Hydrol. Earth Syst. Sci. Discuss., 9, C1266–C1272, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C1266/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "A new approach to model the variability of karstic recharge" *by* A. Hartmann et al.

A. Hartmann et al.

andreas.hartmann@hydrology.uni-freiburg.de

Received and published: 3 May 2012

We thank anonymous referee #2 for accepting large parts of our proposed modifications and for the two additional comments. First of all the authors want to apologize for the confusion that was caused by the inappropriate citations that were mentioned in the last comment of referee #2. In the following we will try to respond more thoroughly to the two remaining concerns of reviewer.

1. Comparison to other models

It is true that both cited studies (Lindström et al., 1997;Perrin et al., 2003) include a comparison with an older version of the model (HBV-96 vs. the standard HBV and the

GRJ4 vs. GR3J plus a series of other models). The authors wanted to provide examples of studies that introduced a new model which were actually Bergström (1972) for the HBV model, and Edijatno and Michel (1989) for the GR3J model, the predecessor of the GR4J model. Consequently, the HBV-96 and the GRJ4 were further developments of already existing models. The modifications, which resulted in an increase of complexity in both cases, had to be justified by model intercomparisons that were included in the papers we originally cited. For our model, however, there is no former version. It is a completely new development based on the physical understanding of the epikarst functioning. As mentioned in the first reply to referee #2, the focus of this study was on introducing a new model type for karst aquifer recharge and to show that its physics and concepts are correct and suitable, rather than proving that it is superior to other recharge models. We are confident that for our purpose, i.e. using variability and hydrochemical information to constrain model parameters in karst hydrological models, there is no existing alternative model type. Hence, a development of a new model type was necessary.

Either way, we understand the point of referee #2: for practical applications the performance of a newly developed model is hard to evaluate without model intercomparison. With the information provided by the manuscript there is only proof that our model works fine under present conditions and data availability. But there is no information about its transferability. If referee #2 agrees, we will clearly point out in the revised manuscript that our approach is completely new, and as such has to be tested and compared with other established approaches before it can be applied for practical purposes, e.g. water resources management.

2. Split-sample test

Showing that after calibration step 4 the model is able to simulate simultaneously drip rates, tracer concentrations and information about variability, the manuscript provides proof that the model works physically plausible. To show the applicability of the approach for future predictions probability distributions of the parameters "(...)may be used directly to evaluate the uncertainty limits for future events (. . .)"(Beven and Binley, 1992). But Beven and Binley (1992) also mention the value of a split-sample test. However, in our study, it is difficult or even impossible to simply separate our heterogeneous time series into a calibration and a validation period: The series include different time scales (some days for the sprinkling experiment and almost 1.5 years for the seasonal time scale) and different types of information (drip rates and information about variability at both time scales but tracer concentrations only during the sprinkling experiment). Even if all information would be available for both the experiment and the seasonal time scale, the strong climatic variability of semi-arid regions (Bull and Kirby, 2002) would require a much longer observation time to reflect all possible states of the hydrological system. Hence, splitting the available seasonal data into the first rainy season and the second rainy season, both of them different in rain amount and frequency, will most probably result in different parameter sets after calibration, as shown by Merz et al. (2011) for 273 catchments in Austria under even less climatic variability. The only way the authors see to perform a split-sample test with the current data availability, would be a random bootstrapping of two subsets from the whole observation set and to use one of them for calibration and the other one for validation. This we perform as follows:

Split-sample test by bootstrapping

To perform a split sample test by bootstrapping, half of the elements of each time series (Qexp,Qseas,Cexp,CV) were selected from the whole data set by a random draw and the model was calibrated as in calibration step 4 (section 3.3). For validation, the calibrated the model was applied to the excluded data. To account for variability, the bootstrapping procedure was repeated five times.

The resulting five parameters sets and the resulting validation efficiencies, compared to the parameters and efficiencies derived by all the data in calibration step 4, are shown in Figure 1. The positions of the parameters of the bootstrap samples (Figure 1a) indicate that for the parameters adepth, dsoil,max, log Kvert,max, A, Cold and Vold,max the automatic calibration results in similar parameters values as for the whole

set of observations. The same is true for depi,max except for one bootstrap sample. The other parameters show a much higher spread, sometimes covering the entire (e.g. log Klat,max) or a certain part of the parameter range (e.g. nepi). These results are in accordance with the parameter distributions shown in Fig. 5 in the manuscript: parameters that show a larger spread in our bootstrap test also show a low identifiability in their parameter distributions: for instance, Klat,max shows a uniform distribution, while nepi has a constant slope only at the second two thirds of its range, similar to its positions visible by the bootstrap calibration. Hence, the split-sample test by bootstrapping did not provide new information about the identifiability of the parameters and the reasons for a larger or smaller parameter spread are already included in the manuscript (section 5.3).

The individual efficiencies in Figure 1b show that for most of the bootstrap samples, the efficiency of the validation data set is inferior to the efficiencies in calibration step 4. A certain decrease of efficiency for the validation data set is very probable for every hydrological model (e.g. Perrin et al., 2001), but a large decrease of efficiency indicates a model deficiency. This seems to be true for NSC,exp: even though the hydrochemical parameters, Cold and Vold,max, do not change much compared to calibration step 4. The reason for this can again be found in the heterogeneity of our data series. During the first day of the sprinkling (see Fig. 2 of the manuscript) only two tracer observations gave proof that there was a tracer reaction already at the first day of sprinkling. If these two measurements are not part of the bootstrap sample, they are not considered in calibration and the model tends to simulate a first tracer reaction only when the second day of sprinkling began. In that case, these two measurements result in a strong decrease of NSC,exp for the validation period. Hence, the split-sample test by bootstrapping showed that especially tracer predictions are very sensitive to changes in the model input. This was already suggested by varying the number of model compartments (see section 5.4).

Considering these results we think that additional information on model performance by the performed split-sample test using random bootstrapping is rather limited. It rather emphasizes artefacts caused by model calibration using a reduced time series and increases the length of the paper. This might add to confusion of non-experienced readers. Hence, we prefer to leave the bootstrapping results in this HESS-D comment and as such to constrain the final HESS paper. However, if referee#2 insists, we are ready to include the analysis also into the revised version for HESS.

References

Bergström, S.: The application of a simple rainfall-runoff model to a catchment with incomplete coverage, Sveriges meteorologiska och hydrologiska institut, Notiser och preliminara rapporter, Serie hydrologi 26, 1972.

Beven, K. J., and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrological Processes, 6, 279-298, 1992.

Bull, L. J., and Kirby, M. J.: Dryland Rivers – Hydrology and Geomorphology of semi-arid channels, John Wiley & Sons, Ltd., 2002.

Edijatno, and Michel, C.: Un modèle pluie-débit journalier à trois paramètres, La Houille Blanche, 113-122, 1989.

Lindström, G., Johannson, B., Perrson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, Journal of Hydrology, 201, 272-288, 1997.

Merz, R., Parajka, J., and Blöschl, G.: Time stability of catchment model parameters: Implications for climate impact analyses, Water Resources Research, 47, 10.1029/2010wr009505, 2011.

Perrin, C., Michel, C., and Andréassian, V.: Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, Journal of Hydrology, 241, 275-301, 2001.

Perrin, C., Michel, C., and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, Journal of Hydrology, 279, 275-289, 2003.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 2443, 2012.

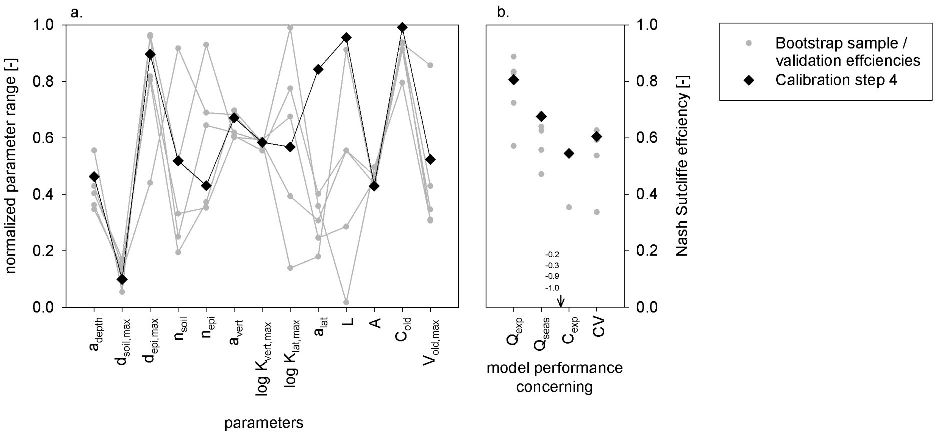

Fig. 1. Results of split-sample test and calibration step 4 using five bootstrap samples from the observations: (a) calibration parameters and (b) efficiencies with the validation data sets