



SCC Publishing
Michelets vei 8 B
1366 Lysaker Norway

ISSN: 2703-9072

Correspondence:

ronan@ceres-science.com

Vol. 6.3 (2026)

pp. 16-39

Adjusting the Data or Rewriting History?

The Hidden Costs of Temperature Homogenization

Ronan Connolly

Center for Environmental Research & Earth Sciences, Ireland

Abstract

Annual “hottest-year-on-record” headlines rely on global temperature-anomaly indices built from weather-station data that are incomplete, geographically biased and contaminated by non-climatic artefacts (urbanization, poor siting, instrument changes). Automated homogenization algorithms, especially NOAA’s Pairwise Homogenization Algorithm, are widely believed to overcome these challenges. However, in practice they often rewrite station records on a daily (or monthly) basis. Adjustments applied to the same record often change from one day to the next and rarely match documented station events. Meanwhile, current statistical homogenization techniques blend urban warming into rural series. The resulting homogenized network therefore converges on a more urbanized average rather than removing bias. In contrast, it is shown that a rural-only Northern Hemisphere reconstruction that avoids automated homogenization shows 38 % less warming than standard estimates and also implies a more cyclical alternation of warming and cooling periods since the late-19th century – a period that coincidentally marked the end of the Little Ice Age.

Keywords: Global warming; non-climatic biases; temperature homogenization.

Submitted 2026-09-17, Accepted 2026-09-17. <https://doi.org/10.53234/scc202608/20>

1. Introduction

Almost every year, the public is told that the year just ended was globally “the hottest on record” (Stevens, 1996; McCarthy, 1998; Bhattacharya, 2006; UN News, 2025) – sometimes even the “hottest in 100,000 years” (Associated Press, 2016; Hess, 2024). The dramatic nature of such headlines coupled with the fact that it seems to be continually recycled can lead to opposing reactions depending on the reader: for some, cynicism (“how many times has this boy cried wolf?”); for others, despair (“how many times must we hear this before something is done about it?”). However, in this article¹, we will investigate how this “global temperature record” is constructed and discover that all of these headlines were based on flawed datasets that were never fit for purpose – due to a combination of what are called “non-climatic biases” in the individual temperature records and the failure of the various “homogenization algorithms” that have been implemented in a foolhardy attempt to correct for these biases. Along the way, we will discover that these homogenization algorithms effectively “rewrite history” on a monthly or even daily basis by continually altering the temperature records of each station according to the latest iteration of various computer algorithms. We will conclude by comparing the standard “global temperature

¹ This article is adapted from a presentation given by the author at EIKE’s 17th International Climate and Energy Conference (IKEK-17), Halle (an der Saale), Germany, 26-27 June 2026. At the time of writing, this presentation can be viewed on Youtube at either the CERES-Science channel at <https://youtu.be/9mdFR04FfJM?si=oMIR1PYWTEnVEXjx> or the EIKE channel at <https://youtu.be/vH91b5toNLo?si=87gq8bjVonWd1a59>.

records” to the much less dramatic or concerning temperature changes observed when you try to reduce these non-climatic biases *without* relying on automated statistical homogenization computer algorithms.

2. The limitations of the “global temperature data”

2.1 Is there such a thing as a “global temperature”?

Beforehand, however, we should first address a surprisingly basic question that is usually overlooked in most discussions on this topic – does the Earth even have a “global temperature”? Nearly 20 years ago, Essex et al. pointed out in their article, “Does a global temperature exist?” (Essex, McKittrick and Andresen, 2007) and in a pop science book (Essex and McKittrick, 2007) that the Earth doesn’t actually have a “global temperature”. We all know this intuitively, it is warmer: in the summer than the winter; in the Sahara Desert than the Arctic tundra; at sea level versus at the top of the mountain; in the Earth’s magma versus in the stratosphere. However, Essex et al. demonstrate on physical, mathematical and observational ground that “there is no physically meaningful global temperature for the Earth” (Essex, McKittrick and Andresen, 2007). Recently, another speaker at this conference, Jonathan Cohler, has extended the arguments of Essex et al. to conclude that even the concept of a global mean surface temperature (GMST) anomaly index is “physically meaningless” and a “meaningless statistical artifact” (Cohler, 2025).

While I agree with Cohler that Essex et al.’s observations on the concept of a “global temperature” for the Earth are valid and important, I think there is still potential scientific value in the concept of a GMST anomaly index – provided it is correctly recognized that it is not a “global temperature” *per se*. Indeed, Essex et al. emphasized that “an index need not be a physical variable explicitly” and that it can potentially be scientifically useful if “the statistic usefully encapsulates a possibly large set of data into one dimension for a specific purpose” (Essex, McKittrick and Andresen, 2007). However, they also caution that, “Proposing a statistic to be an index can be risky. What is regarded as ‘useful’ may prove to be quite imprecise and subjective”.

When my late father (Dr. Michael Connolly, 1949-2025²) and I published our first analysis on the “global temperature” datasets through a series of working papers that we shared for public open peer review on our Open Peer Review Journal³ website (Connolly and Connolly, 2014c, 2014d, 2014e, 2014b, 2014a), we explicitly grappled with the issues raised by Essex et al. in the Supporting Information for (Connolly and Connolly, 2014e)⁴. We concluded that if temperature changes are experienced simultaneously across multiple parts of the world, then averaging together the temperature *changes* (as opposed to the temperatures themselves) into a regionally-averaged (and potentially, globally-averaged) index can potentially be a scientifically useful way to collectively analyse these changes.

This is typically done by converting each temperature record into an “anomaly” record of changes relative to the average temperature for that station. For instance, in Figure 1, we can see that the annual temperature changes experienced over the period 1960-2014 across the island of Ireland were remarkably similar – even though the temperatures themselves increased from the south of the island to the north, and each location experienced slight differences (Soon, Connolly and Connolly, 2015, 2024).

² <https://www.ceres-science.com/post/dr-michael-connolly-1949-2025>

³ <https://oprj.net/>

⁴ See the file called “statistics_for_gridded_analysis.pdf” in the Supporting Information files at <https://doi.org/10.6084/m9.figshare.1004125>.

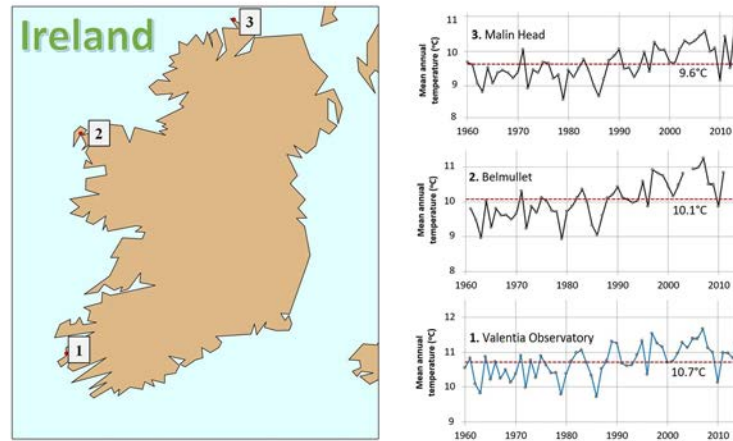


Figure 1: Comparison of the annual average temperature records for three different weather stations on the island of Ireland over a common period of reference (1960-2014). Adapted from Soon *et al.* (2015). See also Soon *et al.* (2024).

Of course, having a lot of coherence between the temperature changes across a relatively small country is quite intuitive. However, this does not necessarily mean that the correlations hold on the other side of the world. Indeed, several studies have attempted to quantify the distances over which temperature changes are correlated, e.g., (Jones, Osborn and Briffa, 1997; Christiansen and Ljungqvist, 2017). It is found that the correlations between temperature records fall off with distance, but that stations even 1000s of kilometres apart often have meaningful correlations – especially over longer time scales – decadal changes are better correlated than annual changes which are better correlated than daily changes (Jones, Osborn and Briffa, 1997). Meanwhile, it has been observed that some temperature changes seem to affect multiple regions across the world simultaneously, e.g., cooling following major volcanic events or warming during El Niño events (Chylek *et al.*, 2020).

Hence, if we remember that a globally-averaged index of local temperature anomaly records is not “a global temperature record”, it could still be useful for investigating whether similar local temperature changes might be occurring simultaneously across large regions of the planet.

As we will discuss, this optimistic desire needs to be tempered by the lack of good geographical coverage for much of the globe before the mid-20th century and the problem that most long temperature records are plagued by various non-climatic biases. In my experience, most people using “global temperature” data either have never considered these problems, or they mistakenly assume that the groups in charge of these datasets have adequately resolved the problems. Hence, establishing accurately how globally-averaged temperature anomalies have changed over the last century or longer is a much more challenging problem than most assume.

Moreover, identifying that there have been globally-averaged temperature anomaly changes over a given period, does not in itself prove the cause. Current global climate models calculate that human-caused greenhouse gas emissions should have been causing unprecedented “anthropogenic global warming” and so this seems to have led to the presumption that any long-term “global warming” trend that is identified must be human-caused. However, it is important to remember that the model-based “attribution” studies that claim to have proven this, e.g., (Gillett *et al.*, 2021), have substantially underestimated potential natural drivers of long-term climate change and overlooked the possibility of non-climatic biases in the data, see e.g., (Soon, Connolly and Connolly, 2015, 2024; Connolly *et al.*, 2021, 2023; Soon *et al.*, 2023). Indeed, temperature proxies suggest that the 19th century marked the end of a relatively cold period across much of the globe that is known as the “Little Ice Age” (Soon and Baliunas, 2003; Soon, Baliunas and Legates, 2003; Soon *et al.*, 2003; Connolly and Connolly, 2014a; Christiansen and Ljungqvist, 2017). This suggests that if there has been a long-term “global warming” since then, at least some of this could represent a “recovery from the Little Ice Age” (Akasofu, 2010).

2.2 The monthly average temperature is not the average monthly temperature

Some of the earliest weather station records were actually recorded indoors, e.g., Hohenpeißenberg (Winkler, 2009), although typically in uninsulated rooms. However, especially since the late-19th century, outdoor measurements rapidly became the standard. In order to protect the thermometers from the elements, they are normally housed in a Stevenson screen⁵ or a similar louvered shelter that blocks direct radiation and precipitation while allowing ventilation.

The classic thermometer system is actually a pair of thermometers: a maximum–minimum thermometer system (MMTS). The maximum thermometer has a marker that is pushed upwards whenever a new high is reached and the minimum thermometer has a marker pushed downwards by new lows. This means that if an observer takes the readings every day and then resets the markers for the next day, he/she can know the two temperature extremes that had been reached in the preceding 24 hours.

By convention, the arithmetic mean of this daily maximum and minimum is treated as the “daily mean” even though it is clearly not the average temperature of the day – rather, it is just the average of the two extremes of the previous 24 hours. This can also lead to a subtle “time of observation bias” because if the observation time occurs near the time of day/night when temperature extremes tend to occur, the extreme for the previous 24 hours might get doubled up for the next 24 hours (Karl *et al.*, 1986; Connolly and Connolly, 2014e). Furthermore, if the observation time is later changed to a different time, this could introduce a non-climatic bias in the record (Karl *et al.*, 1986; Connolly and Connolly, 2014e).

Nonetheless, the fact that some average of the temperature variability of the preceding 24 hours period could be recorded with one daily observation was revolutionary before the invention of modern automated electronic thermometer systems. Hence, most “daily average temperature” records are of this form. In the schematic in Figure 2, we illustrate how these daily “averages” are typically then averaged into monthly averages and then afterwards the monthly temperature records are often converted into “temperature anomalies” relative to a 30-year reference period at that station.

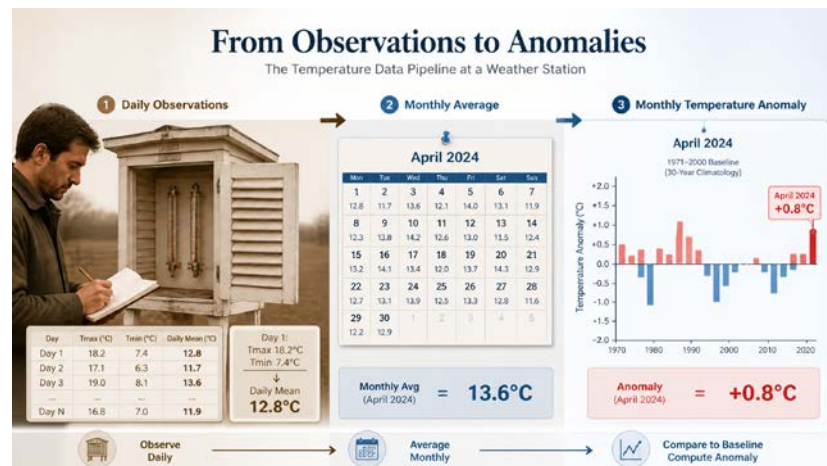


Figure 2: Schematic infographic illustrating how daily minimum and maximum temperature readings are converted into a monthly temperature anomaly series for a given weather station. This infographic was generated through iterative prompting of various AI LLMs.

These individual anomaly records then need to be averaged further to give us regional or even “global” averages. Several averaging techniques have been used, e.g., statistical kriging (Rohde and Hausfather, 2020), but one of the simpler approaches is the gridded average approach (Jones *et al.*, 1986; Jones, Wigley and Wright, 1986): The globe is divided up into grid boxes. Then for

⁵ Invented in 1864 by Thomas Stevenson, the father of the well-known author, Robert Louis Stevenson.

a given month, all of the available stations in a given grid box are averaged to yield a “gridded average”. These grid averages are then typically area-weighted and averaged further to produce regional, hemispheric or global indices.

Some readers might be wondering what should you do scientifically with the grid boxes without data? This is unclear, although in the Supporting Information for (Connolly and Connolly, 2014e), we considered several different approaches. However, a more challenging problem arises when stations are only available in many grids for some parts of the record.

2.3 Geographical biases over time

One of the main monthly temperature datasets used for calculating global land temperature anomalies is NOAA’s Global Historical Climatology Network (GHCN) (Lawrimore *et al.*, 2011; Menne *et al.*, 2018). When we first began our research on this topic, Version 3 of this dataset was available and this was an improved and updated version of the previous versions (Vose *et al.*, 1992; Peterson and Vose, 1997). Usefully, this dataset provided two different urbanization metrics (associated population and night-lights) for each station (Peterson *et al.*, 1999). Although each urbanization metric was quite crude and approximate, we realized that by combining these metrics we could distinguish between the most rural stations (rural by both metrics) and the most urban stations (urban by both metrics) and all other stations (“semi-urban”).

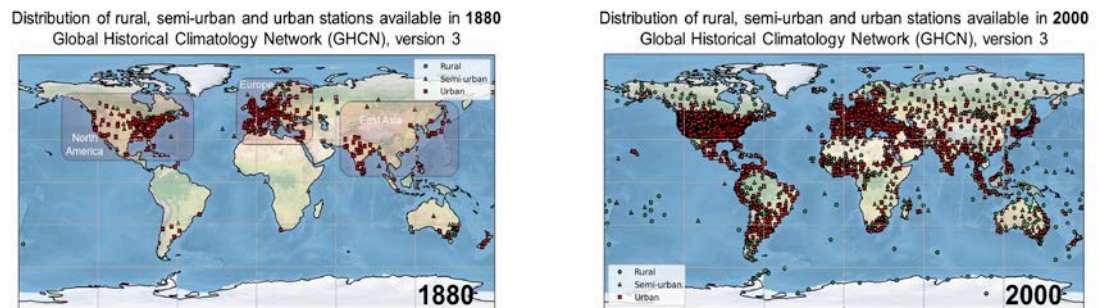


Figure 3: Maps showing the stations available in version 3 of the GHCN dataset in 1880 (left) compared to 2000 (right). Degree of “current” urbanization (at the time of Peterson *et al.* 1999) is indicated using the legend. See text for details.

From Figure 3, we can already see several problems for any groups attempting to calculate “global temperature changes” since 1880 (or earlier), i.e., the groups whose “global temperature record” were used to justify the headlines we discussed in the introduction:

- While there is reasonable geographical coverage for analysing the modern era, e.g., 2000 AD, there was almost no coverage in the southern hemisphere for the earlier years, e.g., 1880 AD.
- The northern hemisphere coverage in the earlier years was mostly confined to Europe and North America – with some East Asian coverage.
- The vast majority of the stations from either hemisphere with data in 1880 are now urbanized. As we will discuss later, urban stations tend to be warmer than the surrounding countryside due to the “urban heat island” (UHI) effect (Oke, 1973; Mills, 2008; Connolly and Connolly, 2014c; Stewart, 2019). Hence, as these key temperature stations with long records have become urbanized over the years since 1880, we would expect their temperature records to have experienced a long-term localized warming due to the growing urban heat island that is not representative of the climate outside the urban areas, i.e., an “urbanization bias” (Ren *et al.*, 2008; Connolly and Connolly, 2014c; Soon, Connolly and Connolly, 2015; Connolly *et al.*, 2021; Scafetta, 2021; Zhang *et al.*, 2021).
- While there are more rural stations available for the modern era, e.g., 2000, most of the modern stations are also urbanized to some extent.

Version 3 comprised 7200 stations, 1200 of which were part of a high-density, relatively rural, dataset specifically for the contiguous United States (i.e., the states excluding Alaska and Hawai'i) known as the United States Historical Climatology Network, USHCN (Menne, Williams and Vose, 2009; Katata, Connolly and O'Neill, 2023). However, as we saw from Figure 3, the data coverage was much less impressive in 1880. Indeed, most of the station records only began after 1950. A surprising legacy of the dataset compilation is that the sources for many of the stations (especially the rural stations) were only compiled up to 1990. This meant that there was a marked drop-off in station data before about 1950 and after 1990 (Soon, Connolly and Connolly, 2015; Soon *et al.*, 2018; Katata, Connolly and O'Neill, 2023).

When NOAA switched to Version 4 in 2018, these drop-off problems were partially reduced, but only partially. They built the dataset up independently from the previous versions – even though many of the sources used for compiling Versions 1- 3 were re-used. This means that Version 4 is not backwards compatible with the earlier versions, e.g., a completely different station ID system is now used. Unfortunately, NOAA have so far neglected to estimate the degree of urbanization of the stations in the new dataset. So, the type of analysis in Figure 3 in terms of urban subsets is not as straightforward for Version 4.

Nonetheless, analysing the spatial coverage including all stations regardless of urbanization remains straightforward. From Figure 4 we can see that while the total number of stations has nominally increased dramatically from 7,200 to ~20,000, most of the extra stations are from the contiguous United States – a region that had already been relatively well covered in Version 3. The data for the earlier years remains mostly confined to Europe, North America, and some of East Asia, and the southern hemisphere only represents a small minority of the stations for the entire record.

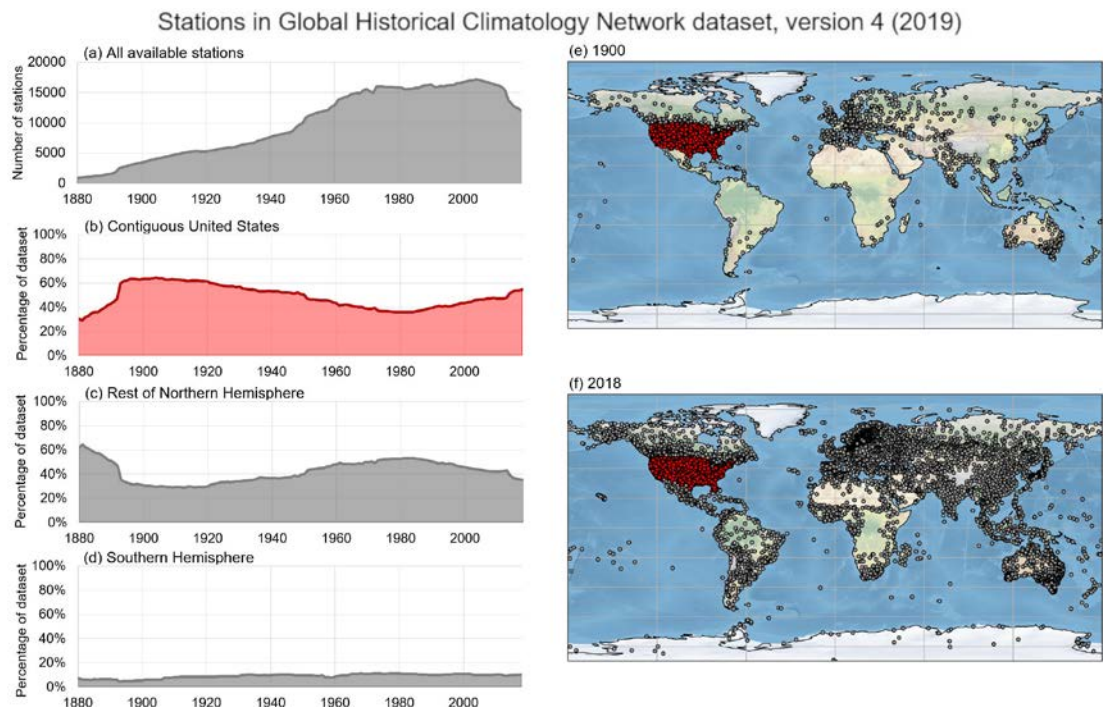


Figure 4: The stations available in version 4 of the GHCN dataset. (a) number available for each year since 1880; (b)–(d) percentage of the dataset in (b) Contiguous U.S.; (c) the rest of the northern hemisphere; (d) the southern hemisphere; maps of stations with data available for (e) 1900 or (f) 2018). Contiguous U.S. stations are shown in red.

While the infamous “1990 drop-off” associated with previous versions has been partially alleviated and is no longer obvious from the station counts in Figure 4(a), a close inspection of the

available data reveals that for many regions, this has been achieved by introducing new short station records for the post-1990 period after the longer records have dropped out – see e.g., (Soon *et al.*, 2018; Katata, Connolly and O’Neill, 2023). In some cases, if the old and new station records are suitable, it might be plausible to combine them into composite records, e.g., see (Christy, 2026) – indeed, this was done for many stations in the original USHCN dataset (Karl, Diaz and Kukla, 1988). However, this should be done with caution and awareness that it could introduce additional non-climatic biases.

So, when we see a “global temperature record” covering the period 1880 to present, 1850 to present or even longer, we need to recognize that most of the stations used for calculating the post-1950s part of the record were not available for the bulk of “the record” – just the most recent decades. We have also touched on another challenge – non-climatic biases in each of the individual temperature records.

2.4 Non-climatic biases and the bias in the biases that are considered biases

Figure 5 is a schematic infographic that provides an overview of the problem of non-climatic biases. The bottom panel shows a hypothetical temperature record that experienced multiple non-climatic events over the course of the record that altered the temperatures afterwards. For simplicity, this hypothetical record only experienced one bias at a time. However, it shows how a temperature record can be altered by many factors that have nothing to do with the regional climate: station relocation, change of instrument type or screen, change in observation time, growth of trees or buildings around the site, and urbanization of the surrounding landscape.

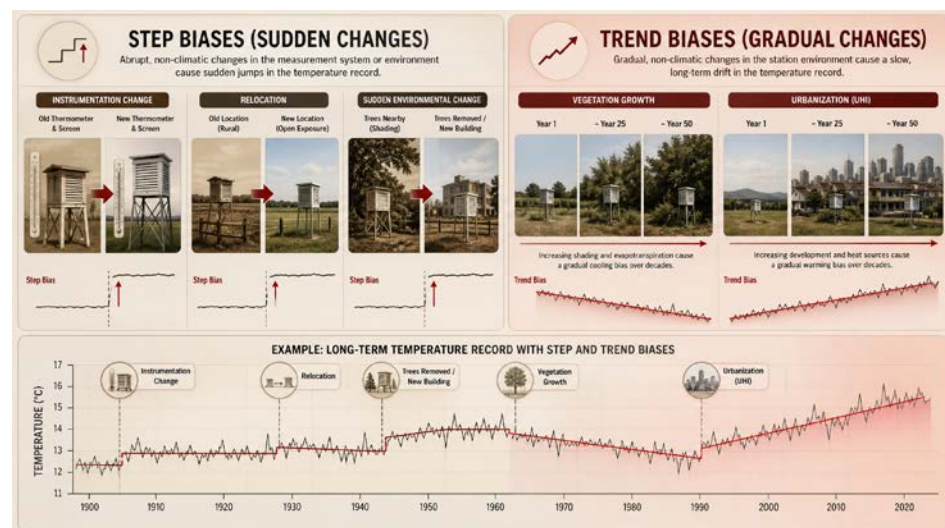


Figure 5: Schematic infographic for a hypothetical weather station record illustrating how various non-climatic factors can introduce non-climatic biases into a temperature record. This infographic was generated through iterative prompting of various AI LLMs.

In theory, these non-climatic biases could lead to either warming or cooling – or even a mixture of warming or cooling. However, anecdotally, something I have personally noticed from carefully reading the literature on non-climatic biases is that whenever the non-climatic biases lead to a long-term cooling trend, e.g., time of observation biases in the U.S. (Vose *et al.*, 2003) or both signs, e.g., some changes in thermometer systems (Quayle *et al.*, 1991; Hubbard and Lin, 2006), the groups compiling global temperature records seem to insist that these particular non-climatic biases need to be corrected for, i.e., adding extra warming to the data (Balling and Idso, 2002; Vose *et al.*, 2003; Böhm *et al.*, 2010). Yet, whenever the non-climatic biases lead to a long-term warming trend, there is considerable effort made to suggest that these non-climatic biases are trivial and/or not especially concerning (Jones *et al.*, 1990; Li *et al.*, 2004; Parker, 2006; Menne, Williams Jr. and Palecki, 2010). That is, there seems to me to be a bias in the types of biases that

are considered to be biases.

With that in mind, let us briefly discuss two of these non-climatic biases that seem to have been unfairly dismissed due to them being of the wrong sign, i.e., introducing long-term warming biases.

2.4.1 Urban warming bias

We mentioned earlier that urban areas tend to be warmer than the surrounding countryside (the urban heat island). This was already noted for London, England as early as the early-19th century (Mills, 2008). Therefore, as a station's surroundings urbanize, there is a serious risk that its record will experience a gradual, but long-term warming trend due to the growth of that UHI.

This urban warming is a real climatic change, and human-caused. Since the majority of the world's population live in urban areas, it is also a form of climate change that is experienced by many of us. Indeed, I have noticed that urban dwellers tend to be more concerned about unusual climate change than rural dwellers. But, urban warming is a localized climate change, and is not due to greenhouse gas emissions. So, if you are using urbanized stations to calculate global temperature trends, or even regional temperature trends, then this UHI growth can introduce "urbanization bias", i.e., a long-term warming that is not representative of the climatic changes that occur outside of the urban areas.

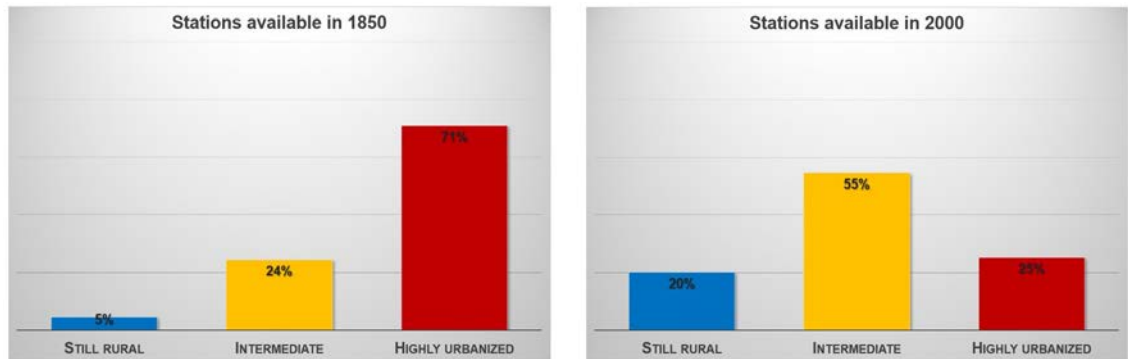


Figure 6: Breakdown of the stations in Version 3 of the Global Historical Climatology Network dataset with data available for 1850 (left) compared to 2000 (right), according to their current degree of urbanization.

We saw from Figure 3 that most of the stations with data available for the earlier parts of the record are urbanized. We can see this more starkly in Figure 6. Only 5 % of the stations in GHCN version 3 with data in 1850 are still rural, while 71 % of the stations with data in 1850 are now highly urbanized. Counterintuitively, this breakdown shows that while the rates of urbanization have accelerated in recent decades, because there has also been a considerable increase in the available rural stations, it is in many ways the earliest parts of the record that are most affected by urbanization biases (Connolly *et al.*, 2021; Soon *et al.*, 2023). This is because when looking at the temperature trends since 2000, only 25 % of the stations are highly urbanized and 20 % of the stations are still rural. But, if you are looking at the temperature trends since 1850, 71 % of the stations available to you are now highly urbanized and only 5 % are still rural.

Despite this there have been multiple papers published over the years insisting that urbanization bias is a small or negligible problem, e.g., (Jones *et al.*, 1990; Li *et al.*, 2004; Parker, 2006; Wickham *et al.*, 2013). In particular, the UN's Intergovernmental Panel on Climate Change (IPCC) reports have repeatedly asserted since their 4th Assessment Report (IPCC, 2007) to their most recent 6th Assessment Report (IPCC, 2021) that the magnitude of urbanization bias is less than 10% of the reported global temperature trends – with the implication that this is not a major concern.

However, in one of our first working papers, we investigated each of these studies in turn and

found fundamental flaws in each case that either negated their conclusion or showed their conclusion to be unjustified by their analysis (Connolly and Connolly, 2014c). In another working paper, we demonstrated that urbanization bias was a significant problem for the U.S. at least – where there was enough data for a rigorous analysis (Connolly and Connolly, 2014e). In later peer-reviewed papers, we confirmed these preliminary findings for the U.S. also apply more generally across the northern hemisphere and shown that urbanization bias is greater than the IPCC’s “<10 %” optimistic estimate (Soon, Connolly and Connolly, 2015; Soon *et al.*, 2018, 2023; Connolly *et al.*, 2021; Katata, Connolly and O’Neill, 2023). Other groups have also independently shown that urbanization bias is more substantial than the IPCC maintains, e.g., (Scafetta, 2021; Zhang *et al.*, 2021; Spencer, Christy and Braswell, 2025).

One blind spot seems to be that most of the “homogenization” algorithms used by the various groups in an attempt to correct for non-climatic biases only focus on “step change biases” and ignore “trend change biases”, e.g., (Easterling and Peterson, 1995; Menne and Williams, 2009). As we saw from Figure 5, urbanization biases are generally of the latter type.

Hence, it turns out that NASA GISS is the only major group that has attempted an explicit urbanization adjustment, using night-lights or population to classify stations and then forcing urban trends to match nearby rural trends (Hansen *et al.*, 2001, 2010). Yet, their GISTEMP global temperature record looks remarkably similar to the global temperature records of the other groups that do not include any explicit correction for urbanization bias. This appears to have convinced the other groups that correcting for urbanization bias is not especially important. However, when we downloaded and analysed all of NASA GISS’s urbanization adjustments for another of our early working papers, we discovered that their algorithm was completely flawed and was apparently never stress-tested (Connolly and Connolly, 2014d).

We found that the great majority of the station adjustments involved “urban cooling” adjustments; that the net effect on the global series was unrealistically small (especially after 1990, when urbanization should have increased), and that highly urbanized stations were left essentially unadjusted in recent decades. Another feature that might surprise many readers is that because their algorithm was designed to adjust the earlier “less urbanized” part of the records to match the most recent parts of the records, every month their algorithm applies a fresh set of adjustments to the ~50 % of their stations they identify as “rural” (Connolly and Connolly, 2014d). Hence, each month, they effectively “rewrite” the temperature history of about half of their records.

2.4.2 Siting biases – a different warming bias from urbanization bias

Urban heat islands can increase the temperatures of large regions – covering many square kilometers depending on the size of the urban areas. Some researchers have mistakenly assumed that if a station is located within an urban park, for instance, that this would remove the problem (Peterson, 2003). However, even if an urban park is relatively cool compared to the non-green areas (“the urban park cooling effect”), it can still be warmer than outside of the urban area (Gaffin *et al.*, 2008). Indeed, (Stewart and Oke, 2012) have cautioned that many “rural/urban” comparative studies are really urban parks vs. urban non-green areas.

We should avoid conflating the large-scale urban heat island effects with local climate differences associated with the station’s microclimate, i.e., the local environment within 30-100 metres of the station. However, it turns out that non-climatic biases can also arise from changes in the microclimate. From 2007 - 2011, the “Surfacestations” project, a voluntary “crowd-sourced” survey of the microclimate of most of the stations in the U.S. Historical Climatology Network revealed that a large majority of these stations fail standard siting requirements. Poorly sited stations included those located in parking lots, near air-conditioning exhausts, surrounded by asphalt, or even on top of roofs. When the temperature trends of the worst-sited stations were compared to the best-sited stations, it became apparent that poor siting typically leads to warming biases relative to good siting (Fall *et al.*, 2011; Connolly and Connolly, 2014b).

A high profile study by the NOAA group in charge of the USHCN conceded this, but argued that

after NOAA had applied their automated statistical Pairwise Homogenization Algorithm (PHA) to the data, there was very little difference between the badly-sited and well-sited stations (Menne, Williams Jr. and Palecki, 2010). They argued that this meant their homogenization algorithm was automatically able to correct for these non-climatic biases. A study involving several of the organizers of the survey conceded that the homogenized trends were similar for all stations, but argued that even after homogenization there were non-climatic biases in the differences between the maximum and minimum temperatures – suggesting that the homogenization algorithm was not entirely removing the biases (Fall *et al.*, 2011).

In another of our working papers, we independently evaluated the results from this Surfacestations project (Connolly and Connolly, 2014b). The key results are reproduced here in Figure 7. We confirmed that the warming trend associated with each subset of the stations generally increased the worse the siting quality (top panel). Although correcting for time-of-observation biases seemed to add some long-term warming, it also seemed to make the siting biases more pronounced (middle panel). However, while we nominally confirmed (Menne, Williams Jr. and Palecki, 2010)’s finding that the trends were similar among all subsets after homogenization, we showed that this arose by adding extra warming to the well-sited stations and slightly removing warming from the worst-sited stations until all stations had similar trends to the most common subset, i.e., the poor quality stations that made up 64 % of the network (bottom panel).

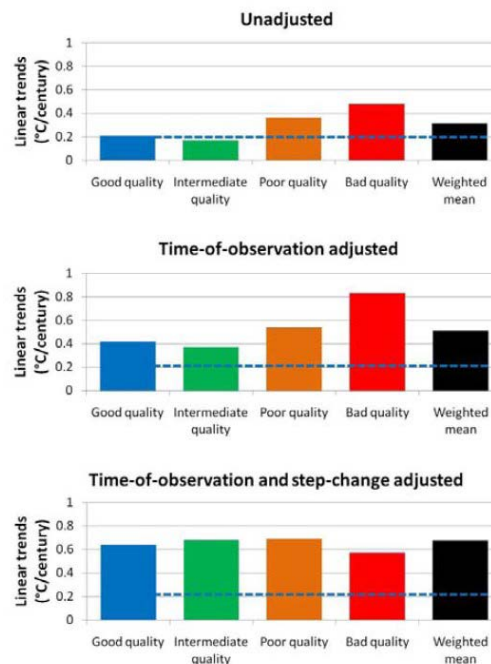


Figure 7: Comparison of the 1895-2011 linear temperature trends of each of the subsets of USHCN stations sorted by the Surfacestations team according to their siting quality. Top panel shows the trends for the raw unadjusted data; middle panel is after applying time-of-observation bias adjustments; bottom panel has also had NOAA’s PHA adjustments applied. Adapted from Connolly & Connolly (2014b).

That is, we showed that NOAA’s homogenization adjustments were effectively spreading the siting biases between all stations until they all showed similar trends to the poor quality stations.

While our analysis was strictly confined to the USHCN dataset since those were the stations that had been surveyed, preliminary surveys of other stations around the world at the time were suggesting that poor siting was a worldwide problem (Connolly and Connolly, 2014b). But, let us now consider in more detail these homogenization adjustments.

3. The current “solution”: Automated statistical homogenization

The principle behind temperature homogenization is the recognition that multi-decade long temperature records are likely to contain multiple non-climatic biases, e.g., some of those described in Figure 5. From a philosophical perspective, some scientists might argue that it is better to analyse these records as they were – despite the biases – since those were the measurements that were actually recorded, e.g., (Courtillot *et al.*, 2008). Indeed, if the biases are equally likely to be of either sign and occur at relatively random intervals for each station, then one might argue that “the random component of such errors tends to average out in large area averages and in calculations of temperature change over long periods” (Hansen *et al.*, 1999).

Other scientists (including myself) would argue that if the observers noted that they had made a systematic change to how they were measuring temperature after a specific date and that this altered all subsequent temperature readings, then carefully applying a suitable correction to the station’s temperature record can be scientifically justifiable and improve the reliability of your subsequent analysis.

For example, for (Soon, Connolly and Connolly, 2015), when analysing the station record for Valentia Observatory, Ireland – a significant record in that it is one of the longest and most complete stations (1869–present) that still remains relatively rural – we visited the station and consulted the station observers who provided us with “station history metadata” that documented known changes in how the temperatures were recorded over the preceding 145 years. For the two most recent events (a station move in 2001 followed by a change in 2012 to an automated electronic thermometer system), the observers also kindly provided us with 2 years of parallel measurements taken before and after the change. This allowed us to calculate that the 2001 station move led to a slight cooling bias, but that the 2012 instrument change gave almost identical readings (Soon, Connolly and Connolly, 2015).

I would also argue that if we know that some non-climatic biases are systematically of one sign and/or are not randomly distributed, that we should take this into account in our analysis. For instance, we discussed above how most of the longest station records available are now highly urbanized and therefore possibly contaminated by a long-term urban warming bias (Soon, Connolly and Connolly, 2015; Connolly *et al.*, 2021; Soon *et al.*, 2023). Similarly, the Surfacestations project revealed that the majority of stations (in the USHCN at least) are affected by siting biases due to poor siting (Fall *et al.*, 2011; Connolly and Connolly, 2014b).

Yet, what should we do when we suspect that there are non-climatic biases in a station’s temperature record but we do not have much or any relevant station history metadata? Or what do we do when we have station history metadata documenting a significant non-climatic change but we do not have any parallel measurements to quantify what bias, if any, this introduced?

The solution that most of the groups involved in compiling regional or global temperature series have converged on is to use statistical homogenization algorithms to statistically identify non-climatic jumps (“breakpoints”) in each station record and to apply corrections to the records to remove these jumps (Easterling and Peterson, 1995; Reeves *et al.*, 2007; Menne and Williams, 2009; Skrynyk *et al.*, 2019; Domonkos, 2021). I should note that researchers homogenizing regional temperature datasets will often combine their analysis explicitly with station history metadata, e.g., (Skrynyk *et al.*, 2019), but most of the groups generating global temperature records rarely even consider the use of station history metadata (Jones, Wigley and Wright, 1986; Hansen *et al.*, 2010; Menne *et al.*, 2018; Rohde and Hausfather, 2020; Osborn *et al.*, 2021).

3.1 What is temperature homogenization?

The infographic in Figure 8 describes the general outline of a typical temperature homogenization algorithm, e.g., NOAA’s Pairwise Homogenization Algorithm (PHA) that is currently used to homogenize the GHCN dataset. The exact algorithm varies between different groups, e.g., (Easterling and Peterson, 1995; Reeves *et al.*, 2007; Menne and Williams, 2009; Skrynyk *et al.*, 2019;

Domonkos, 2021). However, the basic principle is usually something similar to the following:

- The monthly temperature record of each station is analysed by comparison with the temperature records of neighbouring stations. In NOAA’s original homogenization algorithm for the GHCN dataset, this was done by constructing a composite reference series that was the average of five neighbours (Easterling and Peterson, 1995). The current method compares each station iteratively to up to 100 neighbours, one pair at a time (Menne and Williams, 2009).
- A temperature difference series is created between the target station’s record and its reference record.
- A statistical test for any abrupt jumps in the difference series is applied and any jumps that pass the test are considered as potential “breakpoints”, i.e., non-climatic biases.
- If a breakpoint is identified as (a) being statistically significant and (b) associated with the target series as opposed to the reference series, then it is assumed to be a non-climatic bias. The sign and magnitude of this bias is determined by comparing the values before and after the breakpoint.
- If station history metadata are used, then either (a) only breakpoints that coincide with a documented event are adjusted, or (b) breakpoints that are associated with a documented event are given a lower threshold to be considered as requiring adjustment. Note that most of the global temperature groups rarely, if ever, use station history metadata.

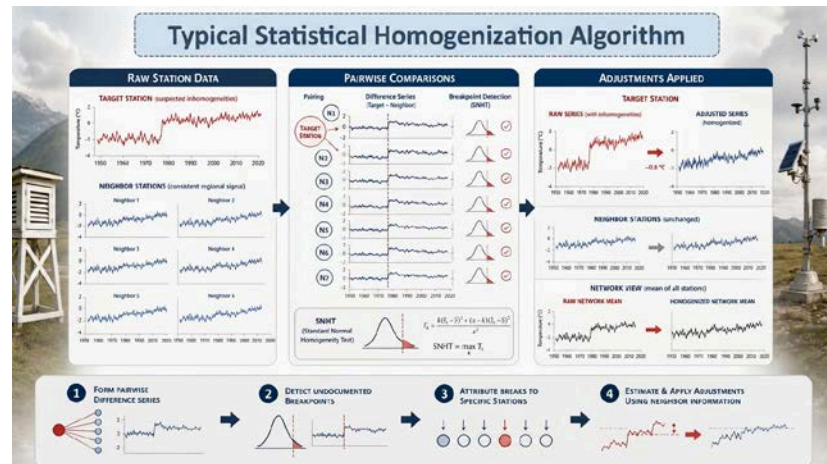


Figure 8: Schematic infographic illustrating the basis and workflow for a typical automated statistical temperature homogenization algorithm. This infographic was generated through iterative prompting of various AI LLMs.

In principle, the adjustments could be applied to either (a) bring the most recent temperatures into line with the temperatures at the start of the record, or (b) bring the earliest temperatures into line with the temperatures at the end of the record. The former approach means that all new monthly temperature values will need to be adjusted, while the oldest measurements will remain constant. However, the latter approach means that every time there is a new data point added to the record, the entire record might have to be recalculated so that the earliest data points match the latest. Since these temperature records are typically then converted into temperature anomaly records, the choice is often somewhat arbitrary. But, philosophically, the latter approach effectively means that the earlier parts of the homogenized temperature record are continually being rewritten whenever new breakpoints arise.

One thing that has been repeatedly confirmed is that after such a temperature homogenization algorithm is applied to a temperature dataset, the temperature records all converge to having similar trends to their neighbours (Menne, Williams Jr. and Palecki, 2010; Fall *et al.*, 2011; Hausfather *et al.*, 2013; Connolly and Connolly, 2014e; Katata, Connolly and O’Neill, 2023). In other words, the temperature records become much more “homogeneous” after homogenization. However, does that mean that the homogenization adjustments have correctly identified and removed

the non-climatic biases from the data? Or are the homogenization adjustments merely making the station records more homogeneous with respect to each other?

Until recently, most of the attempts to address this have focused on assessing how well the algorithms perform when applied to synthetic temperature records that have had artificial biases added to them. These are often referred to as “benchmarking” experiments (Venema *et al.*, 2012; Williams, Menne and Thorne, 2012). These experiments are typically carried out as collaborations where one researcher will synthesise a dataset of multiple physically plausible temperature records and then artificially introduce breakpoints into a large number of the records. The homogenization team will be simply handed the dataset and asked to use their algorithm to identify and correct any non-climatic biases. The other researcher then compares the proposed homogenization adjustments to their secret list of actual breakpoints. He/she then calculates how many true positives (correctly removing a bias), false positives (incorrectly applying an adjustment when none was needed), and false negatives (failing to identify a bias). Such synthetic experiments have shown that NOAA’s PHA, and various other homogenization algorithms, are able to successfully homogenize these synthetic records with high true positives and low false positives and negatives, e.g., (Venema *et al.*, 2012; Williams, Menne and Thorne, 2012).

This seems to have convinced many in the homogenization community that the algorithms are doing a good job and improving the accuracy of the temperature records (Reeves *et al.*, 2007; Venema *et al.*, 2012; Menne *et al.*, 2018; Lund *et al.*, 2023). However, how do these algorithms fare when applied to real temperature records?

3.2 Does it work on real temperature data?

NOAA re-run their PHA on their GHCN datasets roughly once per day. They overwrite the older version with the new one. In 2015, we noticed when we compared the homogenized data that we had downloaded on different days, that the homogenization adjustments that NOAA applies to each station often changed radically (Soon, Connolly and Connolly, 2015). We also noticed that the homogenization adjustments NOAA was applying to the Valentia Observatory record we had specifically analysed had no match to the station history metadata we had been provided (Soon, Connolly and Connolly, 2015).

Afterwards we discovered that a colleague, Peter O’Neill, had independently noticed the inconsistency in NOAA’s homogenization adjustments and had begun archiving the GHCN datasets from 2011 – initially intermittently, but eventually on a daily basis. We also noticed that, in recent years, there has been considerable effort placed into digitizing station history metadata for many stations, e.g., (Przybylak, 2010; Brázdil *et al.*, 2012; Wypych, Ustrnul and Schmatz, 2018; Skrynyk *et al.*, 2019; Chimani *et al.*, 2020).

In (O’Neill *et al.*, 2022), we decided to combine these two efforts to carry out a more practical assessment of NOAA’s PHA adjustments. We collaborated with some of the groups compiling station history metadata in Europe and also identified any other published station history metadata we could find in the literature or public archives. In total, we found station history metadata for more than 800 European stations. We then analysed O’Neill’s archive to identify all the different breakpoints NOAA had applied to these stations over the ~3,600 different GHCN datasets he had downloaded.

The key results are summarized in Figure 9:

- Only about 17 % of the adjustments to a given station were applied consistently.
- Only about 19 % of the breakpoints coincided with a documented station event.

We would not have been surprised if the algorithm was identifying some undocumented events, but it was surprising that only a small minority of the adjustments coincided with documented events. However, an especially concerning aspect of their adjustments was that the adjustments applied to a given station often changed wildly from one day to the next. This means that, for example, a researcher who downloaded the dataset on Tuesday might have a completely different

temperature record for a station than their colleague who had downloaded the dataset on the previous Friday.

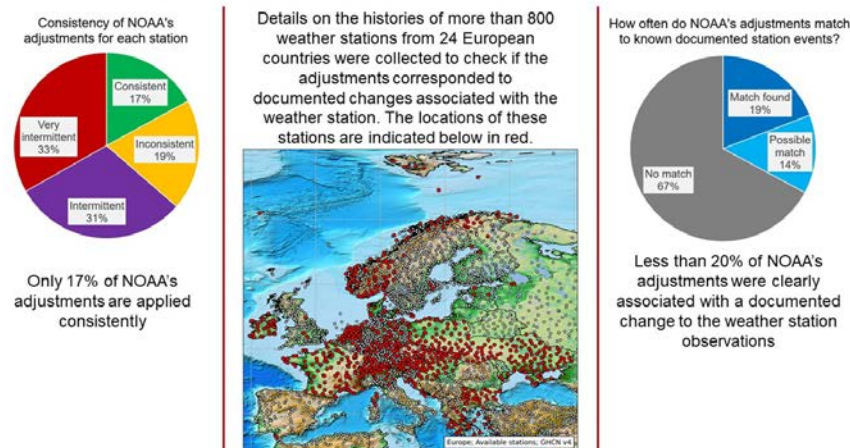


Figure 9: Summary of the key findings of our analysis in O'Neill *et al.* (2022). The left panel indicates how consistently a given set of homogenization adjustments were applied to each station. The middle panel shows the location of the stations used for our analysis. The right panel indicates how often a homogenization adjustment coincided with a documented station event (“Match found”) or occurred within 1-3 years of a documented event (“Possible match”).

The findings of (O'Neill *et al.*, 2022) revealed that NOAA's PHA adjustments rarely coincide with documented station history events and are remarkably inconsistent – despite the algorithm performing very well in the synthetic benchmarking experiments (Venema *et al.*, 2012; Williams, Menne and Thorne, 2012). This is important since NOAA's PHA-adjusted version of the GHCN dataset is the main dataset used for NOAA, NASA and JMA's global temperature series and a secondary dataset for CMA and CRU's equivalents. However, in a separate study, we demonstrate that there is a more systemic statistical flaw that is common to all of the current temperature homogenization algorithms, i.e., “urban blending”.

3.3 The urban blending problem

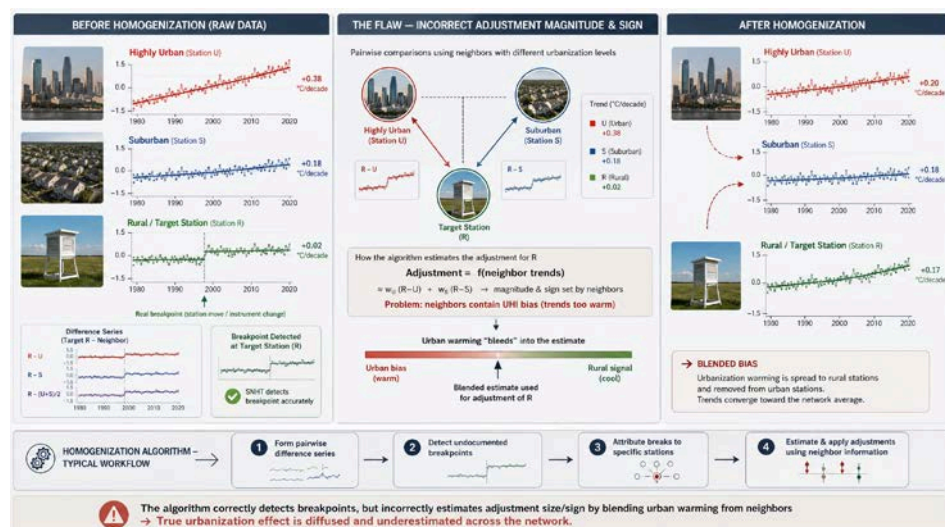


Figure 10: Schematic infographic illustrating how current temperature homogenization algorithms lead to urban blending as currently implemented. This infographic was generated through iterative prompting of various AI LLMs.

In (Katata, Connolly and O'Neill, 2023), we showed that all of the current temperature

Science of Climate Change

<https://scienceofclimatechange.org>

homogenization algorithms are prone to a statistical problem that we call “urban blending. The schematic in Figure 10 summarises the essential problem:

- When a breakpoint is identified in a station, the adjustments applied to “correct” for the apparent non-climatic bias are calculated by comparing the difference series before and after the breakpoint.
- However, if the neighbours used for calculating the difference series are more urbanized than the target station, the calculated adjustments may incorporate any urbanization bias experienced by that neighbour over the comparison period.
- This means that during homogenization, rural stations will tend to have urban warming added to their records to better match the trends of its more urbanized neighbours.
- The opposite process also occurs for the most urbanized stations, i.e., if an urban station experienced more urban warming around the time of a breakpoint than its neighbours, the homogenization adjustment will have additional cooling added.
- The net result is that homogenization will tend to spread urban warming biases among all stations until all stations in a region have about the same degree of urban warming to the average of the neighbours in that region.
- This process can potentially occur every time a breakpoint is identified in the target station and it is independent of whether the breakpoint is a genuine non-climatic bias or not. So, the more breakpoints are identified in a dataset, the more widespread this urban blending.

As well as demonstrating the theoretical reason for this phenomenon, (Katata, Connolly and O’Neill, 2023) specifically evaluated the extent of this effect for two different countries: Japan and United States, and confirmed that it is real and substantial in both cases. This explains why so many studies had incorrectly concluded that homogenization was removing urbanization bias, e.g., (Li *et al.*, 2004; Menne and Williams, 2009). After homogenization, rural and urban stations all show similar trends, but not because the urbanization biases had been removed. Rather, the homogenization process leads to urban blending.

We noted that a similar process could also potentially apply to other non-climatic biases that are common to multiple neighbours, such as siting biases. This therefore explains how the siting biases also get blended between stations after homogenization (Menne, Williams Jr. and Palecki, 2010; Fall *et al.*, 2011; Connolly and Connolly, 2014b).

We suggested two potential ways that current homogenization algorithms could be adapted to reduce the urban blending problem:

1. Stations could be homogenized only using neighbours that have a similar degree of urbanization, i.e., using rural neighbours to homogenize rural stations and urban neighbours to homogenize urban stations.
2. Not calculating the homogenization adjustments from the difference series, but assuming that any temperature jump in the target station’s record after the breakpoint was a non-climatic bias.

In principle, all stations could potentially be used for identifying the breakpoints. The key issue is whether to use them for also calculating the sign and magnitude of the bias.

The first method has the disadvantage that if there are a lot of additional non-climatic biases in the dataset, e.g., siting biases, they might potentially be blended during homogenization. However, the second method risks conflating any genuine temperature jumps that might have coincidentally occurred during the breakpoint as “non-climatic”.

4. Bypassing the above problems: our rural Northern Hemisphere temperature reconstruction

Alternatively, another approach might be to not rely on these automated statistical homogenization algorithms. This was a path we had also worked on for more than six years (Soon, Connolly and Connolly, 2015; Connolly, Connolly and Soon, 2017; Soon *et al.*, 2018, 2019; Connolly *et al.*, 2021) yielding results that led to insightful collaborations with dozens of researchers (Connolly *et al.*, 2023; Soon *et al.*, 2023).

4.1 Using GHCN version 3 to reconstruct rural temperatures from 1850-2018

Rather than rely on automated homogenization we identified the longest, most rural stations in GHCN version 3, applied only well-documented, metadata-supported corrections, and constructed a Northern Hemisphere rural-only land series (Connolly *et al.*, 2021; Soon *et al.*, 2023).

In Figure 11, we first compare the conventional “all-stations” (urban-plus-rural, homogenized) series we get without taking these extra steps to the equivalent time series for the other main groups. Our series (solid black) agrees with all of the other estimates. This suggests that any other differences between our methodology and those of the other groups are relatively trivial (Connolly *et al.*, 2021).

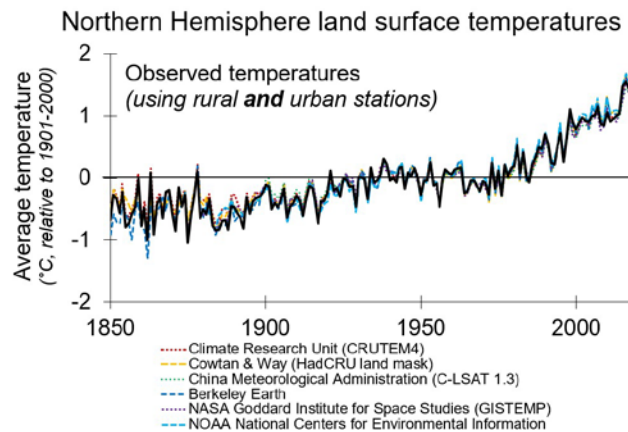


Figure 11: Comparison of the Northern Hemisphere land surface temperature series our methodology produces if we include the homogenized versions of all stations (urban and rural) – black solid line – to the equivalent series by other groups. Adapted from Soon *et al.* (2023). See Connolly *et al.* (2021) for details on each temperature series.

Yet, we can see from Figure 12 overleaf that our rural-only approach (bottom panel) leads to noticeably different results from the standard approach of using automated homogenization and all stations regardless of urbanization. For instance, the long-term (1850–2018) linear trend is approximately $0.55\text{ }^{\circ}\text{C century}^{-1}$, compared with $0.89\text{ }^{\circ}\text{C century}^{-1}$ for the standard all-station homogenized series. This implies urbanization bias and/or homogenization artefacts contributed up to 38 % of the apparent warming in the standard version. Meanwhile, the rural series suggests an almost cyclical variation in temperatures over the course of the record – warming from the late-19th century until the 1940s; cooling until the 1970s; warming since the 1980s.

4.2 How does our reconstruction compare to other temperature reconstructions?

In Figure 13 overleaf, we compare our rural-only series to various other independent estimates of Northern Hemisphere temperature changes since 1850 (and a few other climate change indicators). Surprisingly, while each estimate is constructed from very different data sources, they each imply qualitatively similar trends to the rural-only record (with the exception of worldwide precipitation which does not show much trend if any), i.e., they all are consistent with a general warming from the end of the 19th century to the 1940s, followed by a general cooling until the 1970s, followed by a general warming to present.

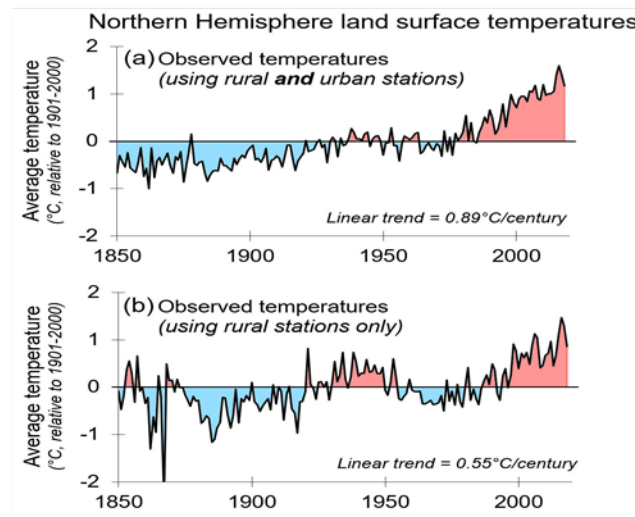


Figure 12: Comparison of the Northern Hemisphere land surface temperature series our methodology produces if we include (a) the homogenized versions of all stations (urban and rural) vs. (b) our rural-only series. Adapted from Soon et al. (2023). See Connolly et al. (2021) for details on how each temperature series was constructed.

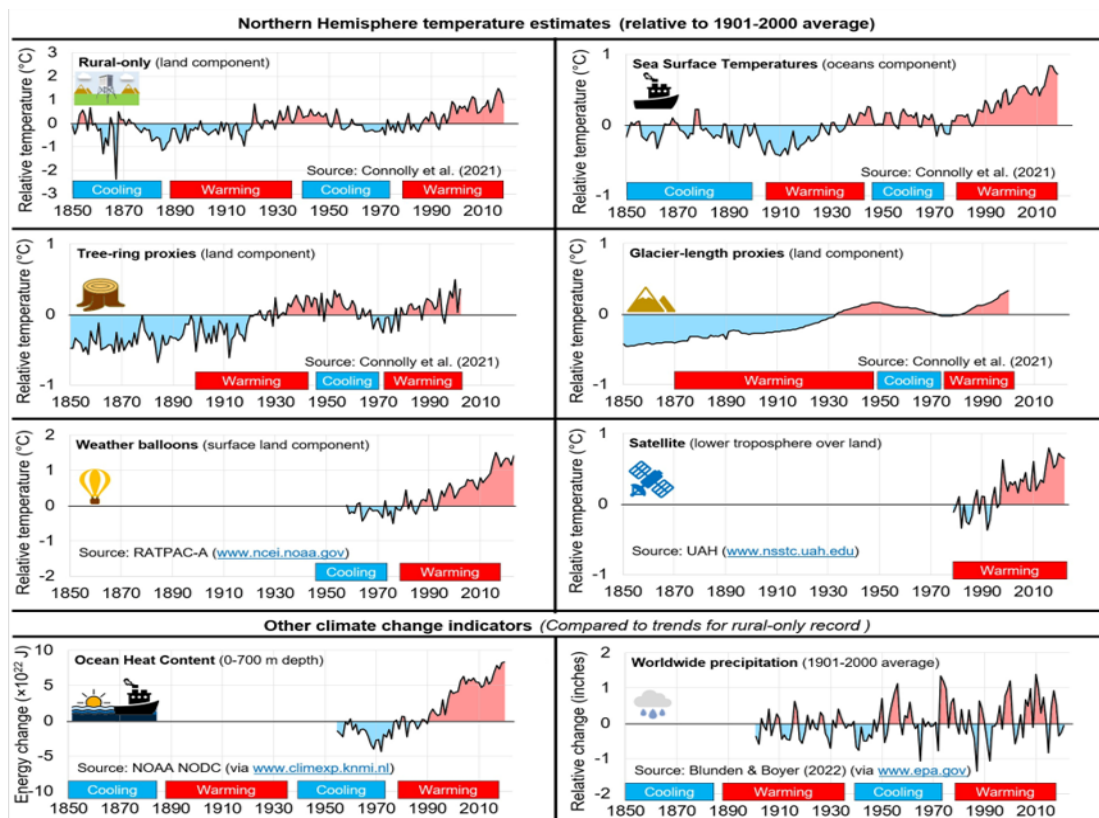


Figure 13: Comparison of our rural-only series to various other independent estimates of Northern Hemisphere temperature changes since 1850. See the sources indicated in each panel for details on how each temperature series or other climate change indicator was constructed.

4.3 Ongoing efforts to update and improve our analysis using GHCN version 4

Version 3 of the GHCN was discontinued in 2019. And Version 4 is not backwards compatible with version 3 and does not provide any urbanization metrics. Therefore, to update our rural-only

record and improve on it, considerable ongoing work is necessary. We have already carried out some work on developing new (improved) urbanization metrics for the Version 4 stations (Soon *et al.*, 2018; Katata, Connolly and O'Neill, 2023). However, we would also like to expand on the compilation of station history metadata that was begun in (O'Neill *et al.*, 2022). We also are considering availing of Peter O'Neill's archive of daily PHA runs to identify potential undocumented but persistent breakpoints. Finally, we think that the two statistical approaches to minimising urban blending we outlined in Katata *et al.* (2023) should be considered if statistical homogenization is to be continued.

5. Conclusions

There is no “global temperature” for the Earth, but averaging together temperature anomaly records from across the globe can potentially provide an index for describing widescale multi-decadal climatic changes. Unfortunately, until the mid-20th century, most of the weather station data available is heavily biased towards North America and Europe with some East Asian coverage. Most of these earlier stations were in areas that are now highly urbanized – and therefore potentially contaminated by urban warming. Nonetheless, this limited data suggests that the 20th century was warmer than the late-19th century. However, temperature proxies from across the globe suggest that on average the late-19th century coincided with the end of the Little Ice Age, i.e., a relatively cold period globally.

When we exclude stations that are potentially contaminated by urban warming, the data is too limited for the southern hemisphere to consider a global analysis. However, the rural data suggest that the northern hemisphere underwent a widespread cooling from the 1940s to the 1970s. The rural data confirm a general warming since then. Although this description of multidecadal temperature changes disagrees with the common narrative of a continuous “global warming” since the mid-19th century, it is surprisingly consistent with the trends implied by multiple independent estimates of northern hemisphere changes since the 1850s, i.e., sea surface temperatures, temperature proxies, weather balloons, satellites, ocean heat content measurements, etc.

The statistical homogenization techniques that current groups are relying on to remove any non-climatic biases from the temperature data might perform well on synthetic data, but they perform poorly on the real, incomplete, urban-contaminated datasets that are available. In particular, the adjustments from NOAA's widely used Pairwise Homogenization Algorithm change from day to day, rarely match the written station history, and systematically blend urban warming into rural records.

Claims that each new year is “the hottest on record” or “the hottest in 100 000 years” therefore rest on a short, geographically biased, heavily adjusted instrumental series whose adjustments change from day to day and only rarely match documented station events.

Funding

No funding specified.

Co-Editor: Stein Storlie Bergsmark

Acknowledgements

Much of the material covered in this proceedings article reviews a collection of collaborative papers that I have co-authored over the last 12 years: (Connolly and Connolly, 2014b, 2014c, 2014d, 2014e, 2014a; Soon, Connolly and Connolly, 2015; Connolly, Connolly and Soon, 2017; Soon *et al.*, 2018, 2023; Connolly *et al.*, 2021, 2023; O'Neill *et al.*, 2022; Katata, Connolly and O'Neill, 2023). I would like to thank all my co-authors on these papers, while noting that the views expressed specifically in this article are my own and not necessarily theirs. I would especially like to thank my late father, Dr. Michael Connolly, who worked with me on this research since we first

began pulling the thread of the temperature homogenization and non-climatic biases problems in 2010 until his untimely death on 25th August 2025. He is sorely missed. Dr. Willie Soon has also been an integral part of this research project since we first began collaborating in late-2014, and his unwavering passion for genuine scientific inquiry has been instrumental in keeping this project, and our other scientific collaborations (see www.ceres-science.com for more), going. Several of the figures in this paper were generated through the iterative prompting of various AI Large Language Models (LLMs). These are indicated in the relevant captions. Finally, I would like to thank the organizers of EIKE's 17th International Climate and Energy Conference (IKEK-17), Halle (an der Saale), Germany, 26-27 June 2026, for inviting me to present this work at their conference.

References

- Akasofu, S.-I. (2010) "On the recovery from the Little Ice Age," *Natural Science*, 2(11), pp. 1211–1224. Available at: <https://doi.org/10.4236/ns.2010.211149>.
- Associated Press (2016) "Study: Earth's roughly warmest in about 100,000 years," *Daily Mail*, 26 September. Available at: <http://www.dailymail.co.uk/~/article-3808373/index.html> (Accessed: September 12, 2026).
- Balling, R.C. and Idso, C.D. (2002) "Analysis of adjustments to the United States Historical Climatology Network (USHCN) temperature database," *Geophysical Research Letters*, 29(10), pp. 25-1-25-3. Available at: <https://doi.org/10.1029/2002GL014825>.
- Bhattacharya, S. (2006) "2005 was the warmest year on record," *New Scientist*, 25 January. Available at: <https://www.newscientist.com/article/1922623-2005-was-the-warmest-year-on-record/> (Accessed: September 12, 2026).
- Böhm, R. *et al.* (2010) "The early instrumental warm-bias: a solution for long central European temperature series 1760–2007," *Climatic Change*, 101(1), pp. 41–67. Available at: <https://doi.org/10.1007/s10584-009-9649-4>.
- Brázdil, R. *et al.* (2012) "Temperature and precipitation fluctuations in the Czech Republic during the period of instrumental measurements," *Theoretical and Applied Climatology*, 110(1), pp. 17–34. Available at: <https://doi.org/10.1007/s00704-012-0604-3>.
- Chimani, B. *et al.* (2020) "Compilation of a guideline providing comprehensive information on freely available climate change data and facilitating their efficient retrieval," *Climate Services*, 19, p. 100179. Available at: <https://doi.org/10.1016/j.cliser.2020.100179>.
- Christiansen, B. and Ljungqvist, F.C. (2017) "Challenges and perspectives for large-scale temperature reconstructions of the past two millennia," *Reviews of Geophysics*, 55(1), pp. 40–96. Available at: <https://doi.org/10.1002/2016RG000521>.
- Christy, J.R. (2026) "Declines in hot and cold daily temperature extremes in the conterminous US, 1899–2025," *Theoretical and Applied Climatology*, 157(5), p. 309. Available at: <https://doi.org/10.1007/s00704-026-06200-3>.
- Chylek, P. *et al.* (2020) "CMIP5 Climate Models Overestimate Cooling by Volcanic Aerosols," *Geophysical Research Letters*, 47(3), p. e2020GL087047. Available at: <https://doi.org/10.1029/2020GL087047>.
- Cohler, J. (2025) "The Father of Lies Hijacking Climate Science: Global Mean Surface Temperature Does Not Exist," *Journal of American Physicians and Surgeons*, 30(4), pp. 122–126. Available at: <https://doi.org/10.5281/zenodo.18256765>.

Connolly, R. *et al.* (2021) “How much has the Sun influenced Northern Hemisphere temperature trends? An ongoing debate,” *Research in Astronomy and Astrophysics*, 21(6), p. 131. Available at: <https://doi.org/10.1088/1674-4527/21/6/131>.

Connolly, R. *et al.* (2023) “Challenges in the Detection and Attribution of Northern Hemisphere Surface Temperature Trends Since 1850,” *Research in Astronomy and Astrophysics*, 23(10), p. 105015. Available at: <https://doi.org/10.1088/1674-4527/acf18e>.

Connolly, R. and Connolly, M. (2014a) “Global temperature changes of the last millennium,” *Open Peer Review Journal* [Preprint]. Available at: <http://oprj.net/articles/climate-science/16> (Accessed: January 21, 2020).

Connolly, R. and Connolly, M. (2014b) “Has poor station quality biased U.S. temperature trend estimates?,” *Open Peer Review Journal* [Preprint]. Available at: <http://oprj.net/articles/climate-science/11> (Accessed: November 11, 2019).

Connolly, R. and Connolly, M. (2014c) “Urbanization bias I. Is it a negligible problem for global temperature estimates?,” *Open Peer Review Journal* [Preprint]. Available at: <http://oprj.net/articles/climate-science/28> (Accessed: November 11, 2019).

Connolly, R. and Connolly, M. (2014d) “Urbanization bias II. An assessment of the NASA GISS urbanization adjustment method,” *Open Peer Review Journal* [Preprint]. Available at: <http://oprj.net/articles/climate-science/31> (Accessed: November 11, 2019).

Connolly, R. and Connolly, M. (2014e) “Urbanization bias III. Estimating the extent of bias in the Historical Climatology Network datasets,” *Open Peer Review Journal* [Preprint]. Available at: <http://oprj.net/articles/climate-science/34> (Accessed: November 11, 2019).

Connolly, R., Connolly, M. and Soon, W. (2017) “Re-calibration of Arctic sea ice extent datasets using Arctic surface air temperature records,” *Hydrological Sciences Journal*, 62(8), pp. 1317–1340. Available at: <https://doi.org/10.1080/02626667.2017.1324974>.

Courtillot, V. *et al.* (2008) “Response to comment on ‘Are there connections between Earth’s magnetic field and climate?, Earth Planet. Sci. Lett., 253, 328–339, 2007’ by Bard, E., and Delaygue, M., Earth Planet. Sci. Lett., in press, 2007,” *Earth and Planetary Science Letters*, 265(1), pp. 308–311. Available at: <https://doi.org/10.1016/j.epsl.2007.09.031>.

Domonkos, P. (2021) “Combination of Using Pairwise Comparisons and Composite Reference Series: A New Approach in the Homogenization of Climatic Time Series with ACMANT,” *Atmosphere*, 12(9), p. 1134. Available at: <https://doi.org/10.3390/atmos12091134>.

Easterling, D.R. and Peterson, T.C. (1995) “A new method for detecting undocumented discontinuities in climatological time series,” *International Journal of Climatology*, 15(4), pp. 369–377. Available at: <https://doi.org/10.1002/joc.3370150403>.

Essex, C. and McKittrick, R. (2007) *Taken by Storm: The Troubled Science, Policy, and Politics of Global Warming*. Toronto: Key Porter Books.

Essex, C., McKittrick, R. and Andresen, B. (2007) “Does a Global Temperature Exist?,” 32(1), pp. 1–27. Available at: <https://doi.org/10.1515/JNETDY.2007.001>.

Fall, S. *et al.* (2011) “Analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends,” *Journal of Geophysical Research: Atmospheres*, 116(D14), p. D14120. Available at: <https://doi.org/10.1029/2010JD015146>.

Gaffin, S.R. *et al.* (2008) “Variations in New York city’s urban heat island strength over time and space,” *Theoretical and Applied Climatology*, 94(1), pp. 1–11. Available at: <https://doi.org/10.1007/s00704-007-0368-3>.

Gillett, N.P. *et al.* (2021) “Constraining human contributions to observed warming since the pre-industrial period,” *Nature Climate Change*, 11(3), pp. 207–212. Available at: <https://doi.org/10.1038/s41558-020-00965-9>.

Hansen, J. *et al.* (1999) “GISS analysis of surface temperature change,” *Journal of Geophysical Research: Atmospheres*, 104(D24), pp. 30997–31022. Available at: <https://doi.org/10.1029/1999JD900835>.

Hansen, J. *et al.* (2001) “A closer look at United States and global surface temperature change,” *Journal of Geophysical Research: Atmospheres*, 106(D20), pp. 23947–23963. Available at: <https://doi.org/10.1029/2001JD000354>.

Hansen, J. *et al.* (2010) “Global Surface Temperature Change,” *Reviews of Geophysics*, 48(4), p. RG4004. Available at: <https://doi.org/10.1029/2010RG000345>.

Hausfather, Z. *et al.* (2013) “Quantifying the effect of urbanization on U.S. Historical Climatology Network temperature records,” *Journal of Geophysical Research: Atmospheres*, 118(2), pp. 481–494. Available at: <https://doi.org/https://doi.org/10.1029/2012JD018509>.

Hess, P. (2024) “Earth surpassed ‘doomed’ 2.7F warming limit for the first time in 2023,” *Daily Mail*, 9 February. Available at: <https://www.dailymail.co.uk/sciencetech/article-13066483/earth-breaches-temperature-warning-limit-record.html> (Accessed: September 12, 2026).

Hubbard, K.G. and Lin, X. (2006) “Reexamination of instrument change effects in the U.S. Historical Climatology Network,” *Geophysical Research Letters*, 33(15). Available at: <https://doi.org/10.1029/2006GL027069>.

IPCC (2007) *Climate Change 2007 - The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC*. Cambridge ; New York: Cambridge University Press.

IPCC (2021) *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Jones, P.D. *et al.* (1986) “Northern Hemisphere Surface Air Temperature Variations: 1851–1984,” *Journal of Applied Meteorology and Climatology*, 25(2), pp. 161–179. Available at: [https://doi.org/10.1175/1520-0450\(1986\)025%3C0161:NHSATV%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025%3C0161:NHSATV%3E2.0.CO;2).

Jones, P.D. *et al.* (1990) “Assessment of urbanization effects in time series of surface air temperature over land,” *Nature*, 347(6289), pp. 169–172. Available at: <https://doi.org/10.1038/347169a0>.

Jones, P.D., Osborn, T.J. and Briffa, K.R. (1997) “Estimating Sampling Errors in Large-Scale Temperature Averages,” *Journal of Climate*, 10(10), pp. 2548–2568. Available at: [https://doi.org/10.1175/1520-0442\(1997\)010%3C2548:ESEILS%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010%3C2548:ESEILS%3E2.0.CO;2).

Jones, P.D., Wigley, T.M.L. and Wright, P.B. (1986) “Global temperature variations between 1861 and 1984,” *Nature*, 322(6078), pp. 430–434. Available at: <https://doi.org/10.1038/322430a0>.

- Karl, T.R. *et al.* (1986) “A Model to Estimate the Time of Observation Bias Associated with Monthly Mean Maximum, Minimum and Mean Temperatures for the United States,” *Journal of Applied Meteorology and Climatology*, 25(2), pp. 145–160. Available at: [https://doi.org/10.1175/1520-0450\(1986\)025%3C0145:AMTETT%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025%3C0145:AMTETT%3E2.0.CO;2).
- Karl, T.R., Diaz, H.F. and Kukla, G. (1988) “Urbanization: Its Detection and Effect in the United States Climate Record,” *Journal of Climate*, 1(11), pp. 1099–1123. Available at: [https://doi.org/10.1175/1520-0442\(1988\)001%3C1099:UIDAEI%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(1988)001%3C1099:UIDAEI%3E2.0.CO;2).
- Katata, G., Connolly, R. and O’Neill, P. (2023) “Evidence of Urban Blending in Homogenized Temperature Records in Japan and in the United States: Implications for the Reliability of Global Land Surface Air Temperature Data,” *Journal of Applied Meteorology and Climatology*, 62(8), pp. 1095–1114. Available at: <https://doi.org/10.1175/JAMC-D-22-0122.1>.
- Lawrimore, J.H. *et al.* (2011) “An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3,” *Journal of Geophysical Research: Atmospheres*, 116(D19), p. D19121. Available at: <https://doi.org/10.1029/2011JD016187>.
- Li, Q. *et al.* (2004) “Urban heat island effect on annual mean temperature during the last 50 years in China,” *Theoretical and Applied Climatology*, 79(3), pp. 165–174. Available at: <https://doi.org/10.1007/s00704-004-0065-4>.
- Lund, R.B. *et al.* (2023) “Good Practices and Common Pitfalls in Climate Time Series Change-point Techniques: A Review.” Available at: <https://doi.org/10.1175/JCLI-D-22-0954.1>.
- McCarthy, M. (1998) “1998 warmest year on record worldwide,” *Irish Independent*, 18 December. Available at: <https://www.independent.ie/world-news/1998-warmest-year-on-record-world-wide/a/149787070.html> (Accessed: September 12, 2026).
- Menne, M.J. *et al.* (2018) “The Global Historical Climatology Network Monthly Temperature Dataset, Version 4,” *Journal of Climate*, 31(24), pp. 9835–9854. Available at: <https://doi.org/10.1175/JCLI-D-18-0094.1>.
- Menne, M.J. and Williams, C.N. (2009) “Homogenization of Temperature Series via Pairwise Comparisons,” *Journal of Climate*, 22(7), pp. 1700–1717. Available at: <https://doi.org/10.1175/2008JCLI2263.1>.
- Menne, M.J., Williams, C.N. and Vose, R.S. (2009) “The U.S. Historical Climatology Network Monthly Temperature Data, Version 2,” *Bulletin of the American Meteorological Society*, 90(7), pp. 993–1008. Available at: <https://doi.org/10.1175/2008BAMS2613.1>.
- Menne, M.J., Williams Jr., C.N. and Palecki, M.A. (2010) “On the reliability of the U.S. surface temperature record,” *Journal of Geophysical Research: Atmospheres*, 115(D11). Available at: <https://doi.org/10.1029/2009JD013094>.
- Mills, G. (2008) “Luke Howard and The Climate of London,” *Weather*, 63(6), pp. 153–157. Available at: <https://doi.org/10.1002/wea.195>.
- Oke, T.R. (1973) “City size and the urban heat island,” *Atmospheric Environment (1967)*, 7(8), pp. 769–779. Available at: [https://doi.org/10.1016/0004-6981\(73\)90140-6](https://doi.org/10.1016/0004-6981(73)90140-6).
- O’Neill, P. *et al.* (2022) “Evaluation of the Homogenization Adjustments Applied to European Temperature Records in the Global Historical Climatology Network Dataset,” *Atmosphere*, 13(2), p. 285. Available at: <https://doi.org/10.3390/atmos13020285>.

- Osborn, T.J. *et al.* (2021) “Land Surface Air Temperature Variations Across the Globe Updated to 2019: The CRUTEM5 Data Set,” *Journal of Geophysical Research: Atmospheres*, 126(2), p. e2019JD032352. Available at: <https://doi.org/10.1029/2019JD032352>.
- Parker, D.E. (2006) “A Demonstration That Large-Scale Warming Is Not Urban,” *Journal of Climate*, 19(12), pp. 2882–2895. Available at: <https://doi.org/10.1175/JCLI3730.1>.
- Peterson, T.C. *et al.* (1999) “Global rural temperature trends,” *Geophysical Research Letters*, 26(3), pp. 329–332. Available at: <https://doi.org/10.1029/1998GL900322>.
- Peterson, T.C. (2003) “Assessment of Urban Versus Rural In Situ Surface Temperatures in the Contiguous United States: No Difference Found,” *Journal of Climate*, 16(18), pp. 2941–2959. Available at: [https://doi.org/10.1175/1520-0442\(2003\)016%3C2941:AOUVRI%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016%3C2941:AOUVRI%3E2.0.CO;2).
- Peterson, T.C. and Vose, R.S. (1997) “An Overview of the Global Historical Climatology Network Temperature Database,” *Bulletin of the American Meteorological Society*, 78(12), pp. 2837–2850. Available at: [https://doi.org/10.1175/1520-0477\(1997\)078%3C2837:AOOTGH%3E2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078%3C2837:AOOTGH%3E2.0.CO;2).
- Przybylak, R. (2010) “The Climate of Poland in Recent Centuries: A Synthesis of Current Knowledge: Instrumental observations,” in R. Przybylak *et al.* (eds.) *The Polish Climate in the European Context: An Historical Overview*. Dordrecht: Springer, Berlin Heidelberg New York, pp. 129–166. Available at: https://doi.org/10.1007/978-90-481-3167-9_5.
- Quayle, R.G. *et al.* (1991) “Effects of Recent Thermometer Changes in the Cooperative Station Network,” *Bulletin of the American Meteorological Society*, 72(11), pp. 1718–1724. Available at: [https://doi.org/10.1175/1520-0477\(1991\)072%3C1718:EORTCI%3E2.0.CO;2](https://doi.org/10.1175/1520-0477(1991)072%3C1718:EORTCI%3E2.0.CO;2).
- Reeves, J. *et al.* (2007) “A Review and Comparison of Changepoint Detection Techniques for Climate Data,” *Journal of Applied Meteorology and Climatology*, 46(6), pp. 900–915. Available at: <https://doi.org/10.1175/JAM2493.1>.
- Ren, G. *et al.* (2008) “Urbanization Effects on Observed Surface Air Temperature Trends in North China.” Available at: <https://doi.org/10.1175/2007JCLI1348.1>.
- Rohde, R.A. and Hausfather, Z. (2020) “The Berkeley Earth Land/Ocean Temperature Record,” *Earth System Science Data*, 12(4), pp. 3469–3479. Available at: <https://doi.org/10.5194/essd-12-3469-2020>.
- Scafetta, N. (2021) “Detection of non-climatic biases in land surface temperature records by comparing climatic data and their model simulations,” *Climate Dynamics*, <https://doi.org/10.1007/s00382-021-05626-x>. Available at: <https://doi.org/10.1007/s00382-021-05626-x>.
- Skrynyk, Oleg *et al.* (2019) “Quality control and homogenization of monthly extreme air temperature of Ukraine,” *International Journal of Climatology*, 39(4), pp. 2071–2079. Available at: <https://doi.org/10.1002/joc.5934>.
- Soon, W. *et al.* (2003) “Reconstructing Climatic and Environmental Changes of the Past 1000 Years: A Reappraisal,” *Energy & Environment*, 14(2–3), pp. 233–296. Available at: <https://doi.org/10.1260/095830503765184619>.
- Soon, W. *et al.* (2018) “Comparing the current and early 20th century warm periods in China,” *Earth-Science Reviews*, 185, pp. 80–101. Available at: <https://doi.org/10.1016/j.earscirev.2018.05.008>.

<https://doi.org/10.1016/j.earscirev.2018.05.013>.

Soon, W. *et al.* (2023) “The Detection and Attribution of Northern Hemisphere Land Surface Warming (1850–2018) in Terms of Human and Natural Factors: Challenges of Inadequate Data,” *Climate*, 11(9), p. 179. Available at: <https://doi.org/10.3390/cli11090179>.

Soon, W. and Baliunas, S. (2003) “Proxy climatic and environmental changes of the past 1000 years,” *Climate Research*, 23(2), pp. 89–110. Available at: <https://doi.org/10.3354/cr023089>.

Soon, W., Baliunas, S. and Legates, D. (2003) “Comment on ‘On past temperatures and anomalous late-20th century warmth,’” *Eos, Transactions American Geophysical Union*, 84(44), pp. 473–476. Available at: <https://doi.org/10.1029/2003EO440007>.

Soon, W., Connolly, R. and Connolly, M. (2015) “Re-evaluating the role of solar variability on Northern Hemisphere temperature trends since the 19th century,” *Earth-Science Reviews*, 150, pp. 409–452. Available at: <https://doi.org/10.1016/j.earscirev.2015.08.010>.

Soon, W., Connolly, R. and Connolly, M. (2024) *The Unreliability of Current Global Temperature and Solar Activity Estimates and Its Implications for the Attribution of Global Warming*. 305. Washington, D.C: The Heritage Foundation, p. 31. Available at: <https://www.heritage.org/climate/report/the-unreliability-current-global-temperature-and-solar-activity-estimates-and-its> (Accessed: September 12, 2026).

Soon, W.W.-H. *et al.* (2019) “Reply to Li & Yang’s comments on ‘Comparing the current and early 20th century warm periods in China,’” *Earth-Science Reviews*, 198, p. 102950. Available at: <https://doi.org/10.1016/j.earscirev.2019.102950>.

Spencer, R.W., Christy, J.R. and Braswell, W.D. (2025) “Urban Heat Island Effects in U.S. Summer Surface Temperature Data, 1895–2023,” *Journal of Applied Meteorology and Climatology*, In press. Available at: <https://doi.org/10.1175/JAMC-D-23-0199.1>.

Stevens, W.K. (1996) “95 Is Hottest Year on Record As the Global Trend Resumes,” *The New York Times*, 4 January. Available at: <https://www.nytimes.com/1996/01/04/world/95-is-hottest-year-on-record-as-the-global-trend-resumes.html> (Accessed: September 12, 2026).

Stewart, I.D. (2019) “Why should urban heat island researchers study history?,” *Urban Climate*, 30, p. 100484. Available at: <https://doi.org/10.1016/j.uclim.2019.100484>.

Stewart, I.D. and Oke, T.R. (2012) “Local Climate Zones for Urban Temperature Studies,” *Bulletin of the American Meteorological Society*, 93(12), pp. 1879–1900. Available at: <https://doi.org/10.1175/BAMS-D-11-00019.1>.

UN News (2025) “Confirmed: 2024 was the hottest year on record, says UN weather agency | UN News,” 10 January. Available at: <https://news.un.org/en/story/2025/01/1158891> (Accessed: September 12, 2026).

Venema, V.K.C. *et al.* (2012) “Benchmarking homogenization algorithms for monthly data,” *Climate of the Past*, 8(1), pp. 89–115. Available at: <https://doi.org/10.5194/cp-8-89-2012>.

Vose, R.S. *et al.* (1992) *The Global Historical Climatology Network: Long-term monthly temperature, precipitation, sea level pressure, and station pressure data*. ORNL/CDIAC-53; NDP-041. Oak Ridge National Lab., TN (United States). Carbon Dioxide Information Analysis Center. Available at: <https://doi.org/10.2172/10178730>.

Vose, R.S. *et al.* (2003) “An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network,” *Geophysical Research Letters*, 30(20), p. 2046. Available at: <https://doi.org/10.1029/2003GL018111>.

Wickham, C. *et al.* (2013) “Influence of Urban Heating on the Global Temperature Land Average using Rural Sites Identified from MODIS Classifications,” *Geoinformatics & Geostatistics: An Overview*, 01(02). Available at: <https://doi.org/10.4172/2327-4581.1000104>.

Williams, C.N., Menne, M.J. and Thorne, P.W. (2012) “Benchmarking the performance of pair-wise homogenization of surface temperatures in the United States,” *Journal of Geophysical Research: Atmospheres*, 117(D5). Available at: <https://doi.org/10.1029/2011JD016761>.

Winkler, P. (2009) “Revision and necessary correction of the long-term temperature series of Hohenpeissenberg, 1781–2006,” *Theoretical and Applied Climatology*, 98(3), pp. 259–268. Available at: <https://doi.org/10.1007/s00704-009-0108-y>.

Wypych, A., Ustrnul, Z. and Schmatz, D.R. (2018) “Long-term variability of air temperature and precipitation conditions in the Polish Carpathians,” *Journal of Mountain Science*, 15(2), pp. 237–253. Available at: <https://doi.org/10.1007/s11629-017-4374-3>.

Zhang, P. *et al.* (2021) “Urbanization Effects on Estimates of Global Trends in Mean and Extreme Air Temperature,” *Journal of Climate*, 34(5), pp. 1923–1945. Available at: <https://doi.org/10.1175/JCLI-D-20-0389.1>.

Editorial note: The following disclaimer is standard journal text and will be maintained by the editorial office. Please do not edit or delete.

Disclaimer/Publisher’s Note:

This article reflects solely the views, analyses, and conclusions of the author(s). Science of Climate Change, its editors, and its publisher make no representation or warranty as to the accuracy or completeness of the content and disclaim liability for any loss, damage, or other consequence arising from the use of, or reliance on, any material published herein.