Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-472-RC1, 2017

Hydrology and Earth System Sciences Discussions

HESSD

Interactive comment

Interactive comment on "Hydro-stochastic interpolation coupling with Budyko approach for spatial prediction of mean annual runoff" *by* Ning Qiu et al.

M. Mälicke (Referee)

mirko.maelicke@kit.edu

Received and published: 27 October 2017

Summary

I would like to thank the authors for this interesting contribution, I appreciated reading it. The authors presented a new interpolation approach for mean annual runoff depths. It was presented as a proof-of-concept comparing the interpolation results to long-term observations in a catchment in China.

The proposed method is, to my understanding, an extension of the commonly used estimation approaches based on the Budyko curve. The authors describe these approaches as semi-empirical, or deterministic, and try to account for an assumed stochastic error on runoff depth observations by using a geostatistical interpolation method to regionalize errors observed in gauged basins to ungauged basins. By describing deterministic shares of runoff by the Budyko curve and utilizing observed covariance structures in the deviations, the authors could show, that the combination of both can yield significant better results, than each method on its own.

Evaluation

In my opinion, the proposed method is of relevance and adds value to recent issues in runoff depth interpolation. The results seem promising and the methods were presented largely in a clear and transferable way. Except from some technical remarks, the figures were relevant and descriptive.
I would recommend to propose some major revisions for the methods, as these need clarification or extension in some parts. The results were presented in figures, tables and in word, where I would propose to add more numbers in order to obtain an even more comprehensive insight. A major revision is also proposed for the discussion, as this section does only summarize the results in most parts. My remarks for revising this work can be found below.

I kindly ask to consider my remarks and finally I would like to see this work published in HESS.

Major points

- l.63 – 71: This paragraph consists of only two sentences, which are way too long and thus, were confusing for me. In the first sentence the authors make two different points. First, streamflow is a combined landscape information and second, that climate-landscape variability leads to non-stationary runoff observations. I kindly ask the authors to separate these points and reword the following statements in order to foster the structure. Especially the term "deterministic term"

(l. 65) needs more and clearer introduction. This is in the following work also referred to as "deterministic trend" and is of fundamental importance for the proposed method. Introducing this term in more detail will significantly increase my text comprehension for the entire work.
The second sentence in this paragraph (l. 68 – 71) does in my opinion not connect to the first one and it was not obvious what this sentence shall emphasize. What trend does the "the spatially nonstationary trend of runoff" (l. 68) refer to? And how is a runoff trend interpretable as "hydrological regionalization in terms of hydro-climate and landscape data" (l. 68 – 69)? What I read out of this sentence is that non-stationary runoff is caused by heterogeneity in hydro-climate and the landscape and can be described by empirical relationships as done in the presented studies (l. 71). But this is not exactly what is written down in this paragraph.
In my opinion, the authors shall rewrite the whole paragraph in shorter, non-nested sentences.

- I would strongly recommend to completely rework the whole section 5 from line 408 to 451, due to many factors. Above all, this whole section is neither a discussion nor a conclusion in my opinion.
  The first paragraph (l.409-424) basically lays the framework for coupling "deterministic and statistical models" (l.420), which is used as a justification for the proposed method. The paragraph itself seems to be helpful and relevant but should thus be moved to the introduction, somewhere located (and linked) to the paragraph l.105-122.
  This paragraph is followed by two paragraphs that summarize major parts of the publication. l.425-434 summarizes the proposed method; while l.435 - 447 summarizes the reported result.
  The only conclusions drawn can be found in the last paragraph (l.448-451). In my opinion, these conclusion are way too general. Furthermore the authors pre-

sented a new interpolation method, while long-term climate change impacts are modeled into the future, which would require an extrapolation. Thus, the proposed method is not appropriate to predict climate change impacts.
As the authors presented some interesting results in this publication, it should be easy to draw some more immediate and definite conclusions.

- l.136 – 141: To me it is not clear why the authors have chosen Fu's equation. In the introduction to Budyko approaches (l. 129 – 136) the authors introduced a number of adjustments and improvements to the original approach suggested by other studies and highlighted their importance. Fu's equation does not incorporate any of these, but rather a "dimensionless model parameter" (l. 144), which does only control the "partitioning of precipitation into runoff" (l. 145). The authors are kindly asked to give more insights on this decision. Additionally, the calibration of this parameter is just mentioned in l. 146, but not further described.

- l.240 - 244: For my understanding, this is the key paragraph of the methodology as it describes the actual coupling of Budyko with hydro-stochastic interpolation. I would summarize this as: 1.: $R_d(x)$ in equation (18) is substituted with equation (2) by setting $R_d(x) = R$. and calculated for all basins. 2.: The residuals between $R_d(x)$ and observed R is calculated for all gauged basins. Further, these residuals are interpolated for all ungauged basins by "residual kriging" (l.243). and set as $R_s(x)$ 3.: Equation (18) applies as the final result of this study.
Following the cited "residual kriging" from Sauquet (2006) it was not clear to me, how exactly the "residual kriging" is performed on the ungauged basins. The residuals from this study would be described by a first oder polynomial ("Accounting for spatial heterogenity", last paragraph, in Sauquet (2006)), and be combined with ξ_q, the error in residuals. But, for me, it is not clear how this ξ_q or the g from Sauquet (2006) were calculated. From my point of view, the interpolation scheme described in Sauquet (2006) seems to be closely related to the general approach presented by the authors. Then, the delimitation between the

two studies was not clear to me from the introduction. In any case a clarification of how $R_s(x)$ is calculated, how section 2.2 sets in and is linked here would be highly appreciated.

- l.219 - 220: Which scatter diagram are you referring to, here? Furthermore I can hardly imagine how such a diagram would look like. For my understanding, an empirical covariogram relates the separating distances of lag classes the the inner-class covariance observed in the data. Please describe how a diagram like this shall be scattered over the distances between all sub-basin combinations.
  Furthermore, equation (17) presented in line 222 is used to derive a theoretical covariogram. From my understanding (and in fact I am not sure what u_1, u_2, du_1, du_2 are referring to here, see minor point below) this will yield a single value $Cov(A, B)$ for sub-basins A and B. Does the theoretical covariogram then relate $Cov(A, B)$ to $d(A, B)$ defined in (16) (l.217)? If so, a more descriptive and clear explanation in the respective paragraph would be highly appreciated.
  Additionally, do $Cov(A, B)$ (l.220) and $Cov(u_i, u_n)$ (l.178-180) describe the same thing?

- The authors should consider to report their result more consistently and comprehensive. Beside a cross-validation, the authors compare the three different interpolation approaches by comparing the errors each method yielded. This error reporting in line 377-379; 355-356 and 328-331 shall be harmonized and report the same numbers.
  I would suggest reporting the overall minimum, maximum and mean error found in a single sub-basin, along with the minimum, maximum and mean relative error (as share of basin-specific runoff) found in any sub-basin. Both kind of errors can be reported as a absolute (in mm) and relative (in %) number. In my opinion this makes sense as, for example, the sub-basin yielding the biggest absolute error in equation 2 (which is HWH), does not show the biggest relative error (as eg. SQ shows a bigger relative error).

Beside reporting these important numbers, the authors should consider to report these numbers in table 2, as well.

Minor points

- l.322 - 323: This observation is not supported by fig. 3. From my point of view it is not possible to derive the location of a sub-basin from this figure.
- l.83 – 88: The authors make different points here within one long confusing sentence. They are kindly asked to break this sentence down to the core statements of: 1.: runoff is an integrated spatial continuous process, not a field like precipitation; 2.: runoff interpolation must take the stream network into account; 3.: the stream network constraints the water balance up- and downstream.
Furthermore, please clarify the connection between a water balance constraint and assumed runoff properties that can be traced back to field properties.
- l.90 – 91: Please explain "lateral streamflow" (l. 90). What is that and how is it connected to the topic? None of the two presented studies, that shall explain the link between runoff overestimation and "neglecting lateral streamflow" contain the term "lateral streamflow". Please clarify what the two studies actually indicate.
- l.92 – 96. For my understanding, this part is not linked to the other parts of this paragraph or the introduction as far as I read it at that point. Why is this important? Additionally, "hydro-stochastic interpolation" (l. 92) was not clear to me at that point and the authors might consider some more explanation.
Furthermore, the difference between "Euclidean distances" (l. 94) used in "conventional stochastic methods" and the "spatial distance" (l. 95) is too vague for me. Consider adding an explanation.
- Please clarify what "sample" refers to in line 171.

- l. 362-364: The authors are asked to clarify what "trend removal" refers to here, as no kind of trend removal was reported in the methods. From that, what kind of assumptions do you "justify" from applying a trend removal? Do you assume the residuals to be spatially autocorrelated or do you assume an existing spatial autocorrelated random error underlying the residuals themselves, as the "hydro-stochastic interpolation" is performed on the residuals?
Consider extending the corresponding methods part.

- is the du_1, du_2 used in (17) (l.222) and (18) (l.236) the same thing, or does the d from (17) refer to the $d(A, B)$ calculated in (16) (l. 217)? If not, what is (16) then used for? If yes, please clarify the difference of the two used du_1, du_2.

- Please describe what "spatial variance" (l. 259) exactly means here and how it is defined.

- The used precipitation data is described to be a "climatological dataset" (l.287). What kind of data product is this? An interpolated and aggregated map from a observation network? A radar product?

- l.290 How was this interpolation conducted? "ArcGIS" is capable of more than one interpolation method. Please name the method, not the tool.

- What is the "relative error [of] 91 mm"? Is this the absolute error at XZ station, where the relative error is the largest observed of 81.6%?

- l. 340 - 348: Why was HRB divided into a grid? The corresponding methodological description of these results (l. 212 - 217) did not mention this step. Furthermore, for me the link between equation 24 (l.337), equation 25 (l. 349) and figure 4 is not clear. Both equations describe a empirical covariance $C(d)$, while figure 4 shows a "covarinance function" along with an "empirical covariogram". Which one does refer to what here? The authors are kindly asked to make this clearer and the notation more distinct.

- l.351 How shall equation 25 be used to "calculate the theoretical covarinace matrix $Cov(A,B)$"? In line 220 $Cov(A,B)$ was described as a "theoretical covariogram", not a matrix. Is Cov_p in equation 17 than the same as $C(d)$ in equation 25? Is the d in equation 25 then derived from equation 16 for each sub-basin pair A,B? Are u_1, u_2 in equation 16 then the grid points mentioned in l.340 - 348 or the "samples" mentioned in line 171? Clarifying this specific step in the methods wherever appropriate would be highly appreciated.

- l.403-404 Did you mean that figure 7 (a) and (b) overestimate runoff, instead of underestimate, as stated? Because (a) ranges from 145mm - 280mm in the north and (b) ranges from 140mm - 280mm, in contrast (c) ranges from 60mm to 250mm in the north. Adding another sub-figure to figure 7 showing the measured runoff values can make figure 7 even more meaningful. Additionally, I would strongly recommend using the same value ranges for the color codes in figure 7, this will make the sub-figures more comparable and consistent.

Technical points

- In my opinion all the figures should be revised. The figure captions shall be extended and describe all figure elements. This is especially true for figures 3,4,5 and 6. Consider adding legends to figures 3 and 6.

- The authors are kindly asked to revise all their equations. Please make sure, that all used symbols are explained beneath the equation. This is especially true for μ^* and μ^i (l. 189); The sub- or superscripted T used in e.g. in l 192; the undefined symbols u_1, u_2, du_1, du_2 (l.217); $Cov(u_i, u_n)$ in l.178-179, 194-197) . Wherever possible the symbol description shall also include the used unit. The unit was only given in a single case.

Printer-friendly version

Discussion paper

- l.129 – 136: This part is in fact a literature review on Budyko approaches and should thus be moved from the methodology part into the introduction.
- l.334 - 339: In my opinion, these are methods and should be moved to the correct section.
- l.314 - 316: Consider moving this to the methods (l.147-148), where the "calibration" is not further described.
- What exactly is meant by "drainage basin" in line 224? In the preceding text the authors referred to basins and sub-basins.
- Consider replacing "method with semi-empirical approach" (l. 112) with "method with semi-empirical Budyko approach", in order to be even more clear here.
- l.405 The authors should consider replacing "area above BB" with "area upstream of BB" or "area south of BB", to be more precise here.
- In line 388, I would not state that "0.93 [is] much larger than 0.81 and 0.54", as 0.93 - 0.81 is in fact smaller than 0.81 - 0.54. I would rather sayr "cross-validation outcome R^2_{cv} performed best for the coupled method (0.93)..." or something similar.
- The authors are asked to consider adding an overview map locating HRB in China. This could be added to figure 1 or as a fourth sub-figure to figure 7.

References

Sauquet, Eric (2006). "Mapping mean annual river discharges: Geostatistical developments for incorporating river network dependencies". In: it Journal of Hydrology 331.1-2, pp. 300-314.

Printer-friendly version

Discussion paper

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-472, 2017.