Dear editor,

We thank the reviewer for the valuable comments. We have addressed the comments in bullet points below. Please note that the lines mentioned are referring to the track-changed document.

Best wishes, and on behalf of the co-authors,

Ileen Streefkerk

This text has certainly clarified why certain modelling choices were made. Thanks for that! I would suggest bringing some elements of the author's response in the main text – especially on the PMT on page 7 – but the model setup is clear now.

- Good to hear the modelling choices are more clear now. The following text on the PMT is now added into the manuscript. Lines 158-162:
  "PMT is a common psychological theory to model people's intentions to adapt that has been applied in many ABM studies on natural hazard management (Heliegiorgis et al. 2018; Wens et al. 2020; Michaelis et al. 2020; Moradzadeh and Ahmadi, 2024). Recent household survey studies with farmers and pastoralists in East-Africa show that show that PMT is a suitable theory to explain adaptation behavior under drought risk conditions (Wens et al. 2021; Schrieks et al. 2024; Gebrehiwot and Van der Veen, 2021)."

The text describes how a deterministic model – as inputs and responses are prescribed in the model setup – yields results that could be meaningful because of the many runs with slightly different parameter sets. That approach is adequate.

- Thank you for the positive acknowledgment of our modeling approach. Indeed, while the model is deterministic in structure, we employed multiple runs with varied initialization sets to explore the randomness of the system and to provide a more robust understanding of the range of possible outcomes. We appreciate the reviewer's recognition of the adequacy and relevance of this methodology.

Both discussion and conclusion are much more specific and as such better suited. I would suggest starting the discussion with 5.3, followed by 5.2 and 5.1 – as this zooms out from model results to implications. Now, the discussion reads like "we have impacts (5.1), but do not believe them (5.3)".

- Thank you for your suggestions. We agree and in the revised manuscript we now start with the uncertainties of the model and then go into results implications.

My main concern with the current text is that the model results are presented as distinct enough to be separated and to be seen as relevant, even when the quantitative model results seem to be hardly different (2% change is not a change, given the model uncertainty, I would argue) and the difference between scenarios is low. What we see is that drought has an impact? I am not sure all the modelling efforts are required to claim that. However, as soon as we want to know where the impact will be, modelling results that show such extremely small changes between model settings may be inadequate. I may have missed something very important on the results, which is why I share more details on my doubts below.

- We understand your concern. We have addressed the specific elements below.

The model fit may be relatively high with correlations of 0.7 and the like – but again, that leaves plenty of space for the minor changes that are modelled to be random. Modelled milk production remains interestingly different from the observed production. The model underestimates milk production in dry periods – which suggests again that the drought effects that the model produces may not be very useful.

- We acknowledge that, despite the relatively high correlation coefficients (around 0.7), discrepancies between modelled and observed milk production remain, particularly the underestimation during dry periods. This issue indeed suggests limitations in the model's ability to fully capture the complexity of drought impacts on livestock productivity and milk production. As noted, these discrepancies may arise from factors not explicitly represented in the model. We have added a more detailed discussion of these limitations and their implications for interpreting the drought-related results in the revised manuscript (see Section 5.1, lines 501-507).

On page 14, we are confronted with some results. Fair enough, stream flow seems to show something, but I find it hard to see why soil moisture of 0.255 and 0.257 can be seen as distinct, or why a groundwater level of 65.98 m versus 66.17 m is distinct. Distances between household and water show a 0.2 km variation on a 3 km distance – again, not distinct. Figure 5 is mainly an illustration of this observation - and the scales are selected to show at least some differences (which I would advise against). The authors are well aware of this issue, as they write on page 14 (lines 356-358) "While […] periods." If we know that, why still defend the outcomes as meaningful? We encounter the same caution on page 17, line 420-421 – even when "relatively low" should perhaps become "negligible"?

- The specific examples cited indeed reflect minimal variation at the aggregate level, which may not appear distinct or impactful. We appreciate the reviewer's observation that the visualization in Figure 5, though technically accurate, may unintentionally emphasize these small differences due to the chosen scale (aggregated outputs - averaged over the entire area and across all agents) has limitations in conveying meaningful variability, particularly when the resulting differences are numerically small. We have revised Figure 5 by separating the drought hazard and impact by climate zone (representing up/downstream and livelihood differences). We have adjusted the scale to avoid any potential overstatement, and we now present a clearer narrative about the limitations of interpreting such small variations.

- With regard to our previous use of terms like "relatively low," we agree that the phrasing could be sharpened. Where appropriate, we have revised such descriptions to reflect the limited magnitude or practical insignificance of the differences observed using terms like "minimal" or "negligible" where justified.

Page 20, lines 505-506 suggest that the model shows "the importance of heterogeneity and individual characteristics of agents", but I find it tricky to link this (useful claim in itself) with the averaged, lumped model setup when it comes to agents. The modelling results are based on lumped agents on patches, as explained on pages 5 and 6 – which must mean that the minor differences between the model scenarios become even less meaningful.

- Thank you for your valuable comment, we made some major changes to the manuscript to address this. To show the heterogeneity at agent-level dynamics, in the revised manuscript (see Section 4.2.3, lines 469–485, Figure 9), we disaggregate selected indicators by climate zone and show the variety among agents. This additional analysis reveals more nuanced patterns that are otherwise masked in the averages and has strengthened our understanding of the variety in impacts across the agent population, differentiated by climate zone. We have reflected on these results in the discussion in lines 584 to 590.

The major result that seems to be modelled is that more commercial farming upstream results in lower groundwater tables – as in the effect seems to be one meter difference between scenarios with and without commercial farming (figure 6)? If the absolute groundwater levels are the 66 meter mentioned earlier, we are talking about 1/66 = 1.5% difference? This may be one of the higher changes in terms of percentages, but still not terribly high?

- These results were averaged over the four sub-catchments, while the results on lower groundwater tables is most clearly seen in Figure 7. There, up to 7 meters of difference are observed at location of the commercial farms, and 1 meter downstream. In the revised version we have clarified that the groundwater changes are not primarily downstream, but mainly at the location of the farms itself. We agree that we should indeed make it more clear here what the reference is. In the revised version we are referring to Figure 5 as a reference (see lines 425-427). It can be observed in the revised Figure 5 that groundwater depths are 50 meter in the upper climate zone where the commercial farms are located (although around 25 m at the commercial farms locations itself).

Do we need to read this paper to learn how we can model these processes (the key novelty mentioned on page 2), or do we need to read this paper to discover the implications of certain events and/or interventions (the effect of commercial farming, as mentioned on page 2)? The first reason might not require this paper (as the model setup has been published elsewhere), the second reason would probably need much more explanation why such minor differences between scenarios are meaningful.

- Mostly the second aim, as we have indeed explained the model in a previous paper. However, there was no performance test involved in that paper, so we included those results in this paper. On second aim; we believe that with the extra/adjusted plots we can bring more meaningful and nuanced findings and discussions to the paper.