Notes concerning

Can local climate variability be explained by weather patterns? A multi-station evaluation for the Rhine basin

Aline Murawski, Gerd Bürger, Sergiy Vorogushyn and Bruno Merz

Hess-2016-286

There was an initial reaction upon studying the manuscript: 40 patterns?? Are you serious??

Then another question arose: Are pressure fields as classification a good basis for the stratification of data? There is good experience with using relative topography instead (Spekat et al., 2010 – reference is given at the end).

The authors do not go into seasonality when analyzing their data. There is contrary experience from the COST733 Action on classification methods (in the meantime, the results of that COST Action have been made publicly available, see reference at the end). Plus there is experience towards great usefulness of seasonality in classification, indicated, e.g., in Spekat et al. (2010).

Usage of E-OBS data – Fig. 1 only shows 8 E-OBS grid points. Text mentions 10. Moreover in Fig. 1: Poor choice of colours for dots.

From page 5 on the line numbering is irritatingly confusing. Just look at the repetition of line number 5 on page 5…

Section 3.2 page 7 line 12: Maybe relative topography is not directly available, but contributing geopotentials can be easily extracted and the retop could be easily computed.

Page 7 line 14 – and I mean the second appearance of this line (sigh…) beginngin with "(extents " it must be made more clear that these numbers refer to geographic degrees of latitude and longitude.

Page 10 figure 3: Not clear what the numbers "12" , "14" and "33" on the right-hand side of the array of figures mean. I would furthermore recommend to deliberately use different colour schemes for different parameters, so they can be better distinguished. Even if the authors would not follow this suggestion at least they should reverse the assignment of the colours (left side red, right side blue) for PREC and HUMID since blue would then point to wet condition which suits the intuition better. More a thought than a substantial comment: All selected patterns are rather cold/have rather low radiation, so perhaps one would like to see examples for classes which denote different conditions.

Page 10 line 8 I have doubts if retaining the absolute values is a good approach when you have season-specific classes. Particularly temperature is way different from season to season. So *au contraire* to what the authors write, anomalies are a good way forward because they cover a similar value range in all seasons. Moreover, there is the experience that a further reduction in the number of classes is possible by using anomalies (this would be favourable in the light of the big set of 40 classes used by the authors).

Page 11 figure 4: I suggest to add y-axis labels on the right-hand side, too. Further suggestion: Use open triangle and open circles which are better visible in case of overlaps, and those do appear frequently.

Page 12 figure 5: It would be could to have the results for four different class number shown at least for one more extent (or domain size, as one might better call it).

Page 12 line 14: "increasing EV" - this is very minute if you look at it. Therefore I suggest to write "almost no change".

Page 14 figure 7: I am amazed how relatively even the frequency distribution is. Expectation would be that some classes would be quite rare. Furthermore: is the property displayed really the quotient of frequency and cumulative EV? Furtherfurthermore, what does the (-) at the end of the y axis label mean?

Page 14 lines 30 and following as well as Fig. 8: It is remarkable and should be pointed out that for numerous classes the reanalyses (black dashes) mark either the lowest or the highest frequency so in those cases ALL GCMs are unanimously indicating either higher or lower frequencies, respectively. Isn't that an odd behavior?

Page 14/15 Section 4.2.2 I assume that Fig. 9 on Page 16 is meant to visualize this, right? Then make a reference to that figure!

Section 4.2.2 again: A definition needed is needed as to what is considered a *good* reproduction. Imagine that there could be ties in the months of most frequent occurrences - or months with very similar frequency. Would that still be good/superior/inferior reproductions, then?

Page 15 Section 4.2.3 The text points to Figure 8, wheras the reference should point to Figure 10.

Page 18 around line 35, the aspect of stratification skill was presented in Spekat et al. (2010), too. Perhaps this needs to be mentioned in the text.

Page 18 line 12 (bottom): More like a comment - this discussion opens up a whole philosophical can of worms, i.e., universality versus optimization. Should the goal be to find a classification able to cover "everything but with a variable degree of fidelity" or should the goal be to find a classification that is region- and variable-specific, yet has a high skill? Maybe the authors could be drawn into discussing this for a bit, too?

General comment on the "Discussion" section: It is rather long (no criticism concerning the length, mind you) and could benefit from the insertion of subsections.

Page 34 Table 1: Is it "Runs" or "Run", i.e., did the authors use all 10 CNRM runs or all 10 CSIRO runs (for example) or did they use just a specific one of those? Then this particular run should be specified. This refers back to page 3 line 12 where it is ambiguous if ALL or SOME runs were analyzed in this paper.

General comment concerning Figs. 11 thru 16: It is amazing to me that those 40 classes, some of which are *visually* quite similar to each other, apparently constitute sets of necessary distinctiveness. Just from looking at them the, admittedly subjective, estimate would be that much fewer classes should be sufficient.

General comment on the line numbering: It should be uninterrupted, starting with 1 and end in the high several hundreds. The numbering in the draft here is misguided and misguiding.

General comment on the figure placement: Particularly for Figs. 7 thru 10, a better proximity to their mentioning in the text and respective paragraphs to which they belong should be found.

General comment on Literature – one could of course think of a "me too" effect… - but there is a paper from 2010 which covers or complements several aspects of the manuscripts's reasoning:
Spekat, A., F. Kreienkamp and W. Enke, 2010: An impact-oriented classification method for atmospheric patterns. – Physics and Chemistry of the Earth **35**, 352-359.
Also, perhaps unknown to the authors of the manuscript, the final report for the COST733 Action is now available. The link to the final report is: https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/index/index/docId/3768

That link is permanent. The URN is urn:nbn:de:bvb:384-opus4-37682 (it can be found using a web search engine).

So, bottom line: Something in between minor and major revision. Some reasoning needs to be better shaped, some needs to mention a bit more what alternative paths have been pursued. There is some potential to improve technical aspects (figures, mostly) and general understandability.