Geosci. Model Dev. Discuss., 6, C2839–C2848, 2014
www.geosci-model-dev-discuss.net/6/C2839/2014/

Geoscientific Model Development Discussions

Interactive comment on "The regional MiKlip decadal forecast ensemble for Europe" *by* S. Mieruch et al.

S. Mieruch et al.

sebastian.mieruch@kit.edu

Received and published: 21 March 2014

Review answer 2

The regional MiKlip decadal forecast ensemble for Europe

S. Mieruch, H. Feldmann, G. Schädler, C.-J. Lenz, S. Kothe, and C. Kottmeier

referee comments in **red**, author reply in **black**

Main issues:

No details are given regarding the downscaling method, hypothesis, results. After reading several times the experimental design, it remains very obscure to me how the outputs from the MPI-ESM decadal system were used to construct the regional forecast system. Which fields were exactly used for the forcing? How are the lateral boundary conditions dealt with?

The method used for the CCLM ensemble is direct dynamical downscaling of the global forcing over Europe, analogue to the CORdinated Downscaling EXperiment (CORDEX) for the IPCC AR5. Initial and boundary values are taken from the global model. Only for the initialization of the soil the soil fields from a re-analysis driven CCLM simulations are used (p5715, l25ff). The MPI-ESM ensemble is generated as described in Müller et al. (2012) and Matei et al. (2012); further details can be found there. The coupled model MPI-ESM consists of the atmospheric model ECHAM6 and the ocean model MPI-OM. The ocean is initialized using an ocean-only simulation with atmospheric boundary conditions from the NCEP re-analysis and then nudging sea temperature and salinity towards this simulation. The first of January of the folowing year is set as starting date of the simulation (e.g. 2001-01-01 for decadal2000) with a simulation period of 10 years each. The boundary conditions for CCLM are taken from the 3-dimensional fields of MPI-ESM and are updated every 6 hours.
We will rewrite the experiment description to achieve more clarity on the topic.

It is not even clear which dates were chosen (p. 5715 l. 22). I can only guess this from Fig. 2. (or wait p. 5722). This example reflects the lousiness of the description of the experimental protocol. See other questions below in minor points.

At page 5715 we say that "Therefore, all 10 available realizations of the MPI-ESM-LR for 5 starting dates (decadal1960-decadal2000) were downscaled covering the whole

50 yr period." To make it clear, we will add a sentence saying explicitely that the simulations started at Januar 1st 1961, Januar 1st 1971, Januar 1st 1981, Januar 1st 1991, Januar 1st 2001 and ran freely each for 10 years.

Two regional models are announced in the introduction and in the experimental design, but the REMO is never discussed again in the rest of the paper. I am not even sure what was the aim of the authors in using two models. Was the REMO model used at all? Where?

In the framework of the MiKlip project, two RCM's have been used, CCLM and REMO, but the focus of this work is on CCLM. At page 5714, line 12 it is said that "The focus of this paper is on the description and the assessment of regional predictions with COSMO-CLM." And again on page 5715, line 6: "The focus for this paper is on the simulations with CCLM. This RCM is used in the same model version and as for CORDEX (Panitz et al., 2013)." However, it is planned to combine CCLM and REMO realizations in the next project phase. We will add this information in the revised version.

After an interesting section on metrics, the results are then presented as a catalogue of statistics. No interpretation is ever given, either in terms of regional climate dynamics (why does the filtering have no effect over the Iberian Peninsula, Italy and the Balkan? I would guess it has something to do with the different dynamics over northern and southern Europe? Is there really no explanation?) nor with respect to the performance of the full MPI-ESM decadal system over similar regions. From my understanding, predictability of the RCM is expected only if and where there is predictability already in the initial system. Is that true? I would expect at least a discussion of this point. Finally, significance (which are a crucial point for the evaluations of decadal predictions and signals are often very weak) is not convincingly addressed (see below).

The aim of our study is to analyse the feasibility and prospects of regional decadal prediction. This is one of the first tries worldwide. A thorough investigation and interpretation of the observed predictability patterns is appealing and of course a topic of high interest. However, such an analysis would be beyond the scope of this paper. Nevertheless, we already have included some thoughts into this direction:

- page 5713: "... predictability must come from the large scale processes and interactions (like AMO, ENSO, quasi-biennial oscil- lation), who must be captured by the global models, and ..."
- page 5726: "This could be a hint to the causes of the decadal predictability observed here, arising from a north-atlantic large scale low frequency variability."
- 5730: "The reason for such a pattern could be e.g. a large scale atmospheric teleconnection."

With respect to the global MPI-ESM-LR driving simulations, we will include a new section in the revised version and discuss the predictability of the global system and compare it with the results from CCLM. The significance of the correlation is addressed on page 5719.

Because of these weaknesses which prevent a proper interpretation and any learning from the results presented, I have to recommend a rejection of the paper. Nevertheless, as I underlined above, the topic is of course very promising, and the experiments most certainly valuable. Therefore, I strongly encourage the authors to revise the work by addressing, at least, the points above and resubmit.

As explained above we will address all concerns of the referee, except the analysis of a thorough interpretation, because this is much beyond the scope. However, we think that we have learned a lot from the analysis, which was not known before:

- There exists a potential of regional decadal prediction of detrended summer and winter temperatures, especially in northern Europe.
- Temporal smoothing is beneficial for the skill assessment and it seems that predicting annual means is to ambitious, whereas low pass filtered data (9 year moving average) can be well predicted.
- The climatological distributions of CCLM and E-OBS observations are not significantly different.
- The reliability of the 10 member CCLM ensemble is reasonable. This is in contrast to numerical weather forecasting or seasonal forecasting, where 10 member ensembles are mostly overconfident, i.e. the spread is too small.
- ...

Minor comments:

P5713 l 13 (or line 20) initialization: consider citing the more recent Garcia Serrano et al 2012 GRL for example rather? Or Doblas Reyes 2013 Nature Communications

We will cite these papers in the revised version.

L14 who $->$ which

Changed in the revised version.

L17 so therefore : this is not English

Corrected in the revised version.

P 5714 – l.14: "1 day time lagged initialization for the atmosphere?" I don't understand this statement. I have the impression there is a confusion between what is usually called "initialization" and "perturbation" in forecasting systems. Please clarify.

Yes, the reviwer is correct. The MPI-ESM uses perturbed atmospheric initial states as ensemble generation method. 10 ensemble members are created with a 1-day time lag from one to the next. For instance, the atmospheric conditions from January 2nd are set to the first of January for realisation 2 and so on.
We will clarify this in the manuscript.

P5715 l 25: How is the transition between ERA 40 and ERA interim treated?

An ERA40-driven simulation with CCLM was started in 1959 (using the first 2 years as spin-up) until 1979. Starting 1st 1979 ERAInterim was used as boundary condition. The state of the soil from the end of the ERA40 driven CCLM simulation was taken as initial soil condition, analogue the the soil initialisation of the decadal hindcast simulations.

P5716 l 4: I don't understand why "this can be regarded as anomaly initialization". And again, I don't see how this relates to the MPI-ESM decadal system.

All initial conditions for the RCM except the soil are obtained from the global forcing. The soil conditions are taken from the above mentioned re-analysis driven simulation, to achieve a dynamic equilibrium between the soil model TERRA-ML of CCLM and the analysed atmospheric fields. With this method the anaomalies of the 3-dimensional soil conditions w.r.t the model climatology are applied as initial conditions but not for example the full fields from any observational data set. We will clarify this in the new manuscript.

p. 5716 section 3: What procedure is used for the detrend?? Linear? Is it applied grid point wise?? What are the justifications?

Yes, simple linear regression, grid point wise. We will include this information in the revised version. The justifications are given on page 5716, line 20-22.

P 5717 l 3: "these high frequency fluctuations cannot be predicted using decadal model initializations": this is part of what we would like to be shown here. Or else, give a reference (do you refer to works from Boer et al for example?)

This statement is not based on hard statistics, it is based on visually interpreting the thin curve in Fig. 1 including what is known up to date about decadal predictions. However, to weaken our statement we will change "cannot be predicted" to "are unlikely to be predictable".

P5717 l 21: add "spatial"?

Will be included in the revised version.

P5718 l 20 bias adjustment: this has not be discussed before: how is the bias removed? (and which bias?)

This is explained on page 5716, line 20: "... we decided to remove the long term means and trends from the time series ...". By removing the long term means from the model data and the observations, also the long term bias between model and observation is removed automatically. However, to avoid misunderstanding we will change "bias" on page 5718, line 20 to "mean".

P5718 l 21 What do you mean "the interpretation of the value of the correlation coefficient is very individual"?? You mean it depends on the scientist himself? But there

are some statistical tools (significance evaluation) to diminish as much as possible this subjectivity, isn't it?

No, as explained on page 5718, line 20-27, the question of what is good and what is bad depends on the experiment, data, null-hypotheses and more. Similar the significance. What is good a p-value of 0.05, 0.01 or 0.001?

P5718 l 25: "common interpretation" of the significance level? Common for who, for what? Please check a statistics textbook.

On page 5718, line 25 it is not about the significance level, it is about the correlation coefficient and its common interpretation regarding statistics textbooks.

P 5720 l 14: How accurate is it to average p values? Please give a reference. Fig 4 legend. "here": this is not proper language for a figure legend.

Given that the data are in accordance with the null-hypothesis, p-values are equally distributed from 0 to 1, thus averaging seems to be not problematic. We will correct the language of the figure caption in the revised version.

p. 5723 l 21: it would also probably change the atmospheric circulation etc. I find this statement quite speculative.

That "... the grainy structure ... can be attributed to orographic differences between E-OBS and CCLM." and "... a orography correction ..." would decrease the differences is a well known effect. Basically reflects the higher variability at higher resolution.

On the use of moving average of 9 yrs: given that the hindcasts are 9 years long, I don't really understand how you do that.

The hindcasts are 10 years long. The application of the moving average filters is discussed in detail on page 5717, lines 16-21 and Tab. 1.

P5724 l 15-20 “but simultaneously smears out the resolution”: isn’t that expected? Rephrase this passage. + check repetition with data/methods.

Yes, this is expected. However, our intention was to explicitely mention the impacts of the smoothing. We will stick to this passage.

4.1.3: fidelity. Is this result surprising? Are their models in which it is NOT “fidel”. How it is in the decadal system for Europe averages for example? Please discuss.

The result is not necessarily surprising. But it is a result, which was not known before the analysis was made. Thus, the outcome could have been also different. Regarding the fidelity, as we use it, of other models, we refer to DelSole and Shukla (2010). The fidelity of the filtered data is already discussed in Sect. 4.1.3.

Section 5 P 5729 l 12: “predictability increases with low pass filtering”. I have the impression this sentence is not fully consistent with the end of section 4.1.4, am I wrong? Here, your interpretation of predictability seems to rely only on the correlation. However, as you explained it in section 3, “predictability” is rather the result of all (or some) of metrics developed above, no?

Right, actually we separate the predictability into predictive skill and reliability. Thus, we will change “predictability” here to “predictive skill”.

Section 6: I disagree that this study has made “a large step in understanding” of regional predictability. In terms of feasibility, yes, perhaps, if the experiments were more clearly and precisely described.

Ok, we will slightly weaken our statement to “a step forward in understanding”. However, it has to be considered that it is the first time world wide of dynamically downscaling decadal predictions and the results presented in this study are absolutely novel and unique. The list of new research insights is long as explained above in the “main issues” section. Additionally, the results reported here are of high impact for the community and could yield as a benchmark for future regional decadal predictions and will hopefully stimulate future research in that direction.

P 571 l 21: should come much before.

It seems that the last character in P 571 is missing. Thus we cannot figure out the page under consideration.

References

DelSole, T. and Shukla, J.: Model fidelity versus skill in seasonal forecasting, J. Climate, 23, 4794–4806, 2010.

Matei, D., Pohlmann, H., Jungclaus, J., Müller, W., Haak, H., and Marotzke, J.: Two tales of initializing decadal climate prediction experiments with the echam5/mpi-om model, J. Climate, 25, 8502–8523, 2012.

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. S., and Marotzke, J.: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, Geophysical Research Letters, 39, L22 707, 2012.

Interactive comment on Geosci. Model Dev. Discuss., 6, 5711, 2013.