Geosci. Model Dev. Discuss., 7, C2958–C2963, 2015
www.geosci-model-dev-discuss.net/7/C2958/2015/

Geoscientific Model Development Discussions
Open Access

GMDD
7, C2958–C2963, 2015

Interactive Comment

Interactive comment on "An observation-constrained multi-physics RCM ensemble for simulating European mega-heatwaves" *by* A. I. Stegehuis et al.

J Fernández (Referee)

fernandej@unican.es

Received and published: 16 January 2015

* General comments:

This work analyses an unprecedented (to my knowledge) multi-physics ensemble consisting of 216 summer seasonal simulations, focusing on heat waves over France (2003) and Russia (2010). It provides a fully systematic approach towards the selection of an optimal subensemble to represent heat waves. The paper presents fairly novel concepts and ideas and a new dataset which will probably feed subsequent work. Therefore I suggest the paper to be published after a minor revision taking care of the specific comments below, which mainly refer to further discuss some points and solve

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

some doubts to improve the reproducibility of the results.

* Specific comments:

1) The authors intend to create an optimized WRF ensemble for heat waves (7873:9-10). What would be the use of such ensemble? The experimental setup used should be taken into account. The simulations shown were run for a few months nudged towards the observed flow. For climate change simulations, nested into a GCM, nudging the atmospheric flow could be a problem. Also, long-term simulations could build up biases not arising in a few months (e.g. related to soil moisture). For seasonal forecasting, the authors recognize problems (7863:15-19) to reproduce observed events due to the caotic nature of the atmospheric circulation.

2) The 5-member sub-ensemble was only tested for heatwaves. The 2007 "normal" season is not shown in Figure 3 or any of those in the supplemenary material (even though it is stated that they are in the Suppl. material in 7871:2). If these members are the best in any physically meaningful sense, they should also perform well in the "France 2007" case study. Is that the case?

3) There are already examples in the literature of "sub-ensembles" breaking model democracy, in which the subensemble outperforms the full ensemble. For instance, Herrera et al (2010) selected a sub-ensemble using mean precipitation and show that this sub-ensemble is also well fitted for extreme precipitation regimes. This result is close to the results found in this work (subensembles selected for a heatwave work well for other regions or regular seasons –if this is the case–), and could be added to the discussion, given that it extended the idea to multiple models.

4) The potential implications of the study for climate modelling (7873:17-) need to be discussed in a wider framework. The authors constrained the ensemble to a particular season, variables and error metrics. In this way, they were able to find an "optimized" set of configurations. However, It has long been recognized (Fernandez et al, 2007), that in a climate simulation an optimal configuration cannot be chosen. Biases and

GMDD

7, C2958–C2963, 2015

the best-performing configurations heavily depend on the season (Garcia-Diez et al, 2013), variable and even on the metric used (Jerez et al, 2013).

5) Moreover, observational uncertainty was not considered. It has been shown that the reference observations affect model rankings (Gomez-Navarro et al, 2012). This needs to be discussed at some point in the paper.

6) Jerez et al (2013) did not use WRF (7865:15). Other potential references here are Awan et al (2011) and Mooney et al (2012). Also, multi-physics ensembles did not start with WRF. There are a few other works with its predecessor, MM5.

7) Weisheimer et al (2011) did not use WRF (7870:15). Remove or rephrase.

8) The authors find probable (7872:18) that the inclusion of another land surface model would increase the ensemble spread. This statement can be accompanied by a cite to Mooney et al (2012), where they show strong differences when changing the LSM (see their Fig. 2b, e.g. Sim 9 vs. 11). It is not clear at all why the LSM sensitivity was left out of the study. There is plenty of literature (even cited by the authors) highlighting the role of soil-atmosphere interactions in the development of heat waves and the authors themselves recognize it (7869:16).

9) I don't agree with the sentence (7870:12) "By contrast, heatwave temperatures do not seem very sensitive to the planetary boundary layer and surface layer physics schemes". Figure 2d seems noisier than the rest because there are more PBL options tested. However, there is a clear, systematic temperature dependence on the PBL. If you imagine a regression line for each PBL scheme, all of them preserve the relationship (slope) with soil moisture, but the the heat wave average temperature is clearly different.

10) The caption of Figure 2 says correlation where scatterplot is meant. These particular plots show that there is indeed (negative) correlation, but the plots are scatterplots.

11) X-axis labels in Figs 1 and 3 read "Time (DOY)". I assume it means Day Of the

GMDD

7, C2958–C2963, 2015

Year but, please, define. Also, in the panels with this axis, two vertical lines showing the heatwave period considered would help, given that diferent periods were chosen for each event. Also, if Fig 1bc shared the Y-axis with Fig 1a, they could be directly compared with each other. Currently, the normal year seems as hot as the 2003 heat wave, when in fact it is 5K colder.

12) In Fig 3c-August, the cyan circle is missing (probably hidden behind other member). Using non-overlapping symbols would help. The same happens with the pink circle in Fig. 3b-August.

13) The resolution is stated to be 50km (approx 0.44deg). Was a Lambert grid projection in Kms used? or the Euro-Cordex standard 0.44 rotated lat-lon grid? Please, clarify.

14) Observational data is not fully described. Which E-OBS version was used? Was any interpolation carried out in the analyses?

15) The pre-screening of the simulations considering only those within 1K of the E-OBS temperature might be problematic. RCMs have biases. With the method proposed, a fairly physically-consistent simulation could be disregarded, while a simulation unrealistically compensating temperature biases might get in. The latter can easily happen (Garcia-Diez et al. 2014).

16) The ranking metrics are not fully clear to me. Daily temperature differences are used (7867:26). But, which score was built out of them? RMSE? Why was temperature considered at a daily scale and precipitation at monthly scale?

17) For radiation data, was the model interpolated to the stations to compute the spatial averages? which interpolation method was used? How many radiation stations were available in each region?

18) Regarding the rejection of the members differing in only one scheme (7868:08): How many of these members were disregarded to get the top-5? What is the interest

GMDD

7, C2958–C2963, 2015

of "keep[ing] a large range of different realistic physics combinations between the simulations" (7868:11)? I see also an interest in the single-step ensemble members. In these ones, the differences can be traced to the single scheme that changed. For instance, "The two simulations that are largely overestimating latent heat flux" (7871:21) are those not using Tiedke, but this could be just by chance, given that they also differ in other schemes and the schemes interact in a non-linear way (Awan et al. 2011).

19) How were the extreme configurations selected (7868:22)?

20) By "the middle of the simulations" (7869:22), I guess you mean the "median".

21) At some point (7870:18), the effect of convective clouds on radiation is invoked. However, note that in WRF the interaction of radiation with sub-grid clouds has only recently been implemented (Alapaty et al. 2012) and included in WRF3.6 for certain combination of radiation and cumulus schemes. It was not included in the version used in this work (WRF 3.3.1).

22) The discussion in 7871:18-25 seems to imply (althouth it is not explicitly stated) that the good performance of the Tiedke scheme just during the heatwave is just by chance.

23) "We found a large spread" (7872:11) I would highlight, just at the beginning of this sentence "Even though the simulations were constrained by grid nudging,"

24) The journal recommendations suggest that "The model name and number should be included in [the title of] papers that deal with only one model". Replace RCM by WRF in the title.

* Technical corrections:

(beware I'm not a native speaker)

7862:17, "together with varied physics scheme." sounds odd to me. Please, rephrase. 7866:04, "temperatures differ by less among one another than 0.5C" sounds odd.

GMDD

7, C2958–C2963, 2015

7867:01, "Tawary" should read "Tewary". 7869:26, "better" -> "well" (or "better than [what?]") 7872:27, missing period "heatwaves Changes"

* References:

Alapaty et al. (2012) "Introducing subgrid-scale cloud feedbacks to radiation for regional meteorological and climate modeling" Geophys. Res. Lett. 39:L24808

Fernández et al (2007) "Sensitivity of the MM5 mesoscale model to physical parameterizations for regional climate studies: annual cycle" J. Geophys. Res. 112:D04101

Gomez-Navarro et al (2012) "What is the role of the observational dataset in the evaluation and scoring of climate models?" Geophys. Res. Lett. 39:L24701

Herrera et al. (2010) "Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel" J. Geophys. Res. 115:D21117

Mooney et al (2012) "Evaluation of the Sensitivity of the Weather Research and Forecasting Model to Parameterization Schemes for Regional Climates of Europe over the Period 1990–95", J. Clim. 26:1002-1017

Interactive comment on Geosci. Model Dev. Discuss., 7, 7861, 2014.