Dear Dr. Xu,

Thank you very much for your insightful comments on our manuscript. We would like to take this opportunity to address your comments by clarifying the unclear points and proposing how we will improve our manuscript.

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. Knowledge is typically recognised as a system. Scientific knowledge represents "ordered knowledge of phenomena and the rational study of the relations between the concepts in which those phenomena are expressed" (Dampier, 1944). Recently, scientific knowledge is increasingly recognised as a complex and dynamic network in which scientists, disciplines and phenomena to be "weaved together into an overarching scientific fabric" (Coccia, 2020; Shi et al., 2015). The complex interdependencies in the fabric are considered as the structure of knowledge. In our context, we define the knowledge structure in a river basin as a co-evolutionary process involving scientific disciplines and management issues which have their respective evolutionary dynamics. We will include more detailed definition of knowledge structure in the introduction section.

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. As we stated in our response to Professor Savenije, we will reorganise our classification of knowledge structure into four types using the two commonly used metrics in the system network theory: centrality and diversity (Figure 1).

Centrality measures the number of connection a node has in a knowledge network system, reflecting the level of knowledge concentration: the greater the centrality, the more connected a discipline is and thus more concentrated. Diversity measures the inverse sum of connecting distances to all other nodes, reflecting the extent to which a node is isolated within the knowledge system: the greater the diversity, the fewer extended connections a discipline has and thus forming more confined small groups in the network. Empirical analyses have demonstrated that concentrated knowledge structures facilitate dissemination of existing knowledge, whereas isolated structure can increase adaptivity to different disciplinary knowledge and facilitate radical innovations to knowledge development through looking from divergent angles (Bodin & Prell, 2011; Foray, 2018; Schot & Geels, 2008).

Based on the differences between the centrality and diversity values, we will classify the knowledge structure of river basins into four types of knowledge structures (Figure 1). They are:

1. Ideal structure with high centrality and high diversity. With this structure, the river basin should have high research intensities in core disciplines to provide solid theoretical foundations, while at the same time sufficient cross-disciplinary collaborations to ensure knowledge innovations to address unexpected, emerging river basin management challenges.
2. Innovation-inclined structure with high diversity but low centrality, which could have a risk of discipline hollowing-out (marginalization of influence of core discipline). For the river basins with this structure, the connection with core disciplines (centrality) should be strengthened.
3. Legacy-driven structure with high centrality and low diversity, which discourages knowledge innovation. In the river basin with this structure, the cross-disciplinary collaborations (diversity) should be strengthened to increase the potential of knowledge pattern transformation against emerging management challenges; and
4. Underdeveloped structure with low centrality and low diversity, indicating that the knowledge development is still at its early stage and the knowledge development should be strengthened comprehensively.

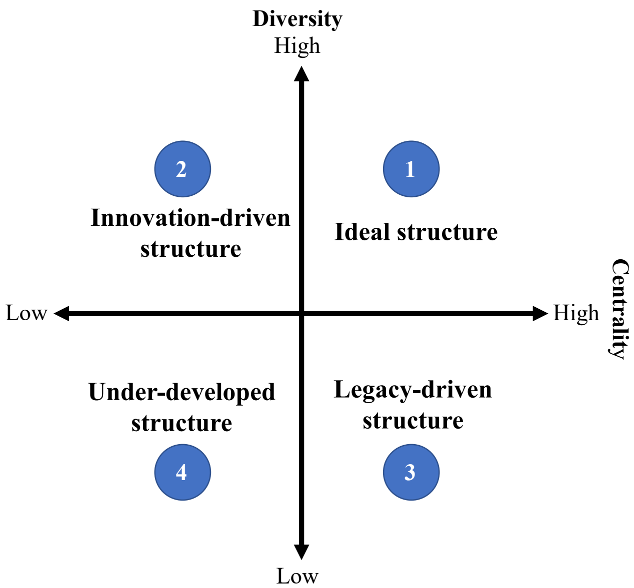

Figure 1 Four different knowledge structures based on their structural metrics

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 will include a separate section "Framework/Definition", then data sources and data processing, followed by our analysis methods.

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 will revise our key findings based on the results to be updated from our proposed revision of definition of knowledge structure of river basin. Then, we will discuss the implications of our findings on structural deficiency of water resources knowledges to complement to existing findings by professional knowledge and research experience in hydrology. For example, we will assess the development of socio-hydrology and eco-hydrology from the links of hydrology with other relevant disciplines, and recommend if these links should be strengthened and/or new links should be established according to both diversity and centrality at each river basin and all 95 river basins.

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 will discuss the limitations of our study including only journal papers in WoS as you pointed out.

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 comment. As we stated in our response to Professor Savenije, we chose river basin as the spatial unit for analysis as it represents the territorial unit of water cycle linking to other cycles of the Earth system (e.g. nutrients, energy, and carbon), which are commonly adopted by researchers to understand the integrated impacts of water use, land use and environmental management (Newson, 2008; Warner et al., 2008). 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). But we agree that we may have missed the publications on general conceptual/theoretical development without specific spatial links and those publications at global scale. Thus, we will revise our manuscript title as "Gaps of water resources knowledge structure in river basins".

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 systematic documentations of knowledge development across a broad range of disciplines. Large online publication databases enable consistent data retrieval for a long timeframe. However, we do recognise that some river basins may receive fewer academic attention and focus more on practice-driven management. We will acknowledge this limitation in the discussion section.

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. The network connections were established based on the co-occurrence principle. Two disciplines were connected if they were linked to the same key word in a publication; and two management issues were connected if they appeared in the same publication. As a result, the disciplinary network contained different disciplines as nodes, whereas the management issue network contained issues as nodes. The weighting of links were the number of publications. We will include these details in the method section.

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. First, key words have widely been used to express the research topics of articles and considered a basic element in understanding the content and structure of disciplinary knowledge (Cheng et al., 2020; Khasseh et al., 2017). In addition, we have used **all** key words retrieved in the title, abstract and key words sections of each article after natural language processing. We will make these clearer when we revise our manuscript. Second, the nine groups of key words we identified were derived from our data rather than pre-set. Grouping was based on broadly recognised river basin management concerns and the processes underlying them. In addition, as we stated in our response to Professor Savenije, we will re-examine those newly appeared key words in each temporal stage and may categorise them in newly defined groups to more precisely reflect the evolution of management issues. To keep consistency of key words grouping, the two independent coders who did the grouping will be asked to code the key words with any ambiguity thoroughly discussed. Finally, we will include a table in the methods section to list our identified key words groups and give examples for each group. Surface water and groundwater management referred to the general water resources management issue (e.g "water level fluctuations", "drainage"); Water policy refer to the specific policy initiatives and instruments (e.g. "flood management", "integrated management"). All groundwater related issues were grouped into surface water and groundwater management. Agricultural irrigation referred to the specific irrigation methods and techniques. The word co-occurrence in the same publication was used to indicate the relationships between agricultural irrigation and groundwater depletion.

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 removed five river basins: the Lawrence River, the St. Lawrence River, the Red River, the Lena River and the Missouri River as the retrieved river names could not differentiate the rivers (i.e. the Lawrence River and the St. Lawrence River), the retrieved river name did not have clear connection with specific location (i.e. the Red River), and the retrieved river name did not have full-length data during the study period (i.e. the Lena River and the Missouri River). We will acknowledge this in more details in the method section.

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 retrieved and used for analysis. A total of 9128 publications from 1970-2017 were finally used for analysis. The publications per year will be attached as supplementary materials in the edited manuscript.

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 will revise this sentence as "the interactive impacts between water quality and ecosystem degradations have been a major focus of scientists during our study period".

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

Thank you for your comment. We will separate the discussion section (as briefly outlined above) and the conclusion section (briefly summarise the key findings and implications) in our revised manuscript.

Yours sincerely,

The authors team:

Shuanglei Wu, Yongping Wei, Xuemei Wang

References

Bodin, Ö., & Prell, C. 2011. *Social Networks and Natural Resource Management : Uncovering the Social Fabric of Environmental Governance*. Cambridge, UNITED KINGDOM: Cambridge University Press.

Cheng, Q., Wang, J., Lu, W., Huang, Y., & Bu, Y. 2020. Keyword-citation-keyword network: a new perspective of discipline knowledge structure analysis. *Scientometrics, 124*(3), 1923-1943. doi:10.1007/s11192-020-03576-5

Coccia, M. 2020. The evolution of scientific disciplines in applied sciences: dynamics and empirical properties of experimental physics. *Scientometrics, 124*(1), 451-487. doi:10.1007/s11192-020-03464-y

Dampier, W. C. 1944. *A shorter history of science*: Cambridge University Press.

Foray, D. 2018. Smart specialization strategies as a case of mission-oriented policy—a case study on the emergence of new policy practices. *Industrial and Corporate Change, 27*(5), 817-832. doi:https://doi.org/10.1093/icc/dty030

Khasseh, A. A., Soheili, F., Moghaddam, H. S., & Chelak, A. M. 2017. Intellectual structure of knowledge in iMetrics: A co-word analysis. *Information Processing & Management, 53*(3), 705-720. doi:https://doi.org/10.1016/j.ipm.2017.02.001

Newson, M. 2008. *Land, water and development: sustainable and adaptive management of rivers*: Routledge.

Schot, J., & Geels, F. W. 2008. Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technology Analysis & Strategic Management, 20*(5), 537-554. doi:10.1080/09537320802292651

Shi, F., Foster, J. G., & Evans, J. A. 2015. Weaving the fabric of science: Dynamic network models of science's unfolding structure. *Social Networks, 43*, 73-85. doi:https://doi.org/10.1016/j.socnet.2015.02.006

Warner, J., Wester, P., & Bolding, A. 2008. Going with the flow: river basins as the natural units for water management? *Water Policy, 10*(S2), 121-138. doi:10.2166/wp.2008.210