Review of "Comparison of land-surface humidity between observations and CMIP5 models" by R.J.H. Dunn and co-authors

This paper provides a welcome comparison between a hierarchy of datasets based on observations and models. The choice of datasets is a reasonable one, although a little more reference to other datasets could be given. I also find the level of discussion reasonable for a paper of this type. I thus consider the paper worthy of publication subject only to minor revisions to take into account the points listed (in no order of priority) below.

1. Page 2, lines 29 and 30. It would perhaps be better to add a word such as approximately, so the text reads "Relative humidity over oceans from reanalyses appears approximately constant ...". Whilst the reanalyses carried out to date probably do not assimilate moisture data over the oceans well enough to detect reliably a trend over the ocean, Hersbach *et al*. (2014; doi: 10.1002/qj.2528) do show a slight decline in dew-point depression in the ERA-20CM ensemble of model integrations with prescribed SST and CMIP5 forcings. Simmons *et al*. (2017; doi:10.1002/qj.2949) show a slight change over time in the difference between marine air temperature and SST in both the ERA-Interim and the JRA-55 analyses. This is a feature also of CMIP5 models (Cowtan *et al*., 2015; doi: 10.1002/2015GL064888). There may thus be a slight shift in the relative humidity of near-surface air over the oceans over time, even if it cannot be reliably detected directly in the reanalyses.

2. Page 3, lines 13 and 14. The reference here to thermal comfort of humans and livestock, and productivity, is rather repetitive of what is stated in the third paragraph on page 2.

3. Page 3, line27. The term *historicalNAT* should be explained here, where it first appears, rather than later.

4. Page 3, line 33. It would be better if the sentence referring to Sect 2 were to appear before the sentence referring to Sect 3, rather than before a sentence that begins "Finally ...".

5. Page 6, line 15 and 16. A stronger justification of the selection of ERA-Interim as the reanalysis to be included in this study could be given. Only ERA-Interim and JRA-55 of the major atmospheric reanalyses provide direct analyses of 2m temperature and humidity observations. Willett *et al*. (2016) shows much better agreement between ERA-Interim and JRA-55 than between either of these reanalyses and MERRA-2, especially for relative humidity. It probably would add little to include JRA-55 as well as ERA-Interim in this study, given that its main focus is on the comparison with CMIP5 models rather than the intercomparison of reanalyses, and the humidity analyses from ERA-Interim are the better documented in the literature. MERRA-2 cannot be recommended for use in this study as it does not give a reliable time series of two-metre temperature (Simmons *et al*., 2017; doi:10.1002/qj.2949). Its inhomogeneity is much larger than that which arises from the ERA-Interim SST changes.

6. Page 6, lines 24 to 28. It's more complicated than stated in that there was an ERA-Interim SST change in January 2002 as well as one in June 2001. It was the combined effect of these two changes that shifted SST about 0.1K colder, a shift we now adjust for in studies such as Simmons *et al*. (2017). The other changes in source of SST analysis are more minor in their impact. Apologies for more self-citation, but reference could be given to the summary of the SST and sea-ice changes given by

Simmons and Poli (2015; doi: 10.1002/qj.2422), which notes the January 2002 change as well as the June 2001 one.

7. Page 6, line 29. “inhomogeneities” might be a better word than “instabilities”.

8. Page 8. The three columns of panels in Figure 1 could be headed “Air temperature”, “Specific Humidity” and “Relative Humidity” to help someone glancing through the paper. At present these words appear only in the headings of the lowermost panels.

9. Page 12, line 1. m/m-1. is zero under the usual convention of doing the divide before the minus. Do the authors mean m/(m-1.)?

10. Page 12, line 19. Whilst the CMIP5 models all have larger positive trends than observations indicate for the period 1996-2015, the observed warmth of 2016 (and perhaps 2017, given it has started warm and another El Nino is forecast) makes one wonder what the conclusions will be when we have CMIP6 integrations – even if the models do not change much. Some discussion relevant to this is given in section 7 (see also comment 20) and perhaps a reference to the discussion later in the paper could be given on page 12.

11. Page 14, line 2. Missing word “be” before “expected”.

12. Page14, line 17. Reference here could be made to Hersbach *et al*. (2014), as the finding for HadGEM3-A mirrors that already made for ERA-20CM. Hersbach *et al*. showed that in model runs for the period 1900-2010, the driest conditions occurred in the final decade, both for surface air humidity and for soil moisture. Although the dryness of the ensemble mean was not as large as that in ERA-Interim, a few of the ten ensemble members reached levels of dryness similar to those reached by ERA-Interim.

13. Page 17. The final sentence of the figure caption states that all climatologies were calculated over the 1975-2009 period. This cannot be the case for ERA-Interim.

14. Page 21, line 27. Missing word “in” before “ingested”.

15. Page 21, line 27. The change in January 2002 as well as June 2001 is relevant in this regard, as discussed in comment 6.

16. Page 21, line 32. I baulked a little when I read of relative humidity becoming more arid. Is this a correct use of the word “arid”? Land and air can become more arid, but relative humidity?

17. Page 24, page 26 and page 27. The colours referred to in these figure captions are not all correct. The captions state that ERA-Interim is shown in blue, whereas I think from the figure that the colour is deep pink (or red). The captions of Figures 8 and 10 refer to a yellow that should be orange, and a red that should be blue.

18. Page 28, line 10. “ERA-interim” should be “ERA-Interim”.

19. Page 29, line 28. “spare” should be “sparse”.

20. Page 30, lines 15 to 25. A more nuanced discussion should be given, and should reflect that we now have the figures for 2016. Reference could be given to Simmons *et al.* (2017), who compare the temperatures from several datasets for the period up to July 2016, with numbers until the end of 2016 presented at

http://climate.copernicus.eu/resources/data-analysis/average-surface-air-temperature-analysis

The datasets show some variability when it comes to how much warmer 2016 was than 2015, with HadCRUT4 showing little difference, and datasets that provide values over more of or all the Arctic and Antarctic giving warmer values in 2016 than 2015, the difference approaching an unusually large 0.2°C in the case of ERA-Interim and JRA-55. However, these differences are probably of limited relevance in the context of the paper under review, as they stem mainly from regions over or near to sea-ice, which has been of record low extent in both the Arctic and the Antarctic in recent months.

Simmons *et al*. (2017) is also a relevant reference in that the paper shows that the latest estimates of the trend from 1998-2012 are (for all datasets examined) higher than the central estimate made in IPCC AR5, suggesting less of a "hiatus" or slowdown in warming than first indicated.

21. Page 31, line 20. I am puzzled by the sentence that the reanalyses do not show as clear a warming trend as the observations. This is not what is concluded by Simmons *et al*. (2017) when comparing ERA-Interim (adjusting for its warmer SSTs prior to 2002) and JRA-55 with GISTEMP, HadCRUT4 and NOAAGlobalTemp.

22. page 32, line 4. The text here should refer to "the drying identified in ERA-Interim and HadISDH". It should be in this order as the drying was identified first in ERA-Interim (published in 2010, confirmed then by an intermediate Hadley Centre dataset) and confirmed later by HadISDH (published in 2014).

Adrian Simmons
16 March 2017