Hydrol. Earth Syst. Sci. Discuss., 10, C286–C291, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C286/2013/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Optimal depth-based regional frequency analysis" *by* H. Wazneh et al.

Anonymous Referee #2

Received and published: 11 March 2013

GENERAL

The paper is on regional estimation of flood quantiles by multivariate linear regression. It is a continuation of an earlier paper by the second and third author (Chebana and Ouarda, 2008), where a weight function that measures the similarity between sites was introduced to avoid the delineation of homogeneous regions. In the present paper the parameters of the weight function are optimized, which makes the method more objective. The method is applied to flood quantiles in three different regions.

The optimization of the weight function is an important contribution to the regional estimation of flood quantiles. However, the method is poorly explained. It took me quite some time to understand a bit what the authors have done. My difficulties are explained below. The presentation needs much improvement.

MAJOR COMMENTS

1. In the introduction the authors state that they are interested in design event quantiles at ungauged sites (p. 520, l. 23, 24). In line with this, the quantiles at the start of their procedure are unknown (p. 531, l. 17). Then an iterative procedure is described on p. 532, which ends with an estimate of the quantiles at the target site on top of p. 533. However, on the same page two performance indices are given, in which there are N sites for which a local quantile estimate is available. This suggests that the authors have applied the iterative procedure N times to the quantiles of gauged sites rather than an ungauged target site.

2. The definition of vectors and matrices in the description of the weighted least-squares procedure needs more care. From p. 528, l. 4-9, I understand that Y and log Y are row vectors with s elements (a 1 * s matrix) and log X is a row vector with r+1 elements [a 1*(r+1) matrix]. Then, if beta is an (r+1)*s matrix, the product log X * beta is indeed a 1*s matrix. However, it is not possible to write beta * log X as is done on p. 528, line 15, unless s = 1. Furthermore, log X and log Y on p. 528, l. 16 and p. 528, l. 20 [Eq. (11)] seem to be no longer row vectors but matrices with N rows (l.16) or N columns (l. 20). Moreover, if Yi is an 1*s matrix in Eqs. (17) and (18), the question arizes how the ratio of two such matrices is defined. I had the same difficulties with the Chebana and Ouarda (2008) paper on depth-based regional frequency estimation.

3. Section 2 is a rather long section on background material which is sometimes difficult to read (in particular sections 2.1 and 2.4) and it leads to questions about the connection between the various topics. For instance, the sentence "The subset A represents the neighborhood or the region in the classical RFA approaches" (p. 527, l. 6,7) would be immediately clear to the reader if he/she realizes that the argument x of the weight function is the Mahalanobis depth. Weight functions are defined in section 2.2 and weights wi appear in the next section 2.3. The reader may ask at that stage how these two are related. It is not explained in section 2.4 that the optimization refers to the coefficients of the chosen weight function. This is even not clear in the later

section 3.1 where the general procedure is outlined. A flow chart may be useful. I have the impression that step i is an iteratively weighted least-squares procedure within the optimization step iii.

4. To be more specific on the readableness of section 2.1, why should the reader be bothered with terms like affine invariance, simplicial volume, halfspace? (p. 524). It is curious that if F is replaced by the empirical distribution function, there appears a mu in the left-hand side of Eq. (2) instead of the empirical distribution function, and that the right-hand sides of Eqs. (1) and (2) are identical. The discussion of the depth function should be reduced. The need for a separate subsection should be questioned. It may be better to introduce the depth function in the discussion of the weights.

5. It is not necessary to explain the Nelder-Mead and pattern search optimization methods (sections 2.4.1 and 2.4.2). A reference to the relevant literature suffices. The discussion on p. 523, l. 3-11 on optimization algorithms should also be removed.

MINOR COMMENTS

The authors should bring their reference list in agreement with the main text. For instance, the often cited paper Chebana and Ouarda (2008) is not found in the reference list. This is also the case for Singh and Bardossy (2008), p. 521, l. 28.

p. 520, l. 3. Please replace “correspond to” by “lead to” or “result in”.

p. 521, l. 16. “method” should read “methods”.

p. 522, l. 2. “depth function” should read “a depth function” or “depth functions”.

p. 522, l. 29 and p. 525, l. 12. “three families of weight functions”. However, four different weight functions are discussed in section 2.2.

p. 525, l. 1. “Mahalanobis depth”. Should it not be “Mahalanobis depth of x with respect to mu”? This is more in line with p. 532, l. 2, 3.

p. 525, l. 21 and p. 526, l. 14. What is a “derivable” function?

p. 526, l. 4. The coefficient b in Eqs. (3) and (4) is a scale parameter. It determines the spread and not the shape of the curve. For the Gompertz function this is demonstrated in Fig. 1c. Note that the caption of Fig. 1c is not correct. It should read b varies with fixed a and c (and the caption of Fig. 1b should read a varies with fixed b and c).

p. 527, l. 10. What is “the quantile of order alpha for p degrees of freedom?”. Does it relate to some quantile of the chi-square distribution?

p. 528, l.3. The variable epsilon in Eq. (7) is not defined. Note that in contrast to Eq. (8), epsilon is not a vector in Eq. (7).

p. 528, l.9. Please reformulate “matrix formed by “r” vector”.

p. 528, l. 11. “error of model” should read “model error”.

p. 528, l. 11. The mean of epsilon is the null vector. It is not clear what the authors mean by “null mean vectors”.

p. 528, l.12 and p. 528, l. 18-20.The authors assume that the covariance matrix of the quantile estimates is the same for all sites. How valid is this assumption? Is it better fulfilled after the log-transformation?

p. 528, l. 17. Please change “wi and wi” into “wi where wi”.

p. 528, l. 21, 22. The fact that the regional estimate contains a transformation bias may question the use of RB in Eq. (17) as a criterion for optimizing the coefficients in the weight function. RB will differ from zero if the local quantile estimate is unbiased.

p. 531, l.23. Please change “identity matrix of dimension N” into “N*N identity matrix”.

p. 532, l.4. The second argument in the depth function seems to be no longer a location parameter mu. In contrast to Eq. (2), the two arguments of the depth function are separated by a comma instead of a semicolon.

p. 532, l. 4. I understand why the quantity D has been enclosed by large round

brackets. A consequence of this is that more round brackets are needed in the right-hand side of Eq. (15) to indicate that the depth is the argument of the function phi.

p. 534, l. 17. It is uncommon to refer to section 3 within section 3.

p. 535, l.4. “wil” should read “will”.

p. 535, l. 14. “performed”. Should it not be “used”?

p. 535, l. 17. “precipitations” should read “precipitation”.

p. 536, l. 21. The fact that alpha refers to the neighborhood coefficient should be mentioned on p. 536, l. 20 and not on p. 536, l. 21.

p. 537, l. 3,4. The sentence has no verb.

p. 537, l. 9. The word “consequently” could be omitted.

p. 537, l. 19, 20. What is R?

p. 538, l. 9. In statistical theory the term “efficient” refers to the variance of an estimator. Therefore an estimate cannot be efficient in terms of the relative bias RB.

p. 538, l.13. Please delete “in terms of values”.

p. 538, l. 21. What is a high S-curve?

p. 539, l. 9-11. How can we see from the results in Table 2 that the optimal function keeps the S shape?

p. 540, l. 1,3. Do the results in Fig. 8 really imply that there is an underestimation by a factor 2-5 for some sites? For Southern Quebec and Arkansas the largest underestimation is strongly reduced by the depth-based approach. How much contributes this to the reduction in the average relative bias?

p. 550. Fig. 3 is unnecessary.

p. 552. Please change “with respect RRMSE” into “with respect to RMSE” and “with

respect RB” into “with respect to RB”.

p. 555. Please use the same vertical scales in the top and bottom rows of Fig. 8.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 519, 2013.