Clim. Past Discuss., 3, S550–S555, 2007
www.clim-past-discuss.net/3/S550/2007/

Climate of the Past Discussions

CPD

3, S550–S555, 2007

Interactive Comment

Interactive comment on "Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe" *by* M. Vrac et al.

Anonymous Referee #2

Received and published: 21 August 2007

Summary

Generalised Additive Models (GAMs) are used to downscale mean monthly temperature and precipitation output from the CLIMBER model to the 10' grid-box resolution. The GAMs are calibrated using the CRU observed dataset over Western Europe and are then applied to N America and N Europe for the present day, and for the LGM. The LGM results are compared with palaeoclimate reconstructions and PMIP2 GCM output for 10 Western European locations

General comments

The paper addresses an aspect of downscaling, i.e., how to downscale from rather low-resolution models (models of intermediate complexity, EMICs) and for long-term past (e.g., LGM) or future time periods, which has been much less studied than the 'conventional' problem of downscaling output for the next century from the current generation of GCMs. The need for higher-resolution information for anthropogenic climate change impact studies is widely recognised, but the need for downscaling on longer timescales less so. The latter need is not particularly well presented by the authors - but is nonetheless beginning to emerge, e.g., for assessing high-resolution palaeoclimate reconstructions, and where long-planning horizons are needed - such as radioactive waste disposal safety assessment.

This focus of the study does not, however, come out in the abstract or introduction. The latter includes an outline of the conventional GCM/RCM downscaling approach to downscaling which is not particularly relevant here. Forcing a RCM directly with output from CLIMBER would not be a feasible option - some intermediate/initial downscaling would be needed in order to provide appropriate boundary conditions. Similarly, many of the statistical downscaling methods listed would not be appropriate. Downscaling from EMICs and for simulations on glacial timescales raises a number of particular challenges - scale issues, the potential large changes in predictors and the need to consider sea level and consequent geographical changes. I would like to see these issues and challenges brought out more clearly in the introduction and then elsewhere in the paper.

The need for robustness is mentioned in the introduction and elsewhere, but what the authors mean by robustness is not defined. To my mind, testing for robustness would involve an assessment of sensitivity to choice of predictors and domain sizes, calibration/validation period etc., etc. Also, consideration of potential dangers such as overfitting. On page 902, it is stated that the inclusion of physical as well as geographical variables gains robustness under climate change conditions. There is another rather clumsily constructed sentence on page 918 about physical variables bringing some stability and robustness. Though the previous sentence says it is the geographical variables that provide robustness. I do not really understand any of these sentences/claims. What is needed is clear statements at the beginning about exactly what the authors mean by stability and robustness. And much clearer presentation of the sensitivity analyses aimed at evaluating this.

For example, on page 918, there is some mention of predictors being outside the calibration range. This is an important issue, particularly for LGM simulations. It would be informative to present some examples showing the calibration range and the values from the LGM simulations to demonstrate just how much of an issue this is.

Like the other reviewers, I struggled with the language and level of presentation, and agree that it is very difficult, often impossible to understand what was done or what we are looking at. e.g., Figure 2 - are these averages across all grid points?

There also seem to be stages in the study where a number of rather arbitrary and subjective decisions seem to be made. E.g., BIC is used to develop the monthly GAMs - whereas the choice of predictors etc for the annual models seems much more subjective.

I am also concerned about the lack of discussion of underlying physical processes throughout the paper. Presumably it is concern about such processes that lead the authors to include the continentality and W-slope variables as potential predictors. Unfortunately, the discussion of how and why these variables are constructed is rather poorly written and very difficult to follow. Maybe some schematic representation would help. More importantly, there does not seem to be any discussion of whether the selected predictors make physical sense. Figure 1, for example, suggests that there may indeed be some strong non-linearities in predictor-predictand relationships. Where these occur, such as for Q in the toy model - do they make physical sense?

I also share concerns expressed by other reviewers that the abstract and conclusions do not seem to reflect the main body of the paper and both give a more optimistic overview of GAM performance than I think is justified. Similarly, the figures need substantial improvement to improve their legibility and interpretability.

CPD
3, S550–S555, 2007

In conclusion, the paper needs substantial work in order to allow readers to make a proper evaluation of the method presented.

Specific comments

These are not exhaustive, e.g., parts of the text where rewriting is required due to language difficulties are not listed.

Page 902, line 10. The issues of linearity and stationarity are different. Some models are linear - others (not just GAMs) are non-linear (eg. Neural networks). All statistical models, including GAMs, have to make the stationarity assumption.

Page 903, section 2.1. The description of CLIMBER and the simulation used for this work is inadequate. More details and supporting references are needed - especially on the forcing used, time period of the simulation and validation work. How many years of output are used to calibrate the GAMs?

Page 904, line 24. This paragraph implies that a number of methods were considered/evaluated but rejected - leaving only GAMs. Unless, a range of potential non-linear models were properly considered, this is misleading and rewording is needed.

Page 905, lines 12-18. Should give some references supporting choice of distribution. Are these distributions appropriate for all regions - or just W Europe?

Page 908, bottom line. Give a bit more detail about interpolation. To what extent do the GAM results improve on this simple interpolation?

Page 909, Section 3.2, first paragraph. The language/terminology used here e.g., 'light climate change' seems inappropriate to a scientific paper.

Page 913, line 13-15. I find the statement that the results are 'actually quite acceptable' difficult to understand given that only 17.5% of variance is explained.

Page 913, line 20. What is meant by 'more efficient'? A 5°C error in mean monthly temperature seems fairly large to me.

CPD

3, S550–S555, 2007

Interactive Comment

Page 914. The whole discussion of results is generally rather poorly worded throughout the paper. There are a couple of examples of phrases I find particularly uninformative on this page: Line 7 ‘very unrealistic structures’ Lines 19/20 ‘..show similar results than previously despite some differences’.

Page 914, lines 21/22 - at first I thought a decimal point must be missing - but it seems not. These are huge not just high residuals!

Page 915, second paragraph. The selection of predictors for the annual GAMs here appears rather arbitrary and not really informed by the earlier analyses.

Page 915, line 15. Presumably CLIMBER does have ice sheets!

Page 915. I would like to see more discussion of Figure 7. Is there, for example, any (qualitative) evidence for these differences - e.g., increased rainfall in the NE. Are the spatial and seasonal patterns produced realistic, e.g., smaller cooling in summer than other seasons.

Page 916, first paragraph. Somewhere, need to give more details about these GCM simulations - how does their resolution and forcing used compare to that for the CLIMBER simulation used here?

Page 917, line 16. I fully agree with the other reviewers that it is not acceptable to remove these two stations.

Page 918, lines 4-6. Given what is said here about the applicability of the monthly GAMs and the fact that the performance of monthly/annual GAMs is not directly compared, I wonder if it is really worthwhile including the monthly GAMs at all? I think it would be more interesting to focus on the novel aspect of the paper, which is the palaeoclimate and long timescale perspective.

Page 918, bottom line. I think that the statement that ‘the downscaled values behave well’ gives on over optimistic view of performance. Certainly you don’t see any systematic improvement in the downscaled values.

Figures.

All the figures need work to improve their clarity. Axis labelling all needs improving and captions need checking to ensure that they give sufficient detail such as information about units not given on the figures themselves.

Figure 9 - The x action goes from 0 to 40 - this should be properly labelled from 1-10 for the 10 locations.

Interactive comment on Clim. Past Discuss., 3, 899, 2007.