Reply to referee 1 – Report #1 from 2025-09-15

Comments of referee 1 are in black.

The comments from the online report mask and those annotated in the pdf of the referee report were identical. The bold highlighting in some comments were taken from the pdf referee report.

The referee structured the comments in three sections. To simplify cross-referencing, the authors added the section titles: A - Introductory comments, B - Comments on the last set of author replies, and C - Comments on the last version of the manuscript; and added the respective capital letter to the enumeration of the comments. Thus, for example comment B-3 refers to the third comment in section B.

Replies of the authors (AR) are in blue.

Comments of referee 1

Review of latest revision of Technical Note: An illustrative introduction to the domain dependence of spatial Principal Component patterns by Lehr and Hohenbrink.

A - Introductory comments

I appreciate all of the work put into this manuscript. Given the latest responses to the last revision, I believe that the present manuscript is about as complete as I can hope for, with the exception of cleaning up and further explaining a few issues, mostly in the section on rotation. Given the hesitation to investigate fully the domain effects on rotated solutions, I believe there are three good options. 1. Remove the incomplete rotation section. 2. Leave the very limited experiments using the single algorithm (Varimax) over a limited number of PCs AND issue appropriate caveats, rather than the vague sentences in this last version (see my comments in the annotated manuscript). 3. Complete a comprehensive section of rotation using several algorithms and testing for the optimal number of PCs to retain and test those for DD and for validity. I believe that options 1 and 2 are more viable given the comments in the last response about the manuscript size and your intended scope. That is fine as there are other paths to finalize this manuscript.

AR: Thank you very much for your positive evaluation and appreciation. We likewise appreciate all the work you spend for reviewing and improving our manuscript. We also see mainly the three options you are pointing out and agree that options 1 and 2 are more suitable for our scope. We decided to go for option 2. See also our further replies.

The present version could stand as an important contribution to the literature, with appropriate caveats issued. If you wish to proceed to a larger manuscript, it is important to be fully prepared to investigate all the possible permutations arising from each step when rotated solutions are investigated. Those permutations would involve running a much more detailed set of experiments for the number of EOFs/PCs to retain, as that number now effects the pattern morphology (unlike the unrotated solutions where increasing the number retained adds a new pattern but does not change the previous ones), assessing the degree of simple structure in their data sets and then selecting the optimal rotational algorithm from a sizable number of algorithms (again, this step requires validating many sets of patterns to the largest correspondence to the similarity matrix from which those patterns were derived). Given all the domain shapes, such a set of experiments is cumbersome so, unless you are really motivated to go that route, I don't feel comfortable prescribing that as the required bar and deem the present manuscript is acceptable after the issuance of some caveats. Specifically, issue caveats that the analyses presented for Varimax have no way to deconvolute the effects of keeping too few PCs or too many PCs from DD and temper comments such as "the Varimax way of displaying DD" (lines 751-752). For all we know, it has nothing to do with Varimax but is a function of keeping too few PCs and forcing multiple correlation modes onto too few PCs. Acknowledge the possibility. Not investigating it is okay but, then, either don't mention it or list it as one of a number of possibilities.

AR: Thank you for the positive assessment of our work and providing context for the caveats you suggest. The outline of the experimental study you are suggesting is what we meant by “full-scale rotation study” in our last replies (see also Lines 758–759). We agree that such work would be cumbersome, but valuable addition. Given the extent of the work we find it more suitable as standalone future study. Correspondingly, and in line with your suggested option 2 above, we issued several caveats in the revised manuscript (see our replies to your further comments).

In particular, we included (i) a caveat that the presented analyses did not cover the effects of keeping too few or too many PCs (Lines 761–766), (ii) acknowledged the possibility that multiple correlation modes were forced onto too few PCs (Lines 761–763) and (iii) deleted the phrasing “Varimax way of displaying DD”. For the details regarding these points, please see our replies to your comments C-8 to C-12.

B - Comments on the last set of author replies

Here are my comments on the last set of replies:

B-1. “If PCA is used purely for data reduction, DD is of no interest as the patterns are never examined; they serve only as an efficient set of basis vectors. If, however, the subsequent use of the PCs requires an adequate description of multiplet subspaces, for example if PCA is used as preprocessing step for other analyses, care should be taken that no multiplet is split by the selection of retained PCs.”

Response: I think as a general statement that is solid advice. However, there are at least two other considerations:

(a) Small magnitude eigenvalues are thought to be associated with noise. Rarely, if ever, if that assumption tested. Small scale signals that extract little variance would be indistinguishable from noise through investigating eigenvalue magnitude, particularly without resampling (there is no resampling in the North et al. test).

(b) Ignoring point (a), the effects of the infusion of noise into the signal+noise PC patterns may or may not be an issue depending upon the amount of variance extracted by those retained. For example, if the retained PCs account for 95% of the total variance, the remaining 5% would be expected to have a minimal effect on the patterns retained, particularly if rotated. In contrast, if that percent variance were 60%, there exists a larger risk of rotation being effected by the noise introduced.

That said, the conservative advice may be safe as it should prevent problems regardless of the total amount of residual noise or if some signal is being discarded.

AR: We addressed your point (b) by extending the sentence in Lines 561–563 to

“In particular, special care has to be taken that the truncation point of a PCA does not split a multiplet (North et al., 1982), especially when the amount of variance associated with the excluded PCs is relatively large compared to the amount of variance extracted by those retained.”

and added a sentence at the end of the second paragraph in the conclusion (Lines 826–827):

“This applies in particular when the amount of variance associated with the excluded PCs is relatively large compared to the amount of variance extracted by those retained.”

B-2. “If the spatial PC patterns do not differ significantly from DD reference patterns, we recommend to report that and stop any interpretation of individual spatial PC patterns as distinct hydrological features.”

Comment: This is good advice, though providing some thoughts on "significantly" or how to assess that would be helpful.

AR: We added a specification in the first sentence of the paragraph from the addressed sentence (Lines 833–834) to clarify that we mean significant differences in the sense of statistical testing here:

“To test whether spatial PC patterns differ significantly from DD patterns, reference patterns can be used as null hypothesis in pairwise statistical tests.“

A similar specification was added in the Methods Section 3.2, please see the revised manuscript in Line 240. In our study, we used simple t-tests (see the last sentence of Section 3.3).

B-3. "What we are suggesting is that if spatial PC patterns are used for interpretation, the patterns should be checked for DD before. We state this in Lines 803–806 of the conclusions. We recommend “visual comparison of the spatial PC patterns from subdomains with markedly different shapes and/or sizes … as quick qualitative check” (end of 3rd paragraph of the conclusion). And we recommend DD reference patterns as null hypothesis to “test whether spatial PC patterns differ significantly from DD patterns” (beginning of the 4th paragraph of the conclusion)."

Comment: Yes, this is good advice as the first step. In practice that null hypothesis would have an alternative that the patterns are different (the alternative hypothesis), a statistical test is than brought to bear on the sample pattern and the probability of a Type I error arises is produced (perhaps with a yes/no decision). One possibility (or many) to avoid a purely subjective determination would be to relate each EOF/PC pattern to the patterns derived in your manuscript for the best matching domain shape. One could apply the aCC statistic and (because it arises from an unknown distribution) apply a permutation test to determine the p-value.

AR: In principle yes, but there is one important difference. What we suggest here is not to compare the PC patterns from other analysis with the sample patterns derived in our manuscript for the best matching domain shape. Instead, we suggest to fit DD reference patterns to the specific spatial domain that is analysed because "DD patterns are original for every combination of spatial domain and spatial correlation properties" (Lines 112–113, 830–831). We present two methods to do so (fifth paragraph of the conclusion). Then a statistical test can be used to assess the PC patterns for significant differences with the specifically fitted DD reference patterns. We agree that applying a permutation test to determine the p-value of the significance test would be one of the possibilities. Here, we used simple t-tests.

B-4. "AR: We highly appreciate the sincereness and precision of your comments and the amount of work you spent. It considerably helped to improve our manuscript, in particular with respect to completeness, precision and explicit statements. However, some of your suggestions point conceptually to a different direction than it is our intention for our study."

Comment: That is a fair statement. To run a proper set of DD tests on rotated solutions would require being fully prepared to investigate all the possible permutations arising from each step when rotated solutions are investigated. Those permutations would involve running a much more detailed set of experiments for the number of EOFs/PCs to retain (k), as that number now effects the pattern morphology (unlike the unrotated solutions where increasing the number retained adds a new pattern but does not change the previous ones), assessing the degree of simple structure in their data sets and then selecting the optimal rotational algorithm from a sizable number of algorithms (again, this step requires validating many sets of patterns to the largest correspondence to the similarity matrix from which those patterns were derived). Given all their domain shapes, such a set of experiments is cumbersome. Perhaps list this as an area of investigation for others or for a future paper?

AR: We agree. We provided a suggestion for future research in the last paragraph of Section 5.2.2. Compared to the previous version, we slightly extended the wording at the end of the paragraph to:

"… The experiments could be used to compare the performance in hydrological feature identification of unrotated versus rotated PCA, different orthogonal and oblique rotation methods, but also of S-mode versus T-mode PCA (Richman, 1986; Compagnucci and Richman, 2006; Isaak, et al., 2018) and different scaling of the eigenvectors (Jolliffe, 1995; Wilks, 2006)."

Our suggestion is revisited in prominent position as final sentence of the conclusion:

"Building on this study, a next research task could be to conduct systematic experiments with synthetic test data derived from hydrological simulation models to evaluate which PCA modes, rotation methods and scaling of the eigenvectors work best for hydrological feature identification."

Given the above statements it seems there are three possibilities to satisfy Occam's razor in two broad categories:

1. Pure data reduction with no physical interpretation: keep the process simplest and don't bother with DD investigations or with post processing with rotation. That is now discussed in the manuscript.

2. Data reduction with physical interpretation:

a. No longer can the simplest process be assumed to work without further investigation, so test for DD. If DD is present in a "significant" amount, either stop, try step 2b or try a methodology other than EOF/PCA.

b. Investigate if the patterns reflect the patterns of variation in the similarity matrix to ensure physically valid patterns.

(i) If there is not a significant amount of DD present (that is why we are at step 2b) but there is a "significant" correspondence for each unrotated pattern retained to those in the similarity matrix, then interpret each pattern in the set of k patterns or try 2biii to determine if the matches become more significantly matched to the similarity matrix.

(ii) If there is not a significant amount of DD present (again that is why we are at step 2b), but there is not a "significant" correspondence for every unrotated pattern retained to those in the similarity matrix, then (a) reduce the number of EOF/PCs retained and retest each in the set of patterns for correspondence to the similarity matrix or (b) try 2biii to determine if the matches become more significantly matched to the similarity matrix patterns.

(iii) If there is not a "significant" correspondence for every unrotated pattern retained for any set of k EOFs/PCs retained, or one is interested in determining if the matches are improved by post-processing the unrotated EOF/PC patterns, then try rotating the patterns and assessing those rotated patterns for a "significant" correspondence for each pattern retained to those in the similarity matrix, then interpret each of the rotated patterns.

(iv) If either step 2bii or step 2biii fails to find a set of k PCs where each shows insignificant amount of DD and each is significantly related to the patterns on the similarity matrix, then try a different non-EOF/PC approach or stop.

I believe some comments on this could be added to Lines 774–780 and at the end of the conclusions.

AR: Thank you for summarizing the possibilities you see for including the aspect of DD in a PCA in a decision tree. We agree, that this often helps to clearly see the available possibilities. However, in the decision tree you are suggesting the assessment of the physical validity of the spatial PC patterns (2b) is an integral part. We did not cover this here. Given your earlier comments (second part of the introductory comment A, first part of this comment here), we think that you agree that including such work is material for a standalone manuscript and not required here. Therefore, we think that the remainder of the content you are summarizing is already explicit and clear enough in our manuscript. Thus, instead of repeating that content and adding another paragraph in the conclusion, we rather like to keep it as it is.

We think your point 1 is clearly covered in the conclusion in Lines 823–824:

“If PCA is used purely for data reduction, DD is of no interest as the patterns are never examined; they serve only as an efficient set of basis vectors.”

Focusing on what we presented in our manuscript here, that is, leaving aside the test for physically valid patterns (your matches with the similarity matrix), the aspects of your point 2 are covered in Lines 820–823:

“Consequently, DD should be considered for any application in which the PCs are used to draw conclusions about spatially distinct properties of the analysed system. In other words, it should be checked whether the spatial PC patterns differ significantly from patterns that result from the trivial case of nearby locations being homogeneously more related than those further apart.”

and in Lines 834–836:

“If the spatial PC patterns do not differ significantly from DD reference patterns, we recommend to report that and stop any interpretation of individual spatial PC patterns as distinct hydrological features.”

The possibility to apply another method than PCA, we did not explicitly state. We assume that this is obvious to the reader.

B-5. "Table 4 provides the comparison of the spatial PC patterns from the precipitation data (Figure 16) and the corresponding DD reference patterns (Figure S7). The reasoning is to show how well the patterns of the precipitation data match with those of the DD reference. The DD reference patterns here and throughout the paper are calculated for unrotated correlation matrix based PCs. An equivalent comparison with varimax rotated PCs would require to calculate DD reference patterns for varimax rotated PCs, depending on the k number of PCs selected for rotation."

Comment: Because the analysis of precipitation or any other field cannot be assumed to be well-rendered by Varimax, such an experiment, if it were made, should examine a range of rotations. Investigation of various Python and R libraries (e.g., the R package GPArotation) suggest dozens of possibilities (other than Varimax).

AR: Thank you for this addition. We understand it as a valuable hint in the context of future studies.

B-6. "To our knowledge, calculating DD reference patterns for varimax rotated PCs has not been done before. It could be an interesting objective in a future study focussing on rotated PCA (and physical interpretation of PC patterns)."

Comment: See the comment above where eliminating DD or reducing DD to an "insignificant" amount is a necessary but not sufficient step for physical interpretation. To meet the more stringent level of sufficient for physical interpretation, each of the retained EOF/PC patterns must reflect the correlation (or covariance) patterns well. However, because that sufficiency applies to both unrotated and rotated EOFs/PCs, some mention of that could be made because it is possible a set of patterns may not show DD but may not reflect well the patterns of data similarity (e.g., recall the heavy lift issues of maximum variance and orthogonality that are perhaps related to but not the same as DD).

AR: We agree. We have statements on this in the last paragraph of the conclusion (Lines 843–848):

“Passing the check for DD and accounting for effective multiplets in the selection of the PCs are necessary but not sufficient conditions to assure physical meaningfulness. When single PCs, or combinations of PCs, are assigned to distinct hydrological features, it should be carefully considered whether the S-mode PCA constraints (i) successive maximization of variance on the PCs, (ii) orthogonality of spatial PC patterns and (iii) linear uncorrelatedness of temporal PC patterns support such interpretation. The spatio-temporal PC patterns should not only be checked for resemblance with the postulated features, but also the invariance of the spatial and temporal PC patterns against subsampling should be approved."

It resumes the discussion about physical interpretation of spatial PC patterns in paragraphs four to six of the introduction (Lines 68–104). See also our reply to your comment B-4 above.

Regarding the suitability of spatial PC patterns for hydrological feature identification (what you call here sufficiency for physical interpretation) we included a suggestion for future studies performing "systematic experiments with synthetic test data derived from hydrological simulation models" in the last paragraph of Section 5.2.2 and the last sentence of the conclusion.

B-7. "See the comment above where eliminating DD or reducing DD to an "insignificant" amount is a necessary but not sufficient step for physical interpretation. To meet the more stringent level of sufficient for physical interpretation, each of the retained EOF/PC patterns must reflect the correlation (or covariance) patterns well. However, because that sufficiency applies to both unrotated and rotated EOFs/PCs, some mention of that could be made because it is possible a set of patterns may not show DD but may not reflect well the patterns of data similarity (e.g., recall the heavy lift issues of maximum variance and orthogonality that are perhaps related to but not the same as DD)."

Comment" Thanks for the clarification and note that in some fields unit length eigenvectors (e.g., EOFs) are more often used than the scaled version. This has led to a morass in terminology.

AR: We assume your comment refers to our reply to the "scaling of the eigenvector" topic (It seems there was a copy-paste mistake regarding the reference of your comment. The text in quotation marks here is your previous comment). However, thank you for your acknowledgment and hint regarding the different standards in different fields. Regarding the terminology: Yes, we agree. A morass, indeed.

B-8. "AR: There was no testing to optimize the selection of the k number of PCs for rotation performed. The purpose of the varimax rotation experiment and section in our manuscript is not to identify the one set of k PCs that is best suited for physical interpretation of the rotated precipitation PC patterns or alike. It seems your suggestion is pointing in this direction. We also do not want to perform or include a full-scale rotation study here."

Comment: Understood. I left some comments earlier in case you decide to try such an experiment. It would involve more than testing Varimax over a broader range of PCs retained.

AR: Thank you for your comment. We feel appreciated and understood. We share the same understanding that such study would require a systematic and much more detailed analysis than what we did here and what we wanted to present. The extent of such a study gets clear from the second part of your introductory comment A and the first part of your comment B-4 ("… investigate all the possible permutations arising from each step"). Nevertheless, we agree that it would be an interesting task to perform such study in the future and appreciate your comments for this purpose.

B-9. "Thus, in our simple experiment here, varimax rotation was not successful in resolving DD"

Comment: Yes, that is true for the limited scope of your design. I would suggest adding the caveat that your design was not comprehensive, so you can't say if Varimax could reduce DD to an insignificant level.

AR: Thank you. We are aware of the limitation of our design that we have chosen for our scope here and agree that it is good to clearly point this out. We stated this explicitly in Section 5.2.2, Lines 758–761.

“Note however, that for the introductory scope here, the experiment with the three varimax rotation variants was kept deliberately simple. It is not a full-scale rotation study that would involve finding the best suitable set of rotated PCs for physical interpretation or alike. We did neither investigate which number of rotated PCs resulted in more or less DD, nor did we aim to find an optimum number of rotated PCs with respect to DD. Therefore, the results and their significance are limited. …”

See also our replies to your comments C-8 to C-12.

B-10. “Note however, that for the introductory scope here, the experiment with the three varimax rotation variants was kept deliberately simple."

Comment: Read my comments in the revised manuscript why the deliberately simple experiment cannot deconvolute between the way varimax creates patterns or the way keeping a specific subset of k PCs creates patterns.

AR: For our replies to your comments in the manuscript, please see section C below.

B-11. "Also, we did not investigate which number of rotated PCs resulted in more or less DD, nor did we aim to find an optimum number of rotated PCs with respect to DD.“.

Comment: I provided comments in the manuscript about this too. Not having DD patterns (unless that DD pattern happens to be the correlation pattern) is necessary for physical interpretation. However, the sufficiency comes from relating the EOFs/PCs to the correlation matrix. If the EOF/PC patterns are not supported by the correlations, then some other factor(s) are creating them. Most likely these are mis-specifying the domain size (or shape) that fails to capture the data variability scale, maximum variance of the PCs, orthogonality of the PCs or combinations thereof.

AR: For our replies to your comments in the manuscript, please see section C below. That not having DD patterns is necessary but not sufficient condition for physical interpretation is stated explicitly in the last paragraph of the conclusion.

The effects of mis-specifying the domain size with respect to the data variability scale are described in the first paragraph of Section 4.3 and some newly added lines at the end of Section 4.3 in the revised manuscript (see our reply to your comment C-1).

That the PCA features maximization of variance and orthogonality of the PCs constrain physical interpretation of PCA results (beyond DD) is stated in the introduction and the last paragraph of the conclusion. Assessing the suitability of spatial PC patterns for hydrological feature identification (what you address here as sufficiency for physical interpretation) is addressed in our suggestion for future studies in the last paragraph of Section 5.2.2 and the last sentence of the conclusion. See also our replies to your comments B-4 and B-6.

B-12. "AR: We agree that, usually, rotation of PCs is applied for physical interpretation of the PCs and their patterns. However, in our study, this is not the case. We did not perform any physical interpretation of the PCA results in the paper and we never meant to. The focus is to introduce DD to the PCA users in the hydrological community."

Comment: Yes, but my understanding of the conclusions in the manuscript is that if the EOFs/PCs contain significant amounts of DD, they should not be interpreted. I offer up an additional thought on this. If there is not significant DD, there may or may not be validity for other reasons.

AR: We agree. We included many of your thoughts and mentioned further requirements for (the validity of) physical interpretation in the manuscript, for example in the last paragraph of the conclusion. In addition, we included a suggestion for future studies focusing on the identifiability of hydrological signals from PCA results in the last paragraph of Section 5.2.2. In addition, we addressed the aspect of underfactoring/overfactoring in the revised manuscript (see our reply to your comment C-8).

C - Comments on the last version of the manuscript

Comments in the revised manuscript:

C-1. Line 508: "4.3. Effects of the domain size and spatial correlation length"

Comment: Yes, and consider the following. For domain size less than or equal to the correlation length, one cannot apply rotation as the goal of rotation is to identify patterns that are subsets of the domain. If the correlations span the domain, there can be no meaningful simplification. Application of rotation in such cases will attempt to simplify patterns that should not be simplified. This is why examining the PC loadings and comparing them to the correlation patterns is so important. Only in cases where the domain size exceeds the correlation length can rotation be examined in a meaningful way for improving the resolving of correlation modes, as finding spatially simplified configurations of the data is supported by the data. [As an aside, small domains, relative to the spatial correlation length, is rarely an issue in weather and climate studies but may be an issue for other fields of study]

AR: We agree. Thank you for this precision. We added a statement on this in Lines 540–542 at the end of Section 4.3.

“When rotation of PCs (Section 5.2.2) is applied to improve the identifiability of the features of interest it should be considered that the domains size must exceed the correlation length of the respective features. Otherwise, the simplification of the PC patterns by the rotation will not be meaningful for this purpose.”

C-2. Lines 558-559: "In particular, special care has to be taken that the truncation point of a PCA does not split a multiplet (North et al., 1982)."

Comment: This is reasonable general advice but the likelihood of problems arising is a function of the amount of variance explained by those PCs retained. If there is a substantial percentage of variance beyond the truncation point, there is a probability of more noise contaminating those eigenvectors retained. In such cases, the North et al. test is more critical. If the retained eigenvectors explain a large majority of the total variance, the small amount of residual variance (thought to represent noise) is much less and the application of the test becomes less important. Further, because rotation is immune to degenerate multiplet distortion for closely spaced eigenvalues, the problem is lessened in that situation.

AR: Thank you for this precision. We added a statement on this in Lines 561–563 and at the end of the second paragraph of the conclusion. See also our reply to your comment B-1.

C-3. Line 719: "Analysing a subsampled data set..."

Comment: Assuming this refers to spatial subsampling rather than time subsampling. Please clarify.

AR: Yes, you are right. For clarification, we changed the title of the subsection in “Subsampling of the spatial domain” and included the specification “analysed spatial” in the second sentence (Line 723):

“Reducing the symmetry of the analysed spatial domain can remove effective multiplets.”

C-4. Line 730: "5.2.2. Rotation of PC eigenvectors"

Comment: At this juncture, given the comments in the previous response about a full rotation analysis being beyond the intended scope, it is easier to clean up the details as listed below rather than embark on a full rotation analysis for this particular manuscript. However, should you decide to go there now or later in a separate manuscript, it is important to be fully prepared to investigate all the possible permutations arising from each step when rotated solutions are investigated. Those permutations would involve running a much more detailed set of experiments for the number of EOFs/PCs to retain, as that number now effects the pattern morphology (unlike the unrotated solutions where increasing the number retained adds a new pattern but does not change the previous ones), assessing the degree of simple structure in their data sets and then selecting the optimal rotational algorithm from a sizable number of algorithms (again, this step requires validating many sets of patterns to the largest correspondence to the similarity matrix from which those patterns were derived). Given all their domain shapes, such a set of experiments is cumbersome so, unless the authors are really motivated to go that route, I don't feel comfortable prescribing that as the acceptable bar and deem the present manuscript can be made acceptable. What I do request is that you issue caveats that you have no way to deconvolute the effects of keeping too few PCs or too many PCs from DD and they need to temper comments such as "the Varimax way of displaying DD" (lines 751-752). For all we know, it has nothing to do with Varimax but is a function of keeping too few PCs and forcing multiple correlation modes onto too few PCs. Please acknowledge the possibility. Not investigating it is okay but, then, either don't mention it or list it as one of a number of possibilities.

AR: Ok. Please see our replies to the second part of your introductory comment A and your specific comments to Section 5.2.2 below.

C-5. Lines 739-740: "No multiplets were split by the rotations (Figure S7) to ensure that the results of the rotation were not affected by multiplet effects (Section 4.4)."

Comment: Although it is fine to say this, keep two things in mind:

1. Eigenvalues are a property of unrotated EOFs/PCs. The property is destroyed by rotating. After rotation, the variance on individual PC loading vectors can be tallied by summing the

squared PC loadings for that vector. The total variance for the k rotated PCs will be identical to the k unrotated total variance, but the variance of individual rotated PCs will differ from the variance defined by the eigenvalues on individual unrotated PCs.

2. Rotated PCs are immune from the effects of closely spaced eigenvalues. It is a good practice not to select the truncation point, k, in the middle of a degenerate multiplet if the total variance explained by those k PCs is not large. I have commented on this in other parts of the paper.

AR: Thank you for advice. Regarding the effect of the ratio between the amount of variance associated with the retained PCs versus those excluded (signal to noise ratio) we added statements in Lines 561–563 and at the end of the second paragraph of the conclusion. See also our replies to your comment B-1 and your comment C-2.

C-6. Lines 741-742: "Note, that the newly assigned fractions of variance do not any longer decrease continuously with the PC ranks in all cases."

Comment: That depends on the software being used. Some packages will sort the rotated PCs by their variance explained.

AR: Thanks for the hint. We were not aware of that. We added this information here after a slightly rephrased version of the addressed sentence (Lines 744–746):

"The newly assigned fractions of variance did not any longer decrease continuously with the PC ranks in all cases. However, that depends on the software being used. Some packages will sort the rotated PCs by their variance explained. Note that the fractions of variance …"

C-7. Line 746: "simple structure"

Comment: Yes. In fact, this fits into the comments made earlier about domain size and rotation. If the domain is too small to expect near-zero values on some subset of locations, rotation should not be applied.

AR: Yes. In our example here, we took care that the correlation length is smaller than the domain size. Regarding your hint that the ratio of spatial correlation length and domain size of the data should be considered when applying rotation, we included a statement at the end of Section 4.3. See also our reply to your comment C-1.

C-8. Line 749: "2rPCs variant"

Comment:

I think there are two issues here, in the abstract. Only one might apply to your study, but both should be mentioned because they involve domain shape.

1. A. If the data correlation scale is larger than the spatial domain selected, the shape of the domain will affect the EOF/PC patterns because on has sampled a subset of the correlation pattern (think of it as, for example, having a large circular correlation pattern and then applying a triangular cookie cutter to that pattern, distorting the original shape). One cannot determine DD as envisioned in this paper in such a situation. Obviously, the correlation patterns should be examined prior to deciding on a domain shape.

1. B. If the data correlation scale is approximately the same spatial scale the domain selected, the shape of the domain may affect the EOF/PC patterns because on has sampled a subset of the correlation pattern (think of it as, for example, having a circular correlation pattern and then

applying a triangular cookie cutter, of approximately the same size, to that pattern), distorting the original shape. One can test for DD in such cases but there could be a competing effect of truncating the physical correlation patterns. Obviously, the correlation patterns should be examined prior to deciding on a domain shape.

In cases 1.A. and 1.B., rotation cannot work as it requires correlation scales to be smaller than the domain size.

1. C. If the data correlation scale is smaller than the domain selected, DD can be tested as suggested in the manuscript. Obviously, the correlation patterns should be examined prior to deciding on a domain shape.

2. For your experiment with varimax, it is possible 2 rPCs was insufficient and you are forcing more than 2 unique signal patterns onto 2 PCs. We know that keeping too few PCs (known as "underfactoring") forces unrelated signals on a single PC, which could be mistaken for DD. Similarly, keeping too many PCs (known as "overfactoring") splits the correlation patterns (e.g., waves with positive and negative loadings, into two separate PCs, each with one piece of the pattern). Although the overfactored rotated PCs may not show DD, then may be non-physical just the same. **That is why optimizing the k in rotated PCs is so important.** You need to add a caveat that your experiment did not involve this optimization step, so claiming that DD is present is not testable in your framework and is why I object to you claiming (for now, until a full rotated test is performed) that it is a "varimax way of displaying DD". **For all you know, it is a 2rPC way of portraying DD, where there is an unfortunate choice of k, and not the rotation method. The same might hold for 3rPC, ...**

AR: Based on your enumeration, we assume the two issues you ask us to address are:

(1) the effect of domain size versus correlation length (1.A–1.C) and
(2) that the missing optimization step for varimax can cause underfactoring or overfactoring that cannot be distinguished from DD with our simple experiment.

In addition, you emphasize once again the importance of checking the PCA patterns for physical validity („Obviously, the correlation patterns should be examined prior to deciding on a domain shape.“). We clarified already in the previous rounds of the review process that we did not include this in our work here. Please see our replies there and to your comments B-4 and B-6.

Leaving the aspect of testing for physical validity aside, we addressed issue (1) in its own Section 4.3. The aspects you are adding to issue (1) in the subitems 1.A–1.C are very specific. Adding them as claims to the manuscript would require (a) further experiments in our manuscript to demonstrate the claimed aspects first or (b) references from the literature. We are not aware of any literature in this regard and we do not want to add any further experiments to our study. Therefore, we stay with issue (1) on the level that is covered in Section 4.3.

Regarding issue (2) we declared that we did not perform a full-scale rotation study that would involve the optimization step you are mentioning (Lines 759–761):

“It is not a full-scale rotation study that would involve finding the best suitable set of rotated PCs for physical interpretation or alike. We did neither investigate which number of rotated PCs resulted in more or less DD, nor did we aim to find an optimum number of rotated PCs with respect to DD. Therefore, the results and their significance are limited.”

However, the possibility of underfactoring/overfactoring and the possible interference with DD, were not mentioned so far. To point out the importance of optimizing the number of k rotated PCs in (varimax) rotation also in this regard, we added in the revised manuscript in Lines 762–765 the following lines immediately after the aforementioned lines on the missing optimization step.

“It cannot be ruled out that the DD of the presented results might be an effect of keeping too few PCs (underfactoring). In other words, unrelated signals might be forced on a single PC causing the observed DD. Keeping too many PCs (overfactoring), on the other hand, might split the correlation patterns, respectively the representation of a hydrological feature. However, overfactoring is not an issue here due to the small number of PCs retained.”

However, acknowledging the limitations of our experiments and study, we think that we still can say something based on our experiment. Namely, that simply taking the first few PCs and performing varimax rotation - that is without any optimization of the number of rotated PCs - is not per se sufficient to resolve DD. We think this finding is of practical value because we assume that (varimax) rotation is regularly applied without the extensive optimization set up you were outlining.

Therefore, we added the following sentence after the issued caveats in Lines 766–769:

“Despite its limitations the experiment shows that the application of varimax rotation per se – that is without optimizing the number of rotated PCs – is not necessarily sufficient to resolve DD. For practice this implies that, whereas rotated eigenvectors are generally considered to be less prone to DD than unrotated ones (Richman, 1986; Wilks, 2006), it cannot be taken for granted that simply taking the first few PCs of an analysis and varimax rotating them suffices to resolve DD.”

All together the former third paragraph of Section 5.2.2. was revised substantially. It now reads (Lines 751–769):

“In all three varimax rotation variants, the patterns were clearly dependent on the domain geometries (Figure 17). While the dominant PC 1 monopole of the unrotated PCA disappeared, the new dominant patterns are gradients reflecting the domain shape. For example, the patterns of the 2rPCs variant showed gradients from southwest to northeast in the square domain, from west to east in the rectangular domain and from north-west to south-east in the triangular domain. The gradients of the square domain from the 4rPCs variant reflect the rotational symmetry of the square (Figure 17a, right panel). The gradients of the rectangular and triangular domain associated with the major fractions of variance (Table 5) depict in all three rotation variants the longest extent of the domain (Figure 17bc). Thus, here, varimax rotation was not successful in resolving DD.

Note however, that for the introductory scope here, the experiment with the three varimax rotation variants was kept deliberately simple. It is not a full-scale rotation study that would involve finding the best suitable set of rotated PCs for physical interpretation or alike. We did neither investigate which number of rotated PCs resulted in more or less DD, nor did we aim to find an optimum number of rotated PCs with respect to DD. Therefore, the results and their significance are limited. It cannot be ruled out that the DD of the presented results might be an effect of keeping too few PCs (underfactoring). In other words, unrelated signals might be forced on a single PC causing the observed DD. Keeping too many PCs (overfactoring), on the other hand, might split the correlation patterns, respectively the representation of a hydrological feature. However, overfactoring is not an issue here due to the small number of PCs retained.

Despite its limitations the experiment shows that the application of varimax rotation per se – that is without optimizing the number of rotated PCs – is not necessarily sufficient to resolve DD. For practice this implies that, whereas rotated eigenvectors are generally considered to be less prone to DD than unrotated ones (Richman, 1986; Wilks, 2006), it cannot be taken for granted that simply taking the first few PCs of an analysis and varimax rotating them suffices to resolve DD."

As an aside, PCA is regularly used to identify dominant processes or alike in hydrometric monitoring data. That is, non-synthetic data that does not stem from controlled experimental set ups. In this application, the number of dominant processes is usually not known beforehand. Thus, analogue to the selection of the truncation point in unrotated PCA, the question arises how to identify the optimum number of retained PCs there? We think one possibility might be to start with k = 2 and stepwise increase k until the patterns of the retained PCs remain rather stable. However, this is one of the aspects that can be analyzed in future studies.

C-9. Lines 751-752: "seemed to be the varimax way of displaying DD."

This is both vague. and not tested (see earlier comment). What is "the varimax way of displaying DD"? Such a statement would suggest that no matter what the underlying correlations, varimax PCs would give the same set of patterns. Is that the case? From what I can see, it might as well be "the k Varimax way" (whatever that is) for a single example and not generalizable to other correlation functions as it is for unrotated EOFs/PCs? See earlier comments on the distortions known to occur when underfactoring/overfactoring. Of course, the added step (not examined in this manuscript) of relating the PC patterns to the correlation patterns will instantly confirm if DD is a potential issue for any analysis (unrotated, rotated). If that comparison has a poor match between the PCs and the correlations, then some other factors (perhaps including DD) might play a role. We just can't tell from this experiment.

AR: Thank you for pointing this out. We agree that it is not possible to generalize from these few examples to a general "varimax way of displaying DD". Our intention was not to make such general claim here. It was meant as a descriptive phrase to summarize that – for the examples we showed – the shape of the domain was clearly reflected in the patterns of the varimax variants (2rPC, 3rPC, 4rPC). So, it was meant as description of our results, not as a generalization. Now that we read the paragraph again, we realized that we stated this already in the first sentence of the third paragraph of Section 5.2.2.:

"In all three varimax rotation variants, the patterns were clearly dependent on the domain geometries (Figure 17)."

Thus, to prevent giving the misleading impression that it would be meant as generalization, we deleted the sentence you were addressing in the revised manuscript. Furthermore, we rearranged some sentences in the paragraph to adjust for the missing sentence. Please see our reply to your precious comment C-8 for how the paragraph now reads.

Regarding the caveats of our simple experiment, including the possible effects of underfactoring/overfactoring, please see our reply to your previous comment C-8.

The issue of the missing check for physical validity ("Of course, the added step (not examined in this manuscript) of relating the PC patterns to the correlation patterns will instantly confirm if DD is a potential issue for any analysis (unrotated, rotated).”) was clarified already in the previous rounds of the review process. Please see our replies there and in this document here (i) the beginning of our reply to your comment C-8 and (ii) our reply to the first part of your introductory comment A.

C-10. Line 758: "deliberately simple."

Comment: You need to tell the reader what "deliberately simple" means. Hopefully, I have left sufficient comments about the critical need to test each k PCs when rotated to determine if any of those sets gives a valid result.

AR: We explicitly listed the limitations of our study in Lines 758–765. To be even more clear, we restructured the paragraph such that the limitations follow immediately after the sentence which states that our experiment is simple. Furthermore, we added in the revised manuscript another caveat addressing the aspect of underfactoring/overfactoring. Please see also our reply to your comment C-8.

The issue of the missing check for physical validity (“Hopefully, I have left sufficient comments about the critical need to test each k PCs when rotated to determine if any of those sets gives a valid result.”) was clarified already in the previous rounds of the review process. Please see our replies there and in this document here (i) the beginning of our reply to your comment C-8 and (ii) our reply to the first part of your introductory comment A.

C-11. Line 759: "are limited."

Comment: Suggested addition to this sentence: "or misattributed to the rotation method rather than underfactoring/overfactoring."

AR: In the revised manuscript, we moved the addressed sentence and added a caveat regarding the aspect of underfactoring/overfactoring in the directly following sentences. Please see our reply to your comment C-8 for how the paragraph now reads.

C-12. Lines 760-761: "Also, we did not investigate which number of rotated PCs resulted in more or less DD, nor did we aim to find an optimum number of rotated PCs with respect to DD."

Comment: Lines 760-761: What you have not deconvoluted is the DD effect of selecting a different number of PCs to retain from the rotation method applied. Equally important, because the PCs are not compared to the correlation patterns, the physical validity cannot be established. [I realize you don't want to go there, though if the paper is concluding that EOF/PC patterns with DD should not be interpreted; therefore, failing to tell the reader when the patterns should be interpreted (and the method to support the interpretation) is less than satisfying.] Once again, the conclusion in this paragraph could be because the data are not well represented by the k PCs retained, or the data are not well represented by a varimax rotation, or by a combination of both.

AR: In the setting of the option 2 you were suggesting in your first introductory comment A, we issued further caveats in the manuscript and restructured the paragraph in which the addressed sentence appears. In particular, we issued caveats regarding the aspects of (i) underfactoring/overfactoring (“What you have not deconvoluted is the DD effect of selecting a different number of PCs to retain from the rotation method applied”) and (ii) physical validity (“Equally important, because the PCs are not compared to the correlation patterns, the physical validity cannot be established. …”).

For (i) and how the paragraph of the addressed sentence reads now, please see our replies to your comment C-8. For (ii), please see our replies of the previous rounds and to your previous comments in this document here.

C-13. Lines 762-763: "be more robust against spatial"

Comment: It may be the spatial instability is inter-related to DD. Can you comment on that from these experiments?

AR: Yes, it may be. We assume you are referring to Richman (1986)´s statement that "subdomain instability …. is a corollary of domain shape dependence". We are not sure which experiments you mean. We assume it is the one from Section "Subdomain stability" in Richman (1986). However, we included the aforementioned aspect as a general statement (Lines 770–773):

"Except from being less prone to DD (Richman, 1986; Wilks, 2006), rotated PCA results were found to be … more robust against spatial (Richman, 1986) and temporal (Cheng et al., 1995) subsampling. Note that spatial instability may be inter-related to DD. In particular, subdomain instability can be a corollary of DD (Richman, 1986)."

C-14: Line 763: "and less sensitive to degeneracy (Richman, 1986)."

Actually that study showed essentially no sensitivity to degeneracy for rotated PCs, as the eigenvalues were the same to many decimal places. It did show some sampling variability at very small sample sizes.

AR: Thank you for the precision. We assume you mean that the eigenvectors ("loadings" in Richman (1986)) were the same to many decimal places and that you are referring to the experiment described in the sub-section "Sampling errors" with the results being summarized in Table II (Richman, 1986). There, unrotated PCA and an oblique rotation method (Direct Artificial Personal Probability Factor Rotation (DAPPFR criterion)) were compared in their performance regarding the recovery of population eigenvectors with nearly equal eigenvalues (degeneracy) when using different sample sizes. The DAPFR method was found to be robust against sampling errors even for very small sample sizes. We changed the wording in Lines 770–772 to:

"Except from being less prone to DD (Richman, 1986; Wilks, 2006), rotated PCA results were found to be robust against sampling errors in case of eigenvalue degeneracy (Richman, 1986), and more robust against spatial (Richman, 1986) and temporal (Cheng et al., 1995) subsampling."

C-15. Line 765: "drawbacks of rotation"

These are undefined in this manuscript. Jolliffe's comments applied to the loss of uncorrelatedness and orthogonality in the spatial and/or temporal patterns and the extra work involved in running a rotated analysis. You have previously criticized orthogonality in this paper as one factor hindering interpretation, so it's not clear what is meant here. I do mention the situation when the domain size is smaller than or equivalent in size to the correlation scale as factors against rotation. However, in such cases of small domains (relative to the correlation scale), Jolliffe's suggestion of rotating select PCs will not help physical interpretation, regardless of the eigenvalue spacing.

AR: We wanted to point out that when rotating only multiplet members, it is also only the multiplet members that are affected by the changed PC properties, compared to standard PCA. This means the loss of uncorrelatedness and orthogonality in the spatial and/or temporal patterns (depending on the applied scaling of the eigenvectors), and in particular that the PC patterns depend on which and how many PCs are rotated (Section 3.5). In the context of Jolliffe (1989; 1995) these changes were addressed as “drawbacks”. We agree that in our context it makes more sense to use the neutral “changed PC properties”. However, Jolliffe (1989; 1995) was also pointing out that if the rotation is restricted to a multiplet the effect of

(i) rather even redistribution of variance among the rotated PCs is not much of an issue, as the variance was already rather evenly distributed in the multiplet before rotation, and
(ii) the scaling of the eigenvectors on the results is diminished when the rotation is restricted to PCs that exhibit similar eigenvalue sizes (which is the case for multiplets).

Therefore, we rephrased the sentences in Lines 774–780 to:

“Rotating only multiplet members limits thereby the changes in the PC properties (Section 3.5) to the multiplet (Jolliffe, 1989; 1995), in particular the dependency of the rotated PC patterns on which and how many PCs are rotated. That rotation results typically in a rather even variance distribution between the PCs is not much of an issue, because the variance in the multiplet is already rather equally spread between the multiplet members before rotation (Jolliffe, 1989). Also, the effects of the scaling of the eigenvectors (meaning the loss of uncorrelatedness and orthogonality in the spatial and/or temporal patterns, see Section 3.5) are diminished, because the eigenvalues of the multiplet are of similar size (Jolliffe, 1989; 1995).”

Jolliffe, I. T.: Rotation of Ill-Defined Principal Components, Applied Statistics, 38, 139–147, https://doi.org/10.2307/2347688, 1989.

Jolliffe, I. T.: Rotation of principal components: choice of normalization constraints, Journal of Applied Statistics, 22, 29–35, https://doi.org/10.1080/757584395, 1995.

C-16. Lines 818-820: "If the spatial PC patterns do not differ significantly from DD reference patterns, we recommend to report that and stop any interpretation of individual spatial PC patterns as distinct hydrological features."

Comment: Yes, though I believe you can say more about this. See my previous comments.

AR: We assume you are addressing the issue of “significance” regarding the differences from DD reference patterns. In the revised manuscript, we clarified that we mean “significant differences” in the sense of statistical testing. Please see our replies to your comments B-2 and B-3.

I hope my comments are useful in finalizing the manuscript.

AR: Yes, they were. Thank you for your work and spending so much of your time!