Hydrology and Earth System Sciences

Discussions

Open Access

HESSD

Interactive comment

Interactive comment on "Comparing hydrological modelling, linear and multilevel regression approaches for predicting baseflow index for 596 catchments across Australia" *by* Junlong Zhang et al.

Anonymous Referee #1

Received and published: 22 January 2018

The manuscript by Zhang et al. proposes to use multilevel regression for large-scale baseflow index (BFI) prediction. Using 596 catchments in Australia, the BFI prediction accuracy of multilevel regression is evaluated against benchmark BFI calculated by baseflow separation methods. It was found that multilevel regression outperforms classic linear regression, with both models using same explanatory variables. In addition, both regression models outperform two calibrated hydrologic models. The results suggest that climate variability should be considered in order to better understand the effects of explanatory variables on BFI.

The topic is pertinent for the scope of HESS. As far as I am aware of, the application of multilevel regression for BFI is novel. However, in my opinion, the present manuscript is not completely convincing. The assumptions involved in the experiments are not discussed, and the results do not sufficiently support the conclusions. In addition, the manuscript is hampered by a lack of clearness in the description of methods and experiments. So my suggestion is to reject the paper. My justifications are listed below.

1. The implication of using the BFI ensemble mean as "observed" BFI is not discussed. The average value of BFI calculated by four methods ("ensemble mean") is used as benchmark for evaluating the hydrologic models and regression methods. I agree that this seems to be the best choice given that no true observed BFI is available. However, a plot or statistics should be added to show the agreement among the BFI values given by the four methods. Table 3 seems to suggest significant discrepancy among the four products (My interpretation about Table 3 might be wrong, as no explanation is given in text for this table). Note that the discrepancy among the four products could be an indicator of the uncertainty associated with the ensemble mean, and should be taken into account when performing regression. For example, how does the regression residual compare with the uncertainty of ensemble mean?

2. The calibration of the two hydrologic models A major conclusion is that hydrological models overestimate baseflow in Australian catchments. In this study, the two models are calibrated using streamflow. What if the models are calibrated using the BFI as objective function? In that case, would they still be outperformed by the regression models? I missed from my reading why the SIMHYD and Xinanjiang models are selected as representatives of various hydrologic models. It is mentioned in the manuscript and also my understanding that the Xinanjiang model is widely used for humid regions, and it might not be suitable for all Australia catchments. It would be helpful to include results of the calibration goodness-of-fit.

3. The description of the multilevel regression method needs to be improved. This is related to comment # 5. I also made suggestions in specific comments in the hope to

improve the clarity of this section. What is the physical meaning in the BFI context of the correlation coefficient (rho) between alpha and beta? I missed from my reading a discussion about rho in the results. The definition in Eq (6) is contradictory to line 284. Line 284 says rho is a between-group correlation, but it is defined in Eq. (6) as the correlation of intercept and beta within a same group. Please provide more information about how intercepts and slopes/coefficients are estimated, the software package you used or scripts developed, along with other relevant algorithm configurations/settings so that interested readers can follow up your work. The equations in section 3.3 are all developed for the univariate regression case. Since multiple explanatory variables are used, I suggest including a matrix-vector equation in which the slope is a vector.

4. Discussions centering around estimated coefficients/slopes are not convincing. While I appreciate the discussion section which connects findings from this study to literature, I do not think the reasoning there is convincing. For the multilevel regression, the estimated intercept and coefficients show only small differences among climate groups. Given the wide error bar (Figure 8), these differences are not statistically significant. The similarity of estimated intercept and slopes suggests that the effects of explanatory variables (including climatic controlling factors) are similar across climate groups. It is risky to conclude that the effects of P, ET, and F vary among climate groups. In terms of understanding the controlling factors in different climate groups, it seems that multilevel regression provides limited advantage over the classic linear regression.

5. Discussions centering around cross-interaction are unclear to me. Cross-interaction seems to be an important concept and a major strength of multilevel regression. The term is brought up in several places but never clearly defined for the context of BFI prediction. For example, I do not understand why interactions crossing various group levels are primary drivers to influence baseflow processes (line 255). What is the interaction refer to? Interaction between which and which?

6. I like it that the manuscript is well-structured. The presentation quality does not

Interactive comment

Printer-friendly version

Discussion paper

meet the standard of HESS. As mentioned in my comments above, the description of methods and experiments is not complete or precise for reproduction. The manuscript would benefit from a thorough proofreading and English language editing. There are repeated sentences (line 405-410), incorrect reference to figures (e.g., line 448 Figure 9) and equations (line 275), typos, and grammar mistakes.

Specific comments

Line 225, do you mean that in leave-one-out cross validation, the parameters of the models are filled in by either taking the calibrated parameter values in the closest catchment or a combination of parameter values from several basins that are both spatially and hydrologically close?

Eqn. (2). Given line 247, "we further assumed that the effects of those predictors on BFI vary with climate zones...", alpha and beta should be replaced with alpha_j and beta_j, respectively. Then what is the difference between (2) and (3)?

Line 356. Figure 7 shows the validation results, which support the conclusion that multilevel regression outperforms classic linear regression. However, this is not the case for Figure 6, which is based on the data used for regression. Multilevel regression uses more parameters, so it is anticipated that it will fit the data better.

Line 360. What are possible reasons that the two regression models perform similarly for the winter rainfall climate group?

Line 383 "model structure is more important than parameterisation" The reasoning here is unclear to me. Did the three parameterisation schemes lead to different parameters?

Table 3 is not referred to in text. How is the statistics calculated? For example, is the value of 114 calculated by fitting the regression model to UKIH results, or using the regression model fitted to the ensemble mean?

737, 2017.