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

S. Mieruch et al.

sebastian.mieruch@kit.edu

Review answer 1

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**

C2827

Main issues:

1) ... how the downscaled regional forecasts compare to the set of global MPI-ESM-LR based forecasts ...

We agree with the referee that the comparison of the global driving model with the regional model is very important. However, the main goal of the paper is to show the prospects and the feasability of regional decadal predictions and an accurate comparison, global vs. regional, is another topic. Nevertheless, we will include a section on the skill and reliability of the global MPI-ESM-LR model to elucidate agreements and differences. However, this can only be a snapshot in the context of this paper. The comarison with GCM and added value will be the main topic for an upcoming paper.
Further, to save space, we will remove the analysis of the "fidelity", since the fidelity is good for all regions, seasons and averaging times. Thus, we can not learn a lot from the fidelity analysis and will concentrate on skill and reliability.

a) To what extent is the predominantly cold bias in Figs. 3 and S1 (upper right) attributable to biases in MPI- ESM-LR vs biases in CCLM? (Compare with MPI-ESM-LR and ERA40-forced CCLM biases.

We will discuss this shortly in Sect. 4.1.1 Bias and trend and Sect. 4.2 Winter temperature.

b) To what extent are the weak CCLM trends (Figs. 3 and S1 second row) attributable to inaccurate trends in MPI-ESM-LR vs inaccuracies originating in CCLM? (Compare with MPI-ESM-LR and ERA40-forced CCLM trends.)

Similar to the bias, we will discuss this in Sect. 4.1.1 Bias and trend and Sect. 4.2 Winter temperature.

C2828

It is true, CCLM shows slightly too weak trends to MPI-ESM-LR. Figure 1 (in this review reply) shows the MPI-ESM-LR bias and trends. We will discuss that.

We will include the figures and the discussion on the skill and reliability of the MPI-ESM in the new manuscript (ms).

We will include a discussion on the "added value" in the new ms.

C2829

Anomaly initialization minimises drifts as shown e.g. by Smith et al. (2013). However, drifts can occur even using the anomaly initialisation. To avoid such drifts we describe from page 5715 line 25 to page 5716 line 5 how we used soil variables from an ERA driven run started in 1959.

The discussion about Fig. 1 is indeed intended as an illustration and it is not our aim to perform a study on potential predictability. However, regarding the ideas about "poten-tial predictability", it is essentially the question of signal to noise. To apply an analysis on potential predictability (Zwiers and Kharin, 1998) one has always to define the sig-

C2830

nal and the noise in the data, which is somehow arbitrary and should, in the best case, be checked using real observations. This is also done in Boer (2004) (cited by the referee), who define pentadal, decadal and 25 year means as the "potentially predictable signal". Thus, choosing the 9 year moving average filtered data as our signal is absolutely ok. Nevertheless, to avoid any misunderstanding we remove the use of the phrase "potentially predictable signal", because this is not the topic of the paper.

Regarding the white noise example, we have to say that it is not possible to generate a figure such as Fig. 1 using white noise as stated by the referee. The signal to noise ratio in Fig. 1 is 0.3, for white noise surrogates it is expected to observe a signal to noise ratio of about 0.1 ± 0.04. Thus, the "signal" in Fig. 1 is far away ($\approx 5\sigma$) from being random.

Minor comments:

p. 5712 line 6: This sentence gives one the impression that the primary focus of MiKlip is specifically prediction on time scales of "decades" (evidently meaning 20-30 years minimum), whereas there is little evidence at present that useful predictions of natural climate variability on such time scales can be made. On the other hand, fona-miklip.de uses terminology such as "time frames of years to decades" and decadal (rather than multi-decadal) predictions. Therefore I would recommend changing "time scales of decades" to "time scales of years to decades" or "time scales of years to a decade or longer".

Right, we are interested in predicting time scales from years to a decade. But not longer. We will correct that in the revised version.

p. 5712 line 25: Should inform the reader here how projections differ from predictions, e.g. "as opposed to projections, which do not take into account the influence of initial conditions".

C2831

Will be included in the revised version.

p. 5713 lines 9-10: The authors seem to be asserting here that decadal predictions based on statistical methods cannot be useful. This sounds overstated, but if that is indeed what the authors mean then they need to provide one or more supporting references.

This text refers to global climate predictions and we are not aware of any publication describing global climate predictions (and projections) with statistical methods (there is also no mention of that in the IPCC Report). However, regional downscaling may well be possible with statistical methods. To make the context clearer, we change the text as follows: "... (i) global climate predictions (and projections) need to be produced by coupled atmosphere-ocean models, ..."

p. 5713 line 11: Off hand I am not aware of any evidence for multi-year predictability specific to soil. On the other hand, there is some such evidence for sea ice (Blanchard-Wrigglesworth et al., GRL 2011). Therefore, suggest rewording e.g. to "i.e. the oceans and possibly sea ice and soil" perhaps including also "slow phenomena in the atmosphere such as the quasi-biennial oscillation" to be in accordance with the subsequent discussion.

There are a few new studies coming out soon on the predictability of soil: Chikamoto et al. (2014); Khodayar et al. (2014). We will include sea ice and QBO in the revised version.

p. 5715 line 24: Should add that the 5 starting dates are separated by 10 year intervals

We will add that information in the revised version.

C2832

Re-analysis forcing is applied the way as GCM forcing, namely at the boundaries. The update frequency of ERA driven simulations is 6-hourly, as for MPI-ESM simulations. No interior constraints are applied, but for sea-surface temperature, which is taken from the driving model. The initial values are also taken from the driving model. The exception are the soil parameters in the hindcast simulations which are taken from the re-analysis driven simulation by CCLM (cf. 5715 25ff)

We will apply that change in the revised version.

We will apply that change in the revised version.

We will include that information in the revised version.

C2833

Actually, it is not the prediction of the means, rather the tendencies. The reason is, as explained e.g. on page 5719, that the correlation coefficient is invariant to location and scale transformations. Thus, in this sense it is not needed to exactly forecast the absolute value of the observation, but the tendency, e.g. being above or below zero.

The interpretation of the observed statistics has a large component of human judgement. That means, that e.g. the correlation coefficient has to be interpreted with a lot of knowledge on the experimental design, the expected results given some null-hypotheses and so on. Further, the human visual system performs often excellent in comparison tasks. Regarding Fig. 2 it is easy to see that the Lodz data agree well and the Rome data agree badly. Thus, we stick to the word "unsatisfactory".

Will be corrected in the revised version.

The potential predictability exhibits indeed two aspects, first it is a mathematical concept, which can be applied e.g. on model data solely. Second, as stated on page 5719, the correlation coefficient, using model data and observations, can be

C2834

used to estimate the potential predictability (Murphy, 1988; Kharin and Zwiers, 2003). However, to avoid misunderstanding we will replace the "potential predictability" here, but stick to it in the discussion on page 5719.

p. 5731: Does "a.o." mean something in English?

Probably not. We will replace that with "amongst others".

Fig. 2: the correlation values are so small that they are difficult to read

This is due to the format of the discussion paper. The figure is originally of high quality and will be quite larger, and thus readable, in the final version.

Fig. 2: in addition to correlation values include also N_eff or p values?

We will discuss the typical size of N_{eff} in Sect. 3. The p-values are in principle given by the stippling, i.e. stippling means a p-value < 0.5 and no stippling means p-value > 0.5.

Fig. 4: the stippling is very difficult to see; one suggestion would be to include a high resolution version of the warm season temperature results, comparable to current Figs. S1-S3, as a new Fig. S1 in the supplementary material (renumbering the current figures to S2-S4).

As suggested by the referee, we will include an additional supplementary figure in the revised version.

C2835

References

Boer, G.: Long time-scale potential predictability in an ensemble of coupled climate models, Climate Dynamics, 23, 29–44, doi:10.1007/s00382-004-0419-8, http://dx.doi.org/10.1007/s00382-004-0419-8, 2004.

Chikamoto, Y., Timmermann, A., Stevenson, S., and DiNezio, P.: Decadal predictability of soil water, vegetation, and wildfire frequency over North America (submitted), J. Climate, 2014.

Kharin, V. V. and Zwiers, F. W.: Improved seasonal probability forecasts, J. Climate, 16, 1684–1701, 2003.

Khodayar, S., Sehlinger, A., and Feldmann, H.: Sensitivity of soil moisture initialization for decadal predictions under different regional climatic condititons in Europe, Int. J. Climatol., submitted, 2014.

Murphy, A. H.: Skill scores based on the mean square error and their relationships to the correlation coefficient, Mon. Wea. Rev., 116, 2417–2424, 1988.

Smith, D., Eade, R., and Pohlmann, H.: A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction, Climate Dynamics, 41, 3325–3338, doi:10.1007/s00382-013-1683-2, http://dx.doi.org/10.1007/s00382-013-1683-2, 2013.

Zwiers, F. W. and Kharin, V. V.: Intercomparison of interannual variability and potential predictability: an AMIP diagnostic subproject, Climate Dynamics, 14, 517–528, doi:10.1007/s003820050238, http://dx.doi.org/10.1007/s003820050238, 1998.

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

MP I _EnsemblePa

Fig. 1. MPI-ESM-LR long-term bias and trends.

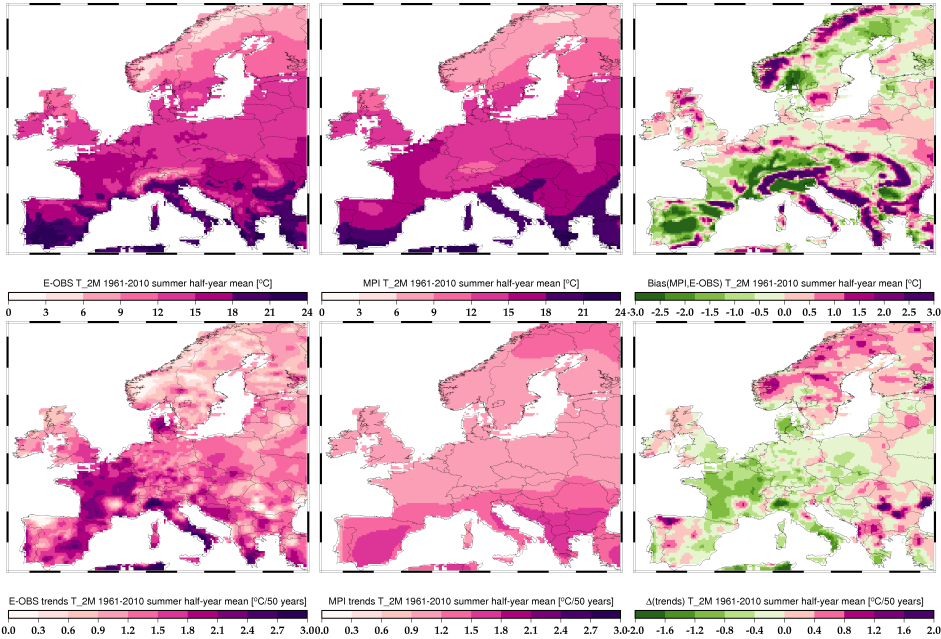E-OBS T_2M 1961-2010 summer half-year mean [°C]

MPI T_2M 1961-2010 summer half-year mean [°C]

Bias(MPI,E-OBS) T_2M 1961-2010 summer half-year mean [°C]

0	3	6	9	12	15	18	21	24

0	3	6	9	12	15	18	21	24

-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0

E-OBS trends T_2M 1961-2010 summer half-year mean [°C/50 years]

MPI trends T_2M 1961-2010 summer half-year mean [°C/50 years]

Δ(trends) T_2M 1961-2010 summer half-year mean [°C/50 years]

0.0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0

0.0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0

-2.0	-1.6	-1.2	-0.8	-0.4	0.0	0.4	0.8	1.2	1.6	2.0

Fig. 2.