

Similar to clustering studies that meticulously compare cluster patterns, the paper is at times hard to follow for the reader, due to the tedious nature of going through detailed dissemination of specific patterns found in data. While this reflects the subjective impression of me as a reader, from the standpoint of the reviewer I can attest that the paper has a clear structure and exhibits straight and deep thinking about governing processes that is valuable for publication.

Its scientific significance is good: There are numerous studies tracing streamflow and groundwater time series patterns to overall governing processes – but few compare streamflow and groundwater side-by-side.

The scientific quality is fair, but can be good once the main general comment below addressing selective reporting of results is addressed.

Presentation quality is good and could be excellent if language and reasoning would be more concise and on point, and more aware of clearly communicating complicated mental concepts and connections instead of assuming the reader being “in the know” – at least in some parts of the results section. However, I would not draw necessity for improvement. It is good enough.

General comments

There is a fundamental problem of selective use of analysis methods in this paper: for example, the author does not report correlations of PC2-6 with raw time series means, but only the correlation of PC1 raw time series means is reported (line 173). As another example goes the analysis of autocorrelation (only PC2 & PC4, Figure 3 & Figure 6) or correlation with trends (only PC2, Figure 4) or the reporting of distribution of loadings w.r.t. aquifer type (only PC4, figure 6) or overall the selective reporting of correlations to individual catchment attributes throughout the paper, and several other examples. This is to say that in this paper, the (shown) application of methods is selective, and while selective reporting makes for a good storytelling, the conclusions drawn in the study (PC1 captures “mean behavior”, PC2 the “dampening” etc.) can only be drawn from comparative analysis, e.g. when correlation with time series mean is much better for PC1 than for PC2-6, and analogously for the other examples. It may or may not be that comparative analysis has been done in the background and only the significant results are shown (line 230 indicates that – stating that autocorrelation “not shown” for PC3). But to make sure that the results are not selective constructions, uniform method application and reporting across all PCs and is required. For the most part, at best in the appendix, to not blow up the paper, although some referencing in the main text of the manuscript will be needed. To be clear, I am not suggesting scientific misconduct. The displayed results seem to be reasonable, however for reasons of scientific rigor and for the reader to be able to reconstruct the conclusions properly, the paper simply needs verification via negative elimination through comparative analysis for reasons of completeness.

Minor comments

Introduction section

Line 33: WFD citation missing; maybe add half a sentence of explanation what WFD is for intercontinental audience; explain that for WFD the definition of water bodies includes GW

Line 33-45: In addition to citations to studies diagnosing the pitfalls of heterogeneities in groundwater, the author is advised to include citations to papers that also try to grasp it, e.g. this Gothenburg-based research group – but happily also others:

- Barthel, R., Haaf, E., Giese, M., Nygren, M., Heudorfer, B. and Stahl, K., 2021. Similarity-based approaches in hydrogeology: proposal of a new concept for data-scarce groundwater resource characterization and prediction. *Hydrogeology Journal*, 29(5), pp.1693-1709. <https://doi.org/10.1007/s10040-021-02358-4>
- Giese, M., Haaf, E., Heudorfer, B. and Barthel, R., 2020. Comparative hydrogeology–reference analysis of groundwater dynamics from neighbouring observation wells. *Hydrological Sciences Journal*, 65(10), pp.1685-1706. <https://doi.org/10.1080/02626667.2020.1762888>
- Haaf, E., Giese, M., Heudorfer, B., Stahl, K., & Barthel, R. (2020). Physiographic and climatic controls on regional groundwater dynamics. *Water Resources Research*, 56, e2019WR026545. <https://doi.org/10.1029/2019WR026545>

Line 79-80: for signature-based analyses of hydrologic similarity, essential (and more suitable) citations that need to be included to connect the reader to the large body of literature of earlier decades on hydrologic similarity, are the following; they are also good evidence for the bold claim that signatures are “hardly ever checked for relevance”.

- Olden, J. D., & Poff, N. L. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications*, 19(2), 101–121. <https://doi.org/10.1002/rra.700>
- Olden, J.D., Kennard, M.J. and Pusey, B.J., 2012. A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. *Ecohydrology*, 5(4), pp.503-518. <https://doi.org/10.1002/eco.251>

Line 68-71: This attempt can brilliantly be motivated based on literature as well, e.g. by this older call to interdisciplinary studies on hydro(-geo)logical similarity:

- Barthel, R., 2014. HESS Opinions" Integration of groundwater and surface water research: an interdisciplinary problem?". *Hydrology and Earth System Sciences*, 18(7), pp.2615-2628. <https://doi.org/10.5194/hess-18-2615-2014>

Line 78-83: Maybe reiterate the method of choice (PCA) here.

Data section

Line 101-114: Would be informative to state the total number of samples and total number or share of interpolated samples that you had in there in the end.

Line 108: which analogous approach? Standard? Or citation? Unknown to me.

Method section

Line 130: “it has rarely been used in hydrology” -> that is simply untrue, as per the very large body of literature that opens up to the reader once they trace the literature starting from the Olden 2003/2012 paper provided above.

Line 140-143: Belongs to data section.

Line 145: “unlike in some other studies”: cannot be stated like this without citations. Citation or rephrasing necessary.

Results section

See general comment about selective vs. comparative analysis

Discussion section

Reading through this discussion, I can only double my suggestion above to read the works of the referenced Gothenburg group, especially since they use data from the adjacent region. Haaf 2020 can be brilliantly connected to the overall discussion points, and the Giese 2021 paper to the subsurface section 5.3 especially, which is lacking discussion with literature overall. Also the Barthel 2014 paper belongs next to the Berkowitz and Zehe 2020 in line 444. Crazy how people who write about the same can be completely unaware of each other (Berkowitz/Zehe and Barthel, that is).