

Comment on hess-2021-137

Li Xu (Referee)

Referee comment on "Structural gaps of water resources knowledge in global river basins"
by Shuanglei Wu et al., Hydrol. Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/hess-2021-137-RC2>, 2021

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.

- 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?
- 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).

- 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".
- 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.
- 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.

Specific comments

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

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