

Response Letter

Dear Prof. Savenije, Dr. Xu, and Dr. Bogaard,

Thank you very much for your time and effort in reviewing the manuscript initially titled “Global water resources knowledge gaps”.

We have revised the manuscript title to “**Structural gaps of water resources knowledge in global river basins**” to better reflect the focus of our study.

We really appreciate all of your insightful comments. To fully address them, please see the revised version of the manuscript (a clear version and a track-changed version). As the manuscript has been substantially rewritten, we will not include all details but refer to specific sections of the manuscript by line numbers in this point-by-point responses to each of your comments below:

COMMENTS FROM PROF. SAVENIJE:

The authors have analysed papers categorised in the WoS under Water Resources and limited the analysis to articles that deal with river basins or catchments in a broad sense. They then looked for connections between disciplinary fields of WRM and analysed the connections between these fields and how they developed over time. They then classified the patterns of interconnection, or lack thereof into knowledge structures with the following names: Isolated, Innovative-inclined, Legacy-inclined and Centralised. To me, these classifications have hardly any explanatory power. I have gone through the description several times, but I fail to see what these terms actually mean or imply in relation to WRM. I can't see whether they have a positive or negative connotation. To me, Isolated and Centralised sounds rather negative; Innovative-inclined sounds positive; and Legacy-inclined may be both positive or negative, depending on one's perspective. That in recent years more basin studies are legacy-inclined may be evidenced by the data, but I have difficulty to see what it means.

Thank you for your comments. We have revised our method and data section to clarify the knowledge structure metrics and their significances (**Line 60-100**) and the revised knowledge structure types have been graphically illustrated in **Figure 1** of the revised manuscript.

*I do understand that this is a data-mining exercise, and that the authors did not necessarily familiarize themselves with the field of Water Resources Management and its development over time. I recommend looking at the paper on "Evolving water science in the anthropocene" (<https://hess.copernicus.org/articles/18/319/2014/>) and the huge body of papers that have recently been published under the IAHS research initiative "Panta Rhei", e.g. "Global perspectives on hydrology, society and change", *Hydrological Sciences Journal*, 61:7, 1174-1191, (DOI:10.1080/02626667.2016.1159308).*

Thank you very much for your comments. We have fully revised the introduction section of our manuscript to clarify the motivation of this study and review existing literature gaps that

can be addressed by our approach. The detailed introduction can be found in **Line 20-60** of the revised manuscript.

In addition, we have also revised the discussion section of our manuscript with the recommended references and additional ones to discuss the implications of our findings from the perspective of structurally complementing existing professional knowledge and research experiences (**Line 280-365**).

By taking river basins as the entree point, I fear the authors have missed a huge body of conceptual and global research. Not all WR research is done at river basin level. Much happens at the global scale, national scale, policy scale or conceptually.

Thank you for your comments. We have provided more details on the justifications why river basin was chosen as the spatial unit for analysis in the method section (**Line 100-115**). In addition, as explained in the method section, we merged those publications focusing smaller spatial units (e.g. sub-catchment, or wetland or lake into the river basin which they are affiliated with). We have also recognised the limitations in the river basin scale chosen by this study in the discussion section (**Line 355-360**), which may miss general conceptual/theoretical development without specific spatial links and those studies on global scales.

By choosing to analyse traditional disciplinary fields, such as: Agricultural irrigation; Erosion and sedimentation; Water pollution and treatment; Surface water and groundwater management; Ecological degradation; Droughts and floods; Climate variability and change; the results obtained are hardly pointing towards stronger societal linkages. I miss emerging new fields, such as: demand management; decentralisation; participation; international water law; and new technologies such as Remote Sensing, New observation technology, Global modelling, Artificial intelligence, If you look for traditional terms, you are bound to find traditional results.

Thank you for your comments. We have clarified in the method and data section how to retrieve and group key words with the classical bibliometric approach (**Line 120-140**). In addition, we have also re-examined and highlighted those newly appeared key words in each temporal stage to precisely reflect the evolution of management issues in the result section (**Line 225-250**). We have also included a table in **Appendix A** that lists the identified management issue groups and give examples for key words classified in each group. It should also be noted that the computer-based text-mining technique was adopted, so some key words might be missed.

Section 4. Discussion and Conclusion, is hardly a discussion. It is rather a set of three recommendations where certain lines of research are "encouraged": 1) investigation of "new" river basin phenomena; 2) spatial diversity of Water Resources research; and 3) strengthening collaborations with social sciences. These are rather obvious and general recommendations and hardly a discussion. The conclusion that "the stationarity of the water resources knowledge system persist" is not supported by the large body of work that is recently being produced as a result of and as part of the "Panta Rhei" initiative. This large body of work is hard to detect if one constrains oneself to the 95 most studied river basins in the world and the connections to traditional fields.

Thank you for your comments. We have substantially revised our discussion section based on the results and draw the conclusion from the perspective of knowledge structure of river basin. Detailed discussions can be found in **Line 280-370**.

COMMENTS FROM DR. XU:

General comments:

This work aimed to discover knowledge gaps in water resources research at the river basin scale through looking into the knowledge structure and disciplinary connections over time. The starting point of this paper is very interesting and the topic is important as river management and governance are highly fragmented. Generalizing knowledge patterns for research and management practices at the basin scale is challenging but should be done. Identification of knowledge gaps through investigating the knowledge structure is an innovative approach. Tracing the knowledge development patterns could also help identify gaps between science and policy, which is critical for the knowledge mobilization that promotes science-based decision-making for water systems. The synthesis of such fragmented knowledge would be benefited from large data analytics such as text mining approaches and content analysis. Text mining is an efficient way for the synthesis of knowledge which otherwise will be buried in the large number of texts.

This paper used academic literature obtained from the Web of Science as the main source and made use of a text-mining approach to extract key terms from the literature. The authors then used two indicators (degree and closeness) to measure connections among knowledge domains defined in this study. Overall, the methodology is designed in a reasonable manner and discussions are fair. However, some revisions are required to make it more readable and informative.

Thank you very much for your positive comments.

Knowledge structure is a keyword of the paper and it is a cognitive concept/science which needs to be carefully defined. It has been well defined in many other disciplines such as education, psychology, etc. What does it mean in water science at the basin scale?

Thank you for your comment. We have clarified the definition of knowledge in the context of the water resources discipline in this study in the introduction (**Line 25-50**) and method section 2.1 (**Line 60-70**).

It will be beneficial for the paper to list definitions of terms in a table (i.e., limited development, isolated development, innovative-inclined development, legacy-inclined development, centralised development). As these terms are not commonly used in the context of water sciences, nor is it in knowledge evolution, one might need to go back to read definitions a few times before he/she could understand and remember them. If they are new to the field, the authors should make them clearer to be understood. A diagram that distinguishes them from each other would be helpful as they are now ambiguous. Alternatively, the authors may need

to rephrase them into terms that are more common (e.g., “lack of knowledge”, “disciplinary”, “multidisciplinary”, “interdisciplinary”, “transdisciplinary”, etc. Tress et al., 2005. Clarifying integrative research concepts in landscape ecology. Landscape Ecology, 20, 247-493).

Thank you for your comments. We have reorganised our classification of knowledge structure into four types and clarify their implications on water resources knowledge development (**Line 60-100**) and graphically in **Figure 1** of the revised manuscript.

Reorganizing the methodology section is needed to make it easier to follow. In its current state, the section starts with definitions, which is fine, but the rest is discussed all around how the data was processed with methods inserted in the text. It will be better to split up section 2 into three sub-sections “definition”, “data” and “methods”.

Thank you very much for your comments. We have reorganised the methods and data section to ensure a more logical flow (**Line 60-165**). We added separate sub-sections in the method and data section that first defined the structure of the water resources knowledge system (2.1), followed by data sources (2.2), key words analysis (2.3), knowledge network analysis (2.4) and temporal periods division (2.5).

The discussion section would be valuable if some thoughts were put in ways to make water research more interdisciplinary than “isolated/centralised knowledge” as defined, for example how gaps identified could contribute to the framings of socio-hydrology, eco-hydrology, etc.

Thank you for your comments. We have revised our discussion section based on the key findings from our results from the perspective of knowledge structures of river basins. Detailed discussions can be found in **Line 280-370**.

The limitation of the paper should be acknowledged in some aspects. To be specific, the data for the knowledge synthesis does not cover grey literature which usually has reported management efforts that are not covered in academic papers. Papers that are not indexed in WoS could have also contributed to the field and be worth acknowledging. The absence of studies is not evidence of the absence of issues/development.

Thank you for your comment. We have discussed the limitations of our study in **Line 355-365**, in recognising the limited study scope of using only journal papers in WoS as well as other limitations.

Specific comments:

The authors may want to rename the title of the paper as it now does not cover the whole water resources system.

Thank you for your comments. We have revised our manuscript title as “Structural gaps of water resources knowledge in global river basins” to better reflect the focus of our study.

We provided details on the justifications about choosing river basin as the spatial unit for analysis in the method section (**Line 100-120**). We have also discussed the limitation of only choosing river basins in **Line 355-365**.

Section 2.1: using the availability of studies to define the knowledge status/gaps, in particular management of rivers, may not be appropriate as management practices could have been implemented to some river basins that have not drawn much academic attention. The absence of studies does not necessarily mean the absence of knowledge development for the basins. The authors should acknowledge its limitations.

Thank you for your comment. We chose academic publications as our data source as it provides systemic documentations of knowledge development across a broad range of disciplines. Large online publication databases enable consistent data retrieval for a long timeframe (**Line 100-115**). However, we do acknowledge that some river basins may receive fewer academic attention and focus more on practice-driven management. This limitation has been discussed in **Line 355-365**.

The authors used network indicators to measure knowledge connections. However, how the network was built is not well explained. What are nodes and links in the network are not clearly defined in the main text.

Thank you for your comments. We have added a separate sub-section in the method and data section to clarify the establishment of knowledge network (**Line 140-155**).

Section 2.2: First, using the keywords-based approach to retrieve records sometimes is controversial, because the results are significantly affected by the words selected for data collection. Some justifications of words selection should be added. Second, how groups of concerns were defined (i.e., agricultural irrigation, climate variability, etc.) and how each publication was classified into a specific group will need more explanations. For example, how studies on water policy were distinguished from management, how the overlaps were treated? What about studies of groundwater depletion and agricultural irrigation, were they included in agricultural irrigation or groundwater management? Some examples given may be helpful.

Thank you for your comments. We have clarified in the method and data section about how to retrieve and group key words using the classical bibliometric approach (**Line 120-140**). In addition, we have also re-examined and highlighted those newly appeared key words in each temporal stage to precisely reflect the evolution of management issues in the result section (**Line 225-250**). Finally, we have included a table in **Appendix A** that lists the identified management issue groups and give examples for key words classified in each group (**Line 375**).

Section 2.2: Which 5 basins, except for St Lawrence River basin, were removed? Justifications should be added to improve the robustness of data. St Lawrence River is a large river basin in

North America which connects to the Great Lakes Basin draining all the way up to the Atlantic Ocean. The drainage basin of ST Lawrence River has been ranked 13th largest in the world, providing millions of population and wildlife with water resources. A series of management strategies and actions have been planned since the 1980s, which have made significant progress on the protection of the ecohydrological systems of the basin. <https://www.planstlaurent.qc.ca/en/our-history>

Thank you for your comment. We have explained the reason why the five river basins including the ST Lawrence River were removed in the method section (**Line 100-115**).

Section 3: the total number of publications retrieved was not given in the text. Were all those publications included for the analysis or if any criteria were applied to clean the dataset?

Thank you for your comment. All publications after the filtering of key words were included and used for analysis. A total of 9128 publications from 1970-2017 were finally used for analysis. The details were added in **Line 165-170**.

Section 3, line 150: This would indicate that scientists started to focus on/realized synergistic impacts from water quality issues to ecosystems. Less previous studies do not mean that the impacts were not important.

Thank you for your comment. We have removed the results about synergistic impacts of management issues including this sentence.

Section 4: It would be good to separate discussion and conclusion sections.

Thank you for your comment. We have separated the discussion section (**Line 280-365**) and the conclusion section (**Line 365-370**) in our revised manuscript.

Yours sincerely,

The authors team:

Shuanglei Wu, Yongping Wei, Xuemei Wang