

Earth Syst. Dynam. Discuss., referee comment RC1  
<https://doi.org/10.5194/esd-2021-24-RC1>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on esd-2021-24**

Corey Lesk (Referee)

---

Referee comment on "Impacts of compound hot-dry extremes on US soybean yields" by  
Raed Hamed et al., Earth Syst. Dynam. Discuss.,  
<https://doi.org/10.5194/esd-2021-24-RC1>, 2021

---

This paper examines the dependence of county level historical soybean yields on a suite of climate variables, with a particular eye to the influence of compound extremes. The paper is well motivated by recent climate and crop science literature, and nicely illustrates the particular relevance of hot and dry extremes for soybean yields. It further outlines historical trends in such extremes in relation to their univariate temperature and soil moisture components (using proxies to extend the time series). The manuscript is well written and the discussion raises a lot of interesting points. Overall, I think the authors did a great job and the paper should be considered for publication with revisions.

To me, the weaknesses of the paper are a 1) that there are a few methodological concerns and 2) that the paper doesn't extend that much beyond what is already fairly well established, even though it could using its data and methods. I elaborate below on these critiques and suggest some ways the authors could make the paper more compelling in revision.

### General comments

#### Methodological concerns:

The statistical modeling framework is quite detailed and meticulous and there is much attention paid to many sources of confusion, error, or interpretability issues. I commend these solid methods. However, I think the contradiction in scale between the model calibration (pooled national data) and its application (county-level) data greatly limits the soundness of the otherwise meticulous method. This to me is one of the main things to address in revisions.

The use of pooled national data to calibrate the model and then the subsequent use of that model to assess yield variability at county scales seems a bit inconsistent. All county-year pairs of yield values were combined into a larger dataset as in the Troy et al. (2015) paper referenced, but then the resulting model was applied for individual counties, whereas Troy et al. ran all their analysis at national scale. It doesn't seem common or intuitive to me to calibrate and run models at such different scales, and this isn't justified, discussed, or acknowledged in any detail. What the authors did here is an expedient way to make their analysis applicable at a county level, but leverage a wider dataset to increase degrees of freedom and enable the testing of more complex models. But the cost is that the mismatch in scales raises questions on whether some results are a result of the mismatch, or truly robust results.

A few ideas on how this might be influencing results: First, it could be that the particular relevance of concurrent low August soil moisture and high Tmax in Illinois (nicely illustrated in Fig. 5) is a result more of the suitability of the nationally fitted model at that location, rather than the relevance of compound extremes more generally. This concurrent temperature-moisture result does not strongly agree with the results of Mourtzinis et al. (2015) which found larger temperature impacts in southern states, nor Zipper et al. (2016) which found stronger drought impacts on soy in southern states. It could very well be that the compound impact is larger in Illinois, but the ambiguity induced by the contradiction in scales in the modeling makes the result not as robust as it could be.

Second, there is strong variation in the significance and model  $r^2$  across counties (Fig. 3). Generally, this itself raises the question of whether models should be calibrated locally (indeed we wouldn't expect nationally-consistent model to be optimal everywhere, even if parameters of said model are estimated locally). More specifically, the model performs quite strongly in Illinois, which to me raises the question of whether the strong compound Tmax-SM impact is really just because the national model works very well in Illinois (i.e., a result determined by methods, so not very robust).

The data-driven approach also is attractive because it allows the 'most important' months to be identified. However, I have doubts about the methods for this around the selection of the earlier among collinear climate signals (see detailed comments). Further, it's not only collinearity in time, but among variables at the same time, that matters, as you discuss in lines 280-90. Leaf- to field-scale experiments show that light is very important for crops, so that it is excluded from the modeling is a methodological choice that requires careful interpretation. Another example is the idea that temperature is a strong predictor because it encapsulates many moisture and heat related stressors, as in many of the cited papers. Point is, as a result of these assumptions regarding variable exclusion in the methods, the model specification is actually not as data-driven in the end, so worth considering other approaches with their own strengths/weaknesses:

I think an alternative approach could be to compromise a bit on data-driven model specification and simply prescribe the model structure a priori. This is appropriate because you cite much literature in your introduction on why and how compound extremes should matter, so you have a prior to base the specification on. You can then run the stepwise model selection on a smaller set of predictors for each county and see if results are robust, e.g. Illinois/August still pops out. I understand there is a compromise in this

alternative, but it might complement and add confidence given concerns about the original approach. It might also add some confidence to run the panel regression for the full national model (i.e. what is the national value of the coefficients in Fig. 4a).

Novelty and advancing understanding:

I really like how this paper clearly puts data and nice visualization to the idea and existence of examples of compound extreme impacts on crops. It also goes into some detail on soybean, a crop for which there is somewhat less attention on the topic. However, I think the core conclusions of this paper have essentially already been established. For instance, the Illinois case study in Figure 5 is an excellent visualization, but its message essentially quite similar to Kent et al. (2015, Fig. 2c, a great paper on maize which might be a handy reference to include) combined with what was published in Matiu et al. (2017), namely that such compound impacts do occur in places. It's useful that this paper points out that this occurs for this particular crop and location, but a similar point has been made in Ortiz-Bobea et al. (2019, also a great paper that probably needs to be referenced in this paper). Examining trends in concurrent heat-drought is also a useful topic, but has been covered in some detail in e.g. Sarhadi et al. (2019) and Lesk and Anderson (2021). Overlap with past research is a great contribution, but I think it does demand that this study go a bit deeper.

For instance, I think the study could choose one result to go into some more detail on to really gain some new insight. One could be how exactly these extremes are impactful in some places, less so in others, and some of the uncertainties and challenges around understanding this (see minor comments). If the particular importance of compound impacts in Illinois turns out to be a robust finding, why exactly might this be? The authors hypothesize a link to a reversal of the crop induced land-surface cooling during dry episodes, leading to compound impacts (as suggested by Mueller et al. 2016). Many papers have recently speculated about this, and you have the data to examine this in great detail for this location (e.g. compositing and examining coevolution of AET, SM and Tmax timeseries over hot-dry events) and add valuable concreteness to the speculation. Another direction could be to assess drivers of the trends in Figure 6 more concretely, possible roles of agriculture itself in influencing those trends, roles of modes of climate variability and aerosols (e.g. Fig. 6c probably shows some dependence of on the changepoint selection as visible in Fig. 6d, why that might be, and does it say anything meaningful about future change?).

References:

Kent, C., Pope, E., Thompson, V., Lewis, K., Scaife, A. A., & Dunstone, N. (2017). Using climate model simulations to assess the current climate risk to maize production. *Environmental Research Letters*, 12(5), 054012.

Lesk, C., & Anderson, W. (2021). Decadal variability modulates trends in concurrent heat and drought over global croplands. *Environmental Research Letters*, 16, 055024.

Ortiz-Bobea, A., Wang, H., Carrillo, C. M., & Ault, T. R. (2019). Unpacking the climatic drivers of US agricultural yields. *Environmental Research Letters*, 14(6), 064003.

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

#### Detailed comments

Lines 25-30: Introduction has great context for why we should care about US soybean. I think it would give helpful context to readers to stay somewhere here that a large portion of soybean is produced for feed.

Cassidy, E. S., West, P. C., Gerber, J. S., & Foley, J. A. (2013). Redefining agricultural yields: from tonnes to people nourished per hectare. *Environmental Research Letters*, 8(3), 034015.

Line 46: Agreed that more attention is needed especially for soy, but probably should cite a few missing studies that have been written on the topic, consider: Rigden et al. (2020), Ortiz-Bobea et al. (2019), Haqiqi et al. (2021)

Haqiqi, I., Grogan, D. S., Hertel, T. W., & Schlenker, W. (2021). Quantifying the impacts of compound extremes on agriculture. *Hydrology and Earth System Sciences*, 25(2), 551-564.

Rigden, A. J., Mueller, N. D., Holbrook, N. M., Pillai, N., & Huybers, P. (2020). Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting US maize yields. *Nature Food*, 1(2), 127-133.

Line 91: There's some evidence that 30mm/day is not a high enough rainfall amount for negative impacts on soy yields in the US (Lesk et al. 2020). I wonder if heat/extreme rainfall would pop out as a compound (possibly positive/compensating) impact on crops if you used a higher threshold.

Lesk, C., Coffel, E., & Horton, R. (2020). Net benefits to US soy and maize yields from intensifying hourly rainfall. *Nature Climate Change*, 10(9), 819-822.

Line 115: Selecting the earlier among collinear monthly predictors raises an interesting question of whether the signal for one variable preceding the others in time necessarily means that variable is the driver of the crop response. That is, the later signal could easily have caused the real impact on the crop, and the earlier one is predictive because of its correlation with the later. This is worth justifying more, or at least acknowledging as an important assumption (because it partly determines what variables ultimately can be considered drivers of compound impacts in your methodology). Could be an angle for going to deeper on why Illinois pops out for example.

Lines 123-5: Would be good to see more detail on which/why other interactions were left out, and exactly how much 'better' the selected interaction was than other candidates, as this is key to your conclusion. A weaker alternative could be to simply assert that this interaction is one you have a good reason to care about (i.e. the hypothesized interaction is the motivation of your analysis).

Line 148: Ref needed for energy limited AET

Lines 165-6: I'm surprised SM and Tmax are not more strongly collinear in August given the land-atmosphere feedbacks and their involvement in the compound extreme. This should be discussed more and possibly examined in depth. E.g. – are the feedbacks really setting up earlier in the season, so SM and Tmax are more collinear then, and thus get excluded from the analysis? If so, this raises questions of whether August then really is the most important for yield, or just popping up because of this methodological decision (although some other papers you cite do support August being important). There's something deeper to understand here.

Lines 176-8: I don't see this result supported by data in Fig. 3A, please explain.

Line 185: interesting that model predicts yields better in south (as in Schaubberger) – crops here not necessarily ‘decoupled’ from climate, as warmer seasons benefit yields...

Fig 3b: The question this raises for me is if the north-south gradient in  $r^2$  relates to a gradient in suitability of the nationally tuned model. Indeed, since Illinois is a major soybean producer, its contribution of data to the pooled sample is particularly high (meaning the strength of the prediction could be because the national model fits best there, while other models would fit just as well if calibrated on smaller scales).

Lines 192-3: how does this square with e.g. Li et al. 2019 who show both very high and very low soil moisture are damaging?

Line 220: Do you consider AET as a climate variable, or a plant/crop variable (because carbon gain comes with water loss necessarily).

Fig 5: very nice figure. Interesting that there is some tail dependence for hot and dry extremes, in that in this bottom-right quadrant you see very extreme joint temp/sm anomalies compared to the others. does this raise questions of causality around the fact that such extreme low SM values can only be reached with very high Tmax? in other words, is the yield impact especially severe because of the compound impacts of temp and moisture, or simply because of extreme moisture impacts that can only happen if T is also high?

Also could clarify in panel b that the slopes of those lines are the  $t_{max}$  slope + interaction slope \* 5-50-95 percentile soil moisture value. Also, given the low sample of hot and wet events, I wonder if it even makes sense to draw the blue line beyond 2sigma Tmax anomalies (there are no such events observed as you say, probably for an important climate reason).

Line 255: there is a strong role of relatively few years in these time series, and possible some signal of climate oscillations, that may be worth at least referring to (Lesk and Anderson 2021 ERL and/or refs therein might be useful)

Line 290: pun intended?

Lines 300-310: It’s also worth noting that Schlenker and Roberts (2009) and Schaubberger et al (2017) too found that the crop damages beyond the ~30 degree threshold were mitigated when moisture was sufficient (either from irrigation or rain). So your findings are in loose agreement with those studies too, in addition to Carter et al. (2016), Siebert et al. (2017), and Troy et al. (2015). I also think it’s worth acknowledging that wet

conditions may simply prevent very high temperatures, thus reducing exposure rather than sensitivity to heat (see my comments on Fig. 5).

Lines 309-311: I have a paper in review showing evidence for this globally. If it is accepted in time, it would be a good reference.

Lines 311-313: Again, I think you're overstating the lack of attention a bit, see suggested refs above.

Nice work thanks!