is most needed.~~ fully understand the impact of water extremes on crop yields, as this will be important in both fundamental understandings of the crop-water system, and in predicting the impacts of future extreme events.

Further, crops obtain their water directly from soil moisture, yet~~Other metrics of~~ extreme water metrics based on soil moisture ~~conditions~~ have been only minimally explored (Fishman, 2016). Several~~ome~~ studies ~~also~~have highlighted the need for irrigation to compensate for soil moisture deficits (Li et al., 2017; McDonald and Girvetz, 2013; Meng et al., 2016; Williams et al., 2016), further pointing to soil moisture as a potentially more important crop water availability metric than precipitation. However, ~~C~~current statistical studies have had limited success in statistically capturing the yield response to soil moisture metrics (Bradford et al., 2017; Peichl et al., 2018; Siebert et al., 2017). There are several potential reasons for the limited success of previous statistical studies in capturing yield response to soil moisture. Direct measures of soil water availability include complex biophysical and hydrological processes that are difficult to capture in a rather simple statistical model. On the other hand, seasonal mean soil moisture is highly correlated to seasonal precipitation. Thus, including an average metric of soil water content may not add value to a statistical model. Another~~One~~ barrier has been limited availability of daily fine-scale soil moisture data and inconsistency of soil moisture data with heat information. It has become a standard practice ~~for current studies~~ either to focus on a limited geographical area (Rizzo et al., 2018; Wang et al., 2017) or to employ a proxy variable like precipitation, evapotranspiration, or vapor pressure deficit estimates (Comas et al., 2019; Roberts et al., 2013). The recent work by Ortiz-Bobea et al. is an ~~notable~~ exception~~. It~~ that highlights the importance of mean soil moisture metrics for estimating crop yields in the US (Ortiz-Bobea et al., 2019).~~However, it ignores the daily interaction of soil moisture and heat and the significant role played by soil typesIn the physical sciences, researchers have discussed the impact of climate change on soil moisture (Feng and Zhang, 2015; Jung et al., 2010; Marshall et al., 2015; McDonald and Girvetz, 2013; Rodell et al., 2018; Taylor et al., 2013).~~

A~~However, a~~ key unknown is the extent of the benefits of soil moisture in buffering ~~the~~ heat damage to yields. Despite existing theoretical frameworks and controlled experiments, we currently lack a comprehensive understanding of the ~~conditional marginal~~ impact of heat on yields while controlling for water (Bradford et al., 2017; Ortiz-Bobea et al., 2019). The problem is that current studies tend to separate the impact of heat from water stress. These studies estimate the average impact of heat stress on corn yields without distinguishing between a hot-dry day (dry heat) and a hot-wet day (wet heat). There is no robust predictive framework that captures the implications of ~~daily interactions of soil moisture and heat~~ compound extremes in the determination of national crop yields. Also, the current literature is focused mainly on the impacts of dry-heat and ignoring the impacts of wet-heat stress (Ribeiro et al., 2020). ~~In short, standard measures of heat and water stress are missing important temporal, spatial, and vertical dynamics.~~

~~However, t~~The growth effects of heat and soil moisture are mutually interdependent. Beneficial heat is less beneficial without sufficient soil moisture. On the other hand, soil moisture is not beneficial without sufficient heat for plant growth. Harmful heat can be less harmful when there is enough soil moisture (Hauser et al., 2018). While the amount of daily water requirement depends on the biophysical properties of soil and crop, it changes with temperature, solar radiation, humidity, and wind speed. In this framework, daily weather variability, which is expected to change in the future with climate change, can affect both soil moisture supply and demand by altering the abundance and frequency of precipitation and by increasing the water required to compensate evapotranspiration and evaporation. If the temperature is high and there is not enough soil moisture for a long period (drought conditions), this may cause severe damage to crops (Denmead and Shaw, 1960). Therefore, consideration of the daily ~~interaction~~ compound impacts of soil moisture and heat is necessary to capture the impacts on natural supply and plant demand for soil moisture.

In this paper, we investigate the significance of ~~the~~ compound heat and water conditions in predicting crop yields, including dry-heat and wet-heat. We focus on corn as the major field crop in the US. We also compare the ~~indicator~~metrics of compound extremes versus individual extremes (i.e. only heat stress or only water stress). This study also demonstrates the advantages of using soil moisture metrics over current proxy variables in capturing climate-driven variations in heat and moisture availability.

~~Marginal impact and conditional marginal impact are two statistical concepts equivalent to partial derivatives in mathematics. When the partial derivative of one variable does not depend on other variables, we use the term "marginal impact". When it depends on other variables, we use "conditional marginal impact". A conditional marginal impact shows the impact of a compound extreme. A non-conditional marginal impact can show the impact of individual extremes. This study improves our understanding of the value of water management in crop production. It contributes to the socio-hydrology literature by providing a robust framework for studying the climate-agriculture-water links (Ertsen et al., 2013; Fernald et al., 2015; Van Emmerik et al., 2014; Di Baldassarre et al., 2019). It also serves to bridge the gap between statistical studies of climate impacts on crops and their biophysical counterparts. Understanding the true value of water management for agriculture is critical in the face of a warming climate. Fluctuations in precipitation can lead to drought or flooding. They account for more than 70% of crop indemnities in the US during the 2001-2015 period (USDA-RMA). As economic agents, farmers may choose to adjust to climate change depending on the likely benefits and costs of alternative options. Current studies suggest a significant impact from climate change on rainfed agriculture (Ortiz-Bobea and Just, 2013; Annan and Schlenker, 2015; Liu et al., 2017; Sesmero et al., 2017; Hsiang and Kopp, 2018; McCarl and Hertel, 2018). This study sheds light on the benefits of adaptation options including full-scale irrigation or supplemental irrigation by showing how water can reduce heat damages to crops. Although converting to complete irrigation is sometimes an attractive solution, a more challenging question involves the likely benefits and costs of supplemental irrigation. While the biophysical information necessary for these calculations is offered, at least in part, by the hydroclimatic, biophysical, geospatial, earth, and atmospheric sciences, this study transforms this information into economic terms that are useful for both farmers and policymakers.~~

In this paper, we will show how the results can be used to economically quantify the marginal value of water, in the form of soil moisture, for corn production in the US under different hydroclimatic conditions.

The remainder of this paper is organized as follows. The next section provides an overview of the empirical concerns in this type of study. Then we introduce two models to explain the significance of soil moisture for estimating crop yields. We also describe different data sets used in the study. Section 4 provides estimation results. Section 5 contains a discussion on the implications of the findings for climate impact research. And Section. 6 concludes.

2 Empirical concerns

In this section, we will review some of the empirical concerns in investigating the individual and compound impacts of heat and water on crop yields in statistical models. Before starting our discussion, we will briefly describe the basic model as introduced by Schlenker and Roberts (2009). The model assumes that the effects of heat on corn yields are cumulative over the growing season. In other words, the end-of-season yield is the integral of daily heat impacts over the growing season. This relationship can be demonstrated via Eq. (1):

$$y_{it} = \int_{\underline{h}}^{\overline{h}} g(h)\varphi_{it}(h)dh + z_{it}\delta + c_i + \epsilon_{it} \qquad (1)$$

where $\varphi_{it}(h)$ is the time distribution of heat (h) over the growing season in county i and year t, while the heat ranges between the lower bound $\underline{h}$ and the upper bound $\overline{h}$; indicators of water availability and other control factors are denoted as z_{it}, and c_i is a time-invariant county fixed effect.

When using high-resolution data on heat and water, there are empirical concerns regarding the data generating processes as well as estimation itself. Here we will briefly talk about choosing the variables, the level of aggregation, endogeneity issues, measurement errors, model specifications, and standardization.

2.1 Degree of spatial and temporal aggregation

The impact of water on crop production depends on local hydroclimatic conditions. There is considerable heterogeneity in the geographical distribution of water resources and the location of producers. There is also significant spatial heterogeneity in soil properties, which modulate temperature and precipitation signals through their capacity to hold water as soil moisture. It is important to have a sample with extreme conditions and a pattern of spatial heterogeneity which is representative of the entire market (population).

:

An important aspect of estimation of the impact of compound extremes is choosing the right spatial scope of the study. A simple spatial aggregation can eliminate the extreme conditions (tails of the distribution) from the sample. On the other hand, farm-level data is limited to specific locations. Although estimates based on geographically limited observations can be informative for those locations, a more comprehensive analysis of yield response to climate is necessary for market predictions (Kucharik, 2003; Kudamatsu, 2018; Martin, 2018). We construct our database at the county level in the US. To ensure a correct data generating process, the hydroclimate data are constructed such that nonlinear transformations are taken at the grid cell level before being aggregated across time and space (Hsiang, 2016). The county-level aggregation has a couple of benefits. First, the corn yield data reported by USDA at the county level and we do not need to have heavy data processing on the yield data. Also, this will provide enough heterogeneity for local market analysis and local climate impact reports.

~~Another empirical challenge is~~Note ~~that yields data are reported annually, while weather data have a higher temporal resolution (e.g. hourly, daily, or weekly). Consequently, some empirical studies employ annual or monthly weather indicators like average temperature. However, many studies utilize~~ aggregate the ~~daily climate information by introducing growing degree days and harmful degree days through the growing season (D'Agostino and Schlenker, 2016; Mueller et al., 2012). One standard solution is the use of the growing degree days approach along with an index of cumulative rainfall to proxy for water availability as in Schlenker and Roberts (2009); or as mean monthly or seasonal soil moisture in Ortiz-Bobea et al. (2019).~~

~~However, cumulative indices, monthly mean, or seasonal average metrics do not capture extreme events during the season (e.g. early-season floods and late-season droughts can cancel out when taking the average). The mean variable can be misleading as the plants respond to day to day variability. Furthermore, the mean water index may not represent hydrological extremes (D'Odorico and Porporato, 2004; Lobell and Burke, 2010; Schaffer et al., 2015; Werner and Cannon, 2016). While the average conditions are important, exposure to extreme water stress can cause permanent unrecoverable damage to the plant (Denmead and Shaw, 1960). In addition, too much water can cause flooding, waterlogging, or may wash out soil nutrients and fertilizers (Kaur et al., 2018; Schmidt et al., 2011; Urban et al., 2015). Therefore, it is necessary to introduce indicators of extreme soil moisture stress. This will be even more important in the future, as climate scientists are predicting more extreme drought and precipitation events (Myhre et al., 2019). In other words, mean variables can create biases in future climate impact analysis by ignoring the extreme events. It is important to introduce different metrics of daily water availability to measure the value of water at the time which is most needed. Panel (a) in Fig. 1 visualizes four soil moisture conditions that are unfavorable for crop yield. Both too much water [i] and intense moisture stress [ii] can cause severe damage to crop yields. Similarly, a long period of mild moisture stress [iii] or a short period of severe moisture stress [iv] can also cause significant yield loss. These measures can help to understand the need for artificial drainage or irrigation as shown in panel (b).~~

~~2.2 Water availability index~~

~~While soil moisture plays a crucial role in determining climate impacts on agricultural yields, there have been only a few successful statistical studies in measuring this relationship. Many researchers have acknowledged the need for soil moisture data to predict the response of crop yields to variations in water availability. Some studies also highlight the need for irrigation to compensate for soil moisture deficits (Li et al., 2017; McDonald and Girvetz, 2013; Meng et al., 2016; Williams et al., 2016). One barrier has been limited availability of daily fine-scale soil moisture data and inconsistency of soil moisture data with heat information. It has become a standard practice for current studies either to focus on a limited geographical area (Rizzo et al., 2018; Wang et al., 2017) or to employ a proxy variable like precipitation, evapotranspiration, or vapor pressure deficit estimates (Comas et al., 2019; Roberts et al., 2013). While cumulative precipitation is significant in previous studies, it may not be a good representation of available water for plants in many places due to irrigation, runoff, or evaporation. Indeed, it is only relevant if the precipitation is stored in the soil for plant use during the season. Following these studies, one might be able to estimate the marginal impact of change in "mean precipitation" or "mean evapotranspiration". However, this will not necessarily provide appropriate coefficients for future climate impacts as the distribution of precipitation across space and time is estimated to change, leading to more frequent extreme events (Myhre et al., 2019). To undertake climate impact analyses of water availability required further information.~~

~~In this study, we will show that, although cumulative precipitation and mean soil moisture are correlated, their performance can be different in predicting corn yields. We will show how an empirically validated, high-resolution hydrological model, such as WBM, can provide valuable information for estimating the marginal value of water.~~

2.3 Interaction of soil moisture and heat

To accurately measure the marginal impact of soil moisture, we need to draw on biogeochemistry, hydrology, and plant physiology perspectives on crop yields and soil moisture. We treat soil moisture as an integrative variable that contains information on precipitation, temperature, and soil types, as well as the behavior of the crops themselves. Crop yields depend on daily growth during the season (Hatfield and Prueger, 2015). Plants require water for germination, transpiration, nutrient transport, and to buffer against temperature fluctuations (Maharjan et al., 2016; Teixeira et al., 2014). Therefore, timely irrigation can play an important role in boosting yields (Carter et al., 2016; Siebert et al., 2017; Tack et al., 2017; Troy et al., 2015).

However, the growth effects of heat and soil moisture are mutually interdependent. Beneficial heat is less beneficial without sufficient soil moisture. On the other hand, soil moisture is not beneficial without sufficient heat for plant growth. Harmful heat can be less harmful when there is enough soil moisture (Hauser et al., 2018). While the amount of daily water requirement depends on the biophysical properties of soil and crop, it changes with temperature, solar radiation, humidity, and wind speed. In this framework, climate change can affect both soil moisture supply and demand by altering the abundance and frequency of precipitation and by increasing the water required to compensate evapotranspiration and evaporation. If the temperature is high and there is not enough soil moisture for a long period (drought conditions), this may cause severe damage to crops (Denmead and Shaw, 1960). Therefore, consideration of the daily interaction of soil moisture and heat is necessary to capture the impacts on natural supply and plant demand for soil moisture.

2.4 Importance of soils

In a large literature in the statistical estimation of corn yields in the US, water availability is represented by cumulative precipitation and its square term in a fixed effect panel regression. The estimated coefficients suggest a positive impact from cumulative precipitation and a negative impact from its square term. This leads to a universal optimum precipitation level, $\hat{p}*$, which is the same for all the observation locations. However, this is not necessarily equal to the true optimum level of water for production in each location. According to the agronomic literature, the optimum amount of water depends on the moisture stored in the soil, soil type, and heat (Fang and Su, 2019). Thus, many studies utilize other metrics of water availability including estimated evapotranspiration, standardized precipitation, and drought indices.

On the other hand, standard measures of volumetric soil moisture are not the best indicator of water availability. In the agronomic literature, the water available to plants depends on volumetric soil moisture as well as soil type. for the same volume of soil moisture, different soil types imply different wilting points and different field capacity which result in different water availability to crops. Figure 2 shows the difference between soil moisture content, water available to plants, and unavailable water. This illustrates that sand and sandy-loam soil types have the lowest field capacity (and water availability) while clay and clay-loam have the highest. As soil moisture metrics (volumetric or fraction) vary over the space, we need to look at soil type, crop cover, and other biophysical variables. Generally, soil moisture thresholds are defined in terms of the soil available water for plants, or soil wilting point, not a constant depth of water.

As a simple solution, one can capture the differences in soil type by introducing dummy variables. However, at the county level aggregation which has been chosen by many studies, it is challenging to select a soil type for a county. While a dominant soil type can work, it is not necessarily the best option. As we prefer to take care of differences at the grid cell level before aggregation. Another solution is to standardize the soil moisture indicator. Introducing the soil moisture fraction can help as it takes the ratio of soil moisture content to the field capacity. However, interpretation of the results is not straightforward. A better measure is the soil moisture deviation from normal. This is defined as daily soil moisture deviation from historical average soil moisture at each

~~location. In a standard Schlenker-Roberts type model, the coefficient on this indicator would show the percentage change in corn yields in response to one mm higher soil moisture deficit (or surplus). We use deviation from normal levels as this can remove the location-specific features of soil moisture. While irrigation is taken into calculations, variation in this metric is higher in non-irrigated areas and is lower in irrigated areas as the irrigating farmers try to keep the soil moisture around a normal range.~~

~~2.5 Measurement errors and endogeneity concerns~~

~~While there exist remotely sensed metrics of soil moisture (e.g. NASA's Gravity Recovery and Climate Experiment or the European Space Agency's Climate Change Initiative), they are coarse in spatial and temporal resolution. Also, they are relatively new and therefore give rise to a short length for the panel data. Also, there is in situ observed soil moisture data that suffer from missing data points and requires a significant amount of interpolation as the stations are irregularly scattered in time and space (Ford and Quiring, 2019).~~

~~On the other hand, simulated soil moisture data from hydrological models can be problematic in different ways. Also, if the simulation involves a time-varying yield input, the estimations will be biased due to serious endogeneity problems. Besides, if a model employs a simulation framework based on specific parameters and functional forms, there is a likely systematic measurement error due to the correlation of unobservable determinants of soil moisture over space and time. For studies covering the continental US, encompassing both a highly irrigated West and a less-irrigated East, irrigation is also important in estimating the soil moisture. If the simulated soil moisture metric ignores the irrigation inputs, the estimation will suffer from a key omitted variable. Irrigation inputs will be correlated with the soil moisture, as irrigation water will be applied when precipitation inputs are insufficient for optimal crop growth. Thus, when soil moisture is low, irrigation is more likely to be high. This challenge is the reason that most of the literature linking corn yields to temperature and precipitation across the US has relied only on counties east of the 100th meridian, where corn is rarely irrigated.~~

~~The average soil moisture output from WBM is informed mainly by daily soil moisture memory, exogenous heat data, time-invariant crop cover, time-invariant soil features, precipitation, irrigation, and carefully calibrated parameters. Furthermore, the output from the WBM has been validated against observational data (Grogan et al., 2017). This ensures that the model performs well in replicating the observations. Also, as it includes irrigation in generating the soil moisture, it is a reliable data source for both Eastern and Western US (in which irrigation is dominant). As the model does not use yield data, the soil moisture is invariant to changes in yield. That said, we will test the performance of soil moisture in predicting corn yields.~~

~~Typically, clustering the standard errors (for example by state) is a standard practice to minimize the remaining concerns. We will also use clusters by state. In addition, we will employ the soil moisture deviation from normal, to select the bins and generate the soil moisture extreme metrics. This deviation can eliminate the likely systematic measurement errors in soil moisture data which can happen due to simulation. As discussed in Sect 2.4, we will introduce metrics for soil moisture deficit and soil moisture surplus by calculating the daily deviation from normal soil moisture levels. This will tackle two problems at the same time: choosing a water availability index that provides extreme conditions while taking into account different soil types.~~

~~3~~2 Methods

Technically, we extend the models in Schlenker and Roberts (2009) and Ortiz-Bobea et al. (2019) by assuming the growth effects of heat and water are mutually interdependent. ~~The model captures the impacts of compound extremes (e.g. hot-dry or hot-wet conditions) as well as individual extremes (excess heat, excess water, and water deficit).~~ We use detailed soil moisture information available from ~~recent developments in the Water Balance Model (Grogan, 2016; Wisser et al., 2010), hereafter WBM. We show~~

~~that the coefficient on excess heat in the estimation of corn yield is significantly different when we consider the daily interaction with soil moisture. T~~a physical hydrologic model. ~~We find that the soil moisture index, and its daily interaction with heat, perform better in predicting corn yields compared to the commonly used proxy variables such as cumulative precipitation.~~ Specifically, we estimate 1) the marginal impacts of heat stress (individual extreme) on crop yields; 2) the marginal impact of daily soil moisture stress (individual extreme) on crop yields, and 3) the marginal impact of heat and soil moisture (compound extremes) on crop yields. Marginal impact and conditional marginal impact are two statistical concepts equivalent to partial derivatives in mathematics. When the partial derivative of one variable does not depend on other variables, we use the term "marginal impact". When it depends on other variables, we use "conditional marginal impact". A conditional marginal impact shows the impact of a compound extreme. A non-conditional marginal impact can show the impact of individual extremes.

~~In this section,~~Here, we introduce two statistical models of crop yield as a function of heat and soil moisture. For each model, we consider different parameterizations of heat and ~~water~~ soil moisture to estimate the impacts of water availability on corn yields ~~of corn~~ in the US. Model 1 assumes the impacts of heat and water on corn yields are separable. This model considers metrics of individual extremes (heat stress and ~~different measures of~~ water availability). ~~Panel (a) in~~ Figure~~.~~ 1 visualizes four soil moisture conditions that are unfavorable for crop yield. Both too much water [i] and intense moisture stress [ii] can cause severe damage to crop yields. Similarly, a long period of mild moisture stress [iii] or a short period of severe moisture stress [iv] can also cause significant yield loss. These measures can help to understand the need for artificial drainage or irrigation as shown in panel (b). Within this framework, we investigate which metric 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 ~~in~~metrics of ~~teractions of heat and soil moisture with individual and~~ compound extremes. ~~For each model, we will describe the relevant variables and their measurement. After describing all the models, data sources are introduced in detail. We construct the models taking into account the fact that plants generate biomass each day using the available resources like heat and water (we assume no change in soil nutrients).~~

2.1 Data

In estimating the marginal impact of soil moisture on corn yields, we employ information about soil moisture, temperature, precipitation, and corn yields for counties of the United States for the 1981-2015 period as summarized in Table 1. The data on yield is obtained from USDA-NASS (United States Department of Agriculture-National Agricultural Statistics Service) at the county level. The yield is defined as the corn production (in bushels) divided by harvested area (in acres). Precipitation is defined in millimeters as accumulated rainfall during the growing season (Apr-Sep). It is calculated based on PRISM (Parameter-elevation Regressions on Independent Slopes Model) daily information at 2.5 x 2.5 arcmin grid cells over the continental US for 1981-2015. It is aggregated to each county according to cropland area weights. Compound metrics ~~Daily interaction~~ of heat and soil moisture~~is~~ are also calculated daily at the gridded level. Then we aggregate the metrics to the growing season and county level.

3.3.1 Degree days (index for heat)

Following D'Agostino and Schlenker (2015), the daily distribution of temperatures is approximated assuming a cosine function between the daily minimum and maximum temperature. Let $\bar{t} = \text{acos}\left(\frac{2b - T_{max} - T_{min}}{T_{max} - T_{min}}\right)$, then degree days (dday) at each day is defined using

$$D(b) = \begin{cases} \frac{(T_{max}+T_{min})}{2} - b & if\ b \le T_{min} \\ \frac{\bar{t}}{\pi}\left[\frac{(T_{max}+T_{min})}{2} - b\right] + \frac{(T_{max}-T_{min})}{2\pi}\sin(\bar{t}) & if\ T_{min} < b \le T_{max} \\ 0 & if\ T_{max} < b \end{cases} \qquad (1)$$

where b is the base for calculating degree days and can take the base values as well as critical values. we consider a piecewise-linear function to aggregate the degree days. The major assumption is that plant growth is approximately linear between two bounds. Degree days between two bounds is simply degree days above the smaller bound minus degree days above the larger bound.

We calculate county-level seasonal degree days based on daily weather information. The weather information on daily maximum and minimum temperature are obtained from PRISM at 2.5 x 2.5 arcmin grid cells over the continental US for 1981-2015. Degree days are initially calculated for each day at each 2.5 x 2.5 arcmin grid cell during the growing season (Apr-Sep). Then they are aggregated for the whole growing season from the first day of April through the last day of September. Finally, they are aggregated to the county level using cropland area weights.

3.3.2 Soil moisture (index for water availability)

Daily soil moisture content and soil moisture fraction are obtained from the Water Balance Model (Grogan, 2016; Wisser et al., 2010) based on daily simulations using PRISM data at 6 x 6 arcmin grid cells for the 1981-2015 period over the continental US. Here, we briefly describe WBM's soil moisture module. However, the model is much more complex and employs a large list of inputs. Full documentation for WBM can be found in Wisser et al. (2010) with updates in Grogan (2016). In WBM, crop-specific soil moisture balance within each grid cell is calculated with an accounting system that tracks a location's water inputs and outputs and is limited by the soil moisture pool's water holding capacity.

$$\frac{\delta W_s}{\delta t} = \begin{cases} g(Ws)(I - PET)\ if\ I < PET \\ I - PET\ if\ PET\ \le I\ and\ (I - PET) < (W_{cap} - W_s) \\ W_{cap} - W_s\ if\ PET\ \le I\ and\ (W_{cap} - W_s) \le (I - PET) \end{cases} \qquad (2)$$

where W_s is soil moisture, t is time, I is the sum of all water inputs to the soil moisture pool, PET is potential evapotranspiration, and W_{cap} is available water capacity. Water inputs to the soil come in the form of precipitation as rain and as snowmelt. Water intercepted by the canopy reduces precipitation reaching the soil. Here, we use the Hamon method for estimating PET (Federer et al., 1996; Hamon, 1963)~~(Hamon, 1963; Federer et al. 1996)~~, and $g(W_s)$ is 1 for all crops, while it is an exponential function of soil moisture depth for non-crop soil areas. Crop-specific potential evapotranspiration values, $PETc$, are calculated following the FAO-recommended crop-modeling methodology (Allen et al., 1998):

$$PET_c = k_c \cdot PET \qquad (3)$$

where k_c [-] is a crop-specific, time-varying scalar. Crop scalar values are from Siebert and Döll (2010), and crop maps that identify the area of each rainfed crop type within a grid cell are from the Crop Data Layer (CDL, USDA NASS, 2017). When soil moisture is insufficient for crops to extract water equal to PET_c, actual crop evapotranspiration is limited to available soil water volumes. Available water capacity, W_{cap}, is a function of vegetation-specific rooting depth, a crop-specific depletion factor, soil field capacity, and soil wilting point:

$$W_{cap} = D_c R_c (F - W_p) \qquad (4)$$

where D_c is the depletion factor for crop c, R_c is the rooting depth of crop c, F is the soil field capacity, and W_p is the soil wilting point. Here we use the Harmonized World Soil Database (Fischer et al., 2008) as model input for all soil properties. Corn rooting depth is set to 1 meter~~ and the corn depletion factor is 0.55~~; and the depletion factor is 0.5, following Siebert and Döll (2010). Once the soil moisture content reaches field capacity, no further water is added to the soil moisture pool; excess inputs move to the groundwater pool via percolation and the river system via runoff.

~~The soil moisture data used here is not a simple linear transformation of precipitation data, as evidenced by their simple correlation coefficient of 0.44 (scatter plot shown in Fig. 3). We have also investigated other correlations including the correlation between mean soil moisture and evapotranspiration as illustrated in Fig. A2; and the correlation between mean soil moisture and mean daily soil moisture fraction as shown in Fig. A3.~~

~~WBM and PRISM grid cells have different extent, different resolution, and non-matching centroids. Therefore, we interpolate WBM to PRISM using nearest neighbor and bilinear methods, providing soil moisture information at 2.5 x 2.5 arc-minute grid cells. The regression results are reported for bilinear interpolation; results using the nearest neighbor interpolation method are very similar (Table S6).~~

The ~~S~~soil moisture metrics used in both statistical~~ the~~ models are~~is~~ calculated as the mean of soil moisture content (in mm for the 1000 mm topsoil) or cumulative deviations from normal levels during the growing season (Apr-Sep) for each 2.5 x 2.5 arcmin grid cell. For the compound metrics of ~~interaction of~~ soil moisture and heat, we sum up degree days for each temperature interval (5ºC) for each soil moisture deviation interval (10 mm) for each 2.5 x 2.5 arcmin grid cells for the 1981-2015 period. We employed Crop Data Layer from the US Department of Agriculture ~~These fine-scale metrics are checked with satellite scans of cropland area~~ 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)~~. Finally, we aggregate all the grids in each county using cropland area weight~~.

We have constructed the soil moisture metric relative to the "normal" levels. In this study, we define normal as the 25-year average soil moisture in the growing season. ~~Average growing season soil moisture is heterogeneous across the Continental US, with distinct regional patterns (see Fig. 4). For the corn belt, the soil moisture level is relatively high compared to other regions.~~ T~~In the agronomic literature, t~~he water available to plants depends on volumetric soil moisture as well as soil type. For the same volume of soil moisture, different soil types imply different wilting points and different field capacity which result in different water availability to crops. Figure S1 shows the difference between normal soil moisture content, water available to plants, and unavailable water. This illustrates that sand and sandy-loam soil types have the lowest field capacity (and water availability) while clay and clay-loam have the highest. To operationalize ~~this~~ the soil moisture metric, we consider soil moisture deviation from normal. Soil moisture deviation is defined as daily soil moisture minus the normal soil moisture levels. The soil moisture level is considered extreme if it is ~~25 mm~~ below/above~~normal condition~~ a threshold. The threshold is obtained by testing the impacts of 5-mm intervals of soil moisture deviation from normal. ~~In general, variation in soil moisture average is higher than in that of precipitation (Fig. 5), again showing how this new water metric is different from previous approaches. Heat and soil moisture combinations vary through the growing season (Fig. 6) The data shows significant month-to-month variation, with the second half of the season facing hotter and dryer days. Also, July has the highest variation in soil moisture deviation.~~

We construct our water metrics based on soil moisture conditions shown in Fig. 1 (extreme surplus = A, surplus = B, around normal = C+D, deficit = E, extreme deficit = F). Three types of metrics are constructed for each condition. A simple metric is the number of days during the growing season with each condition. To show the intensity of each condition, the second metric is defined based on cumulative deviation from normal for each condition. Finally, a compound metric is defined as the sum of degree days for each observed soil moisture condition.

~~3~~2.2~~1~~ Model (1) ~~cumulative precipitation, mean soil moisture, and~~ individual extremes

~~We describe these metrics in Sect. 3.3. Supplementary Materials provide further metrics including the mean evapotranspiration and the mean of soil moisture fraction. The estimation strategy is described in Sect. 3.4.~~

~~In this section, we will review some of the empirical concerns in investigating the individual and compound impacts of heat and water on crop yields in statistical models. Before starting our discussion, we will briefly describe the~~Model 1 is a basic model ~~with~~that uses individual extremes, ~~as introduced by~~following a similar approach as Schlenker and Roberts (2009). ~~The basic model~~Model 1 assumes that the effects of heat on corn yields are cumulative over the growing season and separable from water. In other words, the end-of-season yield is the integral of daily heat impacts over the growing season. This relationship can be demonstrated via Eq. (5~~1~~):

$$y_{it} = \int_{\underline{h}}^{\overline{h}} g(h)\varphi_{it}(h)dh + z_{it}\boldsymbol{\delta} + c_i + \epsilon_{it} \qquad (5\cancel{1})$$

Where y_{it} is crop yield, $g(h)$ is ~~…~~a function showing yield as a function of heat, $\varphi_{it}(h)$ is the time distribution of heat (h) over the growing season in ~~county~~location i and year t, while the heat ranges between the lower bound $\underline{h}$ and the upper bound $\overline{h}$; ~~indicators~~metrics of water availability (e.g., precipitation or soil moisture) and other control factors are denoted as z_{it}, and c_i is a time-invariant ~~county~~ fixed effect~~.~~ All other unobserved variables are in the ϵ_{it} term. 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. Note that this form of equation with fixed effects and unobserved variables is a standard econometric method. We evaluate the accuracy of this model, compared to historical data, using first cumulative precipitation, then mean soil moisture as the water availability metric z_{it}.

~~3.2~~2.3 Model (2) compound extremes ~~and daily interaction of soil moisture and heat~~

Here, we introduce a new statistical model to focus on the ~~daily interaction~~ compound metrics of available water and heat as the major indicators of plant growth to evaluate if including the conditional marginal impact of heat and water on yields provides improved yield estimates. ~~This will show the significance of the conditional marginal impact of heat and water on crop yields:~~Model 2 is:

$$y_{it} = \int_{\underline{m}}^{\overline{m}} \int_{\underline{h}}^{\overline{h}} g(h,m)\varphi(h,m)dhdm \cancel{+ z'_{it}\boldsymbol{\delta}} + c_i + \epsilon_{it} \qquad (6\cancel{2})$$

where y_{it} is the crop yield, $g(h,m)$ is the yield response function to ~~the crop growth is different for~~ each combination of soil moisture level, m, and heat, h; $\varphi(h,m)$ is the distribution of soil moisture and heat;~~,~~ $\overline{m}$ and $\underline{m}$ are upper and lower thresholds of soil moisture; $\overline{h}$ and $\underline{h}$ are maximum and minimum heat;~~other control factors are denoted as z_{it}, and~~ c_i is a time-invariant county fixed effect; and ϵ_{it} is the residual. ~~y is average corn yields.~~ Here, we do not separate the impact of heat from water. In other words, the marginal impact of heat depends on water; and the marginal impact of water depends on heat. ~~We will consider~~The ~~different approaches~~ used ~~to estimate this model as described in Sect. 3.4.~~

~~3.3 Data~~

~~In estimating the marginal impact of soil moisture on corn yields, we employ information about soil moisture, temperature, precipitation, and corn yields for counties of the United States for the 1981-2015 period as summarized in Table 1. The data on yield is obtained from USDA-NASS (United States Department of Agriculture-National Agricultural Statistics Service) at the county level. The yield is defined as the corn production (in bushels) divided by harvested area (in acres). Precipitation is defined in millimeters as accumulated rainfall during the growing season (Apr-Sep). It is calculated based on PRISM (Parameter-elevation~~

~~Regressions on Independent Slopes Model) daily information at 2.5 x 2.5 arcmin grid cells over the continental US for 1981-2015. It is aggregated to each county according to cropland area weights. Daily interaction of heat and soil moisture is also calculated daily at the gridded level. Then we aggregate the metrics to the growing season and county level.~~

~~3.3.1 Degree days (index for heat)~~

~~Following D'Agostino and Schlenker (2015), the daily distribution of temperatures is approximated assuming a cosine function between the daily minimum and maximum temperature. Let~~ $\bar{t} = \mathrm{acos}\left(\frac{2b - T_{max} - T_{min}}{T_{max} - T_{min}}\right)$~~, then degree days at each day is defined using~~

$$D(b) = \begin{cases} \frac{(T_{max} + T_{min})}{2} - b & if\ b \leq T_{min} \\ \frac{\bar{t}}{\pi}\left[\frac{(T_{max} + T_{min})}{2} - b\right] + \frac{(T_{max} - T_{min})}{2\pi} sin(\bar{t}) & if\ T_{min} < b \leq T_{max} \\ 0 & if\ T_{max} < b \end{cases}$$

~~where b is the base for calculating degree days and can take the base values as well as critical values. we consider a piecewise-linear function to aggregate the degree days. The major assumption is that plant growth is approximately linear between two bounds. Degree days between two bounds is simply degree days above the smaller bound minus degree days above the larger bound.~~

~~We calculate county-level seasonal degree days based on daily weather information. The weather information on daily maximum and minimum temperature are obtained from PRISM at 2.5 x 2.5 arcmin grid cells over the continental US for 1981-2015. Degree days are initially calculated for each day at each 2.5 x 2.5 arcmin grid cell during the growing season (Apr-Sep). Then they are aggregated for the whole growing season from the first day of April through the last day of September. Finally, they are aggregated to the county level using cropland area weights.~~

~~3.3.2 Soil moisture (index for water availability)~~

~~Daily soil moisture content and soil moisture fraction are obtained from the Water Balance Model (Grogan, 2016; Wisser et al., 2010) based on daily simulations using PRISM data at 6 x 6 arcmin grid cells for the 1981-2015 period over the continental US. Here, we briefly describe WBM's soil moisture module. However, the model is much more complex and employs a large list of inputs. Full documentation for WBM can be found in Wisser et al. (2010) with updates in Grogan (2016). In WBM, crop-specific soil moisture balance within each grid cell is calculated with an accounting system that tracks a location's water inputs and outputs and is limited by the soil moisture pool's water holding capacity.~~

$$\frac{\delta W_s}{\delta t} = \begin{cases} g(Ws)(I - PET)\ if\ I < PET \\ I - PET\ if\ PET\ \leq I\ and\ (I - PET) < (W_{cap} - W_s) \\ W_{cap} - W_s\ if\ PET\ \leq I\ and\ (W_{cap} - W_s) \leq (I - PET) \end{cases}$$

~~where W_s is soil moisture, t is time, I is the sum of all water inputs to the soil moisture pool, PET is potential evapotranspiration, and W_{cap} is available water capacity. Water inputs to the soil come in the form of precipitation as rain and as snowmelt. Water intercepted by the canopy reduces precipitation reaching the soil. Here, we use the Hamon method for estimating PET (Hamon, 1963; Federer et al. 1996), and $g(W_s)$ is 1 for all crops. Crop-specific potential evapotranspiration values, PETc, are calculated following the FAO-recommended crop-modeling methodology outlined in Allen et al (1998):~~

$$PET_c = k_c \cdot PET$$

~~where k_c [-] is a crop-specific, time-varying scalar. Crop scalar values are from Siebert and Döll (2010), and crop maps that identify the area of each rainfed crop type within a grid cell are from the Crop Data Layer (CDL, USDA NASS, 2017). When soil moisture is insufficient for crops to extract water equal to PET_c, actual crop evapotranspiration is limited to available soil water volumes.~~

~~Available water capacity, W_{cap}, is a function of vegetation-specific rooting depth, a crop-specific depletion factor, soil field capacity, and soil wilting point:~~

~~$W_{cap} = D_c R_c (F - W_p)$~~

~~where D_c is the depletion factor for crop c, R_c is the rooting depth of crop c, F is the soil field capacity, and W_p is the soil wilting point. Here we use the Harmonized World Soil Database (Fischer et al. 2008) as model input for all soil properties. Corn rooting depth is set to 1 meter and the corn depletion factor is 0.55; and the depletion factor is 0.5, following Siebert and Döll (2010). Once the soil moisture content reaches field capacity, no further water is added to the soil moisture pool; excess inputs move to the groundwater pool via percolation and the river system via runoff.~~

~~Fig. 3 shows the scatter plot of cumulative precipitation versus mean soil moisture for US counties for the growing season from 1981 to 2015. The simple correlation coefficient between them is 0.44. This ensures the soil moisture output is not a simple linear transformation of precipitation data. We have also investigated other correlations including the correlation between mean soil moisture and evapotranspiration as illustrated in Fig. A2; and the correlation between mean soil moisture and mean daily soil moisture fraction as shown in Fig. A3.~~

~~One limitation for historical analysis is the inconsistency of WBM and PRISM grid cells as they have different extent, different resolution, and non-matching centroids. Therefore, we interpolate WBM to PRISM using nearest neighbor and bilinear methods. The main regression results are reported for bilinear interpolation. However, the regression results using the nearest neighbor interpolation method are very similar (Table A6). The interpolation provides soil moisture information at 2.5 x 2.5 arc-minute grid cells.~~

~~Soil moisture in the model is calculated as the mean of soil moisture content (in mm for the 1000 mm topsoil) during the growing season (Apr-Sep) for each 2.5 x 2.5 arcmin grid cell. For the interaction of soil moisture and heat, we sum up degree days for each temperature interval (5°C) for each soil moisture deviation interval (10 mm) for each 2.5 x 2.5 arcmin grid cells for the 1981-2015 period. These fine-scale metrics are checked with satellite scans of cropland area to exclude grid cells with no cropland. Finally, we aggregate all the grids in each county using cropland area weight.~~

~~Fig. 4 displays "normal" soil moisture, which is the temporal average of daily soil moisture data over Apr-Sep over 1981-2015, for the Continental US. This map shows the rich heterogeneity of these data across the nation. However, there are distinct regional patterns. For the Corn Belt, the soil moisture level is relatively high compared to other regions. To operationalize this metric, we consider soil moisture deviation from normal. Soil moisture deviation is defined as daily soil moisture minus the normal soil moisture levels. The soil moisture level is considered critical if it is 25 mm below/above normal condition. The threshold is obtained by testing the impacts of 5-mm intervals of soil moisture deviation from normal.~~

~~This implies that the variation in seasonal mean soil moisture may not follow the variation in seasonal mean precipitation. Figure 5 illustrates the year-on-year variation of the precipitation and soil moisture indexes aggregated over the corn growing areas in the US. In general, variation in soil moisture average is higher than in that of precipitation.~~

~~Fig. 6 shows the bivariate density of daily temperature and soil moisture ratio to normal for all the grid cells in the Corn Belt for 1981-2015 by the month of the year, capturing the daily variation of the heat and soil moisture combinations. The data shows significant month-to-month variation, with the second half of the season facing hotter and dryer days. Also, July has the highest variation in soil moisture deviation.~~

~~3~~2.4 Estimation strategy

For Model (1), we build on Schlenker and Roberts (2009) by including different representations of water variables. In Model (1-a), z_{it} includes cumulative precipitation from the first day of April to the last day of September and its square term; this will evaluate the standard way yields have been estimated in previous studies. In Model (1-b), z_{it} is the seasonal mean soil moisture index and its square term, used to evaluate the use of soil moisture instead of precipitation. Model (1-c) includes the number of days with low soil moisture as well as the number of days with high soil moisture, evaluating the importance of extreme soil moisture events (Fig. 1). In Model (1-d), z_{it} includes metrics of soil moisture below or above normal levels, evaluating the importance of extreme soil moisture intensity (Fig. 1).

For Model (1), we assume a piece-wise linear form for $g(h)$. We include degree days above 29°C ~~(~~as a ~~indicator~~ metric of extreme heat~~)~~ as well as degree days from 10 to 29°C ~~(~~as a metric ~~indicator~~ of beneficial heat~~)~~. 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):

$$\cancel{y_{it} = \alpha D_{it}^{10-29C} + \beta D_{it}^{29} + z_{it}\delta + \lambda_s T_t + \lambda_s T_t^2 + c_i + \epsilon_{it}} \qquad \cancel{(1')}$$

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \delta_a P_{it} + \delta'_a P_{it}^2 + \lambda_s t + \lambda'_s t^2 + c_i + \varepsilon_{it} \qquad (7)$$

where i is an index for counties, t is the index of time, s is the index for states, y_{it} ~~shows~~ is the log corn yields, D_{it} represents growing degree day variables, ~~z_{it} includes is metrics indicators of water conditions~~ P shows cumulative precipitation over the growing season, ~~T~~ t shows the time trend variable (~~T~~ 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:

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \delta_b M_{it} + \delta'_b M_{it}^2 + \lambda_s t + \lambda'_s t^2 + c_i + \varepsilon_{it} \qquad (8)$$

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:

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \delta_c N_{it}^{def} + \delta'_c N_{it}^{sur} + \lambda_s t + \lambda'_s t^2 + c_i + \varepsilon_{it} \qquad (9)$$

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):

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

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

We assume the errors are serially correlated due to unobservable and systematic measurement errors, and we consider clustering US counties by the state which has been a standard approach in the literature (Blanc and Schlenker, 2017; Hsiang, 2016; Lobell and Burke, 2010). In this study, the models are estimated using a panel fixed-effect approach. 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 those data collected 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. The panel consists of 35 years (1981-2015) for all US counties with corn production. For purposes of model comparison, we provide adjusted R^2, Akaike's information criterion (AIC), and Bayesian information criterion (BIC).

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. ~~The~~ A key empirical challenge arises when estimating the model with daily interaction of heat and soil moisture. A simple multiplicative interaction of soil moisture variable and heat variables will be problematic (Hainmueller et al., 2019). It implies a linear interaction effect that changes at a constant rate with heat. However, as will be shown below, soil moisture has a non-linear marginal effect. We take two approaches here to calculate the conditional marginal impact of heat on corn yields to address the challenges of aggregating daily soil moisture to seasonal water availability metrics.

First, we construct a binning estimator based on daily interaction on heat and soil moisture in model (2-a). We define several intervals of soil moisture (SM) represented by daily dummy variables and we interact these dummy variables with the daily excess heat index~~.~~ of ~~We considered~~ 29ºC~~ a critical temperature for heat~~. Also, we take 25 mm intervals for soil moisture deviation from normal. In other words, we split the degree days into degree days conditional to soil moisture conditions. This includes dday29˚C & SM 75+ mm below normal (extreme deficit), dday29˚C & SM 25-75 mm below normal (deficit), dday29˚C & SM 0-25 mm around normal (normal), dday29˚C & SM 25-75 mm above normal (surplus), and dday29˚C & SM 75+ mm above normal (extreme surplus). We estimate a coefficient for each combination of excess heat and soil moisture~~. In other words~~; ie., we ~~will~~ estimate a model with metrics ~~indicators~~ of degree days while controlling for soil moisture. The model ~~will~~ provides the conditional marginal impact of excess heat as:

$$y_{it} = \alpha D_{it}^{10-29} + \left\{ \sum_m \beta_m D_{mit}^{29} \right\} + \delta M_{it} + \delta' M_{it}^2 + \lambda_s t + \lambda_s' t^2 + c_i + \varepsilon_{it}$$

(11~~2-a~~)

where i is the county index, t is the time index, m is an index of soil moisture condition (high, low, normal), s is an index for states, y ~~shows~~ is average corn yields, D represents conditional growing degree day variables, M shows the seasonal mean soil moisture content, T stands for the time trend variable, c_i is a time-invariant county fixed effect. Here, β is indexed by m~~. In other words~~; i.e., the marginal impact of heat is conditional to soil moisture conditions. α, β, δ, λ are the regression parameters showing the marginal impacts.

Second, we estimate a model with metrics ~~indicators~~ of soil moisture while controlling for temperature in model (2-b). We define an index of soil moisture when the temperature is above the threshold and an index of soil moisture when the temperature is below the threshold. In this model, the soil moisture is separated by a temperature threshold H^*.

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \left\{ \sum_m \delta_m \left. M_{mit} \right|_{H<H^*} + \delta_m' \left. M_{mit} \right|_{H>H^*} \right\} + \lambda_s t + \lambda_s' t^2 + c_i + \varepsilon_{it}$$

(12~~2-b~~)

where i is the county index, t is the time index, m is an index of soil moisture condition, s is an index for states, y shows average corn yields, D represents growing degree day variables, M shows conditional seasonal mean soil moisture, T stands for the time trend variable, H is the average daily temperature, H^* is the temperature threshold, and c_i is a time-invariant county fixed effect. Here, we define δ and δ' to test whether the marginal impact of soil moisture depends on heat. The soil moisture metrics ~~indicators~~

are calculated from daily gridded data and aggregated to county and growing season. This includes the index of normal soil moisture (*SM* 0-25+ mm around normal) when $H > H^*$, the index of normal soil moisture when $H < H^*$, the index of moisture deficit (*SM* 25+ mm below normal) when $H > H^*$, index of moisture deficit when $H < H^*$, the index of moisture surplus (*SM* 25+ mm above normal) when $H > H^*$, and the index of moisture surplus when $H < H^*$. $\alpha, \beta, \delta, \lambda$ are the regression parameters showing the marginal impacts.

2.5 Decomposition method

To show the significance of weather variation for crop yields, we estimate the historical impacts of heat and water. In a general form, we can decompose the impacts by taking the total derivative from the yield function. The general form is:

$$dy = \underbrace{\frac{\partial y}{\partial h} dh}_{\text{heat impacts}} + \underbrace{\frac{\partial y}{\partial m} dm}_{\text{water impacts}} \tag{13}$$

where dy shows the deviation of crop yields from the trend, dh is the deviation of heat from thehistorical mean; and dm is the deviation of soil moisture from normal levels. We apply this to Model (2-a) while the trend is estimated assuming no variation in heat and water availability. We predict the overall variation in yields using the estimated coefficients of Model (2-a):

$$d\hat{y} = \frac{\partial y}{\partial D^{10-29}} dD^{10-29} + \sum_m \frac{\partial y}{\partial D_m^{29}} dD_m^{29} + \frac{\partial y}{\partial M} dM + \frac{\partial y}{\partial M^2} dM^2 \tag{14}$$

where d shows the differential, $d\hat{y}$ is the predicted variation of crop yields, and partial derivatives are the estimated coefficients. Then, we re-predict the yields using the estimated coefficients of Model (2-a) for normal soil moisture. Thus, the predicted variation in crop yields is driven only by the variation in observed heat.

$$d\hat{y}^{heat} = \frac{\partial y}{\partial D^{10-29}} dD^{10-29} + \frac{\partial y}{\partial D_{nl}^{29}} dD^{29} \tag{15}$$

Finally, the difference between (14) and (15) shows the predicted impact of variation in water.

$$d\hat{y}^{water} = d\hat{y} - d\hat{y}^{heat} \tag{16}$$

Note that the deviations are calculated for each year.

3~~4~~ Results

The overall simulation results from WBM are illustrated in Fig. 2-4, showing gridded historical mean for the cultivated continental US, average annual variations for the cultivated continental US, and bivariate distribution of soil moisture and heat for the corn growing grid cells. To illustrate the spatial heterogeneity, Fig. 2 shows the growing season mean soil moisture content (in mm in 1000 mm topsoil) as calculated based on daily root-zone soil moisture level from Apr-Sep for 1981-2015 at 2.5 x 2.5 arcmin grids excluding non-cultivated area. Average growing season soil moisture is heterogeneous across the Continental US, with distinct regional patterns (see Fig. 42). For the corn belt, the soil moisture level is relatively high compared to other regions. The mean of volumetric soil moisture ranges from below 50 mm in southern California to above 250 mm in the Corn Belt and around Mississippi.

To compare the variation of simulated soil moisture and precipitation, Fig. 2 illustrates the weighted average soil moisture and precipitation over the cultivated US for 1981-2015. In general, variation in soil moisture average is higher than in that of precipitation (Fig. ~~5~~3), ~~again~~ showing how this new water metric is different from previous approaches. One interesting finding is that for some years the mean precipitation and the mean soil moisture move in opposite directions. For example, in 1990 the mean precipitation is declined by around 5% while the mean soil moisture is increased by around 13%.

To show the dynamics of soil moisture and heat, Fig. 4 shows their bivariate distribution by month based on daily information for all the cultivated grid cells in the US Corn Belt for 1981-2015. Heat and soil moisture combinations vary through the growing season (Fig. ~~6~~4) The data shows significant month-to-month variation, with the second half of the season facing hotter and dryer days. Also, July has the highest variation in soil moisture deviation with a high probability of compound extremes as the distribution moves toward the lower right.

Below, ~~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 metrics 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 variation in U.S. corn yields. This point is illustrated in Fig. 5~~ure 7~~, which compares Model 1a ~~(a-d range)~~ to Model 2a: each model estimates the percentage change in corn yields assuming additional 10 degree-days above 29˚C and no change in mean soil moisture. ~~. The error bars show the 95% confidence interval when simply taking the standard errors of the estimation. The shaded area is the 95% confidence interval when using Model 2.~~ The figure shows that Model (1) would significantly underestimate the damage for conditions with extreme water surplus or extreme water deficit.

3.1 Model (1): predicting yield responses to ~~cumulative precipitation, mean soil moisture, and~~ individual extremes

~~This section provides estimation results for different representations of Model (1). We will discuss the implications of these results in Sect. 5. Regression coefficients, standard errors, R-squared, AIC, and BIC values for Models (1-a), (1-b), (1-c), and (1-d) are reported in Table 2.~~ The ~~first column~~ results from Model (1-a) show a strong relationship between corn yields and heat and precipitation (Table 2 column 1-a). The marginal impact of a degree-day within 10-29˚C is significantly positive while that from an additional degree day above 29˚C is strongly negative, confirming the seminal findings of Schlenker and Roberts (2009).

The results from Model (1-b)~~second column~~, excluding precipitation, shows the marginal relationship with soil moisture is also significant (Table 2 column 1-b). This confirms the findings of Ortiz-Bobea et al. (2019). It shows that the marginal relationship with soil moisture is increasing up to ~92 mm in 1000 mm topsoil and decreasing for higher values.

In Model (1-c), we consider the number of days that soil moisture is either too high or too low. The model with metrics of soil moisture extremes further improves the fit, revealing a negative marginal relationship associated with the number of days with low/high soil moisture. Regarding Model (1-c), the coefficient on the number of days with low moisture is also significant and negative. Our estimation sample shows 26 days of high soil moisture and 27 days of low soil moisture on average. The implication is that eliminating 25 days of high soil moisture and 25 days of low soil moisture can improve the corn yields by up to 12.6%.

Model (1-d) shows the estimated coefficients when considering surplus and deficit (soil moisture deviation from normal) instead of average seasonal soil moisture. Here, we consider two thresholds for low and high soil moisture. Returning to Fig. 1, we evaluate the area of all blue bars and the area of all red bars. It shows that the marginal impact of the moisture deficit (cumulative negative soil moisture deviation) is significant and positive. This indicates the positive contribution of additional soil moisture when the soil moisture levels are below normal. On the other hand, the marginal impact of additional soil moisture in a wet period – i.e., a positive soil moisture deviation -- is negative. In other words, this measure captures the fact that plants will benefit from reductions in soil moisture when the soil moisture levels are above normal. This is an indicator of the value of sub-surface drainage for agriculture. Note that the Model (1-d) decreases the marginal relationship with extreme heat (~~DD29˚C~~dday29˚C). However, this effect is not statistically different from that produced by the first model. ~~A central finding is that metrics of soil moisture extremes are statistically significant.~~

The coefficient of the deficit in Model (1-d) is significant and positive. On the other hand, the coefficient of the extreme deficit is also significant and positive. The estimation sample shows this metric ~~indicator~~ is around 2300 mm on average. It indicates that reducing the deficit ~~and~~ by 2300 mm and reducing the surplus by the same amount can improve the corn yield by up to 21.2% on average. Note the mean soil moisture can stay unchanged in this scenario.

3.2 Model (2): predicting yield responses to compound extremes ~~and daily interaction of soil moisture and heat~~

~~We introduce heat-soil moisture interactions to test whether soil moisture availability changes the marginal impact of heat. In Model (2-a), we estimate a model while splitting the heat stress index according to soil moisture conditions. Here we construct the heat index for different intervals of soil moisture deviation. Table 3 shows the estimation results. As shown in this table,~~ In Model (2-a) we introduce heat-soil moisture interactions to test whether soil moisture availability changes the marginal impact of heat on yields (estimation results are in Table 3). We find that the average marginal impacts of dday29˚Cs (heat stress) are all significant. The coefficient on dday29˚C ~~(heat stress)~~ combined with the extreme deficit is -0.0082. The coefficient of ddays29˚C (heat stress) combined with extreme water surplus is -0.0140. These figures are significantly different compared to Model (1). ~~Figure 7 illustrates these coefficients while they are translated to the percentage change in corn yields assuming additional 10 degree-days above 29˚C and no change in mean soil moisture. The error bars show the 95% confidence interval when simply taking the standard errors of the estimation. The shaded area is the 95% confidence interval when using Model 2. The figure shows that Model (1) would significantly underestimate the damage for conditions with extreme water surplus or extreme water deficit.~~

W~~Finally, w~~e estimate a model with soil moisture while controlling for temperature (2-b). The results are presented in Table 4. The coefficient of degree days from 10˚C to 29˚C is significant and positive. This is not significantly different from previous models (1-a, 1-b, 1-c, 1-d, and 2-a). The coefficient on degree days above 29˚C is significant and negative. It is close to the estimated values from Model (2-a) but slightly lower than Model (1). This indicates that the average damage from extreme heat index (dday29˚C) is around 25% lower than Model (1). The estimated parameters show the yield response to changes in soil water content. Comparing the parameter values can show the difference in yield response to soil moisture in hot weather and moderate weather. The coefficient on normal soil moisture conditional to hot weather is 0.00012. The coefficient on normal soil moisture conditional to moderate weather is 0.00003. This indicates that the yield response to ~~-~~water is up to four times ~~more valuable~~ higher in hot weather. The marginal impact on soil moisture deficit index is 0.00009 in hot weather and is 0.00002 in moderate weather. This also supports the finding that water is up to four times more ~~valuable~~ beneficial to corn yields in hot weather. Also, the results ~~suggest~~ show that the damage from excess water is up to two times ~~bigger~~ larger in hot weather.

3.3 Model comparison

~~We have presented estimated coefficients for different models for predicting corn yields with individual and compound extremes.~~ A comparison of model performance metrics is given in Table 5, along with a description of the water metric and extreme metric used in each model. We find that for Model 1b-d and Model 2a-d the coefficients on the soil moisture metrics are significant and with expected signs. ~~We also find that *the average damage from excess heat has been up to four times more severe when combined with water stress.*~~ 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 both previous models (with cumulative precipitation or with mean soil moisture). T~~Finally, t~~he best corn yield predictor ~~performance~~ is from Models (2-a) and (2-b), considering compound extremes through the daily interaction of heat and soil moisture~~.~~.

We find that using a seasonally averaged soil moisture metric is insufficient for capturing yield extremes; i.e., the temporal resolution of the soil moisture metric is important for estimating corn yield variability. Figure 6~~8~~ illustrates the difference by comparing the modeled impacts of average soil moisture (Model 1-b) on corn yields (Panel a) to the impacts considering the deviation from normal soil moisture (Model 1-d) estimated for a sandy soil type (Panel b) and a clay soil type (Panel c). In other words, when parametrizing the soil moisture as a deviation from normal, we get a specific piece-wise linear yield response to water depending on soil types (and normal levels of soil moisture), the extremes of which are completely missed by the model that only uses mean soil moisture. We find that *the average corn yield damage from excess heat is up to four times more severe when combined with water stress.* This damage can only be estimated when including soil moisture and metrics of extreme water stress (e.g., Models 2a-d).

3.4 Decomposing the variation in US corn yields

We have decomposed the changes in the US corn yields from 1981 to 2015 to understand the relative roles of soil moisture and heat in interannual corn yield variation. Figure 7~~10~~ illustrates a decomposition based on our findings while aggregated for the whole US. With no climate variation, the US corn yield is expected to have a smooth positive trend as shown in green color. The deviation from the trend occurs due to changes in water and heat stressors. The blue bars are showing the expected changes in US corn yields due to changes in the water stress while the orange bars are demonstrating the expected yield changes due to changes in heat stress. While there have been years in which the stressors have moved together (e.g. 2011 and 2012), for several years water and heat have offset each other's benefit or damage. For example, in 1992 the damage from heat is partially offset by benefits from water. Or in 2010, the damage from water stress is partially offset by benefits from heat.

3.5 Robustness checks

The Supplementary Material provides 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. Overall, the findings remain robust to alternative soil moisture metrics from WBM including the mean of soil moisture fraction (soil moisture content divided by field capacity), the seasonal mean of evapotranspiration as well as within season standard deviation of them. We also ~~report~~ look at the results ~~for~~ 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, ~~for~~ time separability, we re-estimate Model (1-b) for two-month intervals (Apr-May, Jun-Jul, Aug-Sep) ~~as well as for the whole season~~, 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 geographical area, we re-estimate the models for Eastern US and 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.

~~5~~4 Discussion~~ and robustness checks~~

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 (Bevacqua et al., 2019; Manning et al., 2019; Myhre et al., 2019; Poschlod et al., 2020; Potopová et al., 2020; Wehner, 2019; Zscheischler et al., 2018)~~ (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 Fig.~~ure~~ 1. As ~~There is an increasing body of literature in climate sciences about the changes in the likelihood of compound extreme events (Zscheischler et al., 2018; Manning et al., 2019; Bevacqua et al., 2019; Poschlod et al., 2020; Potopová et al., 2020; Wehner, 2019). Therefore, this research is critical for evaluating the impacts of future climate change as w~~we ~~found~~find that the coefficient on extreme heat is significantly different when considering soil moisture,~~. We have not investigated the size of the overestimation or underestimation of climate impact studies. However, predictions of significant changes in precipitation and soil moisture within the growing season suggest that the impact could be substantial.~~ 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~~ing climate impact studies~~ to over-estimating~~g~~ yield damages under future climate ~~due to~~ extremes by not accounting for human adaptations designed to conserve soil moisture.

~~We have presented estimated coefficients for different models for predicting corn yields with individual and compound extremes. We find that the coefficients on the soil moisture metrics are significant and with expected signs. We also find that *the average damage from excess heat has been up to four times more severe when combined with water stress.* 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 both previous models (with cumulative precipitation or with mean soil moisture). Finally, the best performance is from Models (2-a) and (2-b), considering compound extremes through the daily interaction of heat and soil moisture.~~

~~The appendix of this paper provides several robustness checks. Overall, the findings remain robust to alternative soil moisture indicators from WBM including the mean of soil moisture fraction (soil moisture content divided by field capacity), the mean of evapotranspiration as well as within season standard deviation of them. We also report the results for an alternative interpolation of WBM data to PRISM resolution (nearest neighbor versus bilinear). To test for time separability, we estimate Model (1-b) for~~

~~two-month intervals (Apr-May, Jun-Jul, Aug-Sep) as well as for the whole season. Finally, we have provided different interaction models. Here we discuss the implications of our findings for climate-agriculture-water studies, as well as in broader literature of climate impact studies.~~

~~5.1 Should we replace precipitation with soil moisture in climate-related studies?~~

~~Cumulative precipitation and mean soil moisture can be strongly or weakly correlated. The main factors in the difference between the two are runoff, drainage, and irrigation. If the runoff is rare and there is little or no irrigation, there is a high chance that cumulative precipitation and mean soil moisture are strongly correlated. If they are expected to stay highly correlated, then adding soil moisture to the model may not benefit the researchers. Precipitation may still be a valid measure of future water availability for locations with small runoff, drainage, and irrigation.~~

~~However, as drainage becomes more attractive in the Eastern US and irrigation dominates in the Western US, the correlation between cumulative precipitation and mean soil moisture weakens considerably. A quick test for a given locality would involve looking at the share of irrigated area, the area equipped with drainage systems, the share of intensive precipitation, or the number of days without precipitation. If any of these are dominant or expected to be dominant in the future, we recommend the use of soil moisture data or other metrics of water availability as suggested in the literature but not cumulative precipitation.~~

~~The studies using degree days above a critical threshold may capture part of the damages from low soil moisture. This can be due to feedbacks and dynamics of temperature and soil moisture as abundant soil moisture can reduce the temperature and extreme heat can reduce the soil moisture (D'Odorico and Porporato, 2004; Seneviratne et al., 2010). While previous studies are unable to capture the different impacts of dry heat versus wet heat (Feng and Zhang, 2015; Schoof et al., 2017), our study suggests that the impact of excess heat can be significantly different while considering soil moisture.~~

~~5.2 Should we use mean soil moisture or deficit-surplus metrics?~~

~~Does it make a difference to consider seasonal mean soil moisture or metrics of extreme conditions? Figure 8 illustrates the difference by comparing the impacts of soil moisture on log corn yield using the estimated coefficients. The black curve in Fig. (8-a) shows the relationship between soil moisture and log corn yield from Model (1-b), without the extreme conditions and the interaction term. This indicates a more general relationship that looks like an envelope to local functions. Panel b and c show the relationships considering the deviation from normal in Model (1-d) drawn for a clay soil type (c) and sandy soil type (b). In other words, when parametrizing the soil moisture as a deviation from normal, we get a specific piece-wise linear yield response to water depending on soil types (and normal levels of soil moisture).~~

~~In addition, an examination of extreme conditions can improve our understandings of climate impacts with intensive extreme events. 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 a bias.~~

~~Here is an example from our sample. Consider the case of Bureau County in Illinois (FIPS: 17011). Our data show that the mean soil moisture is almost the same in 1983 and 1992 and is around 125 mm. However, soil moisture deficit and surplus metrics are quite different. They are -5,934 mm and +4,937 mm for 1983, and -2,451 mm and +2,137 mm for 1992, respectively. As a result, Model (2) predicts almost no impact from the change in soil moisture, while Model (3) predicts a 24.7% increase in corn yield due to changes in soil moisture. While this may be counted as a rare incidence historically, this may not be the case in the future. As climate models are warning about significant changes in the frequency and intensity of extreme precipitations (Myhre et al., 2019),~~

~~the mean metrics of water availability are not enough in capturing the impacts of water on yields. It is necessary to consider the metrics of extreme events in the models.~~

~~5.3 Should we use daily interaction of soil moisture with heat metrics?~~

~~Recall Fig. 6 which illustrates the historical dynamics of change in the daily distribution of heat and simulated soil moisture over the Corn Belt for the 1981-2015 period. Throughout the growing season, the density moves in the direction of lower soil moisture and warmer conditions. If a location is expected to face minimal changes in the bivariate distribution of heat and water availability, adding interaction terms will benefit the analysis relatively little. However, if a significant change is expected in compound extreme events, then the use of models with interaction terms is inevitable.~~

~~There is an increasing body of literature in climate sciences about the changes in the likelihood of compound extreme events (Zscheischler et al., 2018; Manning et al., 2019; Bevacqua et al., 2019; Poschlod et al., 2020; Potopová et al., 2020; Wehner, 2019). Therefore, this research is critical for evaluating the impacts of future climate change as we found that the coefficient on extreme heat is significantly different when considering soil moisture. We have not investigated the size of the overestimation or underestimation of climate impact studies. However, predictions of significant changes in precipitation and soil moisture within the growing season suggest that the impact could be substantial.~~

~~However,~~ A~~a~~pplying this framework to climate impact studies ~~may~~ will face a key challenge – namely projecting the future compound extremes with the high temporal resolution of Model 2. It requires collaboration between hydrologists, climate scientists, and statisticians (Zscheischler et al., 2020). For future yield projections, we need reliable future projections of daily temperature (maximum and minimum) and soil moisture. Unfortunately, to the best of our knowledge, available data sets including predictions of future soil moisture have a relatively coarse spatial and temporal resolution, and rely on climate model projections with known difficulties representing daily temporal resolution events (Hempel et al., 2013). ~~While various climate products are projecting future daily temperatures, the choice of climate model requires extreme caution and should be compatible with the special needs of each study. Although there are some projections of future levels of soil moisture, there is a great deal of inconsistency among the models regarding this variable.~~ Further research is required to improve the ability of climate models and impact models in projecting the bivariate distribution of heat-moisture (Sarhadi et al., 2018).

~~In places predicted to face higher mean precipitation coupled with more extreme water stress, adaptation through soil moisture management will be beneficial to farmers. This may motivate investments in supplemental irrigation. Also, farm management practices such as no-till farming, cover cropping, and soil conservation can increase soil moisture without (or in addition to) irrigation. Farmers may also consider improvements in water use efficiency, both by crops and by irrigation systems, as one way to address the need for increased irrigation. However, the expansion of irrigation can increase the stress on global water resources.~~

~~5.4 Implications for irrigation water demand and subsurface drainage~~

~~Considering the estimated coefficients for Model (2), we construct the daily marginal value product of soil moisture conditional on a given soil type and temperature. The economic literature on the value of water offers a variety of techniques to estimate the value of irrigation water in agriculture (Aubuchon and Morley, 2013; García Suárez et al., 2018; Gemma and Tsur, 2007; Griffin, 2016; Mesa-Jurado et al., 2010; Mukherjee and Schwabe, 2014; Rigby et al., 2010; Young, 2010). The agronomic literature also considers deficit and supplemental irrigation (Hargreaves et al., 1989; Hargreaves and Samani, 1984) and crop water productivity (Kang et al., 2009; Zwart and Bastiaanssen, 2004). Here we employ the marginal value product (MVP) approach to estimate the value of water (Costanza et al., 1997; Griffin, 2016; Young, 2010). In this approach, the impacts of the change in the water input~~

are estimated while assuming other inputs are constant. Assuming a general form of production function as $Y = Y(L, W, H)$ where Y is the output, L is augmented land, W is water, and H is heat. The MVP of water is given by:

$$P_Y \frac{\partial Y(L, W, H)}{\partial W} = MVP_w$$

where P_y shows the price of output. If farmers make their decisions about other inputs before the planting date, then the variation in Y is mainly due to variation in W and H. Figure 9 shows one example of this marginal productivity assuming clay soil type (with the normal moisture around 200 mm) and temperature around 25C. In the left panel, the marginal contribution is displayed. In the right panel, the marginal value product is illustrated assuming the price of corn is \$3.5 per bushel. Note that even at normal soil moisture levels, the value of water is positive.

5.5 Decomposing the variation in US corn yields

We find that *the average damage from excess heat has been up to four times more severe when combined with water stress.* To illustrate the significance of this finding, we have decomposed the changes in the US corn yields from 1981 to 2015 considering soil moisture and heat. Figure 10 illustrates a decomposition based on our findings while aggregated for the whole US. With no climate variation, the US corn yield is expected to have a smooth positive trend as shown in green color. The deviation from the trend occurs due to changes in water and heat stressors. The blue bars are showing the expected changes in US corn yields due to changes in the water stress while the orange bars are demonstrating the expected yield changes due to changes in heat stress. While there have been years in which the stressors have moved together (e.g. 2011 and 2012), for several years water and heat have offset each other's benefit or damage. For example, in 1992 the damage from heat is partially offset by benefits from water. Or in 2010, the damage from water stress is partially offset by benefits from heat.

5.6 Implications for climate studies

The results emphasize the value of soil moisture management as an effective means of adaptation to climate change. This adaptation can moderate production damages from a hot future climate. Thus, we predict that supplemental irrigation will be more beneficial to farmers. However, the expansion of irrigation in many areas may lead to further increases in unsustainable groundwater withdrawals. Such trade-offs are inevitable as environmental stresses in agriculture increase in the future. Furthermore, we confirm that excess soil moisture is damaging for corn and it is intensified when combined with heat stress. This emphasizes the importance of subsurface drainage for crop production in the future.

We have examined the possible impacts of climate change on global corn yields by the mid-century. We employ information from NASA Earth Exchange (NEX) Global Daily Downscaled Projections (NEX-GDDP) product for Representative Concentration Pathway (RCP) 8.5 scenario at 15 x 15 arcmin at the global level. We consider grid-specific growing season according to SAGE growing calendar (Sacks et al., 2010) and carefully calculate the degree days considering leap years. According to this projection, heat stress conditions are expected to increase sharply in the US by the mid-century. Figure 11 shows the climate impact on maize yields at the global level for irrigated and non-irrigated corn based on the CCSM4 model. This figure shows a heterogenous impact around the world. The critical finding is that adaptation through irrigation can significantly reduce the damage of heat stress on corn yields.

These findings are important for assessing the regional resilience of agroecosystems, global food security, and as well as future climate impacts. This framework can help farmers quantify the daily importance of soil moisture for future climate adaptation which can indirectly enhance food security. At the policy level, this study improves our understanding of the implications of

~~compound hydroclimatic extremes which are critical to economic assessments undertaken at the local, national, and global levels. The estimation framework also provides a better measurement of climate-related variables which is also valuable for economic studies. Our findings also provide a significant contribution to the climate impact literature through the estimation of the monetary value of damages from compound hydroclimatic extremes for agriculture.~~

~~Finally, this paper demonstrates the value of fine-scale hydroclimatic information for research in the economics of climate change, global environmental changes, and coupled human and environmental systems. A strength of our findings is that they can be used widely by the research community, as many hydrology and land surface models can simulate soil moisture. Also, this method can be tailored for use with different climate model outputs as well as different soil maps. It can also accommodate the analysis of hypothetical situations (e.g., drought) which may vary by study location and research question at hand.~~

5 Conclusions

This study serves to bridge the gap between statistical studies of climate impacts on crops and their biophysical counterparts by recognizing the central role of soil moisture – which is not a simple linear transformation of precipitation – in understanding crop yields. ~~and underscores that findings of statistical models based on county-level data are in line with experimental agronomic studies (Lobell and Asseng, 2017).~~ We employ a fine-scale, high temporal resolution dataset to investigate the conditional marginal value of soil moisture and heat in US corn yields for the 1981-2015 period employing a statistical framework. The major contribution of this study is showing that the coefficient on extreme heat (dday~~DD~~29˚C) is significantly different while considering daily interactions with soil moisture, emphasizing the importance of compound hydroclimatic conditions.

Our first key finding is that seasonal mean soil moisture performs ~~well~~ better than average precipitation in statistically predicting corn yield. While the majority of current empirical studies employ precipitation as a proxy of water availability for crops, we show that the precipitation coefficient may not be always an appropriate measure of water availability. This study suggests that soil moisture content should be used in estimating crop yields instead of cumulative rainfall for locations with high runoff, drainage, or irrigation (e.g. Western and Central US).

Also, the metrics ~~indicators~~ of soil moisture extremes can explain a portion of the damages to corn yield. On average, farmers can improve corn yields by up to 24% only by avoiding extreme water stress. We also find that the coefficient of excess soil moisture is negative. This is in line with the current agronomic literature (Torbert et al., 1993; Urban et al., 2015) which points out that high soil moisture content can result in nutrient loss through excess water flows. In addition, at high humidity, the plants may have difficulty remaining cool at high temperatures. There is also a risk of waterlogging soils. With a few notable exceptions (e.g., rice), most crops do not grow well in inundated conditions as the plant roots need oxygen, so the direct impact of excess water stress is because of the anoxic conditions.

Finally, the marginal impact of heat index on crop yields depends on the soil moisture level. We show the average yield damage from heat stress ~~has been~~is up to four times more severe when combined with water stress; and therefore the value of water in maintaining crop yield ~~has been~~is up to four times ~~bigger~~ larger on hot days.

~~**Appendix**~~

~~This appendix provides some robustness checks on the results and the model variables. First, we illustrate the relationship between mean soil moisture and other seasonal variables in this study. This includes mean seasonal evapotranspiration, mean seasonal soil moisture fraction, and the degree days above 10˚C. We provide some examples to demonstrate the seasonal mean soil moisture~~

~~shows no linear relationship with the seasonal heat index (degree days above 10˚C). However, it has a positive correlation with evapotranspiration and soil moisture fraction. Then we provide alternative models controlling for irrigation, growth periods, spatial scope of the study, and other measures of individual and compound extremes.~~

~~A.1. Correlation of mean seasonal soil moisture and other variables~~

~~The soil moisture output from WBM is informed mainly by soil moisture memory, heat, precipitation, and many other time-variant and time-invariant information. We have taken two other variables from WBM including soil moisture fraction and evapotranspiration (ET). Also, we have interpolated WBM soil moisture using an alternative method (nearest neighbor method). Section A.6 will provide the estimation results when using these variables to show the robustness of the results to variable selection. Here we plot these variables against the volumetric soil moisture content to illustrate the correlation and differences. As shown in Fig. A1 two interpolations of soil moisture are closely correlated by R= 0.9997. Figures A2 and A3 are the scatter plots of seasonal ET and seasonal mean soil moisture fraction against volumetric soil moisture. The figures show the seasonal variables are not following a simple linear relationship. Figure A4 shows the scatter plot of cumulative growing degree days above 10˚C versus mean soil moisture for US counties for the growing season from 1981 to 2015. This indicates the soil moisture output is not a simple linear transformation of heat data.~~

~~A.2. Robustness check: controlling for normal soil moisture (2-c)~~

~~Here we introduce Model (2-c) trying to control for compound stresses. The idea is that the daily excess heat may come with or without water stress. Thus, the coefficient on the excess heat indicator (DD29˚C) shows the impact of heat stress which usually involves some water stress. We estimate a model with the interaction of heat indicators and index of soil moisture defined as the "share of heat at normal moisture".~~

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \alpha' S_1 D_{it}^{10-29} + \beta' S_2 D_{it}^{29} + \left\{ \sum_m \eta_m MD_{mit} \right\} + \lambda_s T_t + \lambda_s' T_t^2 + c_i + \varepsilon_{it} \qquad \text{(2-c)}$$

~~where i is the index for counties, t is an index for year, y_{it} shows log corn yields by county and year, m is an index of soil moisture bins (high, low, normal), s is an index for the US states, D represents growing degree day variables, MD shows soil moisture metrics, T is a time trend index. For calculating S, we split the degree days according to soil moisture bins. Then, S is defined as the share of degree days around normal soil moisture over seasonal degree days.~~

$$S_1 = \frac{\sum_d D_d^{10-29}\big|_{M \sim \bar{M}}}{\sum_d D_d^{10-29}}, \qquad S_2 = \frac{\sum_d D_d^{29}\big|_{M \sim \bar{M}}}{\sum_d D_d^{29}}$$

~~Here, S1 is the share of DD10-29˚C with normal soil moisture for all DD10-29˚C calculated by county and by growing season; and S2 is the share of DD29˚C with normal soil moisture for all DD29˚C at each county each growing season. This model makes the interpretation easier for decomposing heat stress from water stress. With no water stress, $S_1 = S_2 = 1$, and soil moisture is always around normal levels. Thus, $\alpha + \alpha'$ will show the marginal impact of additional DD10-29˚C; and $\beta + \beta'$ will show the marginal impact of an additional DD29˚C. On the other hand, for water-stressed corn, $S_1 = S_2 = 0$, and soil moisture is not around the normal levels. Therefore, α will show the marginal impact of additional DD10-29˚C combined with water stress; and β will show the marginal impact of an additional DD29˚C combined with water stress. We will discuss this while looking at the estimation results.~~

~~Table A1. provides the estimated coefficients, standard errors, R-square, and AIC and BIC for Model (2-c). The coefficient on the beneficial heat (DD10-29˚C) is significant and positive. The coefficient on its interaction with soil moisture (share of heat at normal soil moisture) is also significant and positive. With no water stress (S1= 1), the point estimate of the marginal impact of the beneficial heat is 0.00036 (+0.00025 + 0.00011). The marginal impact will be 0.00025 when the soil moisture is not at normal levels. The finding that heat is less beneficial with soil moisture deficit is not a surprise and is in line with agronomic literature. However, it has crucial implications for climate impact studies as it suggests a likely overestimation of the benefits of global warming in the context of more erratic and concentrated rainfall events.~~

~~The coefficient on the extreme heat (DD29˚C) is significant and negative. However, the coefficient on its interaction with soil moisture (share of heat at normal soil moisture) is significant and positive. In other words, with no water stress (S2= 1), the marginal impact of extreme heat is -0.0040 (-0.0057 + 0.0017). While with water stress, it is -0.0057. This is another critical finding and approves the results from Models (2-a) and (2-b). It shows when soil moisture is not at normal levels, the average damage from excess heat is around 43% more severe than normal conditions of soil moisture. Note that this is the average damage and the actual damage can be more/less severe depending on the degree of water stress as we show in Model (2-a).~~

~~The estimated coefficients on soil moisture metrics (sum of daily deviations below and/or above a threshold) are significant. This coefficient is -0.000023 for soil moisture surplus and is +0.000050 for normal soil moisture. The important finding is that the signs are as expected and consistent with other models.~~

~~A.3. Robustness check: controlling for irrigation in Model (2-d)~~

~~Here we estimate Model (2-d) controlling for irrigation. While irrigation ensures the soil moisture at normal levels, it can have additional benefits than just providing water. One major impact is the *cooling effect* of irrigation technologies as used in the Western US. Sprinkler irrigation not only reduces the water stress but also makes the air temperature lower near the surface. This will reduce heat stress and is not captured by GDD29˚C based on PRISM. Another point is that the panel focused on the West includes a little variation in soil moisture metrics as suggested by various water deficit and surplus indicators in Table A3. As a result, the soil moisture data for the Western US provides little information on soil moisture variations.~~

~~Here, we investigate whether irrigation has other benefits than just providing soil moisture. To control for irrigation, we re-estimate Model (2) including a term for interaction of DD29˚C and share of irrigated area. Specifically, we estimate this model:~~

$$y_{it} = \alpha D_{it}^{10-29} + \beta D_{it}^{29} + \beta' S_2 D_{it}^{29} + \beta'' S_3 D_{it}^{29} + \left\{ \sum_m \eta_m MD_{mit} \right\} + \lambda_s T_t + \lambda'_s T_t^2 + c_i + \varepsilon_{it} \qquad \text{(2-d)}$$

~~Here, if S2 =1, S3=0 then *β + β'* will show the marginal impact of an additional DD29˚C at normal soil moisture for non-irrigated corn; if S2 =1, S3=1 then *β + β' + β''* will show the marginal impact of an additional DD29˚C at normal soil moisture for irrigated corn. Table A2 shows the results. All the estimated coefficients are significant. In summary, the marginal impact of additional DD29˚C: for non-irrigated corn and with water stress is -0.0062; for non-irrigated corn and without water stress is -0.0044; and for irrigated corn without water stress is -0.0020; In other words, we strongly reject the hypothesis that irrigation benefits are limited to providing soil moisture. This also explains the differences between Western and Eastern US as discussed in Sect. S.5.~~

~~A.4. Robustness check: West versus East in Model (1)~~

~~In this section, we estimate the main models separately for Eastern and Western US. Those counties with centroids on the left of 100th meridian are considered West. The idea is that water stress is less severe in the Western US as it is mostly irrigated. Table A3. provides the main descriptive statistics to compare these regions. Overall, Western US experiences more excess heat by 82~~

~~versus 58 DD29˚C in the East. On average, Eastern US receives 601 mm of cumulative precipitation while it is only 271 mm in the Western US. On the other hand, within-season SD of soil moisture is 39 mm in the East while it is 13 mm in the west. Looking at the number of days with high/low soil moisture, only 11 days in the West soil moisture is not at normal levels, while this is 59 days in the East.~~

~~Table A4. shows the estimated coefficients, standard errors, adjusted R-squared, AIC, and BIC statistics for four models for Eastern US. Model (1-a) includes cumulative precipitation. Model (a-2) includes mean soil moisture metrics. The third model, similar to Model (3b), considers soil moisture extremes. And Model (4) considers the interaction terms. The results suggest that the coefficient on the extreme heat is not significantly different from the estimations for the whole US.~~

~~Table A5. shows the estimated coefficients, standard errors, adjusted R-squared, AIC, and BIC statistics for four models for the Western US. The results suggest that the coefficients on the extreme heat are significantly different from the estimations for the whole US and the Eastern US. For example, the coefficient on DD29˚C is -0.0020 in Model (1) for the West, while it was estimated -0.0056 for the East. This is around 65% lower damage for a given degree day above 29˚C. Also, the AIC and BIC statistics would reject the hypothesis that models with interaction perform better compared to the model with cumulative precipitation. The difference can be a result of the "*cooling effect*" as discussed in Sect. S.4.~~

~~A.5. Robustness check: bi-monthly metrics of soil moisture~~

~~Table A6. provides the estimation coefficients, standard errors, AIC, BIC, and R-squares statistics for Model (1-b) for Eastern, Western, and the continental US with bi-monthly mean soil moisture. The results suggest that the coefficients on extreme heat (DD29˚C) are not significantly different from the model with seasonal mean soil moisture.~~

~~The results suggest that the marginal impact of mean soil moisture is higher in June-July. This is in line with agronomic literature as it suggests the water stress during pollination and the silking stage is more damaging. These stages are the most critical stage of development for corn. Water stress during this stage can cause higher yield loss than almost any other stage in the crop's development.~~

~~The marginal impact of mean soil moisture is not significant in August-September. This suggests that additional soil moisture can have a positive or negative impact on yield. This also makes sense as a high level of moisture can hurt the maturity and drying stage. High soil moisture at the end of the growing season can cause delayed grain maturity and may lead to delay in the harvest.~~

~~In Addition, the marginal impact of mean soil moisture in April-May is negative for the whole US and the Western US and significant at 90% confidence interval. This can be a result of the negative impacts of excess soil moisture on germination and early crop developments as a result of flooding and waterlogging.~~

~~A.6. Robustness check: other metrics from WBM outputs (soil moisture fraction and ET)~~

~~Here, we re-estimate Model (1) with other related metrics of water availability to crops including simulated daily evapotranspiration of rainfed corn (ET) from WBM; daily soil moisture fraction (SMF) from WBM; and soil moisture content from different spatial interpolation of WBM grid cells to PRISM (nearest neighbor method versus original bilinear method).~~

~~The soil moisture fraction index considers the volumetric soil moisture content divided by field capacity. We have also considered the within-season standard deviation of ET and SMF. Note that we keep the degree days above 29˚C as an indicator of heat stress and the degree days from 10˚C to 29˚C as an indicator of beneficial heat to corn.~~

~~Table A7. reports regression results for these models. Columns 1 and 2 show a significant relationship with the mean of soil moisture fraction, its square term, and its within season standard deviation. Columns 3 and 4 with mean ET and within-season SD of ET also show a significant relationship. Column 5 shows that the other interpolation of soil moisture has a very close marginal~~

~~coefficient and standard error compared to our original Model (1). The important finding is the marginal relationship for beneficial and harmful heat remains significant and not significantly different from Model (1).~~

~~A.7. Robustness check: East and West in Model (2)~~

~~Her we re-estimate Model (2). The results are presented in Table A8 and A.9 for the US, West, and East. We see a similar pattern for East versus West. The coefficient on heat stress is smaller for the West which can be a result of the cooling effect.~~

~~The results of Model (2-a) are presented in Table A8. Column 1 shows the results for the whole US while columns 2 and 3 contain the results for the Western US and Eastern US, respectively. According to column 2, the coefficient on dday29°C and the extreme deficit is -0.0074 in the Western US which is significantly different from all other estimations for the Western US. This is another evidence of the cooling effect. These results indicate that, even in the Western US,~~ *~~the damage from heat stress can be up to four times higher when combined with water stress~~*~~. The coefficient on excess heat and the extreme surplus is not significant (note that this is a very rare condition in the West).~~

~~As in column (3) of Table A9, the coefficient on normal soil moisture conditional to hot weather is 0.00010. The coefficient on normal soil moisture conditional to moderate weather is 0.00002. This indicates that water is up to four times more valuable in hot weather. The marginal impact on soil moisture deficit index is 0.00008 in hot weather and is 0.00002 in moderate weather. This also supports the finding that water is up to four times more valuable in hot weather. Also, the results suggest that the damage from excess water is up to two times bigger in hot weather.~~

Code availability. The codes are available at DOI:10.4231/Q07D-J369.

Data availability. The historical weather data (PRISM) is available at http://www.prism.oregonstate.edu. ~~The future weather data (NEX-GDDP) is available at https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp.~~ The input data for estimations are available at DOI:10.4231/0M14-EY38.

Author contribution. All authors contributed to conceptualization, methodology, formal analysis, and writing- review & editing. IH and DSG collected model input, performed the simulations, and contributed to the investigation, resources, software, and validation. IH contributed to writing the original draft and visualization. TWH contributed to supervision and funding acquisition.

Competing interests. The authors declare that they have no conflicts of interest.

Acknowledgement. This work was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program, Earth and Environmental Systems Modeling, MultiSector Dynamics, Contract No. DE-SC0016162. The database construction process is done employing high-performance computing solutions of Institute for CyberScience Advanced CyberInfrastructure (ICS-ACI) at Penn State University.

References

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, Fao Rome, 300(9), D05109, 1998.

Bevacqua, E., Maraun, D., Vousdoukas, M. I., Voukouvalas, E., Vrac, M., Mentaschi, L. and Widmann, M.: Higher probability of compound flooding from precipitation and storm surge in Europe under anthropogenic climate change, Sci. Adv., 5(9), eaaw5531, doi:10.1126/sciadv.aaw5531, 2019.

Blanc, E. and Schlenker, W.: The use of panel models in assessments of climate impacts on agriculture, Rev. Environ. Econ. Policy, 11(2), 258–279, 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.

Bradford, J. B., Schlaepfer, D. R., Lauenroth, W. K., Yackulic, C. B., Duniway, M., Hall, S., Jia, G., Jamiyansharav, K., Munson, S. M., Wilson, S. D. and Tietjen, B.: Future soil moisture and temperature extremes imply expanding suitability for rainfed agriculture in temperate drylands, Sci. Rep., 7(1), 12923, doi:10.1038/s41598-017-13165-x, 2017.

Burke, M. and Emerick, K.: Adaptation to climate change: Evidence from US agriculture, Am. Econ. J. Econ. Policy, 8(3), 106–40, 2016.

Comas, L. H., Trout, T. J., DeJonge, K. C., Zhang, H. and Gleason, S. M.: Water productivity under strategic growth stage-based deficit irrigation in maize, Agric. Water Manag., 212, 433–440, doi:10.1016/j.agwat.2018.07.015, 2019.

Denmead, O. T. and Shaw, R. H.: The Effects of Soil Moisture Stress at Different Stages of Growth on the Development and Yield of Corn 1, Agron. J., 52(5), 272–274, doi:10.2134/agronj1960.00021962005200050010x, 1960.

Diffenbaugh, N. S., Hertel, T. W., Scherer, M. and Verma, M.: Response of corn markets to climate volatility under alternative energy futures, Nat. Clim. Change, 2(7), 514–518, doi:10.1038/nclimate1491, 2012.

D'Odorico, P. and Porporato, A.: Preferential states in soil moisture and climate dynamics, Proc. Natl. Acad. Sci., 101(24), 8848–8851, doi:10.1073/pnas.0401428101, 2004.

Federer, C. A., Vörösmarty, C. and Fekete, B.: Intercomparison of methods for calculating potential evaporation in regional and global water balance models, Water Resour. Res., 32(7), 2315–2321, 1996.

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H. T., Verelst, L. and Wiberg, D.: Global agro-ecological zones assessment for agriculture (GAEZ 2008), IIASA Laxenburg Austria FAO Rome Italy, 10, 2008.

Fishman, R.: More uneven distributions overturn benefits of higher precipitation for crop yields, Environ. Res. Lett., 11(2), 024004, 2016.

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

Hainmueller, J., Mummolo, J. and Xu, Y.: How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice, Polit. Anal., 27(2), 163–192, 2019.

Hamon, W. R.: Computation of direct runoff amounts from storm rainfall, Int. Assoc. Sci. Hydrol. Publ., 63, 52–62, 1963.

Hauser, M., Thiery, W. and Seneviratne, S. I.: Potential of global land water recycling to mitigate local temperature extremes, Earth Syst. Dyn. Discuss., 1–18, doi:https://doi.org/10.5194/esd-2018-48, 2018.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J. and Piontek, F.: A trend-preserving bias correction–the ISI-MIP approach, 2013.

Hsiang, S.: Climate econometrics, Annu. Rev. Resour. Econ., 8, 43–75, 2016.

Kaur, G., Nelson, K. A. and Motavalli, P. P.: Early-Season Soil Waterlogging and N Fertilizer Sources Impacts on Corn N Uptake and Apparent N Recovery Efficiency, Agronomy, 8(7), 102, doi:10.3390/agronomy8070102, 2018.

Li, G., Zhao, B., Dong, S., Zhang, J., Liu, P. and Vyn, T. J.: Impact of controlled release urea on maize yield and nitrogen use efficiency under different water conditions, PLOS ONE, 12(7), e0181774, doi:10.1371/journal.pone.0181774, 2017.

Lobell, D. B. and Burke, M. B.: On the use of statistical models to predict crop yield responses to climate change, Agric. For. Meteorol., 150(11), 1443–1452, doi:10.1016/j.agrformet.2010.07.008, 2010.

Lobell, D. B., Hammer, G. L., McLean, G., Messina, C., Roberts, M. J. and Schlenker, W.: The critical role of extreme heat for maize production in the United States, Nat. Clim. Change, 3(5), 497–501, doi:10.1038/nclimate1832, 2013.

Manning, C., Widmann, M., Bevacqua, E., Loon, A. F. V., Maraun, D. and Vrac, M.: Increased probability of compound long-duration dry and hot events in Europe during summer (1950–2013), Environ. Res. Lett., 14(9), 094006, doi:10.1088/1748-9326/ab23bf, 2019.

McDonald, R. I. and Girvetz, E. H.: Two Challenges for U.S. Irrigation Due to Climate Change: Increasing Irrigated Area in Wet States and Increasing Irrigation Rates in Dry States, PLOS ONE, 8(6), e65589, doi:10.1371/journal.pone.0065589, 2013.

Meng, Q., Chen, X., Lobell, D. B., Cui, Z., Zhang, Y., Yang, H. and Zhang, F.: Growing sensitivity of maize to water scarcity under climate change, Sci. Rep., 6, 19605, doi:10.1038/srep19605, 2016.

Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., Sillmann, J., Schaller, N., Fischer, E., Schulz, M. and Stohl, A.: Frequency of extreme precipitation increases extensively with event rareness under global warming, Sci. Rep., 9(1), 1–10, doi:10.1038/s41598-019-52277-4, 2019.

Ortiz-Bobea, A., Wang, H., Carrillo, C. M. and Ault, T. R.: Unpacking the climatic drivers of US agricultural yields, Environ. Res. Lett., 14(6), 064003, 2019.

Peichl, M., Thober, S., Meyer, V. and Samaniego, L.: The effect of soil moisture anomalies on maize yield in Germany, Nat. Hazards Earth Syst. Sci., 18(3), 889–906, doi:https://doi.org/10.5194/nhess-18-889-2018, 2018.

Poschlod, B., Zscheischler, J., Sillmann, J., Wood, R. R. and Ludwig, R.: Climate change effects on hydrometeorological compound events over southern Norway, Weather Clim. Extrem., 28, 100253, doi:10.1016/j.wace.2020.100253, 2020.

Potopová, V., Trnka, M., Hamouz, P., Soukup, J. and Castraveț, T.: Statistical modelling of drought-related yield losses using soil moisture-vegetation remote sensing and multiscalar indices in the south-eastern Europe, Agric. Water Manag., 236, 106168, doi:10.1016/j.agwat.2020.106168, 2020.

Ribeiro, A. F. S., Russo, A., Gouveia, C. M., Páscoa, P. and Zscheischler, J.: Risk of crop failure due to compound dry and hot extremes estimated with nested copulas, Biogeosciences Discuss., 1–21, doi:https://doi.org/10.5194/bg-2020-116, 2020.

Rizzo, G., Edreira, J. I. R., Archontoulis, S. V., Yang, H. S. and Grassini, P.: Do shallow water tables contribute to high and stable maize yields in the US Corn Belt?, Glob. Food Secur., 18, 27–34, doi:10.1016/j.gfs.2018.07.002, 2018.

Roberts, M. J., Schlenker, W. and Eyer, J.: Agronomic Weather Measures in Econometric Models of Crop Yield with Implications for Climate Change, Am. J. Agric. Econ., 95(2), 236–243, doi:10.1093/ajae/aas047, 2013.

Sarhadi, A., Ausín, M. C., Wiper, M. P., Touma, D. and Diffenbaugh, N. S.: Multidimensional risk in a nonstationary climate: Joint probability of increasingly severe warm and dry conditions, Sci. Adv., 4(11), eaau3487, doi:10.1126/sciadv.aau3487, 2018.

Schaffer, B. E., Nordbotten, J. M. and Rodriguez-Iturbe, I.: Plant biomass and soil moisture dynamics: analytical results, Proc R Soc A, 471(2183), 20150179, doi:10.1098/rspa.2015.0179, 2015.

Schlenker, W. and Roberts, M. J.: Nonlinear temperature effects indicate severe damages to US crop yields under climate change, Proc. Natl. Acad. Sci., 106(37), 15594–15598, 2009.

Schmidt, J. P., Sripada, R. P., Beegle, D. B., Rotz, C. A. and Hong, N.: Within-Field Variability in Optimum Nitrogen Rate for Corn Linked to Soil Moisture Availability, Soil Sci. Soc. Am. J., 75(1), 306–316, doi:10.2136/sssaj2010.0184, 2011.

Siebert, S. and Döll, P.: Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation, J. Hydrol., 384(3–4), 198–217, doi:10.1016/j.jhydrol.2009.07.031, 2010.

Siebert, S., Webber, H., Zhao, G. and Ewert, F.: Heat stress is overestimated in climate impact studies for irrigated agriculture, Environ. Res. Lett., 12(5), 054023, doi:10.1088/1748-9326/aa702f, 2017.

Torbert, H. A., Hoeft, R. G., Vanden Heuvel, R. M., Mulvaney, R. L. and Hollinger, S. E.: Short-term excess water impact on corn yield and nitrogen recovery, J. Prod. Agric., 6(3), 337–344, 1993.

Urban, D., Roberts, M. J., Schlenker, W. and Lobell, D. B.: Projected temperature changes indicate significant increase in interannual variability of U.S. maize yields, Clim. Change, 112(2), 525–533, doi:10.1007/s10584-012-0428-2, 2012.

Urban, D. W., Roberts, M. J., Schlenker, W. and Lobell, D. B.: The effects of extremely wet planting conditions on maize and soybean yields, Clim. Change, 130(2), 247–260, doi:10.1007/s10584-015-1362-x, 2015.

USDA-NASS: USDA-National Agricultural Statistics Service, Cropland Data Layer, U. S. Dep. Agric. Natl. Agric. Stat. Serv. Mark. Inf. Serv. Off. Wash. DC [online] Available from: http//nassgeodata. gmu. edu/Crop-Scape, 2017.

Wang, R., Bowling, L. C., Cherkauer, K. A., Cibin, R., Her, Y. and Chaubey, I.: Biophysical and hydrological effects of future climate change including trends in CO2, in the St. Joseph River watershed, Eastern Corn Belt, Agric. Water Manag., 180, 280–296, doi:10.1016/j.agwat.2016.09.017, 2017.

Wehner, M.: Estimating the probability of multi-variate extreme weather events, in Workshop on Correlated Extremes, Columbia University., 2019.

Werner, A. T. and Cannon, A. J.: Hydrologic extremes – an intercomparison of multiple gridded statistical downscaling methods, Hydrol. Earth Syst. Sci., 20(4), 1483–1508, doi:https://doi.org/10.5194/hess-20-1483-2016, 2016.

Williams, A., Hunter, M. C., Kammerer, M., Kane, D. A., Jordan, N. R., Mortensen, D. A., Smith, R. G., Snapp, S. and Davis, A. S.: Soil Water Holding Capacity Mitigates Downside Risk and Volatility in US Rainfed Maize: Time to Invest in Soil Organic Matter?, PLOS ONE, 11(8), e0160974, doi:10.1371/journal.pone.0160974, 2016.

Wing, I. S., Monier, E., Stern, A. and Mundra, A.: US major crops' uncertain climate change risks and greenhouse gas mitigation benefits, Environ. Res. Lett., 10(11), 115002, 2015.

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), Hydrol Earth Syst Sci, 14(1), 1–24, doi:10.5194/hess-14-1-2010, 2010.

Wooldridge, J. M.: Introductory econometrics: A modern approach, Nelson Education., 2016.

Zscheischler, J., Westra, S., Van Den Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch, D. N., Leonard, M. and Wahl, T.: Future climate risk from compound events, Nat. Clim. Change, 8(6), 469–477, 2018.

Zscheischler, J., van den Hurk, B., Ward, P. J. and Westra, S.: Multivariate extremes and compound events, in Climate Extremes and Their Implications for Impact and Risk Assessment, pp. 59–76, Elsevier., 2020.

Table 1. Yield, heat, and water metrics for 1981-2015 (Apr-Sep)

Variable	Mean	Std.Dev	Min	Max
Corn yield (bushels / acre)	109.8	37.8	4.5	246.0
Cumulative precipitation (mm)	564	183	1	1469
Mean daily soil moisture content (mm)	47	39	0.1	262
in Apr-May (mm)	60	48	0	270
in Jun-Jul (mm)	48	45	0	270
in Aug-Sep (mm)	30	30	0	264
Degree days from 10˚C to 29˚C	1848	434	693	3083
when soil moisture is low	397	430	0	2629
when soil moisture is normal	1112	572	0	3044
when soil moisture is high	330	346	0	2665
Degree days above 29˚C	61	61	0	723
when soil moisture is low	18	31	0	400
when soil moisture is normal	37	48	0	680
when soil moisture is high	5	9	0	140
Index of soil moisture above normal levels (mm)	2370	2135	0	20319
Index of soil moisture below normal levels (mm)	-2384	2147	-23978	0
Number of days with moisture deficit > 25 mm	27	30	0	182
Number of days with moisture surplus > 25 mm	27	35	0	183
Mean daily soil moisture fraction	0.71	0.18	0.01	1.00
Mean daily evapotranspiration (mm)	0.55	0.58	0.00	2.95
Number of observations	69923			

Notes: Table reports descriptive statistics for major variables in this study. The mean and standard deviations are calculated over US counties for the 1981-2015 period. All the weather data are calculated for each 2.5 x 2.5 arcmin grids, averaged over the time interval, and then averaged to counties using cropland area weights. Soil moisture seasonal normal is defined as the average of 1981-2015 daily soil moisture level from the first day of April to the last day of September.

Table 2. Corn yield estimation without the interaction of heat and soil moisture in ~~(~~Model 1 (a-d).

	(1-a) Log CornYield	(1-b) Log CornYield	(1-c) Log CornYield	(1-d) Log CornYield
Degree Days 10-29°C Apr-Sep	.000336*** (.000087)	.000343*** (.00008)	.0003486*** (.0000725)	.0003083*** (.0000683)
Degree Days above 29°C Apr-Sep	-.005307*** (.000673)	-.005114*** (.000691)	-.005277*** (.0006678)	-.005041*** (.0005999)
Precipitation Apr-Sep	.000658** (.000254)			
Precipitation Apr-Sep Square	-5.16e-07** (-9.35e-07)			
Seasonal Mean Soil Moisture Content		.003593*** (.000664)		
Seasonal Mean Soil Moisture Content Square		-.000017*** (3.000e-06)		
Number of days with SM 25+ mm above normal			-.001838*** (.0003816)	
Number of days with SM 25+ mm below normal			-.002089*** (.0002817)	
Index of Soil Moisture above Normal (mm)				-.000040*** (2.800e-06)
Index of Soil Moisture below Normal (mm)				.000044*** (7.100e-06)
Obs.	69923	69923	69923	69923
R-squared	0.4686	0.4714	0.4795	0.4914
AIC (Akaike's information criterion)	-21238.1	-21612.3	-22696.8	-24303.4
BIC (Bayesian information criterion)	-21201.4	-21575.7	-22660.2	-24266.8

Standard errors are in parenthesis & adjusted for state clusters
*** p<0.01, ** p<0.05, * p<0.1

Notes: Table lists regression coefficients and shows standard errors in brackets. ~~Temperature is in degree Celsius, precipitation in mm, soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output and interpolated to 2.5 arcmin, while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA.~~ The constant term and coefficients on the interaction of each state and time trends are not reported.

Table 3. Corn yield estimation while splitting heat stress index in~~,~~ Model 2a

	(2-a) Log CornYield
Degree days from 10˚C to 29˚C	.0003083*** (.0000685)
dday29˚C & SM 75+ mm below normal (extreme deficit)	-.0082398*** (.0014372)
dday29˚C & SM 25-75 mm below normal (deficit)	-.0062069*** (.0009793)
dday29˚C & SM 0-25 mm around normal (normal)	-.0037559*** (.0004045)
dday29˚C & SM 25-75 mm above normal (surplus)	-.0055709*** (.0012041)
dday29˚C & SM 75+ mm above normal (extreme surplus)	-.0140295*** (.0019083)
Mean daily soil moisture content (mm)	.0026635*** (.0008153)
Square of mean daily soil moisture content	-.0000161*** (2.600e-06)
Observations	69923
R-squared	.4921
Akaike's Crit	-24401.6
Bayesian Crit	-24328.3

Standard errors are in parenthesis & adjusted for state clusters
*** p<0.01, ** p<0.05, * p<0.1

Notes: Table lists regression coefficients and shows standard errors in brackets. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA.~~ The constant term and coefficients on the interaction of each state and time trends are not reported.

Table 4. Estimation of corn yields while splitting the soil moisture metrics in Model 2b ~~indicators~~

	(2-b) log CornYield
Degree days from 10˚C to 29˚C	.0003154***
	(.0000689)
Degree days above 29˚C	-.004044***
	(.0005384)
Index of normal soil moisture when T > T*	.0001199***
	(.0000342)
Index of extreme moisture surplus when T > T*	-.0000628***
	(.0000151)
Index of extreme moisture deficit when T > T*	.000092***
	(.0000234)
Index of extreme moisture deficit when T < T*	.0000209***
	(7.100e-06)
Index of extreme moisture surplus when T < T*	-.0000326***
	(3.200e-06)
Index of normal soil moisture when T < T*	.000028**
	(.0000105)
Observations	69923
R-squared	.5006
Akaike's Crit	-25582.4
Bayesian Crit	-25509.2

Standard errors are in parenthesis & adjusted for state clusters
*** p<0.01, ** p<0.05, * p<0.1

Notes: Table lists regression coefficients and shows standard errors in brackets. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA.~~ The constant term and coefficients on the interaction of each state and time trends are not reported.

Table 5: Performance metrics for Models 1(a-d) and 2(a-b).

Model	Water metric	Extreme metric	R-squared	AIC (Akaike's information criterion)	BIC (Bayesian information criterion)
1-a	Avg~~e~~. precipitation	Precipitation sqr	0.469	-21,238	-21,201
1-b	Avg. soil moisture	Soil moisture sqr	0.471	-21,612	-21,576
1-c	Avg. soil moisture	Number of days with low/high soil moisture	0.480	-22,697	-22,660
1-d	Avg. soil moisture	Avg soil moisture deficit/surplus	0.491	-24,303	-24,267
2-a	Avg. soil moisture	T binned by extreme deficit/surplus	0.492	-24,402	-24,328
2-b	normal soil moisture x T	extreme deficit/surplus x T	0.501	-25,582	-25,509
~~2-c~~	~~index of normal soil moisture~~	~~Avg soil moisture deficit/surplus, T x share of heat at normal moisture~~	~~0.503~~	~~-25,900~~	~~-25,836~~
~~2-d~~	~~index of normal soil moisture~~	~~Avg soil moisture deficit/surplus, T x share of irrigated area~~	~~0.510~~	~~-26,840~~	~~-26,776~~

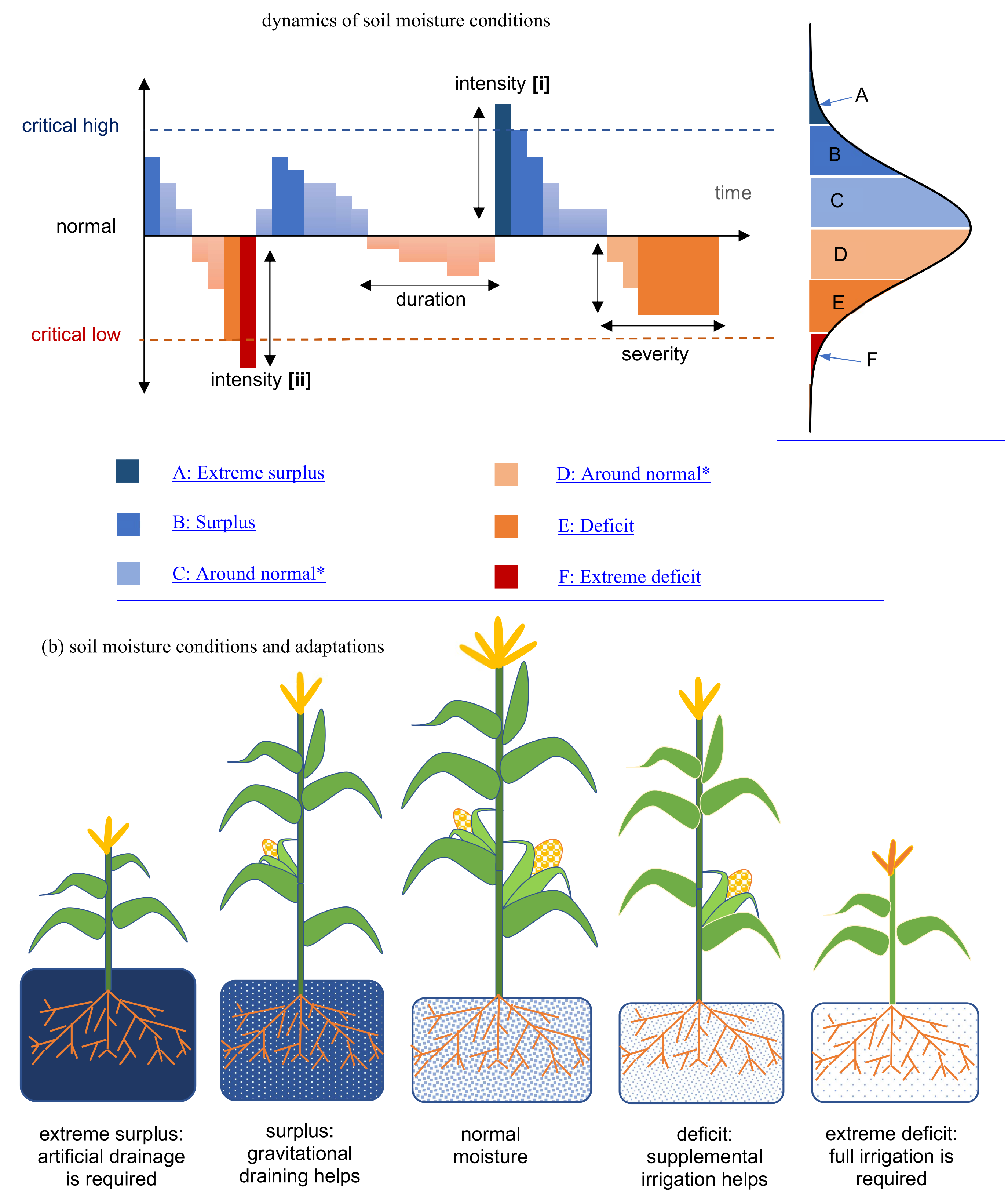
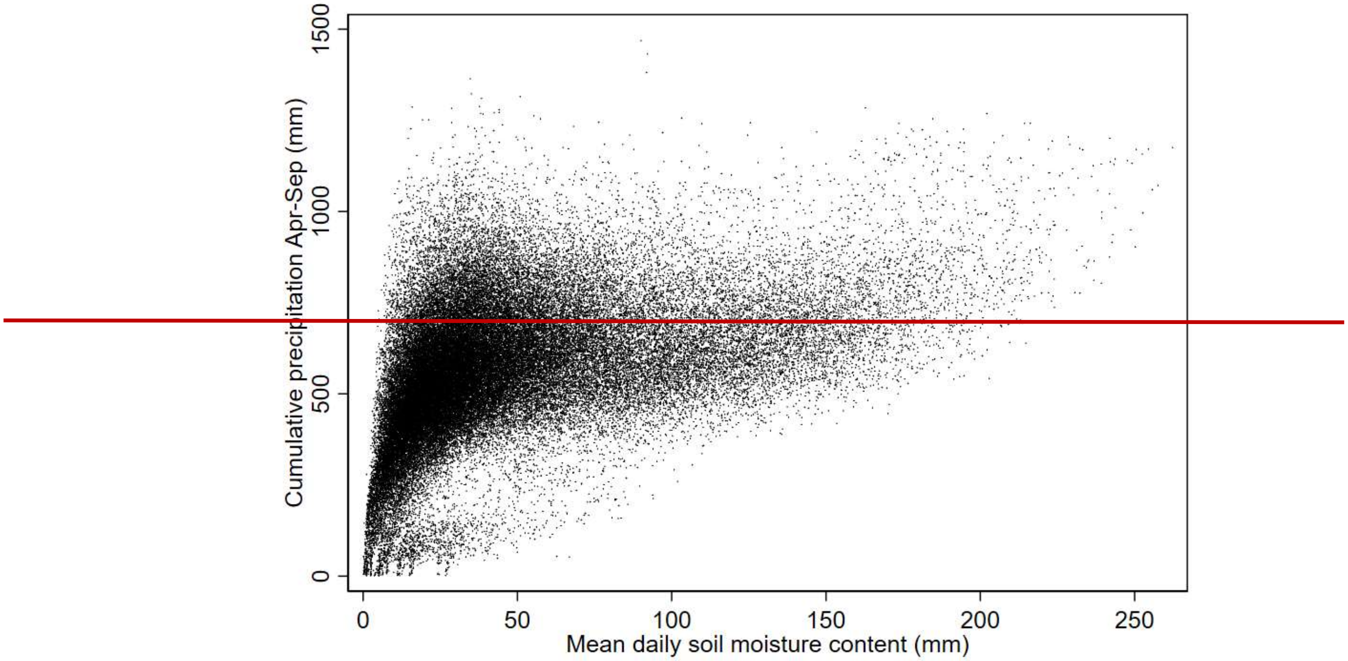


Figure 1. ~~(a)~~ **Soil moisture dynamics within a typical growing season. Some soil moisture conditions can be harmful to crops including excess wetness [i], moisture stress intensity[ii], duration of moisture stress [iii], and severity of soil moisture stress [iv].** ~~**(b) Adaptation mechanisms can reduce the damage to crops. As flood can cause severe damage to corn, artifitial drainage is required; excess water may slow down the growth; normal soil moisture makes optimum growth; water deficit can limit the growth, while supplemental irrigation can help; during an extreme water deficit, irrigation is necessary.**~~

~~**Figure 2. Soil texture affects the wilting point, field water-holding capacity, and the moisture available to plants. This suggests that sandy soil has the lowest wilting point as well as low field capacity. As most of the water infiltrates, this leaves a little amount of moisture available to plants (Tsoar, 2005).**~~



~~Figure 3. WBM mean soil moisture versus PRISM cumulative precipitation for 1981-2015 by US counties.~~

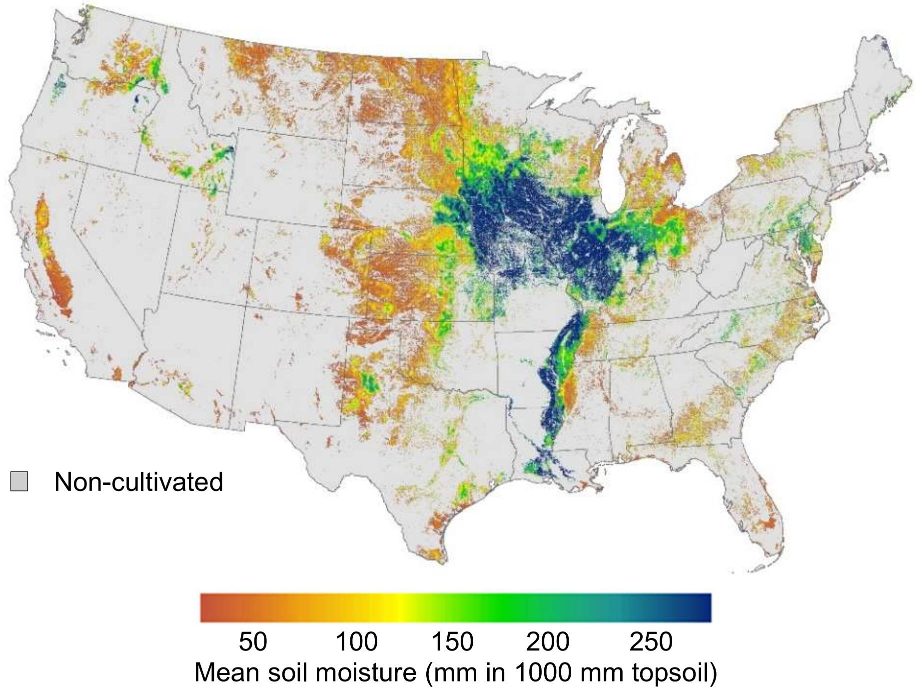


Figure 24. Growing season mean soil moisture content (in mm in 1000 mm topsoil) as calculated based on daily root-zone soil moisture level from Apr-Sep for 1981-2015 at 2.5 x 2.5 arcmin grids excluding non-cultivated area. The soil moisture level is obtained from the Water Balance Model (WBM) and non-cultivated area information is from USDA National Cultivated Layer. This map illustrates the heterogeneity of simulated soil moisture over the Continental US and even within states.

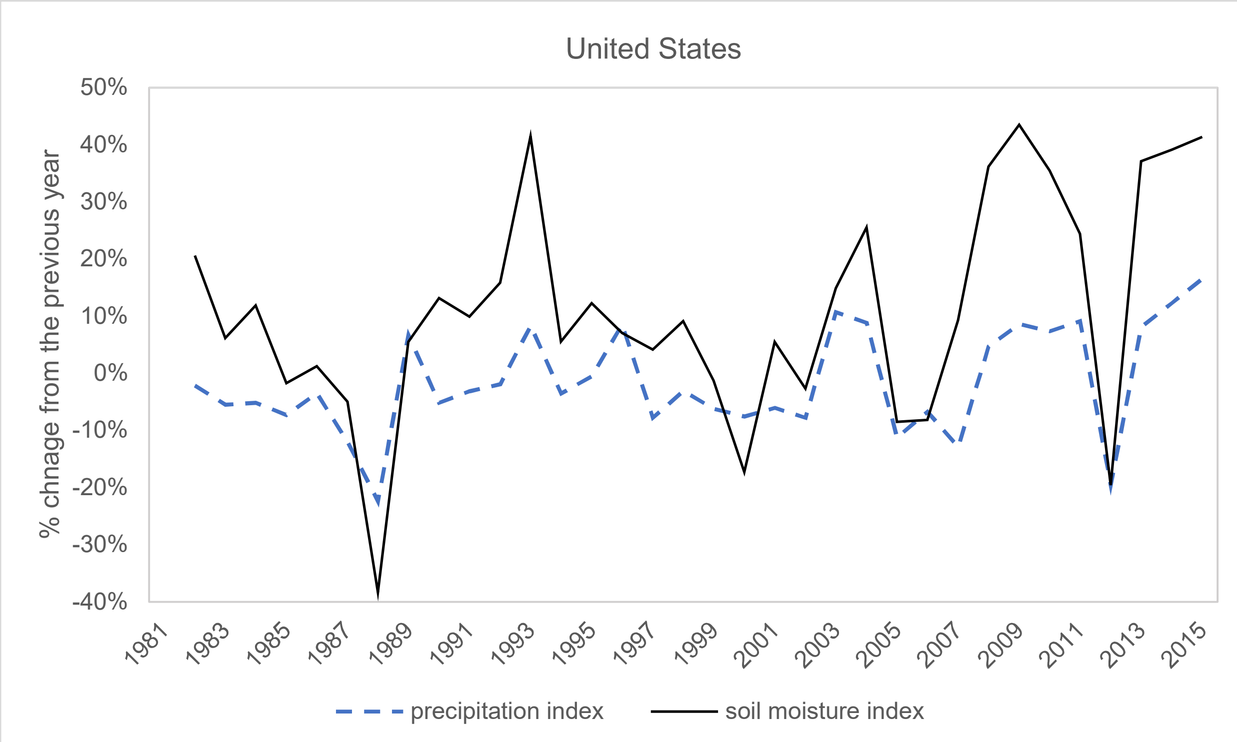


Figure ~~5~~3. Variations of average precipitation versus average soil moisture over corn areas in the United States. The precipitation is aggregated from PRISM and soil moisture is aggregated from WBM from 2.5 arcmin grid cells weighted by cropland area.

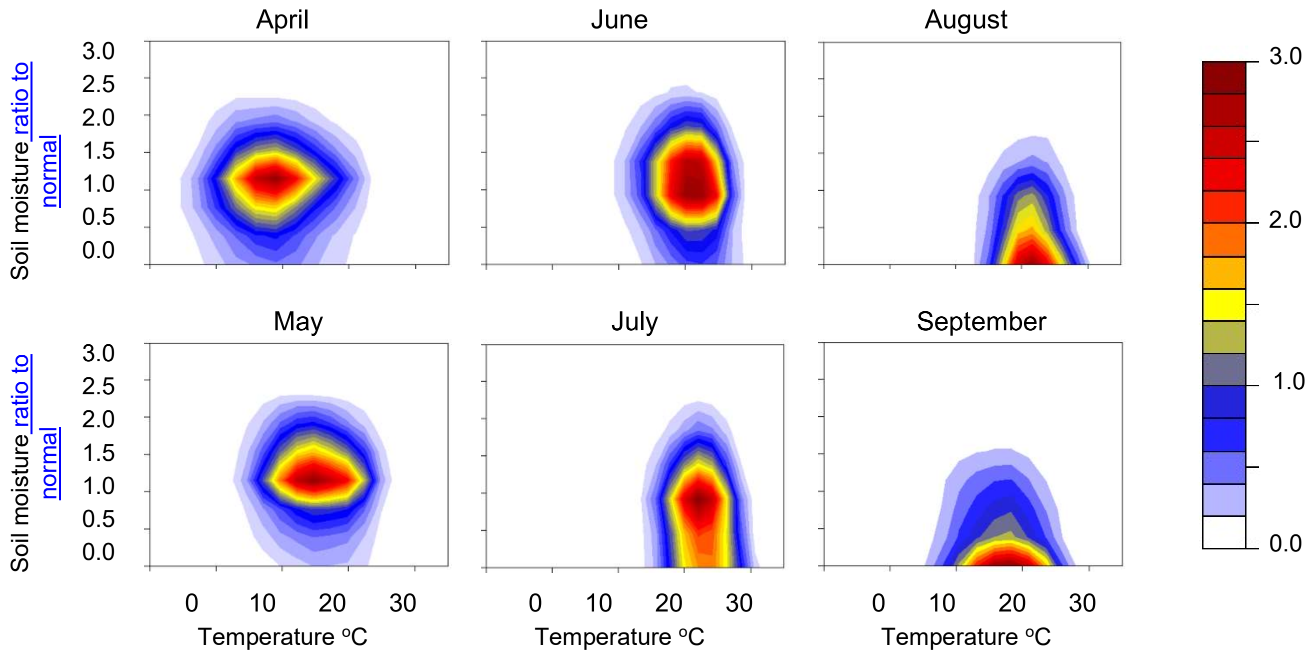


Figure 4~~6~~. The bivariate density of heat and soil moisture for 1981-2015 For all the grid cells in the US Corn Belt. The precipitation is aggregated from PRISM and soil moisture is aggregated from WBM based on 2.5 arcmin grid cells.

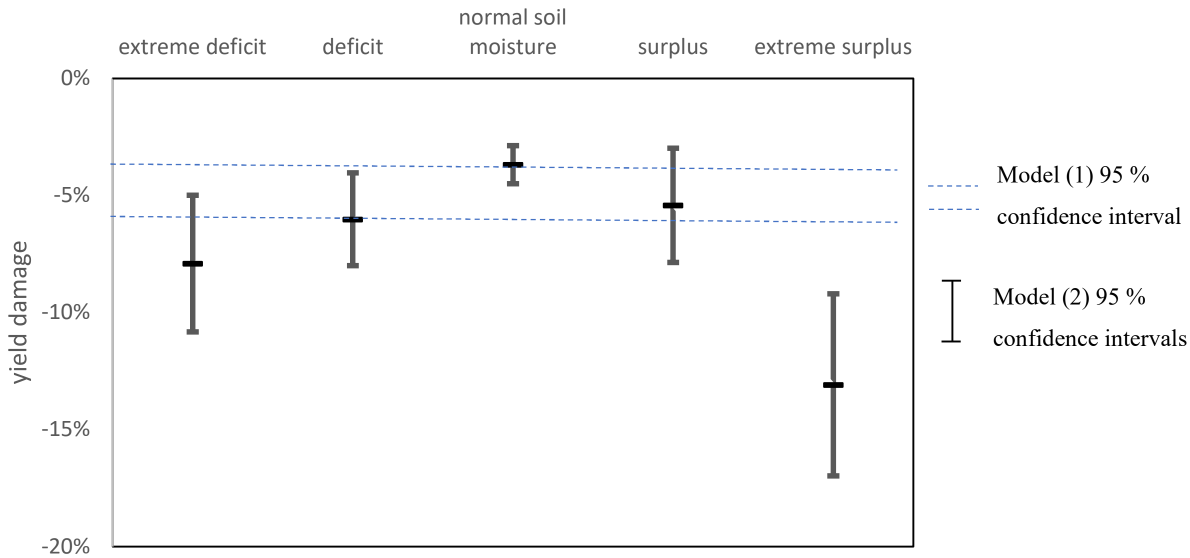


Figure 5~~7~~. Estimated damage to corn yield from an additional 10 degree-days above 29˚C and no change in seasonal mean soil moisture.

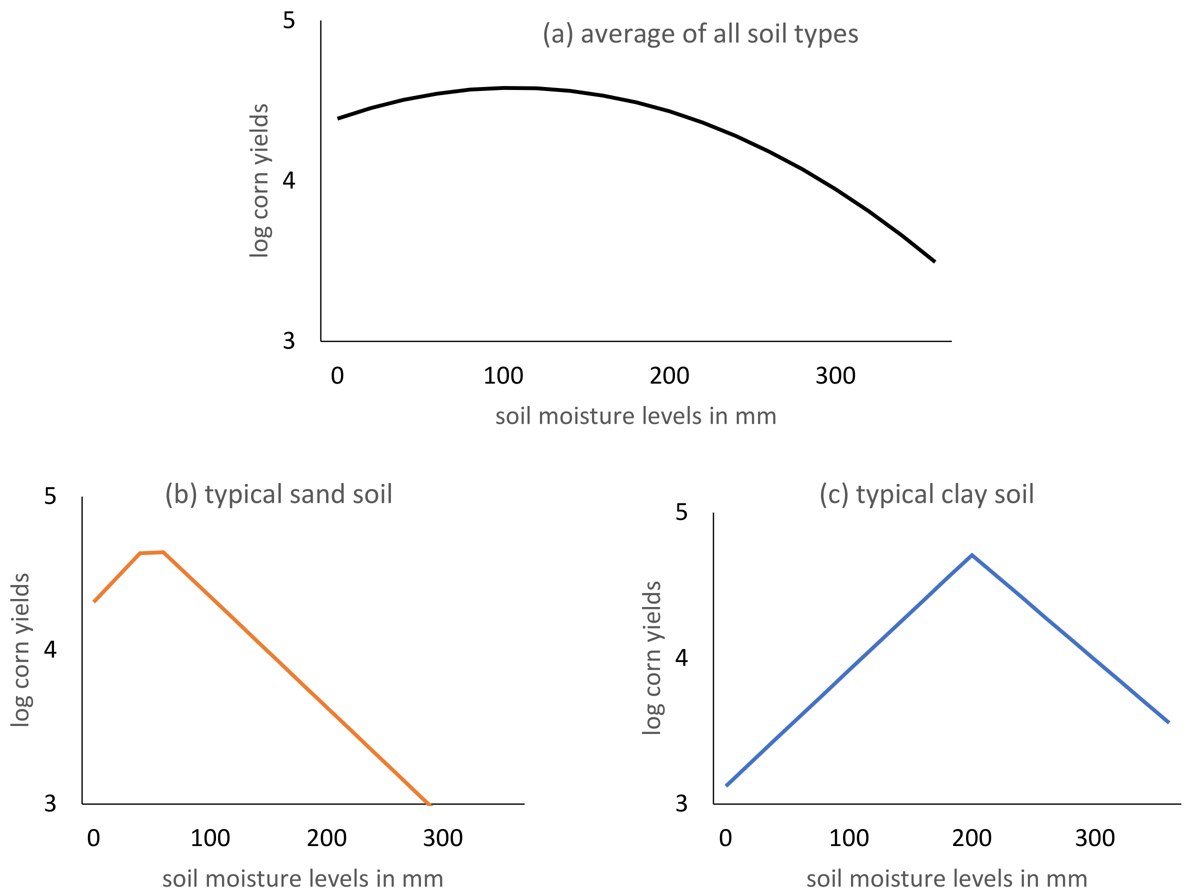
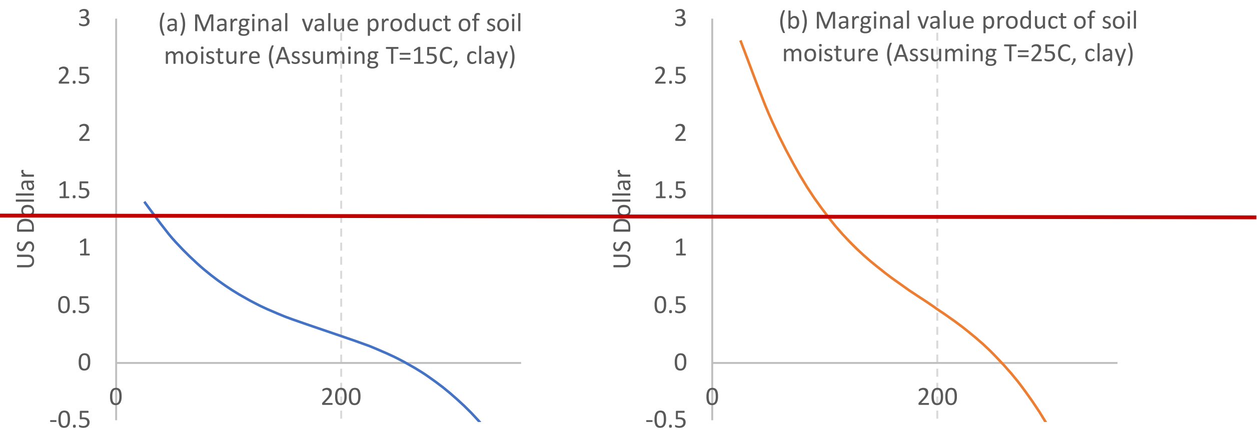


Figure 6~~8~~. Estimated impact of soil moisture on log corn yields. Including soil moisture in the regression and its square term, as in model 1-b, will give us a quadratic relationship between soil moisture and yields as in panel (a). A piece-wise linear parametrization, as in model 1-d, can provide location-specific piece-wise linear relationship based on soil moisture deviation from normal as in panels (b) and (c). This will cause the maximum of the response curve to be in lower soil moisture levels for sand and in higher soil moisture levels for clay soil texture.



~~**Figure 9. The marginal value product of soil moisture in clay soil with normal soil moisture of 200 mm for hot days (average temperature = 25°C) vs moderate days (average temperature = 15°C).**~~

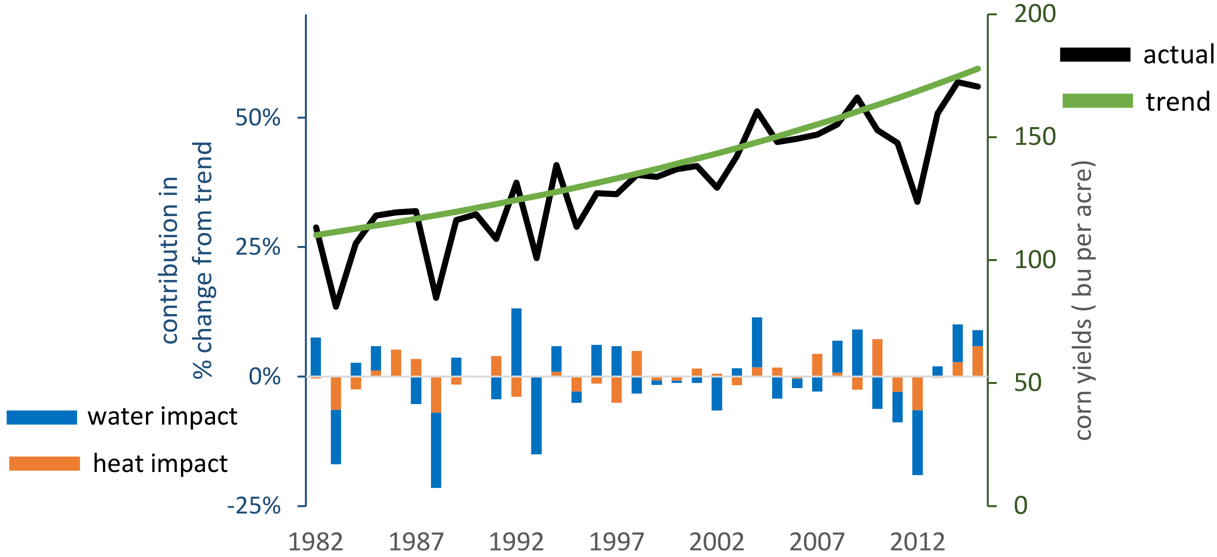
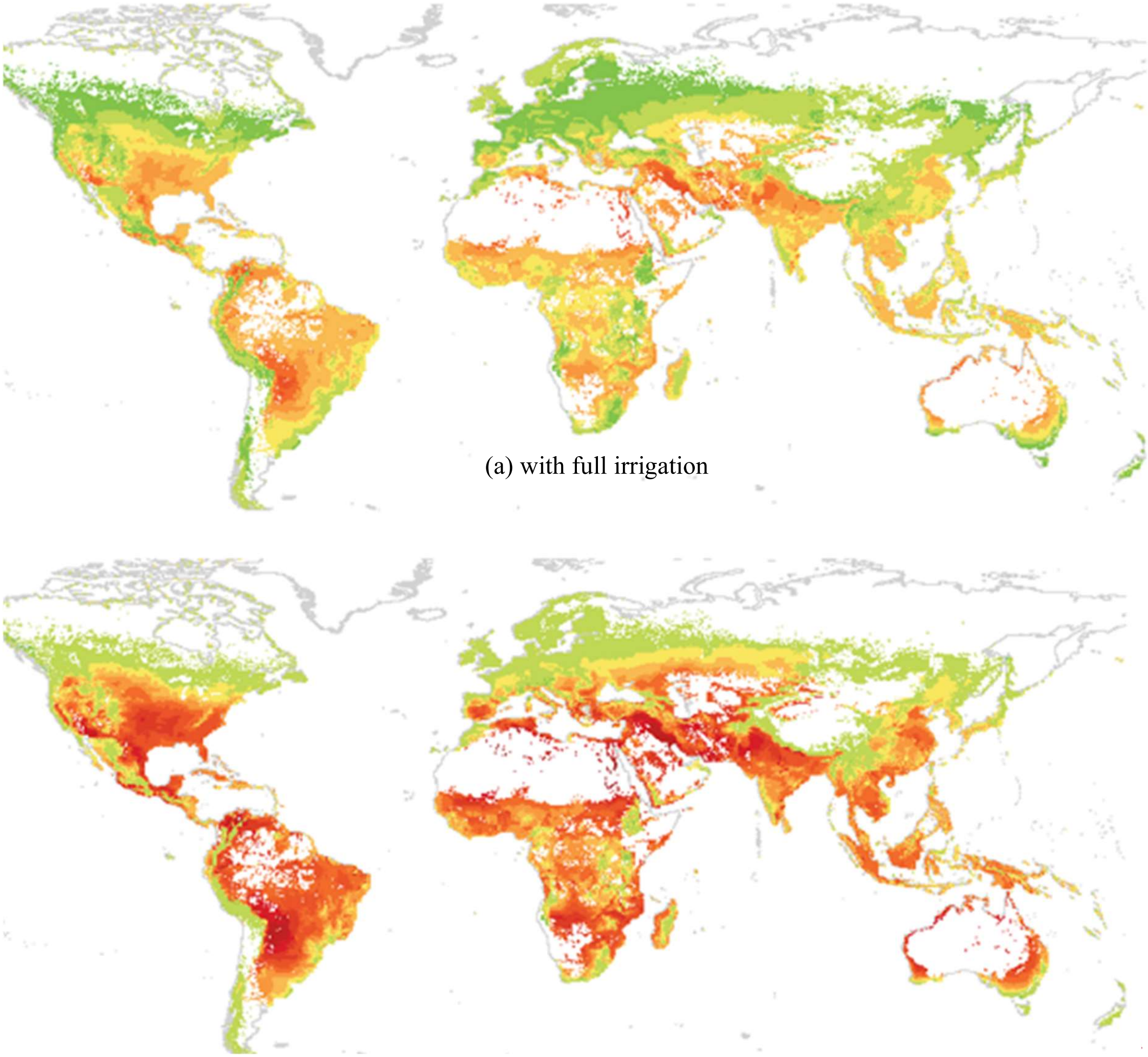
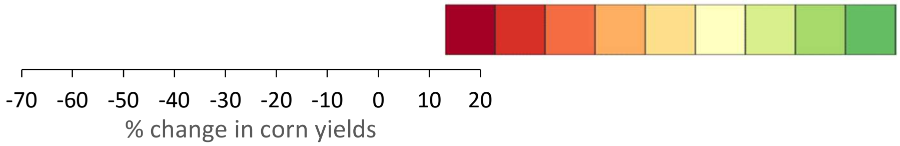


Figure 7~~10~~. The bars show the “contribution of water” and “contribution of heat” in variation of US corn yields (left axis). The lines illustrate actual yields and trend (right axis). ~~**The impact of other factors is not reported.**~~



(a) with full irrigation



~~Figure 11. The impacts of climate change on corn yields with irrigation adaptation illustrated in panel (a) and without irrigation adaptation is shown in panel (b). The maps show the percentage change in corn yield due to climate change from 1976-2005 to 2036-2065 based on CMIP5 RCP 8.5 from NEX-GDDP climate product and estimated parameters in this study.~~

Supplementary Material

S.1. Overview

This material ~~appendix~~ provides some robustness checks on the results and the model variables. First, we illustrate the correlation ~~relationship~~ between mean volumetric soil moisture and other potential seasonal variables that can be used as indicators of water availability ~~in this study~~. This includes cumulative precipitation, mean seasonal evapotranspiration,~~,~~ and mean seasonal soil moisture ~~fraction, and the degree days above 10˚C~~. We provide some examples to demonstrate the seasonal mean soil moisture shows no linear relationship with the seasonal heat index (degree days above 10˚C). However, it has a positive correlation with evapotranspiration and soil moisture fraction. Then we provide alternative models controlling for irrigation, growth periods, spatial scope of the study, and other measures of individual and compound extremes.

S.2.~~A.1.~~ Correlation of mean seasonal soil moisture and other variables

The soil moisture output from WBM is informed mainly by soil moisture memory, heat, precipitation, and many other time-variant and time-invariant information. In a statistical study, a natural first step is to look at the correlation between these variables. To show that mean soil moisture is a different metric than mean precipitation, we have plotted the annual mean soil moisture versus annual cumulative precipitation in Fig. S2. This figure is a scatter plot for US counties for the growing season from 1981 to 2015. The simple correlation coefficient between them is 0.44. This rejects the hypothesis that soil moisture is highly correlated with precipitation. As mean precipitation has a linear relationship with cumulative precipitation, it shows that mean soil moisture is a different metric than cumulative or mean precipitation.

We have taken two other variables from WBM including soil moisture fraction and evapotranspiration (ET). Also, we have interpolated WBM soil moisture using an alternative method (nearest neighbor method). ~~Section A.6 will provide the estimation results when using these variables to show the robustness of the results to variable selection.~~ Here, we plot these variables against the volumetric soil moisture content to illustrate the correlation and differences. As shown in Fig. S3~~A1~~ two interpolations of soil moisture are closely correlated by R= 0.9997. Figures S4 ~~A2~~ and S5~~A3~~ are the scatter plots of seasonal ET and seasonal mean soil moisture fraction against volumetric soil moisture. The figures show the seasonal variables are not following a simple linear relationship. Figure S6 ~~A4~~ shows the scatter plot of cumulative growing degree days above 10˚C versus mean soil moisture for US counties for the growing season from 1981 to 2015. This indicates the soil moisture output is not a simple linear transformation of heat data.

S.3. Are the results different with alternative water metrics? ~~A.6. Robustness check: other metrics from WBM outputs (soil moisture fraction and ET)~~

~~Here, w~~We re-estimate Model (1) with other related metrics of water availability to crops including simulated daily evapotranspiration of rainfed corn (ET) from WBM; daily soil moisture fraction (SMF) from WBM; and soil moisture content from different spatial interpolation of WBM grid cells to PRISM (nearest neighbor method versus original bilinear method).

The soil moisture fraction index considers the volumetric soil moisture content divided by field capacity. We have also considered the within-season standard deviation of ET and SMF. Note that we keep the degree days above 29˚C as an indicator of heat stress and the degree days from 10˚C to 29˚C as an indicator of beneficial heat to corn.

Table S1~~A7~~. reports regression results for these models. Columns 1 and 2 show a significant relationship with the mean of soil moisture fraction, its square term, and its within season standard deviation. Columns 3 and 4 with mean ET and within-season SD of ET also show a significant relationship. Column 5 shows that the other interpolation of soil moisture has a very close marginal coefficient and standard error compared to our original Model (1). The important finding is the marginal relationship for beneficial and harmful heat remains significant and not significantly different from Model (1). Overall, the main findings of the paper remain robust to the choice of alternative seasonal metrics of water availability.

S.4. Are the estimates different when considering the stages of plant growth? ~~A.5. Robustness check: bi-monthly metrics of soil moisture~~

How critical is separating the stages of plant growth in the yield function? We re-estimate the Model (1-b) considering bi-monthly metrics of seasonal soil moisture. Table S3 ~~A6~~. provides the estimation coefficients, standard errors, AIC, BIC, and R-squares statistics for Model (1-b) for Eastern, Western, and the continental US with bi-monthly mean soil moisture. The results suggest that the coefficients on extreme heat (DD29˚C) are not significantly different from the model with seasonal mean soil moisture.
The results suggest that the marginal impact of mean soil moisture is higher in June-July. This is in line with agronomic literature as it suggests the water stress during pollination and the silking stage is more damaging. These stages are the most critical stage of development for corn. Water stress during this stage can cause higher yield loss than almost any other stage in the crop's development.
The marginal impact of mean soil moisture is not significant in August-September. This suggests that additional soil moisture can have a positive or negative impact on yield. This also makes sense as a high level of moisture can hurt the maturity and drying stage. High soil moisture at the end of the growing season can cause delayed grain maturity and may lead to delay in the harvest.
In Addition, the marginal impact of mean soil moisture in April-May is negative for the whole US and the Western US and significant at 90% confidence interval. This can be a result of the negative impacts of excess soil moisture on germination and early crop developments as a result of flooding and waterlogging.

S.5. Do the main findings change if we alter the geographical scope of the study ~~A.4. Robustness check: West versus East in Model (1)~~

In this section, we estimate the main models separately for Eastern and Western US. Those counties with centroids on the left of 100th meridian are considered West. The idea is that water stress is less severe in the Western US as it is mostly irrigated. Table S2~~A3~~. provides the main descriptive statistics to compare these regions. Overall, Western US experiences more excess heat by 82 versus 58 DD29˚C in the East. On average, Eastern US receives 601 mm of cumulative precipitation while it is only 271 mm in the Western US. On the other hand, within-season SD of soil moisture is 39 mm in the East while it is 13 mm in the west. Looking at the number of days with high/low soil moisture, only 11 days in the West soil moisture is not at normal levels, while this is 59 days in the East.
Table ~~A~~S4. shows the estimated coefficients, standard errors, adjusted R-squared, AIC, and BIC statistics for four models for Eastern US. Model (1-a) includes cumulative precipitation. Model (1-b~~a-2~~) includes mean soil moisture metrics. The third column ~~model~~, similar to Model (1-d~~3b~~), considers soil moisture extremes. ~~And Model (4) considers the interaction terms.~~ The results suggest that the coefficient on the extreme heat is not significantly different from the estimations for the whole US.
Table ~~A~~S5. shows the estimated coefficients, standard errors, adjusted R-squared, AIC, and BIC statistics for four models for the Western US. The results suggest that the coefficients on the extreme heat are significantly different from the estimations for the

whole US and the Eastern US. For example, the coefficient on DD29˚C is -0.0020 in Model (1-a) for the West, while it was estimated -0.0056 in Model (1-a) for the East. This is around 65% lower damage for a given degree day above 29˚C. ~~Also, the AIC and BIC statistics would reject the hypothesis that models with interaction perform better compared to the model with cumulative precipitation. The difference can be a result of the "*cooling effect*" as discussed in Sect. S.4.~~

~~A.7. Robustness check: East and West in Model (2)~~

We also~~Here~~ re-estimate Model (2) for Eastern and Western US. ~~The results are presented in Table A8 and A.9 for the US, West, and East. We see a similar pattern for East versus West. The coefficient on heat stress is smaller for the West which can be a result of the cooling effect.~~

The results of Model (2-a) are presented in Table S6~~A8~~. Column 1 shows the results for the whole US while columns 2 and 3 contain the results for the Western US and Eastern US, respectively. According to column 2, the coefficient on dday29˚C and the extreme deficit is -0.0074 in the Western US which is significantly different from all other estimations for the Western US. ~~This is another evidence of the cooling effect.~~ These results indicate that, even in the Western US, *the damage from heat stress can be up to four times higher when combined with water stress*. The coefficient on excess heat and the extreme surplus is not significant (note that this is a very rare condition in the West).

The results of model (2-b) for Eastern, Western, and whole US are shown in Table S7. As in column (3) of Table S7~~A9~~, the coefficient on normal soil moisture conditional to hot weather is 0.00010. The coefficient on normal soil moisture conditional to moderate weather is 0.00002. This indicates that yield response to water ~~water~~ is up to four times more ~~valuable~~ in hot weather. The marginal impact on soil moisture deficit index is 0.00008 in hot weather and is 0.00002 in moderate weather. This also supports the finding that the yield response to water is up to four times more ~~valuable~~ in hot weather. Also, the results suggest that the damage from excess water is up to two times bigger in hot weather.

~~**Table A1. Corn yield estimation controlling for normal soil moisture**~~

	(2-c) Log CornYield
Degree days from 10˚C to 29˚C x S1 (share of heat at normal moisture)	.000111* (.0000584)
Degree days from 10˚C to 29˚C	.0002513*** (.0000827)
Degree days above 29˚C x S2 (share of heat at normal moisture)	.0017202** (.0008421)
Degree days above 29˚C	-.0057224*** (.0009421)
Index of normal soil moisture	.0000504*** (.0000108)
Index of extreme moisture surplus (sum of positive deviations if > +25 mm)	-.0000233*** (6.700e-06)
Index of extreme moisture deficit (sum of negative deviations if < -25 mm)	.0000187*** (5.300e-06)
Observations	69923
R-squared	.502883
AIC (Akaike's information criterion)	-25900.4
BIC (Bayesian information criterion)	-25836.4

~~Standard errors are in parenthesis & adjusted for state clusters~~
~~*** p<0.01, ** p<0.05, * p<0.1~~

~~**Notes: Table lists regression coefficients and shows standard errors in brackets. Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.**~~

~~**Table A2. Estimation of Model (2-d) controlling with irrigation**~~

	~~(2-d)~~ ~~log CornYield~~
~~Degree days from 10˚C to 29˚C, α~~	~~.0003027***~~ ~~(.0000676)~~
~~Degree days above 29˚C, β~~	~~-.0061523***~~ ~~(.0009143)~~
~~Degree days above 29˚C x S2 (share of heat at normal moisture), β'~~	~~.0017508***~~ ~~(.0006243)~~
~~Degree days above 29˚C x S3 (area share of irrigated corn), β''~~	~~.0023809***~~ ~~(.0007472)~~
~~Index of deficit (sum of negative deviations if < -25 mm), η_{lo}~~	~~.0000287***~~ ~~(5.200e-06)~~
~~Index of surplus (sum of positive deviations if > +25 mm), η_{hi}~~	~~-.0000345***~~ ~~(2.800e-06)~~
~~Index of normal soil moisture, η_{nl}~~	~~.000046***~~ ~~(.0000101)~~
~~Observations~~	~~69923~~
~~R-squared~~	~~.5095193~~
~~Akaike's Crit~~	~~-26840.2~~
~~Bayesian Crit~~	~~-26776.1~~

~~Standard errors are in parenthesis~~
~~*** p<0.01, ** p<0.05, * p<0.1~~

~~**Notes: Table lists regression coefficients and shows standard errors in brackets. Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.**~~

Table S1.~~A7.~~ Estimating corn yields using ET and SMF from WBM

	Log CornYield	Log CornYield	Log CornYield	Log CornYield	Log CornYield
Degree days from 10˚C to 29˚C	.0003422***	.0003445***	.0003193***	.0003372***	.0003426***
	(.0000752)	(.0000741)	(.0000801)	(.0000751)	(.0000801)
Degree days above 29˚C	-.005298***	-.005343***	-.005017***	-.004884***	-.005115***
	(.00069)	(.0006681)	(.00064)	(.0006367)	(.0006914)
Mean daily soil moisture fraction	.2533803**	.9821037***			
	(.1107891)	(.2394119)			
Sqr. mean soil moisture fraction	-.1030471	-.777505***			
	(.1166278)	(.2402404)			
SD daily soil moisture fraction		-.509464***			
		(.1156073)			
Mean daily ET~~*~~[1] (mm)			.4901121***	.6357687***	
			(.0735423)	(.0985801)	
Sqr. mean daily ET~~*~~[1]			-.086206***	-.118748***	
			(.0234848)	(.0254433)	
SD daily ET~~*~~[1]				-.2516986**	
				(.0997848)	
Mean moisture content (mm)~~**~~[2]					.0036395***
					(.0006759)
Sqr. mean daily moisture content~~**~~[2]					-.000017***
					(3.000e-06)
Observations	69923	69923	69923	69923	69923
R-squared	.4667911	.4712361	.4755177	.4770727	.4713225
Akaike's Crit	-21005.7	-21589.0	-22159.5	-22365.1	-21602.5
Bayesian Crit	-20969.0	-21543.3	-22122.9	-22319.4	-21565.9

Standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Notes: 1- ET shows the average daily evapotranspiration. 2- It shows the volumetric soil moisture interpolated from WBM to PRISM grid cells using the nearest neighbor method. Table lists regression coefficients and shows standard errors in brackets. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.~~

Table S2.~~A3.~~ Descriptive statistics of main variables for Eastern and Western US

	East		West	
Variables	Mean	Std. Dev.	Mean	Std. Dev.
Degree days from 10˚C to 29˚C	1877.79	433.54	1612.74	363.57
Degree days above 29˚C	58.01	57.13	82.11	80.29
Cumulative precipitation Apr-Sep (mm)	601.13	153.31	271.69	132.12
Mean daily soil moisture content (mm)	50.49	39.49	15.15	13.17
Number of days with high soil moisture	28.89	30.38	8.69	11.57
Number of days with low soil moisture	30.39	35.46	2.97	7.1
Surplus (sum of positive daily deviation, mm)	2546.95	2177.62	964.98	938.69
Deficit (sum of negative daily deviation, mm)	-2563.43	2200.22	-962.27	699.6
Degree days from 10˚C to 29˚C & low soil moisture	442.94	433.88	29.08	94.27
Degree days from 10˚C to 29˚C & high soil moisture	364	351.68	62.88	90.52
Degree days from 10˚C to 29˚C & normal soil moisture	1067.65	573.28	1462.24	426.27
Degree days above 29˚C & low soil moisture	20.19	32.55	.85	3.22
Degree days above 29˚C & high soil moisture	5.17	9.34	.76	2.41
Degree days above 29˚C & normal soil moisture	32.24	41.87	72.91	72.8
Index of extreme deficit	-1823.19	2339.6	-160.91	597.29
Index of extreme surplus	1942.11	2207.68	482.25	770.15
Index of normal soil moisture	-194.99	516.76	-406.16	434.96
Mean daily evapotranspiration (mm)	.6	.59	.15	.19
Mean daily soil moisture fraction	.71	.18	.68	.2
Mean daily soil moisture content (mm), alternative	50.52	39.41	15.17	13.2
Mean daily soil moisture content (mm), Apr-May	21.82	16.5	6.29	6.75
Mean daily soil moisture content (mm), Jun-Jul	17.7	15.77	5.14	4.53
Mean daily soil moisture content (mm), Aug-Sep	10.98	10.74	3.72	3.27
Observations	62094	62094	7829	7829

Table S3.~~A6.~~ Corn yield estimation with bi-monthly soil moisture metrics

	US Log CornYield	West Log CornYield	East Log CornYield
Degree days from 10˚C to 29˚C	.0003176***	.0004543***	.0002921***
	(.0000774)	(.0000853)	(.0000838)
Degree days above 29˚C	-.0044571***	-.0023373***	-.0047849***
	(.0006231)	(.0004904)	(.0006742)
Mean daily soil moisture content (mm), Apr-May	-.0029599*	.0045436**	-.0034124**
	(.0015561)	(.002061)	(.0015243)
Square of mean daily soil moisture content (mm), Apr-May	-9.800e-06	-.0000564	-2.600e-06
	(.000022)	(.0000581)	(.0000216)
Mean daily soil moisture content (mm), Jun-Jul	.0141021***	.0148123*	.013605***
	(.0019928)	(.0071408)	(.0020404)
Square of mean daily soil moisture content (mm), Jun-Jul	-.0001589***	-.0005616**	-.0001562***
	(.0000252)	(.0002422)	(.0000258)
Mean daily soil moisture content (mm), Aug-Sep	.0030501*	.007007	.0026044
	(.001805)	(.0049266)	(.0018059)
Square of mean daily soil moisture content (mm), Aug-Sep	-.0000385	-.000213	-.0000351
	(.0000291)	(.0002114)	(.0000294)
Observations	69923	7829	62094
R-squared	.4884616	.2782172	.515591
Akaike's Crit	-23898.8	-3040.6	-22112.9
Bayesian Crit	-23825.6	-2984.8	-22040.6

Standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Notes: Table lists regression coefficients and shows standard errors in brackets. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.~~

Table S4.~~A4.~~ Estimation of Model (1) for the East

	(1-a) Log CornYield	(1-b) Log CornYield	(1-d') Log CornYield	~~(2-c) Log CornYield~~
Degree days from 10˚C to 29˚C	.0003108***	.0003152***	.0003072***	~~.0002308**~~
	(.0000936)	(.0000868)	(.0000724)	~~(.000088)~~
Degree days above 29˚C	-.0056293***	-.0054707***	-.0052882***	~~-.0056523***~~
	(.0007259)	(.0007343)	(.0006442)	~~(.000946)~~
Cumulative precipitation Apr-Sep (mm)	.0009245***			
	(.0002502)			
Square of cumulative precipitation Apr-Sep	-7.000e-07***			
	(2.000e-07)			
Mean daily soil moisture content (mm)		.00319***		
		(.0006763)		
Square of mean daily soil moisture content		-.0000158***		
		(3.000e-06)		
Index of extreme deficit			.0000379***	~~.0000183***~~
			(5.700e-06)	~~(5.300e-06)~~
Index of extreme surplus			-.0000381***	~~-.0000225***~~
			(2.700e-06)	~~(6.800e-06)~~
Index of normal soil moisture			.0000292**	~~.0000433***~~
			(.0000112)	~~(.0000107)~~
~~Degree days from 10˚C to 29˚C x S1~~				~~.0001296**~~
				~~(.00006)~~
~~Degree days above 29˚C x S2~~				~~.0010785~~
				~~(.000888)~~
Observations	62094	62094	62094	~~62094~~
R-squared	.4997799	.4989592	.5205428	~~.5277292~~
Akaike's Crit	-20126.6	-20024.8	-22756.9	~~-23690.7~~
Bayesian Crit	-20090.4	-19988.6	-22711.8	~~-23627.5~~

Standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Notes: Table lists regression coefficients and shows standard errors in brackets. Model (1-d') is slightly different from Model (1-d) considering extreme deficit and extreme surplus metrics. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.~~

Table S5.~~A5.~~ Estimation of the Model (1) for the West

	(1-a) Log CornYield	(1-b) Log CornYield	(1-d') Log CornYield	~~(2-e) Log CornYield~~
Degree days from 10˚C to 29˚C	.0004426*** (.0000829)	.0004484*** (.0000823)	.0004539*** (.0000862)	~~.0004289*** (.000084)~~
Degree days above 29˚C	-.0020381*** (.000423)	-.0023744*** (.0004911)	-.0022938*** (.0004752)	~~-.0020711 (.0013393)~~
Cumulative precipitation Apr-Sep (mm)	.0005768 (.0003372)			
Square of cumulative precipitation Apr-Sep	-3.000e-07 (5.000e-07)			
Mean daily soil moisture content (mm)		.0078908** (.0027432)		
Square of mean daily soil moisture content		-.0000848** (.0000326)		
Index of extreme deficit			.0000255 (.0000271)	~~.0000217 (.0000342)~~
Index of extreme surplus			-9.800e-06 (7.600e-06)	~~-7.100e-06 (.0000123)~~
Index of normal soil moisture			.0000762** (.0000309)	~~.000077** (.0000334)~~
~~Degree days from 10˚C to 29˚C x S1~~				~~.0000281 (.0001257)~~
~~Degree days above 29˚C x S2~~				~~-.000238 (.001542)~~
Observations	7829	7829	7829	~~7829˚C~~
R-squared	.2784229	.2768284	.2772401	~~.2772526~~
Akaike's Crit	-3050.8	-3033.5	-3035.9	~~-3032.1~~
Bayesian Crit	-3022.9	-3005.6	-3001.1	~~-2983.3~~

Standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Notes: Table lists regression coefficients and shows standard errors in brackets. Model (1-d') is slightly different from Model (1-d) considering extreme deficit and extreme surplus metrics. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.~~

Table S6.~~A8.~~ West versus East in ~~C~~corn yield estimation with the interaction of heat and soil moisture (Model 2-a)

	(US) log CornYield	(West) log CornYield	(East) log CornYield
Degree days from 10˚C to 29˚C	.0003083*** (.0000685)	.0004344*** (.0000847)	.0002963*** (.0000736)
dday29˚C & SM 75+ mm below normal (extreme deficit)	-.0082398*** (.0014372)	-.0074467* (.0035727)	-.0082928*** (.0014365)
dday29˚C & SM 25-75 mm below normal (deficit)	-.0062069*** (.0009793)	-.0033152* (.001627)	-.0061966*** (.0009797)
dday29˚C & SM 0-25 mm around normal (normal)	-.0037559*** (.0004045)	-.0024412*** (.0005053)	-.0041335*** (.0004376)
dday29˚C & SM 25-75 mm above normal (surplus)	-.0055709*** (.0012041)	-.004754* (.0024763)	-.005625*** (.0011677)
dday29˚C & SM 75+ mm above normal (extreme surplus)	-.0140295*** (.0019083)	.0095881 (.0128016)	-.0143573*** (.0018101)
Mean daily soil moisture content (mm)	.0026635*** (.0008153)	.0080027** (.0028858)	.0025636*** (.0008324)
Square of mean daily soil moisture content	-.0000161*** (2.600e-06)	-.0000844** (.0000326)	-.0000156*** (2.600e-06)
Observations	69923	7829	62094
R-squared	.4921263	.2777862	.5149811
Akaike's Crit	-24401.6	-3035.9	-22034.8
Bayesian Crit	-24328.3	-2980.2	-21962.5

Standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

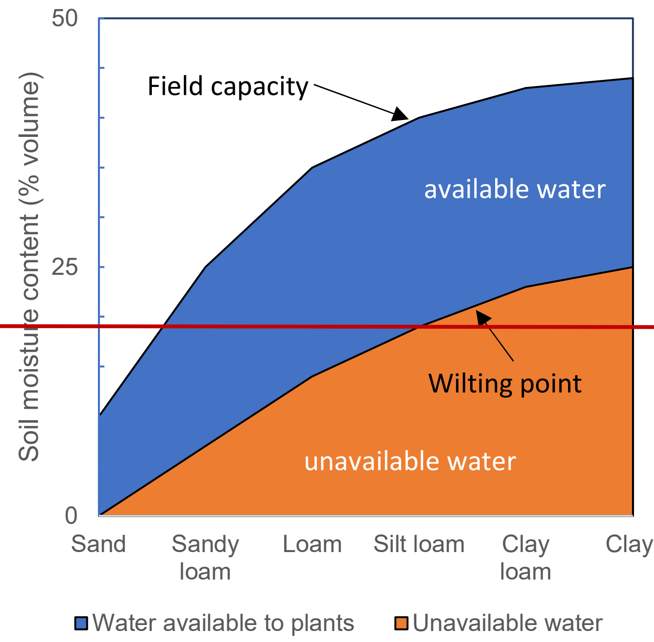
Notes: Table lists regression coefficients and shows standard errors in brackets. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.~~

Table S7.~~A9.~~ West versus East in ~~E~~estimation of corn yields while splitting the soil moisture indicators (Model 2-b)

	(US) log CornYield	(West) log CornYield	(East) log CornYield
Degree days from 10˚C to 29˚C	.0003154***	.0004451***	.0002983***
	(.0000689)	(.0000919)	(.000074)
Degree days above 29˚C	-.004044***	-.0020707***	-.0044516***
	(.0005384)	(.0005793)	(.0005981)
Index of normal soil moisture when T > T*	.0001199***	.0001805	.0001034***
	(.0000342)	(.0001426)	(.0000358)
Index of extreme moisture surplus when T > T*	-.0000628***	-.0001173	-.0000586***
	(.0000151)	(.0001071)	(.0000149)
Index of extreme moisture deficit when T > T*	.000092***	-.0000526	.0000817***
	(.0000234)	(.0000978)	(.0000229)
Index of extreme moisture deficit when T < T*	.0000209***	.0000287	.0000223***
	(7.100e-06)	(.0000337)	(7.000e-06)
Index of extreme moisture surplus when T < T*	-.0000326***	-5.700e-06	-.0000334***
	(3.200e-06)	(6.500e-06)	(3.200e-06)
Index of normal soil moisture when T < T*	.000028**	.000063**	.0000247**
	(.0000105)	(.0000249)	(.0000102)
Observations	69923	7829	62094
R-squared	.5006312	.2782242	.5262193
Akaike's Crit	-25582.4	-3040.6	-23490.5
Bayesian Crit	-25509.2	-2984.9	-23418.2

Standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Notes: Table lists regression coefficients and shows standard errors in brackets. ~~Temperature is in degree Celsius and soil moisture in mm in 1000 mm topsoil. The soil moisture is obtained from WBM at 6 arcmin output while precipitation and temperature are taken from PRISM at 2.5 arcmin. They are aggregated from grid cells to counties based on crop area weight. Yield data is acquired from the USDA. The constant term and coefficients on the interaction of each state and time trends are not reported.~~



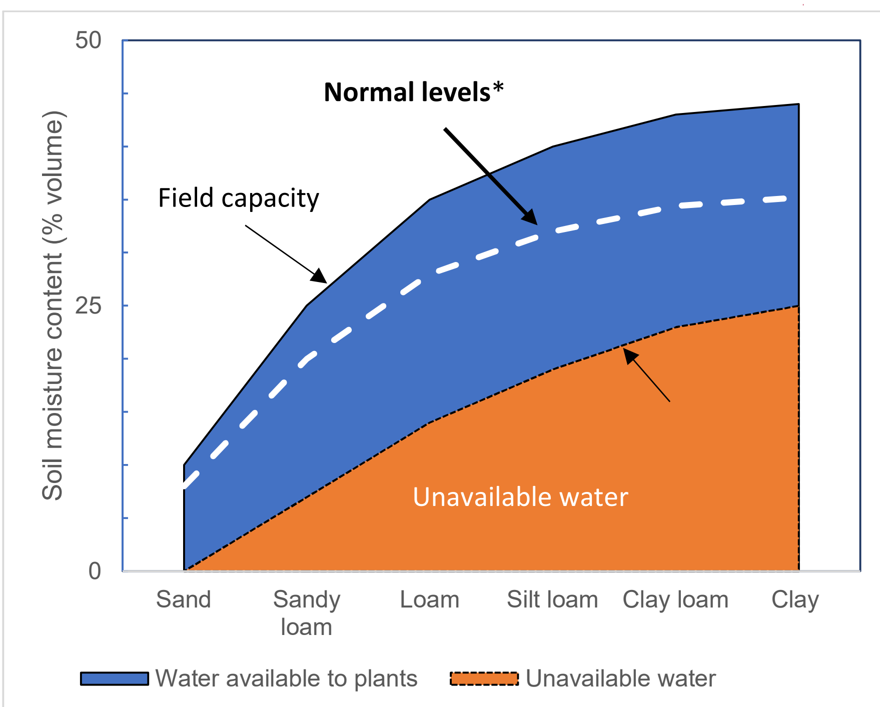


Figure S1~~2~~. Soil texture affects normal moisture levels. The sandy soil has the lowest normal level while the clay has the highest normal levels. ~~Soil texture affects the wilting point, field water-holding capacity, and the moisture available to plants. This suggests that sandy soil has the lowest wilting point as well as low field capacity. As most of the water infiltrates, this leaves a little amount of moisture available to plants (Tsoar, 2005).~~

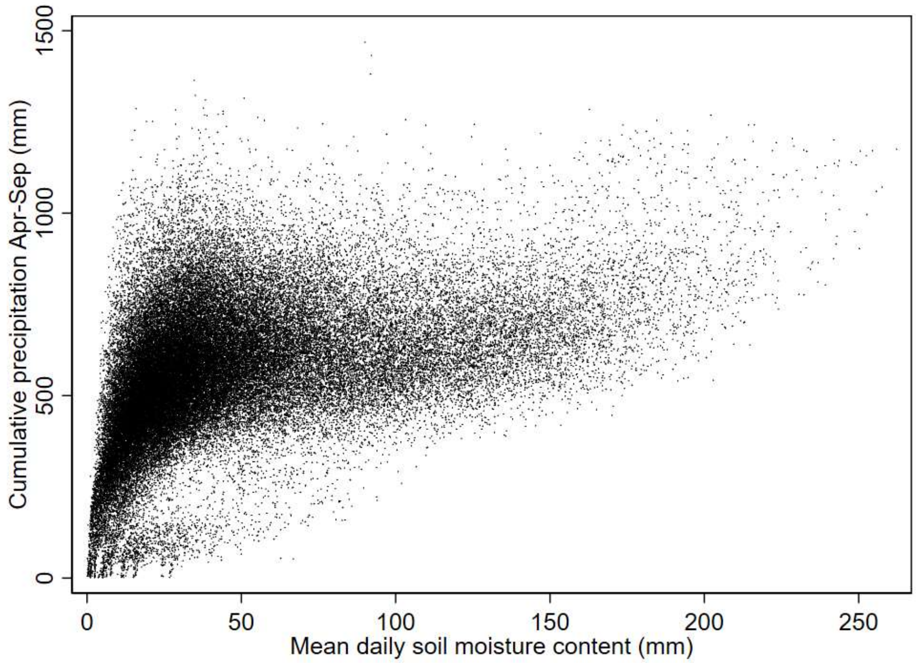


Figure S2.~~9.~~ WBM mean soil moisture versus PRISM cumulative precipitation for 1981-2015 by US counties.

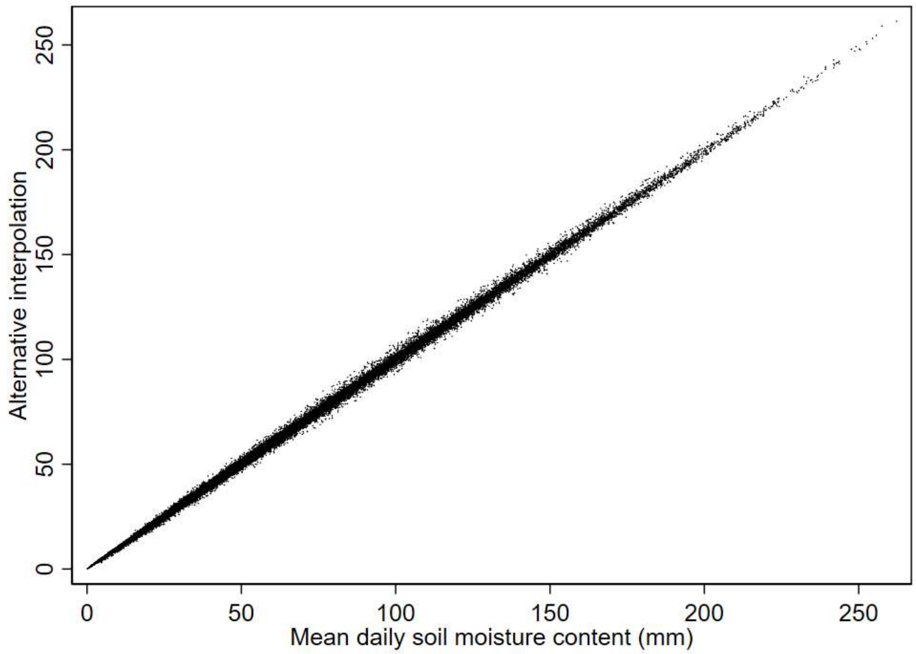


Figure S3.~~A1.~~ County-level mean seasonal soil moisture based on bilinear interpolation versus alternative interpolation (nearest-neighbor) from WBM 6 arcmin grids to PRISM 2.5 arcmin resolution for the 1981-2015 period.

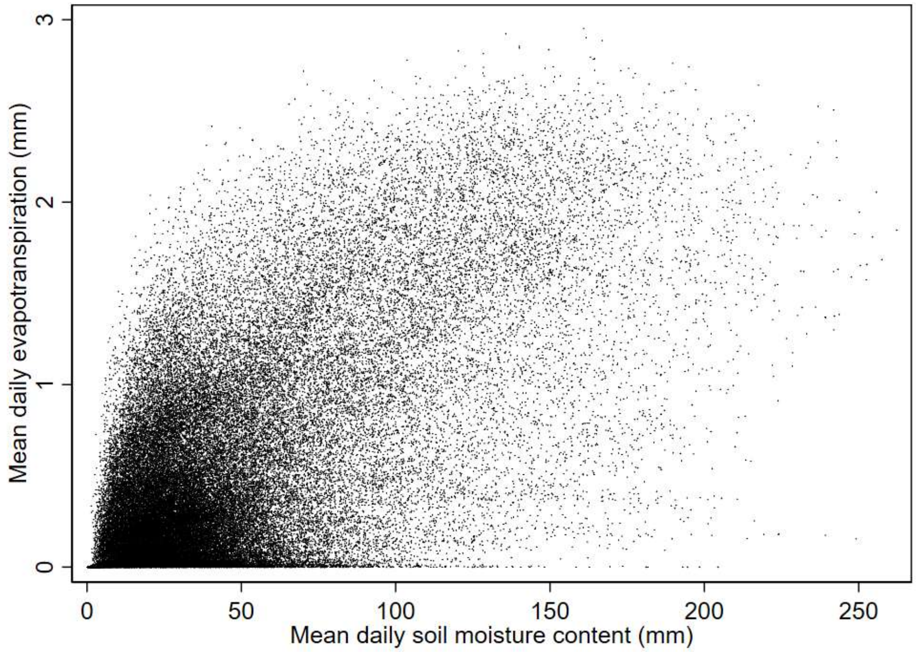


Figure S4.~~A2.~~ County-level mean soil moisture versus mean ET aggregated from WBM for the 1981-2015 period.

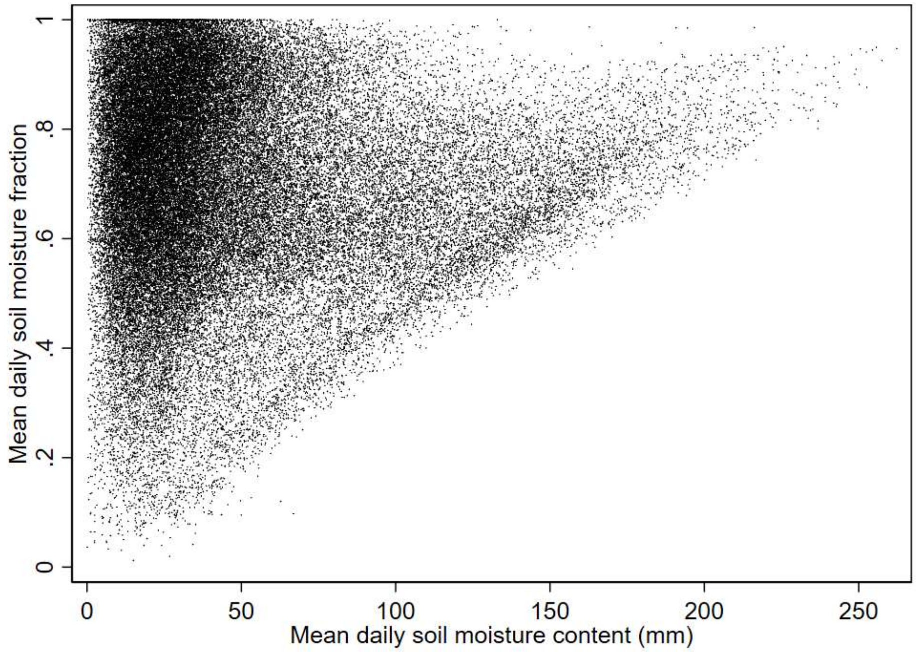
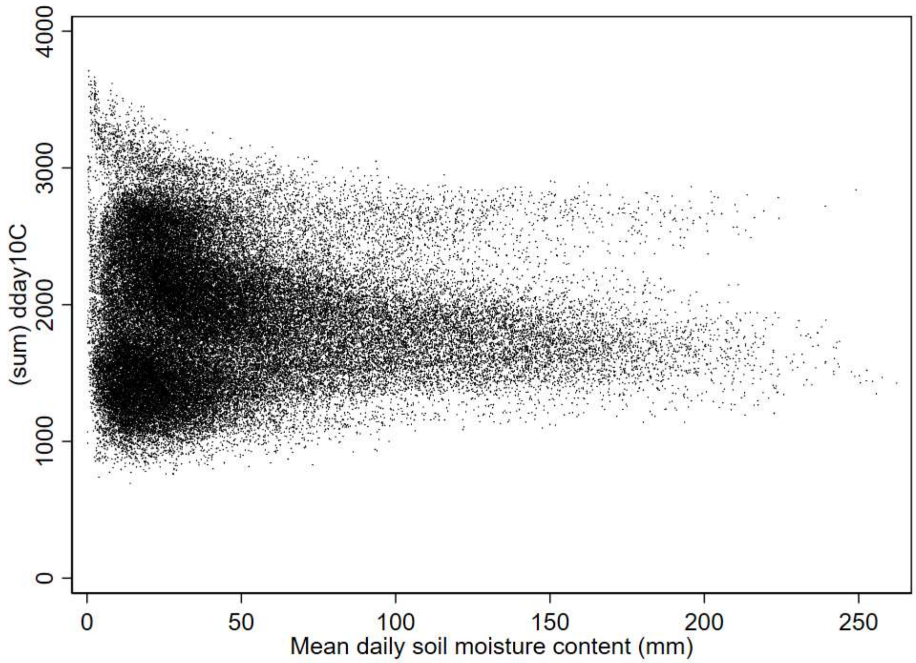


Figure S5.~~A3.~~ County-level mean volumetric soil moisture content versus mean of soil moisture fraction aggregated from WBM for the 1981-2015 period.



1390 **Figure S6.~~A4.~~ County-level seasonal mean soil moisture versus seasonal heat index aggregated from WBM and PRISM for the 1981-2015 period.**