Hydrol. Earth Syst. Sci. Discuss., 10, C7391–C7397, 2014
www.hydrol-earth-syst-sci-discuss.net/10/C7391/2014/

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Geostatistical prediction of flow-duration curves" *by* A. Pugliese et al.

A. Pugliese et al.

alessio.pugliese3@unibo.it

Received and published: 23 January 2014

Reply to Referee Daniele Ganora

We thankfully acknowledge Referee's useful comments, which will significantly help us to improve the presentation of the study. We report below our replies (denoted by "Response", the actual reply, and "Action", a brief illustration of the revisions to the manuscript) to all referee's comments (indicated by "Reviewer").

MAJOR COMMENTS

Reviewer:
1. The Kriging procedure (i.e. the estimation of the weights λ) is applied, for each ungauged basin, to a number on neighbouring stations which is set equal to 6, as indicated on page 13067, after a preliminary analysis. Although this approach can

improve the final outcomes, I think it is in contrast with what is claimed on page 13057 (from line 10), i.e. that the geostatistical approach allows to avoid the identification of homogeneous regions. Of course, strictly speaking, Kriging does not require the homogeneity of the region; however, this pre-selection introduces a subjective element and undermines the robustness of the original method. The point is that Kriging (as well as top-Kriging) automatically provides weights on the basis of the distance of the ungauged site to the donor stations (and according to the correlation structure). Thus, the weights are (automatically) greater for close donors and smaller for donor sites far away. Moreover, weights depend on the location of the ungauged site, so the weighting structure adapts for different ungauged basins. Under this perspective, I would like to see first an application where the whole dataset is considered. The variograms should also be reported.

Response:
The referee is right. The introduction of a fixed number of neighbouring stations considered in order to solve the kriging system could be a strong limitation, however, if on one hand this seems to introduce the hypotheses of an implicit delineation of homogeneous region, on the other hand it is a common practice in several geostatistical applications, as well as in Top-kriging, the use of the so-called "moving neighbourhoods". Moving neighbourhood avoids considering very dissimilar sites in the interpolation, which would be associated with smaller weights, but nonetheless could be detrimental for the accuracy of the prediction. Moreover the spatial variability in the kriging methods relies on empirical variograms and the selected theoretical variogram models, which best fits the experimental points. This computational step involves all stations in the region, so the estimates computed using a limited amount of stations somehow benefit from all the information provided by the dataset.
Action:
In the revised version of the manuscript we will include how the results vary by using moving neighbourhoods with different sizes. Examples of empirical and theoretical variograms will also be reported.

Reviewer:
2. The hypothesis that top-Kriging weights λ can be used to weight empirical FDCs is the core assumption of the method; it is a strong assumption, so I would try to verify it. I suggest one possible way to perform this task, but the authors are free to propose any other reliable method: i) consider one station among the whole dataset of N stations; ii) for that station compute the $N-1$ weights λ; iii) compute the $N-1$ δ values between the empirical FDC of the selected station and the empirical FDCs of the remaining stations; iv) compare the λ values and δ values: large δ (dissimilar curves) should correspond to small λ (small weight) and viceversa; v) repeat points i)-iv) for each station.

Response:
The Referee points out a very important aspect and we thankfully acknowledge his suggestions.
Action:
In order to test and validate the basic assumption introduced with this study, in the revised version of the paper we will include a figure resulting from the analysis of the spatial consistency of the weighting scheme. The figure consists in a scatter diagram relating the distance δ values computed between curves i and j with the corresponding $\lambda_{i,j}$ value obtained in cross-validation, this for all $i, j = 1, \ldots, n_s$ where n_s is the number of stations belonging to the region.

MINOR COMMENTS

Reviewer:
1. Due to the assumption reported in the previous comment the authors should specify, starting from the title, that the prediction of FDCs is somewhat an indirect product of the geostatistical framework. To do so, I suggest to change the title to "Geostatistical weighting scheme for prediction of flow-duration curve".

Response:
We thank the referee for his suggestion. We agree with him that the core of the study is

the introduction of an innovative linear weighting scheme which enables one to predict flow-duration curves, however we also believe that the title proposed by the Reviewer is too unbalanced towards the linear weighting scheme, which is just a mathematical expedient to perform the prediction. Also, the prediction of the FDC that we perform is indeed geostatistical within an index-flow framework. We could modify the title as: "Geostatistical prediction of flow-duration curves in an index-flow framework", if the Handling Editor recommends it.

Reviewer:
2. The point (i) at the top of page 13058, as well as other sentences in the manuscript, describes the TND as a characteristic of the whole curve. Actually, the flood-part of the curve (normalized discharge greater than 1) is not represented by the TND, so I would relax the statements regarding the whole curve by specifying that flood flows are not really accounted for by the TND.

Response:
Point taken, we agree with the Reviewer.
Action:
The revised manuscript will explicitly acknowledge that TDN does not describe the portion of the curve associated with low durations (high flows), also remarking that the model performances and the accuracy of the predicted FDCs are assessed on the whole curve (high flows included).

Reviewer:
3. Page 13060 line 10: the "non-decreasing" property of FDCs actually depends on the way the curve is represented. If ordered discharges are plotted against the non-exceedance probability, the curve is non-decreasing; otherwise (as in this paper), if the exceedance probability is used, the curve is non-increasing. I suggest to use "monotone relationship" to account for both the possible representations.

Response:

We agree with the referee.
Action:
We will adopt "monotone (i.e. non-increasing in this paper) relationship" in the revised manuscript.

Reviewer:
4. In page 13064, from line 14, the authors introduce an operational problem due to the different length of the period-of-record FDCs. The issue regards the lowest value dimensionless duration d which vary for different record lengths and thus affects the computation of the TND. The proposed solution is to fix a maximum d, which is equivalent to cut the right tail of the FDC at the specified d, in order to have the same limit for TND calculation for each FDC. I think this operation would be no longer necessary if the curves were previously resampled at a set of pre-imposed durations. A possible resampling set is $d_i = 1 - i/(N+1)$ with $N = 365$ ($i = 1, \ldots, 365$) if one refers to the equivalent number of days in a year, but it is not the only possible resampling. In fact, a resampling procedure has already been implemented by the authors in section 5, over 20 points equally spaced in the z space. Since this resampling procedure is essential to predict the FDC, it should be applied before the TND computation to keep the framework consistent.

Response:
The Referee raises a good point and we thank him for the possibility to clarify it. We believe that resampling, as the referee proposes, could be a viable solution if the numerical computation of the TND values were computationally demanding, which is not the case. Therefore we still believe that interpolation should be avoided, when possible. We did use interpolation (20 points, which are enough to characterize a curve according to e.g. Shu and Ouarda, 2012), but only for assessing the prediction accuracy and compare the performance of different models at different gauges.
Action:
The revised manuscript will make it clear that the 20-point resampling is used only for

assessing the prediction accuracy and comparing the performances of different models, while empirical TND values are computing using all the information available.

TECHNICAL NOTES AND MISSPELLINGS

Reviewer:
1. Page 13057: the text block of lines 13-16 has basically the same information as the lines 21-25. Please reformulate the paragraph to remove redundant information.
2. Parentheses around citations of equation numbers are often missing throughout the text (se for example P13059 L 14; P13060 L 17; P13064 L16; etc.) and should be added.
3. In plots showing FDCs in the frequency domain I would use "Exceedance frequency" or "Dimensionless duration" rather than "Duration" which, instead, recall a dimensional time variable. The same correction should be done throughout the text, for instance before eq. (4) and on page 13064.
4. Missing punctuation after eq. (7).

Response:
1. We will remove "without the delineation of homogeneous regions" from the statement.
2. Ok. We will rectify.
3. The Reviewer is right, we will clarify in the text that with d we refer to a dimensionless duration equal to the exceedance probability, and then we will use "Duration d [-]" in all Figure (included the central panel of Figure 2).
4. Ok. We will rectify.

REFERENCES

Shu, Chang, and Taha B. M. J. Ouarda. (2012). "Improved Methods for Daily Streamflow Estimates at Ungauged Sites." Water Resources Research 48. doi:10.1029/2011WR011501.