Reviewer #2

Reviewer#2 of "HadISDH land surface multi-variable humidity and temperature record for climate monitoring" by Kate Willett et al.

Following up on the previous years' publication of Willet et al. (Clim. Past, 9, 657–677, 2013) this paper describes the next version of the HadISDH product and its enhancements towards a multi-variable product covering all variables relevant to determine humidity from land-surface observations.

We again look at an excellently written paper and there's no doubt that it merits publication as it complies with the CPD assessment criteria on the levels „good" or „excellent".

The Significance is excellent. A self-consistent multi-variable purely observational data set, ameliorated by two approaches of homogenization is the highly needed information required to access re-analysis data sets like the ERA family and vice versa. A comprehensive comparison between the two worlds of climate reality is not provided here (to the disappointment of reviewer # 1) but I think this would go beyond the scope of this paper.

The Scientific Quality is excellent as well. All methods are described exhaustively and fellow scientist with enough talent and time would be able to reproduce the results presented here, given that all original data would be provided to them.

Presentation Quality is good. Despite many excellent Figures (we are all used to from reading BAMS SOC) there is a little bit of improvement potential remaining as it comes for Figures 2 and 3 were information that should make it into the captions is hidden almost unreadable (the reviewers eyes are obviously weaker than the authors' ones)

Many thanks Andreas for your very positive review. We have taken on board all of your comments and hope that you are happy with our responses. We certainly plan future papers using the HadISDH dataset to make more thorough comparisons with the reanalysis products and also investigate the RH decline. We feel that a data issue is unlikely to be the cause but cannot be ruled out. Your comment about the figures is duly noted. We have decided to move many of the figures for wet bulb temperature, vapour pressure and dew point depression into the Supplementary Material in the hope that this will allow the figures to appear larger in the manuscript. We have also increased the text size in many cases. I'm struggling to find any typos in your review.

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). "

Minor issues:

The homogenization and (rather simple) interpolation methods and their robustness despite the rather limited set of stations take advantage of the fact that temperature and humidity variables behave far more "friendly" than precipitation. Maybe the authors are inclined to insert this caution remark at a suitable place in the manuscript to prevent those colleagues tending to generalize this approach towards precipitation on similar sized (< 10 000 stations) data bases from disappointments.

This is a very good point. We have added a cautionary note in section 3.3:

"It is likely that these methods will be less suitable for types of variables that have shorter spatial correlation distances (e.g., precipitation)."

p2720 l22 – p2721 l4: This is a quite speculative statement, only valid for land-surfaces with "marine" climates.

We're not sure how best to respond to your comment. We hope that this paragraph does represent the lack of clear understanding as to what is going on here. In essence, it is all rather speculative although founded in plausible physical processes. We feel that our use of the word 'Much' covers this to some extent as it is true that most of the water vapour over land has originally been evaporated over the oceans. My textbook estimates 15% of all water vapour is evaporated from the land (including transpiration) with the rest from the oceans. Given that 71% of the earth is ocean we feel safe stating that much of the water vapour over land is transported from over the oceans. We have added these figures in because they may be of interest to the reader and hopefully address your point.

"Much of the moisture over land is transported from the oceans where it was originally evaporated. ***Although the oceans cover just over 70% of the earth surface, ~85% of atmospheric water vapour is evaporated from oceans with ~15% coming from evaporation and transpiration over land (Ahrens 2000).****"*

P2722 l13: If you use superlatives like "first" please add "to the author's knowledge".

Added in two places.

Some sub-headers in the Section 2 (like you have done in Section 3) would help structuring the step by step description of the processing

Added two sub-headers:
2.1 Deriving the humidity variables from temperature and dew point temperature

2.2 Processing of hourly values to monthly climate anomalies

P2723 l14: Maybe you want to make this link to the station list explicit by adding "v102_2013f/files/hadisd_station_info_v102.txt" to the generic one.

We have added the following:
"The full HadISD station list is available at www.metoffice.gov.uk/hadobs/hadisd/v102_2013f/files/hadisd_station_info_v102.txt with variable specific lists alongside the HadISDH data product at www.metoffice.gov.uk/hadobs/hadisdh."

P2726 l20: The COST HOME benchmark (Venema et al., 2012) favoured other homogenization packages (PRODIGE). Maybe you want to justify why you stick to the USGHCN one.

There were four reasons for using PHA: it has a very low false alarm rate; it is easily implementable on a few thousand stations because it is automated; it runs very efficiently (~30 minutes); and it has been quite thoroughly tested (both COST HOME and the Williams et al. 2012 benchmarking exercise). PRODIGE is not as easy to access or implement and we're not very keen on the parallelising of the months in particular. Perhaps something to investigate further for future versions though. We have discussed the four reasons above in the text already so feel that we do not need to add any more words. See text below:

"There are very few automated homogenisation methods able to be applied to networks of the order of several thousand stations ***(Venema et al., 2012).*** *NOAA NCDC's Pairwise Homogenisation Algorithm (PHA; Menne and Williams 2009; Supp. Mat.) has been used on the Global Historical Climate Network – Monthly (GHCNMv3, Lawrimore et al., 2011) and was also used for HadISDH.landq.1.0.0.2012p (see Sect. 3 of Willett et al. 2013a). It has been tested against a set of benchmarks for the USA (Williams et al. 2012) and through the COST HOME benchmarks (Venema et al. 2012). Overall, it was found to bring temperature data closer to their 'truth' but be overly conservative in places; i.e., its adjustments tended to be too few while it avoided making adjustments where none were needed. For the multi-variable HadISDH it is a good choice both because of its previous validation and also because it is computationally efficient. This allows multiple runs during method testing and efficient updating. The code is freely available from http://www.ncdc.noaa.gov/oa/climate/research/ushcn/#phas."*

P2739: The RH decline is really an irritating feature, together with its coincidence with the hiatus. Maybe we only need to wait for the upcoming El Nino, but to what extend could automation of weather station that is still ongoing but started in the early 2000ends play a role?

This is a very interesting question and something that we cannot rule out. Sadly, for this version we had not pulled through the metadata in the ISD database that may allow us to identify whether a station is automated or manual. There are future plans to pull this through to HadISD and then through to HadISDH. In some cases we may be able to tell when a station became automated. We do not think that we will be able to do this for all stations though. The spatial patterns of the drying suggest something at least partially climate related – the drying region covers almost all of the mid-latitudes both north and south. If it was an automation issue we may expect different countries to show features at different times. We may not expect the Northern and Southern Hemispheres to begin drying at exactly the same time. The drying also has reasonable explanations in basic physics although this has not been proved to be the case as yet. This is an important discussion that should be included in the text and so we have added the following in the Concluding discussions section:

"It is possible that the increasing automation of observing instruments, which began in earnest from the late 20th Century, is a contributing factor. However, the widespread signal is temporally

consistent across both hemispheres and physically plausible. Furthermore, the data have been homogenised which should have removed any large artefacts from changes in instruments over time."

P2727 l6: Just proud I found a typo as a non-native speaker reading a native speakers paper: omit "a"

Thanks for spotting this one. A little embarrassing for us but your English is very good.

P2756: Caption Figure 7: The second typo I managed to spot in the entire manuscript: Last sentence should end with "than in the Russia/Eastern Europe group."

As above.