We thank the additional reviewer for their constructive comments, which we addressed in blue below and incorporated into our manuscript. We would also like to thank the editorial team for their thoroughness. However, our paper has been under review with HESS for over a year, during which we have addressed two rounds of revisions. It would be appreciated if a resolution could be reached soon.

I have read this paper with some interest. However, I don't think it is as novel as the authors state. While the authors state that they are not aware of any other work that combines the spatial and temporal ability of OFRs with multi-satellite imagery, I would argue that the paper by Robertson and the associated work on assessing the volume of OFRs is probably more advanced than this study. There is similar work by Xie Yan et al. 2023. (See references below)

We removed the word "novel" in most instances and fully acknowledge the limitations of our work: "It is important to keep in mind that the impact of the OFRs on this study is solely based on modeling scenarios and does not account for OFR management practices, which represents a key limitation of this simulation study."

Thank you for pointing out the studies by Robertson et al. (2023) and by Xie Yan et al. (2023). While we acknowledge the relevance of these studies, they try to understand the problem through a different lens without incorporating on-farm reservoir surface water extent derived based on remotely sensed data in their hydrological model, which we believe is the new perspective we are providing here.

Robertson et al. (2023) quantified how the interaction between climate change and farm dams affects streamflow characteristics. They relied on the farm dam dataset produced in Malerba et al. (2021), which includes point data for farm dams, rather than their varying surface water extent (like in our work presented here). Their use of remote sensing was a straightforward method, using annual time-series of water index thresholds, on three Landsat sensors to determine the first year where water is observed in each farm dam from Malerba et al. (2021), which was then used as a proxy for the year of dam construction in their modeling.

Yan et al. (2023) employed a hydrological model to distinguish the impacts of climate variability, land-use change, and small- to medium-sized reservoirs on streamflow in two river basins in Southeast China.
Their use of remote sensing was ONLY in visually deriving the land use data based on Landsat 5 data. The obtained the reservoir data from one of the local provinces where their study catchments are located (quote taken directly from their paper: "In addition, there were also reservoir data during the study period (1970–2015), including the location, scale (catchment area, storage capacity), type, and construction time, which were obtained from Water Resources Management Department of Fujian Province").

We reviewed and cited the two studies in our manuscript. The new lines in the manuscript are copied below: "For example, other studies have found that OFRs reduce annual and monthly runoffs in southeastern China (Yan et al. 2023) and Australia's Murray-Darling Basin (Robertson et al. 2023)."

The large body of work on this topic generally highlights the major difficulties and uncertainties in assessing the impact of OFRs.
This paper identifies those too, but ultimately they only address one of the major sources of uncertainty i.e the model structural uncertainty in representing spatial and temporal variability of OFRs.

The other major sources of uncertainty identified in the literature i.e surface area volume relationships of the systems and the management thereof for irrigation or other abstractions.

We fully agree with the reviewer and provide the reasons why we were ultimately able to address only one of the sources of uncertainty.

We have not done so for a few reasons:

- Irrigation data is rarely available
- For this type of irrigation (from OFRs), farmers do not have meters on their reservoirs—even if they have meters, the data would be sparse. Hence, calculating water abstract was unfortunately not feasible.
- Surface area-volume relationships are helpful, but also limited, given that the OFRs do not follow a consistent construction pattern and vary from farm to farm. A rough estimate using a trapezoidal equation would probably be as accurate as possible if you do not measure the bathymetry in situ.

This is expressed in the manuscript in the following paragraphs:

"We did not have access to water abstraction data from the OFRs, so all abstractions were modeled using Equation 3, which accounts for water flowing out of the OFR, as well as losses from evaporation and seepage. The total volume of water in the OFR fluctuates in response to changes in surface area and is also influenced by evaporation losses and the operation of the spillway. A reduction in surface area (Equation 4) typically leads to a corresponding decrease in water volume. If inflows are insufficient to fill the OFR, water will not be routed to the downstream channel.".

"For each of the aggregated OFR, the initial water volume (Vstored, see Equation 3) was calculated using the SWAT+ default rule, which is a simple multiplication of the OFR surface area by a factor of 10, similar to other studies based on SWAT+ (Ni and Parajuli, 2018; Zhang et al., 2012). For a scenario where the OFR has a surface area of 1 hectare (10,000 m^2), the corresponding volume would be 100,000 m^3—this is a limitation of our study, as the assumption was necessary due to the absence of available bathymetry data. In addition, since we did not have access to the OFRs' release rates, we used the model's default release rule, which sets the OFRs to release water when the spillway volume is reached—80% of the OFRs' capacity (Bieger et al., 2017)."

• The uncertainties related to the spatial distribution of the reservoirs and how to capture those in hydrological modeling.

o Although I'm not familiar with the latest version of SWAT/SWAT+, based on the paper, it seems that recent advances do provide a useful way of capturing the spatial distribution of reservoirs and simulating their impact.

We have conducted our modeling at the channel level, which incorporates more of the spatial variability in reservoir distribution, illustrated in the following paragraph: "In addition, the latest versions allow for adding more than one OFR per subwatershed by associating the OFR with channels—components of the watersheds, as well as finer divisions and extensions of water stream reaches, enabling modeling analyses at the channel scale. When simulating the impact of the OFRs at the channel scale, there is a higher level of detail of where and when the OFRs are contributing to changes in surface hydrology, unlike the previous versions of the model, which allowed adding only a single OFR per subwatershed placed at the subwatershed outlet as a point (Arnold et al., 2012), and therefore, the analyses were conducted at the subwatershed scale.".

• The difficulty in obtaining and estimating the volume associated with small farm dams and their identification.

o To some extent, the authors address this issue in the paper. However, I do not think that they do so adequately. In effect, they conclude that measured ground data of OFR depth is needed . However, various tools that have been developed globally to address this problem.

For OFRs of the small size we are working with, to the best of our knowledge, there are no accurate ways of doing this, without ground measurements, which we did not have access to. Given that these are small water bodies, a rough trapezoidal equation would likely provide a very *rough* estimation, using a local DEM. But given the small size of OFRs, you can miss the total water volume by a lot, even when using local DEMs (e.g., by 100%). We addressed this in the paper in the following paragraphs:

"Estimating the OFR's volume change can be done by combining the OFR surface area time series with area-elevation equations—these equations describe the OFR's bathymetry, and allow volume estimation by inputting the OFR's surface area (Liebe et al., 2005; Meigh, 1995; Sawunyama et al., 2006). After carefully assessing different methods to derive these equations (Arvor et al., 2018; Avisse et al., 2017; Li et al., 2021; Meigh, 1995; Sawunyama et al., 2006; Vanthof & Kelly, 2019; Yao et al., 2018; Zhang et al., 2016), we concluded that measured ground data of the OFRs' depth—which is not available—is required to estimate the equations with an acceptable uncertainty. Estimating the area-elevation equations entails several challenges, including: 1) even though there are several DEMs available for the study region (Arkansas GIS Office, 2022)—DEMs can be used to estimate the OFRs bottom elevation—the DEMs were collected when most of the OFRs were full (i.e., bathymetry was not exposed), which limits their use in this case; and 2) although the OFRs are located within the same geomorphological region, they have different depth, shape and physical characteristics (Perin et al., 2022; Yaeger et al., 2017). Therefore, even if a generalized area-elevation equation were calculated for our study region—this is a common approach employed by other studies (Mady et al., 2020; Vanthof and Kelly, 2019)—that would still lead to high uncertainties in water volume changes."

"Future improvements should focus on how to better represent OFR's water management (i.e., OFR's inflows and outflows) in SWAT+. Given that each OFR has an independent water balance, accounting for the OFRs water volume change would be a more realistic representation of the OFRs when compared to the three surface area scenarios tested in this study. Estimating the OFRs volume change can be done by combining the OFR surface area time series with area-elevation equations—these equations describe the OFRs' bathymetry, and allow volume estimation by inputting the OFRs' surface area (Liebe et al., 2005; Meigh, 1995; Sawunyama et al., 2006). After carefully assessing different methods to derive these equations (Arvor et al., 2018; Avisse et al., 2017; Li et al., 2021; Meigh, 1995; Sawunyama et al., 2006; Vanthof & Kelly, 2019; Yao et al., 2018; Zhang et al., 2016), we decided that measured ground-data of the OFRs' depth—which is not available—is required to estimate the equations with an acceptable uncertainty. Estimating the area-elevation equations entails several challenges, including: 1) despite the fact that there are several DEMs available for the study region (Arkansas GIS Office, 2022)—DEMs can be used to estimate the OFRs bottom elevation—the DEMs were collected when most of the OFRs were full (i.e., bathymetry was not exposed), which limits their use in this case; and 2) although the OFRs are located within the same geomorphological region, they have different depth, shape and physical characteristics (Perin et al., 2022; Yaeger et al., 2017). Therefore, even if a generalized area-elevation equation were calculated for our study region—this is a common approach employed by other studies (Mady et al., 2020; Vanthof and Kelly, 2019)—that would still lead to high uncertainties in water volume changes. Ideally, each OFR would have its own equation, which was not possible when this study was done. Future work should integrate data on actual evapotranspiration, ET (Kiptala et al., 2014), to quantify the balance between water availability and ET, which determines, in large part, the irrigation system efficiency and crop productivity in the watersheds where OFRs occur."

In particular, the work in Australia, which was based on a large sample of measured reservoirs, is relevant (Malerba et al., 2021). In essence, the authors provide a rebuttal of why they have not addressed this aspect - but in essence, it does not advance the understanding in how to address a major area of uncertainty in the study.

We believe the main contribution of our work is the integration of remotely sensed surface water extent of OFRs in SWAT+. We agree that Malerba's work is great, but they do have access to ground measurements, which unfortunately we did not have access to.

• The major uncertainty around OFRs is how they are managed and how much water for irrigation is actually abstracted from them. Other studies (e.g., Hughes and Mantel) identified this as the major source of uncertainty in such studies - greater than model structure and spatial representation type uncertainties.

o I find this aspect a little bit odd in the paper. The authors identify the importance of OFRs for irrigation in the justification for the paper (Lines 18, 20, 42-43) and recognize that how they are managed has a particular impact on surface water hydrology. However, they never consider irrigation in the model configuration. Section 2.2 deals only with the SWAT+ setup to simulate OFRs but completely neglects any management of the OFRs, such as irrigation abstractions.

We fully agree with the reviewer, however, as mentioned before, we have not done so because this data was not available.
Even if we had access to this data, translating it so that SWAT+ would take those measurements would involve model modifications, which is outside the scope of this work.
Our intention was to model the water bodies as part of the hydrological systems. I.e., what is the impact of their presence?
Having irrigation data would be a plus, but we unfortunately did not have access to such data.

o Ultimately, the authors conclude that this is still a major area of uncertainty and make some recommendations for the enhancement of the SWAT model to deal with this. In essence, the known or estimated inflows and outflows cannot be accommodated in the current form and that the "decision tables" approach of SWAT is too difficult to set up for multiple reservoirs. They do not make new recommendations about how this could be improved. Line 166 highlights the new flexibility in defining management schedules - but the paper does not address this. Had they really got to grips with this aspect, this would provide a strong innovative aspect of the paper.

There is a clear gap that needs to be addressed in SWAT, but we are not SWAT+ modelers per-se and thus our suggestions are primarily centered around the data needs and availability rather than the details in the model. However, we have made some modeling suggestions in the manuscript:

"Efforts should also be made to improve SWAT+ capabilities to receive measured OFRs' inflows and outflows. The latest version of the model has improved the hydrological representation of small water impoundments in SWAT+ (Mollina-Navarro et al., 2018). Nonetheless, at the time of our study, the newest version of the model does not allow users to input measured or calculated OFRs' inflows and outflows. Instead, the model developers recommend simulating the OFRs water balance using decision tables (Arnold et al., 2018; Dile et al., 2022). However, there are very limited guidelines on how to create these decision tables. In addition, the tables would simulate the OFRs water balance instead of using the measured or calculated volume change, which could introduce more uncertainties to the modeling scenarios."

Ultimately, this is a useful technical study combining a number of different techniques, i.e., remote sensing and application of SWAT+ to provide an effective spatial representation of OFRs and a useful framework in which this is done. However, the paper does not really advance this field significantly because the major source of uncertainty, i.e., irrigation abstraction and other management-related aspects of the reservoirs, are completely neglected,

It is not clear whether this is because the model cannot represent them, a lack of data, or other reasons. It is still a major shortcoming in the paper and needs to be better acknowledged as such. Whether the paper is then publishable or not falls back to how unique the combination of the satellite remote sensing coupled to the process-based hydrological model really is.

We thank the reviewer for this comment. We are aware of this shortcoming, which is due to the lack of data availability in our case. We expressed this in the paper in the future improvements section.

Our main improvements are summarized in 4.3 Research implications and applications to other study regions and several other paragraphs: "Studies have employed either data-driven or physically based hydrological model approaches to estimate the effects of OFRs on watersheds. However, combining these approaches provides a better understanding of the spatial and temporal variability of OFR impacts, as it incorporates the dynamic changes of OFRs into the hydrological model.Studies have either used data driven or physically based hydrological model approaches to estimate OFR impacts on the watersheds, despite the fact that combining the two approaches leads to better understanding on what is the spatial and temporal variability of the OFR impacts, given that the dynamic changes of the OFRs are incorporated into the hydrological model. To quantify whether the impact of the OFRS on mean and peak flow varies intra- and inter-annually, and which subwatersheds are more affected, we combined a data-driven remote sensing-based model with the latest improvements in SWAT+ to assess the OFR impacts."

And in this paragraph: "By leveraging the latest improvements in SWAT+ to simulate water impoundments (Molina-Navarro et al., 2018) and combining them with a novel algorithm based on time series of satellite data to monitor OFRs (Perin et al., 2022), we modeled the impact of OFRs on flow at the channel scale. In addition, the surface area scenarios enabled us to account for events when the OFRs were at the lowest, regular, and fullest capacities according to their surface area (see Fig. 2). This is an improvement over previous studies (e.g., Ni et al., 2020; Ni and Parajuli, 2018; Perrin, 2012) that used a single surface area (i.e., one snapshot in time) to represent the OFRs in SWAT."

M.E. Malerba, N. Wright, P.I. Macreadie A Continental-Scale Assessment of Density, Size, Distribution and Historical Trends of Farm Dams Using Deep Learning Convolutional Neural Networks Remote Sens. (Basel), 13 (2) (2021), 10.3390/rs13020319

Hughes, D. A. & Mantel, S. K. (2010) Estimating the uncertainty in simulating the impacts of small farm dams on streamflowregimes in South Africa. Hydrol. Sci. J. 55(4), 578–592.

David E. Robertson, Hongxing Zheng, Jorge L. Peña-Arancibia, Francis H.S. Chiew, Santosh Aryal, Martino Malerba, Nicholas Wright, How sensitive are catchment runoff estimates to on-farm storages under current and future climates?,Journal of Hydrology, Volume 626, Part A, 2023

Xie Yan, Bingqing Lin, Xingwei Chen, Huaxia Yao, Weifang Ruan, Xiaocheng Li, Impacts of small and medium-sized reservoirs on streamflow in two basins of Southeast China, using a hydrological model to separate influences of multiple drivers, Journal of Hydrology: Regional Studies, Volume 50, 2023,

Ignacio Fuentes, Richard Scalzo, R. Willem Vervoort, Volume and uncertainty estimates of on-farm reservoirs using surface reflectance and LiDAR data, Environmental Modelling & Software, Volume 143, 2021,"