Interactive comment on “Towards European-Scale Convection-Resolving Climate Simulations” by David Leutwyler et al.

Anonymous Referee #1

Received and published: 27 June 2016

General Comments:

This is a clearly written manuscript describing world leading efforts to apply high-performance computing to climate studies. With some sharpening of focus, it could make a worthwhile contribution to the literature. Specifically I recommend expanding and deepening section 5, and shortening and making more specific sections 3 and 4.

I suggested minor revisions as my comments and critique should be sufficiently clear for the editor to adjudicate on his/her own, not as an indication of their importance. But also for this reason I would not be inclined to re-review a revised manuscript. This is something the editor can do.

Specific Comments:

1. The conclusion (line 585) makes a concise point of the additional mesoscale structure that becomes representable as the grid is refined to ca 2 km. The manuscript itself is too descriptive and qualitative, and the added value of the 2 km simulations should be more concisely and clearly presented.

2. The most interesting aspect of the manuscript was the proof of concept and computational perspective of the outlook (section 5). Emphasizing the scaling performance, and where and how this might be expected to change either based on different implementations or hardware changes could, when combined with point 1, more sharply guide the community.

Details:

line 8: COSMO (spell out)

line 21: What makes COSMO a climate model? I don’t think this is such and appropriate use of the phrase.

line 29: Fine, but I also think this understates what might be an important role for the mesoscale.

line 32: I never think of Dai and Trenberth as a reference for the uncertainties, or approximations required in cumulus parameterization.

line 40: A case could be made for mentioning super-parameterization here

line 50: ‘partly resolved’ is ‘partly’ true. What about the stable boundary layer, or clouds that are only a few hundreds of meters, or do you mean this to apply to deep convective systems. More clarity would be helpful.

line 60: “narrow”? Perhaps “fine” would be a better choice.

line 146: How adequate is the Tiedtke references? Is this what COSMO really uses . . . I find it hard to believe that the implementation is unchanged from what Tiedtke describes for implementation at ECMWF.

C2
Is it necessary to introduce the DESL acronym?

Same goes for GCL

I am not sure I would lead by saying how you don’t do things.

Traditionally abbreviations, such as “srf” in mathematical text are typed in the roman font. Variables take on the italic font. This helps distinguish between “srf” meaning surface, as opposed to ‘srf’ meaning the entropy times the radius times the free energy... or whatever s*r*f might be.

I appreciate your humanity; but describing this as ‘sad’ might assign unintended value to the situation. For instance, some might argue that ‘sad’ is too passive a voice. I would simply state that it is estimated but there were X storm related casualties and damage estimates were Y (ref).

Check the style guide but I am not used to seeing square brackets to delinate ellipsis.

Is “figure 3” a proper name, i.e., “Figure 3”

Is there no way to look directly at station data so as to infer the core pressures? I guess this depends on where the core is.

Figure 7: Why not continue to show CTRL 50

Is the ‘polar low’ designation correct? Would be good to be specific and connect better to the literature?

But isn’t one advantage of the 2km model that it captures orographic influences, and in this case comparing the valley totals to the mean would help quantify this statement?

Isn’t the ability of convective resolving/permitting simulations to better simulate the diurnal cycle by now an old result?

Maybe an earlier Tompkins and Craig reference is warranted in the discussion of cold pools?

One should distinguish between weak scaling at fixed resolution and weak scaling more generally. The distinction is that the temporal dimension does not exhibit weak scaling.

This discussion would benefit fro some reflection on what the trade offs are of the different ways of comparing things.

Why compare 64x64 with 2620, as it forces the reader to do the multiplication. I prefer to call 64x64 4096.

Some additional performance metrics, like time per price, or time per power would be informative, even if only qualitative.

It may be mention that many CORDEX simulations are performed with very old models (REMO) whose parallel efficiency is quite small.

I would prefer to phrase this as: “what does this mean for global simulations”