RC1

Thank you for inviting me to comment on the manuscript: "Adaptive Behavior of Over a Million Individual Farmers Under Consecutive Droughts: A Large-Scale Agent-Based Modeling Analysis in the Bhima Basin, India". I am an expert on agent-based land use change models. This may limit my expetise on the hydrological aspects of the presented work.

The authors use two coupled models (one ABM and one hydrological model) GEB (please spell out what GEB stands for) to model the land use of presumably 1.4 million farmers in the face of consecutive droughts over several decades. Unfortunately, the text is very rich combing description of drought modelling, theory, ABM model, and many more things. Therefore, at least for me it is impossible to understand the model(s) in detail and to put the results into context. Therefore, I am not able to appreciate the model and its results sufficiently, although the topic is timely and I guess the approach is relevant and promising.

Maybe the use of a protocol such as ODD+D (Mueller et al. 2013, Describing human decisions in agent-based models - ODD+D, an extension of the ODD protocol) could help to present the model in a more digestible way. At least the authors should summarize somewhere (maybe as a table) an overview of the properties of the agents and or agent types.

Thank you for the suggestion, this is indeed helpful as it was difficult to both cover all aspects of the model and keep a concise paper. An ODD+D has been added to the supplementary information that we hope gives better insight into the full working of the model. If you are still missing more technical insights, the paper by De Bruijn et al. (2023) discusses the technical base design of the Geographic Environmental and Behavioural (GEB) model.

The authors present many responses from their model: "Our analysis examines how these adaptations affect profits, yields, and groundwater levels, considering, e.g., farm size, risk aversion and drought perception." Maybe it would be helpful to reduce the number of responses and/or scenarios? Especially in Figures 5 and 7 I would encourage the authors to present less panels and to focus on a more narrow narrative.

We agree that some panels had unnecessary information and that 8 panels is too much, thus we reduced figure 5 to 3 panels and figure 7 to 6 and removed/trimmed paragraphs in the description of the results. Furthermore, we have moved the sensitivity analysis to the appendix.

However, our main finding is that the combination of hydrological and socio-economical factors steer the area and especially certain groups of farmers into a more vulnerable state. For this, combinations of the results of yield, crop choices, income, groundwater and wells are needed, as the interplay between those leads to this result. We also find that being able to model the interplay of these factors is the major strength of this (new) type of model.

I appreciate that the authors follow a theory to justify their decision model. However, the authors need to guide the reader more carefully, since the SEUT is not a standard theory for all Economists and certainly not for all land use change modellers. Fishburn (1981) is a review of several theories and the list of papers suggested as examples of the application of SEUT needs to be critically revised (Groeneveld (review), Haer and Wens do not mention SEUT and do not cite Fishburn). When SEUT is introduced it also should be mentioned that the authors use imitation and "bounded rationality" (line 215) as well in their decision modelling. Later on also prospect theory is considered.

Thank you for your sharp comment. Wens (2020) compares the "economically rational" EUT with the PMT, and concludes that a more bounded rational theory covers behavior better. This was originally used as justification of using the SEUT over the EUT, but this has been lost over subsequent versions. This happened for Groeneveld as well. Haer (2020), however, does not refer to it as the SEUT, but they use the exact same theory (and calculation) of the expected utility for their boundedly rational agents in effect, where the EUT is altered by the changed perception of probabilities due to having experienced an event. The references have been changed and the justification has been expanded in the main text (line 145 to 155) and added in the ODD+D protocol (section 2.1.3).

In general, I am missing an argument why it is useful to model > 1 million agents. Other authors decided to gain knowledge by aggregating actors to agent types in their land use models (e.g. https://landchange.imk-ifu.kit.edu/CRAFTY or the work by Millington et al. https://www.jasss.org/11/4/4.html). It would be great to see an argument developed why this computational demanding approach is seen more appropriate to answer questions of land use change. This is especially critical since the authors argue that it is not computationally feasible to compute for all agents the SEUT for all 300 options ("unique crop rotations"). Would it make sense to compute less agents and therefore consider all 300 options?

There are several reasons why we decided to not aggregate agents (meaning for each farmer in the basin, we also simulate one agent, what we call "one-to-one"). First and foremost, we do not know what a representative agent for our area is (Page, 2012) and by pre-emptively aggregating agents, we may lose interactions that we were not aware existed in the first place (Page, 2012). This is especially true for an area so heterogeneous as the Bhima basin in India, where there are extreme differences in landholder size (Desai et al., 2008) , which factor through in other agent attributes such as which crops they initially cultivate (Department of Agriculture & Farmers Welfare India, 2001), their access to credit or their social factors (Hoda & Terway, 2015; Maertens et al., 2014; Udmale et al., 2015). For example, if we were to aggregate to one agent per grid cell, we would already lose out on the process where larger agents have more funds to invest within similar budget constraints, and tragedies of the commons, where larger farmers extract more groundwater and adjacent smaller farmers are unable to access the deepening groundwater. Instead of one grid cell per agent, we could attempt to scale up all farmers, adaptation costs, etc. and reduce the total number of agents, but this would require many parameter scaling adjustments, and it is unknown if model processes and interactions (including those with the hydrology, for which spatial factors are generally quite important) would remain similar. We agree however that finding what results the model would produce with different levels of aggregation would be a very interesting future study, and we thus added it to recommendations (lines 557-559). However, for such research to be possible, we require the development of these large and efficient models that are able to simulate at this detail and scale in the first place.

The second reason for keeping more agents while arguing that it is inefficient to calculate 300 options for one agent, is that more agents do not scale linearly with computational times in GEB. Due to the high degree of vectorization in the model, many agents doing one operation can be simulated much more efficiently than fewer agents doing many operations (i.e., 1.5 million agents doing 1 action is substantially faster than 5000 agents doing 300 actions). Additionally, as we always need to simulate the full region's hydrology, fewer agents may not bring about the same computational advantages as with non-coupled ABMS. We mention that we do not aggregate agents (lines 111-112) and added additional clarification with regards to why we chose to not aggregate to the initialization section (lines 287 to 290) and to the ODD+D protocol (section 2.7.1).

I have difficulties to understand the results. To my understanding imitating the strategy of more successful agents in the neighbourhood of an agent is at the heart of the presented dynamics. In the abstract this is not mentioned: "n adaptive scenarios, farmers can either do nothing, switch crops, or dig wells, based on each action's expected utility." If my reading is correct the imitation aspect should be mentioned early on and should be discussed in a diffusion of strategy/technology context. What are the updating rules – synchronous or asynchronously? How many neighbours are considered? What is the initial trait distribution of actors. Is there a spatial structure in the inital trait distribution? Is the number of farmers constant over the years? Form Figure 4 it seems that the model does not show much variation between runs. Would have been less agents sufficient? What source/help of A.I. have the authors used for what?

Imitation in combination with calculating the utility using the SEUT is at the heart of crop switching: agents compare the expected utility of a selection of their neighbors' crop rotations and choose the rotation with the highest utility. For choosing whether to dig a well, agents do not look at their direct neighbors. However, for calculating the expected utility of doing nothing or digging a well, it is assumed agents know the "true" added value of a well. Since wells both reduce damages during drought years as well as structurally increase water availability, which affects agents differently based on, e.g., the precipitation they receive, the crops they cultivate, etc. the “true” added value of wells is difficult to empirically predict beforehand. In other models this is often given empirically as, for example, a depth-damage curve in combination with flood maps (flooding). However, we use the agents that already had wells in the model itself to determine this added value, as we believe this gives a more accurate representation given the large differences in contexts for the agents. Therefore we don’t necessarily see this as imitation of other agents, but just as a way to determine the objective added value of wells given a farmer with x crop rotation at x location. We have added more explicitly the difference between the crop switching and well adaptation to the text (161 to 171). Furthermore, the additional factors you have requested are all in the ODD+D protocol (updating rules: 1.3.1; neighbors considered: 1.3.1, 2.4.3, 2.6.3, 2.7.1; initial trait distribution: 2.1.4., added plots of distribution of initialized personal agent factors; spatial structure: 2.1.4., 2.9.1, number of farmers: line 280, 2.1.4, 3.2.1).

There is some stochasticity in the model, which was the original reason for the 60 runs (section 2.9.1 ODD+D) . It does seem this effect is rather low, and could potentially be left out for future studies. This did not have major effects on run times, as these could be run in parallel on a linux supercomputer cluster. Run times were about 20-25 hours for a parallel run and about 15-20 hours for a solo run (without spin-up).

The use of A.I. has been further clarified (lines 703-704).

The role of the “spin-up” period (21 years) needs to be explained in more detail. The model is initialized with data from some point in time (when). Given the substantial temporal dynamics of the responses (see Figure 4) the choice of the length of the “spin up” period should have a strong effect on the results? It is written that the calibration has been done in the period after the spin up from 2001 to 2010? It is difficult for me to understand the evolution of Figure 4. The starting point at 2001 is the result of the spin up period? The period from 2001 to 2010 is calibrated and after that it is the model projection?

We have further explained and addressed these questions about the roles and differences of the spin-up and run (line 327-341).

A full model run consists of a "spin-up” from 1980 to 2001, and a “run” from 2001 to 2015. The spin-up period serves to set-up accurate hydrological stocks in the rivers, reservoirs, groundwater etc., and to establish enough data points for the drought probability – yield relation. At the end of the spin-up, the model state is saved and used as starting point of the run. The start of the run in 2001 was chosen as both the IHDS (Desai et al., 2008) and agricultural census (Department of Agriculture & Farmers Welfare India, 2001) collected data in 2001. As the climate data was available from 1979-2016, the 12-month SPEI was available from 1980. Thus, the spin-up time between 1980 and 2001 was chosen to maximize the duration so that the drought probability-yield relation (the “objective drought risk experience") included as many drought events as possible. Adaptation only occurs during the run. Two scenarios were run: one without adaptation, where agents maintained the same crop rotation and irrigation status as at the start of the model, and another where agents could change their crops or dig wells according to the decision rules outlined in section 2.3. Both scenarios use the same spin-up data.

The model was calibrated between 2001-2010 as we only had discharge and yield data during these years. We would have preferred to calibrate for the full climate data range (meaning until 2016) if the data were available, as the goal of this study was not prediction, but explanation about adaptation and risk under these consecutive droughts.

How are small, medium, and large field farmers defined in terms of hectare?

The ha cutoffs for small/medium and large farms were mentioned in the first section of results, but are now repeated in the figure descriptions for clarity.

Thus, overall I have the impression that potentially great insights are hidden in the current text. More specific and potentially less research questions could help to narrow down the story to allow easier access to the main highlights of the study. And at least for me it would be necessary to have a clearer motivation why it is beneficial to consider so many agents at the very same time.

Specific comments

Abstract: "realistically simulate" – That is maybe personal but I would avoid phrases like "realistically simulate" since it is a model and the best one can do is to model something useful in respect to the research question.

Agreed, it is still far off to be considered truly realistic. Removed the adverb.

The models are written in Python?

Yes, added it in line 104.

Lines 73/74 What means "one-to-one scale"?

Explained what the authors and De Bruijn et al. (2023) refer to as "one-to-one" scale: for every farmer in real life we have a representative agent.

Lines 78: What "simple assumptions of human behaviour"?

Expanded slightly on how behavior was represented before. Lines 77-80.

Figure 1: it is not clear that some boxes are empty – please explain.

This was done to signify the simplification. They are now removed. Line 102

Line 101: reservoir operators – is this agent type considered in the presented study? "reservoir operators" are never mentioned again.

This agent is present but not changed in this study. The ODD+D protocol discusses how the reservoir operators determine downstream and irrigation release.

Figure 2: The land cover map presents less classes (e.g. agricultural land) than used in the results? How did the authors discriminate the different crop types?

Section 2.5 in the main text describes how farmer agents are initialized. These are divided over the land use class "agricultural land". Each farmer controls their own hydrological response unit (HRU, added to section 1.2.4 of ODD+D) (De Bruijn et al., 2023), on which they can make their own land management decision. Thus each agent has their own plot of land where they have their own crop rotation (including the crop's specific characteristics and the implications of these for soil processes, evaporation, crop planting dates etc.).

Lines 132ff – Can you define in term of your indicator (SPEI) what a severe, moderate and so on drought is? I guess there are thresholds? Please specify.

Added the thresholds of the relevant categories of McKee et al. (1993) to the main text. Lines 138-140

Line 156: Why 5 km radius – is this decision based on a sensitivity analysis?

The 5 km was an error, it was actually 1 km, this is changed everywhere now. This was not based on any particular study, but on a general estimate on how far social networks go. A sensitivity analysis on more "meta" model settings (start date, radius, network size, etc.) would indeed be useful for future studies.

Line 165: C_adapt is not considered in equation 4. Does it needs to read "C_input in eq. 4 on current market prices"?

For clarity, we have changed C_adapt to C_well. The C_well signifies the costs of a well. The C_input signifies the costs of agricultural inputs (i.e. cumulative costs of seeds, fertilizers, etc.). Both signify the costs part of the equation, but indeed the line 165 was unclear, so I have added C_input there. (lines 182-184)

Equations 1-4 please explain subscripts x, d, and m.

Added the description in the text. Lines 182 to 187.

Lines 197: crop costs – crop costs C_input are dependent on the type of crop? If so how?

Yes, they are dependent on the farmers' crop type. In the preprocessing of the model, all cultivation costs are sourced from Ministry of Agriculture and Farmers Welfare in Rupees (Rs) per hectare. (https://eands.dacnet. Nic.in/Cost_of_Cultivation.htm, last access: 15 July 2022) for the full model run time. During planting, the crop that the agent is planting is taken from the crop rotation, the crop parameters are set in the model, and the cost at that specific moment in time is sourced from this premade dictionary of cultivation costs.

Equation 9: Where is the “crop coefficient” Kc used? The duration of different harvesting stages are not crop-specific?

All crop factors are crop specific, added clarification. The Kc is used to determine the crop-specific potential evapotranspiration from a reference evapotranspiration. Added this to the main text as well.

“Agent initialization”: Spell out IHDS, the authors should give the ranges of the agent properties. Showing boxplots or other useful representation of the distributions of attributes would be useful. I have no intuition for example how net income is initially distributed among the 1.4 million agents and how it is developing over time.

Spelled out IHDS and added boxplots to the ODD+D in section 2.1.4.

Line 326: (7g)?

7g highlighted the crop market price evolution. It is now in the appendix, as it is not a model produced result, but is relevant for several figures in the results.

Figure 4c: Is it reasonable that one crop is going from 0.05 fraction to the dominating crop?

It is if we look at the decision rules in the model in combination with the strongly risen prices of Groundnut. However, it is not if you compare it to reality. We discuss this in the discussion, and give some explanation and recommendations to resolve it. (Lines 540 to 566). Additionally, we now added some initial argumentation in the relevant results section (lines 359-362).

Figure 5: Very busy graph (6 panels). The authors may want to focus on some panels.

Reduced the figure to 4 panels.

Figure 6 and elsewhere: Specify the unit Rs

This is the Indian currency Rupees. Added the specification.

Figure 8: Too many panels. Legend unreadable. Please focus on the important aspects.

Reduced the panel to 6 panels and increased the font size of legends on each figure.

Technical corrections

The figure labels are too small and therefore hard to read.

Increased the size of all legend/x/y/ labels and titles

Figure 4c and others: colours are too similar – hard to differ crops

The Viridis color palette was chosen to accommodate colorblind individuals. However, it is true that a continuous color palette was not the right choice for categorical data. We now use the OKABEITO color palette (Okabe & Ito, 2002), which should still be colorblind friendly, but better suited to categorical data.

Lines 65-66 and elsewhere: Inconsistent in-text citation of Udmale et al. 2014/2015

Adjusted the references.

Line 114 and elsewhere: "95 %" -> "95%"

Replaced the instances.

Line 263: The authors refer to figure 3 in Jun et al. 2014? Jun et al. 2014 is a comment in Nature without Figures to my understanding. Please check.

Thank you for spotting the mistake! We intended to refer to Jun et al. 2014 to show which agricultural land data we used, and to refer to figure 2 (not 3 as we mistakenly did!) of our paper to show a visualization of the agricultural class in our study area. However, we see that this was quite unclear. We now clarified the text (Lines 318-319).

"Sensitivity Analysis": What are 300 distinct samples. Sampling from what distribution?

The latin.sample function from SALib uses Latin hypercube sampling. In Appendix B.4 is is further explained from what this was sampled. Lines 649-667

Line 306: Where does stochasticity enters the model(s) and how?

We added a description of the stochasticity in section 2.9 of the ODD+D protocol

Citation: https://doi.org/10.5194/egusphere-2024-1588-RC1

RC2

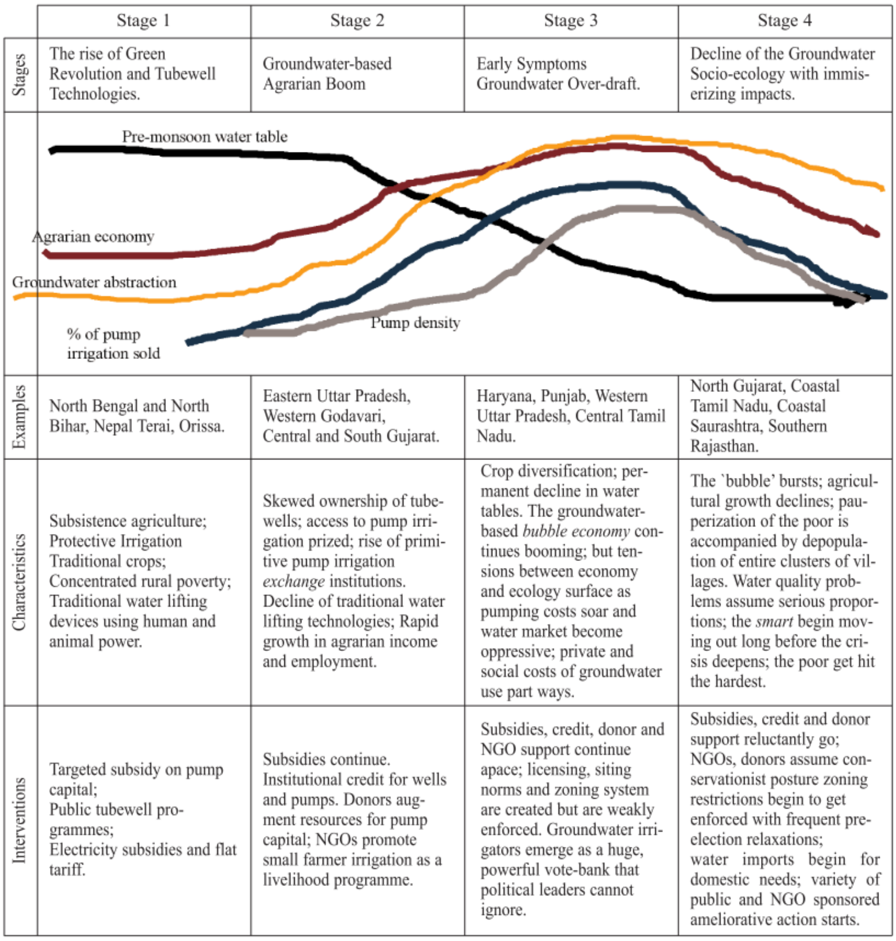
Thank you for the invitation to review this manuscript. In this work, the authors extend the GEB, a coupled agent-based hydrological model, with the Subjective Expected Utility Theory and apply the model for analysis of the Bhima River basin in India under consecutive droughts. The manuscript is impressive for the complexity of model integration and the breadth of analysis conducted. I especially commend the authors for the extensive sensitivity analysis that is conducted using the model, which is often a critical gap of coupled human-water systems analyses. However, the extensiveness of the manuscript is a double-edged sword, with the manuscript very challenging to wade through given the sheer amount of material (as reviewer #1 also noted). In this sense, I reiterate reviewer #1's comments in regards focusing the analysis. I have additional comments in regards to the manuscript:

Thank you for your positive words and overall constructive feedback! The extensiveness has indeed also been referred to by reviewer #1. Therefore, we have reduced the number of panels in figures 5 and 7, removed several paragraphs that did not contribute as much to the discussion points, and trimmed the remaining paragraphs to have a more focused narrative.

However, reiterating the response to reviewer #1, our main finding is that the combination of hydrological and socio-economic factors steer the area and especially certain groups of farmers into a more vulnerable state. For this, combinations of the results of yield, crop choices, income, groundwater and wells are needed, as the interplay between those leads to this result. We also find that being able to model the interplay of these factors is a major strength of this type of model.

1. My first and foremost comment is that the authors should demonstrate the validity and reasonability of the model in relation to real-world observation / understanding. While I understand that a full-scale, spatiotemporal validation of the model isn't likely possible given the sparsity of real-world observations and the complexity of the model, one can still ask the question: does the model better capture real-world patterns of the complex system in comparison to alternative approaches (e.g., the no adaptation alternative). For example, model results indicate that there is a very significant uptake in groundwater wells for large farms (growing from 30 percent of farms to 65 percent of farms) over the course of the model run. Is there any real-world quantitative or qualitative data that supports these model results? The onus in this case would be demonstrating that the adaptive model outperforms the non-adaptive model in replicating these large-scale patterns observed in reality. Similarly, do we in reality see the significant increases in groundwater depletion associated with the adaptive behavior (~10 meters in relation to the non-adaptive version); I would imagine that even apart from point groundwater level measurements, such a stark difference in depletion could be corroborated by GRACE, or *even other qualitative sources. Cropping patterns are another example, the adaptive model shows large-scale crop switching that could likely be corroborated, in a broad scale sense, via agricultural census information or remote sensing data. While the modeling integration and advances are impressive, there are so many choices that are made in regards to theory and implementation (as is the case with nearly all coupled human- natural models), that it becomes nearly impossible to assess the value of these model improvements in the absence of such evaluation.*

This is indeed a valuable suggestion and in the revised version of our paper, we have added a new paragraph to the discussion section where we explicitly verify the modeled trends (e.g., uptake of wells) with literature (lines 564-584). This is mainly centered around the findings of Roy & Shah (2002), which describe multiple stages in a process of well expansion and decline in many locations in India (Figure 1). Additionally, we refer to observed well uptake percentages and observed groundwater decline rates. Regarding crop choices, we discuss how our choice of behavioral theory without sufficient negative feedback effects led to too homogeneous cultivation and propose methods to simulate this more realistically.



	Stage 1	Stage 2	Stage 3	Stage 4
Stages	The rise of Green Revolution and Tubewell Technologies.	Groundwater-based Agrarian Boom	Early Symptoms Groundwater Over-draft.	Decline of the Groundwater Socio-ecology with immiserizing impacts.
Examples	North Bengal and North Bihar, Nepal Terai, Orissa.	Eastern Uttar Pradesh, Western Godavari, Central and South Gujarat.	Haryana, Punjab, Western Uttar Pradesh, Central Tamil Nadu.	North Gujarat, Coastal Tamil Nadu, Coastal Saurashtra, Southern Rajasthan.
Characteristics	Subsistence agriculture; Protective Irrigation Traditional crops; Concentrated rural poverty; Traditional water lifting devices using human and animal power.	Skewed ownership of tubewells; access to pump irrigation prized; rise of primitive pump irrigation *exchange* institutions. Decline of traditional water lifting technologies; Rapid growth in agrarian income and employment.	Crop diversification; permanent decline in water tables. The groundwater-based *bubble economy* continues booming; but tensions between economy and ecology surface as pumping costs soar and water market become oppressive; private and social costs of groundwater use part ways.	The `bubble' bursts; agricultural growth declines; pauperization of the poor is accompanied by depopulation of entire clusters of villages. Water quality problems assume serious proportions; the *smart* begin moving out long before the crisis deepens; the poor get hit the hardest.
Interventions	Targeted subsidy on pump capital; Public tubewell programmes; Electricity subsidies and flat tariff.	Subsidies continue. Institutional credit for wells and pumps. Donors augment resources for pump capital; NGOs promote small farmer irrigation as a livelihood programme.	Subsidies, credit, donor and NGO support continue apace; licensing, siting norms and zoning system are created but are weakly enforced. Groundwater irrigators emerge as a huge, powerful vote-bank that political leaders cannot ignore.	Subsidies, credit and donor support reluctantly go; NGOs, donors assume conservationist posture zoning restrictions begin to get enforced with frequent pre-election relaxations; water imports begin for domestic needs; variety of public and NGO sponsored ameliorative action starts.

Figure 20. Rise and fall of groundwater socio-ecology in India.

Figure 1: Stages of groundwater irrigation (Roy and Shah, 2002))

2. As I understand, the region is also heavily managed in regards to the surface water supply system (reservoirs, diversions, manmade canals, etc.), which influences water availability for irrigation and associated demand for groundwater and farm decisions to install a groundwater well. Can the authors speak to the capabilities or limitations of CWatM in effectively representing surface water deliveries for irrigation in this region and how this may be influencing results?

It is indeed true that the area is heavily managed, and we have therefore included several features in the model specifically to address this supply system. Yet, there are limitations and uncertainties.

First of all, reservoir command areas are included in the model. The delineation of the command areas was obtained from the India Water Resources Information System, and manually linked to reservoirs (De Bruijn et al., 2023). In principle, agents can abstract water from these reservoirs if they are in the reservoir command area and have access to the reservoir based on census data.

However, the current reservoir management module follows relative simple decision rules simulating two types of release: (a) the first is release into the river channel, which is based on protocols for reservoirs that are designed for power generation. (b) The second is a daily fixed proportion of total reservoir storage that gets released to farmer agents to abstract from. In the model, there are no physical canals delivering water to agents; instead, agents directly extract water from the reservoir as long as it remains within the daily allocated budget. Upstream agents have priority in water extraction, simulating the way canal water delivery functions in this region (see section 1.3.1. in ODD+D protocol). The volume of these releases are too low, and we, therefore, see relatively little effects of reservoirs in our results. For future research, we want to improve this module and better represent the different types of reservoirs and their effects on farmer adaptation. However, as you and reviewer #1 have both remarked, there are already many elements in the manuscript, thus we have decided to leave it out of the main text, and left the reservoir agents descriptions in the ODD+D protocol.

3. In this discussion, the authors note that groundwater well drilling is potentially maladaptive, as farmers then rely on wells that can go dry during subsequent droughts. These are important findings that seem to be largely glossed over in the results section. For example, there isn't a figure reporting on the drying of these wells during subsequent droughts.

Thank you for noting this statement was not properly linked to the presented results. The drying of wells is represented by the trend of the well percentage (Figure 7; particularly the 2011-2015). These figures refer to wells *with* groundwater access, or "wet" wells. However, to improve the communication of our findings we renamed the figures and changed the descriptions to make this more clear. This downward trend can be due to the effect of wells not being replaced after their maximum lifespan was exceeded, and the drying of wells. However, the drop is too large to be fully explained by the non-replacement and it coincides with the groundwater decline, thus we can attribute the drop in well uptake fraction mainly to drying wells. While we feel adding an additional figure would increase the amount of material again – which both reviewers noted that we should avoid – we have included these notes much more explicitly in the descriptions of the results (section 3.1).

4. It would seem to me that the imitation technique (described in lines 155-156) would very quickly lead to homogenization of crops across farmers using the same irrigation technology. Is this not the case? Could the authors further comment?

This is indeed a very relevant remark, We include imitation together with the SEUT as this is how adaptation has been observed to spread in real life (Baddeley, 2010). Thus imitation in itself would not directly lead to homogenization. However, we agree that in our case, there is indeed too much homogenization. In the revised version of the paper, we discuss several reasons for this feature in the discussion section. First, there is an absence of economic feedbacks. This is especially important since the economic behavior theory we have implemented is mainly based on utility maximization, thus it would require feedbacks in the same domain. Second, there is no accounting of other factors influencing crop choice, such as cultural factors, intention to behavior gaps, unobserved cost factors (similar to Yoon et al. (2024)), and e.g. the prevalence of subsistence farming in the area. While it's true that once a crop rotation option is eliminated it can no longer be chosen, leading to homogenization, we believe that this elimination itself isn't inherently negative. However, the mechanisms driving it should be modeled more realistically.

We have included recommendations to improve methods and future studies can incorporate either additional economic feedbacks, such as a crop market, ensuring that farmer profits go down as more farmers grow a particular crop. However, due to the already complex methodology, we have reserved this for future work. These options are discussed in lines 584 to 599.

Additionally, reducing the number of crops and crop rotations would allow us to let agents compare the different options, without having to rely solely on imitation for computational reasons. Influence of neighbors could then be translated in an adjustment of the intention factor for example. Additionally, instead of letting agents choose between all possible crops, we may explore the decision between crop variety options, allowing agents to select varieties of their main crop that are more resistant to drought/water-efficient/etc, which also fits with literature (Drugova et al., 2021).

5. *I understand that the political economy of sugarcane is particularly influential on water security outcomes in the region (e.g., https://iopscience.iop.org/article/10.1088/1748-9326/ab9925/meta). Could the authors speak at all to how such considerations factor into the analysis? More broadly, crop prices are a significant driving factor of farm behavior, but the subjective expected utilities are only formulated in relation to subjective drought perception. Can the authors comment on whether/how farmer perceptions of economic conditions might influence results (even if outside the scope of this analysis)?*

Indeed, we refer to economic shocks coinciding with meteorological shocks, but only simulate behavior change in response to the latter. In the model, following the SEUT theory, the behavior of farmers is strongly dependent on crop prices, meaning that if prices drop, agents will start cultivating different crops that are now more profitable, and vice versa. We observe this effect for droughts in the model results, but similar behavior would be exhibited in the case of price drops or increases due to other external factors.

But on a perceptional or behavioral level you would expect that farmers fall back on crops which give security (especially during uncertain times). These could be subsistence crops for smaller farmers (mentioned in the discussion), or indeed crops such as sugarcane which have a guaranteed sale and price set by the government. This may require a second behavioral factor, which instead of reweighting the probability of future events, would reweight future crop prices based on their probabilities (e.g. 50% chance at higher prices, or 100% price of a slightly lower price) and would be similarly reweighted by risk perception and risk aversion. Perhaps this could be implemented alongside forecasts, where crops are weighted based on how well it would do in the current forecast along with the forecasts' probability weights? We included these recommendations in the revised version of our paper (lines 600 to 603)

6. The above article is conducted as part of the Stanford FUSE project, which was an outgrowth of the Stanford Jordan Water Project (JWP) which also introduced a coupled agent-hydrologic model for similar types of analysis (e.g., https://www.nature.com/articles/s41893-023-01177-7; https://www.pnas.org/doi/abs/10.1073/pnas.2020431118). While much of this work was focused in Jordan rather than India, these are important studies to note as part of the literature context. Can the authors speak more to how the current effort relates to and is distinguished from this line of coupled agent-hydrological model?

Thank you for bringing the attention to these papers. There are indeed **many** similarities, and unfortunately we have only seen this research as of now. After carefully examining this literature, we think the differences lie in four areas.

- First, our research is focused on drought *events* specifically: What happens during drought events, what happens over consecutive events, how does the crop yield change (which required, for example, a more extensive crop module), how does this affect profits, etc. This event focus is also a step towards future studies where agents have to react to alternating droughts AND floods, which is a different path compared to these studies.
- Second, is that the focus in our paper is more on the differences in situations and behavioral aspects of farmers: how do they make investment decisions using past experiences of droughts, how is this affected by a risk perception, risk aversion, time preferences, farm size, difference in climate between upstream or downstream and how do all these factors affect their choices? Some of these factors are present in the linked studies (and other spatial factors, like transportation costs are equally important but not of relevance for our study), but they are implemented slightly more rudimentary and are less the focus of the research.
- Third, we simulate all agents and localized abstractions instead of using representative agents.
- Lastly, of course we have very different local conditions which require a different model set up. For example, in our study area electricity is subsidized to cost nothing or almost nothing, which means that the costs are in the loan for the initial

investment, and not in the structural price of water. This makes the differences between access in groups much more strict and leads to other dynamics which are characteristic of the area (like many agents at once losing access during a drought) and is a clear difference between these models. For future studies in the global north we do intend to make it more similar to these papers (and to Yoon et al. (2024)), where it is assumed that if agents could have gotten access to groundwater, they would already have, and now pay a price per volume of water (dependent on the groundwater levels, pumping costs, etc.) they use instead of for getting access to the water. The investment decisions are then focused on different ways of decreasing water use, like switching crops or improving irrigation.

We agree that these are good examples of similar socio-hydrological models and have added references to the socio-hydrological nature of these papers in the introduction.

7. Figure quality throughout could be improved. Resolution is often poor with text difficult to make out and colors often hard to distinguish (e.g., couldn't distinguish crops in the cropping figs). Fig 1 is also difficult to interpret and missing text in boxes.

This was addressed in my comments to the previous reviewer; we now use the OKABEITO color palette (Okabe & Ito, 2002), which should still be colorblind friendly, but better suited to categorical data. All labels have been made larger and figure 1 has been updated. Figure quality decreased when we exported the data to PDFs, but production quality figures (300 dpi) will be made available.

8. Lastly, I agree with reviewer #1's comment regarding the >1 million agents. Even if such # of agents is warranted, headlining the # so prominently throughout the paper (in title, abstract, etc) in my opinion misplaces focus and potentially signals the wrong message (e.g., model complexity for the sake of model complexity). This ability to model of large # of agents was already heavily featured/highlighted in the original GEB paper, so in this case I'd rather see the spotlight placed on the insights drawn from the modeling improvements and analysis, rather than the # of agents that can be modeled.

We agree that this was the main focus of the original GEB paper, while this manuscript focusses more on the analysis and results that can be performed with such an approach. We have altered the title to reflect this, and now reads: "Adaptive Behavior of Farmers Under Consecutive Droughts Results In More Vulnerable Farmers: : A Large-Scale Agent-Based Modeling Analysis in the Bhima basin, India". Thank you for the suggestions, we do believe this is a much more fitting title.

Baddeley, M. (2010). Herding, social influence and economic decision-making: Socio-psychological and neuroscientific analyses. *Philosophical Transactions of the*

Royal Society B: Biological Sciences, *365*(1538), 281–290. https://doi.org/10.1098/rstb.2009.0169

De Bruijn, J. A., Smilovic, M., Burek, P., Guillaumot, L., Wada, Y., & Aerts, J. C. J. H. (2023). GEB v0. 1: a large-scale agent-based socio-hydrological model–simulating 10 million individual farming households in a fully distributed hydrological model. *Geoscientific Model Development*, *16*(9), 2437–2454.

Department of Agriculture & Farmers Welfare India. (2001). *Agricultural Census India*. National Informatics Centre (NIC)| Agriculture Census Division, DAC. agcensus1.da.gov.in

Desai, S., Dubey, A., Joshi, B. L., Sen, M., Shariff, A., & Vanneman, R. (2008). India human development survey. *College Park, Maryland: University of Maryland*. https://doi.org/https://doi.org/10.3886

Drugova, T., Curtis, K. R., & Ward, R. A. (2021). Producer preferences for drought management strategies in the arid west. *Renewable Agriculture and Food Systems*. https://doi.org/10.1017/S1742170521000259

Hoda, A., & Terway, P. (2015). *Credit policy for agriculture in India: An evaluation. Supporting Indian farms the smart way. Rationalising subsidies and investments for faster, inclusive and sustainable growth*. Working Paper.

Maertens, A., Chari, A. V., & Just, D. R. (2014). Why farmers sometimes love risks: Evidence from India. *Economic Development and Cultural Change*, *62*(2), 239–274. https://doi.org/10.1086/674028

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, *17*(22), 179–183.

Okabe, M., & Ito, K. (2002). How to make figures and presentations that are friendly to color blind people. *University of Tokyo*.

Page, S. E. (2012). Aggregation in agent-based models of economies. In *Knowledge Engineering Review* (Vol. 27, Issue 2, pp. 151–162). https://doi.org/10.1017/S0269888912000112

Roy, A. D., & Shah, T. (2002). Socio-ecology of groundwater irrigation in India. *Intensive Use of Groundwater Challenges and Opportunities*, 307–335.

Udmale, P., Ichikawa, Y., Manandhar, S., Ishidaira, H., Kiem, A. S., Shaowei, N., & Panda, S. N. (2015). How did the 2012 drought affect rural livelihoods in vulnerable areas? Empirical evidence from India. *International Journal of Disaster Risk Reduction*, *13*, 454–469. https://doi.org/10.1016/j.ijdrr.2015.08.002

Yoon, J., Voisin, N., Klassert, C., Thurber, T., & Xu, W. (2024). Representing farmer irrigated crop area adaptation in a large-scale hydrological model. *Hydrology and Earth System Sciences*, *28*(4), 899–916. https://doi.org/10.5194/hess-28-899-2024