CHAPTER

5

Global surface temperatures

Peter Thorne

ICARUS Climate Research Centre, Department of Geography, Maynooth University, Maynooth, Ireland

1. Introduction

Global surface temperatures are a broadly used indicator of climate change and defined as an essential climate variable (ECV) by the Global Climate Observing System [2]. There is not one single "surface temperature" but rather a family of closely related but nonidentical temperatures: land surface air temperature (LSAT), land surface temperature, marine air temperature (MAT), sea surface temperature (SST), ice surface temperature, lake surface temperature, etc. [43]. Typical historical global surface temperature change estimates are derived from combining LSAT arising from fixed meteorological sites with SST estimates arising from ships and, more latterly, buoys, whereas model-based projections generally use surface air temperature everywhere—equivalent to observed LSAT and MAT.

Analyses of global surface temperatures have a rich heritage [20], with the first estimate of a globally averaged surface temperature evolution dating from over 80 years ago [7]. Over time, methods, data sources, and computational capabilities have evolved and improved. But even these pioneering efforts at global surface temperature analyses stack up well against modern-day estimates [20].

The Working Group 1 contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [26] concluded that:

> The globally averaged combined land and ocean surface temperature data as calculated by a linear trend, show a warming of 0.85 [0.65 to 1.06]°C, over the period 1880 to 2012, when multiple independently produced datasets exist. The total increase between the average of the 1850–1900 period and the 2003–12 period is 0.78 [0.72 to 0.85]°C, based on the single longest dataset available.

The statement was made as a statement of fact, with no confidence or likelihood statement attached. Such factual statements were few and far between within the assessment and reserved only for the very few most certain aspects of the science. That the world has warmed

Climate Change, *Third Edition*
https://doi.org/10.1016/B978-0-12-821575-3.00005-0

since the start of the instrumental record is unequivocal. The remainder of this chapter outlines the observational and analytical basis that underlies this statement as well as new and emerging science questions.

2. Basic data availability

The basic building block of any analysis is the original observation. But observations never have been and never will be taken continuously across the globe at high resolution or specifically to monitor climate. Rather, observations have been taken where people live and for a myriad of reasons including shipping routing, aviation, agriculture, infrastructure, energy, or just personal or professional interest. Over land, stations appear and are retired continuously, and where, why, and how the measurements have been taken has changed substantially through time. Over the ocean, with the exception of moored buoys or platforms, all measurements are vector measurements whereby the instruments themselves are constantly moving. Drifting buoys move with the currents while ships, and particularly commercial shipping, tend to follow well-worn shipping routes.

Substantial effort has been made to collate holdings of both LSAT and SST data in national, regional, and international holdings with major advances in curation of global data collections in recent years [18,46]. The percentage of both land and ocean observed in available global holdings has generally improved with time (Fig. 5.1). Recent innovations in land data curation have greatly improved access to recent observations (compare red and black traces in the top panel of Fig. 5.1). The impacts of the two world wars on marine activity are clearly evident in the SST panel. The impact is twofold. First, there was a general reduction in shipping. Second, it was rather unwise to turn on a light at night to take a measurement as that may alert unfriendly forces to your location and invite an attack.

There remains considerable scope to further improve these holdings in future. Very many pre-1950 data in particular remain in hardcopy or image form only. There have been a range of activities to rescue these data that have been increasingly successful including citizen science projects such as old weather [3] and weather rescue [21], and classroom-based approaches [50] can augment traditional paid-for-digitization approaches. Substantial efforts are being made to coordinate data rescue activities globally under the auspices of both WMO and the Copernicus Climate Change Service (https://datarescue.climate.copernicus.eu/). There are also many data in digital form that have yet to be integrated including many pre-1850 records [4]. Renewed efforts are being made to improve access to these data and integrate them into global holdings [57].

3. Analyses of land surface air temperature

The basic LSAT data holdings contain myriad data artifacts that arise from factors as diverse as follows:

- Station moves
- Instrument changes
- Observer changes
- Automation

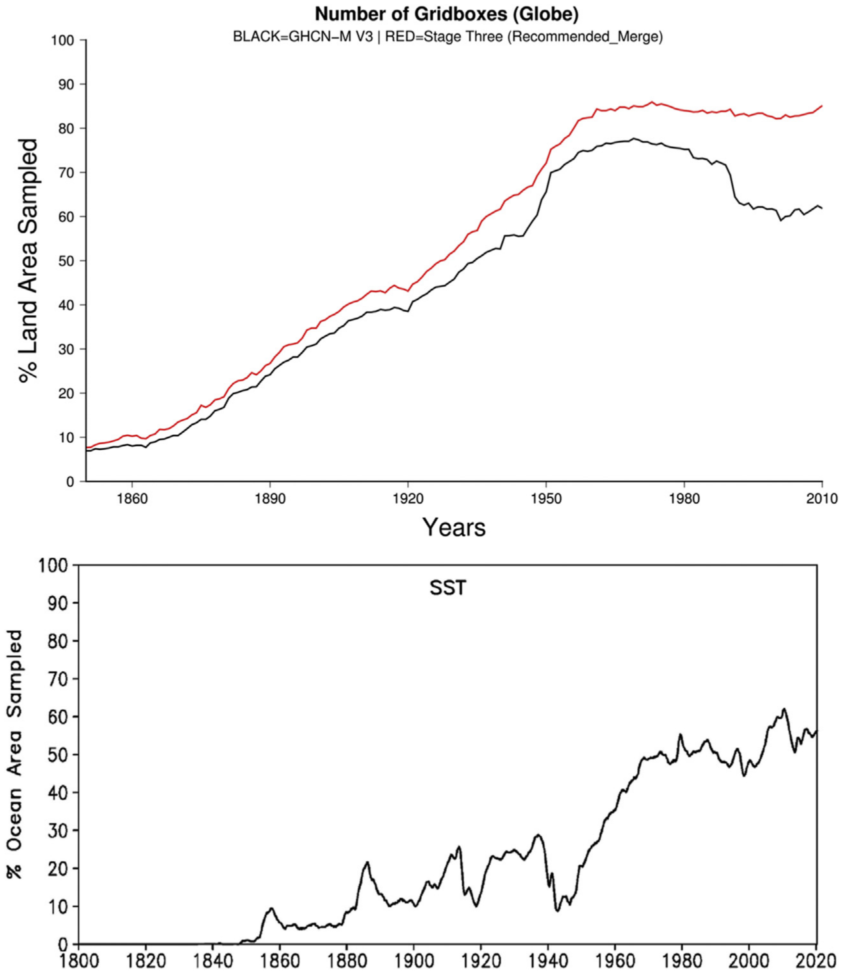

FIGURE 5.1 Change in percentage of possible sampled area for land records (top panel) showing improvement in availability arising from recent data curation efforts (note that data post 2010 continues at c.85% coverage); and marine records (lower panel). Land data come from the ISTI databank and marine data from the ICOADS in situ record. Coverage is defined as data present within a 5° grid box over land and 2° grid box over marine. *Panels courtesy Jared Rennie (CICS-NC) and Chunying Liu (Riverside Inc.), respectively.*

- Time of observation biases
- Microclimate exposure changes
- Urbanization
- And so on.

By way of illustration, an example from Reno, NV, is given in Fig. 5.2. Until the mid-1930s, the station was on the (probably white painted) roof of the P.O. building in the town. In the 1930s, with the advent of aviation, there was a need for observations at the airport so the site was moved to the airport which was a considerable distance from the town and its associated Urban Heat Island. With time, the urban area has expanded, and that encroaching heat island has led to a spurious multidecadal warming trend since the 1970s at the site. Finally, in the

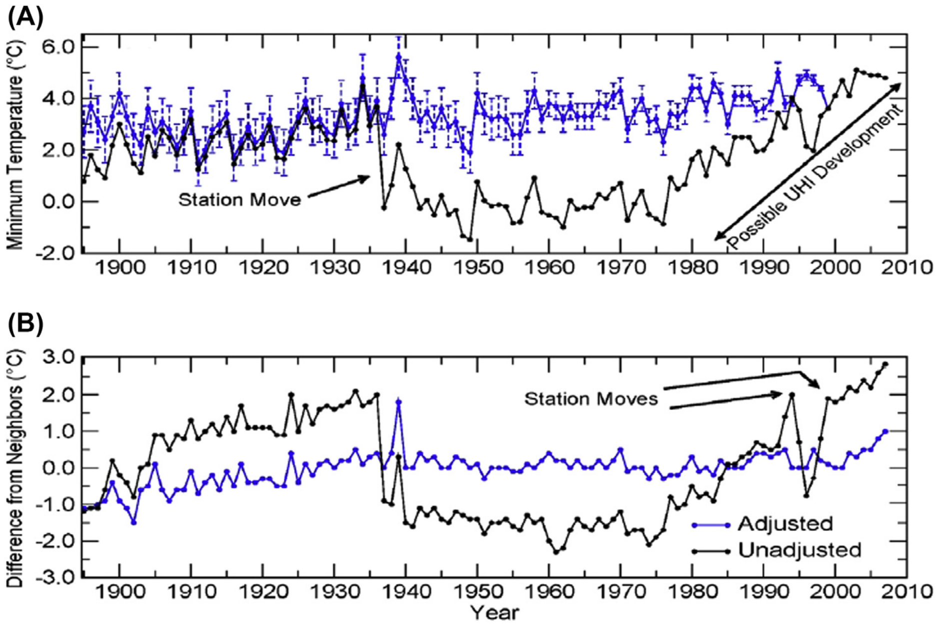

FIGURE 5.2 Example station series arising from Reno, Nevada, United States. Panel (A) is in absolute annual values. Panel (B) is in terms of differences from neighbors. The *black line* is the original series, and the *blue line* denotes the homogenized series after [40]. *M.J. Menne, C.N. Williams, Homogenization of temperature series via pairwise comparisons, J. Clim. 22 (2009) 1700–1717.*

late 1990s, they moved the station from one end of the runway to the other and then back again. There is a clear thermal gradient along the length of the runway, which caused a temporary bias until they relocated the instrument back to its original airport location.

There exist numerous national, regional, and global analyses that attempt to adjust for the nonclimatic artifacts—a process termed homogenization. In general, the regional analyses are in concordance with the global analyses and so are not considered further here [19]. There are a variety of approaches to the challenge of homogenization. Early techniques tended to consider the stations in isolation or their characteristics relative to some composite of series from neighboring stations. Consideration of a station in isolation risks misdiagnosing a real change in climate system behavior as a break in the series, thus adjusting away the real climate signal. Consideration of a neighbor composite has issues if the neighbors themselves contain biases. Such issues become critical when all or a substantial subset of the neighbors contains a common bias.

Most modern techniques utilize some form of pairwise neighbor comparison to identify the breaks [40,61]. These identify breaks in multiple candidate pairs of target minus neighbor station series and then seek to deconvolve the problem. For example, taking a network of 20 stations, it may be that a break is found to occur in 1950 in 15 of the pairwise comparisons. In 15 stations, this break is found once, and in the 16th, it occurs 15 times. In this case, it is this latter station that contains the real break. Once breaks have been attributed, they can either be adjusted [42] or the segments treated as effective stations for each homogeneous segment [49].

Several state-of-the-art homogenization techniques have been assessed against benchmark test cases [61,63]. Such test cases involve presenting to the data set creators with data which have been synthetically produced and where the data originators know what the data issues required to be found and adjusted for are [62]. Benchmarking exercises undertaken to date show that modern techniques tend to improve the consistency and "correctness" of the records but that no technique is perfect. Unsurprisingly, techniques tend to struggle when the data artifacts are small, numerous, or both. It is therefore important to understand fully the likely impacts of common changes that have occurred across the global network. To this end, a number of comparisons have been undertaken between modern and historical instrumentation at several sites [1,5].

There exist myriad ways to assess LSAT changes. Currently, there are five principal global analyses [28,36,42,49,64]. Each analysis takes a distinct approach to one or more of station selection, quality control and homogenization, interpolation, and area averaging. The use of independent approaches serves to highlight the degree of sensitivity of resulting findings to methodological choices. However, the true methodological degrees of freedom are less than the implied five given commonalities across some methods and the use of similar or identical source data. The different estimates are in broad agreement throughout the record, with differences becoming larger earlier in the record. In this early record, data sparsity increases, but also this preceded efforts to standardize temperature scales and methods of observation, which only really began around the turn of the 20th century, and so there is far greater heterogeneity in the individual station records that do exist. Differences between estimates in the global mean are substantively smaller than the long-term warming trend common to all estimates.

4. Analyses of sea surface temperature

The basic SST holdings have arisen from a broad range of measurement platforms using an array of measurement techniques that have changed substantially through time [33]. Biases in SST records are both larger and more systematic in nature than for LSAT, and hence, homogenization is essential. Measurements up until the 1940s were almost exclusively from buckets whereby a sample of the sea water from just below the surface would be hauled onto the ship deck and measured. Since World War II, there has been a preponderance for either engine room intake–based measurement or hull contact sensors. Then since the 1990s, there has been an increasing ubiquity of drifting buoys so that today approximately 90% (by number but not coverage) of all measurements arise from this method. Each of these techniques has a distinct bias relative to the true SST, and failing to account for these effects would add substantial spurious multidecadal variability to the records. Taking these in turn:

- Measurements based on buckets tended to be cold biased due to the effects of evaporative cooling that occurs between sampling the water and its subsequent measurement. Quite how cold biased depends upon the insulation efficiency of the bucket, the ship deck height, the delay between sampling and measurement, and the ambient weather conditions [17]. The effect is greatest when windy and when the atmosphere is substantially warmer or colder than the sea surface. Without accounting for these effects, pre-1942 measurements would be too cold by c.0.3°C globally averaged.
- Engine room intakes and hull contact sensors tend to sample water that has been warmed relative to the ambient temperature by the ship itself and therefore be warm-biased.

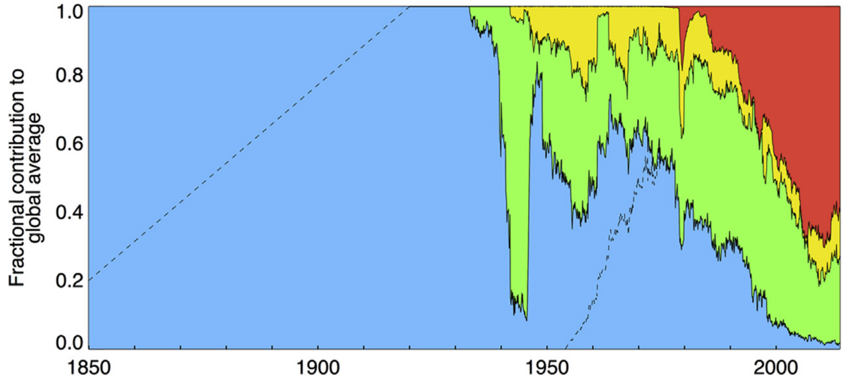

FIGURE 5.3 Best understanding of the changing mix of marine in situ observations since 1850. Blue is buoys with wooden to canvas bucket transition occurring between 1850 and 1920 (*dashed line*, approximate) and from 1954 to 1975 an uncertain switch from uninsulated to insulated buckets. Engine room intake/hull contact measures are in green. Unknown measurement type is yellow and buoys are red. The y-axis is fractional contribution to global average and not observation count. *Modified from J.J. Kennedy, N.A. Rayner, R.O. Smith, M. Saunby, D.E. Parker, Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: measurement and sampling errors, J. Geophys. Res. 116 (2011) D14103. http://ds.doi.org/10.1029/2010JD015218 and courtesy John Kennedy.*

- Drifting buoys exhibit little obvious bias and substantially smaller spread than ship-based measurements. They measure temperatures that are about 0.12–0.18°C colder than the modern ships which are mainly engine room intake or hull sensor–based measurements (Fig. 5.3).

Several global SST analyses have been undertaken which attempt to ascertain and adjust for biases to either some subset of the record or the entire record. Three analyses exist that consider global changes over the entire period of record [23,24,34]. These estimates take substantively different approaches to the problem. Despite this, they are closer to each other than they are to the original basic data on which they are based. Largest differences occur around the times of major transitions within the observing system (Fig. 5.3) or times when the observational record is dominated by ships flying under a single flag (ship-based measurement protocols are broadly dictated on a national basis). For example, in World War II, most measures arise from the US fleet which took measurements that were systematically warmer than most other nations. Having a mix of nations measure prior to 1939 and after 1945, but mainly US measurements in the middle therefore yield a potential spurious SST maximum in the early 1940s in the raw data [54].

Overall, there is a greater sensitivity to data set construction method choices in SST than there is in LSAT. Furthermore, the differences between the raw and adjusted data records are substantially larger at the global mean scales. Whereas LSAT station biases tend to cancel somewhat regionally and globally, SST records afford no such luxury.

5. Global changes

Global surface temperature data sets arise from combining underlying data sets of LSAT and SST. Choices are required as to which underlying data sets are to be merged and how, if at all, attempts are made to account for areas of missing data by interpolation. The choice of whether to interpolate or not can have a significant impact, particularly on decadal timescale behavior. Sampling is not uniform in space or in time, and many regions have never been adequately sampled (deserts, rainforests, polar regions, and regions of seasonal or perennial sea ice). If the temperatures in the unsampled regions are behaving in a way that is not represented in the remainder of the sampled portion of the globe, then a biased estimate will result [11,30,51]. It appears that over the past 20 years, in particular, interpolation has a distinct effect upon apparent global mean behavior with interpolated analyses showing greater warming in the global mean [11,30]. As of 2020, all data sets now undertake interpolation to some extent as a result of these recent insights.

There are at least seven peer-reviewed data sets that estimate global average surface temperatures from direct in situ observations [11,30,36,44,65,66]. These products combine underlying SST and LSAT data sets in different ways and/or make distinct choices in how to then calculate spatial and global averages. However, similarities in data sources and/or approaches mean that there are fewer true degrees of freedom than implied by having seven estimates. For example, these seven estimates are based upon solely two underlying SST data sets and three underlying LSAT data sets. Indeed, two families of data sets differ solely in their postprocessing of the underlying data, sharing the choice of both underlying LSAT and SST products. The different data sets broadly agree in their characterization of global mean changes on the longest timescales. But there exist substantial regional variations driven predominantly by interpolation choices in data sparse regions.

Traditional in situ products are increasingly being supplemented by reanalysis and satellite-based estimates. Reanalyses are modern data assimilation and forecast systems that are run retrospectively using available historical observations. Over successive generations, these products have become increasingly suitable for long-term climate analysis. Recent products compare very favorably with in situ products [22,52], and the ERA5 analysis is now regularly used in climate monitoring monthly statements by the Copernicus Climate Change Service. Sparse-input centennial scale reanalyses have also been shown to reasonably track much longer-term LSAT evolution [9]. Satellite data have also been increasingly used either to inform directly estimation of changes [45] or to validate in situ–based estimates [53]. Available reanalysis and satellite-based estimates build substantial confidence in the in situ records.

The global mean surface temperature has undoubtedly increased since the mid-19th century (Fig. 5.4). The change has not been linear in nature. There exist several decade-plus stretches of either little change or even cooling. This includes the early 21st century. This period had been dubbed a "hiatus" and elicited much scientific and public interest, leading to its inclusion as a box in the IPCC Fifth Assessment Report [16]. As is clear from Fig. 5.4, such periods are not atypical of the longer record [37]. At the time, this elicited much debate around the nature, causes, and implications of the feature ([8,35,39,48,59] and numerous others). This was understandable given that it raised legitimate questions around the short- to medium-term future trajectory of the system.

The suite of literature published on the matter strongly supported the assessment findings of Ref. [16] that the hiatus arose from a combination of natural climate system variability and

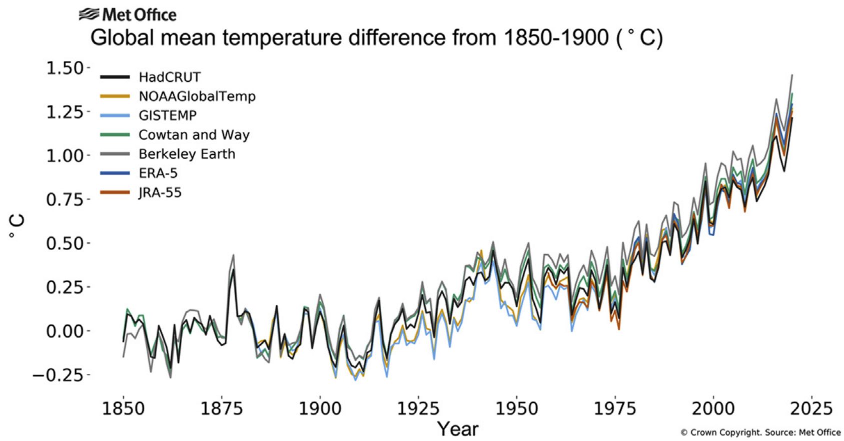

FIGURE 5.4 Global average surface temperature estimates from a range of global data sets including two reanalysis products (JRA-55 and ERA5). *Figure courtesy John Kennedy and sourced from https://www.metoffice.gov.uk/hadobs/monitoring/temperature.html. Figure British Crown Copyright 2020 provided by the Met Office Hadley Centre under Open Government License version 3.*

changes in short-lived, predominantly natural, climate system forcers. In addition, better accounting for modern data biases in SSTs arising from the ship-buoy transition [24,34] and improved station coverage over land [46] reduced the magnitude of the apparent hiatus. Regardless, since 2014, warming has resumed, and the five warmest years on record have all occurred since 2015.

In the same way that the global average change has not been linear, the global pattern of changes has not been uniform. Reasonable global spatial trend estimation is only possible since the start of the 20th century when the southern hemisphere sampling became sufficiently complete to estimate spatial anomaly patterns. It should be noted that estimation of a global average requires substantially fewer observations so long as they are well spaced. This is because anomalies in temperatures have large spatial scales. If it is unusually warm in London, the chances are that it is unusually warm also in Dublin, Edinburgh, Brussels, and Paris. In reality of the order, 150 well-spaced sites would adequately characterize the global mean LSAT [27,58], and similar density would characterize SST. But, obviously, these would not provide local information. Overall, land has warmed faster than the oceans, and the Arctic region has warmed more than any other region of the planet (Fig. 5.5).

6. Uncertainty quantification

Increasing attention is being paid to the quantification of uncertainties within surface temperature estimates. There exist several "flavors" of uncertainty [55]. The most important are structural and parametric uncertainties. Structural uncertainties arise through choices of overall method and can be quantified by comparing the estimates arising from different groups of analysts. Proper quantification would require a large ensemble of data sets that are produced

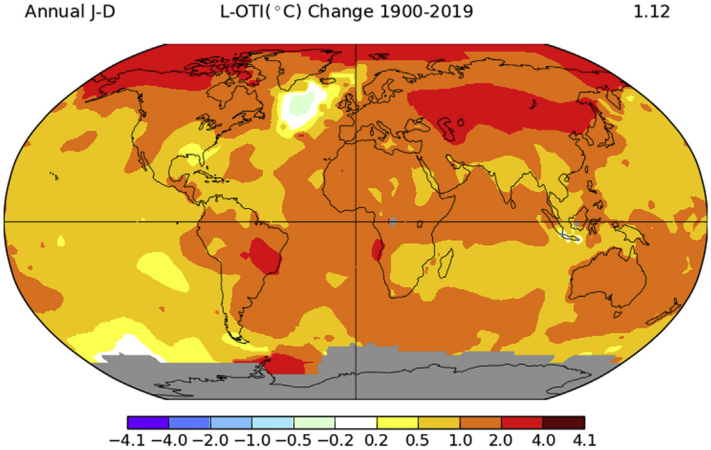

FIGURE 5.5 Spatial pattern of warming trends 1900–2019 taken from NASA GISS using the public plotter available at https://data.giss.nasa.gov/gistemp/maps/index_v4.html. The continents are warming faster than the oceans, and the Arctic is warming fastest of all. The number in the top right denotes the globally averaged trend over the period.

independently. This is only partially true for surface temperature data sets (see earlier discussion). Parametric uncertainty involves the creation of an ensemble of estimates for a given data set assessing sensitivity to uncertain choices within the methodology. For example, whether a break in a station series is assigned at 1%, 5%, or 10% significance threshold for the breakpoint detection statistical test is not a choice with an a priori correct answer. There exist numerous such semisubjective choices in all algorithms.

Several parametric uncertainty estimates have been constructed for SST [25,34,39] and [42,44,63] and LSAT [43,45,64] and combined [25,36,44]. These result in ensembles of possible realizations. Other approaches have also been applied to uncertainty quantification that does not result in ensembles (e.g., Ref. [66]). However, ensembles are intuitively appealing because they allow the expression of uncertainties at various space and timescales to support users who may be interested in more than, e.g., global and hemispheric mean timeseries.

Both structural and parametric uncertainties to the extent thus far quantified are an order of magnitude smaller than the estimated global mean changes since the start of the instrumental record of surface temperatures. It would require a substantial hitherto unrecognized source of uncertainty to be discovered to call into question the conclusion that the globe has warmed on multidecadal timescales. Further support for the conclusion that the world has warmed arises from our understanding of changes in a suite of correlated variables such as tropospheric temperatures, glacier volume, sea ice, surface humidity, and ocean heat content. A variety of data sets produced by a myriad of scientific groups for each of these covariates conclude that they are changing in the manner that would be expected if the world is indeed warming [31] (Fig. 5.6). It is a combination of the confidence in direct measurements and changes in these correlated variables that have led scientists, through the IPCC, to conclude that the warming is unequivocal.

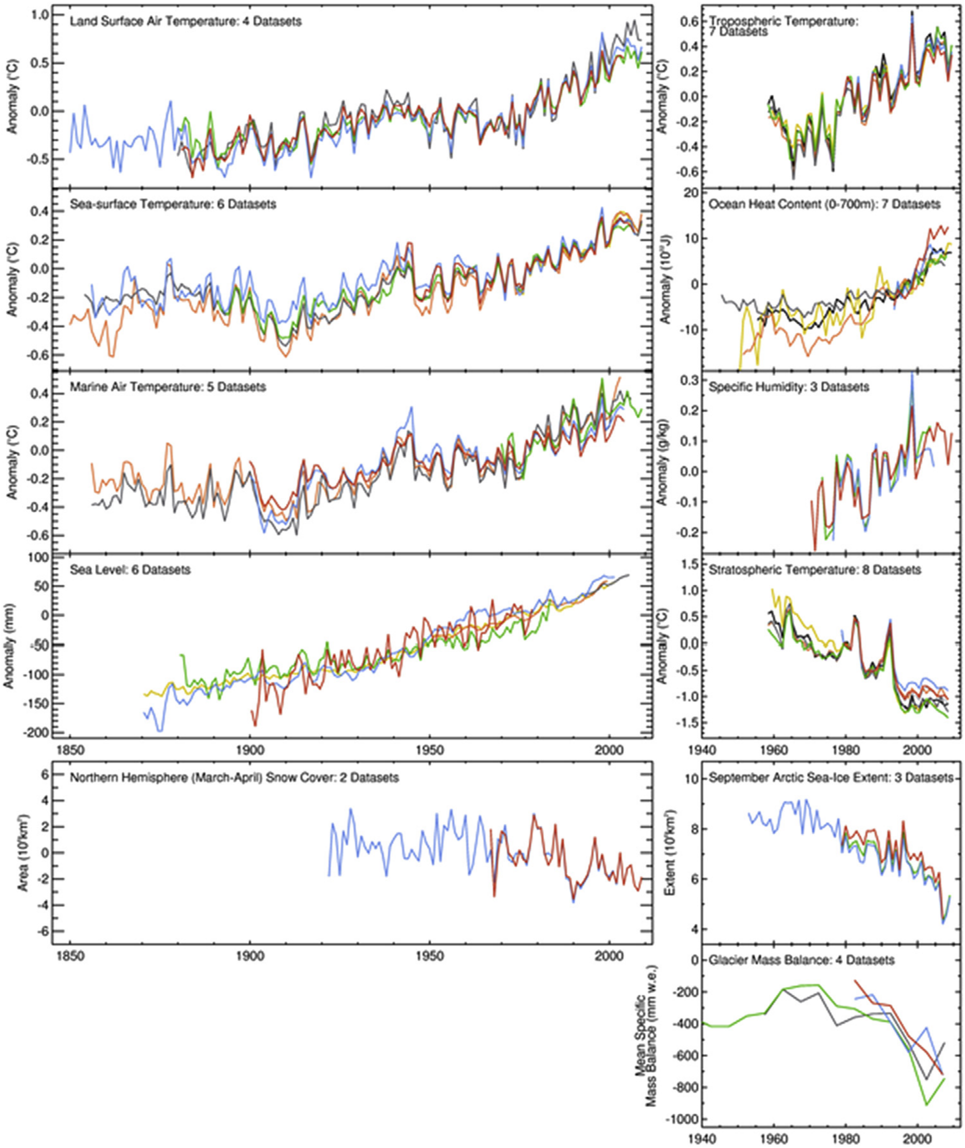

FIGURE 5.6 Changes in a suite of variables consistent with changes in surface temperatures support the contention that the globe is warming. *Figure sourced from https://www.metoffice.gov.uk/hadobs/indicators/11keyindicators.html and courtesy John Kennedy. Figure British Crown Copyright 2020 provided by the Met Office Hadley Centre under Open Government License version 3.*

7. Characterization of extremes and variability

Society experiences the weather and not the climate. Specifically, the major climatic temperature effects upon society relate to extremes of heat or cold or passing thresholds such as the need to heat or cool buildings or being able to grow certain crop types. For

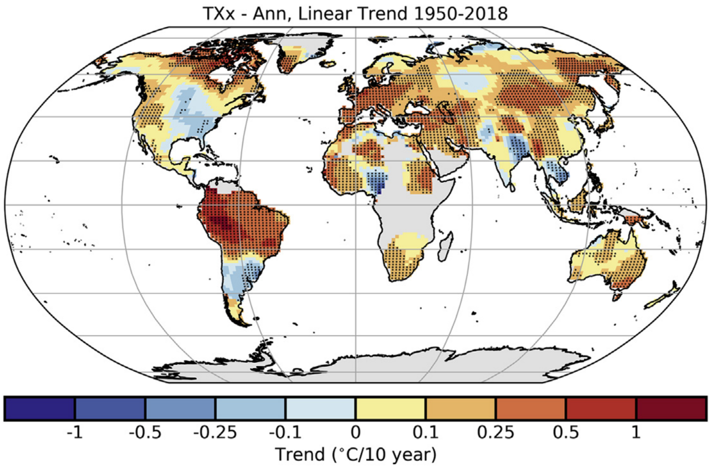

FIGURE 5.7 Trend in the annual warmest day per decade from HadEX3 over 1950 to 2018 (downloaded from metoffice.gov.uk/hadobs/hadex3). Figure British Crown Copyright 2020 provided by the Met Office Hadley Centre under Open Government License version 3.

such cases, global mean anomaly series discussed thus far in this chapter are of limited value. Several daily and subdaily data holdings exist [14,41] and are used to monitor extremes and various societally relevant indices over land [15]. The Commission for Climatology Expert Team on Climate Change Detection and Indices (ETCDDI) has defined 27 core indices of which 16 are directly related to temperatures [67]. Several data sets have been created that composite gridded trends in these ETCDDI (e.g., Ref. [6,15]). These provide useable and actionable information upon aspects such as the changing temperature of the warmest day in the year (Fig. 5.7). Overall, warm extremes have increased, and cold extremes have decreased in both frequency and severity over the period of record. This is consistent with the observed warming in the mean climate. Changes to the tails of the distribution need not follow changes in the mean, however. Daily and subdaily holdings generally do not go as far back as monthly holdings, and so information on these aspects only becomes really globally representative post-1950.

8. Future research directions

In the latter part of the 20th century, the preeminent questions were around whether the mean climate was changing, how much it was changing, and to what extent humans were responsible for those changes. Now demands as well as the expectations are different. Scientifically, it is unequivocal that the world has warmed, and it is certain that humans are primarily responsible. Society, governments, and industry require actionable information at local scales to make informed decisions. Actionable information begets openness and transparency, regional detail, and useable uncertainty estimates as well as information at daily and subdaily timescales. The global science community is addressing this need through activities such as the ETCCDI and the International Surface Temperature Initiative [56] and renewed efforts to improve land and marine data holdings. New and improved data sets as summarized in prior sections are constantly improving our understanding.

Over recent years, there has been an increasing recognition that surface temperatures actually consist of a family of related parameters [43] and in particular that SST and MAT are not equivalent measurements [12,47,52]. The MAT is estimated to warm slightly more than SST in climate models [47], with some limited support from reanalyses [52]. While there are an increasing number of night MAT estimates available [10,29,45], these estimates are less spatially complete and less scientifically mature than available SST estimates. Given uncertainties, the direct observational evidence cannot shed much light on the issue presently [29,34]. The potential nonequivalence arises an issue because analysts regularly use 2 m air temperature over all surfaces from climate models for projections. The mismatch at the join between historical estimates using SST and projections using MAT is potentially becoming increasingly large with time. Whether and if so how to account for these effects has important implications for aspects such as when warming may pass certain thresholds and how much additional carbon can be emitted while avoiding such thresholds.

There is also an increasing need to consider the homogeneity of daily or subdaily land-based data [13,60]. This is a much harder problem than homogenization of monthly timeseries, and there have been only limited efforts at benchmarking to date which have shown mixed success [34]. On monthly timescales, effects of weather on the biases tend to cancel. For homogenization at daily and subdaily timescales, it is likely that more physically based corrections are required. That is, it matters a lot more at these scales whether it was sunny, rainy, windy, or calm as to whether and if so how much to adjust individual values in the series. Initial efforts are being made to build a database of parallel measurements where old and new measurements have been taken side-by-side to better understand these covariate effects [1,5].

9. Conclusions

It is unequivocal that the global surface temperatures have warmed since the instigation of instrumental records. This change has not been linear and has varied substantially geographically. Important uncertainties and challenges remain to be addressed regarding, for example, data availability, measurement understanding, and providing high temporal resolution data suitable for many applications. These challenges may alter important aspects of our understanding of surface temperatures but are very unlikely to affect the bottom-line conclusion that the world has warmed.

References

[1] R. Bohm, P.D. Jones, J. Hiebl, D. Frank, M. Brunetti, M. Maugeri, The early instrumental warm-bias: a solution for long central European temperature series 1760–2007, Clim. Change 101 (2010) 41–67.

[2] S. Bojinski, M. Verstraete, T.C. Peterson, C. Richter, A. Simmons, M. Zemp, The concept of essential climate variables in support of climate research, applications, and policy, Bull. Am. Meteorol. Soc. 95 (2014) 1431–1443, https://doi.org/10.1175/BAMS-D-13-00047.1.

[3] P. Brohan, R. Allan, J.E. Freeman, A.M. Waple, D. Wheeler, C. Wilkinson, S. Woodruff, Marine observations of old weather, Bull. Am. Meteorol. Soc. 90 (2009) 219–230.

[4] S. Brönnimann, et al., Unlocking pre-1850 instrumental meteorological records: a global inventory, Bull. Am. Meteorol. Soc. 100 (12) (2019) ES389–ES413, https://doi.org/10.1175/BAMS-D-19-0040.1.
[5] M. Brunet, et al., The minimization of the screen bias from ancient Western Mediterranean air temperature records: an exploratory statistical analysis, Int. J. Climatol. 31 (2011) 1879–1895.
[6] J. Caesar, L. Alexander, R. Vose, Large-scale changes in observed daily maximum and minimum temperatures: creation and analysis of a new gridded data set, J. Geophys. Res. Atmos. 111 (2006) D05101.
[7] G.S. Callendar, The artificial production of carbon dioxide and its influence on temperature, Q. J. R. Meteorol. Soc. 64 (1938) 223–240, https://doi.org/10.1002/qj.49706427503.
[8] J.L. Cohen, J.C. Furtado, M. Barlow, V.A. Alexeev, J.E. Cherry, Asymmetric seasonal temperature trends, Geophys. Res. Lett. 39 (2012) L04705.
[9] G.P. Compo, P.D. Sardeshmukh, J.S. Whitaker, P. Brohan, P.D. Jones, C. McColl, Independent confirmation of global land warming without the use of station temperatures, Geophys. Res. Lett. 40 (2013) 3170–3174, https://doi.org/10.1002/grl.50425.
[10] R.C. Cornes, E.C. Kent, D.I. Berry, J.J. Kennedy, CLASSnmat: A Global Night Marine Air Temperature Dataset, 1880–2019 in press.
[11] K. Cowtan, R.G. Way, Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends, QJRMS 140 (2014) 1935–1944, https://doi.org/10.1002/qj.2297.
[12] K. Cowtan, Z. Hausfather, E. Hawkins, P. Jacobs, M.E. Mann, S.K. Miller, et al., Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures, Geophys. Res. Lett. 42 (2015) 6526–6534, https://doi.org/10.1002/2015GL064888.
[13] R.J.H. Dunn, K.M. Willett, C.P. Morice, D.E. Parker, Pairwise homogeneity assessment of HadISD, Clim. Past 10 (2014) 1501–1522, https://doi.org/10.5194/cp-10-1501-2014.
[14] R.J.H. Dunn, et al., Expanding HadISD: quality-controlled, sub-daily station data from 1931, Geosci. Instrum. Methods Data Syst. 5 (2016) 473–491.
[15] R.J.H. Dunn, et al., Development of an Updated Global Land In-Situ-Based Dataset of Temperature and Precipitation Extremes: HadEX3 JGR-A, 2020.
[16] G. Flato, J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, M. Rummukainen, Evaluation of climate models, in: T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P.M. Midgley (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.
[17] C.K. Folland, D.E. Parker, Correction of instrumental biases in historical sea-surface temperature data, Q. J. R. Meteorol. Soc. 121 (1995) 319–367.
[18] E. Freeman, S.D. Woodruff, S.J. Worley, S.J. Lubker, E.C. Kent, W.E. Angel, D.I. Berry, P. Brohan, R. Eastman, L. Gates, W. Gloeden, Z. Ji, J. Lawrimore, N.A. Rayner, G. Rosenhagen, S.R. Smith, ICOADS Release 3.0: a major update to the historical marine climate record, Int. J. Climatol. 37 (2017) 2211–2232, https://doi.org/10.1002/joc.4775.
[19] D.L. Hartmann, A.M.G. Klein Tank, M. Rusticucci, L.V. Alexander, S. Brönnimann, Y. Charabi, F.J. Dentener, E.J. Dlugokencky, D.R. Easterling, A. Kaplan, B.J. Soden, P.W. Thorne, M. Wild, P.M. Zhai, Observations: atmosphere and surface, in: T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P.M. Midgley (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.
[20] E. Hawkins, P.D. Jones, On increasing global temperatures: 75 years after Callendar, QJRMS 139 (2013) 1961, https://doi.org/10.1002/qj.2178.
[21] E. Hawkins, et al., Hourly weather observations from the Scottish Highlands (1883–1904) rescued by volunteer citizen scientists, Geosci. Data 6 (2019) 160, https://doi.org/10.1002/gdj3.79.
[22] H. Hersbach, et al., The ERA5 global reanalysis, Q J R Meteorol Soc 2020 (2020) 1–51, https://doi.org/10.1002/qj.3803.
[23] S. Hirahara, M. Ishii, Y. Fukuda, Centennial-scale sea surface temperature analysis and its uncertainty, J. Clim. 27 (2014) 57–75, https://doi.org/10.1175/JCLI-D-12-00837.1.

[24] B. Huang, et al., Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons, J. Clim. 30 (2017) 8179−8205, https://doi.org/10.1175/JCLI-D-16-0836.1.
[25] B. Huang, et al., Uncertainty estimates for sea surface temperature and land surface air temperature in NOAA-GlobalTemp version 5, J. Clim. 33 (2020) 1351−1379, https://doi.org/10.1175/JCLI-D-19-0395.1.
[26] IPCC, Summary for policymakers, in: T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P.M. Midgley (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013, pp. 1−30, https://doi.org/10.1017/CBO9781107415324.004.
[27] P.D. Jones, Land surface temperatures - is the network good enough? Clim. Change 31 (1995) 545−558.
[28] P.D. Jones, D.H. Lister, T.J. Osborn, C. Harpham, M. Salmon, C.P. Morice, Hemispheric and large-scale land-surface air temperature variations: an extensive revision and an update to 2010, J. Geophys. Res. Atmos. 117 (2012) D05127.
[29] R.A. Junod, J.R. Christy, A new compilation of globally gridded night-time marine air temperatures: the UAHNMATv1 dataset, Int. J. Climatol. (2019), https://doi.org/10.1002/joc.6354.
[30] C. Kadow, D.M. Hall, U. Ulbrich, Artificial intelligence reconstructs missing climate information, Nat. Geosci. 13 (2020) 408−413, https://doi.org/10.1038/s41561-020-0582-5.
[31] J.J. Kennedy, et al., How do we know the world has warmed?, in: State of the Climate, 2009 BAMS, 2010.
[32] J.J. Kennedy, A review of uncertainty in in situ measurements and data sets of sea surface temperature, Rev. Geophys. 52 (2014) 1−32, https://doi.org/10.1002/2013RG000434.
[33] J.J. Kennedy, N.A. Rayner, C.P. Atkinson, R.E. Killick, An ensemble data set of sea-surface temperature change from 1850: the Met Office Hadley Centre HadSST.4.0.0.0 data set, J. Geophys. Res. Atmos. 124 (2019), https://doi.org/10.1029/2018JD029867.
[34] R.E. Killick, I.T. Jolliffe, K.M. Willett, Use of generalized additive modelling techniques to create synthetic daily temperature networks for benchmarking homogenization algorithms, Dyn. Statist. Clim. Syst. 3 (2018) 1, https://doi.org/10.1093/climsys/dzz001, dzz001.
[35] Y. Kosaka, S. Xie, Recent global-warming hiatus tied to equatorial Pacific surface cooling, Nature 501 (2013) 403−407, https://doi.org/10.1038/nature12534.
[36] N. Lenssen, G. Schmidt, J. Hansen, M. Menne, A. Persin, R. Ruedy, D. Zyss, Improvements in the GISTEMP uncertainty model, J. Geophys. Res. Atmos. 124 (12) (2019) 6307−6326, https://doi.org/10.1029/2018JD029522.
[37] B. Liebmann, R.M. Dole, C. Jones, I. Blade, D. Allured, Influence of choice of time period on global surface temperature trend estimates, Bull. Am. Meteorol. Soc. 91 (2010), 1485−U1471.
[38] Liu, et al., Extended reconstructed sea surface temperature version 4 (ERSST.v4): Part II. Parametric and structural uncertainty estimation, J. Clim. 28 (2015) 931−951, https://doi.org/10.1175/JCLI-D-14-00007.1.
[39] G.A. Meehl, H. Teng, J.M. Arblaster, Climate model simulations of the observed early-200s hiatus of global warming, Nat. Clim. Change (2014), https://doi.org/10.1038/NCLIMATE2357.
[40] M.J. Menne, C.N. Williams, Homogenization of temperature series via pairwise comparisons, J. Clim. 22 (2009) 1700−1717.
[41] M.J. Menne, I. Durre, R.S. Vose, B.E. Gleason, T.G. Houston, An overview of the global historical Climatology network-daily database, J. Atmos. Ocean. Technol. 29 (2012) 897−910, https://doi.org/10.1175/JTECH-D-11-00103.1.
[42] M.J. Menne, C.N. Williams, B.E. Gleason, J.J. Rennie, J.H. Lawrimore, The global historical Climatology network monthly temperature dataset, version 4, J. Climate 31 (2018) 9835−9854, https://doi.org/10.1175/JCLI-D-18-0094.1.
[43] C.J. Merchant, S. Matthiesen, N.A. Rayner, J.J. Remedios, P.D. Jones, F. Olesen, B. Trewin, P.W. Thorne, R. Auchmann, G.K. Corlett, P.C. Guillevic, G.C. Hulley, The surface temperatures of Earth: steps towards integrated understanding of variability and change, Geosci. Instrum. Method. Data Syst. 2 (2013) 305−321, https://doi.org/10.5194/gi-2-305-2013.
[44] C.P. Morice, J.J. Kennedy, N.A. Rayner, J.P. Winn, E. Hogan, R.E. Killick, et al., submitted, An updated assessment of near-surface temperature change from 1850: the HadCRUT5 dataset. J. Geophys. Res.
[45] N.A. Rayner, et al., The EUSTACE project: delivering global, daily information on surface air temperature, Bull. Am. Meteorol. Soc. (2020), https://doi.org/10.1175/BAMS-D-19-0095.1.

[46] J.J. Rennie, et al., The international surface temperature initiative global land surface databank: monthly temperature data release description and methods, Geosci. Data J. (2014), https://doi.org/10.1002/gdj3.8. http://onlinelibrary.wiley.com/doi/10.1002/gdj3.8/full.
[47] M. Richardson, K. Cowtan, R.J. Millar, Global temperature definition affects achievement of long-term climate goals, Environ. Res. Lett. 13 (2018) 54004.
[48] J.S. Risbey, et al., Well-estimated global surface warming in climate projections selected for ENSO phase, Nat. Clim. Change (2014), https://doi.org/10.1038/nclimate2310.
[49] R. Rohde, et al., A new estimate of the average Earth surface land temperature spanning 1753 to 2011, Geoinform. Geostat. 1 (2013) 1, https://doi.org/10.4172/2327-4581.1000101.
[50] C. Ryan, et al., Integrating data rescue into the classroom, Bull. Am. Meteorol. Soc. 99 (2018) 1757−1764, https://doi.org/10.1175/BAMS-D-17-0147.1.
[51] A.J. Simmons, K.M. Willett, P.D. Jones, P.W. Thorne, D.P. Dee, Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets, J. Geophys. Res. Atmos. 115 (2010) D01110.
[52] A.J. Simmons, P. Berrisford, D.P. Dee, H. Hersbach, S. Hirahara, J.-N. Thépaut, A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets, Q. J. R. Meteorol. Soc. 143 (2017) 101−119, https://doi.org/10.1002/qj.2949.
[53] J. Susskind, et al., Recent global warming as confirmed by AIRS, Environ. Res. Lett. 14 (2019) 044030.
[54] D.W.J. Thompson, J.J. Kennedy, J.M. Wallace, P.D. Jones, A large discontinuity in the mid-twentieth century in observed global-mean surface temperature, Nature 453 (2008), 646−U5.
[55] P.W. Thorne, et al., Uncertainties in climate trends - lessons from upper-air temperature records, Bull. Am. Meteorol. Soc. 86 (10) (2005), 1437-+, http://journals.ametsoc.org/doi/abs/10.1175/BAMS-86-10-1437.
[56] P.W. Thorne, et al., Guiding the creation of a comprehensive surface temperature resource for 21st century climate science, Bull. Am. Meteorol. Soc. (2011), https://doi.org/10.1175/2011BAMS3124.1.
[57] P.W. Thorne, et al., Towards an integrated set of surface meteorological observations for climate science and applications, Bull. Am. Meteorol. Soc. 98 (2017), https://doi.org/10.1175/BAMS-D-16-0165.1.
[58] P.W. Thorne, et al., Towards a Global Land Surface Climate Fiducial Reference Measurements Network, IJOC, 2018. http://onlinelibrary.wiley.com/doi/10.1002/joc.5458/full.
[59] K.E. Trenberth, J.T. Fasullo, G. Branstator, A.S. Phillips, Seasonal aspects of the recent pause in surface warming, Nat. Clim. Change (2014), https://doi.org/10.1038/NCLIMATE2341.
[60] B. Trewin, A daily homogenized temperature data set for Australia, Int. J. Climatol. 33 (2012) 1510−1529.
[61] V.K.C. Venema, et al., Benchmarking homogenization algorithms for monthly data, Clim. Past 8 (2012) 89−115.
[62] K. Willett, et al., A framework for benchmarking of homogenisation algorithm performance on the global scale, Geosci. Instrum. Method. Data Syst. 3 (2014) 187−200, https://doi.org/10.5194/gi-3-187-2014.
[63] C.N. Williams, M.J. Menne, P.W. Thorne, Benchmarking the performance of pairwise homogenization of surface temperatures in the United States, J. Geophys. Res. Atmos. 117 (2012).
[64] W. Xu, Q. Li, P. Jones, et al., A new integrated and homogenized global monthly land surface air temperature dataset for the period since 1900, Clim. Dyn. 50 (2018) 2513−2536, https://doi.org/10.1007/s00382-017-3755-1.
[65] H.-M. Zhang, et al., Updated temperature data give a sharper view of climate trends, Eos 100 (2019), https://doi.org/10.1029/2019EO128229.
[66] R.A. Rohde, Z. Hausfather, The berkeley Earth land/ocean temperature record, Earth Syst. Sci. Data Discuss (2020), https://doi.org/10.5194/essd-2019-259 in review.
[67] T.C. Peterson, et al., Report on the Activities of the Working Group on Climate Change Detection and Related Rapporteurs 1998-2001. WMO, Rep. WCDMP-47, WMO-TD 1071, Geneve, Switzerland, 2001, p. 143.