Reviewer 1

The authors responded well to all reviewer comments and performed additional analyses as requested by the reviewers. The article itself is interesting and has the potential to further advance the scientific understanding of heat and drought effects on crop development and yield and corresponding effects of irrigation. I clearly see the merit of the work presented here. However, I also see a risk that the results will be misunderstood by many researchers and in particular by the media. The three reviewers and the authors agree that heat and drought effects cannot be completely separated because of a considerable interaction and that this interaction cannot be represented by the methodology used by the authors. What if the single effects of water and heat are small and the interaction explains most of the total effect? Consequently, I'm afraid that the central finding of 65% water alleviation effect and 35% heat alleviation effect of irrigation will be communicated in subsequent studies without pointing to the limitations and context of the present study. Clearly this should be avoided. The authors added an explanation to the discussion explaining these limitations but I still think that this will not help much to avoid misinterpretation. I suggest therefore to mention this aspect already in the abstract section of the article to highlight the relevance of this limitation. To make this very explicit: it is not my intention to lower the merit of the present work but to avoid misinterpretation! I suggest therefore to pack it in a way that one final sentence will be added to the abstract explaining that future research should make efforts to consider as well the interaction effects between heat and drought alleviation.

We thank the reviewer for the constructive suggestions, which have significantly improved our study. We have added text to the abstract suggesting that future research consider the interaction effects between heat and drought alleviation. "considering the potentially strong interaction between water and heat stress, future research on irrigation benefits should explore the interaction effects between heat and drought alleviation."

Reviewer 2

Compared to the previous version, the authors have consolidated the statistical analysis with an explicit consideration of collinearities. Nevertheless, I think that the results given for thermal stresses should be more discussed and the analysis of the results should be pushed further. The value of 35% could be compared with the results of the literature. For example, the study by Lobell et al. 2011 gives orders of magnitude of 1% yield loss per ℃ above 30 for well-watered crops and 1.7% per ℃ for poorly watered crops. In the study, the authors find a value of 1% per ℃ in the case of irrigated crops, which is the case where the thermal effect is well isolated from the hydric effect, as shown by the low uncertainty in the thermal sensitivity coefficient shown in Figure 7. Such results is comparable to that of Lobell and thus demonstrate the interest of the method used to aggregate the data of the study. Such a result is somewhat at odds with the strong effect of temperature found by the statistical analysis on yield variations between irrigated and non-irrigated systems, which I still believe suffers from uncertainties related to residual collinearities. A FIV of 2.2, while below the severe collinearity threshold of 5, still shows collinearities that give a 50% increased variance on the parameters. Could this explain a discrepancy between the results obtained under irrigated conditions where the effects of water and heat stress are well deconvoluted, but also the discrepancy with the Lobell 2011 study in case of water stressed crop 5(1.7% per ℃). I encourage the authors to really consider this discussion, which does not questioned the interest of the paper which clearly demonstrates the need to take temperature into account in the operating models.

Thank you for this detailed and constructive review. We have made the suggested changes, as detailed in the point-by-point response below.

We appreciate the suggestion to compare our findings to Lobell et al 2011. We want to clarify an important point -- the estimation reported by Lobell et al. 2011 is based on degree days (GDD_{30+}, which measures sum of temperature exposure above 30 ℃) and not the mean temperature. So the 1% yield loss reported is per degree day, not per ℃. However, Fig 2b in Lobell et al. 2011 shows the yield sensitivity of 1 ℃ warming. We can see the poorly watered maize shows high sensitivity (<-10% per ℃) to warming stress based on the estimated yield sensitivity.

We have added a sentence in line 382-384 to show the similar findings of the previous studies on the weakened temperature sensitivity of yield due to irrigation application. "We found that irrigation not only lowered water and high temperature stress, but also made yield less sensitive to water and high temperature stress (Figure 8a-c), consistent with previous studies (Troy et al., 2015; Tack et al., 2017). For example, field data across Africa suggests that better water management can reduce yield loss due to heat stress from -1.7% per degree days to -1% per degree days (Lobell et al., 2011). "

From a statistical point of view, it would be good to define what the uncertainties reported in the paper correspond to (error bar on figure 7, % given on the contribution of water and thermal effects on page 13.

Thanks for your suggestion. We have added details about what those uncertainties represent for both Figure 7 and the contributions of water and thermal effects reported on page 13.

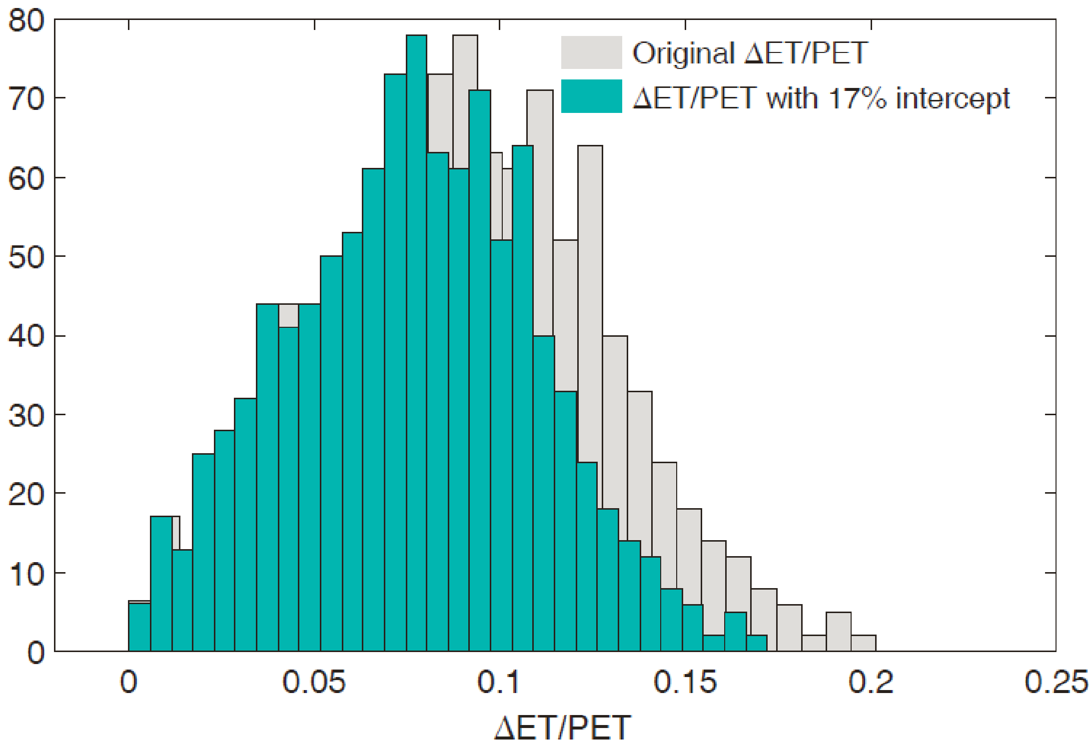
It would be good to give the VIF for the adjustment of the parameters of equation 7. It is likely to be high in the case of rainfed crops.

We added the VIF in Line 369-371 for equation 7 for both irrigated and rainfed maize yield. Then we find the VIF for equation 7 is still not very high with 2.8 and 3.6 for irrigated and rainfed maize yield, respectively. We added this in Line 369-371 "we quantified the variance inflation factor (VIF) in the model; this was found to be well below standard thresholds, with a value of 2.8 and 3.6 for irrigated and rainfed maize yield, respectively"

L200 : I disagree. In figure 5 of supplemental, I saw rather an overestimation of the Modis PET. The MODIS PET data were always at the upper limit of the observation data. Taking time average will exhibit bias (~15-25% from a visual impression). This might explain the low ET/PET ratio obtained with irrigated fields (I expected values closer to 1 and I am surprised that some irrigated fields led to ET/PET <0.6) - What is the consequence of such overestimation on the study

Indeed, we confirmed that across all site-year samples, MODIS PET is on average 17% higher than the observation data, which partially explained the relatively lower ET/PET in irrigated fields. However, the main reason for the lower than expected ET/PET is because current irrigation systems cannot ensure crops are fully irrigated during the growing season as it is not precision irrigation and tends to be run somewhat infrequently due to the lumpy expenditure.

Obviously, overestimation in PET will impact the water stress and heat stress alleviation contribution. We simulate the effect of overestimated PET on our attribution analysis by adding a 17% positive offset to all PET samples. We find the mean of ΔEP/PET between irrigated and rainfed maize changed from 0.088 to 0.075 as the following figure shows.



Then we apply the equation (9) to ΔET/PET samples with 17% positive offset and ΔLST to estimate the water stress and heat stress alleviation contribution. We find the yield improvement due to water and high temperature stress alleviation is 61±10% and 39±5.7%, which is very close to the original estimation that water and high temperature stress alleviation contributed to 65±10% and 35±5.3% yield improvement. With this simulation, it is safe to say that overestimation in PET has a small effect on the attribution analysis and will not change our overall findings.

L340 : I would add here the variety effect.
Thanks for your suggestion. We have added this potential effect in Line 340: "(3) variety differences between irrigated and rainfed maize"

L340-342 not really useful, consider to remove such statement
We followed the suggestion and removed the two lines and updated the figure accordingly.

360 : is 2% really significant ? (further a -2% was found not significant (375-376).
Here the 2% difference is not significant. We have added the significance level for

this comparison. The -2% temperature sensitivity estimation in L375-376 is significant. We added signs in Figure 7 to represent the significance of temperature sensitivity estimation.

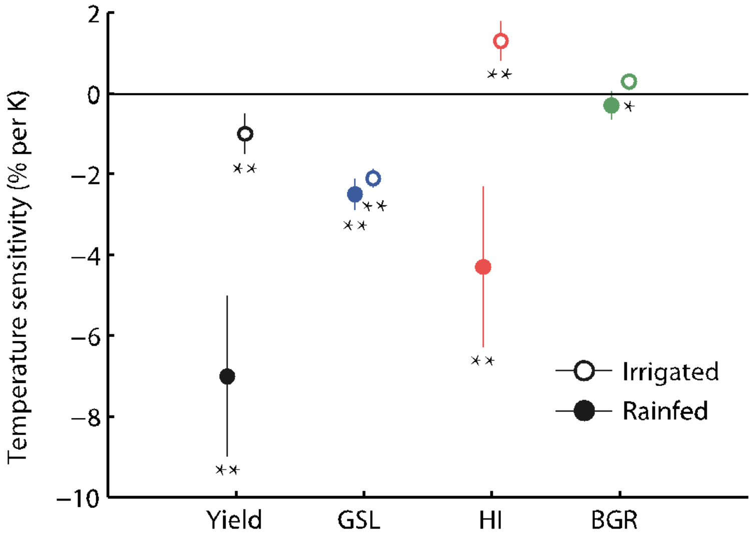


Figure 7: Temperature sensitivity of yield and yield components (GSL, HI and BGR) for irrigated and rainfed maize areas. The error bars represent the 95% confidence interval of estimated temperature sensitivity. ** indicates a significant estimation of temperature sensitivity with $p<0.01$ while * indicates significance with $p<0.05$.

429-431 : I am not sure that scale of study was really helpful to farmers. Heat and drought tolerance are likely analysed in selection plan. I would removed such statement

We followed the suggestion and have removed these statements.

447-452 : I disagree, several models already take crop temperature into account as well as the interaction between heat and water stresses (CLM, STICS, CERES?,). It would have been interesting to analyse in the discussion how heat stress are addressed in such models and if the order of magnitude found if the statistical analysis are retrieved (for instance by making numerical experiments).

Indeed, there are some crop models (like STICS and APSIM) already taking canopy temperature into account when addressing heat stress. We have added some discussion about how canopy temperature stresses are estimated in these models in line 456-468. "Relatedly, recent studies compared heat stress representation in crops models which explicitly simulate canopy temperature (Webber et al., 2017). For example, STICS estimates canopy temperature using canopy energy balances which account for net radiation, soil heat flux, evapotranspiration and aerodynamic resistance (Brisson et al., 2003). In APSIM, canopy temperature is taken as 6 ℃ higher than air temperature when the crop is fully stressed and 6 ℃ cooler than air temperature when the crop is fully transpiring. Between these limits, the basis of the expression for canopy temperature is the relationship between temperature difference (canopy temperature minus air temperature) and the ratio of actual and potential evapotranspiration (Webber et al., 2017). This model comparison study suggests that models using canopy temperature to account for heat stress effects indeed outperform those models depending on air temperature but the model comparison also identified a wide range for the simulated canopy temperature in current crop models. "

For the crop models we used in AgMIP project, CLM-crop is the only crop model taking canopy temperature into account. We also do not find any numerical model experiments that did such analysis like ours to disentangle the heat and water stress effects. Therefore, we suggest crop model communities pay more attention to this issue and set up more model experiments by including crop models explicitly considering canopy energy balance to better address heat stress and water stress alleviation due to irrigation and thus better simulate crop yields with human management practices.