Earth Syst. Dynam. Discuss., 4, C125–C131, 2013
www.earth-syst-dynam-discuss.net/4/C125/2013/

Earth System Dynamics Discussions

Interactive comment on "A trend-preserving bias correction – the ISI-MIP approach" *by* S. Hempel et al.

Anonymous Referee #2

Received and published: 20 March 2013

Summary:

The authors present a bias correction method for GCMs that relies on a quantile mapping approach and is based on the methods of Piani et al. (2010) and Hearter et al. (2011). Additionally, it preserves absolute rends in temperature and relative trends in other variables. The paper gives a very detailed technical description of the bias correction procedure and presents some evaluation results.

Major comments:

1) My major methodological point has already been raised by Reviewer 1 and John Smith: A major caveat of this study is that the authors only evaluate their bias correction method using the same calibration and evaluation period. In such a setup it is easy to

achieve nearly perfect evaluation results regarding the distribution. However, since the bias correction is applied to future climate simulations, some kind of cross-validation, split-sample test, or any other evaluation that demonstrates how your method performs if calibration and evaluation periods are not the same has to be added. I want to add, that the current evaluation (same calibration and evaluation period) is in this specific case still important and shouldn't be removed from the study, since it clearly demonstrates substantial technical issues of the author's implementation of bias correction (see next comment).

2) A quite disturbing feature of the presented bias correction method is its weak performance. After bias correction you find, e.g., remaining deviations of more than 5K in q10-q50 ranges and remaining errors of comparable size in other evaluations. If you consider that in your evaluation setup (calibration period = evaluation period) a proper mapping would trivially lead to perfect evaluation results, these errors are very large. You correctly state that they are caused by the parametric fit (with only one parameter in case of temperature and 3 parameters in case of precipitation). Knowing that, why do you stick to this approach? Don't you expect to achieve better results by other (e.g. non-parametric) methods?. Please refer to, e.g., Gudmundsson et al. [2012] for an evaluation of different implementations of quantile mapping. If you stick to your approach, you have to discuss it compared to other published approaches and to argue why your approach is suitable for your application.

3) Now a quite fundamental issue: The trend-preserving nature of the presented method is achieved by mapping anomalies instead of absolute values. As a consequence, a certain correction value is not attached to a certain absolute value of the corrected variable, as it is the case in "traditional" quantile mapping. One usually argues that a climate model can be expected to have a typical temperature error at, say, -10°C daily mean temperature and another (potentially different) typical error at +25°C. Those temperatures are obviously related to different weather situations and it can be expected that models feature typical errors related to each weather situation. This

concept is of course a wild simplification, but at least it roughly explains why quantile mapping can be successful when it is applied to future simulations. However, in your application, a specific correction value is attached only to a temperature anomaly, i.e. to different absolute temperatures and consequently to different weather situations. Why should the same correction value be appropriate for different weather situations? How can you argue that? This problem is particularly severe when you apply your method to future simulations, where similar anomalies can be related to quite different absolute values. I can't judge how severe this problem is, but since it is a fundamental conceptual issue, it should be analyzed and discussed in your study.

4) The troubles described above follow from the aim to preserve absolute or relative trends. As you correctly note, it is currently not so clear whether preserving trends desirable or not (P51, L25, Ehret et al. [2012]). I want to add that there is strong indication from Christensen et al. [2008] and Boberg and Christensen [2012], and even stronger indication from unpublished work, that certain bias correction methods (including quantile mapping) may modify trends in a way that can be regarded as improvement.

You don't argue clear enough why the trends and relative trends should be preserved in your application. Your arguments on P54, at least in their present form, are not convincing. Please elaborate on why it is so important for impact models to have consistent temperature trends over land and over sea. And if so, why is it then acceptable to have inconsistent trends (= consistent relative trends) in all other variables than temperature?

Otherwise, also taking into account the arguments in the previous paragraphs, I don't see how the presented method is a step forward, compared to what has already been published.

5) You provide a bias correction method for quite a bunch of meteorological variables, but evaluate only temperature and precipitation. Please also show results for the other variables, at least as supplementary material. This is very important, since your data

are used by impact modelers in ISI-MIP and plenty of impact-publications will be based on it. This requires a complete evaluation in order to enable the subsequent studies to rely on a well described basis.

Considering these points, I cannot suggest the paper to be published in its present form, but it should be considered for publication after major revisions.

Specific comments:

Abstract: Results (e.g. your main evaluation results) are missing in the abstract. Abstract: Please streamline. E.g., you mention twice that you present a trend preserving bias correction method – once is sufficient. You could, e.g., apply the following structure: Describe all "introductory information" (general information about bias correction and ISI-MIP) in the first paragraph, followed by a paragraph describing the method (its trend preserving nature and all information that is currently in the last paragraph), and finish with results (which are so far missing).

P51, L14: Bias correction via scaling (applying a "multiplicative constant") does obviously not conserve trends, as you describe in sect. 3.1. Please correct the sentence.

P52, L13: You refer to bias correcftion also as downscaling tool. However, the downscaling ability of most bias correction methods (including quantile mapping) has limitations, particularly regarding variability, as described by Maraun [2013]. Please mention these limitations (see also comment by D. Maraun).

P52, L19: What do you mean with "more detailed altitude-stratified" apart from what you already said in the previous paragraph? Isn't that exactly what downscaling does? Please clarify.

P54, L5, Why exactly is it essential to ensure consistency between global mean temperature change and bias corrected temperature change (i.e. to preserve the trend)? And why is it the essential conserve relative trends in the other meteorological elements? Please argue in more detail. The trend conservation is the key development of your method and it should be clearly explained why it is advantageous or necessary. (See also general comments)

Chapter 2: Please add a sub-section for the model data (as you did for the observations).

P59, L19: Again the issue of variability. Bias correction in the form presented here cannot improve the temporal structure of the time series. You might get a the variance closer to the observations due to "blowing up" the time series (effect of bias), but the sequence of e.g. precipitation events is not improved. See Maraun (2013) and the comment of D. Maraun for much sharper arguments regarding this topic.

P60, L10: "we adjust the residual distribution of the GCM to that of the WFD using a parametric quantile mapping (cf. Eq. 7)." It is unclear why you refer to Eq. 7 here. It doesn't describe quantile mapping.

P60, L10: "Since temperature is well described by a normal distribution, a linear fit is sufficient." Please explain clearer, that you mean a linear fit to the transfer function. The current formulation is a bit confusing.

P60, L11: You state: "Since temperature is well described by a normal distribution, a linear fit is sufficient." As you show later in your study, the linear fit causes quite some error. It is therefore inappropriate to call it sufficient (see also general comments).

P70, L15, Eq. 29: It has been already discussed in section 3 why the method preserves temperature trends. There is no need to repeat that here.

P71: Same for precipitation

Figures 7,8,9, and 10: For easier reading, please clearly specify in the figure captions what exactly you show. You do this mostly in the text of the paper, but it would be helpful for the reader also having it in the figure caption. In particular, you should clearly state what you mean with "deviation" (I assume you mean: model – observation) (fig8), or with "differences in trend" (fig 7) (I assume it is the difference between the trend of the

corrected and the uncorrected model).

Figure 8 and its discussion: After bias correction you find remaining deviations of more than 5K in case of temperature. This is large, if you consider that in your evaluation setup (calibration period = evaluation period) a proper mapping would trivially lead to perfect evaluation results. You correctly state that the remaining deviations from the observations are caused your parametric fit (with only one parameter in case of temperature and 3 parameters in case of precipitation) of the transfer function. Knowing that, why do you stick to this approach? (See also general comments).

Figure 9 and its discussion: You show quite large differences between the basic and the extended version. What consequences does this have for the impact modelers? Please clearly quantify the total errors in the data that has been delivered to ISI-MIP (not only the differences to the enhanced version).

Equations in general: The paper contains a large amount of equations (30), but some of them hardly contain helpful information in addition to what is already very clear from the text, or are more or less duplicates (e.g. Eq. 3 and Eq. 22). Please consider removing some of the not so important equations.

References:

Boberg, F. and J. H. Christensen (2012), Overestimation of Mediterranean summer temperature projections due to model deficiencies, Nature Climate Change, 2(6), 433-436, doi:10.1038/NCLIMATE1454.

Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher (2008), On the need for bias correction of regional climate change projections of temperature and precipitation, Geophys. Res. Lett., 35(20), L20709, doi:10.1029/2008GL035694.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrol. Earth Syst. Sci., 16, 3391–3404, doi:10.5194/hess-16-3391-2012, 2012. 50, 51, 53.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using quantile mapping – a comparison of methods, Hydrol. Earth Syst. Sci. Discuss., 9, 6185-6201, doi:10.5194/hessd-9-6185-2012, 2012.

Maraun, D., Bias correction, quantile mapping and downscaling: revisiting the inflation issue. J. Climate, Online First, 2013.

Interactive comment on Earth Syst. Dynam. Discuss., 4, 49, 2013.