

Interactive comment on “Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modelling” by D. Duethmann et al.

Anonymous Referee #1

Received and published: 7 December 2012

Review of Duethmann et al.: Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrologic modelling

This manuscript presents an approach to correct/adjust downscaled WRF reanalysis precipitation using station observations, and evaluates the resulting precipitation datasets, by first comparing them directly against station data, those derived from interpolation using multiple linear regression (MLR) and inverse distance weighting (IDW), as well as other spatial precipitation datasets, and then comparing their performances (as a source of areal precipitation) on simulated discharge via hydrologic modelling

C5706

and calibration. The authors conclude that the presented downscaling approach is effective especially in data sparse regions, and introducing a precipitation bias factor in the calibration process is useful for comparing and differentiating between different precipitation datasets.

While I agree with the authors on the overall methodology, I do have a few major concerns on the implementation, experimental design and overall presentation. I hence recommend a major revision of the manuscript before it is accepted for final publication in HESS.

Major comments:

1. The interpolation approaches

In Section 2.2.3, the description of interpolation of station data using downscaled reanalysis data is quite confusing. Personally, I think it would be more appropriate to describe the approach as “Adjusting downscaled precipitation data with station observations”, rather than “interpolation of station data using spatial fields from downscaled reanalysis data”. This is because the scaling factor is applied to the daily downscaled reanalysis precipitation data to derive the final products (WRFadj-all and WRFadj-ind), while the station data is only used to derive the scaling factor. Also, there exists a technical issue with the approach: why not using the relationship between monthly station data and monthly downscaled reanalysis, since it is more reliable/stable than the relationship between daily station data and monthly reanalysis?

In addition, for the IDW approach, although it is commonly used, some brief description and references should be provided (e.g., how is the distance calculated? Is it based on x, y, z or x and y only?) This applies to the IDW approach mentioned in Sections 2.2.4 and 2.2.5.

In Section 2.2.4, some explanations for the stepwise backward vs. forward MLR approaches (as well as relevant references) need to be provided.

C5707

Personally, I think it would be more effective if equations are used to summarize the approaches presented 2.2.3, 2.2.4, and 2.2.5.

2.The calibration experiments

My major concern on the experimental setup is that the different precipitation datasets are evaluated for four different calibration periods without any validation. As the paper also points out, calibration tends to have the parameters adjusted toward compensating for the errors from other different sources. As a result, a precipitation dataset that performs best in a calibration period may not be the best for an independent validation period. In other words, the approach by elevating the performance within calibration periods is essentially flawed. Hence, I would recommend that the authors use two of the four time periods for calibration (e.g., 1st and 3rd) and the other two for independent validation (e.g., 2nd and 4th), and evaluate the precipitation datasets based on their performance in the validation.

The results presented indicate that the precipitation bias factor has a dominant influence on the calibration process, making it less effective on constraining the other parameters. My recommendation would be to conduct a two-step calibration experiment by first calibrating all the relevant/important parameters (as it has already been done by the authors), and then follow up with a second calibration with the precipitation bias factor fixed at the value(s) from the first calibration, in order to more effectively constrain the other parameters.

The dominance of the precipitation bias factor also makes it less meaningful to examine the sensitivity of the bias factor to inputs and parameters while it is being calibrated. A more sensible approach would be to assess the sensitivity of various inputs and parameters prior to calibration and identify the most sensitive/important parameters and inputs to be included in the calibration process. Including non-sensitive parameters in a calibration experiment may interfere with the constraining of other parameters, rendering the calibration process ineffective.

C5708

The paper only discusses the parameter distributions from one calibration case (3.2.1). Please be specific about the sub-catchment, precipitation dataset and calibration period for this case. What about the rest calibration cases? Are the parameter distribution patterns in those cases significantly different from (or very similar to) the presented case? And why?

Finally, more explanation/clarification of the optimization algorithm is needed (Page 10735, last paragraph of Section 2.4.4). What does DDS-AU stand for? How are the short DDS runs used to assist the long run?

3.The presentation and overall structure

The overall presentation and structure of the paper need to be improved. For example, the paper loosely wrap up too many pieces of information into Section 2, including study domain, datasets, interpolation approaches, the evaluation approaches, the hydrologic model, the calibration algorithm, and the calibration runs etc. These need to be more tightly re-organized into smaller, more distinctive sections.

The paper first discusses the point based evaluation in Section 2.3 and then presents the results in 3.1.1. For better organization, one could give an overall summary in an earlier section on the evaluation strategy and then discuss the evaluation details and results together in a later section. For example, Section 2.3 can be combined into 3.1.1, to make things easier to follow.

Section 3.1.3 (Comparison to global gridded datasets) seems to be out of place and not making much contribution to the study. More importantly, the downscaled reanalysis precipitation datasets (WRFadj-all and WRFadj-ind), the focus of this study, are not included in the comparison. Hence, this section can be safely removed from the paper.

Minor comments: P17036, L7: shouldn't it be Table 4? P10738, L21: what does "this method" refer to? Please be more specific. Please consider increasing the font size of texts in the plots for better readability in Figures 5, 6, 7, 10, 11, 12. Figure 6: what is

C5709

'WRF_{dir}'?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 10719, 2012.

C5710