

## Review ESSD-2017-93

Interesting and complicated product, describing and presenting a variety of outputs. I read the editor's initial comments in the on-line discussion; I can see how the submission might challenge ESSD. I tend to agree that the submission does not quite fit in ESSD but also does not fit elsewhere. I suspect GMD - the other plausible Copernicus journal - would not take it. Overall I consider the manuscript well-organised and very informative. I will comment on access to and utility of the products - data as well as model outputs - below. In this review I focus on a goal taken directly from the abstract "facilitate and encourage its use within and outside of the WASCAL community" with perhaps a slight preference for the 'outside' users.

In looking for the filename conventions in Table A4 - very useful, by the way - I noticed that the entire table organisation seems odd. Why do we get Table 1 - also essential - in the text right at the top, then tables A1 to A6 as part of an Appendix after text but before references, and then start with table 17 through 19 in the more usual location after references and ahead of figures? What happened to tables numbered 7 to 16? The authors should sort this. Definitely we need Table 1 in the text. Other tables could appear together, under a consistent numbering scheme, after references before figures? Then the Appendix with 6 additional tables?

Starting from the very important file source information in Table 1 (and using the login for WDCC provided by ESSD, which works - thank you - for both access and anonymity) I downloaded:

WRF12\_HADGEM2\_RCP45\_2039-2050\_tsmin\_24H: surface\_temperature

WDCC CERA very well organised, abundant information, clear presentation, almost too many options!

From information obtained in Table A4 I expect 12km, Hadley, RCP4.5, decadal slice 2039 to 2050 (one year spin-up plus 10 years model output), for daily minimum surface skin temperature, at daily (24h) resolution, in netCDF.

I also tried the Pangaea link. Here one needs to go to the bottom of the landing page (links listed in the middle of the page take one directly back to the same list on the WDCC CERA site), view the data offerings in html, and then make a download selection. I downloaded DAC\_wa12clmN\_hadgem2\_rcp45\_pr\_2020\_2049\_DAYSUM. Pangaea files have a slightly different naming convention than the WDCC files but transparent enough so that I think I got 12km, Hadley, RCP4.5, 30-year slice, daily summary, also in netCDF.

I could open, view and manipulate both files in R and Panoply. Confirm formats and accessibility, good options for present and future users.

Page 3, line 7: provide a link here to <http://www.wascal.org> so readers can see which countries, follow the data recovery activities, etc?

Page 3, line 9: not only of terrain-induced phenomena but also landscape and land use, organised convection, soil moisture and surface storage or run-off, etc.? Later the authors mention coastline.

Page 3, line 13, improved representation of the WAM: fact? Documentation? Or prediction? Validated by this work?

Page 3, line 18, northward movement of monsoon system: which suggests Atlantic SST, hemispheric- or continent-scale atmospheric circulation patterns, land-sea temperature differences, seasonal variation of all of the above, plus regional and local factors addressed here. Point: because regional monsoon nested within larger hemispheric features, can't rely on the RCMs in isolation? This point emerges again in the authors' conclusions, e.g page 10 lines 33 and 34.

Page 4, lines 7,8, improved resolution of coastline and topography: here the authors repeat the conclusions of prior work or make their own assessments?

Page 4 line 11, “entire”? Somewhat subjective, others might disagree especially for 12km domain. Say large areas of continental West Africa or just say continental West Africa? Can you have a ‘part’ of a century? Better to say for extensive periods of the 21st century?

Page 4 line 28: One could argue about the ‘reasonable’ scenario of RCP4.5 but the authors have stated and justified their choice.

Page 4, lines 29,30: cover the temp and precip extremes of the <50 km?> CORDEX runs? “two scenarios” refers to 4.5 and 8.5?

Page 4 line 32: First table after Table 1 is Table 17?!?!

Page 5,6, Section 2.2 on WRF model configuration. This section and information from Table 19 will prove very useful and very important to the quality of these model outputs. WRF serves an extensive community of users with many options and versions; some of us run minimalist versions on our laptops. Should the authors ‘freeze’ their version of WRFV3.5.1 with all their option choices modifications and archive that version along with the model outputs, at CERA and/or Pangea? Otherwise, a reader / user two years from now, with a new higher-resolution re-analysis from ECMWF and the CMIP6 GCMs, will not have the ability to evaluate those changes with the identical RCM used here and now. I do not know version control procedures for WRF - probably good - but many users and especially these high-end users develop their own varieties or ‘flavours’ for specific purposes. In this case the authors should make extra effort to preserve their particular version for future re-use?

Thinking again of current and future users, I missed any discussion of the DEM. Did the authors have access to and use of a DEM of sufficient quality and resolution? Did they use a new or improved product that readers should know about? Either way, especially given their focus on topographic resolution, we should know about the DEM?

Page 6, line 24: do readers have access to the ‘un-netcdf’ tool? If not, should they? Same question for the Python post-processing utilities (page 7, line 1): available to present and future users?

Page 7, lines 8 to11: Presumably the authors used UDEL Willmott and Matsuura because of its 0.5 degree resolution and likewise the Ag version of NASA’s MERRA re-analysis for its high resolution. But CRUTEM4, admittedly at only 5 degree resolution, reminds us of the temporal and spatial paucity of valid surface met data in this region (particularly in the northern third or half of the modelling domains, and for that reason WASCAL includes focus on surface data improvement and data recovery). By implication, therefore, both the UDEL observations and the AgMERRA re-analysis necessarily require a substantial degree of interpolation and/or satellite blending over West Africa. Have the authors considered the implications that their 12km products almost in all cases have substantially higher resolution than supporting observations? Readers should know the authors’ assessment of how this mismatch impacts their assessment of RCM performance? Any observed fit or mis-fit of, for example, AgMERRA to UDEL reflects at least in part the fact that both products used similar assumptions and perhaps the same gap-filling modifications in data-sparse regions? Comparison of model products to UDEL or AgMERRA thus involves similar uncertainties?

Page 9, line 14: Here a reader finds reference to the Appendix, which follows after text but before references. Typically for ESSD and Copernicus journals an Appendix appears last, after all other text, references, acknowledgements, figures, etc. The Appendix includes Tables A1 to A6? If so the authors should fix the numbering of the standard (references in text) tables?

Overall the authors have provided a useful assessment of these initial WASCAL products. They offer readers a fair discussion of the outcomes of four forcing strategies (ERA-Interim and three CMIP5-class GCMs) and probably offer reasonable validation comparisons with UDEL surface met and AgMERRA re-analysis (subject to the caveats already mentioned about data abundance and quality). The inclusion of CCLM seems to represent a parallel effort, not (apparently) subject to the same open access expectations applied to these products? For this reviewer, looking

backward to 50 km CORDEX runs based on CMIP5 (or, in some cases, CMIP3?) GCMs offers not much useful information. One might like to see, instead or in addition, some hint of plans for exploiting CMIP6, and especially HiResMIP, outcomes. Nevertheless, with their nice summary (final two paragraphs on page 10 continuing into page 11), these authors clearly understand future challenges. This overall product, with its very clear methods and well-described strengths and weaknesses, even if it represents an interim step in a longer sequence of applications and improvements, establishes a very important benchmark for open access and clear documentation; the regional modelling community should take note! One can anticipate much research, perhaps contentious, going forward, proceeding from this very positive open contribution to the climate research community. I understand hesitation for ESSD, but if the editors agree this reviewer urges publication.