

Response to reviewer A

Abstract

(A1) The last two sentences are difficult to understand in the context / contradict the previous statements. Is exploiting the spatial autocorrelation (i.e. KED) better or as effective as an elaborate predictor set? Do climatological background fields add something to the performance or not as suggested in the sentence stating that circulation types do not add anything (or do you mean something else than circulation types with climatological background fields?)

The seeming contradiction in the abstract is now resolved by a more careful wording in these sentences. Indeed the notion of a “background” field was not meant as alternative to “circulation types”. This was rather referring to the common approach of using climatological mean fields as reference (background) in daily interpolation. This is now explicitly mentioned and should avoid confusion. Exploiting spatial autocorrelation was more beneficial than elaborate predictors with a regression model in our concrete application. However this may depend on region and predictors considered, which is why we chose the moderate formulation “may be as effective”.

Introduction

(A2) As far as I see there are very few references specific for the Alps, for example for KED.

Lines 105-106: We newly include references to Tobin et al., 2011 and Allamano et al., 2009, where KED and regression kriging were used for interpolation in regions of the Alps.

Case study

(A3) What are “composite means”?

The notion was used to describe averages over all members (days) of the classes of a circulation type classification. To avoid confusion we have removed the word ‘composite’ in our revised manuscript in favor of a more explicit formulation (see lines 213-215).

Method

(A4) I do not understand exactly how the bootstrap experiments were completed (how is the sampling error assessed?, what is the “exact” mean referring to? For an observed sample there is no such thing as an exact mean)

We are interested in the minimum number of days a station has to record, so that the average

precipitation value lies close to its climatic mean. The exact mean remains unknown, but stations with the most complete time series serve as a benchmark to study the convergence of the average precipitation when including more days. This sensitivity study starts by randomly sampling n days for the subset of 20 stations with the longest time series available. Then, the error made by estimating the climatic mean from n days instead of the complete time series is quantified. The error metric is based on the relative mean root transformed error. The upper 95% quantile of the error is estimated by bootstrapping the data for many sets of n days. The minimum number of days necessary to have 95% of the bootstrapped climatic estimates below an error threshold of 0.1 is calculated for the 20 most complete stations. The median number of the minimum number of days for these 20 stations becomes the minimum record length that all stations must fulfill in order to be included in the spatial analysis. The error threshold is somewhat arbitrary but is chosen to be sufficiently restrictive to guarantee reliable climatic estimate while retaining enough data. This procedure has been clarified in the text.

(A5) Why do you assume isotropic variograms, what could be expected from using anisotropic variograms?

In the revised version we have introduced a justification for choosing an isotropic instead of an anisotropic variogram in section 3.1 (lines 295-303). Our choice is related to the risk of compensating predictor dependence (in the deterministic model part) by anisotropy in the stochastic model part. This would make our results overtly sensitive to the simple two-dimensional structure of the topography in our region, and, hence, less representative for conditions with a more complex topographic setting.

(A6) Why is the likelihood-based estimation central to the analysis (is this mentioned again)?

The original manuscripts states: “Estimating trend coefficients and variogram parameters jointly means that the procedure implicitly distinguishes between variations in the observations that are better explained by the predictors and variations that are better explained by spatial covariance (spatial continuity)”. We have now added two more sentences (lines 313-317) explaining that the frequently applied two-step procedure where trend coefficients are estimated by linear regression and residuals then interpolated by ordinary kriging, involves contradicting assumptions and risks overfitting.

Indeed the issue of likelihood-based and joint estimation of parameters is mentioned again later in the conclusions (lines 801-811), where it is relevant for our conclusion that regression-based assessments may be a poor guidance for KED predictor choice.

(A7) “Constant variance” is unclear here since it appears to me that you have not mentioned yet whether you estimate the variograms per time step or not (it is mentioned only much later)

By “constant variance” we meant “stationary variance in space”, which we made clear in the text now (line 319).

(A8) Does the transform not have some undesirable effects on high values? What is the empirical distribution of the data? Does the square-root transform suit this?

The main objective of the transformation is to better comply with the assumption of spatial stationarity of the process variance implicit to KED. Precipitation data is skewed towards larger values, which implies larger process variance in areas of large as compared to small precipitation. (See Schabenberger and Gotway, 2005). As mentioned in the manuscript (lines 328-333), the choice of $\lambda=0.5$ is motivated by the analyses of Erdin et al. (2012) showing that a formal estimation of λ (i.e. a data driven transformation) did not significantly alter the best estimates compared to when it was prescribed at 0.5 (i.e. square-root transformation). We think that a more detailed assessment of the setting of λ would inflate the number of experiments in the paper and distract from the core experiments (predictor dependence).

(A9) Math notation: HESS does not like multi-letter variable names, please use standard variable names for eq. 4 and 5.

The name of the errors has been changed to B (for Bias) and E (for rel.MRTE).

(A10) Bias definition: for hydrology a somewhat unusual definition, the bias is more commonly defined as a relative difference value.

Our definition is a variant of the bias defined in Germann et al. (2006). The advantage of the present definition is that values of $B=2$ ($B=0.5$) are immediately recognized as a factor 2 over (under) estimate in domain mean precipitation.

Results

(A12) I think that the presentation of the results could be improved in terms of text structure. Currently the fields are first described and then their performance evaluated (why not having subtitles?); combining the two might help removing the repetitions.

We have considered to combine the description of the fields and evaluation results, but a visual interpretation of the differences between the various models was an important pre-requisite for a condensed description and interpretation of the evaluation results. We have, however, followed the reviewer's suggestion for having more subtitles for better navigation.

(A13) I do not exactly follow what you refer to by stating at the end of 4.2 "because of the omission of stations from . . .".

The length of a time series for a class of a circulation type classification is necessarily much shorter than for the entire seasonal subset. Stations were omitted that did not reach the minimum number of observations required to guarantee a reliable average of their climatic mean (see section 2). We have clarified this point more explicitly in the text (lines 681-683).

(A14) It would have been nice to see somewhere a short discussion of how straight-forward KED is to apply; the estimation of variograms per time step is not necessarily a simple task; did you use any procedure to detect time steps without any exploitable autocorrelation? How were they handled?

The estimation of variograms for the interpolation of climatic averages was not causing technical difficulties. With all our experiments there were non-zero estimates of range and sill parameters. Again, with daily precipitation, the absence of “exploitable autocorrelation” does not cause a technical issue. Such cases are formally represented by a nugget-only variogram, which leads to a residual field that is a constant in space (see e.g. Diggle and Ribeiro, 2007). The only complication that can occur is when the variance in the station values is zero. This occurs when all stations report dry conditions. This case was captured in a preliminary step and the interpolated field was set to a zero constant. We have added a remark on this special case in Section 3.3 (Interpolation of daily precipitation, lines 435-438).

(A15) The readability of the figures such as Fig. 6 could perhaps be improved by replacing the squares with some other object (instead of only differentiation by color, which is ok for the pdf)

New figures 6 and 9 include different symbols instead of squares. This will help distinguishing between the methods.

(A16) Final question: was your work based on existing statistical libraries or did you code everything yourself? In the first case, a reference would be helpful.

At the end of section 3 (Methods and experiments), we cite the geoR package: “All computations are done in R (R Core Team, 2011) using the geostatistics package geoR (Diggle and Ribeiro, 2007)”

Response to Reviewer B

(B1) In the abstract, it is stated first that ‘Again, the stratification by circulation types and the wind-aligned gradient predictor do not improve over the single predictor KED model. Similarly for daily precipitation, information from circulation types is no improving interpolation accuracy.’ Then it says that ‘. . .they support the common practice of using climatological background fields in the interpolation of daily precipitation.’ Is ‘climatological background fields’ referred to the same as ‘circulation types’? If so, the authors need to rethink what information they want to deliver, because it seems that results conflict with conclusions. If not, the authors need to do further work to reword this part so as to avoid potential confusion.

We agree that this part was not clear enough in the original manuscript. In accord with referee’s point A1, we have amended the wording in the second part of the abstract (lines 25-43). See also our response to A1.

(B2) In Section 3.1, it is stated that ‘In all our applications, the semi-variogram is assumed to be exponential with a nugget, sill and range as parameters. The semi-variogram is assumed to be isotropic.’ The authors should at least discuss why it is reasonable to assume an exponential model rather than, for example, spherical, and why it is reasonable to assume an isotropic covariance function. It seems that the zonal patterns of the observed precipitation field as shown in Figures 1, 2 & 4, do not support such an isotropic assumption.

In the revised version we have introduced a justification for choosing an isotropic instead of an anisotropic variogram in section 3.1 (lines 290-303). See also our response to point A5 of referee A. In preliminary sensitivity experiments, we did observe only minor differences in cross-validation results between interpolations with a spherical and with an exponential variogram, which finally led us opt for the more simple exponential model. This is now also mentioned in section 3.1 (lines 303-306).

(B3) Some of the presentations can be optimized. I don’t think in the present study the ‘Box-Cox transformation’ is more important relative to the ‘likelihood-based estimation procedure’ because, as emphasized by the authors, ‘the utilization of a likelihood-based estimation procedure is central in our application.’ I believe, most readers, like me, may be more willing to see what the ‘likelihood-based estimation procedure’ is and why it plays a central role in the current application.

See also our response to point 6 of reviewer A. We have now extended our description in section 3.1 (lines 313-317). The reason why the likelihood-based estimation procedure is important in our study is described in section 6 (conclusions, lines 801-811). The relation of this part of the text to the estimation procedure is now made more evident by a back-reference from section 6 to section 3.1.

(B4) Following the authors, box-cox transformation with a power parameter of 0.5 will transform all wet days to be -2 (though I am not sure whether a fixed value of 0.5 is the best choice or not). This means that after transformation daily precipitation becomes -2-inflated

from zero-inflated. I am wondering how these ‘-2s’ are treated when interpolating daily precipitation because on one hand this is critical for the representation of the intermittent nature of daily precipitation field, and on the other hand daily precipitation time series contains a great amount of zeroes.

The Box-Cox transformation is introduced in our application to improve consistency with model assumptions in terms of non-stationarities related to skewness of non-zero precipitation. As mentioned by the referee, the transformation does not resolve non-stationarity related to intermittency. In the revised manuscript we now mention this explicitly (lines 334-342) and refer to proposals in the literature to account for intermittency in spatial precipitation modeling. Intermittency is violating the stationarity assumption of the stochastic model used for interpolation, but it is not an issue for the interpolation of seasonal mean values.

References

- Allamano, P., Claps, P., Laio, F. and Thea, C.: A data-based assessment of the dependence of short-duration precipitation on elevation, *Phys. Chem. Earth, Parts A/B/C*, 34(10), 635–641, doi:10.1016/j.pce.2009.01.001, 2009.
- Erdin, R., Frei, C. and Künsch, H. R.: Data Transformation and Uncertainty in Geostatistical Combination of Radar and Rain Gauges, *J. Hydrometeorol.*, 13(4), 1332–1346, doi:10.1175/JHM-D-11-096.1, 2012.
- Germann, U., Galli, G., Boscacci, M. and Bolliger, M.: Radar precipitation measurement in a mountainous region, *Q. J. R. Meteorol. Soc.*, 132(618), 1669–1692, doi:10.1256/qj.05.190, 2006.
- Schabenberger, O. and Gotway, C. A.: *Statistical Methods for Spatial Data Analysis*, Chapman and Hall, CRC Press., 2005.
- Tobin, C., Nicotina, L., Parlange, M. B., Berne, A. and Rinaldo, A.: Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region, *J. Hydrol.*, 401(1), 77–89, doi:10.1016/j.jhydrol.2011.02.010, 2011.