Referee 1 Comments – Authors’ responses:

This is an interesting study that I suspect will be of wide interest to ESD readers. The premise for the research is straightforward and clearly articulated, with a generally well-reasoned account of methods and presentation of findings. I have some relatively minor points to consider for a redrafted version of this manuscript, as follows:

R1 : We appreciate the positive comments regarding the overall manuscript and the constructive suggestions to improve the manuscript. Responses to individual comments are given below.

There are quite a large number of acronyms used in the manuscript. Not all of these
seem essential, and in places they detract from the overall readability. Some are rarely used (e.g. COV, PCC), others are arguably not really necessary (PCP for precipitation). I recommend the authors rethink and reduce the number of acronyms accordingly.

R2 : We will seek to reduce the number of acronyms in the final manuscript including some of the suggested that are rarely used. Use of PCP is expected to remain as we apply this in the context of identifying both PCP and PCPA (anomalies) and consider these to be clear, sensible abbreviations.

Inconsistent capitalisation of nouns. Some terms that are converted to acronyms are not capitalised (e.g. global hydrological models, GHMs – p2, line 52), whereas others are (Groundwater Storage, ∆GWS).

R3 : This comment is helpful in identifying necessary corrections that we will make in the course of the review of acronyms.

P2, line 61-62. Presumably this increase of 15%/decade is a global average?

R4 : We will amend to clarify the text as ‘...rapid escalation in global groundwater extraction at an average of ∼15% per decade...’

P7, lines 168 and 189. It is potentially somewhat misleading to provide coefficient of variation statistics for sample sizes of 3 and 4. The mean, standard deviation (and thus the coefficient of variation too) are not particularly informative descriptors of central tendency or distribution of samples that are this small.

R5 : We acknowledge that the population (not sample) size is small for the use of coefficient of variation. We nevertheless consider this parameter has merit in explaining divergence among the respective datasets; we propose to include a phrase acknowledging the small population sizes.

P7, line 192. Improper use of the term “significant” when discussing the magnitude of statistical terms. From my reading of this sentence, “substantial” would be a more appropriate descriptor.
R6: We agree and will amend to read ‘substantial’ and check all other uses of the word, ‘significant’ (3 instances), are appropriate.

Figure 3. The different y-axis scales, and scaling of the data are noted in the caption, but this twin approach to scaling, plus the relatively small size of the panes and faint axis labels makes it somewhat difficult to compare scales of variability between the different aquifers.

R7: We appreciate this comment and will look at ways to improve the clarity of what is being expressed in Figure 3 and improve axis labels. As the primary purpose of the figure is to illustrate correlation between change in GWS and precipitation anomaly for individual systems, we have scaled the data to make these correlations clear. Further information on the variability between aquifers is given in the final two columns of Table 1.

P16, line 400. The reference for this statement is slightly misleading. It could be read that Zwiers and Von Storch (1995) found robust statistical significance, but actually I think the authors mean that robust statistical significance was found using the methods described by Zwiers and Von Storch.

R8: We agree with that this clarification is necessary and will amend to state: ‘when tested using the methods described by Zwiers and Von Storch (ref.)’