

MS No.: gmd-2014-175 revision
Review, 19 June 2015

Title: An observation-constrained multi-physics WRF ensemble for simulating European mega-heatwaves

Authors: A.I. Stegehuis, R. Vautard, P. Ciais, A.J. Teuling, D.G. Miralles, and M. Wild

Recommendation: [Minor Revision]

GENERAL COMMENTS:

The revised manuscript is substantially improved compared to the original version both in clarity as in readability. Thanks to the amendment in the title (WRF instead of RCM) readers know beforehand which family of model formulations is the focus of the paper.

I still find it regrettable that variations in land-surface parameterization are not included in the investigation. Yet, though it came as a surprise when reading the author's reply, I can accept their explanation that apparently none of the alternative land-surface schemes within WRF is capable of providing realistic results. Which makes me wonder, however, why the authors show so much confidence in announcing a future study dedicated to the use of multiple land-surface schemes.

There remain a couple of, mostly textual, issues to be resolved.

MAJOR POINTS:

1.) Abstract, 1st sentence. Actually, interchanging “often not” by “not often” does not take away my concern. In this sentence the authors imply a cause-consequence relation which isn't there. Models are performing poorly in heat wave conditions, not because they have not been evaluated or calibrated with observations, but because the physical processes and their interactions controlling this type of atmospheric conditions are not adequately represented in the models. Evaluation in itself is an essential step, but will not prevent models from performing poorly, while calibration of climate models to certain conditions in the past might probably not be a preferable way to go in climate research.

I recommend to begin the abstract with the perception – which can only come from evaluation – that many climate type models have difficulties in properly reproducing heat wave conditions.

2.) In my opinion the abstract should explicitly mention that only one land-surface scheme has been considered in this study. The ideal place for doing that is the first part of the second sentence which I recommend to rephrase as “*Here we use ... (WRF) regional climate model for a large number of configurations of different atmospheric physics schemes in combination with the NOAA land-surface scheme, with the goal ...*”

- 3.) The last sentence of the abstract is formulated in too general terms. Consider removing this line altogether because it does not highlight your findings. Usually an abstract focusses on what has been done, not on what has not been done. But if you still feel you have good grounds to retain this line, it should be sharpened and make explicit connection with the findings of this paper.
- First of all, the “*varied* physics schemes” is somewhat vague English, please use “*configurations of different atmospheric* physics schemes”. The question is, are these the five configuration found in the paper, or has the research to be carried out all over again.
 - Secondly, instead of “*varied* land surface models” use “*alternative* land surface models”.
 - Finally, the “sensitivity” of what could be included? And what are “land surface processes controlling soil moisture”. Please, clarify.
- 4.) Line 55-57. I recommend to rewrite the 2nd part. “...underlying extreme temperatures during heat waves, because *it is difficult to separate model biases due to deficiencies in the model representation from sensitivities to initial conditions.*”
- 5.) Line 68-69: It is not clear to me what you want to say with this line. Different RCMs respond differently to the same large-scale forcings (re-analysis) because they utilize different “physics packages”. Is that what you mean with “internal variability”? Or are you referring to different “internal degrees of freedom” which allow RCMs with the same physics package to still choose different solutions. Also, it is not clear to me whether “*its*” is meant to refer to “*atmospheric flow*” or to “*events*”. In the latter case it should be “*their*”.
- 6.) Line 91-96: I still miss the point of the term “democracy-driven”, to me it sounds like a misinterpretation of the approach followed in assembling the 2nd type of ensemble. Please, leave it out.

Instead rephrase line 92-94 “... in coordinated experiments (see ...and references therein) *or* combinations of parameterizations...” as “... in coordinated experiments (see ...and references therein), *or by arbitrarily configured* combinations of parameterizations ...”. (Mind the interpunction following the “”).)

Then continue with “In the latter ensemble, ...”

- 7.) The sensitivity tests you performed to quantify the response to changes in the initialization of soil moisture nicely illustrate that the feed back from soil moisture on temperature is potentially large, especially during episodes of blocked circulation with flow coming from the continent.

Regarding the experimental set up I wonder if 20% is with respect to the absolute amount (which it seems to be), or with respect to the range between field capacity and wilting point. The former would indeed be rather radical, because the perturbation of

20% would potentially bring soil moisture outside the physically realistic range. An alternative would have been to initialize soil moisture either at field capacity (“wet” run) or wilting point (“dry” run). Please, mention more specifically how soil moisture is modified in the sensitivity runs.

Line 344-345: rephrase “*a sensitivity test where initial soil moisture was artificially increased and decreased ... was conducted ...*” as “*sensitivity tests in which initial soil moisture was artificially increased or decreased ... have been conducted ...*”

Figure 8 is rather fuzzy and requires upgrading.

Regarding the figure caption of Figure 8 I suggest to rephrase the last two sentences as “*Difference between the perturbed simulations (red indicates 20% reduction of initial soil moisture, blue 20% enhancement) performed with the five highest ranked configurations compared to their corresponding ‘control’ simulations. The darkest lines refer to the simulation conducted with to the best ranked configuration (1), while descending colour shade agrees with descending ranking (1-5).*”

- 8.) Concerning the new Figure 1, the two different shades of blue (light blue = cyan, and blue) are impossible to distinguish. It seems to me only one shade of ‘blue’ was used. Please, use ‘cyan’ for NOAH and ‘blue’ for RUC. (Or if that doesn’t provide enough contrast replace ‘cyan’ by e.g. ‘magenta’ or ‘red’)

MINOR POINTS:

- 1.) Entire manuscript: Consider using “*configuration*” instead of “*combination*” where appropriate. The former is probably more adequate wording in this context than the latter.
- 2.) Line 40: “... problematic ...” → “... impacting ...”
- 3.) Line 53: “... new soil ...” → “... new land surface hydrology ...”
- 4.) Line 55: “... easily ...” -→” ... straightforwardly ...” (Nothing is easy, so *not easily* is a trivial phrase.)
- 5.) Line 97-98: “... because of interacting physical processes and their biases.” Physical processes themselves have no biases, it is the way we describe them. Please rephrase like, for instance, “... because of shortcomings in the representation of physical processes and their interactions”.
- 6.) Line 133: For the sake of completeness, expand “Boundary conditions come from ERA-Interim *including sea surface temperatures ...*”
- 7.) Line 157: “The NOAH scheme seemed more stable in the tests ...”. I suggest to replace “*seemed more stable*” by “*appeared more realistic and robust*” providing more proper wording.
- 8.) Line 172: “..., and *were* not considered.” → “..., and *have therefore not been* considered.”

- 9.) Line 187: Rephrase “ ... is the primary *impacted* variable *and* observations are reliable” as “ ... is the primary *impacting* variable, *while corresponding* observations are reliable”.
- 10.) Line 176: Remove “validate and” , and only retain “we do not use them to rank ...”
- 11.) Line 191: Rephrase “The 1 K threshold *is arbitrary but is* used to avoid ...” as “The 1 K threshold *was arbitrarily chosen and is* used to avoid ...”.
- 12.) Line 203-204: Rephrase “ ... which *is* shown to *be able to* impact ...” as “ ... which *was* shown to *potentially* impact ...”.
- 13.) Line 224: “... during heat waves years, ...” → “... during heat wave years, ...”
- 14.) Line 292: “...to considerable overestimate ...” → “...to considerably overestimate ...”.
- 15.) Line 313: Suggest to rephrase “..., we found a large spread between the different physics for the simulations for ...” as “..., the multi-physics ensemble contained a large spread in ...”
- 16.) Line 314: “..., three variables ...” → “..., *the* three variables ...”
- 17.) Line 320: “ ... probable ...” → “ ... possible ...” (probable is too speculative)
- 18.) Line 320: “ ... largely ...” → “... considerably ...” (largely is too strong, substantially is also possible).
- 19.)
 - a. Panels in figures 3, 5 and 7 have different size. Please, align.
 - b. Also for Fig 3a the text along the horizontal axis is not visible.
 - c. Figure 3d) is mentioned in the caption but not indicated in the Figure
 - d. Labels in Figure 4 are still (d), (e), (f), but should be (a), (b), (c).
 - e. Figure 4c does not have labels along the vertical axis.