General comments

This is a well-written paper that clearly identifies the stakes associated with the approach (I particularly like, in 5.1, the discussion about hydroclimatic uncertainty, about the potential impacts of the use of seasonal forecasts, and of an extension to Bayesian theory), and exemplify its use. Ideally, I would have like a couple of points to be further developed (but this is partly subjective and informed by my own biases, in particular the first two points):

We have included additional discussion of the points identified by the reviewer and hope that the revised manuscript fully addresses the reviewer's comments.

- the limitations of the SWAT modeling framework itself, which is a crucial part of the framework, especially when going to higher temporal resolution (it was interesting to see the calibration and validation results),

More details about SWAT model development, including the calibration and validation results were provided in supplementary data (S2-S4). The SWAT model communicates with the agent-based model annually but runs on a daily basis. This temporal resolution of the SWAT simulation is sufficiently high as well as typical for modeling large-size river basins.

- the limitations of agent-based models (although it is mentioned briefly on line 242), and the fact that they are better as a space explanatory tool (what they are used for in the paper) than for prediction

The discussion of the limitations of agent-based models has been expanded, and can be seen in lines 502-514 (reproduced below).

The development of coupled river basin models needs to carefully address several tradeoffs to ensure that the models are scientifically sound and computationally tractable. The focus of this work is to develop a generalized ABM framework that addresses model transparency and model/module reusability (An, 2012; Parker et al., 2003). To address this, the geographic delineation of our agents are relatively larger than traditional agent-based models (which define individual water users as agents). This is a necessary simplification in order to balance model complexity (or the level of detail of simulated decision processes) with computing resources and data availability. Furthermore, it is pertinent to recognize that agent based models are best used to explain existing relationships or phenomena, rather than as prediction tools. Another related limitation associated with large-scale agent-based models is their reliance on informal validation. For the case studies presented here, we validate the ABM with internal checks, for instance by comparing modeled and observed hydropower generated (Fig. S4). We also address this limitation through the use of surveys to inform agent behavior rules.

- the impact of potential seasonal forecasting capacity (e.g. based on El Nino) on agent decisions,

The potential use of seasonal forecasts and related considerations have been added to the discussion in lines 494-501 (reproduced below).

Another useful extension of this modeling framework would be to incorporate seasonal forecasts of water availability into the decision-making process of agents. Water managers often perceive *the advantages offered by seasonal forecasts are often perceived by water managers as being low (Pagano et al., 2002), even though the economy-wide benefits of seasonal forecasts can be substantial (Rodrigues et al. 2016). This modeling framework can be used to highlight the potential benefits of short-term seasonal forecasts for agents' decisions on water allocation and willingness to cooperate with other agents, and introduce another dimension of stochasticity to the agent decision-making process. The seasonal forecasts used, however, would need to be geographically suitable and temporally appropriate for each agent's operations.*

Pagano, T. C., Hartmann, H. C. and Sorooshian, S.: Factors affecting seasonal forecast use in Arizona water management: A case study of the 1997-98 El Niño, Clim. Res., 21(3), 259–269, doi:10.3354/cr021259, 2002.

Rodrigues, J., J. Thurlow, W. Landman and C. Ringler, R. Robertson and T. Zhu. 2016. The economic value of seasonal forecasts: Stochastic Economy-Wide Analysis for East Africa. IFPRI Discussion Paper 1546. Washington DC: IFPRI. http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/130497/filename/130708.pdf

- the surveys performed and their use for calibration.

We have provided further explanation for the surveys that were performed, and their usage in the modeling framework in lines 144-153 (reproduced below). In addition, we have included a copy of the survey questionnaire in the supplemental material.

As part of this project, we conducted comprehensive electronic surveys across three transboundary river basins (Indus, Mekong and Niger) to identify water use preferences (Khan et al., 2017). A sample survey questionnaire is provided in the supplemental material. The surveys were developed to elicit the perceived importance of various ecosystem services in different parts of each basin under a variety of economic and hydrologic future conditions. The survey sample size ranged from 75-85 for each of the basins. One of the questions in the survey asked respondents to rank different ecosystem services in order of importance for each agent. These responses were then averaged across all the respondents for each agent to obtain a ranking of the importance of the different ecosystem services. These rankings were used in the decision algorithm for the case study models developed and presented in Sect. 4.

The web-app is also intuitive to use (although I could not find the source code on GitHub when going to that page).

Thank you for the comment. The source code for the web-app and the coupled model is now on GitHub (https://github.com/qzhao22/WLE_TOOL_INTERFACE/)

Specific comments

The hydroclimate time series are said to come from historical data. Could the sources of the data made clearer? Does the series chosen conserve temporal cycles? Maybe it would be interesting to have some plots as well to compare to the results given.

Sources for the data used for SWAT model development, including climate data that are used to drive the simulations of the entire SWAT-ABM modeling system, are provided in Table S2 of supplementary data (reproduced below). The data periods are 1983-2007 (the Mekong River Basin) and 1985-2010 (the Niger River Basin). Temporal cycles of climate variables in the two study river basin are represented in the simulations and have been preserved from historical data. Furthermore, we have included plots showing modeled and observed streamflow at different points along the Mekong and Niger in the supplementary data (reproduced below).

Table S2: Data for SWAT model setup

Category	**Data**
Elevation	HydroSHEDS[1]
Land use/land cover	GLC2000[2] & SPAM 2005[3]
Soil	Soil Map of the World[4]
Precipitation	Mekong: APHRODITE[5] Niger: NCEP-CFSR[6] (monthly totals were corrected using monthly precipitation data in CRU TS v. 4.00[7])
Temperatures/solar radiation/relative humidity/wind speed	NCEP-CFSR

1. Source: The SHuttle Elevation Derivatives at multiple Scales (HydroSHEDS) database http://www.hydrosheds.org/
2. Source: Global Land Cover (GLC) 2000 database. European Commission, Joint Research Centre. http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php
3. Source: Spatial Production Allocation Model (SPAM) database for 2005, IFPRI. http://mapspam.info/
4. Source: FAO/UNESCO. http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/
5. Source: Asian Precipitation-Highly Resolved Observational Data Integration Towards the Evaluation of Water Resources (APHRODITE) project. http://www.chikyu.ac.jp/precip/english/conditions.html
6. Source: National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR); downloaded via global weather database for SWAT https://globalweather.tamu.edu/
7. Source: Climatic Research Unit - University of East Anglia. http://www.cru.uea.ac.uk/data

S3 Sub-basin Delineation

(a) **Mekong** (b) **Niger**

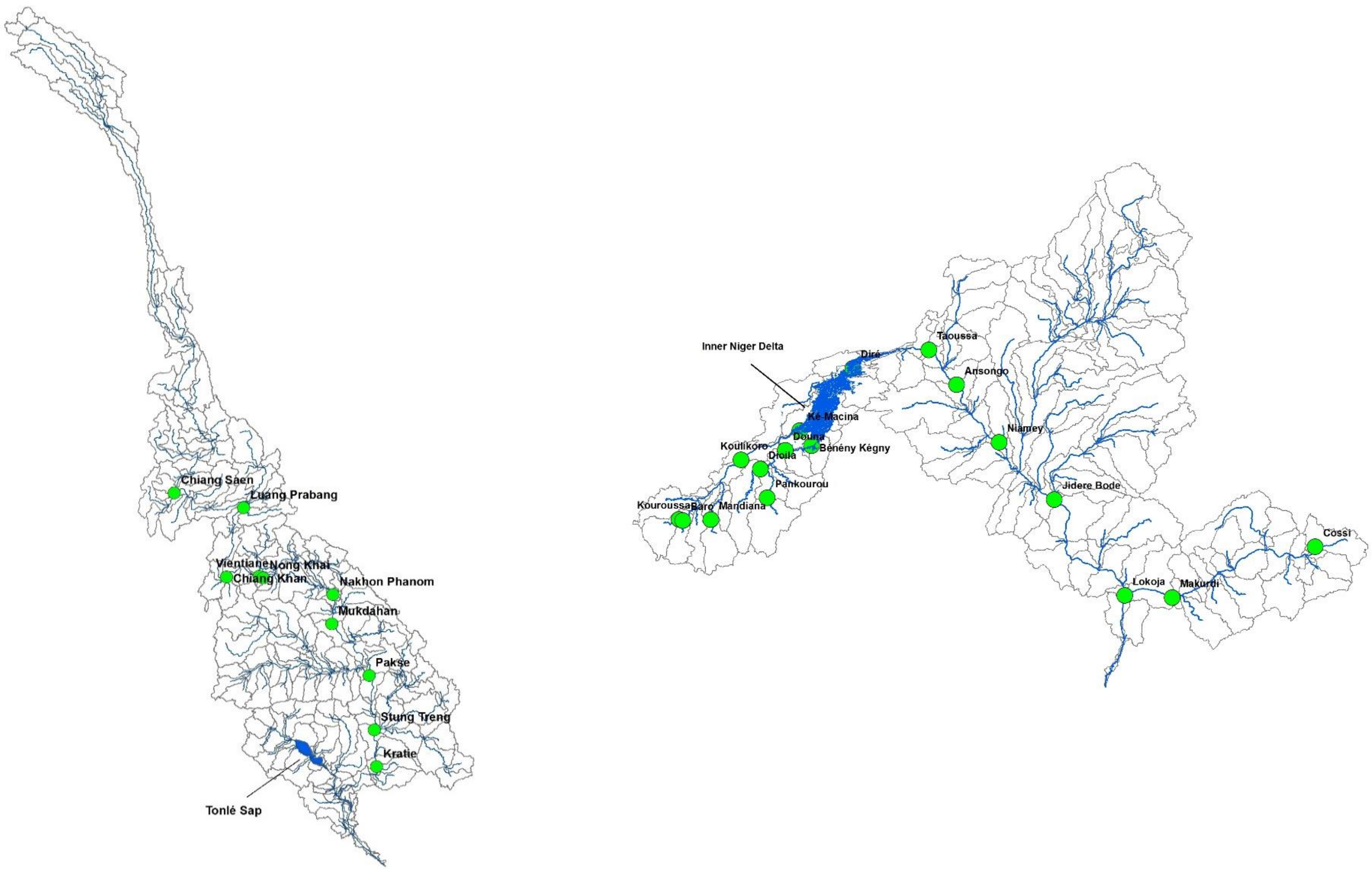

Figure S1 Watershed delineation schemes and locations of streamflow stations used in model calibration/validation

S4 Model Calibration and Validation

The SWAT-Mekong model was calibrated and validated using daily streamflow data from 10 gauging stations, while for the Niger River basin, model calibration and validation was conducted on a monthly basis. The data were obtained from L'Institut de recherche pour le développement (IRD), Niger Basin Authority (NBA) and Global Runoff Data Centre (GRDC). The calibration/validation periods and the model fits achieved by the SWAT model in both case studies are shown in Figures S2 and S3, and Table S3 (a) and (b).

Table S3: Nash–Sutcliffe model efficiency coefficient

Mekong

Station	Calibration (1983-1992)	Validation (1993-2007)
Chiang Saen	0.51	0.62
Luang Prabang	0.73	0.80
Chiang Khan	0.70	0.82
Vientiane	0.71	0.82
Nong Khai	0.74	0.82
Nakhon Phanom	0.80	0.84
Mukdahan	0.85	0.84
Pakse	0.82	0.85
Stung Treng	0.82	0.84
Kratie	0.83	0.85

Niger

Station	Calibration (1985-1994)	Validation (1995-2010)
Ansongo	0.88	0.50
Baro	0.80	0.33
Beneny Kegny	0.68	0.73
Cossi	0.81	0.08
Dioila	0.71	0.67
Dire	0.87	0.83
Douna	0.73	0.81
Jidere Bode	0.89	0.72
Koulikoro	0.92	0.72
Kouroussa	0.81	0.40
Ke Macina	0.88	0.66
Lokoja	0.86	0.72
Makurdi	0.81	0.87
Mandiana	0.65	0.42
Niamey	0.80	0.28
Pankourou	0.35	0.68
Taoussa	0.85	0.40

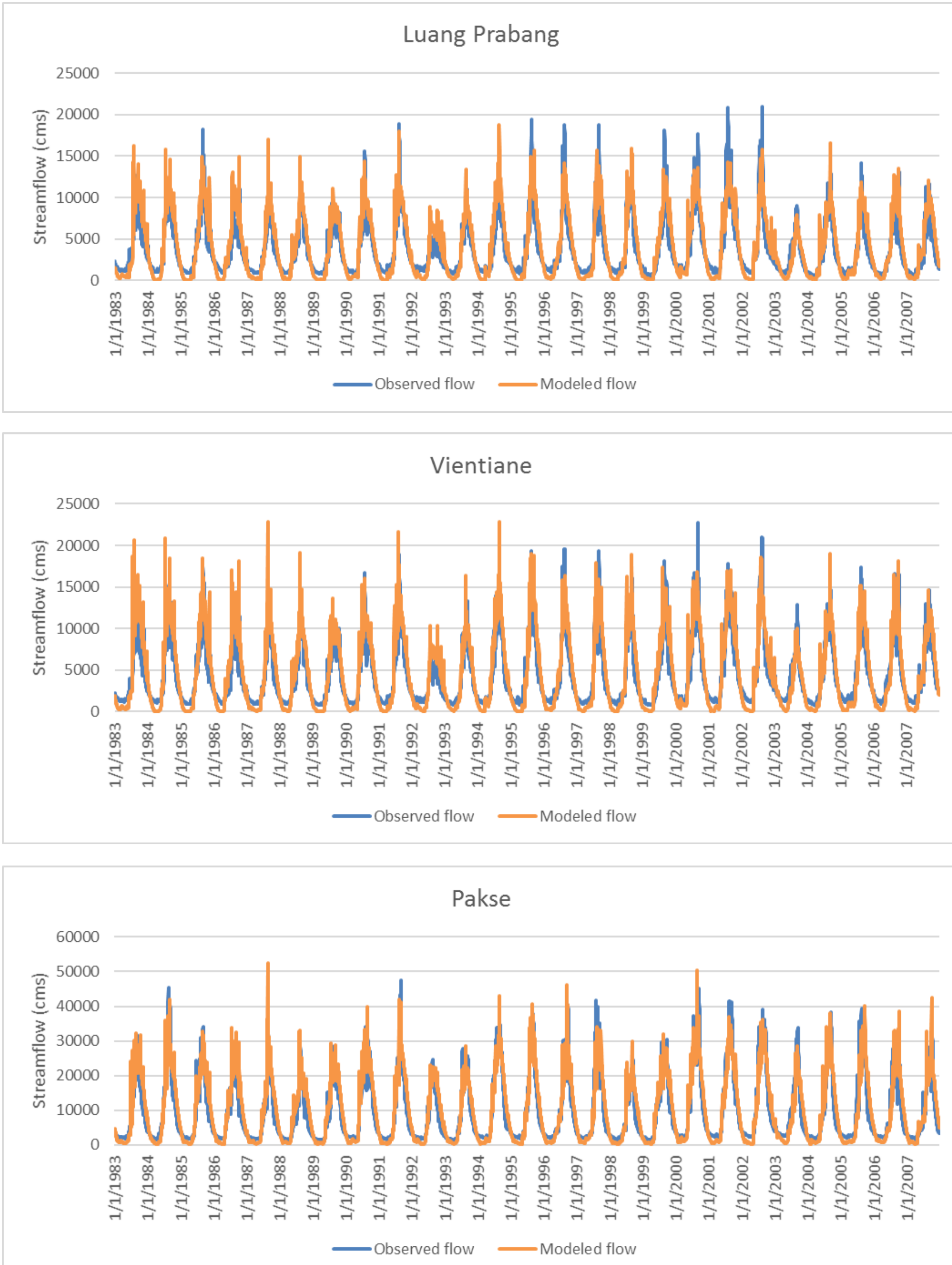

Figure S2: Simulated and observed streamflow at different locations along the Mekong River

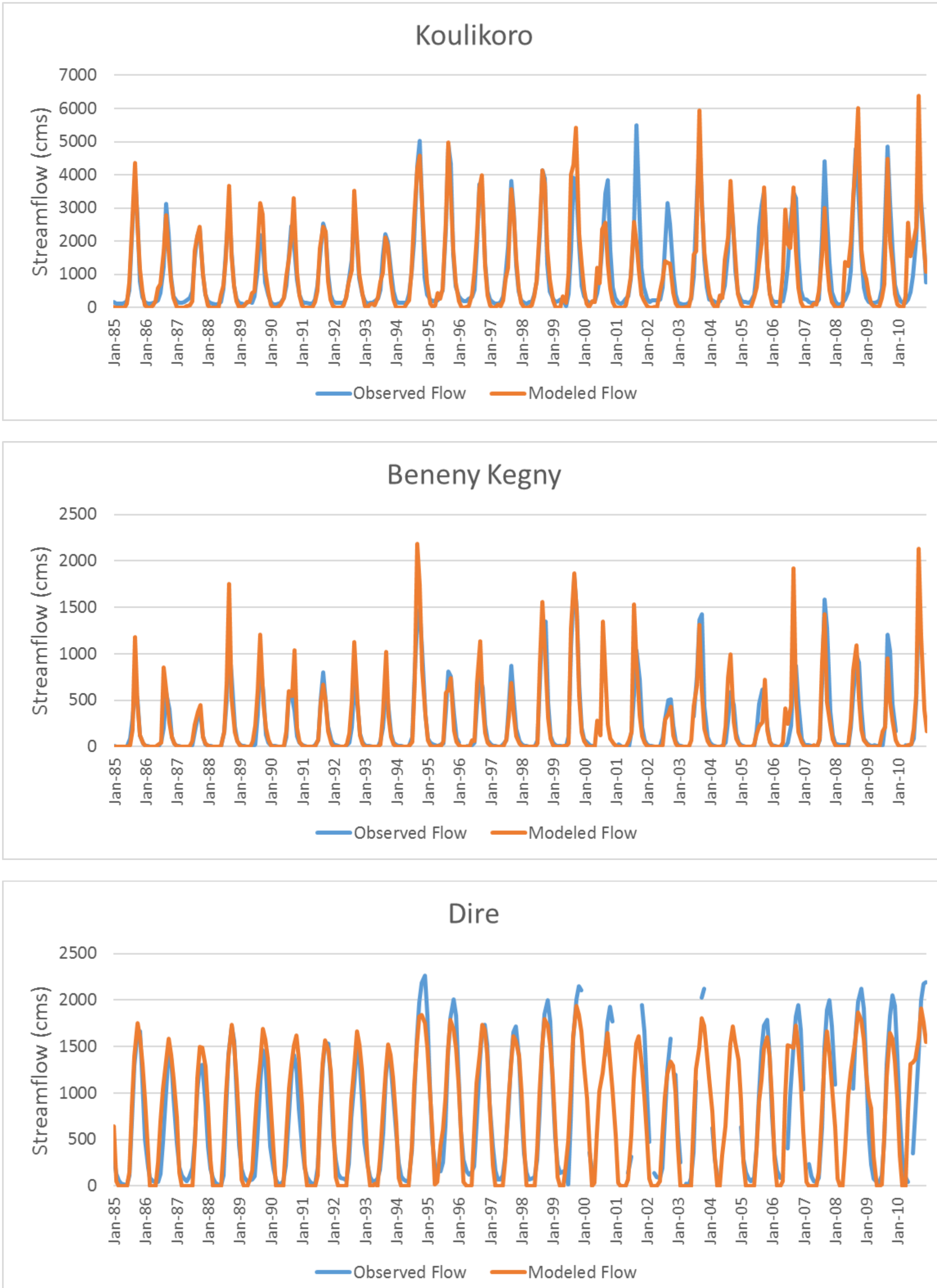

Figure S3: Simulated and observed streamflow at different locations along the Niger River

Similarly, more detail regarding the IHA-EFC data used would be welcome,

We have expanded the discussion of the ecosystem hotspots and the IHA-EFC parameters in paragraph beginning line 196.

An important input to the ABM is identification of ecosystem hotspots. Ecosystem hotspots are specific regions in the river basin that are especially critical to or indicative of the health of the ecosystem in the entire basin. Ecosystem hotspots can be identified in a variety of ways including through a literature review of critical ecological concerns in a basin and/or input from local ecological experts. For this analysis, for each ecosystem hotspot, relevant Indicators of Hydrologic Alteration (IHA) and Environmental Flow Component (EFC) parameters are selected based on expert opinion to measure ecosystem health (Richter et al., 1997, 1996). Baseline values for relevant IHA and EFC parameters, which are streamflow based indicators, are calculated from daily streamflow of the calibrated SWAT model. The IHA and EFC parameters included for the case study application described in Sect. 4 include monthly median flows, 7-day annual maximum flow, small and large flood event duration, timing and duration of extreme low flows etc.

as well as some more explanation as to the potential increase in pollution in the delta mentioned on line 424.

We have further clarified the discussion of the violation of ecological target as a case study for third party impacts. A flow magnitude related ecological target violation occurs, but it does not necessarily imply an increase in pollution as understood by the reviewer. We apologize for the confusion. The current coupled ABM modeling framework does not consider water quality as a driver for decisions or a management target. However, this aspect can be added as a suggested direction for future studies. The revision to the manuscript is shown below.

In particular, the ecological parameter seen to be violated is the IHA parameter for minimum average 7-day flow. Despite the increase in total annual flow due to the additional releases, the change in the flow timing leads to an ecologically inferior outcome for the Outlet Delta. This finding supports the argument that evaluations of ecological health performed at coarse time scales (e.g. annual) may overlook finer time-scale flow parameters that are critical to ecosystems (Palmer et al., 2005). In the absence of detailed data relating flow conditions to aquatic health in the Niger Outlet Delta, it is difficult to ascertain the exact impact that the violation of this target would have on the delta's ecosystem.

Palmer et al., 2005. Standards for ecologically successful river restoration, J. Appl. Ecol., 42(2), 208–217, doi:10.1111/j.1365-2664.2005.01004.x.

Technical corrections

- In the app, for crop yield, the y axis reads “Crop Yeild”
- p.6 l.134: “a level of cooperation (LOC) parameter is included that signifies by” “we include a level of cooperation (LOC) parameter that signifies”
- p.6 l.141: “These input parameters can either be defined by individual users tailored to their specific scenario of interest” by “These input parameters can either be defined by individual users according to specific scenarios of interest”
- p.7. l148: “is defined” by “are defined”
- p.7 l.149: “each of the agents” by “each agent”
- p.7 l.162: “in each agent” by “by each agent”
- p.10 l.217: “in the developing countries” by “in developing countries”
- p.10 l.218: “allow” by “allows”
- p.10 l.220: “the agents” by “agents”
- p.10 l.221: “requests by” by “requests from”
- p.21 l.431: “is conducted” by “was conducted”
- p. 22 l. 178: sentence lacks a verb

Thank you for your careful review. We have made the corrections indicated here in the revised manuscript.