Hydrol. Earth Syst. Sci. Discuss., 9, C5005–C5010, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C5005/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Distributed hydrological modeling in a large-scale watershed of Northern China: multi-site model calibration, validation, and sensitivity analysis" *by* S. Wang et al.

S. Wang et al.

wangshp418@yahoo.com.cn

Received and published: 30 October 2012

Anonymous Referee 1

The authors demonstrate that using the multi-site calibration can have better simulations of streamflow compared to the single-site calibration in a distributed hydrological model, MIKESHE. They use three different kinds of model performance criteria to evaluate the model results; they calibrate at the period of 1991-1995, and validate at the period of 1996-1999. Generally, when more information are used to constrain the model, the model should have better performance and the authors try to demonstrate

this fact. However, the results in this manuscript seems plausible to me. More works need to be done before it can be published. Below are some comments.

Response:

We would like to thank reviewer 1 for the valuable suggestions and comments. Generally, we agree that themulti-site calibration protocol could not improve model performance except for the Dage station in the calibration period. However, considering that the model behavior has reached a compromise between the three stations, and that performance measures were generally satisfied for all the examined stations, we concluded that the multi-site calibration protocol applied in the analysis has advantages to the single-site calibration protocol. We believed that the model performance would be improved when more information with respect to spatial variability of geo-hydrological properties was available. In the revised manuscript, we have given specific explanations for the degraded model behavior of Xiahui and Daiying stations, which included the errors and uncertainties with the representation of spatial variability of geo-hydrological properties, and the errors with the specification of groundwater divide (Please see the sections of "Multi-site model calibration and validation" and "potential source of errors"). Also we have rephrased our conclusions.

Major comments:

1. In the multi-site calibration (section 3.2), the authors change the value of Ks from 2e-6 to 4e-6, but this parameter is fixed in the single-site calibration. Maybe the authors want to add this calibration parameter in the single-site calibration as well. By doing so, the comparison will be more consistent.

Response:

Thanks for the good question. Indeed, a comparison would be more consistent when the parameter set subject to the multi-site calibration protocol was fixed. In our original modeling analysis, the Ks of both multi-site protocol and single-site protocol were all subject to model calibration. The initial value of Ks (for one of the dominant soil profile, i.e. Calcaric Cambisols) in the single-site calibration was 0.68e-007 m.s-1, and the final value was 2e-007 m.s-1. Whilst the final value of Ks in the multi-site calibration protocol was 4e-007 m.s-1. We have a short explanation on this in the revised manuscript (please see the section of "Model calibration, validation, and sensitivity analysis").

2. Both single- and multi-site calibration have the problem when simulating the low flow conditions. I suggest the authors use some criteria having the "LOG" formula that can better constrain the low-flow conditions. Both EF and RMSE focus on the "peak flow" instead of low-flow and they basically tell the same ranking so the authors may consider that just use one of them, and the other criteria use some cost function with "log" formula, such as "log(obs-model)".

Response:

Thanks for the good suggestion. In the revised manuscript, we have removed the RMSE, and kept the EF as one of performance measure. We agree that it is necessary to use some cost function with Log formula to evaluate the model. Accordingly, we have calculated the mean residual (simulation minus measurement) (ME) as one of performance criteria, after the discharges were taken logarithmic transformation. Unlike the EF which usually takes more weight for peak flows, the ME will give an equal weight for all the maximum and minimum discharges, corresponding to the mean model bias of the results (El-Nasr et al., 2004). We believed that, by using the EF and ME as the performance criteria, the model behavior could be comprehensively evaluated.

3. Is there any figure showing the spatial distribution of soil depth? In p5709 line 9 to 16, the authors say "Since the northern of the Chaohe watershed is adjacent to the Inner Mongolia Plateau, the northern part of the watershed is commonly characterized by high soil water storage capacity due to the deep soil profiles, which caused much of water stored in the unsaturated zone available for recharging the groundwater and discharge the river flow subsequently. However, in the middle and downstream area of the

watershed, the thin soil profiles resulted in small soil water storage in the unsaturated zone and less recharge to the saturated zone," I think if there is one figure showing the spatial distribution of the soil depth, it will be great for the readers. Because the authors try to increase the Ks to increase the soil water storage, so it's necessary to let the readers know what is the spatial difference of the soil depth.

Response:

Thanks for the good suggestion. We agree that it is necessary to show the spatial distribution of soil depth in order to get the readers to easily understand our statement (i.e. the upstream area of the watershed is characterized by deep loess profiles with high soil water storage capacity, whilst the middle- and down- stream areas by thin soil profile). Nevertheless, as the soil dataset that we have used in the analysis was derived from HWSD (The Harmonized World Soil Database), which only records the info of the reference soil depth of 0-100cm, there is no detailed info available with respect to the distribution of soil depth, and thus in the manuscript we only gave a reference literature as an argument of our statement.

4. In page 5710, line 15-19:" Generally, the model was insensitive to the parameters of ET and overland flow, but was sensitive to the parameters of unsaturated zone and saturated zone modules (Table 3). This indicated that process of ET and overland flow were less important in affecting streamflow generation of the watershed, whilst unsaturated flow and saturated flow played an important role." This statement has some problems (especially I highlight in red). The reason could be due to the chosen of ET parameters are not sensitive to ET, but it doesn't mean that ET is less important in affecting steramflow generation. On the other hand, it doesn't mean the unsaturated flow can play an important role. The authors need to future work on this part. Maybe some time series of ET, streamflow, and so on.

Response:

Thanks for the comments. We agree that hydrological processes of ET, OL, UZ, and

SZ were all essential for hydrological cycle and water balance. And we have rephrased the whole section of sensitivity analysis, and removed the statement of "the model was insensitive to the parameters of ET and overland flow... whilst unsaturated flow and saturated flow played an important role". Generally, in the revised manuscript, we only focused on comparing the influences of the various parameters on performance measures between the stations.Using the simulated different model response, we tried to emphasize that runoff generation mechanisms for the up- and down-stream areas of the watershed that may be dominated by different groundwater recharge mechanism. Considering the likelihood functions are usually used as an alternation of the predicted variables to examine sensitivities of parameters (Beven, 2001), we did not specifically analyze the influences of the varied parameter values on the time series simulation results (such as ET, unsaturated flow, and so on). Instead, we have only employed performance measures of EF and ME in the analysis.

5. Finally, the most important one, from table 2, I really can't see the benefit of using the multi-site model calibration when compared to the single-site calibration. Especially for Dage station, the RMSE doesn't increase for the validation period, and the R even decreases for the calibration and validation period. Hence, the last sentence of the conclusion is not convincing at all that "Multi-site model calibration protocol would greatly further reduce the modeling errors resulting from the inherent great spatial variability", which is the main conclusion of this paper. Therefore, I suggest the authors should focus on this topic in the revision. Again, the key issue is on the low flow simulations in the Dage station, so I suggest that the author should use other statistical criteria (as the above comment 2) to make this statement more convincing.

Response:

Thanks for the comment and suggestion. In the revised manuscript, we have removed the conclusion of "Multi-site model calibration protocol would greatly further reduce the modeling errors resulting from the inherent great spatial variability", and used the criteria of EF, R and ME to evaluate the model. We agree that the model behavior in

terms of EF, R and ME, from multi-site calibration protocol, did not improve for Xiahui and Daiying stations, and R and ME was even decreased for Dage station. Also, we acknowledged that it was the errors on the representation of the spatial variability that has impeded the model improvement. However, as stated previously, considering that the model behavior, the multi-site calibration protocol, has reached a compromise between the three stations, and that performance measures were generally satisfied for all the examined stations. Thus, we rephrased our conclusion as "The multi-site calibration protocol applied in the watershed has advantages to the single-site calibration protocol, and more improved model behavior could be arrived when the detailed information with respect to spatial variability of geo-hydrological properties was available". As for the modeling errors on low flow simulation, we attributed it to the potential errors of representation of groundwater divide, though we agree that it was necessary to calculate some cost function with "log" formula to evaluate the model behavior.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 5697, 2012.