Further comments:
Comments by reviewer #1:
I have no further substantive comments, but I did notice a few small typos in the revised paper that should be corrected.
The following very minor comments refer to line numbers in the "track changes" version of the revised paper.
138 - there looks to have been a small latex error (missing \ before sigma).
161 - The sentence ending with "bands have ranges that encompass zero, though the most bands suggest decreases" seems grammatically challenged.
206 - "samles" --> "samples"
241 - "but in the next" --> "but we do so in the next"
412 - "changes to at these large scales" ... delete either "to" or "at"
488 - "that large fraction" --> "that a large fraction"

All typographical errors now corrected.

Review – "Changes in statistical distributions of sub-daily surface temperatures and wind speed" by Dunn et al.

General Comments:

This paper provides a detailed analysis of changes and trends in HadISD temperature and wind data. The analysis uses station data that have been quality controlled and, in some cases, homogenized. Apart from the use of quantile regression to describe changes in the upper and lower tails of the distribution, the paper uses simple statistical methods. Findings are, generally, not surprising, although some trends are found that are somewhat surprising. Comparisons are made with similar papers that have used other datasets.

Thank you. We have addressed your comments individually below, and marked in the revised manuscript where we have made changes, except for updated tables and figures, or if paragraphs have been moved into a new section.

While this kind of work is fundamental and absolutely necessary, I found the author's apparent decision to limit themselves only to the description of the data unsatisfying. Unfortunately, the paper does not offer physical insight concerning the changes that are observed, and only speculates about the causes of differences between findings reported here and those reported in previous papers based on other datasets.

This is partially because of the expertise of the authors and our desire as outlined in the introduction to find out what the changes have been, rather than assess these changes against possible explanations. We have now mentioned a number of analyses which have suggested causes for changes in the distribution of temperatures, and hence also the extremes.

It raises the issue of measurement resolution in the context of quantile regression, but does not concern itself with the impact of measurement resolution on the higher order moments, or how changes in measurement resolution might produce changes in those moments.

Given the blending across stations in the zonal analysis we do not address this point in detail there, but do raise it in the summary of that section. A more detailed discussion is now given in the summary of Section 4, which addresses both the measurement resolution and also the discrete temporal nature of HadISD, which also changes over time. We note that there will be an effect from these but are not able to quantify it with the analysis we have performed here. However, given changes in measurement are likely to occur on a national basis over time, we note that individual geo-political regions do not stand out (which could indicate a bias resulting from measurement or processing practices).

Also, while the methods are simple, describing them precisely in prose can be very difficult. The authors do write very clearly, but nevertheless, there is sufficient ambiguity in the description of the methods that it would likely not be possible for another scientist to easily reproduce the analyses that are described in the paper. A technical description of the methods using precise mathematical notation could, and should, be included in the supplement.

We have added a more technical description on the data pre-processing and how the selection criteria apply at the beginning of the supplementary information, which we hope clarifies how we have prepared the data for our analysis. The trend calculation (median of pairwise slopes) and quantile regression are standard algorithms, and so we have not reproduced these.

Detailed comments:

Supplement: The contents of the supplement belong to this paper, but the title does not!

Thank you - corrected

40 – There is a bit of confusion here concerning how to refer to extreme value theory. The correct term is “extreme value theory”, not “Generalize Extreme Value theory”. The GEV (Generalized Extreme Value) distribution emerges from extreme value theory based on one approach to the analysis of extremes – the so-called block maximum approach. Typical practice is to use one-year blocks and to model resulting collection of annual maxima with the GEV distribution. Another approach that is also often used is called the “peaks over threshold” approach, in which exceedances above a high threshold are modelled. This approach leads to the use of the Generalized Pareto (GP) distribution. Both distributions are asymptotic – in the case of the GEV, it is the limiting distribution that is obtained as blocks grow in length without bound (provided there is a limiting distribution). Similarly, the Generalized Pareto distribution is obtained as the level of the threshold increases without bound (again, provided there is a limiting distribution). Thus, it needs to be understood that for a given block length or threshold, the GEV or GP distributions respectively, can only approximate the distribution of the extremes that are identified. The Brown et al paper you cite uses the peaks-over-threshold approach.

Thank you for this clarification - we have amended and expanded this section in light of your comments to read: "The "block maximum" approach models a set of e.g. annual, maxima with a Generalised Extreme Value distribution (e.g. Christidis et al, 2011). Alternatively, the "peaks over threshold" approach models all peaks over a fixed threshold with the Generalised Pareto distribution (e.g. Brown et al 2008)"

44 – Delete “easily”. I think it is debatable whether a full distribution approach is best if the objective is to make inferences about extremes. If this were easy, it is likely that we would not have as much effort as is expended on developing and applying extreme value theory.

Done

56 – “significant” in the statistical sense? Please clarify. To avoid confusion and ambiguity, my suggestion would be to avoid the word “significant” in scientific papers except when discussing statistical significance.

In this case we are reporting the results in Donat et al, 2012, but we have added the significance level from their discussion. The only other mention of significance is in the section on Quantile Regression, where a $p=0.05$ value is used and stated.

83-85 – To what extent are the moments (particularly the higher moments) sensitive to the degree of inhomogeneity that is permitted? Presumably, consideration of that sensitivity should drive the choice of permitted jump. Otherwise the choice of a number like 1°C, while tidy, would be entirely arbitrary. Note that jumps of this size do not have the same impact on higher-order statistics in all climates; the impact might not be easily discernable in a high variability mid- or high-latitude climate, but it might be very large, in relative terms, in a tropical climate with low temperature variability. This is also where the impact of poor measurement resolution would be the largest.

We have added the following sentence to highlight the limitation of the PHA to find jumps for station-variable combinations with high variability: "This used monthly averages derived from the sub-daily HadISD observations to compare station-pair differences to identify potential jumps in timeseries, a process which is more effective for stations and variables with low variability."

At one level, the choice of 1C is arbitrary and tidy. However there are reasons behind the choice. In the homogeneity assessment of HadISD, the distribution of all inhomogeneities is shown in Dunn et al, 2014, Figure 2. The majority of all inhomogeneities are smaller than 1C, and our intention by excluding stations with a jump of more than 1C is to remove those with the greatest inhomogeneities (the most contaminated stations). Around 15% of stations have no breaks in temperature and around 7% for wind speed (Dunn et al, 2016 - HadISD v2), and the average inhomogeneity is around 0.8-0.9C or m/s (the largest of both methods for each variable – average and diurnal temperature range for the temperatures, and average and maximum for the wind speed). In an example assessment shown by Dunn et al (2014) the effect of reducing the number of stations available for an analysis by being more and more restrictive on the maximum inhomogeneity allowed showed that the smallest rms difference between monthly means in HadISD and CRUTEM was reached for a maximum inhomogeneity between 1 and 2C. The result at 0.5C indicated that the resulting small station network limited the analysis. This is what we are attempting to balance in this analysis, a reasonable station network resulting from the exclusion of the most contaminated stations. We now refer explicitly to this assessment at this point in the text. We have also corrected the maximum number of inhomogeneities allowed as well as the selection criteria which used the total number since 1931 (start of HadISD record), but now uses 1974 (and 1973 for the quantile regression section).

The effects of heterogeneity on the higher moments of sub-daily data are unfortunately complex, being . dependent on the weather at the station at each time point, not just the value of the single variable being assessed. Hence, from a monthly inhomogeneity alone, it is in our minds difficult to say how the moments of distributions will be affected from a physical perspective. However, reverting to the distribution of the inhomogeneities in Dunn et al, (2014), the distribution of the estimate jumps (rather than the fitted curve) shows no clear asymmetry for the central portions. The means of both the estimated and fitted distributions are not exactly zero, but correspond to fatter tails on one side or the other, the stations contributing to which are excluded in our analysis as they have inhomogeneities >1C or >1m/s. Therefore any systematic effect of the remaining inhomogeneities can be argued to be small, given this analysis is looking at combinations or larger regions of stations rather than small-numbers where single jumps could cause issues. We have added the following paragraph at this point in the text:
As the selected stations still contain uncorrected inhomogeneities, we note that these will have affected the analysis of the statistical properties of the observations outlined herein. But the distributions of the estimated inhomogeneities as shown in Dunn et al. (2014, 2016) do not have a strongly non-zero mean, and the central portions retained by the approach above have no clear asymmetry. Therefore, as the following analyses use combinations of stations or look for contiguous regions of change that cross national (i.e. observing-practice) boundaries, we do not expect large effects arising from the remaining inhomogeneities themselves.

A second point here is that while temperature is mentioned, nothing is said about winds. Is a similar approach used?

Yes, a similar approach, using 1m/s. We have now noted this at this point.

Also, I assume that both dry air temperature and dewpoint temperature are treated in the same way. This raises a small question about whether the same limit for jumps should be used for both, given that the impact on higher order moments may be relatively larger in one case versus the other.

Yes, they are treated the same way, with the same limit. We have clarified this in the text. The distribution of dewpoint inhomogeneities is in fact a little wider than the temperature one from Dunn et al, 2014, and so this threshold is more restrictive in this case, and hence a smaller network.

114 – Have you thought of using L-moments rather than ordinary moments? L-moments (linear moments) are generally considered to be somewhat more robust (in the statistical sense of the word), and might be less affected by inhomogeneities.

We had not come across L-moments in our reading, so thank you for bringing these to our attention. We had noted the use of Legendre polynomials as basis functions in McKinnon et al (2016) and so were aware that other methods of calculating the changes in shape of a distribution exist (we note this in the text). They are something that we will investigate for future analyses of this type. However, for this study our goals were to be able to easily compare with and build on previous work using the ordinary moments, as used in Donat et al (2012), Cavanaugh & Shen (2014), which we think are also easier for readers without deeper statistical knowledge to understand. We have added the following paragraph at this point in the text to draw attention to the fact that other moments exist, and our reasons for sticking with the ordinary set:

Other statistical moments, e.g. linear (or L-) moments (Hosking, 1990) have also been used for the analysis of changes in the characteristics of distributions (e.g. Fowler et al, 2005, Simolo et al, 2011). However, as noted above, we use the ordinary statistical moments to enable clearer comparison with previous, similar analyses (e.g. Donat et al, 2012).

115 – It is unclear, distribution of what? Is it reasonable, for example, to consider all anomalies (relative to the annual cycle) for the entire year to be part of the same distribution? As an example, the processes that produce variability in summer in the northern mid-latitudes are substantially different from those that do so in winter, so it is not clear, really, what the distribution represents. One might also ask the same question about anomalies that are pooled across latitude bands to produce distributions (should one lump anomalies from different climate types together, and still call it a "distribution"?).

As outlined at the beginning of this section, our aim was to expand upon the study from Donat et al, (2012), who used only 3 zonal bands to compare between two 30-year periods. In this section of the study, we do de-seasonalise the data using a climatology, so that the tails of the distribution are not dominated by e.g. summer and winter for the temperatures in the mid-to-high latitudes. But we agree that the underlying processes are going to be different across the year, and across different longitudes within a zonal band as well. We have added "on an annual basis" at the end of the first sentence of this section, and "at the expense of spatial resolution" at the end of the second sentence.

The limitations of the zonally averaged analysis are addressed in the summary of this section (3.4), which then leads on to the following analysis (section 4) where observations from each station are assessed individually, seasonally and at different times throughout the day. We have expanded the final sentence of 3.4 (was line 199-201) to say the following:
"This analysis combines station anomalies together in zonal bands, separately for local day- and night-time. In doing so, it removes any small scale, regional changes in the characteristics of the distributions of the meteorological variables. We have also not split the analysis up into seasons (neither 3-month or wet/dry), and so large changes occurring in only part of the year will have been

diluted in this analysis. In the following section we improve the assessment of regional and seasonal changes but in doing so reduce the number of observations available to characterise the distributions."

We have added "of these zonally averaged quantities" at the beginning of the 4th paragraph of Section 3 for clarity.

118 – A small quibble here; in statistics, a hat (circumflex) is often added to Greek letters denoting distribution parameters if those parameters have been estimated, as is the case here.

Added

Figure 2 – A small editorial comment is that yellow curves often almost disappear into the background (many people find them hard to see). Another colour scale that doesn't fade away to light colours as you approach the present would be helpful.

This figure has used the "Viridis" colourmap which has been designed for readers who are colourblind, is perceptively uniform and should also work in black and white. Most sequential colourmaps which work well for those with colourblindness use intensity as well as hue, and one end is most often rather light. Of those available as standard in Python, Viridis is one of the darkest at the "light" end (https://matplotlib.org/tutorials/colors/colormaps.html). Other (diverging or qualitative) colourmaps are not appropriate given these plots show a sequence of 5-year values. However, we have increased the line width in the updated plots to assist the clarity.

161 – What aspect of wind speed does the paper consider? Are these hourly mean values?

The values in the HadISD are for example 1 or 2 minute average values over the US, as noted by DeGaetano, (1997) , or 10 minute averages in the UK. in relation to parent datasets to the ISD. We've added to this sentence and include the reference.

163-164 – I think many would dispute that temperature is "normally" (Gaussian) distributed. Indeed, if temperature did have a normal distribution then the study of skewness and kurtosis would be entirely uninteresting, since the normal distribution is fully determined by its first two moments.

Thank you - we have removed this sentence.

167 – Confidence bounds? I'm fine with simply describing +/- one standard deviation as an uncertainty range without a quantified level of confidence, but as soon as you call this a confidence bound, a question arises about the confidence level.

We have replaced this with "an uncertainty range"

178 – How does seasonality play into these findings about changes in the shape of the temperature distribution? Is there a physical or sampling interpretation to the change in kurtosis that is noted? Could this be due to undetected data artefacts (such as changes in instrumentation over time).

As noted in your comment regarding line 115, the analysis in this section is done only for annual data, rather than being split into seasons, and we have now added a mention and discussion of this to the beginning and end of this section respectively. Also, a climatology has been subtracted from the observations to remove the impact of seasonal effects on these results.

Given the small number of stations in the tropical latitude bands, and hence fewer observations to constrain the distribution at in each 5-year interval, we do not want to speculate on possible causes of the kurtosis change, also given the high-order of this moment. We have referred to these changes in kurtosis in Section 4, as these negative changes are concentrated in south-eastern Asia, but have no definitive explanation. As now noted further on in the manuscript, changes in circulation patterns and modes of variability, cloudiness and land-atmosphere coupling could all affect these distributions. More detailed analysis in a regionally focussed paper would be needed to determine the likeliest underlying causes.

Despite the quality control and homogeneity assessment of the HadISD, there will still be undetected data artefacts present which, for small station numbers may have a greater impact. We raise this along with the discussion on resolution.

Figure 3 (and others) – the labelling of the figures could be improved throughout the paper and the supplement. In this case, the key piece of information that tells the reader what is in each panel (mean, stdev, etc), is buried in in the middle of a small-font machine generated label.

We have increased the fontsize, improved the titles and moved some of the other text to help clarify what is shown on these scatter plots.

Figure 3 (and others) – the decision to show only stations with +/- one standard deviation uncertainty ranges that exclude zero seems arbitrary to me. My general preference is not to censor results in this way, since changes in the underlying field are likely relatively smooth and continuous. Imposing a noisy mask (which corresponds to conducting a test with unknown properties) could be argued to be at least as deleterious to describing changes in the dataset as including all of the effects of internal variability and sampling noise at all locations. The locations that are retained are still affected by these uncertainties.

Earlier versions of this analysis did include all the stations, with larger plotted symbols for those where the +/- 1 standard deviation excluded zero. However the result of discussions in the author team and wider, those stations where the +/- range included zero were removed, resulting in the current figures, our thinking at the time being that the trends were less reliable even if they were large in magnitude.

In light of your comments our previous discussions, we have reverted to showing all the stations on these plots. However, we do emphasise those where the range excludes zero as there is greater confidence for the trends in these stations to not be artefacts resulting from (a) the relatively small number of temporal bins being fitted, and (b) the use of a linear model for the change over time.

As a result, all the figures for Section 4 have been updated, along with those in the supplement.

235 – See previous comment.

This has been updated in light of the new figures prepared in response to the previous comment

237-238 – Why is US station density apparently so low?

This is the result of the station selection driven from the homogeneity assessment. As erroneously (and now corrected) we had used the number of inhomogeneities since the start of the HadISD record (1931) rather than 1974 and many of the US stations have long records, these were being excluded. This has been corrected, thank you for bringing this to our attention. As a result, all figures for Section 3 and 4 as well as the tables for Section 3 have been updated.

242 – Replace “Over 90%” with “Approximately 95%” since 2806/2956=0.949.

Done, with updated numbers in light of the correction from the previous comment

257 – “visible in more widely in” à “visible more widely in”

Done

272 – I don’t see how greater station density would increase trends (do stations emit heat?).

We have rephrased this sentence as follows "The region with the largest trends is Europe. However, it also has the highest station densities, and so the eye is more drawn to this region than areas which also show large trends but from a sparser station network."

286-295 – Here and elsewhere, it would be really useful to have more depth in the analysis of differences between datasets (or analyses). There is merit in pointing out differences, of course, but it would be much more useful to those assessing datasets if the authors could delve into the causes of these differences.

In light of comments from reviewer 2, we have re-ordered the paragraphs in the temperature part of Section 4, adding in a discussion sub-heading. We have there expanded the paragraph which outlines the differences between this and the other studies. As well as having its own temporal coverage, station selection, and processing, HadISD is sub-daily and so can show the changes over the course of the day rather than just for the (un-timed) point extremal values.

297-303 – What fraction of stations are affected by expanding urban heat islands, which can induce apparent trends when formerly rural stations come to be affected by urban heat islands as nearby cities develop? This is a huge problem in China, for example. China has a very dense observing network, but it was not built for climate purposes, and only a small fraction of Chinese stations can be classified as being rural throughout their observing history (this is particularly the case in Eastern China). It has been estimated that this difficult-to-detect cause of inhomogeneity has resulted in recorded temperature in China warming substantially more than the country has actually warmed (despite intense development, urban areas still represent only a small fraction of the Chinese land area). See Sun et al., 2016, Nature Climate Change. Other parts of the developing world are, presumably, similarly affected.

Thank you for raising this point. As noted in our introduction, we are interested in the experience of people, environment and infrastructure of extreme events and how these have changed, so the urbanisation effect is part of this. A discussion of this and how it can influence the results is relevant to this work, and we have added a paragraph to the end of section 4.1.1.

328-330 – It would be really useful if the authors could dig more deeply and provide more than speculation on the causes of the differences that are noted.

This paragraph has been expanded in the new subsection discussing these temperature results and also includes the important difference that HadISD is a sub-daily dataset, rather than just a max/min. Hence the using afternoon/morning temperatures as proxies for max/min will not always be appropriate.

337-338 – Is this decrease in relative humidity confirmed in HadISDH?

It is (see Willett et al, 2014, Willett et al, 2019), but as HadISDH is based on HadISD (but undergoes extra homogenisation and gridding) this is comparatively unsurprising. We have added "This decrease in the relative humidity has been observed regionally and globally in the homogenised HadISDH dataset (Willett et al, 2014, Willett et al, 2019), which is based on the HadISD." at this point for completeness.

392 – See a previous comment pointing out that changes in the tails are not necessarily easily inferred from changes in the full distribution.

We have replaced "inferred" with "calculated through, e.g. extreme value theory" to link back to the discussion in the introduction.

395 – Replace "calculates" with "can be used to calculate". Implicitly, the statement here describes quantile regression as calculating linear trends. The method, is in fact, a lot more flexible than that (other trend models can also be fitted, if appropriate).

Done

408 – As mentioned previously, I think the question of measurement (or perhaps better, data recording) resolution should have been introduced much sooner. Also, the statement that temperatures are reported to the nearest 0.1°C or 1°C does not really convey the full complexity of the problem (e.g., associated with conversion of °F to °C, and the subsequent rounding or truncation to increments of 0.1°C or 1°C). See Rhines et al, 2015.

We agree that there are deeper issues than just the apparent reporting to the nearest 0.1, 0.5 or 1 degree. However, we feel that the quantile regression section is not the right place to raise these issues, as the effects of using discrete data are separate to the effect that could occur in the quantile regression algorithm if de-seasonalisation wasn't done. As noted above, we have added extra sections to raise this issue elsewhere in the manuscript.

421 – How many stations in a polygon?

One. We have added a clause to highlight this

435 – Define N.

In this case N is meant to represent whatever upper percentile the reader wishes to insert (1, 5, 10, etc) rather than a denoting a specific quantity. For clarity we have added "where N=1, 5, 10 etc"

451 – This sentence seems to be grammatically challenged.

Thank you. The Franzke 2015 reference at the beginning has been removed as this seems to have been an erroneous paste. It now reads "The mean temperatures show strong regions of increase in both the 5th and 95th percentiles in eastern Europe and western Russia, especially north of the

Black Sea, with the higher percentile region covering a larger area than for the results presented here"

500-501 – Have there been changes in instrumental design or the processing of instrumental readings (e.g., approaches that might have been applied to compensate for variations in cup anemometer drag with velocity) that might have contributed to the apparent "stilling"?

We have added a sentence at this point outlining a number of the possible explanations that have been put forward for the stilling, including issues with wind sensors. Azorin-Molina et al, (2018) show that older instruments measure lower wind speeds than new ones in a comparative field trial but more investigations are needed before being able to confirm whether this is a (partial) explanation for the stilling.

519 – I'm not sure that spatial smoothing would result in "less extreme fields". The idea that the magnitude of extremes could be reduced is reasonable, but the idea that the result could be "less extreme" seems less well founded. "Extremeness", the relative position of an observation in the tail of its distribution, can only be evaluated within the context of the variability that a given data product tries to represent. Saying that the magnitudes of point rainfall observations can be larger than grid box mean rainfall amounts doesn't help one decide whether a given grid box mean value is more "extreme" (situated deeper in the tail) than a given point observation.

Thank you for pointing out this interpretation of our sentence. We had intended to convey that the magnitude of the extremes could be reduced by spatial smoothing compared to an unsmoothed representation, and agree that the relative "extremeness" of the tails of the smoothed fields is not affected. We have rephrased this as "which could result in less variable fields, with smaller extremes"

This manuscript studied the changes in statistical distributions of sub-daily surface temperatures, dewpoint temperatures, as well as wind speeds, using station-based HadISD dataset. Both zonally averaged quantities and the spatial distributions were considered, and a quantile regression analysis was also performed. Besides the changes of the mean values, different statistical moments were also studied. This work provided great details about the changes of the temperatures and wind speed, in context of global warming. Roughly speaking, I think this manuscript can be a good reference for people who studies the effects of global warming. However, to publish this work in ESD, there are several issues that need to be addressed.

Thank you. We have addressed your issues individually below, and marked in the revised manuscript where we have made changes, except for updated tables and figures, or if paragraphs have been moved into a new section.

1, It is difficult to catch the highlights of this work. Many calculations have been done in this work, but by reading the manuscript, it is very easy to get lost. I would suggest the authors to make a better discussion, and the conclusion should be improved.

In light of comments by the other review we have added some extra paragraphs in the discussion/summary sections for each analysis. Given the length of the sections addressing temperature for the station plots, we have reordered these paragraphs, adding in an extra discussion on the temperatures, including references possible causes for the changes observed, before moving on to the other two variables. The summary of Section 4 now also includes aspects of the observational data and how these could have affected the results. Hopefully by including some of the figures from the supplement, as you suggest below, this also helps readers.

We have made some changes to the final summary to highlight some of the other changes further up in the manuscript, and tried to clarify this section.

2, There are many figures in the supplementary materials. But the main text discussed these figures frequently. It seems that the figures are important. Therefore, why not include these figures in the main text? Or maybe the structure of the manuscript needs to be improved. Moreover, for the figures in the supplementary materials, I would suggest the authors use “Fig. S1, S2, etc.”, to distinguish from the figures in the main text.

We were attempting to strike a balance of the number of figures in the manuscript, and not have too many to dominate the text. As we discuss up to three variables, annually and in some cases seasonally, and with a sub-daily dataset are able to split across the day as well for up to 4 moments too, we didn't want to overload the manuscript with figures.

In light of the above comment we have added a number from the supplement into the main body of the paper, but have also retained them in the supplement so that this still has a logical flow to assist readers. This increases the duplication, but we feel this is appropriate to assist readers using the supplement.

We have also updated the numbering as suggested.

3, When studying the changes, what is the statistical significance level? What method was used to do the significance test? Why use 1σ as the threshold?

In this analysis, except for the quantile regression section, we do not assess the statistical significance level of any changes. We had used the +/-1σ range to determine how reliable a trend is by whether this range includes or excludes zero, and only plot those on the maps. In light of the comments by Referee 1, and as noted in our response to their comments about Figure 3 we have reverted to showing all stations in these scatter plots. We emphasise those where the +/-1σ range excludes zero with a larger symbol as the trends are more likely to be reliable but all the stations are now plotted.

Our aim with this approach is to balance occasions where there is a large trend magnitude but also a large spread in the possible trend values from the median of pairwise slopes algorithm. There are many studies which discuss the advantages and disadvantages of any form of significance testing, and the care required to frame both the problem and the test in the correct way (e.g. Ambaum, M.H., 2010: Significance Tests in Climate Science. J. Climate, 23, 5927–5932, https://doi.org/10.1175/2010JCLI3746.1, Wilks, D.S., 2016: "The Stippling Shows Statistically Significant Grid Points": How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It. Bull. Amer. Meteor. Soc., 97, 2263–2273, https://doi.org/10.1175/BAMS-D-15-00267.1, Ziliak, S. and McCloskey, D.N., 2008. The cult of statistical significance: How the standard error costs us jobs, justice, and lives. University of Michigan Press.).

We chose 1σ as this is widely used as the uncertainty of an estimated value, and use this as a way to indicate how more or less likely it is that a trend is different to zero rather than the magnitude of the trend itself. However we are actually not expecting the trends to be zero in many cases and so this has framed how we see the problem. Therefore we do not wish to add formal significance testing to this assessment.

In this study we are (a) fitting a trend to a relatively small number of points, and (b) using a linear trend to summarise changes over the 45 years of the study. We do not expect any changes to be linear, and merely use this as a way of simply quantifying changes over time, as has been done in many other studies of the past climate. Furthermore, the small number of points is the result of balancing sufficient observations per temporal bin to accurately determine the properties of a distribution with a large enough number of bins.

4, Since only data over the past 45 years were analyzed. Are the observed changes influenced by potential decadal variabilities in the climate system? Can the statistical significance test rule out the potential influences from the decadal variabilities?

As noted in the response to (3), we do not use a statistical significance test in this analysis except for the section on quantile regression. We also note that we do not expect any long-term changes to be linear, but use a linear trend to summarise these changes over time in a clear and simple way. We are in effect limited to a maximum of 46 years of data given the drop off in stations available in the HadISD prior to 1973, as discussed in Section 2. So it is quite possible that decadal variabilities could be one of the drivers for these changes in shape, and the length of data we have available may not be long enough to disentangle these effects. Our aim was to investigate what the changes in the distributions were, and point to possible causes rather than determine the most likely cause.

We have in response to comments from the reviewer 1 added discussion into the possible causes of changes in these distributions, and mention decadal variabilities there.

5, The results from this work were compared with the findings from previous studies. When the results are not in line with each other, which results are more reliable? Why? The authors may need to better explain why the results are different.

It is very difficult in this comparison to determine which study has the most reliable results. The limitations with the study done here arise mostly from the amount of data available, e.g. once sub-daily records are split into hourly anomalies and then combined in 5-year periods (Section 4). However we do not have the ability to investigate the details of other studies to pull out where differences lie - which may be from the trend fitting, the data preparation, the time periods covered, the underlying stations etc. We merely explain where our results differ from others and offer possible explanations but cannot say the exact cause or which to take.

In our re-ordering of Section 4 (from your 1st comment) we hope that we've clarified our analyses and have also added sentences outlining why we feel that we cannot say which assessment is best, especially as our opinion could come across as biased.

Changes in statistical distributions of sub-daily surface temperatures and wind speed

Robert J. H. Dunn[1], Kate M. Willett[1], and David E. Parker[1]

[1]Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, UK

Correspondence: robert.dunn@metoffice.gov.uk

Abstract. With the ongoing warming of the globe, it is important to quantify changes in the recent behaviour of extreme events given their impacts on human health, infrastructure and the natural environment. We use the sub-daily, multi-variate, station-based HadISD dataset to study the changes in the statistical distributions of temperature, dewpoint temperature and wind speeds. Firstly, we use zonally averaged quantities to show that the lowest temperatures during both day and night are changing more rapidly than the highest, with the effect more pronounced in the northern high latitudes. Along with ~~an~~ increases in the zonally-averaged mean temperature, the standard deviation has decreased and the skew increased (increasing positive tail/decreasing negative tail) over the last 45 years, again with a stronger, more robust signal at higher latitudes. Changes in the distribution of dewpoint temperature are similar to those of temperature. However, changes in the distribution of wind speeds indicate a more rapid change at higher speeds than at lower.

Secondly, to assess in more detail the spatial distribution of changes as well as across seasons and hours of the day we study each station individually. For stations which show clear indications of change in the statistical moments, the higher the statistical moment, generally the more spatially heterogenous the patterns of change. The standard deviations of temperatures are increasing in a band stretching from Europe through to China, but are decreasing across North America and in the high northern latitudes, indicating broadening and narrowing of the distributions respectively. Large seasonal differences are found in the change of standard deviations of temperatures over North America and eastern China. Temperatures in Eastern Asia also have increasing skew in the winter in contrast to the remainder of the year. The dewpoint temperatures show smaller variation in all of the moments, but similar patterns to the temperatures. For wind speeds, apart from the USA, standard deviations are decreasing across the world, indicating a decrease in variability.

Finally, we use quantile regression to show changes in the percentiles of distributions over time. These show an increase of high quantiles of temperature in eastern Europe during the summer, and also in northern Europe for low quantiles in the winter, also indicating broadening and narrowing of the distributions respectively. In North America, the largest changes are at the lower quantiles in northern latitudes for autumn and winter. Quantiles of dewpoint temperature are changing most in the autumn and winter, especially in the northern parts of Europe.

1 Introduction

The study of changes in the extremes of essential climate variables is vital given their impacts on human health, infrastructure, agriculture and the natural environment. The Intergovernmental Panel on Climate Change (IPCC) released a Special Report on Extremes (SREX, Field et al., 2012), and therein outlined three simple classes of the way changes in extremes occur; shifting the mean, increasing the variability, and changing the symmetry. Past studies have indicated that there is uncertainty around how changes in the occurrence or intensity of extreme values arise. Are extremes changing due to changes in the location of the distribution mean with no change in the distribution shape (Griffiths et al., 2005; Simolo et al., 2011), or are changes in the shape of a distribution the primary driver of changes in extremes? If so, is the change in shape only the consequence of a change in the variance or does it also arise from changes in higher order moments (Della-Marta et al., 2007; Ballester et al., 2010)?

There are a whole host of ways to study climate extremes and determine how these have changed over the recent past. One common approach is to use climate extremes indices to characterise moderate extremes in timeseries of station data, for example those developed by the World Meteorological Organisation (WMO) Commission for Climatology (CCl), World Climate Research Programme (WCRP), and Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM) Expert Team in Climate Change Detection and Indices (ETCCDI, Alexander et al., 2006). A number of datasets comprising these indices now exist, allowing for detailed investigations of past changes and comparisons to model and reanalysis fields (e.g. Caesar et al., 2006; Donat et al., 2013a, b). Another route uses extreme value theory to model the tails of the distributions, and from these few points characterise the occurrence and intensity of extreme events, including those not yet observed in the modern data record. The "block maximum" approach models a set of e.g. annual, maxima with a Generalised Extreme Value distribution (e.g. Christidis et al., 2011). Alternatively, the "peaks over threshold" approach models all peaks over a fixed threshold with the Generalised Pareto distribution(e.g. Brown et al., 2008).

A further method characterises the complete distribution using all available data to establish what are the causes of changes in climate extremes. The advantage of this approach is that it uses all of the available information, and can ~~easily~~ be extended to more or less extremal parts of the distribution. A number of the ETCCDI temperature indices use the 10^{th} and 90^{th} percentile values, which only probe moderate extremes. A recalculation of these indices at 1^{st} and 99^{th} percentile values is possible, but given the diverse sources of data in the HadEX datasets (Donat et al., 2013b), this would be a large undertaking by the international community. Furthermore, as the climate continues to warm, the relative intensity or duration of extreme events against warmer average conditions may be of interest; for example, how warm is the relatively warmest 10 per cent of days now compared to the same fraction selected between 1961-90?

Hence, a number of recent studies have investigated past changes in the distributions of observed land-surface temperatures. Most have used the Global Historical Climate Network Daily (GHCND, Menne et al., 2012) or its gridded version (HadGHCND, Caesar et al., 2006, temperature only) as these provide daily maximum and minimum temperature values. Donat and Alexander (2012) compared two periods to show that the means of both temperature measures have shifted to warmer values. Along with larger changes in the minimum than the maximum temperatures, this has resulted in increases in the skew of the distributions. However changes in standard deviation were less significant (at 10%) and more heterogeneous. Over North America, Shen et al. (2011) show that the standard deviation and skewness were decreasing, but that the kurtoses of the maximum temperatures were increasing, in contrast to the decrease in the kurtoses of minimum temperatures. A subsequent study investigated seasonal changes in the distribution moments over 1950-2010 (Cavanaugh and Shen, 2014). McKinnon et al. (2016) showed that for most stations between 1980 and 2015, over just the summer months, trends can mostly be explained by a shift in the distribution with no change in shape. Reanalyses (e.g. Huntingford et al., 2013; Gross et al., 2018), or coupled models (e.g. Lewis and King, 2017) have also been used in the study of changes in the shapes of current and future distributions of essential climate variables (ECVs).

In this study, we use the sub-daily, station data from a single dataset, as opposed to the maximum and minimum daily temperatures (either station based or gridded). We note that by using a single dataset, we have not investigated the effect of dataset choice on our findings, something which can have a large effect (Gross et al., 2018). In order to compare with a number of previous studies we perform three different assessments to see how the distributions of temperature, dewpoint temperature, and wind speed have changed in recent years. We outline the dataset and station selection criteria in Section 2. The three assessments are outlined in Sections 3 to 5, with a summary in Section 6.

2 Input Data

To study the changes in distributions related to climate extremes, we use the sub-daily HadISD dataset (Dunn et al., 2012, 2014, 2016; Dunn, 2019). This is updated annually, and now covers a period 1931-2018 inclusive. Improved station selection and merging processes as well as minor changes in the quality control tests are features of version 3.0.0.2018f over version 1.0.x. This version contains 8139 unique station locations, with some of these being composited from individual records in the ISD (Integrated Surface Dataset, Smith et al., 2011). As a result of the long standing issues surrounding the free sharing of observational climate data (Thorne et al., 2017), the bulk of these stations are in the Northern Hemisphere, concentrated in North America and Europe. In contrast to some of the previous studies outlined in the Introduction (e.g. Donat and Alexander, 2012), we do not create a gridded form of the HadISD for this assessment, but retain the individual stations as input to the analyses, similar to e.g. Cavanaugh and Shen (2014); McKinnon et al. (2016).

HadISD undergoes a homogeneity assessment using the Pairwise Homogeneity Algorithm (PHA, Menne and Williams Jr, 2009) to identify the number and size of inhomogeneities in its four main ECVs, including temperature and wind speed (Dunn et al., 2014). This used monthly averages derived from the sub-daily HadISD observations to compare station-pair differences to identify potential jumps in timeseries, a process which is more effective for stations and variables with low variability. Although no adjustments are made, the outputs of the PHA allow stations that have few or small inhomogeneities to be selected, i.e. those with the most homogeneous records (see Dunn et al., 2014). They found that for temperature the best balance of station network against the exclusion of stations with the worst inhomogeneities occurred between 1° and 2° C. Therefore in all of the investigations below, we select those stations that have fewer than 5 ~~10~~ inhomogeneities, allowing jumps of up to 1° C (temperature and dewpoint) or 1 m s^{-1} (as the distribution of the inhomogeneities in the wind speed data are similar, see Figure 10 of Dunn et al., 2014). This balances retaining sufficient stations to study the behaviour across the globe with removing those with very large inhomogeneities.

As the selected stations still contain uncorrected inhomogeneities, we note that these will have affected the analysis of the statistical properties of the observations outlined herein. But the distributions of the estimated inhomogeneities as shown in Dunn et al. (2014, 2016) do not have a strongly non-zero mean, and the central portions retained by the approach above have no clear asymmetry. Therefore, as the following analyses use combinations of stations or look for contiguous regions of change that cross national (i.e. observing-practice) boundaries, we do not expect large effects arising from the remaining inhomogeneities themselves.

Although the HadISD v3.0.x contains data from 1931 onwards, there is a large increase in the number of stations from 1973 onwards (the reason why HadISD v1.0.x started then, see Fig. 2 in Dunn et al., 2016). Therefore we restrict our analysis to the time period from 1st January 1973 unless otherwise stated. We also remove data from February 29th for ease of analysis.

3 Zonal Distribution Changes

We firstly look at the changes in distributions of anomalies of the temperatures, dewpoint temperatures and wind speeds for all stations in a specified latitude band on an annual basis. By combining stations together, the quantity of observations available to define the distribution is increased at the expense of spatial resolution. Donat and Alexander (2012) used the HadGHCND gridded daily temperature dataset (Caesar et al., 2006) to compare two 30-year periods of daily maximum and minimum temperatures (1951-80 and 1981-2010) globally, and for three smaller zonal regions (Northern Hemisphere, Tropics and Southern Hemisphere). Here we develop this method further using the sub-daily data available in the HadISD as well as narrower 10-degree zonal bands.

We select those stations that fall within each 10 degree band. An approximation for the whole-hour timezone of the station is calculated from the station longitude, enabling the Universal Time (UTC) observation time to be converted to local time. We then extract those portions of data that correspond to daytime and night-time, using the time periods 09:00-20:00 and 21:00-08:00 respectively, to ensure that the minimum temperatures are on the whole captured during the night-time portion, and the maximum temperatures in the daytime. At high latitudes these definitions of day- and night-time will vary in accuracy throughout the year, but should split the 24 hour day into the hours with most and least insolation. There are few stations at the highest latitudes where 24-hours of day or night occur.

Climatologies for stations for each day of the year for the day- and night-time data are calculated separately, requiring the equivalent of 15 years of 3-hourly observations over the 1981-2010 period. These climatologies are used to create anomalies from the sub-daily data. The creation of climate anomalies also restricts this analysis to stations that have consistent records within this recent climatology period.

To investigate the changes in the distributions of these zonally averaged quantities over time, we split the data up into nine intervals each of five years in length (1974-78, 1979-83, .., 2014-2018). By not including 1973, we can assess changes over a 45 year period. Furthermore, we only take those stations that have sufficient observations spread over the entire analysis

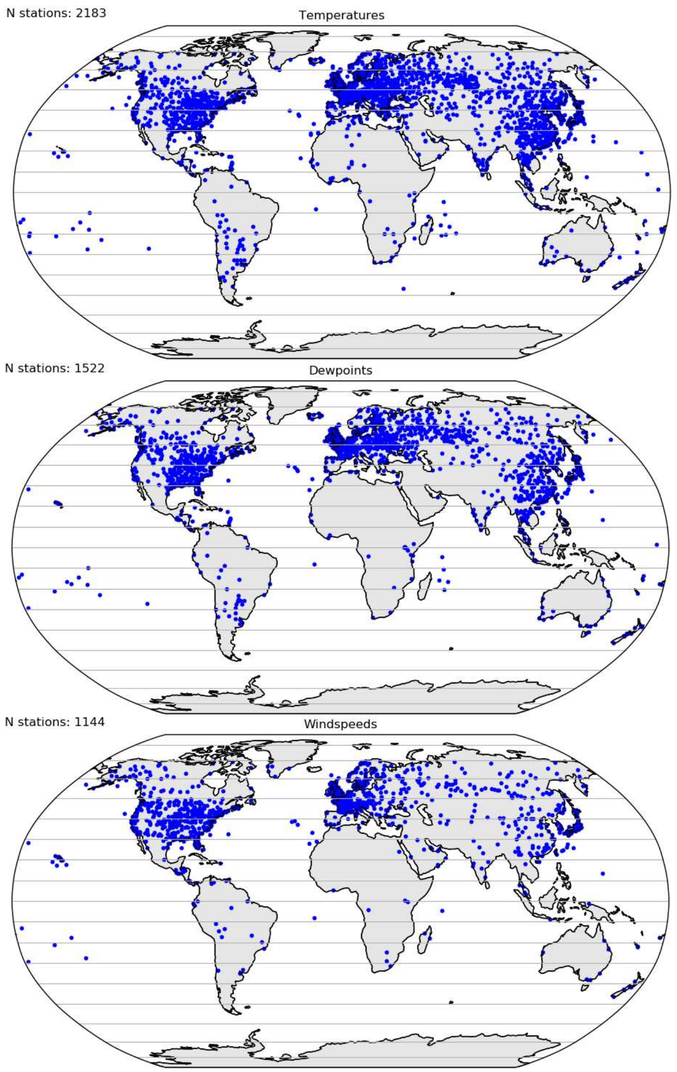

Figure 1. Maps showing the locations of HadISD stations contributing to the zonal analysis along with the 10 degree latitude bands for TOP temperature, MIDDLE dewpoints, and BOTTOM windspeeds.

period. In each interval, we require there to be more observations than the number equivalent to hourly observations for 1/4 of the 5-year interval. Also, at least 3/4 of the intervals require sufficient data for the station to be incorporated into the final distribution. This is to ensure that non-uniform distributions of observations within a station record (both within a five-year interval and across the entire record) do not overly influence the final distributions. The stations which pass these selection criteria are shown in Fig 1. A more technical summary of these processing and selection criteria is shown in the Technical Supplement.

We determine the values for the first four statistical moments (mean, standard deviation, skew and [excess] kurtosis) of the resulting distributions. These in turn represent the central tendency, the spread, the asymmetry and the "peaked-ness" of the distribution. Distributions with non-zero skew or kurtosis indicate departures from a pure Gaussian shape. We note that the skew and kurtosis measures are not fully orthogonal (McKinnon et al., 2016), but will use these measures in this analysis because of their well-understood nature. Other statistical moments, e.g. linear (or L-) moments (Hosking, 1990) have also been used for the analysis of changes in the characteristics of distributions (e.g. Fowler et al., 2005; Simolo et al., 2011). However, as noted above, we use the ordinary statistical moments to enable clearer comparison with previous, similar analyses (e.g. Donat and Alexander, 2012).

The linear change in these ordinary moments ~~values~~, along with $1\hat{\sigma}$ range have been derived from the median of pairwise slopes method (e.g. Sen, 1968; Lanzante, 1996), and are shown for each latitude band in Table 1 and Supplementary Information Table S1 & S2 for temperature, dewpoint temperature and wind speed respectively. We do not expect any changes in the parameters of the distributions to be linear, but it is a useful way to summarise their gross changes over time together with the relative certainty of these changes. Latitude bands where the range in trend values does not encompass zero are highlighted in bold, as for these we can be more confident that there is a true change in the value over time.

3.1 Temperatures

The results from day- and night-time temperature observations for the Northern Hemisphere are shown in Fig. 2 and Table 1. We can be more confident of changes that are seen in latitude bands containing a large number of stations than in those which only contain a handful. The Southern Hemisphere is shown in Supplementary Information Fig. S1 and Table S1, as it includes considerably fewer stations. It is clearly evident from these figures and tables that there has been a shift in the location of the distributions from the early period to the later period, with the most recent period being on average warmer than the earliest. This behaviour agrees with the results presented in Donat and Alexander (2012) who shows increases in the mean of around 1°C between the early (1951-1980) and late (1981-2010) periods in their study (for all their regions and for both maximum and minimum temperatures). We note that their period of study is longer (60 years) compared to the period presented in this work (45 years). It is also unsurprisingly in agreement with the widespread warming of the Earth's surface (Stocker, 2014).

In both day- and night-time observations, positive increases in the mean that are clearly different from zero are seen in all latitude bands (Table 1). The strongest increases are seen in the highest northern latitudes, with increases on the whole becoming smaller when heading towards the equator for both halves of the day. Changes in the southern hemisphere show less

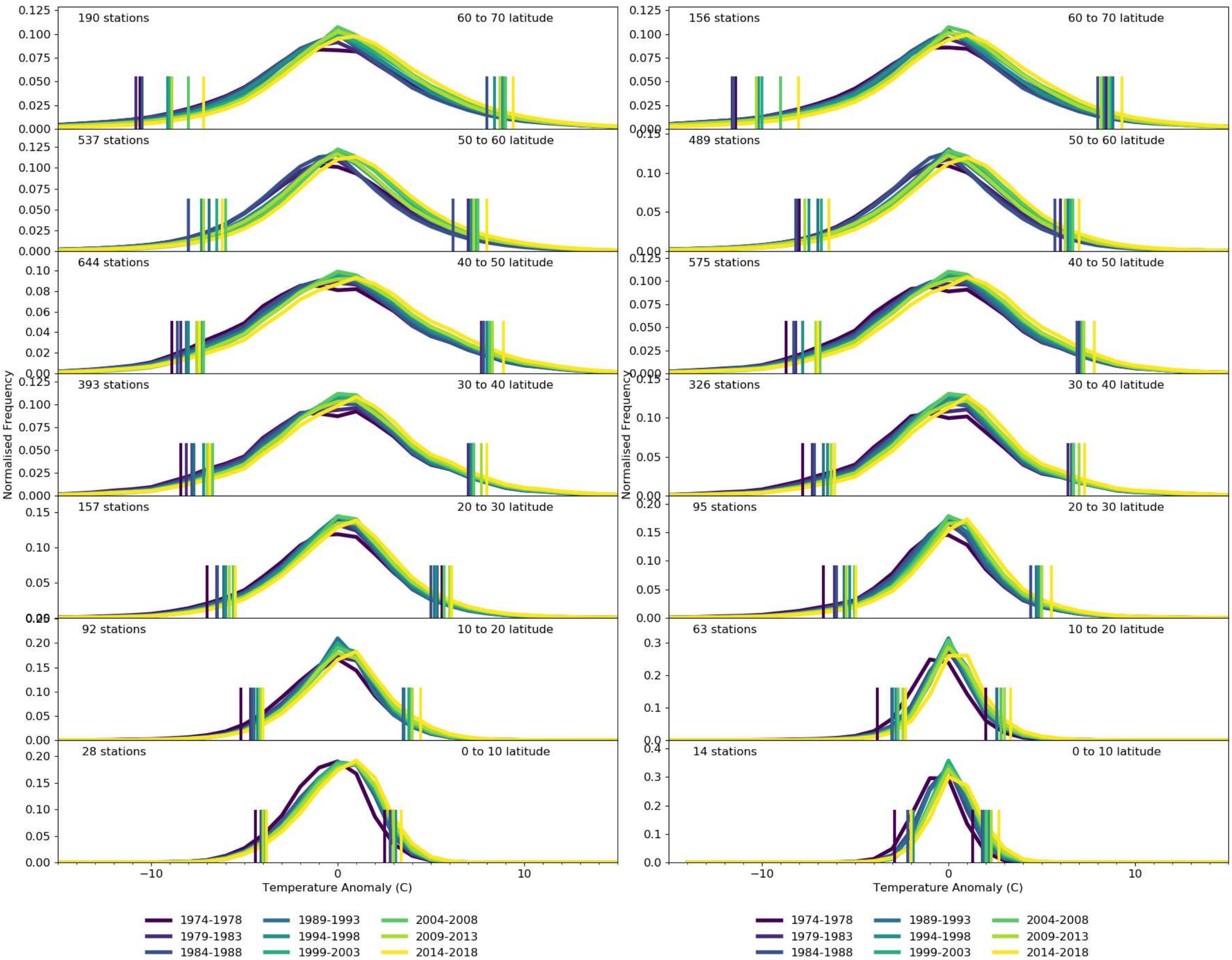

Figure 2. Dry bulb temperature distributions in latitudinal bands for LEFT day-, and RIGHT night-time observations for the Northern Hemisphere stations. The y-axis has been normalised for ease of comparison. The number of stations contributing to each set of distributions is shown in the top left of each sub-panel, with the latitude range in the top right. The short vertical bars indicate the location of the 5^{th} and 95^{th} percentile values for each distribution.

variation with latitude, and are on the whole smaller than in the north, though still showing overall increases. There are few differences in the day- and night-time increases for each latitude band, in most cases the ranges overlap.

The northern mid-to-high latitudes also see a consistent decrease in the standard deviations in both day and night time. However, south of 20°N, the signal is less consistent. In fact, during day- and night-time most tropical and southern-latitude bands have ranges that encompass zero, though most bands suggest decreases. Although Donat and Alexander (2012) also found decreases in standard deviation in the extra-tropics, between 30°S-30°N they found an increase when using zonal averages. This decrease suggests that the maximum temperatures are warming less quickly than the minimum temperatures, reducing the spread of the distribution in the northern hemisphere.

Consistent changes in the skew and kurtosis for both day- and night-time or across latitude bands are not as apparent in Table 1. Slight increases in skew (growth in the positive tail and shrinking of the negative tail) are seen for the daytime high latitudes (north and south). In contrast, during night-time, some tropical regions show decreases in skew. The kurtosis decreases in the tropics at both times of day, with a more mixed signal at the high latitudes. Donat and Alexander (2012) also found an increase in the skew in all regions except for $T_{\max}$ in the tropics, which is broadly consistent with the changes shown in Table 1.

3.2 Dewpoint Temperatures

The changes in the dewpoint temperatures (Supplementary Information Table S1 and Figs. S2-S3) show a slightly more complex picture than for the dry-bulb temperatures. In general, the mean again shows an increase in both day- and night-time values, with larger increases over time in the northern hemisphere high-latitudes, similar to the dry-bulb temperatures. In the southern hemisphere however, and especially at high latitudes, this increasing signal is less strong with few cases where the confidence intervals do not span zero. One southern latitude band even shows a clear decrease in mean dewpoint temperatures over the last 45 years. Also, trends close to the equator are greater than at latitudes immediately to the north and south of this equatorial belt.

The tendency has been for standard deviations in many northern latitudes to decrease, but there are also a number of increases, especially in the southern hemisphere, though determined from few stations. For the skew, in the daytime, these increase in the north and decrease in the south, with a similar, but not as consistent pattern in the night-time. The changes in kurtosis do not show any clear pattern in comparison to the other moments.

3.3 Wind speeds

We also assess changes in the sub-daily wind speeds (for example, 1 or 2 minute average values over the USA DeGaetano, 1997, or 10 minute average values over the UK), but there are fewer stations where this variable has been sufficiently observed (see Supplementary Information Table S2 and Fig. S4). In fact, all latitude bands in the Southern Hemisphere have fewer than ten stations, and so we do not study this hemisphere. We ~~also note that this variable is not normally distributed, unlike the temperatures investigated in previous sections. Despite this we~~ use the same method as for the temperatures but note that the anomalies will be truncated at the negative end by the zero-bounded nature of this variable.

Band	N-stations	Mean (°C decade^{-1})	St. Dev. (°C decade^{-1})	Skew (decade^{-1})	Kurtosis (decade^{-1})
			Day		
$70 \to 60$	190	$\mathbf{0.40\ (0.36 \to 0.49)}$	$\mathbf{-0.19\ (-0.22 \to -0.15)}$	$\mathbf{0.04\ (0.03 \to 0.06)}$	$0.06\ (-0.00 \to 0.14)$
$60 \to 50$	537	$\mathbf{0.33\ (0.29 \to 0.42)}$	$\mathbf{-0.09\ (-0.13 \to -0.04)}$	$\mathbf{0.03\ (0.02 \to 0.05)}$	$-0.04\ (-0.14 \to 0.04)$
$50 \to 40$	644	$\mathbf{0.33\ (0.32 \to 0.39)}$	$\mathbf{-0.06\ (-0.09 \to -0.03)}$	$-0.00\ (-0.01 \to 0.03)$	$-0.02\ (-0.09 \to 0.00)$
$40 \to 30$	393	$\mathbf{0.29\ (0.27 \to 0.30)}$	$\mathbf{-0.10\ (-0.12 \to -0.04)}$	$\mathbf{0.02\ (0.01 \to 0.04)}$	$\mathbf{0.05\ (0.03 \to 0.09)}$
$30 \to 20$	157	$\mathbf{0.24\ (0.23 \to 0.27)}$	$\mathbf{-0.05\ (-0.07 \to -0.01)}$	$\mathbf{0.03\ (0.02 \to 0.06)}$	$\mathbf{0.06\ (0.05 \to 0.09)}$
$20 \to 10$	92	$\mathbf{0.18\ (0.16 \to 0.22)}$	$-0.01\ (-0.03 \to 0.03)$	$\mathbf{0.01\ (0.00 \to 0.04)}$	$0.01\ (-0.02 \to 0.07)$
$10 \to 0$	28	$\mathbf{0.16\ (0.13 \to 0.19)}$	$0.00\ (-0.00 \to 0.01)$	$-0.01\ (-0.03 \to 0.01)$	$\mathbf{-0.10\ (-0.11 \to -0.07)}$
$0 \to -10$	12	$\mathbf{0.18\ (0.17 \to 0.24)}$	$-0.02\ (-0.02 \to 0.01)$	$\mathbf{0.04\ (0.03 \to 0.06)}$	$\mathbf{-0.15\ (-0.16 \to -0.13)}$
$-10 \to -20$	32	$\mathbf{0.15\ (0.13 \to 0.19)}$	$\mathbf{0.04\ (0.03 \to 0.07)}$	$0.01\ (-0.00 \to 0.04)$	$\mathbf{-0.17\ (-0.20 \to -0.06)}$
$-20 \to -30$	37	$\mathbf{0.12\ (0.11 \to 0.15)}$	$\mathbf{-0.08\ (-0.09 \to -0.07)}$	$\mathbf{0.03\ (0.01 \to 0.05)}$	$\mathbf{0.20\ (0.17 \to 0.22)}$
$-30 \to -40$	32	$\mathbf{0.21\ (0.18 \to 0.26)}$	$0.01\ (-0.03 \to 0.04)$	$\mathbf{0.07\ (0.06 \to 0.09)}$	$\mathbf{0.11\ (0.09 \to 0.16)}$
$-40 \to -50$	11	$\mathbf{0.09\ (0.07 \to 0.19)}$	$0.01\ (-0.02 \to 0.05)$	$\mathbf{0.05\ (0.04 \to 0.08)}$	$-0.03\ (-0.05 \to 0.01)$
			Night		
$70 \to 60$	156	$\mathbf{0.39\ (0.33 \to 0.48)}$	$\mathbf{-0.16\ (-0.20 \to -0.14)}$	$0.00\ (-0.01 \to 0.03)$	$0.04\ (-0.03 \to 0.12)$
$60 \to 50$	489	$\mathbf{0.30\ (0.25 \to 0.40)}$	$\mathbf{-0.07\ (-0.10 \to -0.03)}$	$\mathbf{0.03\ (0.01 \to 0.06)}$	$-0.10\ (-0.19 \to 0.00)$
$50 \to 40$	575	$\mathbf{0.32\ (0.30 \to 0.35)}$	$\mathbf{-0.10\ (-0.11 \to -0.07)}$	$0.01\ (-0.01 \to 0.04)$	$-0.01\ (-0.07 \to 0.04)$
$40 \to 30$	326	$\mathbf{0.29\ (0.27 \to 0.30)}$	$\mathbf{-0.11\ (-0.13 \to -0.04)}$	$\mathbf{0.05\ (0.03 \to 0.07)}$	$0.04\ (-0.01 \to 0.12)$
$30 \to 20$	95	$\mathbf{0.27\ (0.25 \to 0.29)}$	$\mathbf{-0.06\ (-0.08 \to -0.00)}$	$0.00\ (-0.01 \to 0.02)$	$\mathbf{0.11\ (0.09 \to 0.14)}$
$20 \to 10$	63	$\mathbf{0.19\ (0.19 \to 0.26)}$	$\mathbf{-0.05\ (-0.05 \to -0.02)}$	$\mathbf{0.07\ (0.06 \to 0.10)}$	$\mathbf{-0.29\ (-0.40 \to -0.20)}$
$10 \to 0$	14	$\mathbf{0.22\ (0.19 \to 0.24)}$	$0.00\ (-0.00 \to 0.03)$	$\mathbf{-0.07\ (-0.08 \to -0.05)}$	$\mathbf{-0.42\ (-0.49 \to -0.38)}$
$0 \to -10$	6	$\mathbf{0.21\ (0.16 \to 0.26)}$	$\mathbf{-0.07\ (-0.08 \to -0.03)}$	$\mathbf{0.02\ (0.00 \to 0.06)}$	$\mathbf{-0.55\ (-0.71 \to -0.32)}$
$-10 \to -20$	8	$\mathbf{0.24\ (0.20 \to 0.28)}$	$-0.02\ (-0.03 \to 0.02)$	$\mathbf{-0.10\ (-0.12 \to -0.01)}$	$\mathbf{0.21\ (0.09 \to 0.32)}$
$-20 \to -30$	21	$\mathbf{0.17\ (0.13 \to 0.20)}$	$\mathbf{-0.05\ (-0.07 \to -0.02)}$	$\mathbf{0.02\ (0.01 \to 0.05)}$	$\mathbf{0.10\ (0.05 \to 0.12)}$
$-30 \to -40$	24	$\mathbf{0.22\ (0.21 \to 0.23)}$	$\mathbf{-0.07\ (-0.08 \to -0.01)}$	$\mathbf{0.04\ (0.04 \to 0.06)}$	$\mathbf{0.14\ (0.11 \to 0.18)}$
$-40 \to -50$	11	$\mathbf{0.09\ (0.06 \to 0.16)}$	$\mathbf{0.01\ (0.00 \to 0.05)}$	$\mathbf{0.01\ (0.00 \to 0.03)}$	$\mathbf{-0.08\ (-0.11 \to -0.02)}$

Table 1. Linear change per decade in fits to parameters in zonal analysis for TOP day-, and BOTTOM night-time temperature observations over 1974-2018. Values in bold show parameters and bands where the 1σ range of the fitted trend does not include zero.

For the northern hemisphere, most latitude bands show reductions in the mean and also the standard deviation over time for both day- and night-time observations, with an uncertainty range ~~confidence bounds~~ distinct from zero. In the lower mid-latitudes, there are indications for a small increase in standard deviations, but at a lower magnitude than other changes. Tropical regions show decreases in the skewness over day- and night-time, but high latitudes an increase during the day.

These decreases in the mean wind speed over land have been noted in past studies (termed “stilling”, Roderick et al., 2007; McVicar et al., 2012), and monitoring studies have also shown decreases over time for the high and low wind speed values (e.g. Azorin-Molina et al., 2019). However, these recent regular global assessments also use the HadISD and so differences are not expected. A number of possible explanations have been proposed to explain the stilling, which may be reversing in recent years (Azorin-Molina et al., 2019), including amongst others changes in the surface roughness because of land use changes (Vautard et al., 2010), variability of the atmospheric circulation (Azorin-Molina et al., 2016), and instrumental issues with wind sensors (Azorin-Molina et al., 2017, 2018).

3.4 Zonal Summary

As can be seen in both Fig. 2 and Table 1 (as well as Fig. S1 in the Supplementary Information) there is a clear increase ~~change~~ in the mean temperatures with time across all latitude bands and times of day. However, this is also combined with a decrease in the standard deviation over the northern hemisphere. The width of the distribution of temperatures is therefore narrowing. Also, the skew of the distribution is becoming more positive, which indicates that the positive tail is growing and the negative tail is shrinking. Together, these three changes indicate that there has been a greater change in the low tail than in the high tail of the temperature distribution. The changes in the kurtosis are mainly increases (the tropical decreases rely on small samples), indicating a change to more peaked distributions, and a reduction in both tails. Although Donat and Alexander (2012) found similar increases in the mean and skew of daily maxima and minima, they showed increases in the standard deviation for tropical regions not found here.

Markers indicating the location of the 5^{th} and 95^{th} percentile values in Fig. 2 track the changing tails of the distributions. Qualitatively for the northern hemisphere, the low tails have changed more than the high tails, as the vertical lines cover a greater spread, and more so for the higher latitudes than the tropical regions. These highlight changes in the standard deviation and also the skew on top of changes in the mean. However, there is no striking difference between the day- and night-time plots for temperature. The number of stations in the southern hemisphere is much smaller, but there are no systematic differences between the behaviours of the low and high tails (see Supplementary Information). A similar pattern is seen for the dewpoint temperatures, with a greater shift for the lower tail, and no striking difference between night- and daytime values. The wind speeds show little change in the lower tails (Supplementary Information Fig. S4), but decrease in the value of the 95^{th} percentile over time. Further investigations into the changes in the 5^{th} and 95^{th} percentile values are presented in Section 5.

The intention of investigating the day and night time values separately was to echo analyses conducted in other studies of daily maximum and minimum temperatures. Many of these show that the minimum temperatures are changing faster than the maxima, and so we expected to find a difference between day and night time, either for the parameters of the distributions or in the tail markers. However, the clearest difference remains between the upper and lower tails. This analysis shows increases in the mean and skew, and decreases in the northern hemisphere standard deviations. These changes in the statistical moments of the temperature distributions indicate that the lower tails are shifting more rapidly than the upper tails. Hence, as well as a change in the location of the distribution, the shape has also changed, resulting in changes to occurrence of extremes. Although fewer stations contribute to the wind speed observations, these also show changes in location and shape, with larger changes in the upper tails caused by decreases in the mean and standard deviations. ~~However, this approach removes any small scale, regional changes in these and the other two characteristics and by focusing on the mome~~

Increases in the mean for the temperature and dewpoint temperatures are expected under a warming world (Stocker, 2014), and so it is the coherent changes in the higher moments which are of greater interest in this analysis. A greater change in the minima than the maxima can drive both the decrease in standard deviation and also the increases in the skew, but interestingly we see no difference between the magnitude of night-time changes compared to daytime ones. Some studies using climate models suggest that changes in extremes in a warming world (from anthropogenic forcings) come from changes in the shape as well as location of the temperature distributions (Clark et al., 2006; Hegerl et al., 2004). More recent work has looked at the effect of moisture on extreme temperatures, in the soil (Whan et al., 2015) or the atmosphere (Sherwood and Huber, 2010; Matthews, 2018), where the separation into latent and sensible heating when moisture is present can reduce the peak temperatures.

We note that observations in HadISD have a finite numerical resolution, arising from the point of recording, through transcription, rounding, conversion and truncation, and in cases combinations and multiples of these. As in this part of our analysis we blend anomalies calculated using daily climatologies from different stations together, and assuming a random distribution of these effects within a latitude band we do not discuss the impact of all these processes at this point, but we do so in the next section (Section 4).

This analysis combines station anomalies together in zonal bands, separately for local day- and night-time. In doing so, it removes any small scale, regional changes in the characteristics of the distributions of the meteorological variables. We have also not split the analysis up into seasons (neither 3-month or wet/dry), and so large changes occurring in only part of the year will have been diluted in this analysis. In the following section we improve the assessment of regional and seasonal changes but in doing so reduce the number of observations available to characterise the distributions.

4 Station Distribution Changes

To investigate the geographical patterns of changes in these meteorological parameters in more detail than in Section 3, we study each station individually. This reduces the amount of data available to characterise the distribution at the gain of increased spatial information. A similar approach was used by e.g. Cavanaugh and Shen (2014); McKinnon et al. (2016) in their investigations into the changing shape of temperature distributions on a daily basis from GHCND. However, the sub-daily nature of HadISD means we are able to investigate the distributions at different times throughout the day, rather than the changes of the maximum and minimum values independently. A large proportion of the stations in the HadISD only report every three hours, rather than hourly. Therefore we use the station longitude to adjust from UTC to local time, collating to the nearest 3-hourly time point (00:00, 03:00, ...).

For each station a climatology over 1981-2010 is calculated for each day of the year for each 3-hourly time point ($365 \times (24/3)$), requiring that at least 15 years of data are present. These 3-hourly climatologies are subtracted from the observations to create 3-hourly anomalies. The analysis is performed using observations over the entire year and also for standard three-month seasons.

The 45 years (1974-2018 inclusive) are split into five 9-year periods, and the anomalies within each period and at each three-hourly time point (0000, 0300, ... etc.), are combined, annually as well as seasonally. Using fewer periods of longer length than in Section 3 allows for better characterisation of the distributions at the individual station level, at the expense of the number of periods. This is in contrast to the previous analysis (Section 3) where combining stations together increases the number of observations contributing to a distribution. For the three-month seasons, using periods nine years in length means there are maximally around 800 observations from which the distribution at a single time point can be characterised, if all observations are present (3 months $\times \sim$ 30 days $\times$ 9 years). The first four moments of the distributions are only calculated if there are more than 300 observations present within a 9-year period (roughly the equivalent of 1 months worth of observations at a given time point). A more technical summary of these processing and selection criteria is shown in the Technical Supplement.

We again use the median of pairwise slopes method (Sen, 1968; Lanzante, 1996) to characterise the change in the moments over time, but only if there are sufficient data in at least four of the five periods. Again, we do not expect any changes to be linear, but this enables the convenient visual inspection of changes over time. Even if our data completeness criteria result in the inclusion of stations which have a non-uniform distribution of observations across the years and seasons, coherent, robust changes (if they are present) will still stand out across the stations.

4.1 Temperature

In Fig. 3 we show changes in the mean, standard deviation, skew and kurtosis for the temperature data for all seasons from HadISD at 15:00 and at 06:00 local time in Fig. 4, to correspond approximately with the times of maximum/daytime and minimum/nighttime temperature respectively. Other times are available in the Supplementary Material. Despite 8139 stations being available in the HadISD v3.0.0.2018f dataset, after selection for homogeneity and length of record, only around 3200 have quantified changes in distributional parameters, the total ~~actual~~ number of stations is shown in parentheses at the bottom right ~~top left~~-hand corner of each panel. ~~Furthermore, only~~ Those stations where the $1\hat{\sigma}$ range of the fitted trend excludes ~~does not include~~ zero are plotted with a larger symbol than those where this range encompasses zero ~~on each panel, as these were the ones where there is some confidence in there being a change over time~~ . As is also clear from these figures, there are many fewer stations in the southern hemisphere than in the north. The highest concentration of stations is in Europe and eastern Asia, and this uneven meridional distribution will have affected the zonal average calculations in Section 3.

~~As the highest station density is in the northern hemisphere, this is where any coherent changes are most clear.~~ For the distributions of daytime/maximum temperature (15:00, Fig. 3), it is immediately apparent that there are large and coherent changes in the mean and standard deviation over the period of record of these stations. Almost all changes in the mean are

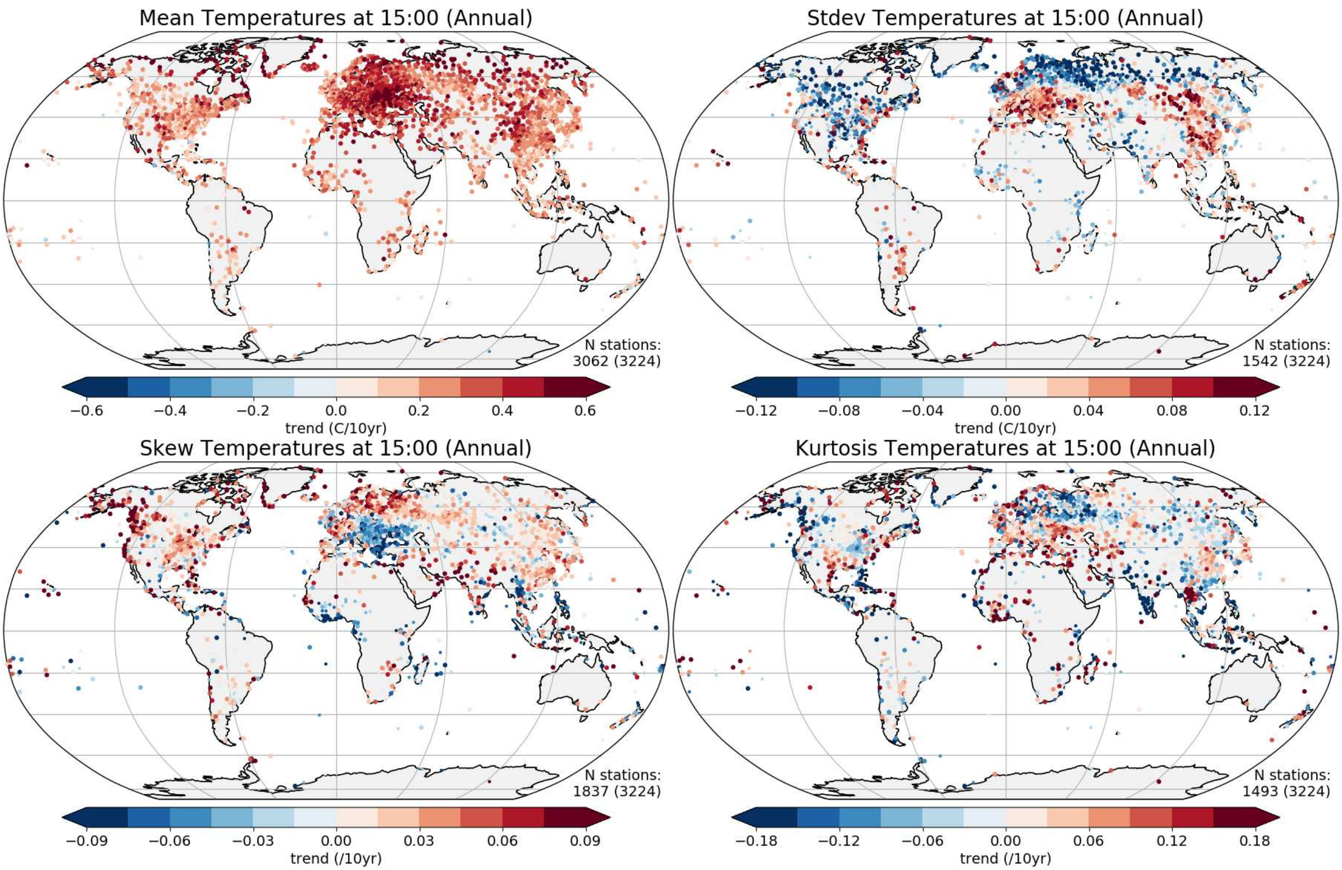

Figure 3. Trend over 1974-2018 in distribution parameters for temperature at 15:00 local time, showing the mean, standard-deviation, skew and kurtosis. The stations where the $1\hat{\sigma}$ range of the fitted trend excludes zero are plotted with a larger symbol (the number of these is shown in the bottom right hand corner of each plot). The total number of stations available is shown in parentheses.

positive, with the strongest signal in Eurasia, and lower values in North America. Approximately 95 ~~Over 90~~ per cent (3062 plotted out of a total of 3224) of the stations have trends where the uncertainty range does not encompass zero. The majority of changes in standard deviation at high northern latitudes are negative, along with most of central North America. There are increases over Europe and parts of China and southern Russia. Fewer stations (around 45 per cent, 1542/3224) have trends where the uncertainty range of the trends do not encompass zero.

Changes in skew and kurtosis have a more heterogenous signal, and again around half the stations exhibit changes where the uncertainty range does not include zero. However, the changes are not completely random, with consistent regional patterns in some areas, but these are smaller than for the standard deviation. The general pattern is for slight increases in the skew, except in a belt around the tropics and in south-east ~~half of~~ Europe. The kurtosis shows decreases in north-central Europe, and a band extending eastwards from the White Sea across to China and India, with a more mixed signal elsewhere with smaller areas showing coherent changes.

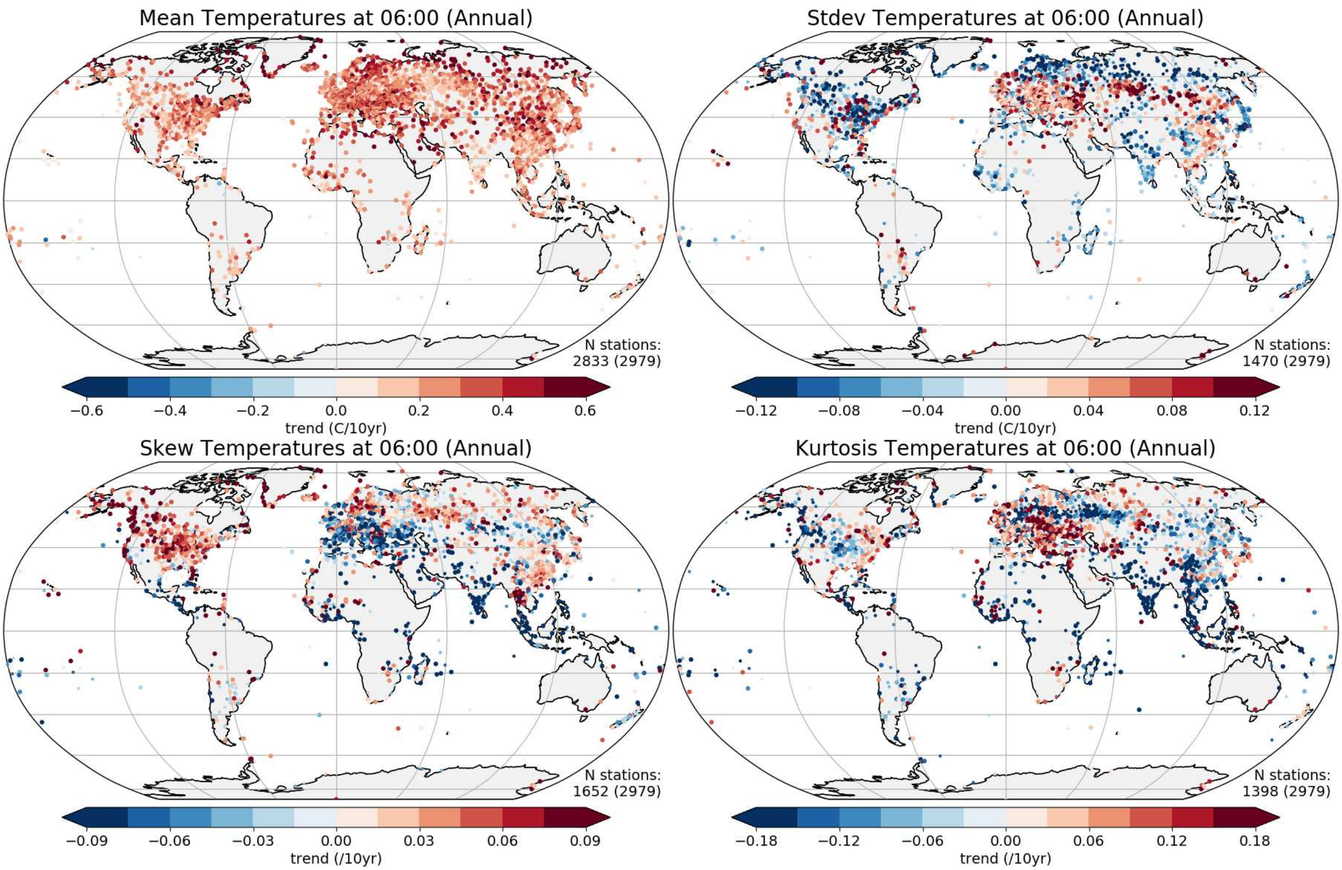

Figure 4. Trend over 1974-2018 in distribution parameters for temperature at 06:00 local time, showing the mean (°C), standard-deviation (°C), skew and kurtosis. The stations where the $1\hat{\sigma}$ range of the fitted trend excludes zero are plotted with a larger symbol (the number of these is shown in the bottom right hand corner of each plot). The total number of stations available is shown in parentheses.

The night time observations (06:00) show very similar general patterns for changes in the mean and standard deviation, with quasi-global increases in the mean, and decreases in the standard deviation at high latitudes and increases for China, Europe and south-eastern Russia. However, the magnitude of the changes in the mean are lower than for the daytime (15:00) observation, especially over Europe.

There is a stronger change between the skewness for the 06:00 observations and those from 15:00, with larger regions showing a coherent signal. Decreases are visible ~~in~~ more widely in Europe ~~than at 15:00~~, and also in India and parts of south-east Asia; and increases in the high latitudes, as well as North America. Regions with a strong decrease in kurtosis are mid-western North America, India through China and a band in north western Russia, but in contrast to 15:00, large parts of Europe show an increase. ~~However, a smaller fraction of stations have trends where the uncertainty ranges do not encompass zero than for the higher mom~~

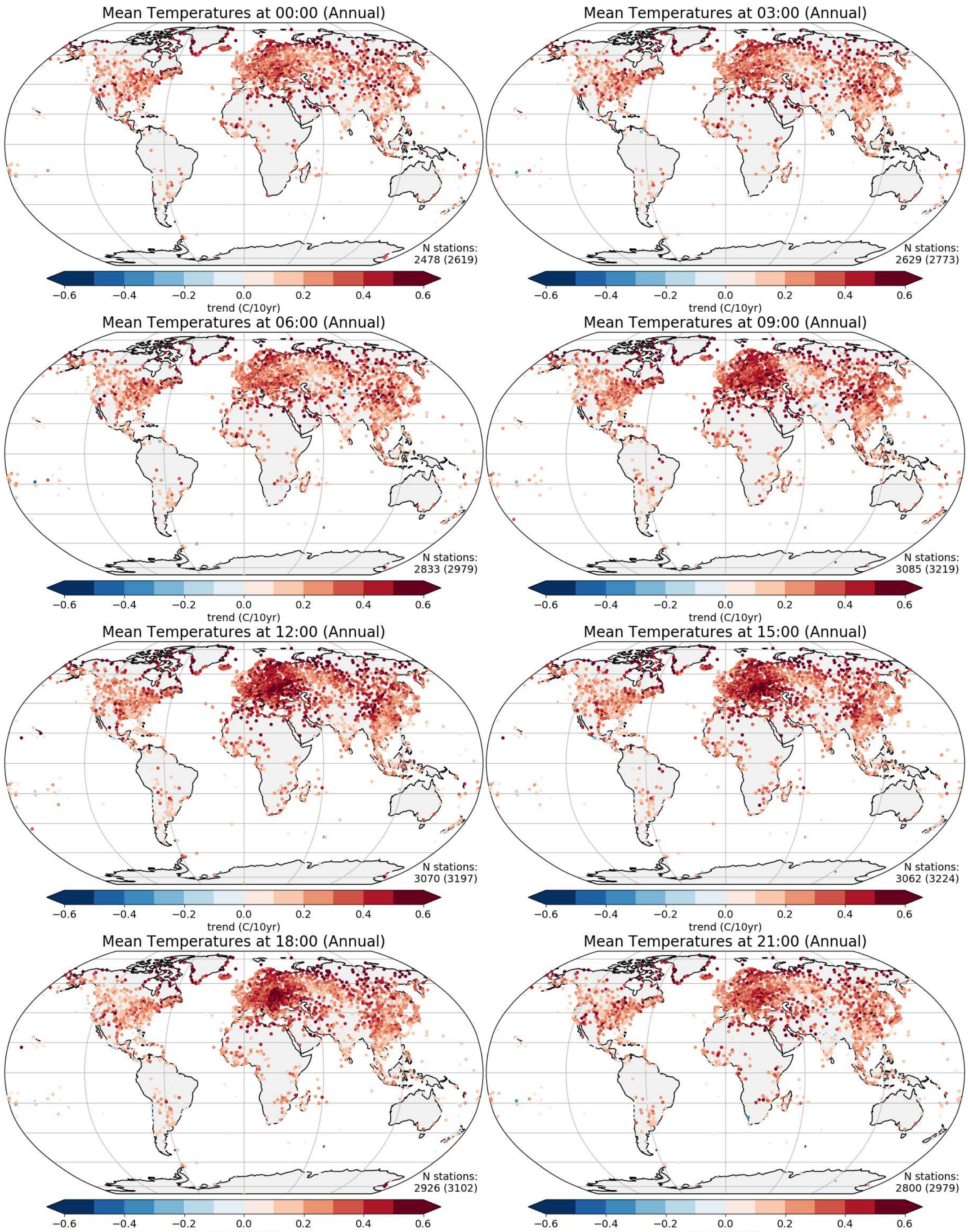

Figure 5. Trend over 1974-2018 in distribution mean for temperature at each three hourly interval. The stations where the $1\hat{\sigma}$ range of the fitted trend excludes zero are plotted with a larger symbol (the number of these is shown in the bottom right hand corner of each plot). The total number of stations available is shown in parentheses.

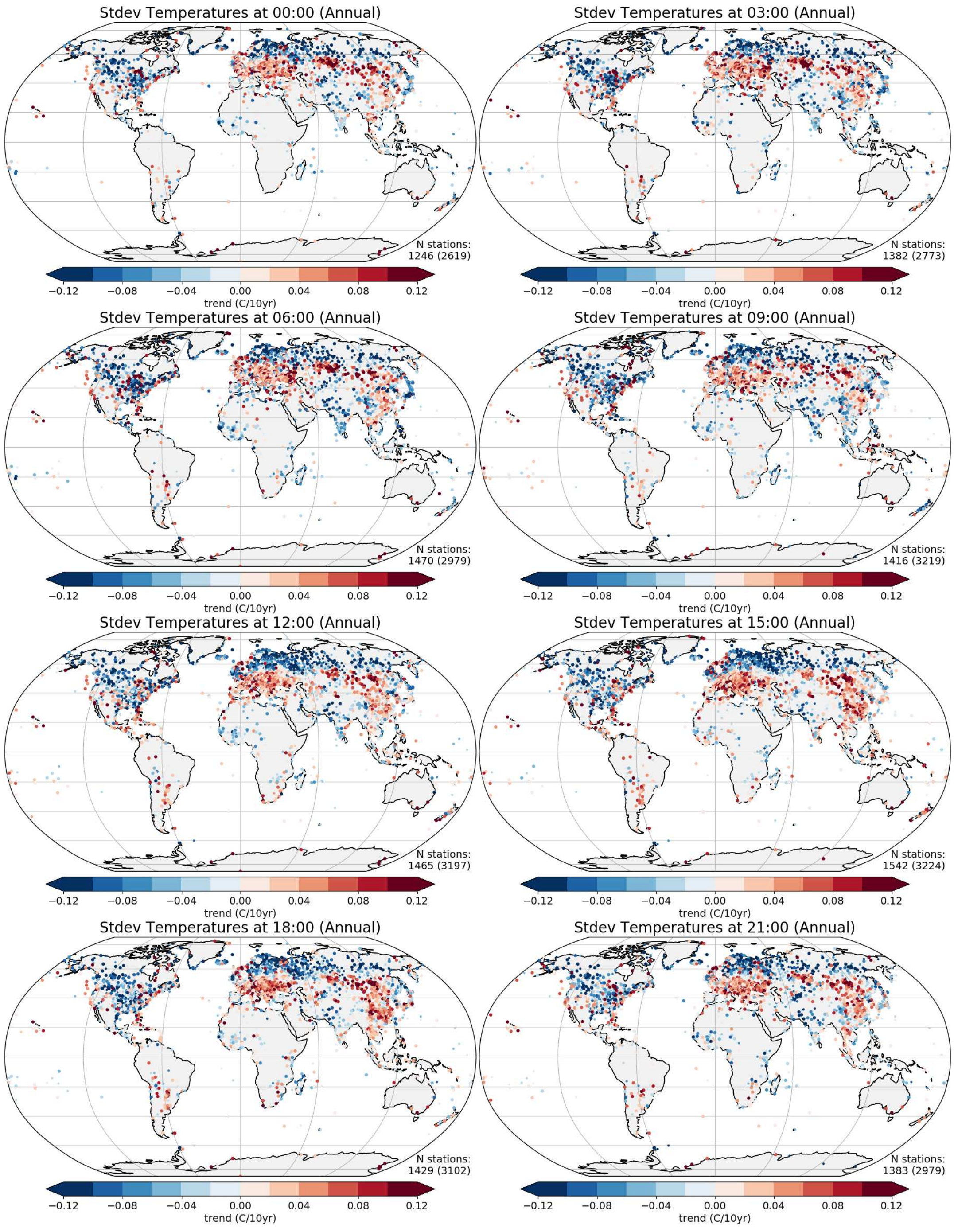

Figure 6. Trend over 1974-2018 in distribution standard deviation for temperature at each three hourly interval. The stations where the $1\hat{\sigma}$ range of the fitted trend excludes zero are plotted with a larger symbol (the number of these is shown in the bottom right hand corner of each plot). The total number of stations available is shown in parentheses.

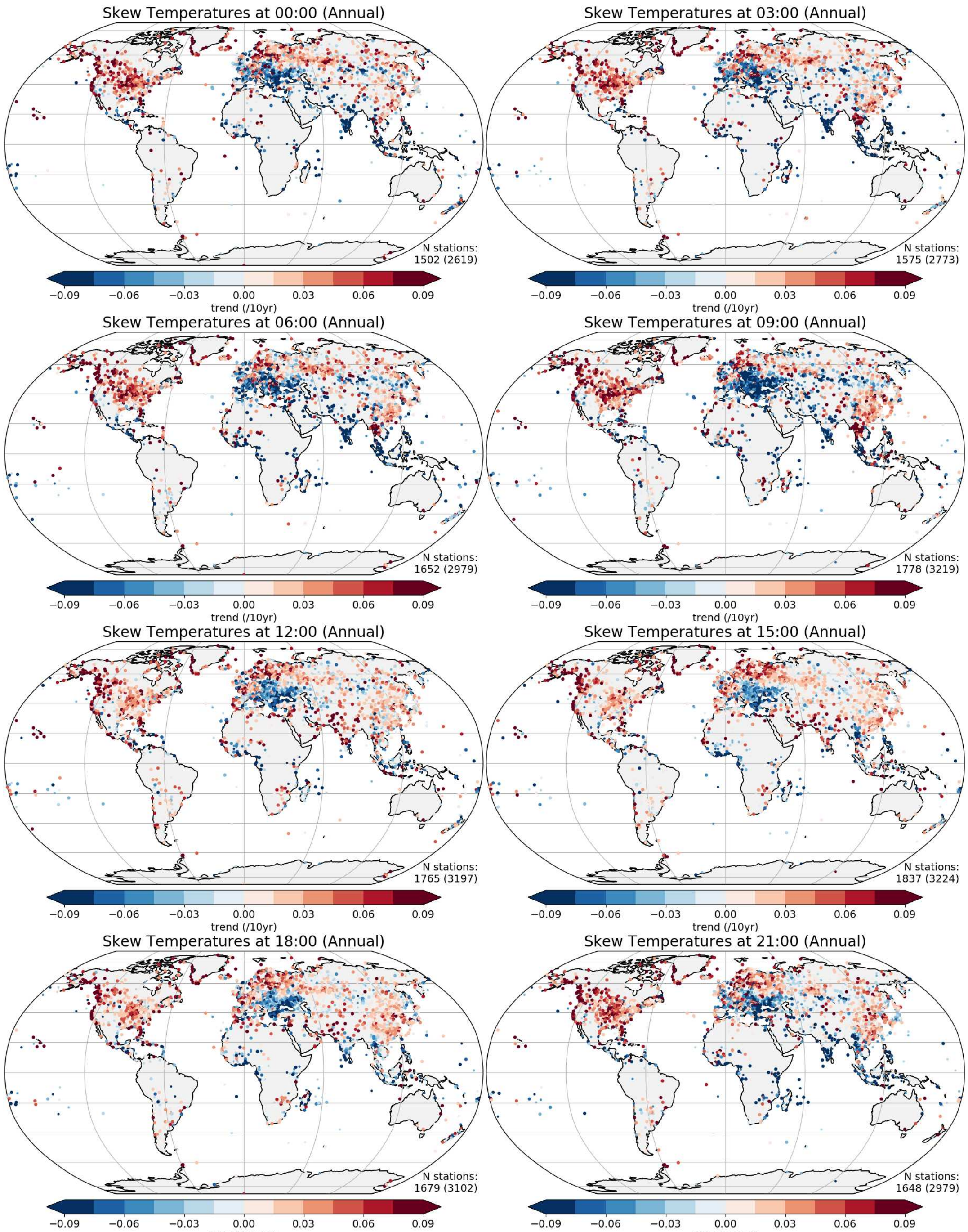

Figure 7. Trend over 1974-2018 in distribution skewness for temperature at each three hourly interval. The stations where the $1\hat{\sigma}$ range of the fitted trend excludes zero are plotted with a larger symbol (the number of these is shown in the bottom right hand corner of each plot). The total number of stations available is shown in parentheses.

4.1.1 Temperature Changes Across the Day

In ~~the Supplementary Information~~ Fig. 5 we show the maps for changes in the mean for all eight of the 3-hourly time-stamps. All hours of the day show an increase in the mean, with the strongest signals during the daytime (09:00, 12:00, 15:00). The region with the largest trends is Europe. However, it also has the highest station densities, and so the eye is more drawn to this region than areas which also show large trends but from a sparser station network. ~~It appears that Europe shows the largest trends, but this may be enhanced to the high station densities in this region compared to regions with a sparser distribution of stations.~~

The trends in standard deviation ~~at 3-hourly intervals from HadISD~~ show decreases at all hours in high northern latitudes, North America, and southern Asia (Indian subcontinent and surroundings), ~~Supplementary Information~~ Fig. 6. Increases in the standard deviation are seen in a band, from Europe across to China. The skew from HadISD (~~Supplementary Information~~ Fig. 7) shows increases in most regions except the majority of Europe and parts of Asia. The decreases in skew over Europe are strongest in the mid-morning (09:00) and weakest in the late evening (21:00). ~~It also appears that the~~ Also, changes in standard deviation and skew are the inverse of each other, when comparing Figs. 6 and 7 ~~of the Supplementary Information~~. ~~The decreasing (changing to more negative) skews indicate a shift from right to left-tailed distributions.~~

4.1.2 Seasonal Temperature Changes

In the Supplementary Information Figs. S9-S12 we show the changes for the first four moments from the 15:00 observations across the four seasons of the year (and Figs. S13-S16 for the 06:00 observations), with the standard deviations at 15:00 being reproduced here in Fig 8. Again, the strongest variations are over Eurasia. Changes in the mean temperature at 15:00 are smallest in boreal winter[1] (DJF), and strongest over a wide area in spring (MAM). However, summer (JJA) temperatures increase strongly with time around the Mediterranean and across continental Europe, and winter (DJF) has strong warming in the western part of Russia and Scandinavia. In North America, the largest trends are seen in SON (especially in higher latitudes), with little variation over the other seasons, especially spring and summer. A similar, but less intense pattern is observed at 06:00.

~~A smaller proportion of stations shows changes in the standard deviation where the uncertainty range does not encompass zero, than for changes in the mean, but there are variations in the patterns of change over the year~~. Although increases in the standard deviation are seen across Asia in springtime and summer at 15:00, a more mixed pattern is present in this region in autumn; decreases are seen in the far east in winter (Fig 8). At 0600, the pattern is somewhat similar, though with the Indian sub-continent and surrounding regions showing consistent decreases throughout the year. Over North America, standard deviations increase during spring, but mostly decrease during the other seasons, except in summer at 06:00. Further increases in the standard deviation are also found around the Mediterranean and in south-eastern Europe for most of the year, with

[1]We will use the seasonal descriptions for the northern mid-latitudes, as the station density is highest and we are able to draw the clearest inferences. We note that these seasonal descriptions are not appropriate for all regions, and hence have defined them in the plots by months as well as the text where they are introduced.

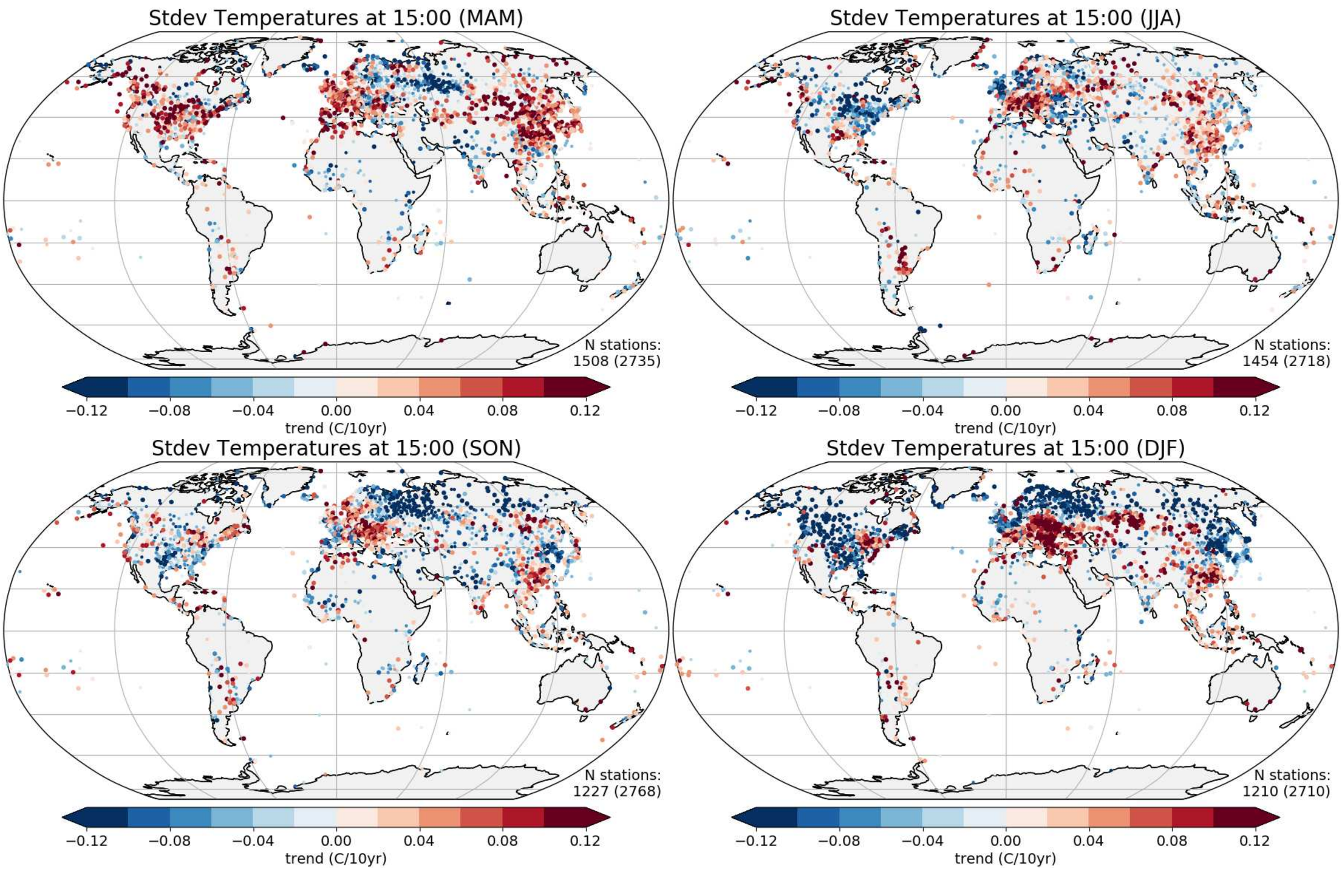

Figure 8. Trend over 1974-2018 in distribution standard deviation for temperature at 1500 local time across the four seasons. The stations where the $1\hat{\sigma}$ range of the fitted trend excludes zero are plotted with a larger symbol (the number of these is shown in the bottom right hand corner of each plot). The total number of stations available is shown in parentheses.

decreases in the north of the continent (British Isles and Scandinavia). Increases in the standard deviation are generally weaker and decreases more prevalent across all seasons at 06:00 then at 15:00.

Patterns of changes in the skew have some clear features, with decreases in China for all seasons except winter. Europe also shows decreases in the spring, but a more mixed pattern for the rest of the year. Increases in the skew at 15:00 are found in northern high latitudes during spring, as well as India for most of the year. Over North America, decreases in skew are observed during spring and autumn in the southern half, but in the northern half during summer.

4.1.3 Temperature Discussion

The results at 15:00 and 06:00 can be compared to the maximum and minimum temperatures in the study by Donat and Alexander (2012) using HadGHCND between 1951 and 2010. Unsurprisingly, the increase in the mean temperatures is also visible in the HadGHCND study, along with numerous other studies on recent climate change. However, HadGHCND shows stronger increases in the minimum temperatures (analogous to 06:00 here), whereas HadISD shows stronger increases in daytime (15:00) temperatures. The HadGHCND dataset uses inverse distance weighting to convert the daily station observations into a gridded dataset. One aspect of this process is that it smooths the data, and also can lead to very few stations contributing to large areas of interpolated values.

For HadGHCND in Donat and Alexander (2012) find larger regions showing stronger increases in standard deviation and a stronger increase around Hudson Bay for both maximum and minimum temperature values than for HadISD. The heterogeneous nature of the changes in the standard deviation, along with the smaller fraction of stations showing clear trends is in line with the study of Huntingford et al. (2013). For the skew, HadGHCND shows decreasing skew in Greenland, northern Africa, the coasts of eastern Asia and parts of Australia and increases over Europe, in contrast to the results of this study, which in some regions may be the result of low station densities in HadISD (Fig 1). The strong decreases in kurtosis observed close to the tropics in Section 3 are now shown to be concentrated in southern India and south-east Asia, and not always linked to an opposite trend in the standard deviation.

Although McKinnon et al. (2016) link the first four statistical moments derived from GHCND between 1980 and 2015 onto orthogonal basis functions, these still are closely tied to the original statistical moments. We also use their maximum and minimum temperatures to compare to the 15:00 and 06:00 local time values respectively. Over Europe, the geographical peak of trends in the daytime/maximum mean are both in eastern Europe. However, HadISD does not show such a strong feature of increasing standard deviations across Asia, and GHCND does not show the decreasing daytime standard deviations in northern Europe, but those in eastern Asia do agree. The results from HadISD do not show daytime cooling in the eastern half of North America observed in GHCND. However the general decrease in variance is observed in both. The increase in the central southern region of North America in GHCND is not present here, though the agreement is better by night Over North America for the minimum/night-time temperatures, the agreement is better for the mean, standard deviation and also skew. And over Europe and Asia, the lower magnitude trends for the mean also agree spatially, however the patterns for the higher moments do not.

Comparing the seasonal plots with those showing changes in the distribution of average daily temperature from Cavanaugh and Shen (2014, 2015) using daily GHCND temperatures over 1950-2010 shows generally good agreement for the changes in the mean. They find the strongest changes over Europe and a clear “warming hole” (Portmann et al., 2009) over the US, especially in JJA and DJF. In HadISD, the strongest changes are over Europe, and the location of the largest values moves between the seasons; but the warming hole is only weakly apparent during JJA over North America. For standard deviation, both Cavanaugh and Shen (2014) and our study find coherent decreases in the Northern Hemisphere, as well as increases in MAM over North America. Over Europe, however, JJA shows the largest differences, with widespread increases in Cavanaugh and Shen (2014, 2015) but changes in both directions are clear in HadISD. McKinnon et al. (2016) also show the decrease in standard deviation over North America, except for a region of increase on the Gulf of Mexico. Also across Eurasia, the patterns are similar in HadISD and McKinnon et al. (2016). The positive trends in skewness found over North America (Cavanaugh and Shen, 2014, 2015) in GHCND are not as prominent in HadISD nor in McKinnon et al. (2016) also using GHCND, with contiguous regions of negative trends in all seasons.

Part of these differences may result from inherent differences in the datasets, for example the differing temporal periods used; HadISD using 1974-2018 compared to ~~and~~ GHCND 1980-2015/1950-2010 and HadGHCND 1951-2010. We note that the number of stations reporting in HadISD prior to 1973 is much lower, restricting the possibility of extending our analyses further into the past. Also, here we and McKinnon et al. (2016) do not combine the stations together, unlike in Section 3, Donat and Alexander (2012) and in Cavanaugh and Shen (2014, 2015). The largest difference is that the sub-daily data within HadISD are not the true maximum and minimum values which comprise GHCND and HadGHCND. Although we have used the morning and afternoon values as proxies to the true daily maximum and minimum temperatures, this will not always be the correct match. This does, however, allow the study of how extremes are changing throughout the day. It is difficult to say which of these analyses is the best at determining the change in the distributions of temperature, given the different temporal periods and processing methods, station selections and level of smoothing. Where a number of studies agree on the sign of the change, results are naturally more reliable than regions where they do not.

Changes in the temperature distributions for stations in China are likely to be the result not only of warming on a global scale, but also from more local effects arising from rapid urbanisation (Sun et al., 2016). Impacts from urbanisation are also likely in other regions undergoing rapid development in recent decades. As only a small fraction of the Chinese station network is unaffected by urbanisation (32% in 2015, Wang et al., 2015), any changes are more representative of the urban areas (and hence a larger fraction of the population). Our study period coincides with the rapid urbanisation over China, and as noted above, we see strong increases in the mean of the temperature distributions at all times of the day. Stronger warming is seen between 06:00 and 12:00 for northern and western parts of China, but not for the southern, coastal regions. Wang et al. (2017) found that the effect of urbanisation was greater on the minimum (morning) temperatures than the maximum (afternoon) ones, which we find in the northern and western parts, but not in the southern, coastal ones (Fig 5). Although China is one example of recent and rapid urbanisation, with resulting effects on the observations as well as people's experience of the weather, other parts of the world will also be affected. This may be to a lesser extent because of a station network better suited to climate monitoring, or the urbanisation occurring over a longer time period.

A number of different causes for changes in temperatures beyond the anthropogenic warming signal, have been suggested. These drivers of change are likely to influence the distributions of observations across the hours of the day in different ways. Large scale circulation changes would change the weather patterns experienced by different regions. Studies across the world have found local and global modes of variability linked to changes in cold and warm extreme temperatures (e.g. Scaife et al. 2008; Kenyon and Hegerl 2008; Barrucand et al. 2008; Sillmann and Croci-Maspoli 2009; Zongxing et al. 2012; Ning and Bradley 2015). Hence changes at these large scale circulation patterns and modes will affect the extremes. However, taking the El Niño Southern Oscillation as an example, there is no consensus on the change under global warming (Stevenson, 2012; Cai et al., 2014; Kim et al., 2014; Vega-Westhoff and Sriver, 2017). That said, as some of the changes in the higher moments shown here have a strong zonal component, especially over Eurasia, this suggests some influence of the poleward expansion of the Hadley Cells (Hu and Fu, 2007) which have been linked to anothropogenic forcings (Lu et al., 2007; Lucas et al., 2014; Tao et al., 2016) on the changes in shape of the temperature distributions.

Other drivers of change in observed temperature distributions have been suggested, including enhanced land-atmosphere coupling and soil-moisture (e.g. Seneviratne et al., 2006; Jaeger and Seneviratne, 2011; Hirschi et al., 2011) or changes in cloudiness and radiation (e.g. Lenderink et al., 2007). In a modelling study of future summer climate over Europe, Fischer and Schär (2008) find that the seasonal cycle as well as interannual and intraseasonal variability have separate effects on the temperature variabiloty, and each of these are affected to a different level by the other drivers.

4.2 Dewpoint Temperature Changes

Comparing the changes in the dewpoint temperature moments across the day and the seasonal cycle (Supplementary Information Figs. S17 to S28) with those from the temperatures, shows that on the whole these two variables are quite closely aligned. Despite fewer stations passing the selection criteria for this variable, there are still coherent patterns of change. Over the eight time points during the day, the changes in the mean dewpoints show less variation than the temperatures, most clearly apparent over Europe in the early afternoon. This suggests that the relative humidity over this region has decreased for this part of the day. A similar but less striking decrease in daytime relative humidity is also seen over North America. This decrease in the relative humidity has been observed regionally and globally in the homogenised HadISDH dataset (Willett et al., 2014, 2019), which is based on the HadISD.

The changes in the higher moments of the dewpoint temperature are also reasonably constant through the day, but have a greater spatial heterogeneity and lower magnitude than the temperatures. Notably, there is a clear increase in skew over south-east Asia during the daytime and early evening, which is not seen in the temperature results, indicating a change in shape away from a low-dewpoint tail towards a high-dewpoint tail.

Seasonally, the greatest change in the dewpoint mean is in SON and DJF over Europe in both the day- and night-time, contrasting with the temperatures of MAM and JJA for the day and also DJF at night. Again the standard deviations and kurtoses are very similar for day and night across all seasons between the dewpoints and temperatures. For the skews, the increase in skew over south-east Asia noted above is most prominent during SON, during the daytime, but is also present at night and differs from the behaviour of the temperatures for all seasons.

4.3 Wind speed Changes

There are fewer stations with sufficient observations to assess changes in wind speed (Supplementary Information Figs. S29 to S33), and very few in the Southern Hemisphere. As a result we focus our analysis on North America and Eurasia. The proportion of stations where changes in the distribution could be assessed varies between each time step to a greater extent than for the two temperature variables. Despite this, regions with sufficient station densities show a diurnal cycle in the behaviour of wind speed trends. There are stronger trends for declining wind speeds during the daytime (09:00-18:00) than for the night-time hours. And there are also no large regions that show large increases (at the same magnitude as decreases) in the wind speeds at any time point. However some individual stations do show some increases, for example in south-eastern Europe. For the other three moments (not shown) there are even smaller differences than for the mean for the annual changes over the day.

We show the seasonal changes at 15:00 for all four moments in Supplementary Information Figs. S30 to S33. For the mean there are no large changes across the seasons. The changes in standard deviation are also generally unvarying compared to the temperature measures. There are increases in the USA, with stronger values during spring and autumn. All other regions with sufficient station density have decreases in the standard deviation. Both skew and kurtosis show no strong seasonal variations. Skews are increasing over North America but there are no coherent, strong changes elsewhere. The kurtosis is mostly decreasing in south-eastern Europe, and also on the whole across Asia.

4.4 Discussion of Station-Level Distributions

Across all seasons, times of day and locations, there are strong increases in the means of both temperature and dewpoint temperature with time as expected in a warming world (Stocker, 2014). Although the geographical distribution of stations is not uniform, there appear to be larger increases in Europe, especially during the spring and summer ~~, than elsewhere in the world~~ The largest increases in ~~the high station density regions of~~ Europe also occur during the daytime rather than night-time, which matches the behaviour seen in Table 1 40-60N (the latitudes corresponding to Europe). ~~However, in most of the other latitude bands in Table 1 the night-time means are increasing faster than the daytime ones.~~ The standard deviations show stronger decreases in the northern latitudes, but more regions with increases in the mid- and tropical latitudes in the northern hemisphere. Seasonally, North America shows decreases in the daytime and night-time standard deviation in winter and summer, but increases in spring (Fig. 8 and Supplementary Information Figs. S10 and S14), which was also observed in Cavanaugh and Shen (2015). This indicates a change towards less variability in winter and summer, but a change towards greater variability in spring for both day- and night-time temperatures.

We also note that in Section 3, standard deviations of the temperatures and dewpoint temperatures were decreasing more or less across the whole northern hemisphere, whereas in this study there are regions with clear increases. The zonal study combines the stations together, even if individually any change in the statistical moments over time would not show clear changes ($1\hat{\sigma}$ range of the trends encompassing zero). In contrast, Figs. 3 and 4 emphasise those stations where the $1\hat{\sigma}$ range of fitted trends do not include zero. As noted, this is only around half of the total number of stations and many of the others show decreasing trends. ~~Earlier versions of these figures plotted all stations, and suggested that many of these exhibit decreases in the standard deviation. However, for clarity, we have decided to only plot those stations in which the trend is clearly different to zero here.~~

Changes in annual skew are different on either side of the North Atlantic, with predominantly increases in North America, and decreases in Europe across the day, and hence a rather small change in the zonally averaged results (Section 3). The seasonal results show increases at higher latitudes (Canada and Scandinavia), and decreases further south (USA and continental Europe) in spring. Other seasons show a more heterogeneous pattern in Europe, but increases across central Asia in summer and autumn, similar to the patterns from GHCND in Cavanaugh and Shen (2015). In China which was not included in their study, all seasons except winter show decreases in the skews. For the kurtoses, Cavanaugh and Shen (2015) found strong decreases in the southern parts of North America in the summer, which are not seen in HadISD. However, the strong decreases over parts of Europe in winter are present in HadISD.

The observations in HadISD have a finite numerical resolution, with more recent observations being more often recorded to 0.1°C, and those from earlier parts of the record to 1°C. Underlying this is a more complex process of conversion between Fahrenheit and Celsius (or knots and miles per hour to metres per second), possible truncation and rounding (or vice versa) as well as human induced errors during transcription, digitisation or recording. In the GHCND, Rhines et al. (2015) found that 65% of all temperature observations were misaligned as a result of these effects, and so it is likely that a large fraction of the HadISD observations also have been affected.

Rounding has an effect on the mean and standard deviation dependent on the rounding resolution and any offset in the two scales from zero (rounding width and shift parameter), with zero bias for some combinations (Schneeweiß et al., 2010; Sheppard, 1897), and likely therefore on higher moments as well.

These changes are combined with the discrete measurement times, which also are more likely to be hourly in the more recent period, and more stations will be recording only every three or six hours in the earlier period. More frequent recording results in a greater likelihood of values close to the true extremes being recorded. Hence, under a stationary climate, this could result in both an increase in the maximum temperatures and a decrease in the minimum temperatures (for a perfect, sinusoidal diurnal cycle). Combined with a warming climate, this could result in an enhancement of the maximum temperatures recorded. There will also have been changes in instruments and recording practices over the period, though the homogeneity assessment should have identified any gross traces these have left in the data record.

It is possible that the effect of these data measurement issues is systematic across the entire global network. We also note that despite the automated quality control and homogeneity assessment of HadISD, there are likely to be data artefacts still present in this global dataset. Therefore further investigation is needed to determine how these two aspects combine to affect the extremes recorded and their change over time in the sub-daily data. However, in our analysis we have looked for contiguous regions of coherent changes for large numbers of stations, spanning geo-political boundaries. Wholesale changes of national networks and their measurement routines causing systematic biases in the results presented here are likely to stand out as political areas, which are not observed.

5 Quantile Regression

Studies of the characteristics of a distribution are useful in describing the overall changes in a meteorological variable. However, they do not make the changes in extremes immediately clear, although they can be calculated through, e.g. extreme value theory ~~inferred~~. An alternative way to study the changes in the climate that show changes in past extreme events is to investigate the changes of the quantiles of a distribution themselves. This enables the question of "how much warmer are the warmest 10 per cent of days now than they were in the past?" to be addressed. Quantile regression can be used to calculate~~s~~ trends in specified quantiles of a distribution, rather than trends in the mean (Koenker and Bassett Jr, 1978; Koenker and Hallock, 2001; Barbosa et al., 2011; Franzke, 2015). This allows, for example, cases where there is little or no mean trend but an increased variance over time to be assessed more completely. In the case of climate observations, and especially given the impact of extreme events on society and infrastructure, quantile regression allows the changes in these extremes over time to be made apparent.

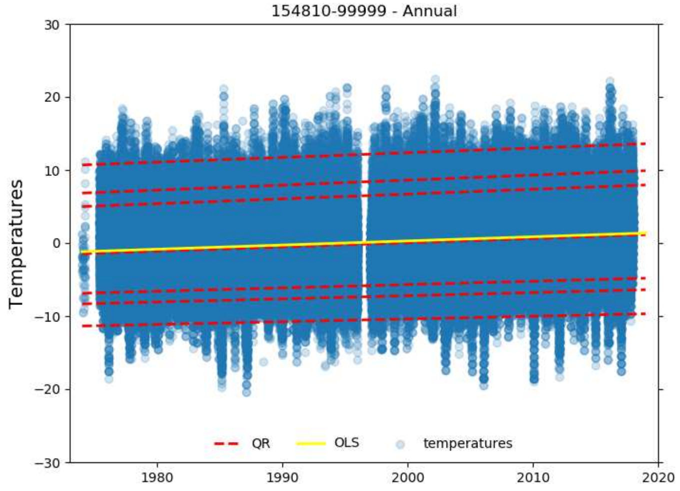

Figure 9. An example of the Quantile Regression analysis for a single station, with the trends in the quantiles in red, and the ordinary least squares (OLS) fit in yellow. This is for station 154810-99999, Constanţa Mihail Kogălniceanu International Airport, Romania.

The observations in HadISD naturally contain both an annual and a diurnal cycle. If run purely on these raw data, then the annual upper quantiles are likely dominated by summer daytime temperatures, and the lower by winter night-time temperatures. Although using standard three-month seasons would enable the more detailed study of winter and summer quantiles, those from the shoulder seasons (spring and autumn) would be dominated by the earlier or later parts of these seasons given the strong annual cycle. For this reason McKinnon et al. (2016) use daily maximum and minimum temperature data from GHCND over July and August only to reduce the effect of the seasonal cycle in their study.

Our approach is to use de-seasonlised data, which removes the seasonal aspect described above and also the issue that quantile regression results can be biased when applied to discrete data (Machado and Silva, 2005; McKinnon et al., 2016). The observations in HadISD are at a finite numerical resolution (e.g. temperatures are reported to the nearest 0.1°C, 0.5°C or 1°C). We use a daily climatology to create hourly anomalies of the HadISD data, retaining the diurnal cycle. Within each calendar day, the numerical resolution remains unchanged by the subtraction of a daily mean from the hourly observations (within-day differences will still have 0.1°C, 0.5°C or 1°C resolution). However, etween each calendar day, the different climatology value will result in a different offset, so that between-day differences would not result in discrete values. Hence, overall, the values passed to the quantile regression algorithm will be less discrete than the raw observations themselves. As discussed in Sections 3.4 and 4.4, we note that the final resolution of HadISD hides the effects of conversion, rounding and truncation (multiple times in various orders) that could occur between measurement and the current use (Rhines et al., 2015).

Daily means were calculated for 1981-2010, requiring at least eight observations in each 24-hour period (equivalent to 3-hourly data). The daily climatology was calculated requiring at least 15 years of data to obtain an average. Finally, the climatology was smoothed using a 5-point binomial filter and subtracted from the sub-daily data for each day to make the climate anomalies. An example of the changes in the percentiles of a timeseries is shown in Fig. 9.

Both Europe and North America have high station densities (see Section 4) compared with other regions, and so our analysis is focussed on these two regions (see also McKinnon et al., 2016). Furthermore, we show the results using a space-filling method, Voronoi tesselation (derived from Delaunay triangulation, Voronoï, 1908; Delaunay, 1934), which uses the perpendicular bisectors of the lines joining each station to split the land surface up into polygons , with each polygon containing a single station. These have then been coloured to show the results of the quantile regression. Stations for which the slope of the quantile shown is not significantly different from zero ($p > 0.05$ as determined by the quantile regression algorithm) are shown in grey, but we do no further significance testing on these trends. It should be noted that areas of high station density will naturally have smaller polygons, but the human eye is drawn to the larger blocks of colour arising from sparse station networks. However, the advantage of this approach is that the underlying station information is retained, as opposed to it being smoothed when applying some form of gridding method.

We show the changes in the quantiles (0.01, 0.05, 0.1, 0.9, 0.95, and 0.99) over Europe for both JJA and DJF in Figs. 10 and 11 respectively. In some sub-regions large numbers of stations show similar patterns, suggesting that there is a coherent change in the quantile in that area. However, within these there may be isolated stations/cells showing changes in the opposite direction. Although we have selected those stations that do not have many or large breaks in their timeseries (as determined by PHA, see Section 2), and the HadISD has undergone a number of quality control checks, there will inevitably be a few data quality issues remaining.

The largest change in the summer (JJA) quantiles is observed in south-eastern Europe and western Asia (Fig. 10), focussed on the area north and west of the Black Sea. The increase observed in these regions becomes progressively less strong towards lower quantiles and also with distance from this region. This indicates that the temperature distribution is broadening over time, with the upper tails changing more rapidly than the lower tails; i.e. the warmest N per cent of events are becoming warmer as the percentile values change, more rapidly than the cool events (where N is 1, 5, 10 etc). In contrast, the regions close to the North Sea and especially Scandinavia show slight increases for the lower quantiles that becomes less strong for higher quantiles. For the north and west of the UK in general and Scandinavia except western Norway, there are even decreases for the highest quantiles.

Conversely, in the winter (DJF), the strongest increases are in the northern regions and the lower quantiles, with Scandinavia, the Baltic states along with Benelux, Denmark and northern Germany indicating a narrowing of the temperature distribution over time. In the higher quantiles, there are increases for eastern Europe but not as much as in the summer. Westernmost Europe shows very little change across all quantiles. We show the panels for the shoulder seasons (MAM and SON) in the Supplementary Information Figs. S34 and S35. There is a general increase for all the percentiles, and the cold quantiles in the north east of the region shown are warming faster than the rest. But in contrast to Figs. 10 and 11 these increases are much slower and the regions are less coherent. Overall, these results show that warm extremes are becoming more common more quickly during the summer in south-eastern Europe, than elsewhere across the continent. Similarly, cool extremes are becoming less common more quickly during winter in northern Europe than elsewhere.

Franzke (2015) performed quantile regression on daily mean, maximum and minimum temperatures between 1950 and 2013. The mean temperatures show strong regions of increase in both the 5^{th} and 95^{th} percentiles in eastern Europe and western

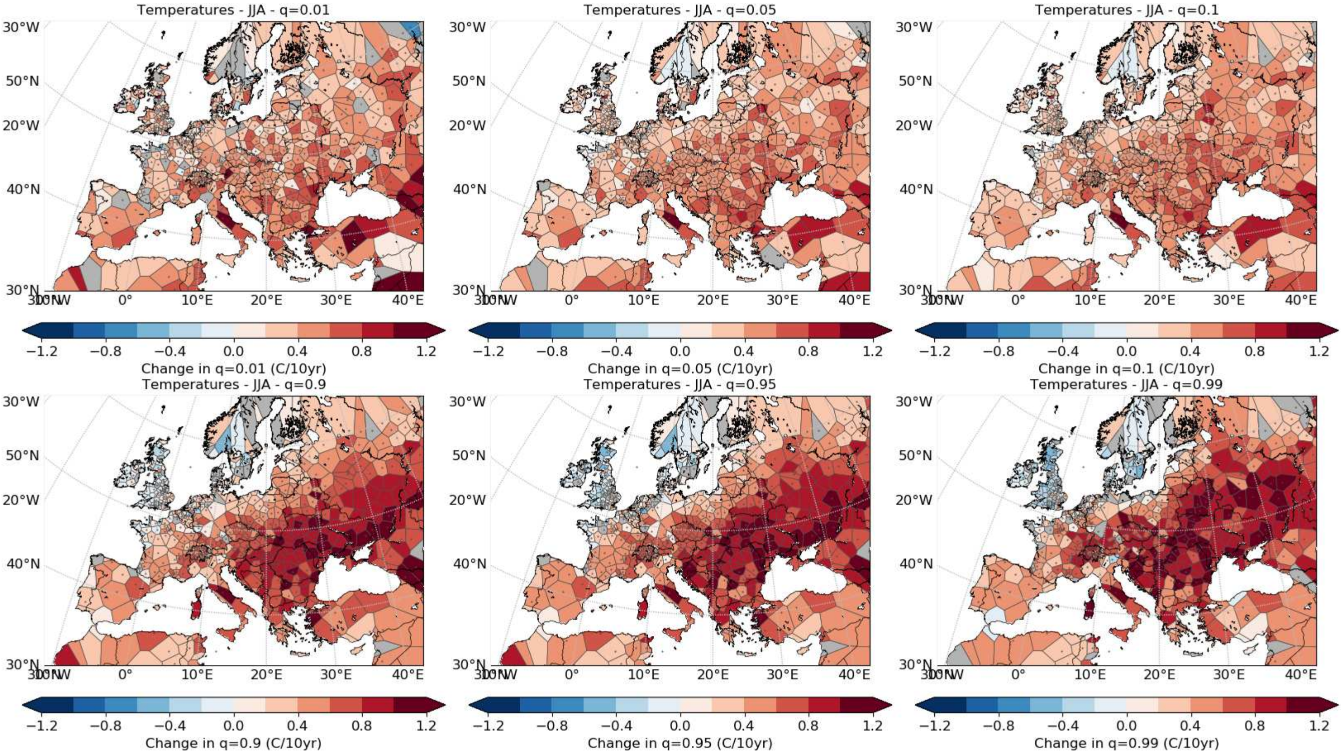

Figure 10. Trend over 1973-2017 in quantiles of temperature ($^{\circ}$C decade^{-1}) at TOP 0.01, 0.05 and 0.10; and BOTTOM 0.90, 0.95, 0.99 over Europe during summer (JJA).

Russia, especially north of the Black Sea, with the higher percentile region covering a larger area than for the results presented here. When using summer maxima, Franzke (2015) shows a more uniform increase over all of Europe at the higher percentile, except Scandinavia, which behaves similarly in Fig. 10. The magnitudes of the change in the winter temperature percentiles from HadISD are consistent with the studies of Barbosa et al. (2011); Franzke (2015). The behaviour over Turkey stands out in Franzke (2015), but not in HadISD, suggesting a difference in the way these data have been processed or effects arising from the difference in temporal coverage or even station selection.

The difference in behaviour of the lower and upper quantiles indicate a change in the variance of the sub-daily temperatures over Europe, which has already been suggested by the results from Section 4. During the summer, the variance in the large region to the north of the Black Sea, has increased. In contrast, during winter the variance over Scandinavia and northern Europe has decreased. Both of these changes are also clearly visible in Supplementary Information Fig. S10 showing the seasonal change of the standard deviation for each station. Franzke (2015) found a more widespread, uniform decrease in variance in the winter months over Europe, but with a stronger decrease over Scandinavia. However, Franzke (2015) found no increase in variance during the summer in eastern Europe.

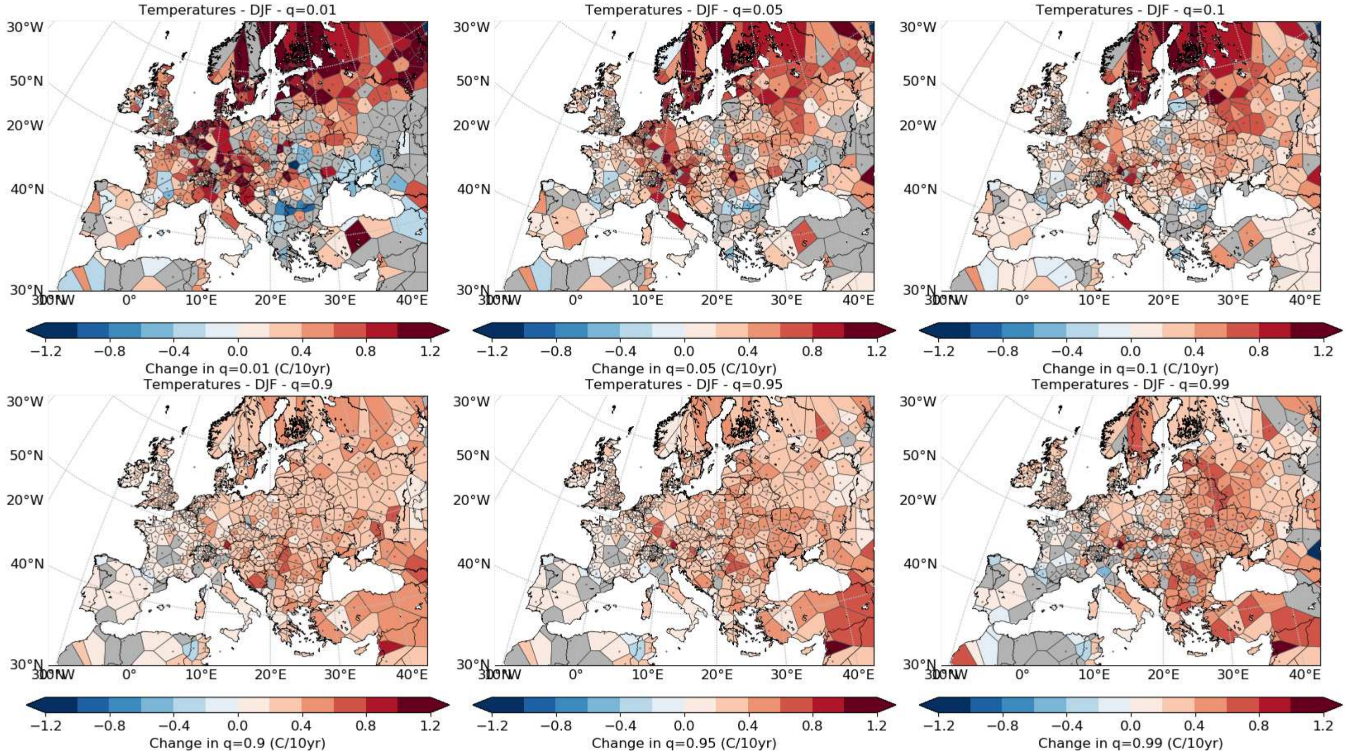

Figure 11. Trend over 1973-2017 in quantiles of temperature (°C decade^{-1}) at TOP 0.01, 0.05 and 0.10; and BOTTOM 0.90, 0.95, 0.99 over Europe during winter (DJF).

Over North America (Supplementary Information Figs. S36 to S39), the trends in quantiles are on the whole smaller than for Europe for most of the year. In the spring (MAM), there are moderate increases in the south-west, which are stronger for the higher quantiles. For the summer (JJA) the higher quantiles show decreases in the northern, central regions, but increases for the lower quantiles in the south west. In the autumn (SON) and winter (DJF) the strongest changes are in the north of the continent at the lower quantiles with more moderate increases in the south and the higher quantiles. Hence the most rapid changes are in the cool extremes during autumn and winter in the north, with less rapid changes in the warm extremes in spring in the south-west. Using GHCND, Rhines et al. (2017) show a similar pattern, with a strong increase in lower percentiles for northern latitudes in winter. In the spring their data show decreases in the lower percentiles in the north and increases in the south, almost the opposite in the autumn, and general warming for the minimum temperatures in the summer. We also find the strongest changes for the lower percentiles during winter, indicating a reduction in the range of temperatures during that season.

We also show changes in the dewpoint temperatures (Supplementary Information Figs. S40 to S43) and wind speeds (Fig. S44) over Europe. The dew point temperatures indicate small increases of the quantiles in most regions with no strong regional differences for MAM and JJA for any quantile. However, in SON and DJF, there are stronger increases in the lower quantiles for the north of the region (Scandinavia and north-western Russia), but small increases for other regions and quantiles. So, although the summer warm extremes are becoming more common in south-eastern Europe, there is no corresponding change in the dew point temperatures, suggesting a decrease in the relative humidity. However, in the northern parts, the dew point temperatures are changing roughly in step with the dry-bulb temperatures.

Only the annual changes in wind speeds are shown (Supplementary Information Fig. S44), with some blank tiles and regions because of the smaller number of stations with sufficient observations. There are no large changes in the lower quantiles but clear and larger decreases in the upper quantiles, which matches the behaviour observed in Section 3 and the decrease in variance observed in Section 4.3. The larger changes in the upper quantiles also suggest the skews are becoming more negative, with decreases in the number and intensities of extremely high wind speeds.

6 Summary

To investigate the changes in the shape of the distributions of temperatures and wind speeds in the HadISD dataset, we have performed three types of assessment. Firstly, by combining the data from the station based observations into zonal averages, the resulting distributions are built from large numbers of observations at the expense of some spatial information. Using five-year bins from 1973 to 2018, we demonstrated the change in the distributions and percentile values. As expected, the mean temperatures increase for both day- and night-time (Stocker, 2014). However, the standard deviations decrease for the northern hemisphere (where there are most stations). Although changes are seen in both tails of the distribution, they are greater in the lower tail. This behaviour is consistent with both the more rapid change in low temperature extremes (Davy et al., 2017) as well as the decrease in diurnal temperature range (Alexander et al., 2006; Thorne et al., 2016a, b). The dewpoint temperatures show similar, if more heterogeneous, behaviour in this analysis. Atmospheric moisture, as well as that in the soil may be acting as a brake on the change in the maximum temperatures, resulting in the consistent changes in shape of the distributions observed.

In contrast, the wind speeds showed decreases in the mean, but also in the standard deviations, with larger shifts in the upper tail than the lower. The changes in distributions observed on a zonal level suggest this may be driven by changes in the upper tail. A number of causes for this "stilling" have been proposed, from changes in land-use and circulation patterns to instrumental degradation. ~~A number of studies have investigated changes in global wind speeds, showing a steady decrease over the past decades, termed "stilling".~~

To improve the spatial discrimination, we ~~also~~ assessed changes in the distributions for each station individually ~~, and show fitted linear trends~~. The sub-daily nature of HadISD allows these to be calculated across the hours of the local day as well as annually or seasonally. The temperature means show the expected increase over time, but with a daily cycle showing a more rapid increase of mean temperatures during the early afternoon than in the early hours of the morning. However, the higher order moments are much less homogeneous, but have a smaller variation throughout the day. Trends in the standard deviations are on the whole negative, except for a band from the Mediterranean across to central Asia. Conversely, trends in the skew of the temperatures are mainly positive, except over Europe and across into western Asia, indicating increasing warm extremes (positive tail) and decreasing cool extremes. These patterns are in some regions different from those reported in previous investigations by Donat and Alexander (2012) and McKinnon et al. (2016) obtained using (derivatives of) GHCND (true maximum and minimum values). For the seasonal changes, which can be compared to Cavanaugh and Shen (2014), there are also differences to the patterns obtained from GHCND.

These disparities may result from a variety of differences in the datasets and methods. Firstly, the period used in this study is 1974-2018 compared to 1950-2010 (Donat and Alexander, 2012; Cavanaugh and Shen, 2014) and 1980-2015 (McKinnon et al., 2016). Also, by using all of the sub-daily observations in HadISD, this work incorporates the effect of changes across the entire distribution, even if we are concentrating on either the moments or specific percentile values (e.g. all temperature observations rather than just the maxima and minima). The studies using GHCND use the true maximum and minimum values, which allow the focussed investigation of the momentary extremes. We also do not perform any spatial smoothing of the underlying stations, which could result in less variable fields, with smaller extremes ~~and less extreme fields compared to station data~~.

A number of different drivers for changes in temperatures have been proposed, including changes in global and large scale circulation patterns, land-atmosphere coupling, and cloudiness and radiation. And a number of these have also been shown to be changing as a result of anthropogenic climate change. Over Eurasia, the quasi-zonal patterns of change of the higher moments suggest a link to the expanding Hadley Cells, but further study is required.

Finally ~~, rather than study the moments of the distributions,~~ we perform a quantile regression analysis to determine how the values of the percentiles have changed, and therefore how the warmest days and coolest nights feel in comparison to others. This approach is in contrast to the ETCCDI indices which use percentiles (e.g. TX90p), which capture the exceedence rate of a fixed percentile value. Here we show that over Europe, the upper percentiles are changing fastest in the large region north-west of the Black Sea in the summer, but the lower percentiles are changing faster in northern Europe and Scandinavia in the winter. Hence the warmest N per cent of summer days are becoming warmer more rapidly than the coolest N per cent in south-eastern Europe, and the reverse for winter days and particularly nights in the north. Coherent changes in dewpoint temperatures are less clear, with increases in the lower percentiles in autumn and winter for the northern parts of Europe. For the wind speeds, as expected from the other analyses, the higher quantiles show a decline over Europe, with stronger decreases in the east.

All the studies presented herein show widespread increases in mean temperature, and less rapid rises in the dewpoints, as expected from numerous other studies on the observed change in global temperatures. However, the changes in the higher moments are of more interest. Although in many cases (seasons and quantities) the patterns of change are very similar to prior studies (e.g. Donat and Alexander, 2012; Cavanaugh and Shen, 2014, 2015; McKinnon et al., 2016), there are still differences in the spatial patterns between all of the studies. These could arise from differences in the input data (daily extrema versus hourly values) as well as the temporal and spatial coverage. Furthermore, we have used a single, station-based dataset, and differences between datasets (Gross et al., 2018) or the effect of gridding (Cavanaugh and Shen, 2015) can change the apparent behaviour of higher moments. Nevertheless, in these investigations, we have shown that there are changes in the mean, the standard deviations and the skews of the distributions studied. And, when using a sub-daily dataset of temperatures, changes in the number and intensity of extremes appear to arise from the combination of changes in all three of these moments. Further investigations are however needed to determine the underlying drivers for the changes in these distributions, and hence the changes in extremes.

Data availability. The HadISD dataset is available under a non-commercial government licence at www.metoffice.gov.uk/hadobs/hadisd

Author contributions. Robert Dunn did the majority of the analysis, plotting and writing. Kate Willett and David Parker provided comments and guidance during this work. All authors have contributed text and edits to the main paper.

Competing interests. The authors declare there are no competing interests

Acknowledgements. We thank Colin Morice for helpful discussions during the course of this work, and Elizabeth Good and Nick Rayner for useful suggestions on the text. This work was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra (GA01101).

References

Alexander, L., Zhang, X., Peterson, T., Caesar, J., Gleason, B., Tank, A. K., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., et al.: Global observed changes in daily climate extremes of temperature and precipitation, Journal of Geophysical Research: Atmospheres, 111, 2006.

Azorin-Molina, C., Guijarro, J.-A., McVicar, T. R., Vicente-Serrano, S. M., Chen, D., Jerez, S., and Espírito-Santo, F.: Trends of daily peak wind gusts in Spain and Portugal, 1961–2014, Journal of Geophysical Research: Atmospheres, 121, 1059–1078, 2016.

Azorin-Molina, C., Vicente-Serrano, S. M., McVicar, T. R., Revuelto, J., Jerez, S., and López-Moreno, J.-I.: Assessing the impact of measurement time interval when calculating wind speed means and trends under the stilling phenomenon, International Journal of Climatology, 37, 480–492, 2017.

Azorin-Molina, C., Asin, J., McVicar, T. R., Minola, L., Lopez-Moreno, J. I., Vicente-Serrano, S. M., and Chen, D.: Evaluating anemometer drift: A statistical approach to correct biases in wind speed measurement, Atmospheric research, 203, 175–188, 2018.

Azorin-Molina, C., Dunn, R., Mears, C., Berrisford, P., McVicar, T., and Nicholas, J.: Surface Winds [in “State of the Climate in 2018”], Bulletin of the American Meteorological Society, 100, S43–S45, 2019.

Ballester, J., Giorgi, F., and Rodó, X.: Changes in European temperature extremes can be predicted from changes in PDF central statistics, Climatic change, 98, 277, 2010.

Barbosa, S., Scotto, M., and Alonso, A.: Summarising changes in air temperature over Central Europe by quantile regression and clustering, Natural Hazards and Earth System Sciences, 11, 3227–3233, 2011.

Barrucand, M., Rusticucci, M., and Vargas, W.: Temperature extremes in the south of South America in relation to Atlantic Ocean surface temperature and Southern Hemisphere circulation, Journal of Geophysical Research: Atmospheres, 113, 2008.

Brown, S., Caesar, J., and Ferro, C. A.: Global changes in extreme daily temperature since 1950, Journal of Geophysical Research: Atmospheres, 113, 2008.

Caesar, J., Alexander, L., and Vose, R.: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set, Journal of Geophysical Research: Atmospheres, 111, 2006.

Cai, W., Borlace, S., Lengaigne, M., Van Rensch, P., Collins, M., Vecchi, G., Timmermann, A., Santoso, A., McPhaden, M. J., Wu, L., et al.: Increasing frequency of extreme El Niño events due to greenhouse warming, Nature climate change, 4, 111, 2014.

Cavanaugh, N. R. and Shen, S. S.: Northern Hemisphere climatology and trends of statistical moments documented from GHCN-daily surface air temperature station data from 1950 to 2010, Journal of Climate, 27, 5396–5410, 2014.

Cavanaugh, N. R. and Shen, S. S.: The effects of gridding algorithms on the statistical moments and their trends of daily surface air temperature, Journal of Climate, 28, 9188–9205, 2015.

Christidis, N., Stott, P. A., and Brown, S. J.: The role of human activity in the recent warming of extremely warm daytime temperatures, Journal of Climate, 24, 1922–1930, 2011.

Clark, R. T., Brown, S. J., and Murphy, J. M.: Modeling northern hemisphere summer heat extreme changes and their uncertainties using a physics ensemble of climate sensitivity experiments, Journal of Climate, 19, 4418–4435, 2006.

Davy, R., Esau, I., Chernokulsky, A., Outten, S., and Zilitinkevich, S.: Diurnal asymmetry to the observed global warming, International Journal of Climatology, 37, 79–93, https://doi.org/10.1002/joc.4688, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.4688, 2017.

DeGaetano, A. T.: A Quality-Control Routine for Hourly Wind Observations, Journal of Atmospheric and Oceanic Technology, 14, 308–317, https://doi.org/10.1175/1520-0426(1997)014<0308:AQCRFH>2.0.CO;2, 1997.

Delaunay, B.: Sur la sphere vide, Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk, 7, 1–2, 1934.

Della-Marta, P. M., Haylock, M. R., Luterbacher, J., and Wanner, H.: Doubled length of western European summer heat waves since 1880, Journal of Geophysical Research: Atmospheres, 112, 2007.

Donat, M., Alexander, L., Yang, H., Durre, I., Vose, R., Dunn, R., Willett, K., Aguilar, E., Brunet, M., Caesar, J., et al.: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, Journal of Geophysical Research: Atmospheres, 118, 2098–2118, 2013a.

Donat, M. G. and Alexander, L. V.: The shifting probability distribution of global daytime and night-time temperatures, Geophysical Research Letters, 39, 2012.

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., and Caesar, J.: Global land-based datasets for monitoring climatic extremes, Bulletin of the American Meteorological Society, 94, 997–1006, 2013b.

Dunn, R.: HadISD version 3: Monthly Updates, Met Office Hadley Centre Technical Note, 2019.

Dunn, R., Willett, K., Thorne, P., Woolley, E., Durre, I., Dai, A., Parker, D., and Vose, R.: HadISD: a quality-controlled global synoptic report database for selected variables at long-term stations from 1973–2011, Climate of the Past, 8, 1649–1679, 2012.

Dunn, R., Willett, K., Morice, C., and Parker, D.: Pairwise homogeneity assessment of HadISD, Climate of the Past, 10, 1501–1522, 2014.

Dunn, R., Willett, K., Parker, D., and Mitchell, L.: Expanding HadISD: quality-controlled, sub-daily station data from 1931, Geoscientific Instrumentation, Methods and Data Systems, 5, 473–491, 2016.

Field, C. B., Barros, V., Stocker, T. F., and Dahe, Q.: Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change, Cambridge University Press, 2012.

Fischer, E. M. and Schär, C.: Future changes in daily summer temperature variability: driving processes and role for temperature extremes, Climate Dynamics, 33, 917, https://doi.org/10.1007/s00382-008-0473-8, https://doi.org/10.1007/s00382-008-0473-8, 2008.

Fowler, H., Ekström, M., Kilsby, C., and Jones, P.: New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate, Journal of Hydrology, 300, 212 – 233, https://doi.org/https://doi.org/10.1016/j.jhydrol.2004.06.017, http://www.sciencedirect.com/science/article/pii/S0022169404002811, 2005.

Franzke, C. L.: Local trend disparities of European minimum and maximum temperature extremes, Geophysical Research Letters, 42, 6479–6484, 2015.

Griffiths, G., Chambers, L., Haylock, M., Manton, M., Nicholls, N., Baek, H.-J., Choi, Y., Della-Marta, P., Gosai, A., Iga, N., et al.: Change in mean temperature as a predictor of extreme temperature change in the Asia–Pacific region, International Journal of Climatology, 25, 1301–1330, 2005.

Gross, M. H., Donat, M. G., Alexander, L. V., and Sisson, S. A.: The sensitivity of daily temperature variability and extremes to dataset choice, Journal of Climate, 31, 1337–1359, 2018.

Hegerl, G. C., Zwiers, F. W., Stott, P. A., and Kharin, V. V.: Detectability of anthropogenic changes in annual temperature and precipitation extremes, Journal of Climate, 17, 3683–3700, 2004.

Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B., and Stepanek, P.: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, Nature Geoscience, 4, 17, 2011.

Hosking, J. R. M.: L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics, Journal of the Royal Statistical Society: Series B (Methodological), 52, 105–124, https://doi.org/10.1111/j.2517-6161.1990.tb01775.x, https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1990.tb01775.x, 1990.

Hu, Y. and Fu, Q.: Observed poleward expansion of the Hadley circulation since 1979, Atmospheric Chemistry and Physics, 7, 5229–5236, 2007.

Huntingford, C., Jones, P. D., Livina, V. N., Lenton, T. M., and Cox, P. M.: No increase in global temperature variability despite changing regional patterns, Nature, 500, 327, 2013.

Jaeger, E. B. and Seneviratne, S. I.: Impact of soil moisture–atmosphere coupling on European climate extremes and trends in a regional climate model, Climate Dynamics, 36, 1919–1939, 2011.

Kenyon, J. and Hegerl, G. C.: Influence of modes of climate variability on global temperature extremes, Journal of Climate, 21, 3872–3889, 2008.

Kim, S. T., Cai, W., Jin, F.-F., Santoso, A., Wu, L., Guilyardi, E., and An, S.-I.: Response of El Niño sea surface temperature variability to greenhouse warming, Nature Climate Change, 4, 786, 2014.

Koenker, R. and Bassett Jr, G.: Regression quantiles, Econometrica: journal of the Econometric Society, pp. 33–50, 1978.

Koenker, R. and Hallock, K. F.: Quantile regression, Journal of economic perspectives, 15, 143–156, 2001.

Lanzante, J. R.: Resistant, Robust and Non-Parametric techniques for the analysis of Climate Data: Theory and Examples, including Applications to Historical Radiosonde Station Data, International Journal of Climatology, 16, 1197–1226, 1996.

Lenderink, G., Van Ulden, A., Van den Hurk, B., and Van Meijgaard, E.: Summertime inter-annual temperature variability in an ensemble of regional model simulations: analysis of the surface energy budget, Climatic Change, 81, 233–247, 2007.

Lewis, S. C. and King, A. D.: Evolution of mean, variance and extremes in 21st century temperatures, Weather and climate extremes, 15, 1–10, 2017.

Lu, J., Vecchi, G. A., and Reichler, T.: Expansion of the Hadley cell under global warming, Geophysical Research Letters, 34, 2007.

Lucas, C., Timbal, B., and Nguyen, H.: The expanding tropics: a critical assessment of the observational and modeling studies, Wiley Interdisciplinary Reviews: Climate Change, 5, 89–112, 2014.

Machado, J. A. F. and Silva, J. S.: Quantiles for counts, Journal of the American Statistical Association, 100, 1226–1237, 2005.

Matthews, T.: Humid heat and climate change, Progress in Physical Geography: Earth and Environment, 42, 391–405, 2018.

McKinnon, K. A., Rhines, A., Tingley, M. P., and Huybers, P.: The changing shape of Northern Hemisphere summer temperature distributions, Journal of Geophysical Research: Atmospheres, 121, 8849–8868, 2016.

McVicar, T. R., Roderick, M. L., Donohue, R. J., Li, L. T., Van Niel, T. G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N. M., et al.: Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation, Journal of Hydrology, 416, 182–205, 2012.

Menne, M. J. and Williams Jr, C. N.: Homogenization of temperature series via pairwise comparisons, Journal of Climate, 22, 1700–1717, 2009.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An overview of the global historical climatology network-daily database, Journal of Atmospheric and Oceanic Technology, 29, 897–910, 2012.

Ning, L. and Bradley, R. S.: Winter climate extremes over the northeastern United States and southeastern Canada and teleconnections with large-scale modes of climate variability, Journal of Climate, 28, 2475–2493, 2015.

Portmann, R. W., Solomon, S., and Hegerl, G. C.: Spatial and seasonal patterns in climate change, temperatures, and precipitation across the United States, Proceedings of the National Academy of Sciences, 106, 7324–7329, 2009.

Rhines, A., Tingley, M. P., McKinnon, K. A., and Huybers, P.: Decoding the precision of historical temperature observations, Quarterly Journal of the Royal Meteorological Society, 141, 2923–2933, 2015.

Rhines, A., McKinnon, K. A., Tingley, M. P., and Huybers, P.: Seasonally resolved distributional trends of North American temperatures show contraction of winter variability, Journal of Climate, 30, 1139–1157, 2017.

Roderick, M. L., Rotstayn, L. D., Farquhar, G. D., and Hobbins, M. T.: On the attribution of changing pan evaporation, Geophysical Research Letters, 34, n/a–n/a, https://doi.org/10.1029/2007GL031166, http://dx.doi.org/10.1029/2007GL031166, 2007.

Scaife, A. A., Folland, C. K., Alexander, L. V., Moberg, A., and Knight, J. R.: European climate extremes and the North Atlantic Oscillation, Journal of Climate, 21, 72–83, 2008.

Schneeweiß, H., Komlos, J., and Ahmad, A. S.: Symmetric and asymmetric rounding: a review and some new results, AStA Advances in Statistical Analysis, 94, 247–271, 2010.

Sen, P. K.: Estimates of the Regression Coefficient Based on Kendall's Tau, Journal of the American Statistical Association, 63, pp. 1379–1389, 1968.

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C.: Land–atmosphere coupling and climate change in Europe, Nature, 443, 205, 2006.

Shen, S. S., Gurung, A. B., Oh, H.-S., Shu, T., and Easterling, D. R.: The twentieth century contiguous US temperature changes indicated by daily data and higher statistical moments, Climatic Change, 109, 287–317, 2011.

Sheppard, W. F.: On the Calculation of the most Probable Values of Frequency-Constants, for Data arranged according to Equidistant Division of a Scale, Proceedings of the London Mathematical Society, 1, 353–380, 1897.

Sherwood, S. C. and Huber, M.: An adaptability limit to climate change due to heat stress, Proceedings of the National Academy of Sciences, 107, 9552–9555, 2010.

Sillmann, J. and Croci-Maspoli, M.: Present and future atmospheric blocking and its impact on European mean and extreme climate, Geophysical Research Letters, 36, 2009.

Simolo, C., Brunetti, M., Maugeri, M., and Nanni, T.: Evolution of extreme temperatures in a warming climate, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL048437, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL048437, 2011.

Smith, A., Lott, N., and Vose, R.: The integrated surface database: Recent developments and partnerships, Bulletin of the American Meteorological Society, 92, 704–708, 2011.

Stevenson, S.: Significant changes to ENSO strength and impacts in the twenty-first century: Results from CMIP5, Geophysical Research Letters, 39, 2012.

Stocker, T.: Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2014.

Sun, Y., Zhang, X., Ren, G., Zwiers, F. W., and Hu, T.: Contribution of urbanization to warming in China, Nature Climate Change, 6, 706, 2016.

Tao, L., Hu, Y., and Liu, J.: Anthropogenic forcing on the Hadley circulation in CMIP5 simulations, Climate dynamics, 46, 3337–3350, 2016.

Thorne, P., Donat, M., Dunn, R., Williams, C., Alexander, L., Caesar, J., Durre, I., Harris, I., Hausfather, Z., Jones, P., et al.: Reassessing changes in diurnal temperature range: Intercomparison and evaluation of existing global data set estimates, Journal of Geophysical Research: Atmospheres, 121, 5138–5158, 2016a.

Thorne, P., Menne, M., Williams, C., Rennie, J., Lawrimore, J., Vose, R., Peterson, T. C., Durre, I., Davy, R., Esau, I., et al.: Reassessing changes in diurnal temperature range: A new data set and characterization of data biases, Journal of Geophysical Research: Atmospheres, 121, 5115–5137, 2016b.

Thorne, P. W., Allan, R. J., Ashcroft, L., Brohan, P., Dunn, R. J. H., Menne, M. J., Pearce, P. R., Picas, J., Willett, K. M., Benoy, M., Bronnimann, S., Canziani, P. O., Coll, J., Crouthamel, R., Compo, G. P., Cuppett, D., Curley, M., Duffy, C., Gillespie, I., Guijarro, J., Jourdain, S., Kent, E. C., Kubota, H., Legg, T. P., Li, Q., Matsumoto, J., Murphy, C., Rayner, N. A., Rennie, J. J., Rustemeier, E., Slivinski, L. C., Slonosky, V., Squintu, A., Tinz, B., Valente, M. A., Walsh, S., Wang, X. L., Westcott, N., Wood, K., Woodruff, S. D., and Worley, S. J.: Toward an Integrated Set of Surface Meteorological Observations for Climate Science and Applications, Bulletin of the American Meteorological Society, 98, 2689–2702, https://doi.org/10.1175/BAMS-D-16-0165.1, https://doi.org/10.1175/BAMS-D-16-0165.1, 2017.

Vautard, R., Cattiaux, J., Yiou, P., Thépaut, J.-N., and Ciais, P.: Northern Hemisphere atmospheric stilling partly attributed to an increase in surface roughness, Nature geoscience, 3, 756, 2010.

Vega-Westhoff, B. and Sriver, R. L.: Analysis of ENSO's response to unforced variability and anthropogenic forcing using CESM, Scientific reports, 7, 18 047, 2017.

Voronoï, G.: Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les parallélloèdres primitifs., Journal für die reine und angewandte Mathematik, 134, 198–287, 1908.

Wang, F., Ge, Q., Wang, S., Li, Q., and Jones, P. D.: A new estimation of urbanization's contribution to the warming trend in China, Journal of Climate, 28, 8923–8938, 2015.

Wang, J., Tett, S., and Yan, Z.: Correcting urban bias in large-scale temperature records in China, 1980–2009, Geophysical Research Letters, 44, 401–408, 2017.

Whan, K., Zscheischler, J., Orth, R., Shongwe, M., Rahimi, M., Asare, E. O., and Seneviratne, S. I.: Impact of soil moisture on extreme maximum temperatures in Europe, Weather and Climate Extremes, 9, 57–67, 2015.

Willett, K., Dunn, R., Thorne, P., Bell, S., De Podesta, M., Parker, D., Jones, P., and Williams Jr, C.: HadISDH land surface multi-variable humidity and temperature record for climate monitoring, Climate of the Past, 10, 1983–2006, 2014.

Willett, K., Berry, D., Bosilovich, M., and Simmons, A.: Surface humidity [in "State of the Climate in 2018"], Bulletin of the American Meteorological Society, 100, S25–S27, 2019.

Zongxing, L., He, Y., Wang, P., Theakstone, W. H., An, W., Wang, X., Lu, A., Zhang, W., and Cao, W.: Changes of daily climate extremes in southwestern China during 1961–2008, Global and Planetary Change, 80, 255–272, 2012.