Response to Reviewers

Reviewer #1

Review of "HadISDH land surface multi-variable humidity and temperature record for climate monitoring" by Kate Willett and co-authors

This paper provides a mostly clear account of a very good piece of work, and certainly merits publication, subject to revision to clarify its comparison with ERA-Interim, and to take into account the other comments given below as the authors see fit.

Many thanks Adrian for a very thorough and constructive review. We have tried to take on board all of your comments in this version of HadISDH and certainly will do for future versions. Your advice and support is highly valued.

Please note that we have made some additional very minor amendments where we felt the text was not clear. One fairly major addition is some discussion in the main text over the removal of stations that had post-homogenisation issues of supersaturation or sub-zero values. We realised that this text was contained only in the supplementary material and so we have added this in to Section 3 with a reference to the supplementary material. The removed stations have always been shown in Figure 2 and counts listed in Table SM2.

"A further requirement is that station data must not show supersaturation (all humidity variables) or sub-zero values (for e, q, and RH only). The application of ID PHA is intended to maintain the physical consistency across variables compared to direct PHA but there are cases where this does not hold true. We have removed all such stations with physically unrealistic data from gridding and further analysis. This is discussed further in Supp. Mat. Sect. 7. Table SM1 lists the number of stations removed because of this for each variable and station locations are shown in Figure 2 (Figure SM1). This is a minor problem for e, q and RH where only 28 stations are removed due to supersaturation and 52 stations removed due to subzero values (e and q only). This is a significant for T_w (808 stations removed) because it is much closer in value to the dry bulb temperature which means that even small adjustments can result in supersaturation. Nearly all stations north of 60 °N are removed from the gridded T_w product resulting in very few stations common to all variables above this latitude (Figure 2a). "

Comments related to ERA-Interim (and JRA)

1. The HadISDH data are described (Abstract, line 31, and Conclusion, page 18, line 45) as a "unique tool" for climate monitoring of a number of humidity variables. It is unique in the sense that it is different to ERA-Interim, but ERA-interim is just as appropriate for studying all these variables. This is because ERA-Interim is made available as six-hourly analysis values from which monthly means of all the humidity values available from HadISDH can easily be computed using standard routines for converting humidity data. Yes, HadISDH takes some of the work away for the user, but ERA-Interim provides the user with six-hourly values (and other meteorological variables), and globally complete values, which HadISDH does not. Anything said for ERA-Interim applies also for the JRA-55 reanalysis, which similarly provides analyses of synoptic humidity observations. So it would be fairer to say that HadISDH offers a tool that is complementary to what is provided by reanalysis, not that it provides a unique tool. Or simply omit the word "unique".

This is a really good point. Please forgive our obs-centric blinkeredness here. We feel that particular attention has been paid to the long-term homogeneity of HadISDH to make it fit for the purpose of

monitoring climate change. This differs slightly from reanalyses that ingest various data-streams as and when they become available although we recognise that great care is taken to minimise this effect in the reanalyses. There are certainly many uses for which ERA-Interim and other reanalyses are suitable where HadISDH is not. We've amended the text to recognise this as follows:

Abstract:
*"HadISDH.2.0.0 is the first gridded, multi-variable humidity and temperature **in situ observations only** climate-data product that is homogenised and annually updated."*
...
*"It provides a tool for the **long-term** monitoring of a variety of humidity-related variables which have different impacts and implications for society. **It is also useful for model and reanalyses validation.**"*

Removed "where they are available" from "HadISDH.2.0.0 is shown to be in good agreement both with other estimates"

Concluding discussions:
*"HadISDH.2.0.0 is the first multi-variable humidity and temperature **in situ observations only land** climate data product that is homogenised and annually updated."*
...
*"It **is designed as a** tool for climate monitoring, **where long-term stability is essential, providing** a variety of **physically consistent** humidity variables which have different uses for understanding of the earth-atmosphere system and societal applications. It is also useful for validation."*

2. In this regard, Figure 12 (and the discussion in the paragraph from lines 27 to 42 on page 16) is misleading. ERA-Interim plots are shown for several humidity variables, but not all of them. There is no logic to this. The only ERA-Interim surface air humidity variable for which monthly-means are archived and made available as standard is dew point. q and RH are also shown in Figure 12 but not Tw and e. Why? The q and RH values must have been calculated from the archived 6-hourly Td and associated non-humidity variables – probably by me in providing values to the lead author for her use in the BAMS State of the Climate article, if I interpret the paragraph on page 16 correctly. But Tw and e could equally have been calculated. So Figure 12 should include time series from ERA-Interim for these values also.

Again a really good point. We are very happy to take this opportunity to calculate ERA-Interim Tw and e. Figure 12 has been amended and the text has been amended as follows:

Section 5.3:
*"The ERA-Interim reanalysis (Dee et al. 2011) is included here for T, q and RH comparison, following comparison with earlier HadISDH humidity and CRUTEM temperature products (Simmons et al. 2010; Willett et al., 2013). ERA-Interim fields for T_d are available and DPD has been derived from the monthly T and T_d fields. **We have also calculated monthly mean e and T_w fields for ERA-Interim using 6 hourly T and T_d fields**."*
...
*"This is not obvious in the difference series although some digression around that time is apparent for **all variables other than T. Interestingly, for q, RH, e and T_w HadISDH exhibits a slightly more humid period between 1997and 2004.**"*

3. Discussion of the geographical distribution of the recent decrease in RH could also have been done a little more carefully. There's nothing to complain about the discussion of mid-latitude values. What is not acknowledged in this HadISDH paper is that it was shown in Simmons et al. (2010) that RH was increasing, overall, in high northern latitudes in ERA-Interim, and that Td was increasing in

the 60N-90N region in both ERA-Interim and the JRA-25/JCDAS predecessor of JRA-55. This leaves the Tropics, where the discussion gets a little difficult. Simmons et al. showed RH to be increasing recently over India and West Africa, which is reproduced in HadISDH. The Caribbean seems to be an area of disagreement. HadISDH misses a large part of the tropical land mass by having almost no values over central Africa. ERA-Interim produces drying there, but this is regarded as unreliable, in view of rainfall shifts discussed by Dee at al. (2011) and by Simmons et al. (2014) with regard to a suspect linked temperature trend. But it should at least be acknowledged that HadISDH does not have sufficient coverage of the Tropics to make a definitive statement of what is happening to relative humidity there in an average sense.

This is a fair comment. We have revisited the papers mentioned above and amended the discussion in the following way:

*"Figures 12f, 13f and 13n show continuation of the decline in RH since 2000 noted in Simmons et al. (2010). This is apparent in the global, Northern Hemisphere and Southern Hemisphere time series although too short to be considered a long-term trend at present. As shown in Figure 11f, it is dominated by the mid-latitudes **while the high latitudes and Northern Hemisphere tropics, including the Indian sub-continent show increasing saturation. For the most part this replicates findings in Simmons et al. (2010) for ERA-Interim. However, aside from West Africa and India, ERA-Interim shows drying across the Northern Hemisphere tropics. ERA-Interim drying over central Africa is thought to be unreliable (Dee et al., 2011) and cannot be compared with HadISDH as there is no data coverage. Given the poor data coverage and large uncertainties for HadISDH over the tropics, the validity of the increased saturation shown in the Caribbean for HadISDH is unclear**.*

*These features of reduced saturation are also shown in the DPD data (Figures 11g and 12g). **Regardless whether the Caribbean is becoming more saturated or not,** the zonal nature of these features suggests that there must be a large scale contributing factor, possibly linked to atmospheric circulation."*

We have also added a paragraph making the sampling limitations of the tropics and Southern Hemisphere much clearer:

*"**For the tropics and Southern Hemisphere decadal trends in q are present only 49% and 59% of gridboxes respectively. For the globe this rises to 72%, largely driven by the relatively well sampled Northern Hemisphere at 84%. As such, there is much larger uncertainty over the tropics and Southern Hemisphere. This is reflected in the spatial coverage uncertainty estimates discussed below and shown in Figures 12 and 13.**"*

Other comments:

4. Page 2, line 49. I suggest adding "availability of" before "water".
Done

5. Page 3, line 32. "Reanalysis products" would be better than "Reanalyses products".
Done

6. Page 4, line 15. One can argue over the words "first operational"; I dislike this wording for the same reason I dislike the reference to HadISDH as a "unique tool". Again, I would prefer HadISDH to be advertised as complementary to reanalysis.

We have added 'only' to this sentence so that it is clear that this is the first 'in situ only' operational product available. We hope that we have been clear enough about the limitations of an obs only product.

7. Page 5, line 17. Although it may not be "possible to identify which instruments were used", it should be possible, for most data at least, to identify whether they came from a manual or an automatic station, and stratify data statistics accordingly. If the metadata are not in HadISD, most of its data can probably be matched with what is in ECMWF archives, where we do have the information. See Simmons and Poli (DOI: 10.1002/qj.2422) for an example of use in study of Arctic temperatures. This is not something I recommend for this paper, but it is something that should be kept in mind for future work, and could be identified as such in the text. It is also something that needs to be brought up with WMO with regard to what they are doing for metadata from now onwards, under WIGOS. The Met Office is well placed to do this.

This is a very good point and certainly something that we will attempt for future versions. Part of the reason for not doing so in this version was the decision to use HadISD as the base product. This was done because of the QC processing that is done as part of HadISD. However, HadISD does not pull through such information. We have done a quick investigation into what information is stored in the ISD that could be used for this purpose. There are a few pieces of information within the code that could be used to imply whether a station was automated but nothing that would guarantee accurate categorisation for all stations. We will investigate whether the ECMWF archive may better suit our need for future versions. A quick look at this left us a little confused on how to access the archive in the first place but we know who to ask to help.

We have added an explicit mention about the value of metadata availability in the conclusions section:
*"Thus, the best way to reduce uncertainty is through: **improving spatio-temporal coverage by** greater data-sharing and increased data-rescue efforts (Thorne et al., 2011, Allan et al., 2011); **and improving homogenisation by increasing the digital availability o**f metadata (i.e., instrument type, station type [manual or automated])."*

8. Page 5, line 27. Why were *monthly-mean* reanalysis data for Pmsl used? Values are usually available at least 6-hourly from reanalyses. If not for 20CR (I've not checked), then ERA-40/ERAInterim, JRA-55 and ERA-20C are alternatives that could be explored. Of course instantaneous reanalysis values of Pmsl will not be perfect, but they surely will be better than monthly means, especially if a reanalysis with higher spatial resolution (better orography) is used. JRA-55 is probably the best for this, as it is the newest comprehensive reanalysis. This is something that should be investigated for fixing at the next release of HadISDH, and it should be acknowledged in the paper that this issue is likely fixable with only a little work. In this context, it should be noted (see comment 2) that if the monthly q values for ERA-Interim shown in Figure 12 are the ones I calculated, then the six-hourly ERA-Interim surface pressures were used in converting from Td to q.

At the time of processing we were limited to using 20CR because we required data back to 1973. We actually used a climatology so one could argue that any of the other reanalyses would have been sufficient assuming that a climatological average over 1979-2005 would not be very different from the chosen 1976-2005 climatological average. The main reason for using a non-varying pressure term was to ensure that any trends in the converted data came solely from the observed data. If we had used 6hrly surface pressure data from reanalyses there is a small (most likely very small) chance that this could influence changes in the converted humidity variables. We assessed the additional error from using climatologically averaged surface pressure for the gridbox mean and found it to be

9. Page 6, line 22. Figures 2 and 3 should be rethought. It is almost impossible to see the differences between many of the panels. For Figure 2, perhaps one could show a single map of the observation locations that provide data for all variables, and then in other maps show just the extra data for each variable. Either way, something needs to be done. Also, the label for map g) in Figure 2 is wrong, as it lacks the word "depression". Same is true for panel g) of Figure 3.

Thanks for the suggestion for Figure 2 – and the corrections. We like your idea and have adopted it. We have also reduced the point size of the 'good' stations and changed the colours of the stations to try and improve clarity. We have decided to move e, T_w and DPD figures to the supplementary material in all cases to allow more space for what we consider to be the key variables: q, RH, T and T_d. Figure 3 has also been revamped both to compress the y-axis to only include stations within HadISDH rather than the whole WMO ID range, and also to now show missing data by WMO ID rather than data present. This paints quite a different picture as it is now far clearer where and when missing data occur. In particular, the 2005 drop off is clearly attributable mainly to the USA station fall out. We have added discussion around this figure in section 2 and in the Supplementary material:

"The tail off at either end of the record and especially post-2005 is partly due to the initial HadISD data selection which maximises coverage over the 1976-2005 climatology period and has remained static since version 1.0.0. of HadISD. This will be revisited in the future as will the list of stations that require merging. The ISD archive is undergoing continual improvements in terms of data coverage which will filter through to HadISD and HadISDH in time. The post-2005 drop off is primarily driven by USA stations (Figures 3, SM2, SM3 and SM4). These stations are reasonably evenly spread across the USA. Any users wishing to focus analysis over the USA should be aware of this issue but given the 600+ remaining stations with reasonably continuous records this should not be a significant issue.

And section 7:
"Efforts have been made to maximise spatio-temporal coverage of the data. However, missing data are prevalent, particularly at the start and end of the period of record and from 2005 onwards for ~120 USA stations. Data are missing for three main reasons: they are not present in the ISD archive; station records are too short or have insufficient data over the climatology period (1976-2005) to be included in our selection; and data have been removed by the QC or homogenisation process. In many cases, short station records often originate from the same station but have different reporting IDs. These can be merged into one long record. The selection and merge process employed here has been static since 2008. A revisit is planned which should improve on spatio-temporal coverage. Although the change in spatio-temporal coverage over time is substantial, it appears to have little impact on the large scale features. There are no obvious signals in the data that would be consistent with the pattern of missing data over time and there is good agreement between HadISDH and other datasets that do not use the same station selection."

10. Page 6, line 24. The drop-off post-2005 in the data entering the ISD archive is very disconcerting. This does not happen for the GTS data received operationally at ECMWF (and doubtless also the Met Office), and these GTS data are used in ERA-interim, which is quite competitive with HadISDH, so the GTS data cannot be that bad. Is a rethink needed as to whether ISD is the most appropriate source of input data for HadISDH? OK the name HadISDH might have to change. If the decision is to stick with ISD, steps need to be taken to address the post-2005 drop-off. This needs a sentence or two of extra discussion in the paper.

We think that there are a couple of reasons for the post-2005 drop off. The station list used for HadISDH was selected in 2008 when HadISD was first worked on. We selected the climatology period of 1976-2005 because this maximised the number of stations that had enough data. The list of stations that could be merged has also been static from that time onwards. This means we're probably missing quite a few stations that could be merged to create longer records. We hope that we will be able to revisit this for HadISD next year which should improve things. At that time we will also assess whether ISD is the best source to continue with.

We have added text to explain both the 2005 drop off to the reader as described in response to point 9 above.

11. Page 6, Section 3. Was consideration given to making use of the background-observation values from reanalysis in the homogenisation? Haimberger has used this with some success for radiosonde data.

No, but this is definitely something worth considering for future versions. Having both a 'simple' PHA homogenised version and a more complex homogenisation using reanalysis background fields, or more likely a combination of the two will make an interesting comparison.

12. Page 7, line 27. There are lines of text running from the a), the b) and the c) in Figure 4 that I could not read in the printed version I worked from. Is the intention simply to delete them in the final version of the paper?

Yes, sorry these should have been removed as they are just date/code stamps to help us keep track of figure production.

13. Page 8, line 2. "are" should be "is".

Done.

14. Page 10, lines 24 to 39. See comment 7. One should be able to do something with metadata identifying the station as manual or automatic.

When discussing future iterations of HadISDH we have added:

"For future versions of HadISD, and HadISDH, it may be possible to identify for some stations whether they are automated or manual, and changes over time from the original ISD code. An assumption could be made that all automated stations use sensors and all manual or non-identifiable stations used psychrometers. This would allow a more specific measurement uncertainty model to be applied."

15. Page 13, lines 27 and 28. The phrase "increasing ... trends" is ambiguous. An "increasing trend in q" could mean that the trend in q is getting larger as time progresses. But I suspect the alternative, better stated as "the trend for q to increase", is what is meant here. I suggest that the sentence be reworded.

We have changed the word 'increasing' to 'moistening' so that is hopefully clearer:

*"The variables that are solely affected by the amount of water vapour in the atmosphere (T_d, q and e) mostly show significant **moistening** homogenised (raw) trends,.."*

16. Page 13, line 44. I suggest that the Clausius-Clapeyron relationship be referred to as "basic physics" rather than "theory".

Done.

17. Page 15, line 46 and 47. This paper indeed shows a continuation of the decline in RH noted by Simmons et al. (2010), but this has also been shown in BAMS State of the Climate articles for the years since 2010, so it would be fairer to refer to the BAMS articles, and state that this paper confirms what is shown there.

Done.

18. Page 16, line 47. "(and ERA-Interim)" could be added after the words "support to the validity of HadISDH".

Done.

19. Page 19, line 12. See comment 16, concerning use of the word "theory".

Good point, thanks.

20. Page 19, line 35. "comes from" is perhaps better than "is driven by".

Good point, thanks.

Adrian Simmons