Interactive comment on "Agricultural management effects on mean and extreme temperature trends" *by* Aine M. Gormley-Gallagher et al.

Anonymous Referee #1

Received and published: 16 August 2020

This paper looks at the effects of prescribed representations of conservation agriculture and irrigation on mean annual 2m and maximum daytime temperature in CESM. There is some interesting analysis and potential for results that could be useful for the community. There are **aspects of the sub-grid scale vs grid scale analysis**, and **possibility for critique of whether more processes enhance model skill**, that are **intriguing**. The figures are generally well presented. However, there are several issues that need to be addressed.

We greatly thank the reviewer for the appreciation of the manuscript and for the constructive comments, which greatly helped to improve the quality of the study. Here below, we provide a point-by-point response to each comment. The modified manuscript text is shown in italics.

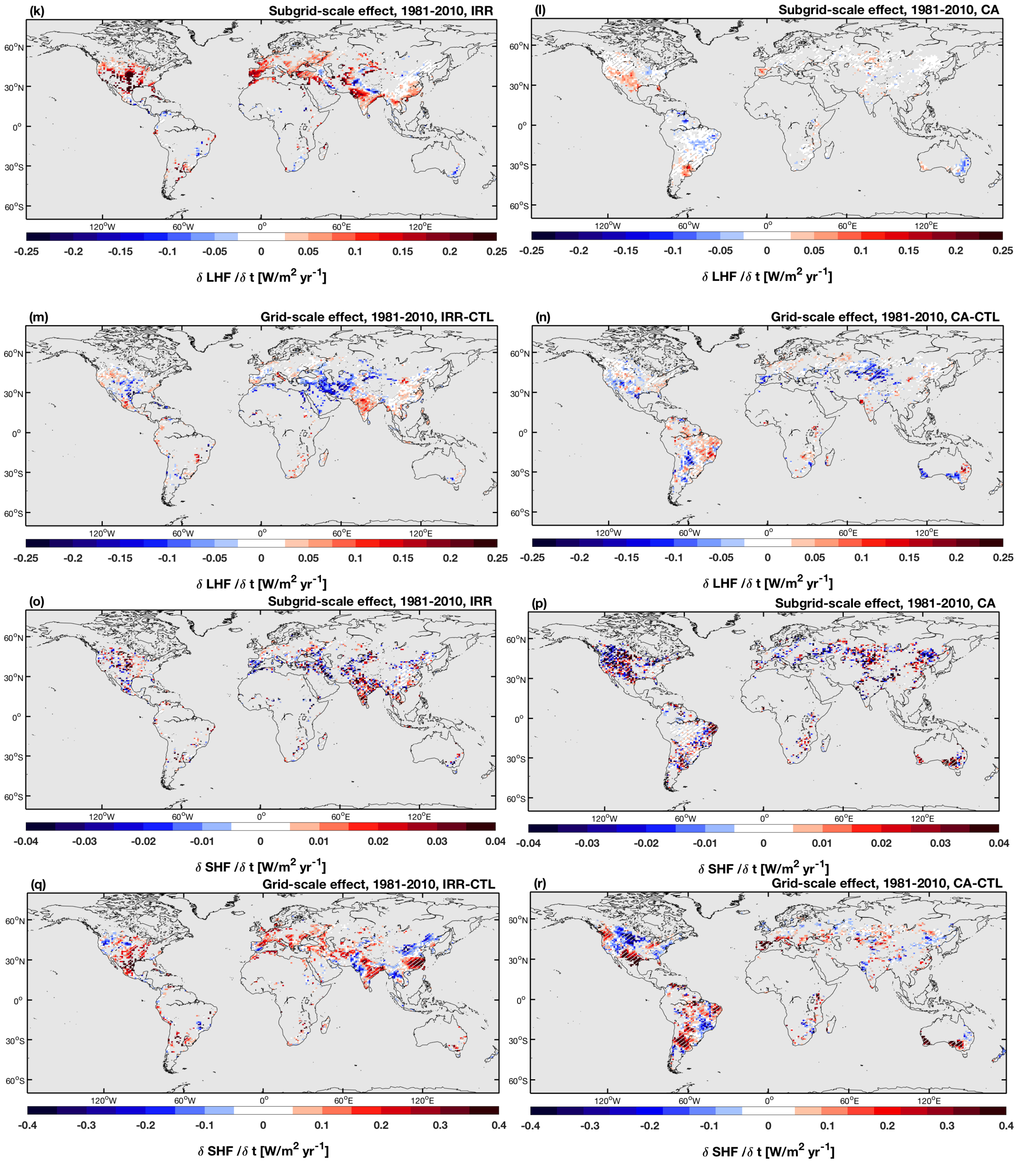
1. The paper reads like a combination of previously published results (specifically, the ensembles used are already published in Theiry et al. (2017) and Hirsch et al. (2018)). That might be unfair, but the regression analyses is simple and it seems unlikely it wasn't done separately for CA and irrigation, and much of the explanatory analysis references these two papers. It is the responsibility of the authors to show clearly why this is novel compared to what has come before.

Reply. We confirm that the analysis presented here is based on simulations that have been published previously. However, we believe that this new study moves beyond the state of the art in three ways. First, the main novelty of the current study lies in the explicit focus on trends, whereas previous studies focused on the influence of land management on the climatology (of means and extreme indicators). Second, the explicit focus on the subgrid versus grid-scale response offers important new insights on the local land surface land surface response to land management. Third, this study for the first time explicitly considers the radiative forcing resulting from realistic land management. We find that the positive radiative forcing signal arising from enhanced atmospheric water vapour is too weak to offset the local cooling from the irrigation-induced increase in the latent heat flux. This is been emphasised more clearly by including new evidence on the latent and sensible heat fluxes (warming/cooling trends as well as spatial averages) and the results and abstract adjusted to reflect these results as well as the reviewer's Point 2 and Point 6.

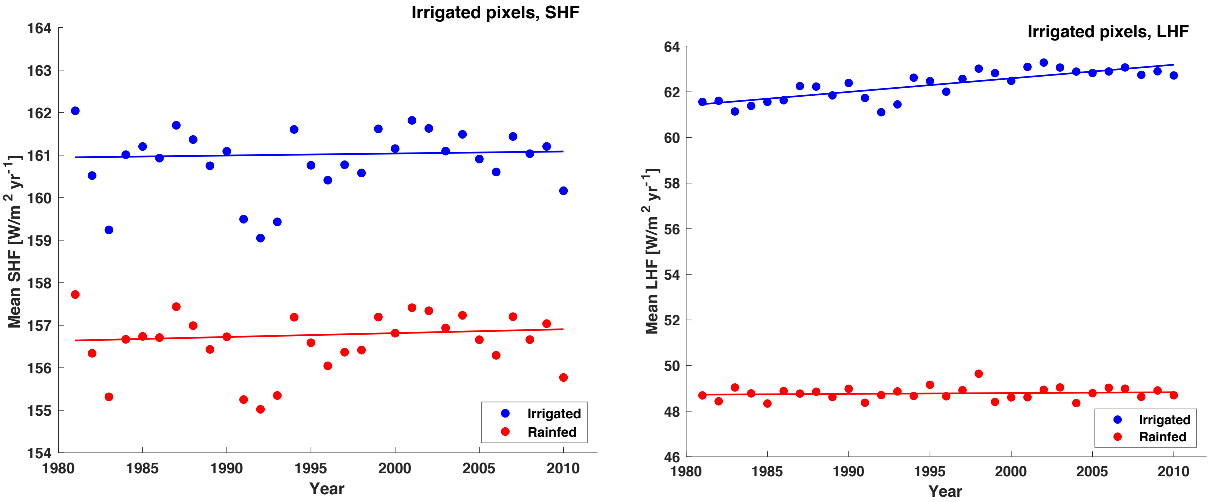
The new abstract reads as:

> ***Abstract.*** *Understanding and quantifying land management impacts on local climate is important for distinguishing between the effects of land management and large-scale radiative forcings at the top of the atmosphere. This study for the first time explicitly considers the radiative forcing resulting from realistic land management and offers new insights on the local land surface response to land management. Regression-based trend analysis is applied to observations and present-day ensemble simulations with the Community Earth System Model (CESM) version 1.2.2 to assess the impact of irrigation and conservation agriculture (CA) on warming trends using an approach that is less sensitive to temperature extremes. At the regional scale, an irrigation- and CA-induced acceleration of the annual mean near-surface air temperature (T_{2m}) warming trends and the annual maximum daytime temperature (TXx) warming trends were evident. Estimation of the impact of irrigation and CA on the spatial average of the warming trends indicated that irrigation and CA have a pulse cooling effect on T_{2m} and TXx, after which the warming trends increase at a greater rate than the control simulations. This differed at the local (subgrid) scale under irrigation where surface temperature cooling and the dampening of warming trends were both evident. As the local surface warming trends, in contrast to regional trends, do not account for atmospheric (water vapour) feedbacks, their dampening confirms the importance of atmospheric feedbacks (water vapour forcing) in explaining the enhanced regional trends. At the land surface, the positive radiative forcing signal arising from enhanced atmospheric water vapour is too weak to offset the local cooling from the irrigation-induced increase in the evaporative fraction. Our results underline that agricultural management has complex and nonnegligible impacts on the local climate and highlight the need to account to carefully represent and evaluate land management in climate models.*

The new evidence, which has been added to the paper's Figure 6, show the Subgrid-scale differences between the irrigated and rainfed crop tile in the IRR ensemble and between CA and conventionally managed (CM) crops, for the latent heat flux (LHF) (k-l below) and the sensible heat flux (SHF) (o-p below). Grid-scale differences between the CTL and IRR ensemble and between CA and CM crops for LHF (m-n) and SHF (q-r) are also included over irrigated/CA pixels for comparative purposes, as detailed below. In addition, to the paper's Figure 7, data on the spatial average of the SHF (e) and LHF (f) warming rates for the irrigated and rainfed crop tiles over irrigated pixels, is now included, as shown in our Response Figure 2.



Response Figure 1 (adding to Figure 6 in the paper). Subgrid-scale differences between the irrigated and rainfed crop tile in the IRR ensemble (irrigated minus rainfed) (a, g, k and o) and between CA and conventionally managed (CM) crops (CA minus CM) (e, h, l and p). For T_s (a-b), ET (g-h), LHF (k-l) and SHF (o-p). Grid-scale differences between the CTL and IRR ensemble (IRR minus CTL) (c, e, I, m and q) and between CA and conventionally managed (CM) crops (CA minus CM) (b, f, h, n and r). For T_s (c-d), ET (i-j), TMQ (e-f), LHF (m-n) and SHF (q-r), displayed over irrigated/CA pixels for comparative purposes. Differences are based on the ensemble mean warming trends of each experiment for 1981–2010. Hatching denotes less than 10% change induced by the model on mean warming trends of lumped ensemble members.



Response Figure 2 (added to Figure 7 in the paper). Spatial average of the SHF (left) and LHF (right) warming rates for the irrigated and rainfed crop tiles over irrigated pixels. Data points specify the mean LHF and SHF values within the crop tiles and pixels specified. The slope was estimated using Sen's slope for the rainfed/CM (red), irrigated/CA (blue) experiments for the years 1981-2010.

2. The results (as shown in Table 1 especially) are difficult to reconcile with the statements made in the abstract and conclusions. Looking at Table 1, if the smallest RMSE (or the anomalies closest to zero) are considered, the Control simulation is better ~ 2/3 of the time. The abstract says, "our results underline. . . the need to account for land management in climate projections". Surely the opposite is true, as the Control scenario does better by the measure most used to assess model skill. Even within the results, there appear to be contradictions. Line 218: "the impact of irrigation and CA on the modeled spatially averaged temperatures. . . is an overall cooling effect". Line 223: "for the IRR and CA models. . . the spatially averaged T2m and TXx warming rates are higher than those of the CTL model".

Reply. The CTL simulation performs when considering the extreme (TXx) temperature results but not when considering mean temperatures (T2m). Particularly in the case of T2m for irrigation, the results in Table 1 show all cases are better all of the time. In one case for CA (all land), the RMSEs are equivalent, but otherwise the CTL RMSE is higher. So we disagree that the opposite is true in general, but we do agree that the statement in the abstract can be refined. It is also an opportunity for raising a critique/reflection point on whether more processes enhance model skill for the mean but not for extreme temperature, which adds to the addressing of the reviewer's introduction paragraph as well as the final point raised regarding the discussion. Therefore the abstract has been adjusted (see the new relevant statement below as well as the full abstract under our response to the reviewer's Point 1) as well as new results provided (also under reviewer's Point 1) so to better reconcile the results with the conclusion and abstract. The adjusted abstract statement reads as follows:

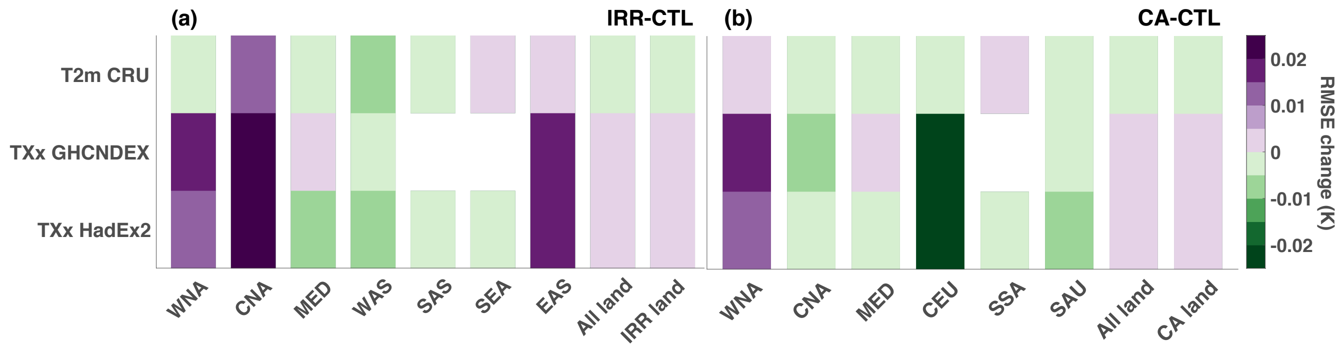
> *Our results underline that agricultural management has complex and nonnegligible impacts on the local climate and highlight the need to carefully represent and evaluate land management in climate models.*

Regarding the apparent contradiction between lines 218 and 223, the 'cooling effect' noted in line 218 (Figure 4) refers to a decrease in absolute temperature, that is, the intercept of the regression. Line 223, on the other hand, does refer to the trend over time, that is, the slope of the regression. This distinction has now been made clear in the text, which now reads:

> *However, the impact of irrigation and CA on the modelled spatially averaged temperatures improves the closeness to that of the observations, i.e. there is an overall decrease in absolute temperature (Figure 4a-d), which is consistent with current theory (Kueppers et al., 2007; Saeed et al., 2009; Kueppers and Snyder, 2012; Thiery et al., 2017, 2020; Hirsch et al., 2018).*

3. Some of the results are presented in such a way as to be somewhat misleading. For instance, the values in Figure 2 (% change in RMSE) with the colored categorization (which, being visual, is much stronger evidence to the reader than a table) can be compared with the equivalent anomaly in Table 1 (RMSE). For irrigation (IRR-CTL) the RMSE T2m (CRU) difference is -0.002, and the figure categorization is - 5-10%. For irrigation (IRR-CTL) the RMSE TXx GHCNDEX difference is +0.004, and the figure categorization is 0-5%. i.e. a difference in RMSE that is twice as big, is categorised as half the size in terms of color. This means that it looks as though the CA and IRR simulations are doing much better than if the simple RMSE is considered.

Reply. We agree with this point and are thus presenting the absolute change in the RMSE data (in K) in Figure 2 of the paper, as detailed in our Response Figure 3.



Response Figure 3 (replacing the paper's Figure 2). Added value of including irrigation and CA in the simulated warming trends over 1981-2010. Absolute change in spatial root-mean-square error (RMSE) for the (a) IRR and (b) CA ensemble relative to the CTL ensemble over different regions (x axis) and with respect to 3 observational products (y axis). Considered regions are the SREX regions where irrigation is extensive (as highlighted in Figure 1a) and where CA is extensive (Figure 1b), in addition to global land, global irrigated land and global CA land. Observational products are for near-surface air temperature T_{2m} (CRU), annual maximum daytime temperature TXx (GHCNDEX and HadEX2). The spatial RMSEs are computed for the ensemble mean warming trend in every pixel, and subsequently averaged over the selected region. Regions with an observational coverage below 50% are marked in white.

4. Some of the results are inconsistent with each other. For instance, on line 182, the range of temperature anomaly compared to observations is given as 0.007 - 0.03 in the text, but in Table 1 it is 0.007 – 0.024 (usually a number is rounded down when the last value is below 5). Or line 179 where the text says 0.004 for the Control, but Table 1 says 0.006. Figure 4b shows TXx HadEX2 on top of IRR and CNT much higher, but Table 1 shows CTL and IRR with differences from HadEX2 of 0.008 and 0.012 respectively.

Reply. The 0.007 - 0.03 in the text refers to the range of temperature anomaly for all three (CTL, IRR and CA) experiments, not just the CTL, which is consistent with the data in Table 1 (the 0.03 K yr^{-1} TXx bias is noted for the GHCNDEX observations and the IRR ensemble). This has now been clarified in the text, as follows:

On average, the CTL, IRR and CA ensembles overestimate TXx warming trends by ~0.007–0.03 K yr^{-1} over all land pixels. Over irrigated pixels, the CTL and IRR ensemble overestimate TXx by ~0.008–0.013 K yr^{-1}. Over CA pixels, the CTL and CA ensemble overestimate TXx by ~0.006–0.013 K yr^{-1}.

For the reviewer's point regarding line 179, indeed the text referring to CTL T2m bias should read according to the data in Table 1 – i.e. 0.006. In order to also address the reviewer's point 7 below, we have addressed this by not including the statement in the text that specifies the CTL result, but displayed this data in Table 1 only.

Regarding the reviewer's final point on Figure 4b, the Table 1 differences stated are consistent with Figure 4b. The slope of HadEX2 in Figure 4b is 0.026 K/yr, the slope of CTL is 0.034 K/yr and the slope of IRR is 0.038 K/yr, which renders a bias of 0.008 K/yr for CTL and 0.012 K/yr for IRR, as detailed in Table 1. To ensure it is clear that Table 1 presents the bias and RMSE of the slopes (and not any other temperature parameter), the caption has been edited to state:

Bias and Spatial RMSE of the Ensemble Mean Warming Trends (Slopes) of the CTL, IRR and CA Experiments Versus the Observational Products for the years 1981-2010[a].

5. The introduction does not do a good job of introducing the main point of the paper. The first paragraph sets up the issue that observations show less warming in TXx than T2m, but models get it the other way around (TXx warms more than T2m). But we basically don't hear about this issue again. Subsequent paragraphs in the introduction are brief summaries of key papers (by the authors) and do not provide the cohesive overview of each topic a reader needs in an introduction, instead being based around a particular reference.

Reply. Thank you for highlighting this area for improvement. The introduction has now been substantially reworked to read:

According to observational and global climate model (GCM) data, temperatures associated with hot extremes have increased consistent with global anthropogenic climate change (Sillmann and Croci-Maspoli, 2009; Donat et al., 2013a, 2013b; Hartmann et al., 2013; Pendergrass and Hartmann, 2014; Fischer and Knutti, 2015). However, hot spots of accelerated warming in annual maximum daytime temperature (TXx) relative to local mean temperature (T_{2m}) simulated by climate models from phase 5 of the World Climate Research Programme's (WCRP) Coupled Model Intercomparison Project (CMIP5) are spatially inconsistent with observations (Donat et al., 2017). This is particularly the case over southeast China, South America, north America and parts of Australia and Europe. In these regions, the modelled TXx warming from the midtwentieth century (1951–1980) to the late 20th/early 21st century (1981–2010) was greater than the modelled T_{2m} warming. In contrast to the models, the observations showed that TXx warmed at a slower rate than T_{2m}. Further analysis of the CMIP5 ensemble over central Europe by Vogel et al. (2018) highlighted that several GCMs overestimate the observed negative correlation between summer precipitation and TXx, resulting in too strong future drying and associated TXx increases under RCP8.5. This *underlines the importance of a correct representation of land-atmosphere coupling for simulating changes in temperature extremes at regional scales. These discrepancies between multiple GCMs and observations raise the questions as to whether: (1) these model results can be used to reliably project changes in local temperature extremes; (2) the discrepancies remain if the rates at which warming occurs over a time period is examined, which less sensitive to outliers common in extreme temperature data than the absolute temperature difference between two time periods, as used in the Donat et al. study; and (3) the inclusion of more processes that represent land-atmosphere coupling would enhance model skill.*

Agricultural land management techniques, including irrigation and conservation agriculture, can have a cooling effect on hot temperature extremes (Davin et al., 2014; Hirsch et al., 2017; Thiery et al., 2017, 2020; Hauser et al., 2019; Jia et al., 2019). Irrigation diverts surface and groundwater resources to agricultural land to increase crop production (Fereres and Soriano, 2007). The addition of this water to the land surface is balanced by the loss of water via runoff, deep percolation, soil storage and/or evapotranspiration (ET) (Fereres and Connor, 2004). Under drier conditions, less evaporative cooling leads to amplified warming because the energy budget becomes dominated by sensible heating instead of latent heating (Donat et al., 2017). If irrigation water is added to the surface, this increases soil moisture as well as latent heat flux over the summer months, leading to more evaporative cooling at the land surface. This irrigation-induced surface cooling, in turn, challenges the radiative forcing concept, which assumes that as radiative forcing increases (from enhanced atmospheric water vapour) so too does surface temperature (IPCC, 2001; Boucher et al., 2004).

Conservation agriculture (CA), which involves crop residue management, crop rotation (Carrer et al., 2018; Lombardozzi et al., 2018) and minimal or no tillage (Kassam et al., 2015), can create climate feedbacks due to the presence of a crop residue over CA land change both the radiative and hydrological properties at the surface (Davin et al., 2014). Hirsch et al. (2018) explored whether applying the no-till component of CA within the Community Earth System Model (CESM) improves the simulation of present-day climate. They found that the surface temperature response was influenced by three competing effects: (1) a surface albedo increase – which reduces the availability of energy for partitioning between the sensible and latent heat fluxes; (2) increased surface resistance (e.g. from mulch) – which reduces soil evaporation; and (3) increased soil moisture retention leading to enhanced transpiration. The local cooling response to CA was somewhat counteracted by grid-scale changes in climate over North America, Europe, and Asia because of negative atmospheric feedbacks. That is, the decrease in evapotranspiration (ET) – both due to higher albedo and higher soil resistance – appeared to activate a decrease in cloud cover in the model that increases incoming shortwave radiation and therefore temperature via enhanced sensible heating. Grid-scale changes in climate counteracting local responses to land use change has also been demonstrated by Malyshev et al. (2015) who showed that the subgrid signal of land use change in near surface temperature was diminished by the averaging with undisturbed portions of the grid cells. The importance of local-scale responses to land cover change has also been indicated in observation-based studies (e.g., Mahmood et al., 2013; Li et al., 2016), yet few global-scale modelling studies examine the local land surface response to land management (Paulot et al., 2018; Meier et al., 2018).

Using GCMs, such as CESM, to simulate land-atmosphere interactions for investigating the effects of irrigation and agricultural conversion has been criticized as insufficient (Niyogi et al., 2002). This is partly because their coarse resolution (e.g. of order 100 km) hampers their performances in describing the present-day climate at the regional scale (Jiang et al., 2016). Furthermore, economic, societal and water resource factors are ignored – a void that initiated the so-called 'bottom-up' approach to evaluating the effects of land-use change (Douglas et al., 2006). Regarding the applicability of the knowledge produced by GCMs, they do not provide the skill required at the spatial scale to offer practical responses at the infrastructure scale (Hossain et al., 2015). Despite these constraints, GCMs remain a prime tool for projecting changes in the climate system (Fajardo et al., 2020; Gupta et al., 2020; Hofer et al., 2020). Examples include the GCMs that are part of the latest Coupled Model Intercomparison Project (CMIP6) and used by the IPCC in consecutive assessment reports (Yazdandoost et al., 2020). However, these GCMs largely exclude agricultural management. In particular, no CMIP5 model incorporates irrigation or CA and only three CMIP6 models include irrigation, while none have CA. Pielke et al. (2011) suggested that landscape change is omitted from the CMIP5 models because the direct radiative impact of global landscape is a lower order than the radiative forcing from greenhouse gas emissions. This constitutes a reason to investigate their inclusion. That is, to distinguish between the effects of land management and other large-scale forcings such as a doubling of CO_2 (Schultz et al., 2016), it is important to evaluate these processes in the GCMs and ultimately gain insight into the contrasts of impacts between regions under different climate regimes.

Considering the potential effects of irrigation and CA on climate (Thiery et al., 2017), it is possible that the discrepancies between climate models and observations regarding temperature changes (Donat et al., 2017) are because the models exclude the effect of agricultural management techniques on temperature. The goal of this study is thus to test the hypothesis that CESM version 1.2.2 overestimates warming trends in some regions because irrigation and CA are excluded. That is, warming rates are hypothesized to increase at a slower rate – showing signs of cooling, in irrigation- and CA-affected regions when climate models do account for a theoretical constant level of these land management practices. To realise this goal, the following objectives were formulated: (1) Determine spatial warming rates using simulations that account for irrigation and CA and inspect whether CESM overestimates warming trends; (2) Compare the observed rates of warming to the modelled rates of warming for irrigated and CA pixels, as well as non-irrigated and non-CA pixels; and (3) Estimate the impact of irrigation on the spatial average of the warming rates over time for all land, selected regions, and irrigated and CA pixels."

6. The abstract doesn't do a good job of summarizing the paper. It goes straight to "we did things" without explaining why the reader should be interested, or what the relevance is, then goes to "it's important" without presenting the evidence of why it is important. It provides no context for what was done, then the emphasis on the "pulse cooling" in the abstract is not followed through in the results. The subgrid scale aspect seems to be the most novel part of the paper, but since the rationale for it isn't explained clearly in the methods or results, it's a minority of the results section, there's no comparison for the scale of the water vapor feedbacks, and no clear explanation of why the water vapor feedbacks are responsible for the differences between the grid scale and subgrid scale, it's not convincing.

Reply. Thank you for this suggestion, we now reworked the abstract. Due to the addition of new data on the latent and sensible heat fluxes (see response to reviewer's point 9 below), we now clarify in the abstract that an increase in the LHF is responsible for the differences between the grid scale and subgrid scale. That is, at the land surface, the positive radiative forcing signal is too weak to offset the local cooling from the irrigation-induced increase in the latent heat flux. The updated abstract is detailed under our response to the reviewer's Point 1. Also regarding this comment, in our abstract, introduction and results section, a rationale for the subgrid scale aspect is now provided. Specifically, because we are looking at irrigation or CA-induced impacts at the land surface, It is important to understand and quantify the effects of land management as such on local climate in order to distinguish between the effects of land management and other large-scale forcings such as a doubling of CO_2.

7. The results section has paragraphs where the numbers in Tables are repeated, (with some unexplained deviations, as discussed above) with little attempt to give a view of what the results mean for the model's performance. There is no point in having a table if the text repeats what it says, and vice versa. There is absolutely a place for straightforward analysis and simple statements, but it needs to enhance clarity, not just be a list.

Reply. The section where we feel this point was most relevant has been substantially condensed (i.e. manuscript Section 3.2), to read:

> *Neither irrigation nor CA has a cooling effect on T_{2m} and TXx warming rates in irrigated/CA or non-irrigated/CA regions (Figure 3 and Table 2). The results suggest a slight irrigation- and CA-induced acceleration of the annual T_{2m} and TXx warming trends, rather than the hypothesised cooling. For instance, irrigation induced an increased T_{2m} warming rate of 0.0023 K yr^{-1} on average over land and 0.004 K yr^{-1} across all irrigated pixels. To put these increases into context, the mean T_{2m} CRU observed warming trend over irrigated pixels was 0.029 K yr^{-1}.*

In addition, the text providing the ranges in Section 3.1 has been written more concisely, as described above under the reviewer' s Point 4.

8. The Tables and Figures all need more detail in the captions, to help explain what they are and why they differ (and why those differences were deemed necessary). For instance: Table 1 – which years go towards the values? Table 2 – what are the "impact of irrigation and CA on various climatological values", because 0.026 K yr-1 for T2m doesn't make any sense as a number for the Control if it's supposed to be IRR – CTL as described in the caption. Figure 4 – presumably when it says "average" it's the mean, but then line 378 says "median", which is a notably different average.

Reply. Thank you for this suggestion. The caption for Table 1 has been updated to resolve this point as well as the reviewer's point 4, to read:

> *Bias and Spatial RMSE of the Ensemble Mean Warming Trends (Slopes) of the CTL, IRR and CA Experiments Versus the Observational Products for the years 1981-2010[a].*

The caption for Table 2 has been updated to read:

> *Impact of Irrigation and CA on Various Climatological Values (Absolute Slope Differences Calculated as IRR Minus CTL and CA Minus CTL for Grid-Scale, IRR_{SUB} Minus RAIN and CA_{SUB} Minus CM for Subgrid-Scale) for the years 1981-2010[a].*

Regarding the spatial average (Figure 4 and 7), these figures show the mean temperature for all the pixels within each mask specified (i.e. all land, irrigated pixels or CA pixels) – plotted on the y-axis, for each year. The slope detailed in the figures was estimated using Sen's slope – so that all slope data used in this study is attained in the same way. As Sen's estimator takes the median slope (it was chosen as the nonparametric alternative to linear regression so that the slope is less sensitive to temperature outliers), the study conclusion thus notes (line 378):

> *Insight into how modelled temperature is affected in its median by irrigation and CA over time was provided.*

The captions for Figures 4 and 7 have been updated to clarify this. Figure 4 caption is:

> *Spatial average of the warming rates for T_{2m} (a, c and e) and TXx (b, d and f) for the CESM ensembles and observations. Data points specify the mean T_{2m} and TXx temperatures for irrigated pixels (a-b), CA pixels (c-d), and (e-f) all land pixels. The slope was estimated using Sen's slope for the CTL (red), IRR (blue), CA (cyan), CRU (purple), HadEX2 (yellow), and GHCNDEX (black) temperatures.*

The new caption for Figure 7 reads:

> *Average of the TS warming rates over (a) irrigated pixels for the irrigated and rainfed crop tiles; (b) CA pixels for the CA and CM crop tiles; (c) all pixels for the irrigated and rainfed crop tiles; and (d) all pixels for the CA and CM crop tiles. Spatial average of the SHF (e) and LHF (f) warming rates for the irrigated and rainfed crop tiles over irrigated pixels. Data points specify the mean TS, LHF and SHF values within the crop tiles and pixels specified. The slope was estimated using Sen's slope for the rainfed/CM (red), irrigated/CA (blue) experiments. For (a), (b), (e) and (f) the regions where less than 50% of the land pixels did not contain a value were excluded. For all land pixels (c and d), the minimum number of land pixels that needed to contain a value in order to be retained in the analysis was 15%.*

In addition, the caption for Figure 1 has been updated to include reasons why the boxes differ and help explain better the distinctions, to read:

> *Figure 1. (a) Percentage of each grid cell equipped for irrigation (%) (Siebert et al., 2005). (b) Potential estimate of CA extent mapped to the CLM crop PFT (Prestele et al., 2018). The red boxes in (a) denote the regional domains where irrigation is extensive and were thus examined in greater detail including Western North America (WNA), Central North America (CNA), south Europe and Mediterranean (MED), West Asia (WAS), South Asia (SAS), Southeast Asia (SEA), and East Asia (EAS). The red boxes in (b) denote the regional domains where CA is extensive and were thus examined in greater detail including WNA, CNA, MED, South-eastern South America (SSA), Central Europe (CEU) and Southern Australia (SAU).*

Figure 2 has been edited to read:

> *Figure 2. Added value of including irrigation and CA in the simulated warming trends over 1981-2010. Absolute change in spatial root-mean-square error (RMSE) for the (a) IRR and (b) CA ensemble relative to the CTL ensemble over different regions (x axis) and with respect to 3 observational products (y axis). Considered regions are the SREX regions where irrigation is extensive (as highlighted in Figure 1a) and where CA is extensive (Figure 1b), in addition to global land, global irrigated land and global CA land. Observational products are for near-surface air temperature T_{2m} (CRU), annual maximum daytime temperature TXx (GHCNDEX and HadEX2). The spatial RMSEs are computed for the ensemble mean warming trend in every pixel, and subsequently averaged over the selected region. Regions with an observational coverage below 50% are marked in white.*

And Figure 6 now reads:

> *Figure 6. Subgrid-scale differences between the irrigated and rainfed crop tile in the IRR ensemble (irrigated minus rainfed) (a, g, k and o) and between CA and conventionally managed (CM) crops (CA minus CM) (e, h, l and p). For T_s (a-b), ET (g-h), LHF (k-l) and SHF (o-p). Grid-scale differences between the CTL and IRR ensemble (IRR minus CTL) (c, e, I, m and q) and between CA and conventionally managed (CM) crops (CA minus CM) (b, f, h, n and r). For T_s (c-d), ET (i-j), TMQ (e-f), LHF (m-n) and SHF (q-r), displayed over irrigated/CA pixels for comparative purposes. Differences are based on the ensemble mean warming trends of each experiment for 1981–2010. Hatching denotes less than 10% change induced by the model on mean warming trends of lumped ensemble members.*

9. The results section has times where it would benefit from showing more evidence. For instance, the two paragraphs starting line 228 speculate that latent and sensible heat partitioning and changes in ET are responsible for the differences between CA and IRR, and the Control. But instead of exploring these and showing how the latent heat changes, there is just references to previous papers. I.e. it is not a result, it is a summary of previously published research. Similarly, the paragraphs line 289 – 315 contain a lot of speculation and references and not enough evidence.

Reply. New sets of results have been added that provide evidence on the changes in the latent heat flux (LHF) and sensible heat flux (SHF). Trend data on the grid- and subgrid-scale LHF and SHF has been provided in Table 2. Figure 6 now shows a comparison between the subgrid and grid scale LHF and SHF changes for irrigated and CA land, while the new Figure 7 details the spatial average of LHF and SHF on irrigated pixels (for the figure additions, please see above under the Reviewer's Point 1). This has also helped to address the reviewer's Point 10 below regarding the discussion section below and the inclusion of a reflection on uncertainties in the partitioning of latent and sensible heat when albedo changes.

10. The discussion section is missing, at the very least: cloud uncertainties (different models do them differently, and they are notoriously difficult to resolve, so that these results rely on them is problematic); uncertainties in the partitioning of latent and sensible heat when albedo changes (i.e. Bowen ratio; the fact that the CA increase in albedo is a huge assumption, as soil albedo is very heterogeneous and dependent partly on soil moisture (thus the CA modeled might be doing the wrong thing in many areas); the representation of transpiration in the model, and the fact that presumably the crop/vegetation cover is the same when in reality these changes would affect the LAI of the crop; the canopy interception and soil interception representation in the model, which affects the evaporation and thus how much the irrigation and CA affect the evaporation.

Reply. The discussion sections has been rewritten to address these points as follows:

> *This study examined the hypothesis of whether excluding a theoretical constant level of irrigation and CA contributes to the overestimation of warming by an Earth System Model. A Sen's slope model was built and applied to ensemble simulations from the Community Earth System Model that include irrigation parameterization to determine if there are spatiotemporal patterns and why they exist. This unexpectedly showed that warming trends are not dampened due to irrigation and CA, except for the subgrid-scale effect of irrigation on the warming trends of T_S.*
>
> *The key findings of this investigation are a net cooling effect of irrigation and CA on the modelled spatially averaged T_{2m} and TXx, but, rather than continuous cooling, the warming trends showed a pulse cooling phase, after which the sensitivity to climatic change remains. Under irrigation, the opposing effects are the result of: (1) evaporative cooling; and (2) atmospheric water vapour strengthening the greenhouse effect. Under CA, the contrasting effects are due to: (1) cooling from a tillage-induced increase in surface albedo; and (2) reduced soil evaporation due to the presence of crop residue, limiting energy partitioning to the latent heat flux. At the subgrid-scale, there was both a cooling effect on T_S and in the dampening of warming trends. This implies that enhanced evaporative cooling is the dominant driver of the subgrid-scale temperature trends.*

Although this study was constructed with great care and built on a state-of-the-art modelling suite, several future developments could improve understanding of the impact of irrigation and CA on climate. Firstly, the quality of the model(s) could be improved by using transient irrigation and CA extents and new land cover datasets from the 6th phase of the Coupled Model Intercomparison Project (CMIP6) (Lawrence et al., 2016). In this study, a static irrigation map for the year 2000 was used for the whole simulation period. This likely contributes to our results being conservative. If, for instance, irrigation expands over time, the cooling effect may become stronger and thus affect the warming trends. Furthermore, the extent to which the increase in surface albedo (i.e. the first competing effect of CA) affects the sensible and latent heat fluxes partly depends on soil moisture, which too is not static. Also, CMIP6 experiments are based on annual emissions, whereas CMIP5 was based on decadal emissions and CMIP6 models were updated with irrigation-related features and land cover maps that incorporate irrigation and CA expansion over time (Goddard et al., 2013; Miao et al., 2014; Boer et al., 2016; Meinshausen et al., 2017; Stouffer et al., 2017). CMIP6 models may therefore improve the dynamics between irrigation, CA and climate change, provided that they represent these land management techniques in their surface schemes.

The second consideration is that all simulations used in this study (5 control, 5 irrigation and 5 CA) were from a single model. Ensembles completed as such with the same model but different simulations (i.e. based on different initial conditions) characterise the uncertainty associated with internal climate variability only, while multi-model ensembles also account for the impact of model differences (Tebaldi and Knutti, 2007; Knutti et al., 2010). This limitation can impact cloud uncertainties. Hirsch et al. (2017) found that the CESM tends to produce large cloud feedbacks over Central Europe, Central North America, North Asia, and South Asia when more energy is reflected at the surface. Irrigation-induced increases in latent heat fluxes led to more water vapor in the lower atmosphere, which generated low-level clouds (see also Sherwood et al., 2017). This limited shortwave radiation and hence the amount of energy available at the surface because the increased cloud cover reflected more downward shortwave radiation above the cloud layer, resulting in surface cooling. This was enhanced by a corresponding decrease in sensible heat fluxes, reflecting the decrease in the amount of energy available at the surface and/or the increase in latent heating. The impact of cloud cover combined with land management change remains challenging to resolve. Therefore, this study should ideally be repeated with other models. Donat et al. (2017), for instance, conducted their study on 20 CMIP5 models, but these models did not incorporate irrigation and CA.

Thirdly, irrigation and CA are the only agricultural management practices considered in this study (and done so individually), whereas other agricultural management practices have been shown as impactful (Luyssaert et al., 2014; Erb et al., 2016, 2018). Trend analysis of integrated land management practices could affect the outcome if there is a lumped effect. Building an additional stochastic model could account for variations in the distribution of the impact of land management practices on warming trends. This would enable sensitivity analyses to ascertain the relative importance of irrigation and CA to the total warming trends (based on all land management practices), as well as the relative contributions of the uncertainty sources (model input, parameter, structure) to the total uncertainty in the model output.

The final consideration is whether regression-based models are suitable for analysing changes in highly variable climate data, particularly annual extreme temperature data (von Storch, 2006). Essentially, the regression slope blends forced temperature change and variability, to provide an estimation of the temperature variation over time – within which variance can be lost due to noisy data. Whether the TXx and T_{2m} temperatures were first spatially averaged and then the slope retrieved or if each slope was estimated for each gridcell and then the overall trends examined, the outcome remains. This is unsurprising considering that in the spatial averaging the noise contributions are averaged out, while the individual regression data suffers from the variance loss related to regression. However, when applied to over 60 years of observational data, the regression model used in this study showed similar trends to using the difference between the past and the present average temperatures (not shown). This implies that the irrigation and CA-inclusive climate system may require a longer timeframe (than the 30 years plus a 5-year spin-up period used) for trends to overtake the natural variability. Additionally, rather than aggregating all months, trends during individual months or seasons could be examined. This can affect, for instance, the influence of irrigation on T_s, which has a clear seasonal pattern, with more cooling during the driest and/or hottest months (Thiery et al., 2017). A smaller magnitude in TXx response to CA at the subgrid-scale has also been noted during the summer season due to a larger leaf area index (LAI) reducing soil surface exposure and thus the contrast between CA and conventionally managed crops (Hirsch et al., 2018). Furthermore, the implementation of CA within CESM does not capture crop planting and harvesting cycling (Davin et al., 2014), which would affect the LAI of the crop and potentially the effect of CA on surface climate.

Thank you for your time and effort in helping to improve our paper. It is great appreciated.