Hydrol. Earth Syst. Sci. Discuss., 12, C1755–C1765, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C1755/2015/

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Spatially shifting temporal points: estimating pooled within-time series variograms for scarce hydrological data" *by* A. K. Bhowmik and P. Cabral

Anonymous Referee #3

Received and published: 26 May 2015

The paper "Spatially shifting temporal points: estimating pooled within-time series variograms for scarce hydrological data" by A.K. Bhowmik and P. Cabral proposes a new method for estimating theoretical variograms using an innovative technique for pooling spatial times series in regions where hydrological data are sparse and the spatial density of gauging stations is low, with special regards to developing countries. This research application could be potentially very interesting because it deals with an issue that has rarely been addressed in the scientific literature and shows how to handle data scarcity through an innovative technique. However, before the final publication, several aspects of the paper need to be improved substantially, including the manuscript sec-

tioning and clarity of language as well as a better discussion of the assumptions implicit in the proposed method and some misleading reasoning.

Major comments

1. Sec. 2.3 describes the core assumptions of the method and how to implement it through an algorithm reported in Eqs. (1). The proposed technique shifts points from time to space in order to increase the sample size available across the study region to meet the threshold of 400 "data points" (see L 3-8 P 2245), all of which are treated as points with different spatial locations. Although establishing a threshold for a minimum number of spatial locations is a reasonable approach, this is actually not easy to generalize and could, to some extent, be contradictory. Indeed, what is questionable is why the authors do not refer to spatial density of the points, i.e. the number of stations divided by the study area. For instance, one can have a lower number of stations with a higher gauge density because the study region is one or two orders of magnitude smaller. On the contrary, one can work with more than 400 stations but with a vast study region, so that the reliability of the estimated theoretical variograms could, in principle, still be inaccurate. Parajka et al. (2015) recently investigated the role of station density and demonstrated how the prediction performances are strictly related to it. Moreover, the literature contains several successful geostatistical applications for predicting a variety of hydrological indices where the number of stations is not as large as proposed by the authors. For instance, Todini et al. (2001) use 82 gauges for the estimation of average yearly precipitation, Castiglioni et al. (2011) 51 for low-flow indices, Castellarin (2014) and Pugliese et al. (2014) use, for the prediction of flow-duration curves, 23 and 18 stations respectively, Archfield et al. (2013) use 61 stations for the prediction of flood in the south-east United States and Laaha et al. (2013) employ 214 stations for river temperatures in Austria. Thus, in the

revised paper, it would be better to directly use station density as a proxy for the reliability of the estimated variograms

2. Since the authors do not cite any already published work specifically on the SSTP (spatially shifting temporal points) technique, one may assume they are introducing a brand new method developed by the authors themselves. If this method has been already published, I strongly recommend the authors to provide as many references as possible for such an approach. If instead they developed this novel method, it would be better to pay more attention to the assumptions, the description and the details of the technique, and provide also a general overview of the method aside from its application to a specific case study, so that anyone could further test and implement the method for a different case study. Including either pictures or schematic diagrams, which graphically illustrate how the method works, would also improve the reader's understanding of the method.

3. The authors claim that spatial stationarity is achieved by analyzing data time series both in the latitudinal (N-S) and longitudinal (E-W) directions employing two different statistical tests, the Pettit-Mann-Whitney test and the Dickey-Fuller test (please report references for such tests). Although these tests might be reliable for detecting change points in time, so that stationarity is reached over a reasonable time window, they do not say any thing about stationarity in space. In geostatistical applications, the analysis of stationarity in space is a well-known concept, which can be pragmatically assessed by looking at the variogram. Indeed, if a random spatial process is stationary in space, this means, at least, that one is assuming the mean of the regionalized variable $Z(\mathrm{s})$ to be constant across the study region and the covariance is a function of the distance between paired spatial locations (well known as "second-order stationarity", see e.g. Cressie, 1993). Thus, the variogram, which is complementary of the spatial covariance, must reach a threshold, known as "sill", at a given distance, meaning that there is no correlation between two points whose distance from each other exceeds a

given threshold, known as "range". Otherwise, if the variogram monotonically increases as the distance between pairs of spatial locations increases without ever reaching a threshold, there is likely/possibly non-stationarity in space due to the presence of a trend in the mean.

The authors adopt power law theoretical models to fit empirical variograms. In doing this, they suggest that there is no stationarity in the spatially random process, so then the stationarity hypothesis they introduced in space seems to be rejected.

4. Furthermore, some considerations about variograms are necessary. Fig. 3 reports empirical and fitted theoretical variograms for three selected time windows as well as variograms from the whole recorded time series. First, in the boxes surrounding the variograms are reported the estimated variogram parameters, but, as long as the authors adopted a power law, it is not clear what the parameter "partial sill" refers to. Secondly, even though the power law model is likely an identifiable parametric model, the "Hole" model is not a common choice in geostatistical applications, therefore its equation along with a description of the parameters are required (a list of commonly adopted variograms can be found in Journel and Huijbregts, 1978). I recommend that the authors specify variogram equations or provide references to them regardless of whether they are commonly used or not.

5. In my view, the authors over-focus on the estimation of an accurate theoretical variogram and insufficiently discuss other issues that may interest the reader. For instance, it might be compelling if the authors would spend more attention to the evaluation of how the proposed method increases the predictive capability at ungauged sites compared to other benchmark methods by using unbiased estimators, such as kriging techniques. Currently, the authors only mention the use of kriging interpolator in a few places in their study, e.g. the first time the authors mention the words "ordinary kriging" is at L 22 P 2251. With the addition of

dedicated section to kriging interpolation, I would support a reasonable focus on the estimation of accurate theoretical variograms in the final version of this paper. Moreover, if non-stationarity in space is confirmed in the revised analysis, using the proper techniques for diagnosing spatial nonstationary described above, the authors should note that using ordinary kriging (OK) could produce misleading and biased results. Even though OK does not require the random process to be second-order stationary, it does not accommodate trends in the mean. If a trend in the mean is detected, the authors should instead consider using a different linear interpolator, such as Universal Kriging, which handles trends in the mean better (Matheron, 1969; Delhomme, 1978).

6. The sectioning of the paper is confusing, misleading and missing in some parts. In Sec. 2 the authors reports "Data and software" first (Sec. 2.1, L 2 P 2247), describing how the proposed method is applied to the available data before introducing the method itself, which is outlined in Sec. 2.3.1 (L 25 P 2248). A general description of the approach before the presentation of the case study would enable the reader to understand what the "SSTP" model is and how the data are used better. Moreover, the manuscript ends with Sec. 4 "Discussion": it is a general rule to end a scientific manuscript with a section dedicated to the conclusions of the proposed research, typically called "Conclusions", presenting a brief recapitulation of the main outcomes, the methods and data used in the study. Finally, the "Results" section (L 16 P 2252) is poorly articulated and overall too short given its importance. I strongly suggest that the authors change the current sectioning of the paper to a more logical and common structure. A possible order of sections could be as follows (this is just a proposal, the authors can definitely modify the sectioning at their convenience):

    1. Abstract
    2. Introduction

    3. Methods
        3.1 Spatially shifting temporal points (SSTP)
        3.2 Averaging empirical variogram (AEV)
        3.2 Kriging methods
    4. Study area and data
    5. Results
        5.1 Stationarity tests
        5.2 Assessment of variograms estimation
        5.3 Kriging interpolation accuracy
    6. Discussion and future research challenges
    7. Conclusions

Minor comments

1. In Sec. 2.3, which describes the SSTP method and the computation of variograms through equations, the mathematical notation should be heavily revised in a way commonly accepted by the scientific community. For instance, vectors must be boldface whereas scalar objects (e.g. model parameters) should not be, and continuous variables must be in italics (e.g. x and y for spatial coordinates, t for time). In Eqs. (1) and (2) parentheses are placed incorrectly. Eqs. (5) compute the minimum and maximum distance between two spatial locations, respectively. There are at least two errors here: (1) the notation "$\mathrm{s}_1, \mathrm{s}_2$" indicating spatial distance between two locations is misleading, and should be changed to something that the reader can deduce more easily, e.g. $D_{i,j}$ or a more complicated $\parallel \mathrm{s}_i - \mathrm{s}_j \parallel$ or perhaps use Greek symbols; (2) the "min" and "max" functions should be applied over scalar quantities, e.g. a set of distances,

instead of the actual notation, which refers to a set of vectors. Also, the radical sign in Eq. (6) should cover the whole expression. Furthermore, the authors should state the domain of variability of indices for each equation with subscript variables, e.g. i or j. For mathematical conventions and other editing rules regarding the typesetting of scientific manuscripts, please refer to IUPAP (1978) or Cohen and Giacomo (1987). Finally, the HESS journal also provides its own quick overview of some general rules the authors should follow before submitting a paper (http://www.hydrology-and-earth-system-sciences.net/submission/manuscript_preparation.html).

2. In my opinion, the authors use parentheses for too many sentences, concepts and details that are fundamental to the manuscript. A few examples are L 18-20 P 2245, L 4-6 P 2246, L 6 and 14 P 2247, L 14-15 P 2255. I suggest avoiding the intense use of nested sentences through parentheses and to linearly articulate concepts in order to smooth out the flow of the reading.

3. In the "Supplement material" section, the authors present 1 table, 2 figures and an R-code script. In my view, the figures could be reasonably included in the body of the manuscript since they are effective for comprehending the text. Moreover, the authors repeatedly cite these figures throughout the manuscript. Fig. S2 could be easily included, while a significant sample can be extracted from Fig. S1, e.g. 15-20 years, which describe how the gauge density varies in time. In addition, the x-axes tick labels in Fig. S1 are illegible. The same goes for Tab. S1: a sample of this table reporting 1st, 2nd , 3rd quartile, min and max values, along with box-whisker plots, can be also reported in the body of the manuscript, e.g. in the section where the study area and data are presented. Regarding the R code, the authors should report a brief description at the beginning of "Supplement material" section highlighting: (*i*) what is the input, (*ii*) what is the output of the code, (*iii*) which version of R is used, (*iv*) any packages dependencies and (*v*) a quick usage example. Finally, the current version of this section starts with

a description of table and figures, which are actually just a "copy and paste" of their respective captions. I recommend that the authors introduce the supplement material with a brief description different from these captions.

4. Fig. 3 reports empirical and theoretical variogram for both the SSTP and AEV methods. It is not clear to what the estimated parameters refer since the power model does not allow any "range" or "sill" parameters, and the "Hole" model seems to have two (or multiple?) sills. Moreover, one of the most important parameters in the power model is the power parameter. Matheron (1971) and Cressie (1993) recommend a further condition to be satisfied by fitted theoretical variogram model, that is:

$$\frac{2\,\gamma(\mathbf{h})}{\parallel \mathbf{h} \parallel^2} \to 0, \quad as \quad \parallel \mathbf{h} \parallel \to \infty$$

where $\gamma(\mathbf{h})$ is the theoretical variogram as a function of the spatial lag $\mathbf{h}$. Given this property, the estimated power parameter should be defined over the interval $[0, 2)$, i.e. the power parameter cannot be equal or larger than 2. In the figure, the only parameter that satisfies this property is the one called "Range". The authors should clarify this crucial aspect.

5. In order to assess the kriging prediction the authors employ the root mean squared error (RMSE). Although this is one of the most broadly used model performance metrics in hydrology because it gives an unbiased evaluation of the goodness of an estimator, in many hydrological applications, it is preferable to employ Nash-Sutcliffe efficiency (see Nash and Sutcliffe, 1970) which provides a better overview of the predictive capacity of a model because it accounts for the variance of the empirical data sample. I would recommend the authors to report this index next to the RMSE in Tab 2.

6. The abstract should highlight the most important findings of the study better, providing quantitative assessments of the main outcomes as opposed to only qualitative ones.

Technical notes and misspellings

- L 18-20 P 2245: would be better to avoid nested parentheses.
- L 25-29 P 2245 and L 1-2 P 2246: verb tenses are contradictory and the inline items list is not coherent with verbs. Please rephrase the whole sentence.
- L 12 P 2246: remove "(semivariance)", otherwise this needs a better explanation.
- L 4 P 2247: define the term "wet days" including any precipitation threshold that distinguishes wet days from non-wet days.
- L 4 P 2247: does "Fig. 1" refer to Fig. 1 of the current manuscript or to the one in Peterson et al. (2001)? Please clarify this point.
- L 5 P 2247: "(DMICCDMP, 2012)" is cited as an acronym, so that in the bibliography would be better to switch and put the acronym first and then its explanation.
- L 8-11 P 2247: one would expect an increased precision due to the increased number of gauges. Please rewrite the sentence.
- L 17 P 2247: substitute "altitude" with "elevation".
- L 24-26 P 2247 and L 1 P 2248: meaningless sentence. Please rephrase.
- L 7 P 2248: substitute "Hereafter" with a more suitable adverb.
- L 11 P 2248: substitute "Next" with a more suitable adverb.

- L 1 P 2249: "touples" should be "3-tuple", but seems a useless specification, please remove it or rephrase.
- L 12 P 2250: avoid the punctuation mark before the equations, revise here and in other parts of the manuscripts.
- L 17 P 2251: "Eq. (5)" should cross-referenced as Eq. (6). The authors use to put references to equations before these are outlined, cross-references should be putted after.
- L 5 P 2252: substitute "means" with "mean".
- L 15 P 2252: please provide bibliographic reference to "PANAGEA".
- L 5 and L 6 P 2253: specify what the authors mean with "$10°$" or "$5°$". These should be geographical measures of latitude or longitude, or the portion of an arc segment, so, please, provide information about them and their geographical datum too (e.g. WGS84).
- L 12 P 2253: it is not clear why the author inverted the sign in the expression "$90° > \varphi > 0°$", please clarify or revert it from $0°$ to $90°$ as usual.

References

Archfield, S. A., Pugliese, A., Castellarin, A., Skøien, J. O. and Kiang, J. E. (2013), "Topological and Canonical Kriging for Design Flood Prediction in Ungauged Catchments: An Improvement over a Traditional Regional Regression Approach?", *Hydrol. Earth Syst. Sci.*, 17 (4): 1575–88. doi:10.5194/hess-17-1575-2013.

Castellarin, A. (2014). "Regional Prediction of Flow-Duration Curves Using a Three-Dimensional Kriging", *J Hydrol*, 513 (May): 179–91, doi:10.1016/j.jhydrol.2014.03.050.

Castiglioni, S., Castellarin, A., Montanari, A., Skøien, J. O., Laaha, G. and Blöschl, G. (2011). “Smooth Regional Estimation of Low-Flow Indices: Physiographical Space Based Interpolation and Top-Kriging”, *Hydrol. Earth Syst. Sc.*, 15 (3): 715–27, doi:10.5194/hess-15-715-2011.

Cohen, E. R. and Giacomo, P. (1987), “Symbols, Units, Nomenclature and Fundamental Constants in Physics.”, North-Holland.

Cressie, N. A. C. (1993), “Statistics for Spatial Data”, J. Wiley.

Delhomme, J. P. (1978), “Kriging in the Hydrosciences”, *Advances in Water Resources*, 1, (5): 251–66, doi:10.1016/0309-1708(78)90039-8.

IUPAP (1978), “Symbols, Units and Nomenclature in Physics”, *Physica A: Statistical Mechanics and Its Applications*, 93, (1–2): 1–60, doi:10.1016/0378-4371(78)90209-1.

Journel, A. and Huijbregts, C. (1978), “Mining Geostatistics”, Academic Press.

Laaha, G., Skøien, J. O. Nobilis, F. and Blöschl, G. (2013), “Spatial Prediction of Stream Temperatures Using Top-Kriging with an External Drift”, *Environmental Modeling and Assessment*, 18, (6): 671–83, doi:10.1007/s10666-013-9373-3.

Matheron, G. (1969), “Le Krigeage Universel”, École nationale supérieure des mines de Paris.

Matheron, G. (1971), “The Theory of Regionalized Variables and Its Applications”, Vol. 5, École nationale supérieure des mines de Paris.

Nash, J.E. and Sutcliffe, J.V. (1970), “River Flow Forecasting through Conceptual Models Part I - A Discussion of Principles”, *J. Hydrol.*, 10, (3): 282–90, doi:10.1016/0022-1694(70)90255-6.

Parajka, J., Merz, R., Skøien, J. O. and Viglione, A. (2015), “The role of station density for predicting daily runoff by Top-Kriging interpolation in Austria”, *J. Hydrol. Hydromech.* 63, doi:10.1515/johh-2015-0024.

Pugliese, A., Castellarin, A. and Brath, A. (2014), “Geostatistical Prediction of Flow–duration Curves in an Index-Flow Framework”, *Hydrol. Earth Syst. Sci.*, 18, (9): 3801–16, doi:10.5194/hess-18-3801-2014.

Todini, E., Pellegrini, F. and Mazzetti, C. (2001), “Influence of Parameter Estimation Uncertainty in Kriging: Part 2 - Test and Case Study Applications”, *Hydrol. Earth Syst. Sci.*, 5, (2): 225–32, doi:10.5194/hess-5-225-2001.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 2243, 2015.