**Hydrology and
Earth System
Sciences
Discussions**

Interactive comment on "Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin" *by* W. Terink et al.

Anonymous Referee #2

Received and published: 10 March 2010

Summary

Terink et al. (2010) present a bias correction for temperature and precipitation of dynamically downscaled reanalysis data. The bias correction corrects the model output for deviations in the mean as well as in the standard deviation and coefficient of variation for temperature and precipitation respectively. The purpose is to correct downscaled reanalysis data in a way that it can be used to calibrate a hydrological model for a climate impact study. They found that the bias correction performs particularly well if it is evaluated on the same data as it has been estimated for and less well if evaluated on independent data.

General comments

The authors have carefully assessed the performance of the ERA15/REMO data in the Rhine catchment. They further thoroughly evaluated the employed bias correction method of Leander and Buishand (2007). In particular, they cross-validated the bias correction model to test for possible overfitting. Such cross-validation has rarely been done in studies of statistical correction of climate model data. Climate impact researchers nowadays become more aware that the ubiquitous assumption of constant model bias might not be true. Cross-validation does not solve the problem but such analysis needs to be done more often in the context of non-constant climate model biases. Thus, this study is worth to be published with some minor revisions.

Detailed comments

Section 1, p. 223: It would be nice if the authors cited some studies of hydrological impact studies that did not focus on the Rhine catchment. For example in the UK, a lot of work has been done that could be worth to mention in a paper focusing on the Rhine catchment.

Section 3.1, p.228: The explanation of the way the correction is applied to basin and daily averages and than further transmitted to the individual grid cells and 3 hourly values is very confusing. Please try to explain it more clearly. I understood it in the way, that the power law transformation has been applied to daily basin mean values but that the spatial disaggregation is a linear function. I had to read it several times and still it is not fully clear to me. I suggest inserting a mathematical formula to prevent any misunderstanding.

Section 3.2, p. 229: It would be very interesting to see some numbers or figures of the block-length sensitivity study. Monthly bias correction factors are commonly chosen

in impact studies. However, it is important to see that the block length does have an effect on the result. More details about this part of the study would be very helpful for the community. Furthermore, Shabalova et al. (2003) did not – at least to my knowledge - show results of the effect of averaging window width on the sampling bias. On p. 240 of Shabalova et al. (2003) they state that they used a 70-day window to reduce the sampling variability. No bias level of 0.5

Section 4.2.4, p. 237: According to figure 5, I understand that for return periods larger than 20 years, the uncorrected data fits the observations better than the corrected one whereas for return periods smaller than 20 years, no difference is visible. In the text it is written the other way around. Please correct this inconsistency.

Section 4.4, p. 240: Since the dominating processes of rainfall generation change throughout the annual cycle, the analysis of correlation between temperature and precipitation should be done on e.g. a seasonal level. If this has been done already and the results do not deviate from the presented analysis, it should be mentioned that the result is robust also on a seasonal level.

Section 5.1, p. 243-244: In the calibration-validation analysis, some overfitting problems were diagnosed. These problems should be addressed in the conclusions as well e.g. in point 6 one could specify what reasonably well means and that some problems particularly in the RMSE-diagnostic exist.

Section 5.2, p. 245: Following the cross-validation analysis, the sentence "Thus correction parameters derived for a current climate can be used to correct precipitation and temperature in a future climate" is not well supported. Of course, one can apply the correction parameters and indeed this is often done. However, Terink et al. (2010) show here nicely, that the assumption of constant model bias might not hold.

Whole text: The word significant is used too often without statistical tests underlying the statement. Such use should be corrected to prevent its misinterpretation as a statistical test result. E.g. what do the authors mean by "the spread of . . . is less significant than

. . ." (section 4.2.3 on p. 235).