Biogeosciences Discuss., 10, C7719–C7723, 2014
www.biogeosciences-discuss.net/10/C7719/2014/
© Author(s) 2014. This work is distributed under
the Creative Commons Attribute 3.0 License.

Biogeosciences Discussions
Open Access

# *Interactive comment on* "What is the importance of climate model bias when projecting the impacts of climate change on land surface processes?" *by* M. Liu et al.

**Anonymous Referee #2**

Received and published: 7 January 2014

Review of Liu et al. "What is the importance of climate model bias when projecting the impacts of climate change on land surface processes?"

In their paper, Liu et al. present a series of climate change impact experiments wherein bias corrected and non-bias corrected regional climate model data are used to drive multiple models as might be used in climate change impact studies. This is a moderately useful study as people often apply bias correction and or downscaling without much thought; however, there are some major areas that could be improved in the current study.

I have two more serious concerns with the present study. First, their statistical tests of



significance are probably inappropriate, and second, their interpretation of the results over HJ Andrews ignores the fact that these results come from a different model.

On the statistical tests, the authors apply a simple student's t-test to test the significance of the difference in the change signal between the BC and NBC results. However, by taking the spatial field of differences as a population, their application of this method is treating real spatial variability as random, independent samples. This is not a fair assumption and likely results in a large number of their tests appearing non-significant due to the large (and non-random) spatial variability. A more appropriate test might include a field significance component where each grid cell (or each small/similar region) is tested independently and then the significance of the differences across the domain are evaluated using a Field test or Walker's test (e.g. Folland and Anderson 2002; Livezey and Chen 1983; von Storch, H., 1982; Wilks, 2006). This is important because the authors are claiming that bias correction does not necessarily matter for some of their results, and I would claim that it would matter for more of their results if they used a more appropriate statistical test. I am not sure that the tests I am suggesting here are the perfect ones and would love to see the authors find better tests if they do not like these methods.

Folland, C., and C. Anderson, 2002: Estimating changing ex- tremes using empirical ranking methods. J. Climate, 15, 2954– 2960. Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. Mon. Wea. Rev., 111, 46–59. von Storch, H., 1982: A remark on Chervin-Schneider's algorithm to test significance of climate experiments with GCM's. J. Atmos. Sci., 39, 187–189. Wilks, D. S., 2006: On "Field Significance" and the False Discovery Rate. J. Appl. Meteor. Climatol., 45, 1181–1189.

On their interpretation of results in HJ Andrews, they state that the ET Climate Change estimates in HJ Andrews are affected by bias correction, while ET in the domain as a whole is unaffected. However, the larger domain ET is simulated using VIC at a lower spatial resolution, while in the HJ Andrews subdomain they use RHESSys and what

appears to be a higher effective resolution in their model domain. These are two very different models with different embedded (or missing) ET, hydrology, and ecological equations/assumptions. As a result, I don't think they can say anything specific about HJ Andrews from the results they have presented so far. A more important point may be that the effect of BC on climate change impact may be different for different hydrologic models. In either case, it would be appropriate to compare the results from RHESSys to the results from VIC output just over HJ Andrews. If the results are identical between models, then they can make the statement they made initially, that regions such as HJ Andrews are affected differently than the domain as a whole. If the results differ, then they are left with the conclusion that the choice of hydrologic model matters. This will also be affected by the fact that the two models have different methods of calculating driving meteorological variables from the temperature and precipitation data supplied. Ideally they should also run RHESSys with the MTCLIM forcing generated by VIC and vice versa. This may be beyond the scope of the current paper, but a short note to point out the met forcing could be an issue would be useful to include. Similarly, were both of these models locally calibrated? I'm fairly certain that RHESSys has been calibrated for HJ Andrews at some point, and VIC has been calibrated for the CRB, were those calibrated parameters used here? Are their results sensitive to model parameters?

While those strike me as the two most important problems with the current manuscript, there are other major areas that need to be looked at. First, they note that BCSD is used to generate forcing data from e.g. CMIP5 data. However, the dataset they cite is producing monthly data, not daily data, and it is based off of the GCM precipitation data instead of a moderate resolution regional climate model as in their case. They should include some discussion of the fact that they are only applying relatively modest bias correction after a detailed dynamic downscaling, while statistically downscaling GCM output requires much larger bias corrections, and as a result, it is more likely that climate change signals would be affected by the bias correction / statistical downscaling step in that case.



Along those lines, I was surprised that there was almost no discussion of the WRF modeling that they performed. This is not an inconsequential component and substantially affects the interpretation of the implication of their results for other studies. At a minimum they should report the physics options used to run the WRF model and mention that the WRF domain was (presumably) that of figure 1. In which case, what did they use for SSTs, are these taken from the driving GCM without bias correction?

The paper as a whole could also use additional discussion and interpretation to make it more useful to a broad audience. For example, their analysis of the relative importance of temperature and precipitation is useful, but more discussion of why they see the results that they see would be useful. For example, they note that precipitation has a greater affect on SWE, but this is probably only true in parts of the domain (this is one of the problems with averaging over a large and heterogeneous domain). In particular, the lower elevation bound of the seasonal snowpack is probably substantially affected by small changes in temperature because of the temperature threshold at 0C. These sorts of thresholds are critical in understanding how the results from this study may affect other variables or locations. If the system response is highly non-linear or crosses some threshold, than BC will matter, if it is linear than BC may not matter.

Similarly, additional discussion of the importance of coupled simulations would be useful. The authors premise and primary conclusion is that there are tradeoffs between using BC data offline and NBC data in a coupled framework. However, the authors do not seem to have presented any results or substantial discussion showing the utility of a coupled simulations. I would assume the primary utility would be for processes that feedback to the atmosphere and affect the local weather patterns, for example, ET and possibly VOCs that form aerosols (if the atmospheric model actually accounts for aerosols explicitly), as both will affect precipitation, but I do not believe this is ever described in the paper.

Finally, the figures as they are presented in the print and discussion versions of this paper are illegible when printed (they are far too small). I suspect this may be due to

formatting potentially applied by the publisher after the authors submitted it, regardless it needs to be fixed at some step in the process. Even after fixing the size issue, figure 6 might be better with different colors for some of the lines, the subtle shades of grey are hard to distinguish.

There are also many minor grammatical errors in the paper too numerous to document here, hopefully the authors can get a proof-reader to fix these errors.

---

Interactive comment on Biogeosciences Discuss., 10, 17145, 2013.