EGUsphere, author comment AC1
https://doi.org/10.5194/egusphere-2022-664-AC1, 2022

Reply on RC1

Jens A. de Bruijn et al.

Author comment on "GEB v0.1: a large-scale agent-based socio-hydrological model – simulating 10 million individual farming households in a fully distributed hydrological model" by Jens A. de Bruijn et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-664-AC1, 2022

We thank the reviewer for taking the time to review our manuscript. Below we list the comments made by the reviewer and suggest changes in italic.

The study by Bruijn et al. with the title of "A large-scale agent-based socio-hydrological model – simulating 10 million individual farming households in a fully distributed hydrological model" is intended to provide a coupled agent-based and hydrological model to simulate farmers' behavior. At the current stage, the paper does not present a proper understanding of agent-based modeling and socio-hydrological model. While the study implies that agent-based modeling benefits this work, I am quite concerned and surprised about the materials of the paper regarding agent-based modeling:

- What is the difference between your current model and the high-resolution water management model? The authors named some components as agents (i.e., reservoir operators, a government, and an NGO) though they act as the same as a traditional water management model. For example, there is a so-called NGO agent, but when we read the information about its action in Section 3.5, there are like "scenarios"! the same example can be seen for the government agent. Thus, another concern can arise regarding the wrong definition of agents in this study.

CWatM does not include the high-resolution water management components which are described in this manuscript i.e., CWatM only has grid-cells, and no sub-grid information / hydrological response units. The differences between CWatM and the high-resolution water management model are described in section 2.1 and 2.2. These sub-grid hydrological response units allow the simulation of these individual farmers agents (of which there are >10 million) and their individual bi-directional connection to the fully distributed hydrological models. The NGO agent, government agent and reservoir operators or entities can make individual autonomous decisions that affect farmer agents or the environment (and thus other agents indirectly), and are thus agents (Bonabeau, 2002).

However, it is indeed correct that the behaviour of farmer agents is rather homogeneous. Our intend was to present the model framework first which includes a framework for computationally optimized ABMs in Python, a heavily adapted hydrological model (e.g., sub-grid hydrological response units), and introduce several stylistic scenarios to showcase the model. However, based on the reviewer's comments (as well as other reviewers) we propose to replace the current scenarios (with rather homogeneous agent *characteristics) with a scenario with more heterogeneous agent characteristics (and resulting behaviour). We hope that this implementation will make the agent-based modelling component stronger, and will thus satisfy the reviewer. To do so, we will*

1. Use the Indian Development Human Development Survey (IHDS), which presents highly detailed information on 41.554 households, including crop types for the different growing seasons, household size, household income, expenditure, irrigation techniques, farm size etc etc.
2. From the Indian agricultural survey, we will collect farm characteristics (i.e., marginal distributions), including farm sizes and crop types at tehsil level (comparative to counties in the US)
3. Use an adapted version of the iterative proportional fitting to create a synthetic farmer population using the micro-level data from the IHDS yet fits the marginal distributions of the tehsil level agricultural survey
Then based on this data, land use data derived from satellite imagery, and the distribution of farm sizes, we will distribute those heterogeneous farmers spatially, which now include heterogeneous characteristics (according to the IHDS), such as crop types for the different growing seasons, household size, household income, expenditure, irrigation techniques, farm size etc.

Then, using historic crop prices, as obtained from the Indian Agricultural Marketing Information System, in combination with simulated yield based on farm size, potential yield and the simulated ratio of actual to potential evapotranspiration, we will simulate individual farmer income. Combined with inflation-adjusted household expenditures (from IHDS data) and crop expenditures (based on data obtained from the Ministry of Agriculture and Farmers Wellfare) we can calculate disposable income.

Next, we can include farmer adaptation behaviour, specifically the construction of irrigation wells, in the model. To simulate this, farmers without an irrigation well, look to (farm-size adjusted) income of neighbouring farmers with similar crop types but with an irrigation well. If the income difference of the agent's farm compared to those surrounding farms is higher than the implementation and upkeep costs of an irrigation well, the farmer will implement an irrigation well.

This also allows us to derive more in-depth conclusions, based on the heterogeneity of these agents (in addition to the presentation of the model framework and potential future applications).

• The agent-based model is well-known for stochastic processes, learning and adaptive procedure, and complex interactions among agents. How do you benefit from each or some of these features in your model structure and equations? Please explain each feature, if there is any, according to your equations.

Stochasticity

In short, the placement of agents and assigning characteristics of those agents are random, while decision rules are largely deterministic. Models that use agents, are on a range between fully stochastic agent-based models where all decisions are stochastic to more deterministic agent-based models, which is frequently the case for spatial models. There are many different definitions of what constitutes an agent-based model. Here, we take the definition that an agent-based model consists of many interacting entities (here: >10 million farmers), in a social setting with relatively simple rules, yet creating a complex system. Because the model is so large, the outcome of the system is unknown beforehand, and the model can provide explanatory insight into emergent behaviour of the system, similar to other agent-based models and environmental models such as traditional hydrological models and climate models. The model now includes a social *component through including individually endogenously adapting agents.*

Learning and adaptative procedure

The agents (both in the 'old' scenarios in the current manuscript and the 'new' proposed scenario as suggested above), agents can adapt over time by implementing adaptation measures such as more efficient irrigation techniques (old scenario) and the construction of groundwater wells (new scenario).

Both in the 'old' and 'new' scenarios there are several behaviours that depend on the neighbours, such as the learning of adaptation techniques and the viewing of income of neighbouring farmers. But most importantly, the behaviour of agents affects other agents through the hydrological modelling. If an upstream farmer abstracts water from a reservoir, or stream it cannot be abstracted by downstream farmers. Moreover, groundwater extraction, lowers the groundwater table, making investing in groundwater wells more expensive.

Complex interactions

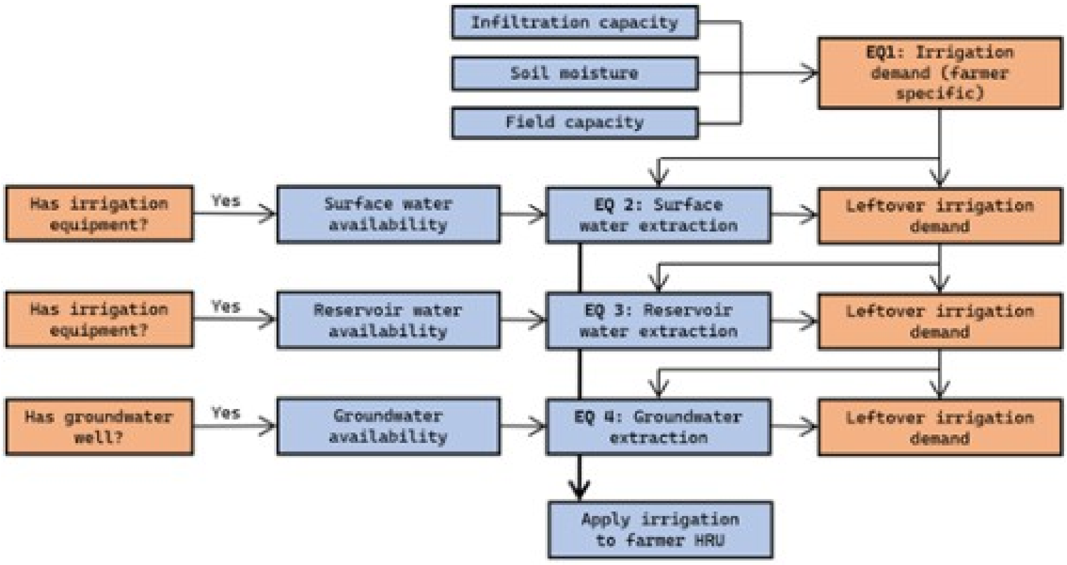
The complex interactions between agents in this model can be subdivided in the parts:
1. Agent-agent interactions: In the 'old' scenario, the agents learn from each other in the NGO adaptation scenario "All farmers with a higher irrigation efficiency have a daily 1% probability of disseminating the knowledge to another farmer within a 5 km range", and in the "new" scenario, a similar procedure is proposed for adaptation of irrigation wells (see above).
2. Agent-environment-agent interactions: The most complex interactions between agents are through the hydrological model (e.g., when one farmer abstracts water, the "same" water cannot be abstracted by another farmer). These interactions go through the groundwater (i.e., MODFLOW), sub-surface and surface (i.e., CWatM). All these models are dynamically coupled thus creating a complex system.

Besides these comments, I also have other major comments:
• The paper lacks a literature review on socio-hydrological models.

We will include an additional paragraph mentioning the benefits that coupled agent-based models have with respect to agent-based models or hydrological models only. However, we would not propose to include a full literature review in line, which is in line with other manuscripts that I have been published in Geoscientific Model Development (but instead focus on presenting the model).

• The socio-hydrology is mainly about the coevolutionary behavior between the hydrological and human systems. Please clarify the bidirectional feedback between a hydrological and social system in a figure. The current figures do not satisfy this need. Please also note that scenarios cannot show the "co-evolutions" between the systems.

In an updated manuscript we will include an additional figure (see below) to show the irrigation water consumption and the effect on hydrology in more detail, see image below. Other feedbacks between the hydrological and human system, such as crop choices, and reservoir management are shown in Figure 2.

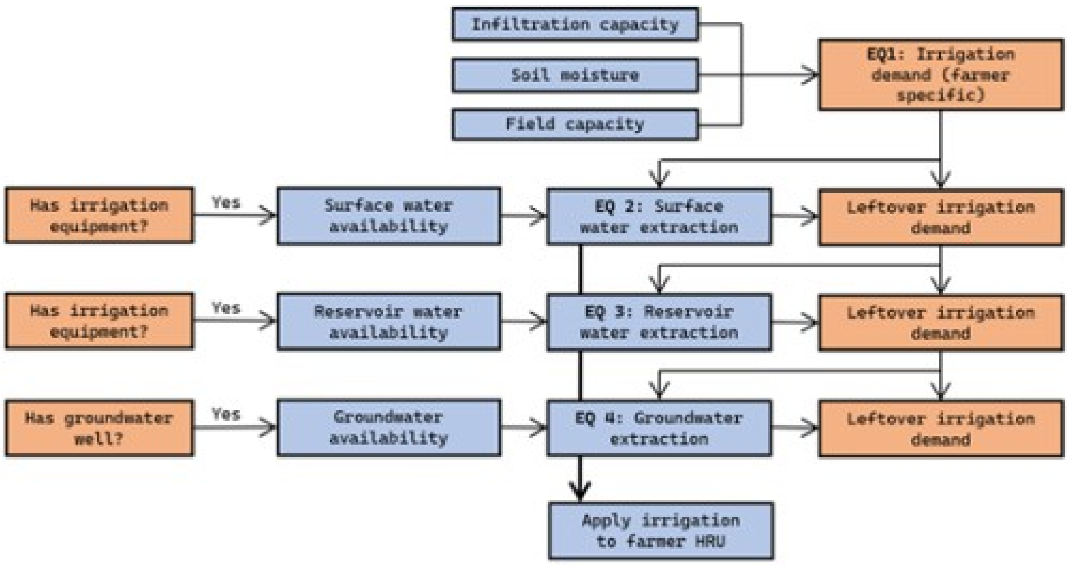


• Another point of socio-hydrology is to involve social factors. In lines 131-134, I am very surprised that the authors just mentioned they will work on it in their future work. Then, what are the social components in this study?

These social components were included to a limited extend in the NGO scenario (lines 337-339), but will be included more in the new storyline as suggested above. Farmers will look at the income of neighbouring farmers with similar crops, and decide based on that, whether to adopt an irrigation well (more info see above).

• I wonder if Figure 2 shows the complete picture of the current paper?! The authors mentioned some components (e.g., characteristics and experience) will be in their future work (lines 131-134)! Also, there should be a meaningful connection between Figure 2 and the equations. You should completely clarify each component of the figure and refer to the corresponding equation.

Here, we propose the same Figure as above, with the equations labelled as EQ1 … EQ4.



• It seems that Equations 1-4 represent all agents' equations in the model. Once again, I wonder how this work benefits from the concept of an agent-based approach. what is the difference between this model and a traditional water management? It seems that the only advantage of this water management work is to provide a high-resolution model.

There are several differences. 1) the ability to include heterogeneous decision making, which would be realized in the newly proposed scenario and 2) the two-way interaction between the hydrological model and the social system (although this can also be included in other types of models, such as system dynamics models) this is a strength of the integration between ABMs and hydrological models).

• I am very surprised by the conclusion section. Around 75% of the conclusion is about future work! what is the take-home message of this study? What are the implications?

The main aim of the manuscript is to present the model, which for the first time enables highly detailed simulation of millions of farmer agents in large river basins. This model can be adapted for a wide range of applications, which is why we suggested several avenues of research. However, in a potential new version of the manuscript we will expand the conclusion to include several take home messages based on the newly proposed scenario, for example focussing on the heterogeneity in farmer characteristics and the influence on adaptation behaviour.

Some minor comments:
Line 55: very confusing. Are you talking about approaches of modeling human systems? what is the traditional hydrological component as an approach? here I suggest you write about approaches of simulating water-human systems (including SD and AB).

We are writing about the hydrological component, which is mentioned at the beginning of sentence. However, we suggest to write this more clearly as follows:
"In general, two approaches can be differentiated in adding a hydrological component: using an agent-based or traditional hydrological component."

-->

"In general, two approaches can be differentiated in adding a hydrological component to coupled agent-based hydrological model: using a hydrological component which is agent-based (e.g., river segments are represented as agents which exchange water) or traditional hydrological component (e.g., a gridded model where water flows from once grid cell to another based on the kinematic wave equation)."

Line 63: The other approach to consider/model what?

The previous paragraph mentions there are two approaches, we first mention the first approach, then the other approach is the second approach. To make this more clear, we suggest to write the following. This also includes merging the paragraphs, making it more clear that these approaches are connected.

In the first approach, the agent-based approach, all the environmental components, such as river segments, are simulated as agents. For example, Becu et al. (2003) simulate farmers, irrigation behaviour, and crop and vegetation dynamics. Their model uses a simple routing scheme that considers water abstraction and water diversions by canal managers. Another example is Huber et al. (Huber et al., 2019), who created a basin-scale coupled model where water flows downstream from river agent to river agent, while other agents such as farmers or water managers can abstract water from the river. In this approach, the hydrological component is usually relatively simple, largely because authors usually build the hydrological component from scratch. The second approach, the hydrological model approach, is to couple an agent-based model with a more traditional hydrological model by allowing the agents to interact with its water storage.

Lines 64-65: Is it an example of coupling the agent-based modeling and a hydrological model? what is the agent-based modeling part about? Please explain more.

The aim of these paragraphs is to talk about the coupling of the hydrological components (which is also the main aim of this manuscript). Therefore, we feel it will confuse the reader to go in depth on the agent-based component of these papers within this paragraph. However, references are provided. We suggest however, to include another paragraph to show a couple of agent-based approaches.

Line 66: What does it exactly simulate?

The irrigation behaviour. We will add this as follows:

"grid-based model at a 270m resolution that simulates the irrigation behaviour of individual farmers in a large basin"

Lines 63-70: this paragraph just provides a list of references and confuses readers. what is the main idea of this paragraph? if you want to mention the advantage of coupling agent-based modeling and hydrological model, explain the agent-based modeling and hydrological part in each study. what was the benefit of agent-based modeling for each study?

The aim of this paragraph is to talk about the coupling mechanisms of the models, which should be clearly with the additions described above. In the updated manuscript, we will, however, include an additional paragraph which mentions some of the benefits and processes that agent-based models have included.

Line 71: what are the "these methods"?

This is indeed phrased in a confusing way, and suggest the following change:

"Many agent-based models with a hydrological component were released using these methods."
-->
"A large number of agent-based models with a hydrological component were released."

Lines 230 and 240: what is section 0?

There was an error in the internal referencing in the document, this should point to section 3.5, and will be corrected.

Line 330: this section should change to scenario analysis

We will adapt the use of the word storyline to scenario in the entire manuscript.

Lines 331-347: what is the argument behind choosing these numbers like 30% probability of switching crops? Do you have any references for them? If not, there is a need to do a sensitivity analysis.

In the new scenario these will be replaced.

Lines 401-404: "for the first time"?

We would not claim to present the first model that makes a coupling to a hydrological model (also see cited models in the introduction), but to the best of our knowledge this is the first model to do so which can include "millions of individual households" in *combination with “bi-directional interactions with the hydrological system”. If we have missed works that do this, we would be happy to adapt this.*