Response to Referee 1: D.A. Hughes

This paper represents the results of a very comprehensive study of data and modelling uncertainties in a relatively data scarse region and therefore makes a valuable contribution to hydrological modelling theory and (potentially) practice. In general terms the paper is also well written, but I found some of the explanations of the methods a bit confusing. However, I am not sure that they can be simplified and perhaps they would become clearer if the previous papers are consulted (something I admit that I did not do). I did, however, find that the discussion section seemed a bit long and somewhat repetitive. I would therefore encourage the authors to look at making the finala section more concise trying not to repeat too much of what is already in other parts of the text.

Reply: We thank D.A. Hughes for his positive words about our paper and the constructive comments that will help to improve and clarify the paper. We agree that the discussion section needs rewriting and will therefore shorten it and add subheadings to give it a better structure and clearer presentation.

A final comment relates to the high degree of uncertainty in the simulations (and some of the observed data). I would like to have seen some comments about this in terms of the practical use of water-balance results. Mention is made of robust predictions under different circumstances and the possible need for more regionalised information. What does this really mean in terms of the use of modelling results for '..effective management of these resources' and can such uncertain results be of any value for water resources management? I realise this is not the main topic of the paper, but I do think that some concluding (possibly even speculative) remarks could be made about this issue.

Reply: This is a very interesting point, especially when it comes to predictions in ungauged basins in a region where data inconsistencies can be expected. We accounted for many different types of uncertainties when making our predictions, and in basins where the data were found to be reliable this resulted in generally reliable simulations where the water balance was constrained according to the regionalised FDCs (where FDCs have a long history of use for different types of water management, e.g. Vogel and Fennessey, 1995). The width of the predicted uncertainty was therefore dependent on the uncertainty in the regionalised FDCs and was in the best cases almost equal to that from the local calibration and in the less accurate cases much wider.

The uncertainty estimates from our method give much more information for water management than deterministic model simulations would have had. Having a prediction with high uncertainty is also much more valuable than having no information at all for an ungauged catchment, but when using that prediction for water-resources management the quality of the information that went into making that prediction needs to be taken into account. In using this method for a completely ungauged basin in this region it would thus be advisable to carefully scrutinise the quality of the precipitation input data to assess potential effects on the predictions.

In the cases where the data were inconsistent, our analyses showed the need for additional information and improved data, which is important knowledge for water-resources management. Since our method would be used for predictions for ungauged catchments in a region with other nearby gauged catchments, much information about the dataset consistency would be found by making the types of analyses for the gauged catchments as we made here and by testing the

method in cross-evaluation for the gauged catchments first to learn about the different types of uncertainties that are affecting the simulations. In this region it was found that for many basins the predictions should often not be expected to be accurate for each individual day because of input data errors, which should be kept in mind when the information is used for water management.

We will add some remarks about this to section "6.4 Concluding remarks" in the revised manuscript.

Other minor comments: The reference to 1000-2500mm lower estimates of precipitation (end of section 4.1) is very important but not perhaps emphasised enough as a major source of uncertainty.

Reply: We agree that this is an important problem for the two basins where it occurred, however this problem of largely overestimated precipitation in the CRN073 dataset only occurred for two Panamanian basins that were clear outliers on the Budyko curve, and was not found to be a general problem. For these two basins no behavioural simulations were found in the local calibration. Since these data inconsistencies were identifiable from the data screening we made, this highlights the value of making such analyses in this type of regional modelling. However, in completely ungauged basins the discharge-dependent data screening methods we used would not be able to identify such data problems, and in the paper we therefore stressed the need to develop data screening methods that do not rely on observed discharge data. We will add a short note to emphasize the precipitation uncertainty in section 4.1.

Line 22 of section 4.5: The sentence 'Simulations with correlation in deviations across successive EPs then obtain a lower weight..' is not very clear to me and perhaps can be better explained.

Reply: This means that a simulation with a systematically over- or underestimated FDC for (part of) the flow range will get a lower weight, but that such simulations are still acceptable. We found it important for allow for such (non-stationary) biases since the data analyses showed that they were frequent in the discharge and model input data. The rating-curve analysis of the Honduran stations showed several stations with under- or overestimated discharge and residuals that varied systematically with flow, and there were also temporally non-stationary rating curves. The screening for dataset inconsistencies and visual analyses of the data series also showed that several stations had likely non-stationary errors in the precipitation data. We will add "…, i.e. a systematically under- or overestimated FDC for (part of) the flow range can still be behavioural but get a lower weight." to the end of the sentence.

Minor errors: Last line of section 5.3 'constraint' should be 'constraints' or 'provide an additional constraint'. Similarly line 4 at the top of the 2nd paragraph of scetion 6 (constraints).

Reply. Thanks, we will change this.

References

Vogel, R. M., and Fennessey, N. M.: Flow Duration Curves .2. A Review of Applications in Water-Resources Planning, Water Resour Bull, 31, 1029-1039, 1995.