Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-286-AC2, 2016

Hydrology and Earth System Sciences Discussions
Open Access

Interactive comment on "Can local climate variability be explained by weather patterns? A multi-station evaluation for the Rhine basin" *by* Aline Murawski et al.

Aline Murawski et al.

murawski@gfz-potsdam.de

Received and published: 19 August 2016

We thank the reviewer for his/her constructive comments on our manuscript and provide point-by-point response hereafter.

Main impression: The paper presents an evaluation of downscaling climate information over the Rhine region by making use of weather patterns. A number of different options are explored, and the conclusion is that best results are obtained with mixed predictors: sea-level pressure in addition to temperature and humidity fields. The main new aspects of this study include the specific emphasis on the Rhine region, large number of stations representing the local climate, and the long time series for searching for weather patterns.

Printer-friendly version

Discussion paper

One question I have with this analysis is whether the evaluation of the method is best done when making use of cross-validation of split-sample for calibration/evaluation. Perhaps the main message has a tendency to get lost in all the details? This could be fixed with some revision and an emphasis/reminder of how the details support the main message. I think the abstract may be rewritten – a bit bolder – to make the paper look more interesting.

With regards to the reviewers comment, we point out that the classification of weather patterns is done based on the ERA-20C reanalysis data, whereas the evaluation of the classifications' capability to stratify local climatological variables is evaluated using an independent data set based on climate station observations.

We shall consider the reviewer's suggestion and make the abstract more appealing pointing out our main messages.

Details:

L15p2: "statistical approaches are comparatively cheap, computationally efficient and relatively easy to apply…". No, it is not always easy to apply statistical downscaling in a good fashion that correctly captures the dependency to large-scale conditions. However, it's easy to apply both dynamical and statistical downscaling to get some output – be it reliable results or non-representative numbers.

We adapt the respective paragraph, deleting the criticised sentence.

L28p2: "The underlying assumption of the downscaling based on weather patterns is that the regional or local behaviour of climate variables is a response to the larger-scale, synoptic forcing." More precisely, downscaling also works if only a fraction f(X) of the variability (which one would expect) is dependent on the large-scale conditions X (local processes n are also usually involved): $y = f(X) + n$. However, both large-scale dependent and local variability must be accounted for. One case in point is precipitation, as discussed in the paper.

We acknowledge the reviewers comment and would change the mentioned sentence to "The underlying assumption of the downscaling based on weather patterns is that the regional or local behaviour of climate variables is partly *a response to the larger-scale, synoptic forcing."*

L31p2: "Statistical downscaling tends to underestimate the variance of regional or local climate and may poorly represent extremes". Some past studies have not accounted for the contribution local processes n, hence the variance in the results will be less than observed. Variance inflation is flawed and a priori gives incorrect results (von Stoch, 1999).

We will adapt the statement accordingly: "Statistical downscaling tends to underestimate the variance of regional or local climate if the contribution of local processes is not considered and may poorly represent extremes. Different methods have been proposed to rectify this problem: variable inflation (Karl et al., 1990), expanded downscaling (Bürger, 1996) and randomisation (Kilsby et al., 1998)."

L12p3: The assumption of stationarity is more severe for GCMs and RCMs, which rely on parameterisation schemes, involving statistically trained equation to represent the bulk description of unresolved quantities (e.g. cloud schemes). In GCMs/RCMs the results of such schemes feed back to the calculation of the large-scales, whereas for statistical downscaling/weather generators, they can produce a trend in biases. Also relevant for L15p19. When mentioning this only in relation with statistical downscaling (SD), the reader gets a distorted picture and thinks it only affects SD – this has resulted in a myth within the downscaling community.

Yes, we agree with the reviewer that dynamical downscaling techniques also rely on the assumptions of stationarity in the empirical relationships incorporated in the climate models. In our study we, however, solely focus on the statistical downscaling and even more narrowly on the weather generator type downscaling conditioned on weather patterns. We thus discuss only the assumptions required for this downscaling *approach.*

"Data:" Gridded observation (EOBS) and station data were mixed? This can introduce artifacts (spatial and temporal inhomogeneities). Furthermore, gridded daily precipitation is no good for analysing extreme precipitation, as the grid points are weighted sums of surrounding observations and hence are expected to exhibit different statistical characteristics (tail of distribution – see attached fig). Also see `http://www.icrc-cordex2016.org/images/pdf/Programme/presentations/parallel_D/D3_Chandler_`

Furthermore, isn't EOBS limited to after 1950? Perhaps it's better to skip France altogether even if the picture is less complete?

The reviewer raised valid concerns regarding the use of E-OBS data. We decided to exclude these data points from our analyses and update all figures and results accordingly. However, the influence of these data points is only minimal, thus the overall messages of the paper remain unchanged.

"3.1 Weather pattern classification" - the paper discloses that the cost733 class software was applied to both reanalysis and GCM data (?) – but does that mean that the weather patterns are the same for the models and reanalyses?

Apparently, our description of the workflow was misleading. We first establish a classification on reanalysis data. The GCM data are then assigned to the existing patterns (by means of minimum Euclidean distance). Thus, the weather patterns are the same for the reanalysis and the GCMs. We adapt the first sentence in section 4.2 accordingly.

"3.2 Finding optimal classification parameters" Keep in mind that with many tests, the likelihood of finding an accidental match increases. The "problem of multiplicity" – See Wilks (2006).

Multiplicity applies to statistical tests. We do not apply multiple tests, but rather evaluate a set of different valid parameters by using two different metrics. Hence, we do not share the concerns of the reviewer.

Eq. 3 – 5: Daily rainfall amount is far from normally distributed, whereas the root-mean-square metric is more appropriate for temperature, which tends to behave more like the normal distribution. TSS, WSS and BSS will be strongly affected by a few heavy precipitation events (acknowledged in 4.1.2), and explains low scores for the metric EV. For precipitation, it may be wise to look at aggregated statistics, eg seasonal wet-day mean (precipitation intensity), wet-day frequency, and probabilities (e.g. Benestad & Mezghani, 2015): The precipitation frequency exhibits a close connection with the circulation pattern (e.g. SLP), whereas the intensity is more complicated and is expected to be strongly affected by local small-scale processes (eg convection, which may be consistent with Fig 7 and mentioned in the discussion), but be somewhat moderated by large-scale conditions.

We thank the reviewer for the suggestion and will consider the mentioned characteristics (wet-day mean, wet-day frequency, and probabilities) in the revised version.

Furthermore, the observations represent a poor sample – a rain gauge represents a few cm^2 capture of a spatially heterogeneous phenomenon with a scale of km^2. Aggregation in time or space may give a clearer picture that is less affected by sampling fluctuations. The alternative to downscaling single station data and then estimate the area average is to estimate the area average from observations and then downscale this index. I suggest adding some text about this possibility and these issues in the discussion, at least.

We share the concern of the reviewer that station records represent only a point estimate of the spatially heterogeneous phenomenon like precipitation. We, however, intend to apply an existing advanced weather generator from Hundecha et al. (2009) which works on a station basis. Hence, we would abstain from interpolation of precipitation and downscaling regionalized indices instead of estimations at station locations.

L3p10: The humidity estimate from reanalyses is difficult to validate – it may have substantial errors?

We generally agree. In the relatively short overlapping period, however, ERA-20C moisture variables such as precipitation and total column water vapour agree fairly well with observations (cf. Poli et al. (2016), DOI: 10.1175/JCLI-D-15-0556.1). Besides, our study can also be viewed as a validation study itself.

"4.2.2 Seasonality" - it's not clear what "the earliest and last months of occurrence in the course of the year" are and how they are specified.

We add another paragraph in the end of section 4.2 to explain more thoroughly, what we mean by seasonality of patterns: "Seasonality is evaluated by the first, last, and peak month of pattern occurrence. All patterns show a distinct seasonality. Each season is characterised by a limited number of consecutive months in which a pattern occurs. We evaluate the beginning (i.e. first month) and end (i.e. last month) of pattern occurrence. The peak month is defined as the month with highest number of days with pattern occurrence. Some patterns show two distinct seasons. In this case both seasons are evaluated separately."

"4.2.3 Persistence" - the duration of phenomenon/event/pattern may follow the geometric distribution, and differences in the models and reanalysis can be gauged based on its statistics. It can provide an estimate of what differences one would expect from randomness and what is likely a systematic bias.

Our point here was solely a comparison of duration times and not the question whether there is any persistence at all. For that one could use the geometric distribution, but it would constitute a rather weak null hypothesis since in that case consecutive days are treated independently, which they are obviously not.

Minor:

I would move the first sentence in the abstract to the beginning of the introduction. You don't need to explain why or provide justification in the abstract.

To move the first sentence from the abstract to Introduction would make this sentence

Printer-friendly version

Discussion paper

stand-alone followed by the description of the basin and the flood problem. In the abstract, it provides an overall idea of the scope of the study. We would prefer to keep the current structure, if the reviewer agrees.

Second sentence in the abstract is a bit difficult, and can be rephrased or moved out of the abstract. It distracts the story away from the main findings. I'd start the abstract with "An objective classification scheme is presented . . ."

Thanks for pointing this out. We shall critically revise the abstract, in particular the first statements and in general make it clearer and 'bolder' as suggested by the reviewer earlier.

L28p4: "For the workflow proposed here three different sets of climate data are needed:" Comma between "here" and "three"?

L1p11: "The selected classification was compared to the Hess-Brezowsky- Grosswetterlagen" - Use "compared with" rather than "compared to" when there was an actual comparison?

Thanks, the corrections will be done.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-286, 2016.