Earth-Science Reviews 150 (2015) 409–452

Contents lists available at ScienceDirect

Earth-Science Reviews

journal homepage: www.elsevier.com/locate/earscirev

Re-evaluating the role of solar variability on Northern Hemisphere temperature trends since the 19th century

Willie Soon [a,*], Ronan Connolly [b], Michael Connolly [b]

[a] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
[b] Independent research scientists, Dublin, Ireland

ARTICLE INFO

Article history:
Received 23 June 2015
Received in revised form 12 August 2015
Accepted 22 August 2015
Available online 1 September 2015

Keywords:
Solar variability
Climate change
Urbanization bias
Northern Hemisphere temperatures
Total Solar Irradiance

ABSTRACT

Debate over what influence (if any) solar variability has had on surface air temperature trends since the 19th century has been controversial. In this paper, we consider two factors which may have contributed to this controversy:

1. Several different solar variability datasets exist. While each of these datasets is constructed on plausible grounds, they often imply contradictory estimates for the trends in solar activity since the 19th century.
2. Although attempts have been made to account for non-climatic biases in previous estimates of surface air temperature trends, recent research by two of the authors has shown that current estimates are likely still affected by non-climatic biases, particularly urbanization bias.

With these points in mind, we first review the debate over solar variability. We summarise the points of general agreement between most groups and the aspects which still remain controversial. We discuss possible future research which may help resolve the controversy of these aspects. Then, in order to account for the problem of urbanization bias, we compile a new estimate of Northern Hemisphere surface air temperature trends since 1881, using records from predominantly rural stations in the monthly Global Historical Climatology Network dataset. Like previous weather station-based estimates, our new estimate suggests that surface air temperatures warmed during the 1880s–1940s and 1980s–2000s. However, this new estimate suggests these two warming periods were separated by a pronounced cooling period during the 1950s–1970s and that the relative warmth of the mid-20th century warm period was comparable to the recent warm period.
We then compare our weather station-based temperature trend estimate to several other independent estimates. This new record is found to be consistent with estimates of Northern Hemisphere Sea Surface Temperature (SST) trends, as well as temperature proxy-based estimates derived from glacier length records and from tree ring widths. However, the multi-model means of the recent Coupled Model Intercomparison Project Phase 5 (CMIP5) climate model hindcasts were unable to adequately reproduce the new estimate — although the modelling of certain volcanic eruptions did seem to be reasonably well reproduced.
Finally, we compare our new composite to one of the solar variability datasets not considered by the CMIP5 climate models, i.e., Scafetta and Willson, 2014's update to the Hoyt and Schatten, 1993 dataset. A strong correlation is found between these two datasets, implying that solar variability has been the dominant influence on Northern Hemisphere temperature trends since at least 1881. We discuss the significance of this apparent correlation, and its implications for previous studies which have instead suggested that increasing atmospheric carbon dioxide has been the dominant influence.

Published by Elsevier B.V.

Contents

* Corresponding author.

http://dx.doi.org/10.1016/j.earscirev.2015.08.010
0012-8252/Published by Elsevier B.V.

1. Introduction

In recent years, there has been considerable debate over what influence (if any) solar variability has had on global and regional surface air temperature trends[1] since the 19th century. Some authors have argued for a large role (e.g., Soon, 2005; Svensmark et al., 2009; Le Mouël et al., 2010; Vahrenholt and Lüning, 2013; Scafetta and Willson, 2014); others have argued that it has only played a minor role in recent decades (e.g., Solanki and Krivova, 2003; Balling and Roy, 2005; Gray et al., 2010; Bindoff et al., 2013); while others have argued it has played little (if any) role (e.g., Foukal et al., 2006; Clette et al., 2014; Tsonis et al., 2015). One of us (WS) has been an active participant in this debate (e.g., Zhang et al., 1994; Soon and Yaskell, 2003; Soon, 2005, 2009, 2014; Soon et al., 2011; Soon and Legates, 2013; Yan et al., 2015), which has become particularly significant lately, since the latest Global Climate Model hindcasts[2] used by the Intergovernmental Panel on Climate Change (IPCC) reports have indicated that solar variability has only had a modest influence on recent temperature trends. As a result, the latest IPCC reports concluded that temperature trends since 1951 are mostly "*…due to the observed anthropogenic increase in greenhouse gas (GHG) concentrations*" (Bindoff et al., 2013).

One reason for the lack of resolution of the debate is that the available information on solar variability is still rather limited, and as a result different estimates for solar trends are often contradictory. For instance, of the three satellite-based estimates for the Total Solar Irradiance (TSI) activity since 1978, one suggests there has been a general decrease (e.g., Fröhlich, 2006, 2012, 2013); one suggests there has been no discernible trend (e.g., Mekaoui and Dewitte, 2008; Mekaoui et al., 2010); and the third suggests an increase until about 2000 followed by a decrease (e.g., Willson, 2014; Scafetta and Willson, 2014). The Total Solar Irradiance (sometimes referred to as "solar activity") is the aspect of solar variability which is most likely to *directly* influence climate. Therefore, in this paper we will usually treat the terms synonymously — although we will briefly consider in Section 2.3 other aspects of solar variability which may indirectly influence the climate, e.g., the strength of the solar wind.

Another problem is that debate exists over the extent to which non-climatic biases in the instrumental records have biased current global temperature trend estimates, e.g., see the debate between Le Mouël et al. (2009, 2011) and Legras et al. (2010). In particular, for many years, there has been concern that the development and expansion of "urban heat islands" (e.g., Stewart and Oke, 2012) around many weather stations may have introduced a warming "urbanization bias" into regional (and possibly global) temperature trend estimates (e.g., Mitchell, 1953; Karl et al., 1988; Balling and Idso, 1989; Ren et al., 2008; Ren and Ren, 2011; Yang et al., 2011, 2013; Li et al., 2013; Symmons, 2014).

Several studies have claimed that urbanization bias has *not* substantially biased current estimates (e.g., Peterson et al., 1999; Parker, 2006; Wickham et al., 2013) and/or that data homogenization has effectively removed the problem (e.g., Menne et al., 2009; Hansen et al., 2010; Lawrimore et al., 2011; Hausfather et al., 2013). However, in a series of three companion papers, two of us (RC & MC) have recently shown that there were flaws in each of these earlier studies, and that urbanization bias is indeed a substantial (and insidious) problem for the current weather station-based temperature estimates (Connolly and Connolly, 2014a,b,c).

With this in mind, we have tried in this collaborative paper to address both problems simultaneously. First, we review the reasons for the ongoing solar variability debate. Second, we construct and assess a new Northern Hemisphere temperature trend estimate derived from predominantly rural stations taken from the widely-used Global Historical Climatology Network (GHCN) dataset (Lawrimore et al., 2011). We then present evidence which suggests that Northern Hemisphere

[1] Typically, the term "temperature trend" refers to decadal to multi-decadal, or longer, secular trends of either warming or cooling, as opposed to shorter-term "temperature variability" associated with events such as ENSO processes, or large volcanic eruptions.

[2] A "hindcast" is the opposite of a forecast, i.e., a retrospective "prediction" of what is expected to have occurred in the past.

temperature trends since the 19th century have actually been heavily influenced by changes in solar variability. However, before we do so, it may be helpful to briefly discuss the caveats associated with analyses that rely on apparent correlations between datasets.

1.1. A cautionary comment on different types of correlations

Much of the research into the possible influence of solar variability on the Earth's climate has relied on the presence (or absence) of apparent correlations between various solar variability datasets and climatic datasets. However, correlation does not necessarily imply causation. That is, there are at least four types of correlations:

1. *Causal correlation*. One of the variables directly influences the other, and so changes in that variable over time will tend to cause a corresponding change in the other variable. Sometimes both variables can influence each other, in which case changes in one of the variables can sometimes trigger a feedback loop. However, if one variable can cause a change in the other, but not vice versa, then we say that the "direction of causation" lies from the former to the latter.
2. *Commensal correlation*. Both variables are influenced by a common factor. So, changes in that common ("parent") factor over time will induce corresponding changes in both variables.
3. *Coincidental correlation*. The two datasets are completely independent of each other. However, due to the variability within both datasets, over a short period of time the trends of both variables temporarily *appear* to be correlated. Often, when these datasets are updated with further data, the apparent correlation will start to break down.
4. *Constructional correlation*. When one of the datasets was being constructed, it might have been assumed that it *should* be related to the other. When subjective decisions are made, researchers may mistakenly allow confirmation bias (e.g., Nickerson, 1998) to affect their decision. As a result, this could have artificially introduced an apparent correlation between the two datasets.

For a given correlation, it is possible that more than one factor might be at play. For instance, one variable might genuinely be causally or commensally correlated to the other, but the apparent strength of the correlation might have been exaggerated by a coincidental or constructional correlation. On the other hand, one variable might directly influence the other, but if the second variable is *also* influenced by other factors, this could reduce the apparent strength of the correlation over short periods of time (i.e., whenever more than one factor is influencing the second variable).

In the case of an apparent correlation between a given solar variability dataset and a climatic dataset, if the correlation is causal then it seems reasonable to assume that the direction of causation lies from the former to the latter. That is, it seems safe to assume that solar variability would be influencing the Earth's climate, rather than the other way around. In some cases, changes in a climatic dataset may appear to precede the changes in solar variability. While this may often indicate that the apparent correlation is spurious, we must caution that this is not always the case. For instance, the variable being used as a solar proxy might lag the actual solar variability. Or, if there are cyclical patterns in both the solar and climate variables, then differences in phase might incorrectly create the impression that the effect precedes the cause.

In many cases, it should not make too much difference to our conclusions whether the correlation is causal or commensal. If a given solar-climate correlation were commensal, then this would indicate that *some* (possibly unknown) factor which is influencing the Earth's climate is also influencing a particular aspect of solar variability. However, if that factor was influencing some aspect of solar variability, it would presumably be some other form of solar variability, and therefore the correlation would still be with solar variability.

It follows that our primary concern should be the possibility that either of the other two types of correlation is involved. For if the apparent correlations are either coincidental or constructional in nature, then an apparently strong link between surface air temperatures (for example) and solar variability can be considered spurious.

1.2. Format of this article

The format of this article is as follows:

- In Section 2, we review the solar variability debate and discuss the evidence for and against various different estimates of solar trends since the 19th century
- In Section 3, we look in detail at four regions in the Northern Hemisphere (China, U.S., Ireland and the Arctic) and determine the regional temperature trends for these areas using data from predominantly rural stations
- In Section 4, we combine our four regional estimates into a single Northern Hemisphere composite covering the period 1881–2014. We then compare and contrast this composite with several other estimates of Northern Hemisphere temperature trends.
- In Section 5, we identify and discuss an apparently strong relationship between our new Northern Hemisphere composite and the updated version of Hoyt & Schatten's estimate of solar activity trends (Hoyt and Schatten, 1993; updated by Scafetta and Willson, 2014).
- Finally, in Section 6, we offer some concluding remarks.

2. Review of the solar variability debate

For thousands of years, researchers have considered the possibility that changes in solar activity can lead to climate change on Earth, e.g., Theophrastus (371–287 BC) suggested there might be a connection between sunspots and rain and wind (see p. 2 of Soon and Yaskell, 2003 and refs. therein). However, without systematic and quantitative measurements and records with which to check these possibilities, any such theories remained mostly speculative. We will consider climate records in Sections 3 and 4, but in this section, we will focus on the various attempts to quantify solar activity and how it has changed over the years.

As technology has improved, so has the ability of astronomers to monitor and observe the Sun's activity (Harvey, 2013). The invention of the telescope in the 17th century provided one such improvement. In particular, this allowed Sun observers to record more accurate and detailed sunspot observations (Hoyt and Schatten, 1998; Arlt et al., 2013; Clette et al., 2014). These sunspot records reveal a pseudo-cyclical pattern of sunspot activity with a period of *around* 11 years (typically varying between 8 and 14 years). At the start and end of each "cycle", very few sunspots are observed, but during the middle of the cycle, large numbers of sunspots are frequently observed.

As will be discussed in Section 2.2, the *exact* relationship between sunspot activity and solar activity is still not entirely clear. However, it is now well established that increases and decreases in sunspot activity are associated with corresponding changes in solar activity (e.g., Solanki and Fligge, 1999; Gray et al., 2010; Fröhlich, 2012; Willson, 2014). Therefore, as we will discuss in Section 2.2, changes in sunspot activity inferred from these sunspot records are often used as a solar "proxy" to ap*proxi*mate solar activity.

At any rate, to some of these early astronomers, changes in sunspot activity were indicative of solar variability, and several researchers used their own sunspot records to study possible connections with the Earth's climate. For instance, the Jesuit astronomer Giovanni Battista Riccioli (1598–1671) speculated about the connection of sunspot activity and weather in his *Almagestum novum* published in 1651. Later, the Mexican astronomer and meteorologist, Jose Antonio Alzate (1737–1799) in 1784 suggested an apparent correlation between

sunspot activity and crop prices (Galindo and Saladino, 2008), as did the German-born British astronomer, William Herschel (1738–1822) in 1801 (Herschel, 1801a,b). More modern sun-climate linkages include the demonstration by Neff et al. (2001) of an apparently strong correlation between fluctuations in cosmogenic radiocarbon (i.e., a proxy for solar activity) and oxygen isotope composition (i.e., a proxy for rainfall) during the mid-Holocene in a speleothem from Oman.

Many of these apparent correlations have since been disputed (e.g., Love, 2013), although others have suggested there may be truth in some of them (e.g., Langley, 1888; Menzel, 1959; Pustil'nik and Yom Din, 2004). However, it must be recognised that the above astronomers were acutely conscious of the limited data that was available to them, and were advocating for more data to be collected, so that more robust studies could be carried out in the future (e.g., Galindo and Saladino, 2008).

In the centuries since the invention of the telescope, and particularly in recent decades, astronomical technology has continued to improve dramatically, e.g., the setting up of permanent astronomical observatories. In particular, starting in 1978, satellites have been able to monitor, *from above the atmosphere*, the Total Solar Irradiance reaching the Earth. Since ground-based measurements are made at the bottom of the atmosphere, by the time the incoming solar radiation has reached the measuring devices, it has already interacted with the atmosphere, e.g., by absorption or reflection. Therefore, the space-based observations from satellites provide more direct measurements of the true solar variability than do the previous ground-based measurements. With this in mind, we next consider the debate over the data during "the satellite era", i.e., 1978–present.

2.1. The satellite era

When we are considering the possible role of solar variability on the Earth's climate, it is first important to distinguish between actual variability in the irradiance emitted by the Sun ("solar variability") and the variability of the solar irradiance *reaching* the Earth ("orbital variability").

The seasonality of the *local* incoming solar irradiance is obvious from the seasons of the year. That is, in the Northern Hemisphere, the incoming solar irradiance is at a minimum in December/January ("winter") and a maximum in June/July ("summer"), while in the Southern Hemisphere, the seasons are reversed. However, because the Earth's orbit of the Sun is elliptical and not circular, in addition to this local seasonality, there is also seasonal variation in the *total* solar irradiance reaching the Earth.

As can be seen from Fig. 1, the total amount of solar radiation reaching the Earth currently is at a maximum in January and a minimum in July, since the Earth–Sun distance is at a minimum in January. This factor is known to change over time-scales of tens to hundreds of thousands of years due to the cyclical variability in the Earth's orbit (Berger et al., 1993). These "Milankovitch cycles" are believed to be a significant factor in the glacial to inter-glacial transitions of the ice ages (Berger, 1988; Roe, 2006). At present, the difference between the January and July solar radiation is quite substantial, at about 6.5% of the average solar radiation (~88 W/m^2).

While this orbital variability obviously has a strong influence on surface air temperatures, it is independent of the actual solar variability. Therefore, when we are specifically considering the effects of solar variability on climate, we are usually interested in the Total Solar Irradiance (TSI) at a fixed Earth–Sun distance. Hence, the daily measured values recorded by satellites are typically rescaled to the values they would have if the satellite was exactly 1 Astronomical Unit (AU) from the Sun, i.e., the *average* distance of the Earth from the Sun.

Fig. 2 shows the estimates of Total Solar Irradiance (TSI) at 1 AU from all seven of the solar monitoring satellites launched since 1978. Several key points which are relevant to our discussion are immediately apparent.

Although each of the satellites implies a *similar range* of solar variability during a given solar cycle, the *absolute* Total Solar Irradiances reported by each satellite are often quite different. For instance, the NIMBUS7/ERB satellite recorded values in the range 1370–1375 W/m^2, while the SMM/ACRIM1 satellite recorded values in the range 1365–1370 W/m^2, even though both satellites were monitoring the Sun at the same time.

This brings us to our first point of contention, i.e., during the satellite era, what has been the mean absolute Total Solar Irradiance (at 1 AU)? The first two satellites seemed to imply the value was somewhere in the range 1365–1375 W/m^2. However, by the mid-1990s, the solar monitoring community began to converge on a consensus that these early estimates were due to problems in instrumentation and the true value was actually about 1365 W/m^2 (e.g., Crommelynck et al., 1995; Mekaoui et al., 2010). Later, however, as new replacement satellites were launched, some researchers began arguing that the true value was actually around 1361 Wm^{-2} (e.g., Kopp and Lean, 2011; Willson, 2014).

This more recent value is similar to a theoretical prediction of about 1360 Wm^{-2} by Monin & Shishkov that was based on assuming an

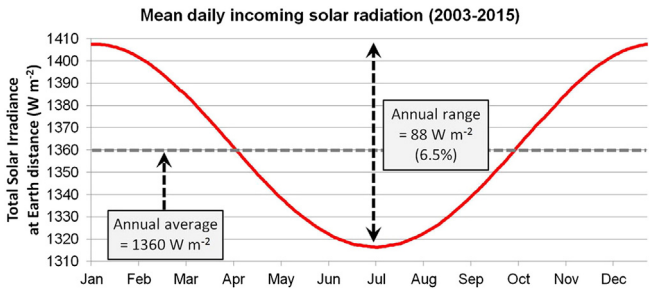

Fig. 1. Mean daily variability of the total solar irradiance at the top of the Earth's atmosphere, over the annual cycle. Values were calculated at Earth distance using data from the SORCE satellite mission (2003–2013).
Downloaded in July 2015 from http://lasp.colorado.edu/home/sorce/.

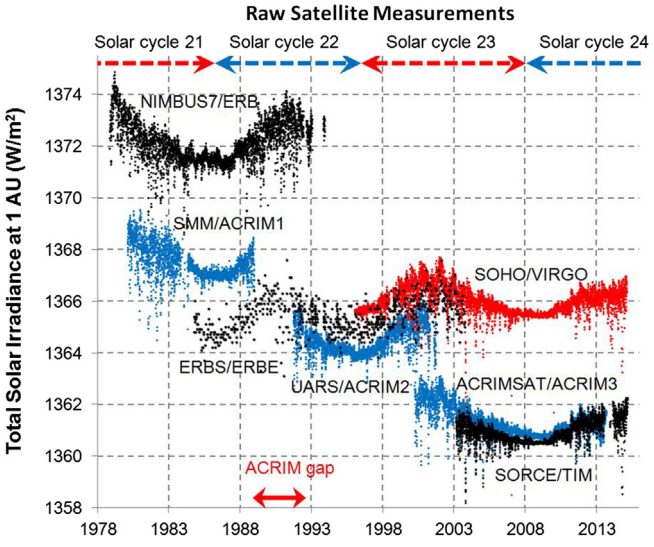

Fig. 2. Estimates of Total Solar Irradiance (TSI) from all seven of the TSI satellites launched since the start of the satellite era, i.e., since 1978. Individual satellite data series were downloaded on 17th March 2015 from the following websites: ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_IRRADIANCE/nimbus.plt (NIMBUS); ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_IRRADIANCE/ERBS2003.TXT (ERBS); http://lasp.colorado.edu/home/sorce/data/ (SORCE/TIM); ftp://ftp.pmodwrc.ch/pub/data/irradiance/virgo/TSI/ (VIRGO); http://www.acrim.com/Data%20Products.htm (ACRIM1-3).

effective temperature of the Sun of 5770 K and luminosity of 3.83×10^{26} W (Monin and Shishkov, 2000; see p. 385). It is also consistent with the suggested calibrated values of about 1361 Wm^{-2} by the PREMOS (Precision Monitoring Sensor onboard the PICARD satellite mission) experiments (W. Schmutz, 2012, private communication with WS). So, maybe the true value is ~1360–1361 Wm^{-2}.

On the other hand, some researchers suggest that the mid-1990s consensus of ~1365–1366 Wm^{-2} is more reliable (e.g., Mekaoui et al., 2010). Meanwhile, a recent analysis by Fontenla et al. suggests a theoretical value, based on physical modelling, of 1379.9 Wm^{-2} (Fontenla et al., 2011). Although Fontenla et al.'s (semi-empirical) model result is higher than any of the satellite measurements, their model (which separately modelled 9 aspects of solar variability) was able to describe the observed solar cycle trends very well over most of Solar Cycle 23, indicating that it is a very plausible result. In other words, after ~35 years of satellite measurements, it is still unclear exactly what the mean absolute value of Total Solar Irradiance has been over the satellite era, and the various values that have been proposed are still the subject of considerable debate (e.g., Mekaoui et al., 2010; Fontenla et al., 2011; Kopp and Lean, 2011; Willson, 2014).

Nonetheless, despite the ongoing debate over the absolute values, all of the satellites appear to agree on the general "rising then falling" trends over each of the "solar cycles". The solar cycle numbers during the satellite era are shown at the top of Fig. 2. These numbers were originally defined based on the sunspot cycles described earlier (Hoyt and Schatten, 1998).

By comparing the Total Solar Irradiance measurements of individual satellites to the contemporaneous sunspot records, it is clear that the sunspot cycles are indeed closely correlated to the solar cycles (e.g., Solanki and Fligge, 2000; Gray et al., 2010; Fröhlich, 2012; Willson, 2014). That is, when sunspot activity is increasing or decreasing, solar activity is rising or falling in like manner. However, while this confirms that sunspot cycles are an important component of solar activity, other solar activity components could also introduce significant longer-term ("secular") trends into the solar activity (e.g., Hoyt and Schatten, 1993; Solanki and Fligge, 2000). In other words, underlying the ~11 year sunspot cycle, there may also be longer-term trends. Hence, it is important to compare the relative magnitudes of the solar activity peaks and troughs between cycles.

So far, none of the satellites have been in operation for more than about 1.5 solar cycles. Therefore, by themselves, they cannot be used for studying the long term trends over the satellite era. Instead, to study long term trends, the individual satellites need to be composited together. Some of the scientific problems involved in the accurate calibration of satellite instruments are discussed by BenMoussa et al. (2013). These include the cavity radiometers used for measuring Total Solar Irradiance, which are strongly affected by in-orbit light, charged-particle radiation exposures, and orbital decay.

With this in mind, some more limitations of the current data are apparent from Fig. 2. A close inspection of the individual satellite measurements in Fig. 2 reveals that there are subtle, but significant differences in the trends between cycles. For instance, while the ACRIM3 and TIM satellites both are currently reporting similar values for Total Solar Irradiance, when TIM was first launched, it was recording a lower Total Solar Irradiance than ACRIM3. When we combine this with the fact that each of the satellites implies a different absolute Total Solar Irradiance value it is unclear exactly how to stitch the different satellite results together.

Additionally, while there are currently three high-precision satellites recording Total Solar Irradiance (VIRGO, ACRIM3 and TIM), there were periods when only one or two satellites are in place. One particularly problematic period is the so-called "ACRIM gap".As a result of the Space Shuttle *Challenger* disaster in 1986, the launch of the ACRIM2 satellite which had been planned to replace the ACRIM1 satellite was significantly delayed. By the time, it was finally launched, the ACRIM1 satellite's lifespan had ended. In order to bridge this gap, researchers have instead had to rely on the ERB and/or ERBE satellites, but neither of these satellites was launched for exclusive solar monitoring, making their data less reliable than the ACRIM measurements. The ERBE and ERB satellites imply slightly different trends during the ACRIM gap, and this introduces considerable uncertainty into the long term trends over the satellite era.

For each satellite, there are uncertainties over the long-term trends due to possible biases, shifts and drifts in individual satellites, and often subjective decisions have to be made over how to composite the data into one single dataset. These decisions can lead to different trends in the final composite. Currently, three groups provide composites from the various satellites and each of these composites implies quite different trends over the satellite era (Fig. 3):

- The RMIB composite implies that there has been almost no long term trend during the satellite era, but simply the ~11 year cycles oscillating about an invariant mean (e.g., Mekaoui and Dewitte, 2008; Mekaoui et al., 2010)
- The ACRIM composite implies that average solar activity increased until the end of the 20th century, but has been decreasing since then (e.g., Willson, 2014; Scafetta and Willson, 2014)

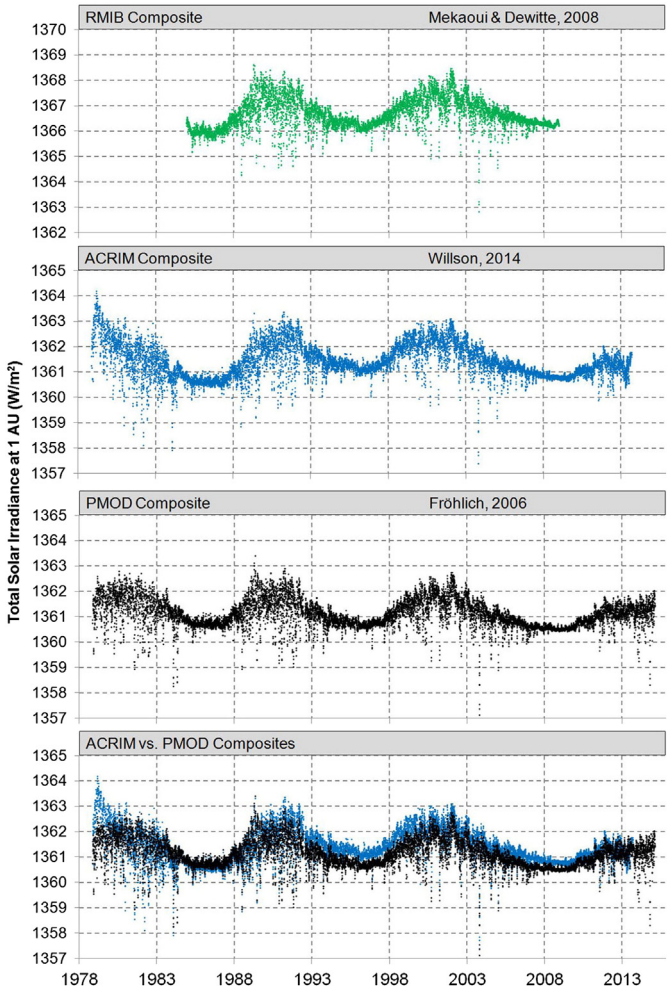

Fig. 3. Comparison of the three different composites of the satellite Total Solar Irradiance data currently available, i.e., RMIB (Mekaoui and Dewitte, 2008); ACRIM (Willson, 2014); and PMOD (Fröhlich, 2008).Note that the RMIB is scaled to a slightly different y-axis. Composites were downloaded from their respective websites on 17th March, 2015: ACRIM (http://www.acrim.com/Data%20Products.htm); PMOD (ftp://ftp.pmodwrc.ch/pub/data/irradiance/composite/DataPlots/); RMIB (http://remotesensing.oma.be/en/2619579-Data.html). Note that the dataset available on the RMIB website only goes up to 2008, while the datasets on the ACRIM and PMOD websites have been updated to 2013 and 2015 respectively. Also, the ACRIM group recommend that the pre-1980 period be treated with caution as this was not derived from ACRIM satellite data (Willson, 2014; Scafetta and Willson, 2014).

- The PMOD composite implies that average solar activity has been decreasing since the late 1970s (e.g., Fröhlich, 2006, 2012, 2013).

From Fig. 2, many readers might contend that too many inconsistencies remain between the individual satellite records, and that there is therefore still not enough data to definitively decide which (if any) of these three composites best reflects Total Solar Irradiance (TSI) trends during the satellite era. However, the American humorist, Arnold Glasow, observed that "*the fewer the facts, the stronger the opinion*", and it seems that the debate over these composites has become highly emotive and opinionated. For instance, in one paper (Fröhlich, 2012), the PMOD group claimed that a particular argument of the ACRIM group "*…has no firm ground and is simply wrong*", and that "*…the ACRIM record during* [cycle 22] *is flawed by ignoring* [an alleged slip in September 1989]". Meanwhile, in a paper by the ACRIM group (Scafetta and Willson, 2014) it is claimed that the PMOD composite "*…relies on postulated but experimentally unverified drifts in the ERB record during the ACRIM gap, and other alterations of the published ERB and ACRIM results, that are not recognized by their original experimental teams and have not been verified by PMOD by original computations using ERB or ACRIM1 data*".

Perhaps another reason for the emotive nature of this debate is the perception that the outcome of the debate may have political consequences. For instance, Zacharias (2014) argued in her conclusions that, "*A conclusive TSI time series is not only desirable from the perspective of the scientific community, but also when considering the rising interest of the public in questions related to climate change issues, thus preventing climate sceptics from taking advantage of these discrepancies within the TSI community by*, e.g., *putting forth a presumed solar effect as an excuse for inaction on anthropogenic warming*". We agree with her first justification for trying to resolve the debate, but have considerable concern over her second proposed justification. As we will discuss in Section 5, it is true that the relative role of solar variability on recent global temperature trends does have important relevance for the relative role of atmospheric carbon dioxide (i.e., "*anthropogenic warming*"). But, if the TSI community intentionally sets out to resolve the debate in order to prevent (*or* enable) climate sceptics from influencing public policies, then this could easily lead to agenda-driven instead of science-driven research.

We note that the political debate over anthropogenic global warming seems to have had a strong influence on the debate over the satellite-era TSI trends for more than a decade. For instance, in an August 2003 article over the TSI debate, Lindsey (2003) quotes Judith Lean (of the PMOD group) as saying, "*The fact that some people could use* [the upward TSI trend of the ACRIM composite] *as an excuse to do nothing about greenhouse gas emissions is one reason we felt we needed to look at the data ourselves.*" Lindsey also quotes Richard Willson (of the ACRIM group) as saying, "*It would be just as wrong to take this one result and use it as a justification for doing nothing as it is wrong to force costly and difficult changes for greenhouse gas reductions per the Kyoto Accords, whose justification using the Intergovernmental Panel on Climate Change reports was more political science than real science.*"

In our opinion, the ACRIM composite is *probably* the most reliable and the PMOD composite the least reliable of the three, for the reasons that have been outlined by the ACRIM group (e.g., Willson, 2014; Scafetta and Willson, 2014). That is, the data used for the ACRIM composite is closest to the original experimental data, and the various further adjustments to the data proposed by the PMOD group seem both speculative and fairly ad hoc. On the other hand, several researchers have formed the opposite opinion and believe that the PMOD composite is the most reliable (e.g., Lockwood and Fröhlich, 2008; Gray et al., 2010). Therefore, we would recommend treating claims that one or other of the composites is most reliable with some caution. For the interested reader, we would suggest reading some of the literature by each of the groups, i.e., PMOD (e.g., Fröhlich, 2006, 2012, 2013), RMIB (e.g., Mekaoui and Dewitte, 2008; Mekaoui et al., 2010) and ACRIM (e.g., Willson, 2014; Scafetta and Willson, 2014), and others (e.g., Chapman et al., 2013; Zacharias, 2014), before forming their own opinion.

One popular approach to attempting to resolve the controversies between the different composites has been to compare the satellite observations with easier-to-collect ground-based observations. For instance, Chapman et al. (2013) recently attempted to compare the three composites to ground-based measurements from the San Fernando Observatory. However, although their results were *slightly* better for PMOD, the differences were quite small, and they found that all three of the composites were fitted almost equally well.

Perhaps the new Daniel K. Inouye Solar Telescope (http://dkist.nso.edu/) which (at the time of writing) is under construction in Maui, Hawai'i will provide us with more detailed ground-based measurements in the future. However, an obvious limitation of using ground-based observations is that the intervening atmosphere absorbs and interacts with the incoming solar irradiance. Indeed, this was the primary motivation for carrying out the satellite measurements in the first place. For this reason, more research into relating the various ground-based observations to the true solar irradiance (e.g., Fontenla et al., 2011) may make the available data more useful.

Another approach may lie in re-assessing the physical instruments used on the various satellites. Zacharias (2014) recommends carrying out more research into re-examining the instrument designs and calibration approaches of each of the Total Solar Irradiance instruments. She suggests that, with more reliable calibration data for all of the instruments, it may be possible to create a fourth composite from the available data which would be less subjective than the current three.

At any rate, the controversies arising from the ACRIM gap and other periods when the data was particularly limited show the importance of having multiple high precision satellites independently measuring Total Solar Irradiance. At present, there are three such satellites, i.e., SOHO/VIRGO; ACRIMSAT/ACRIM3; and SORCE/TIM. However, if we hope to continue our satellite monitoring into the future, then it is important to ensure that replacement satellites for each of these are launched before the end of their lifespan, so that there is enough of an overlap to ensure a continuous composite record.

2.2. The pre-satellite era

Before the satellite era, i.e., pre-1978, we are essentially limited to ground-based data. Broadly, there are four types of data which are relevant:

1. Solar observations and measurements of particular solar phenomena — e.g., sunspot activity (e.g., Solanki and Fligge, 1999), solar rotation rates (e.g., Hoyt and Schatten, 1993; Suzuki, 2014).
2. Observations and measurements of Earth-based phenomena which may be influenced by solar activity, e.g., aurorae (Scafetta, 2012), geomagnetic activity (Cliver et al., 1998; Lockwood and Stamper, 1999; Le Mouël et al., 2009, 2012), slight variations of the order of milliseconds in the length of day (Le Mouël et al., 2010).
3. Cosmogenic isotope records, e.g., ^{10}Be from ice cores or ^{14}C from tree rings (e.g., Bard et al., 2000; Steinhilber et al., 2009, 2012), and other related records, e.g., nitrate concentrations in ice cores (e.g., Ogurtsov and Oinonen, 2014; Soon et al., 2014), which are believed to be indirectly influenced by solar activity.
4. Astronomical observations of other "Sun-like" stars (e.g., Zhang et al., 1994; Hall et al., 2009; Basri et al., 2011).

We will discuss each of these types in more detail below. However, it is important to realise that none of the above data types fully describe all aspects of solar activity, and hence can only approximate the Total Solar Irradiance. Examples of some of these "solar proxies" are provided in Figs. 4 and 5.

There is often considerable agreement between the trends and fluctuations of these different solar proxies (e.g., Gray et al., 2010; Le Mouël et al., 2012), which provides us with some confidence that together they

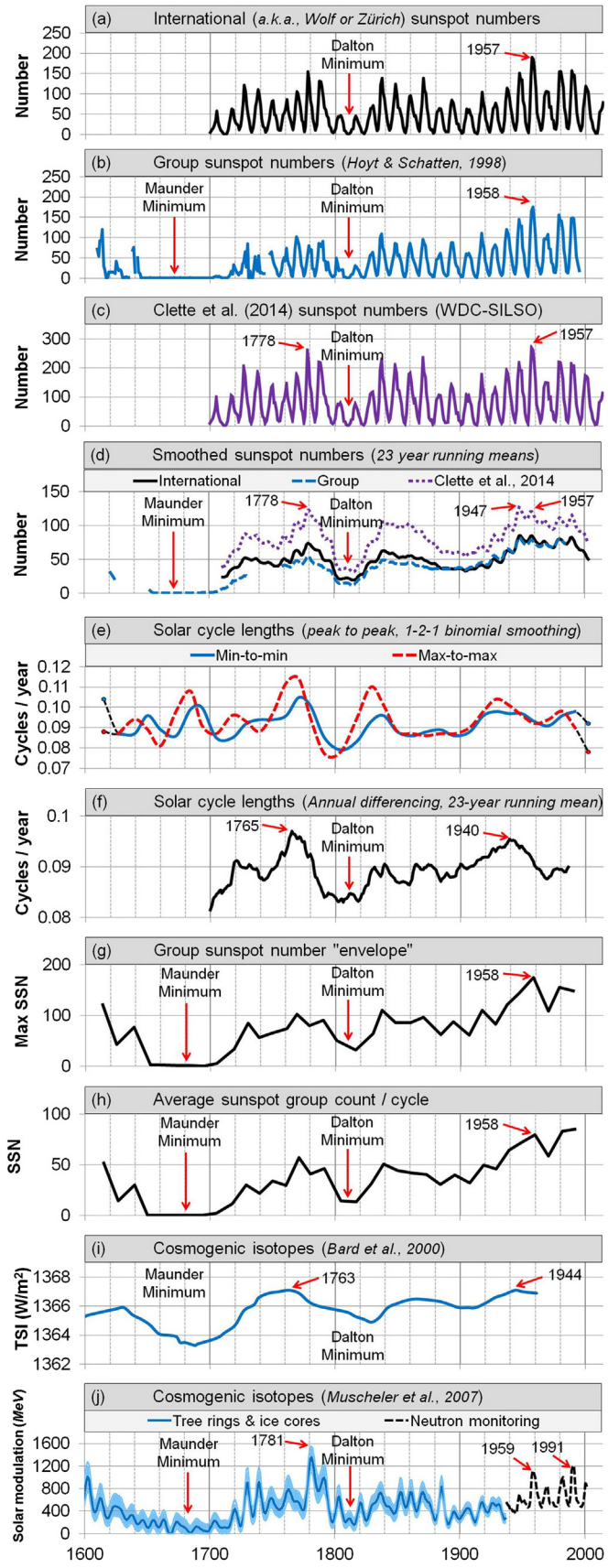

Fig. 4. Some long-term solar proxies with records stretching back to before the 19th century. (a)–(h) are all derived from the sunspot records. (i) and (j) both cover at least a millennium, but are only shown from 1600 onwards for comparative purposes. Peak years for each proxy are labelled, as are the Maunder and Dalton minima. More details of individual records are provided in the Supplementary information.

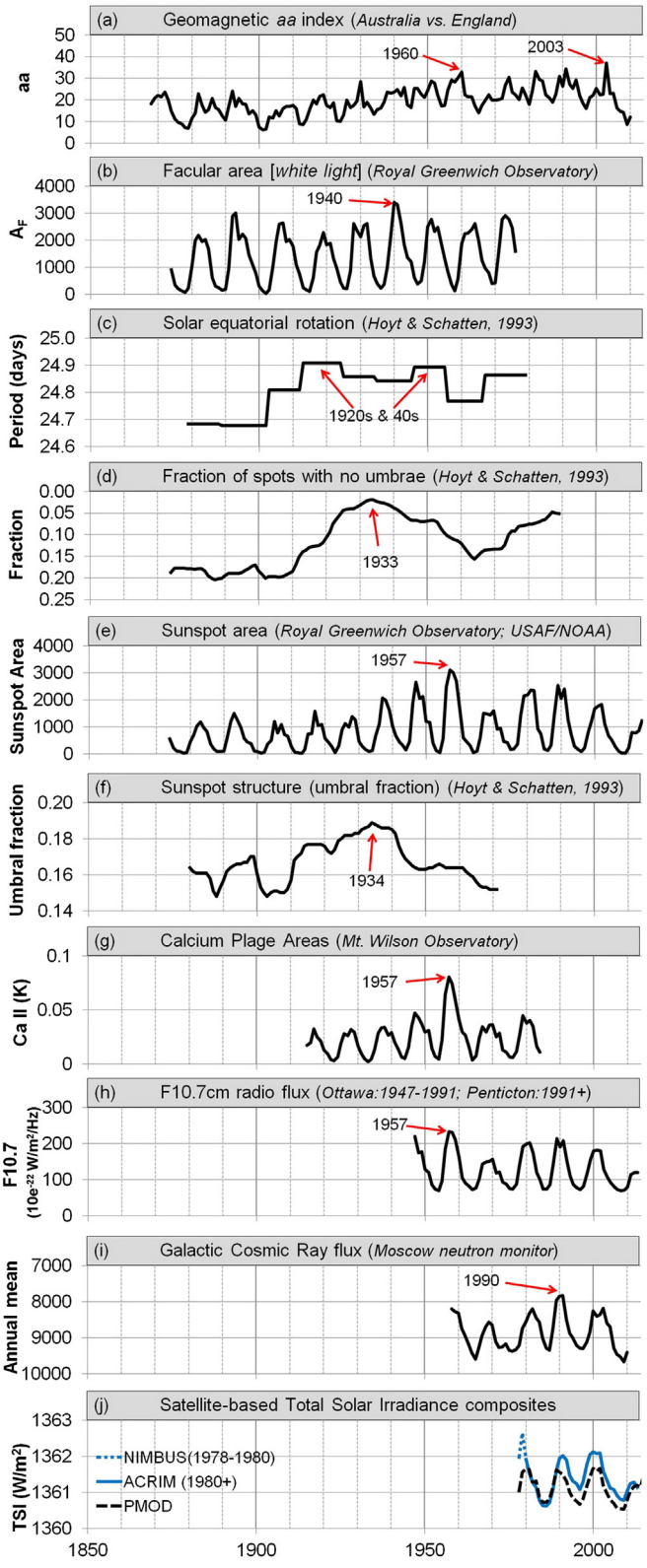

Fig. 5. Some short-term solar proxy records that began after the start of the 19th century. The y-axes for (d) and (h) are reversed, since these proxies are believed to be inversely correlated to solar activity. Peak years for each proxy are labelled. More details of individual records are provided in the Supplementary information.

are managing to capture some common aspects of the total solar variability we are interested in. However, there are also noticeable differences between individual proxies — sometimes subtle but sometimes pronounced (e.g., Hoyt and Schatten, 1993; Solanki and Fligge, 2000;

Le Mouël et al., 2012). For instance, Le Mouël et al. (2009) noted a distinctive "M-shape" for solar variability during the 20th century for several solar proxies. However, while this feature is apparent for the proxies shown in Fig. 4(a, b, c and e), it is not as obvious for the proxies shown in Fig. 4(f, g and h). Also, while most of the proxies in Figs. 4 and 5 imply a mid-20th century local maximum in solar variability, the timing of this peak varies from proxy to proxy, e.g., some suggest it occurred during the 1950s/60s, while others suggest it occurred during the 1930s/40s. Partly, these differences may be due to the individual proxies describing different aspects of solar variability, e.g., plots of sunspot numbers imply slightly different solar trends than plots of sunspot cycle lengths (Solanki and Fligge, 1999). Partly, the differing results stem from different views among researchers as to which versions of individual datasets are most reliable, e.g., compare Clette et al.'s sunspot record (Clette et al., 2014) with Hoyt & Schatten's (Hoyt and Schatten, 1998), or consider the debate over which cosmogenic isotope records are most reliable (e.g., Bard et al., 2000, 2007; Muscheler et al., 2007a,b).

In any case, we can see from Fig. 4 that, aside from the cosmogenic isotope records, our longest solar proxy records are mostly derived from the sunspot records. Although astronomers in the 17th and 18th centuries had already noted the existence of solar phenomena such as faculae and coronal streamers, systematic records of these phenomena only started been made in the late 19th century (Harvey, 2013). More detailed aspects of solar variability such as F10.7 cm radio wave emissions and changes in the solar wind were not even discovered until the mid-20th century (Harvey, 2013). For this reason, many researchers have taken to relying on the sunspot records as our chief solar proxy. So it is important to consider carefully what we know about the relationship between sunspot numbers and total solar irradiance.

2.2.1. Sunspots, faculae and the magnetic network

The first important point to note about the sunspot number records is that the sunspot number (SSN) for a given day does *not* equal the number of sunspots on that day. Instead, the sunspot number is an index which combines information on the number of sunspot groups, the number of individual sunspots and a factor based on who measured the sunspots (Hoyt and Schatten, 1998; Clette et al., 2014). This is because the process of counting sunspots is somewhat subjective and depends on several different factors, e.g., resolution of the observer's equipment, the observer's location, and the individual thresholds used by each observer for defining sunspots.

As a result, when Rudolph Wolf first developed the concept of the sunspot number (R_Z) in the 1850s, he defined it as,

$$R_Z = k(10g + n)$$

where g was the number of sunspot *groups*, n was the number of individual sunspots and k was a correction factor for each observer. Because Wolf did his original calculations at the University of Zürich, and his successors continued his analysis after his retirement, this index is often referred to interchangeably as the Wolf Sunspot Number or the Zürich Sunspot Number. However, more recent updates of this index have been carried out by researchers based at other international observatories, and so the name International Sunspot Number is also used (Clette et al., 2014).

When Hoyt and Schatten (1998) were developing their own version of the sunspot number (which they called the "Group Sunspot Number", R_G), they did not include individual sunspots, but only considered counts of sunspot groups, since they felt these were less subjective. For comparison with the Wolf numbers, they normalised their values to have the same mean value over the 1874–1976 period (when the Royal Greenwich Observatory was the primary observer).

We can see from Fig. 4(a) and (b) that the International and Group sunspot numbers have a lot of similarities, e.g., both imply a period of low sunspot activity in the early 1800s (the "Dalton minimum") and sunspot activity increased during the early 20th century to peak at around 1957/58. However, there are some important differences, e.g., the Group numbers are about 25–50% lower than the International numbers before 1882 (Hoyt and Schatten, 1998).

The definition of a sunspot group is itself somewhat subjective, and each observer might make a slightly different interpretation. In 1966, Patrick McIntosh developed a classification system modified from an earlier classification system by the Zürich group (McIntosh, 1990) — see Fig. 6. This system has since become widely adopted, but it is possible that earlier observers might have recorded different results by using their own classification systems. Hence, there is an uncertainty associated with the observer correction factors.

Recently, Clette et al. (2014) have questioned the accuracy of some of the observer correction factors used for both versions. They have proposed a new third index of sunspot numbers covering the period 1700–present — Fig. 4(c). According to this new index, sunspot activity during the 20th century was less unusual and peaks occurred in the 18th, 19th and 20th centuries.

Although the International numbers only begin in 1700, Hoyt & Schatten also included extra observations which they had compiled, and were able to extend their Group number records back to 1610 (Hoyt and Schatten, 1998). According to the Group sunspots, during the period 1640–1715 sunspots seem to have become very rare, with almost no sunspots recorded. This period is now referred to as the "Maunder Minimum" after a late 19th/early 20th century husband and wife team who highlighted the apparent existence of this prolonged sunspot minimum[3] (Eddy, 1976; Hoyt and Schatten, 1997; Soon and Yaskell, 2003; Usoskin et al., in press).

Not only does there seem to have been very few sunspots during this period, but there is evidence that suggests that those few sunspots which did occur may have had unusually long lifetimes (see p. 21 of Hoyt and Schatten, 1997). The Maunder Minimum seems to have broadly coincided with a period called the Little Ice Age which is believed to have been relatively cold globally (Eddy, 1976). If we treat the sunspot record as a proxy for solar activity, then this seems to agree with the suggestion that a reduction in solar activity leads to a corresponding reduction in global temperatures.

Indeed, by comparing sunspot numbers to the satellite measurements of Total Solar Irradiance during the satellite era, it has been shown that sunspot numbers are closely correlated to Total Solar Irradiance (e.g., Solanki and Fligge, 1999; Foukal et al., 2006; Gray et al., 2010; Fröhlich, 2012; Willson, 2014). However, this is counter-intuitive as sunspots are themselves dark regions of the Sun, and actually have a reduced irradiance — see Table 1. The reason for the increase seems to be that periods of high sunspot activity are *also* typically associated with the formation of other solar features called "faculae" which *are* brighter than average. The increase in solar irradiance from these faculae tends to outweigh the decrease in solar irradiance from sunspots, and so the *net* result is an increase in Total Solar Irradiance (e.g., Solanki and Fligge, 1999; Foukal et al., 2006; Gray et al., 2010; Fröhlich, 2012; Willson, 2014).

In other words, the apparent correlation between sunspot numbers and Total Solar Irradiance is only a commensal correlation, as opposed to a causal one. This introduces a major problem in relying on sunspot numbers as a proxy for solar activity. By comparing parallel measurements of sunspot numbers and solar activity over one or more solar cycles, it is possible to derive a linear relationship between the two. If this relationship is then extrapolated beyond the period of overlap of the two sets of measurements, then the much longer sunspot records could be used to extend our satellite era estimates of solar activity

[3] Technically, Eddy (1976) named the Maunder Minimum after E. Walter Maunder. However, most of Maunder's observations of the Sun and works on synthesizing and interpreting the solar results were collaborative efforts with his wife, Annie Scott Dill Maunder, née Russell (Soon and Yaskell, 2003). At the time, women were discouraged from publishing, so the role and contribution of Annie Maunder is largely hidden. This is the reason why we propose a more correct attribution and honour for this unique husband and wife team of solar observers.

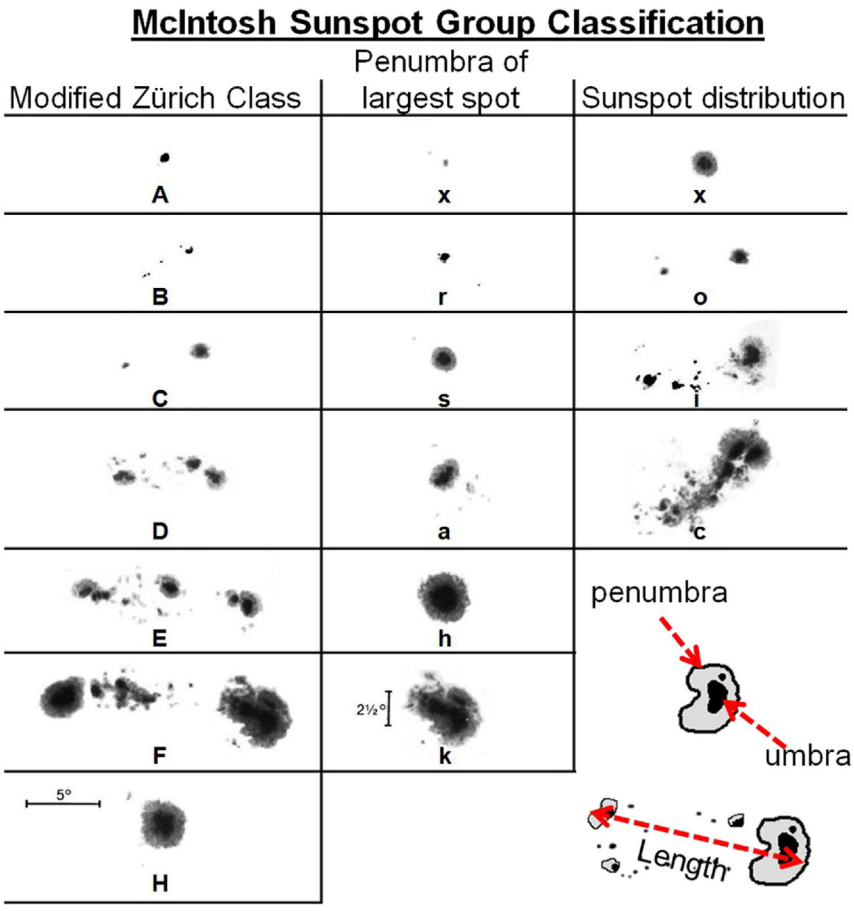

Fig. 6. The three component McIntosh classification system for sunspot groups. Adapted from Fig. 1 of McIntosh (1990).

back to the 17th or 18th centuries. However, if the commensal relationship between sunspot numbers and Total Solar Irradiance varies between solar cycles, then these extended estimates could be unreliable.

Also, when sunspots and faculae form, they only occupy relatively small parts of the Sun, which are known as "active regions". However, outside of these active regions, there is a third manifestation of the varying solar surface magnetic flux. Across the solar surface, there is a large network of small-scale magnetic domains which, like faculae, are brighter than average. This "bright network", sometimes referred to as the "magnetic carpet" (Harvey, 2013), also contributes to the Total Solar Irradiance. However, perhaps because of the ubiquitous, yet less dramatic, nature of this network, it has not been as widely studied as the more obvious sunspots and faculae. At any rate, so far, no consensus has been reached on how the network has changed over time, if at all (Harvey, 2013; Yeo et al., 2013). Therefore, one plausible way in which secular trends could be occurring independently of sunspot activity would be if there are also long-term changes in the magnetic network (Solanki and Fligge, 1999, 2000).

Table 1
Comparison of approximate emission temperatures for various solar regions, taken from Gray et al. (2010), to the average blackbody (BB) temperature of the Sun, i.e., ~5800 K. $\sigma = 5.67 \times 10^{-8}$ W m^{-2} K^{-4} is the Stefan–Boltzmann constant.

Region	Temperature	Solar output	Ratio to BB
	T (K)	σT^4 (W m^{-2})	$\sigma T^4/\sigma T_{BB}{}^4$
Sunspots (central umbra)	~4200	1.76×10^7	0.27 (27%)
Sunspots (penumbra)	~5700	5.99×10^7	0.93 (93%)
Non-active photosphere	~6050	7.60×10^7	1.18 (118%)
Faculae	~6200	8.38×10^7	1.31 (131%)
BB, i.e., Sun as "blackbody"	~5800	6.42×10^7	1.00 (100%)

With all of this in mind, it is important to look carefully and critically at what we currently know about the relationship between sunspot activity and Total Solar Irradiance trends. In Fig. 7, we have compared the International Sunspot Numbers to two of the satellite-based Total Solar Irradiance composites we discussed in Section 2.1, i.e., the PMOD and ACRIM composites.

For Fig. 7(a), we used linear least squares fitting to derive a linear relationship between the PMOD composite and the sunspot numbers,

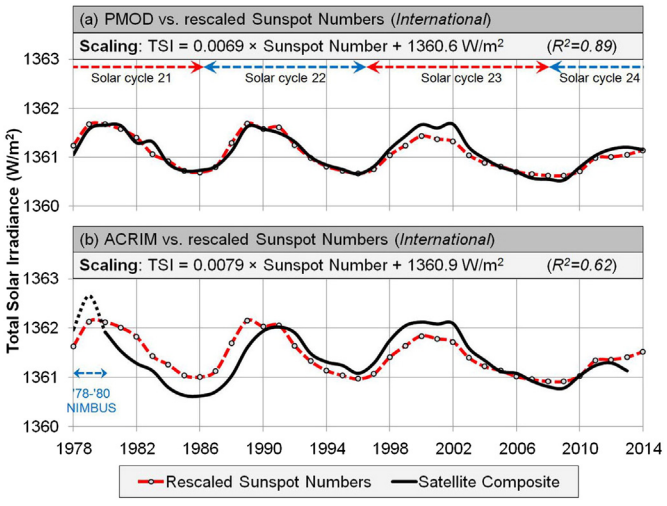

Fig. 7. Comparison of the (a) PMOD and (b) ACRIM estimates of Total Solar Irradiance trends to the International Sunspot Numbers (after rescaling to give the best fit).

and then rescaled the sunspot numbers accordingly. For Fig. 7(b), we repeated the process for the ACRIM composite. The derived linear relationships (and their corresponding R^2 values) are shown in the respective panels.

There are several interesting points to note:

- We can see that the ~11 year solar cycles of *both* satellite composites matches reasonably well with the ~11 year sunspot cycles. That is, when Total Solar Irradiance increases or decreases, sunspot activity also tends to have the same trend, as others have already noted (e.g., Solanki and Fligge, 1999; Foukal et al., 2006; Gray et al., 2010; Fröhlich, 2012; Willson, 2014).
- The relationship between sunspot numbers and Total Solar Irradiance seems to be a lot closer when we consider the PMOD composite ($R^2 = 0.89$) than when we consider the ACRIM composite ($R^2 = 0.62$).
- However, even for the PMOD composite, the relationship is *not* exact, i.e., $R^2 < 1$. For instance, the peak Total Solar Irradiance for Solar Cycle 23 was slightly higher for the PMOD composite than implied by the rescaled sunspot numbers.
- So far, we only have data from less than 3.5 solar cycles data upon which to make our comparison.

Researchers who favour PMOD over ACRIM might use the stronger correlation between sunspot numbers and the PMOD composite over solar cycles 21, 22 and 24 (and to a lesser extent 23), both as evidence in support of the reliability of PMOD, and to argue that the sunspot numbers can be used as a fairly reliable proxy for Total Solar Irradiance. Indeed, as we will discuss later, several researchers who favour PMOD over ACRIM have relied quite heavily on sunspot numbers for their Total Solar Irradiance reconstructions (e.g., Lean et al., 1995; Lean, 2000; Wang et al., 2005; Krivova et al., 2007, 2010; Vieira et al., 2011).

On the other hand, if the ACRIM composite transpires to be more reliable than the PMOD composite, then the small (but noticeable) deviations between Total Solar Irradiance and the rescaled sunspot numbers in Fig. 7(b) indicate that we should be wary of over-relying on sunspot numbers as a direct proxy for Total Solar Irradiance. In other words, it is plausible that the sunspot number records are failing to detect some of the long-term ("*secular*") trends which may be occurring between solar cycles.

The ACRIM group claim that their composite is more "*data driven*" in that it is based on the original experimental data, while the PMOD group is more "*model driven*" in that some adjustments seem to have been made to the data partially to better match certain solar proxy models (Willson, 2014). However, the PMOD group claim that their data adjustments have experimental justification, and that they do *not* make their adjustments only to improve the agreement with their proxy models (Fröhlich, 2012).

We know that the Sun still emits sunlight during the periods when no sunspots occur. That is, there is an underlying Total Solar Irradiance during the minima of the solar cycles. Indeed, the cosmogenic isotope records (e.g., Fig. 4(i) and (j)) imply that there was even solar variability throughout the Maunder Minimum (Bard et al., 2000; Muscheler et al., 2007a,b), even though sunspots seem to have been very rare during that period (Eddy, 1976; Hoyt and Schatten, 1997, 1998; Soon and Yaskell, 2003; Usoskin et al., in press).

In recognition of this fact, some researchers have suggested alternative methods to using the raw sunspot numbers as a solar proxy to avoid implying that zero sunspots means zero solar variability. For instance, Reid (1991) used the peak values of the solar cycle maxima as a measure of the "envelope" of solar activity — Fig. 4(g).

Although Lean et al. (1995) used the raw sunspot numbers as a proxy for the ~11 year solar cycle, they also used a similar technique to generate the background secular trends for their total solar irradiance reconstruction. Instead of using the solar cycle maxima, they calculated the average sunspot numbers for each solar cycle — Fig. 4(h). Comparing Fig. 4(g) and (h), we can see that both of these proxies are fairly similar.

Another approach to avoiding the "zero sunspot" solar cycle minima problem is to apply a running mean (or other smoothing function) to the sunspot numbers to smooth out the minima and maxima, e.g., the 23 year running means in Fig. 4(d). We suspect Hoyt and Schatten (1993) used an approach similar to this for determining the "*mean level of solar activity*" component of their Total Solar Irradiance reconstruction. However, if solar activity does indeed vary significantly below the zero sunspot minimum threshold, then even using such smoothing will not completely remove the problem. That is, because all levels of solar activity with no sunspots are treated as equivalent (i.e., zero sunspots), any solar variability which occurs below the zero sunspot threshold will be undetectable from the sunspot record (whether smoothed or not).

Instead of using sunspot numbers, some researchers have suggested using the lengths of solar cycles as a proxy for solar activity (e.g., Friis-Christensen and Lassen, 1991; Hoyt and Schatten, 1993, 1997; Solanki and Fligge, 1999, 2000; Thejll and Lassen, 2000; Solheim et al., 2012). In general, it seems that during periods of high sunspot activity, the solar cycle length is shorter than during periods of low sunspot activity (Friis-Christensen and Lassen, 1991; Hoyt and Schatten, 1997).

That is, the solar cycle length seems to be inversely related to sunspot activity. Indeed, by studying a sample of Sun-like stars, one of us (WS) found that stellar cycle length also seems to be correlated with stellar brightness (Baliunas and Soon, 1995). However, the relationship between solar cycle length and sunspot activity is not an exact one, and the minima and maxima implied by the solar cycle lengths sometimes differ from those implied by the sunspot numbers.

This suggests that the solar cycle lengths are capturing slightly different (albeit related) aspects of solar variability than sunspot numbers. Because the satellite era has so far only covered about 3.5 solar cycles, the exact relationship between solar cycle length and total solar irradiance is still unclear. However, it is certainly possible that some of these aspects may be important for estimating the secular trends in total solar irradiance (e.g., Solanki and Fligge, 1999, 2000).

This has prompted some researchers to use solar cycle lengths as a solar proxy. In particular, a widely-cited study by Friis-Christensen and Lassen (1991) found an apparently strong correlation between solar cycle lengths and Northern Hemisphere surface air temperatures. However, later studies revealed some problems with that study (e.g., Thejll and Lassen, 2000; Laut, 2003; Damon and Peristykh, 2005; Stauning, 2011).

One of the main problems in relying on solar cycle lengths as a solar proxy is that because each solar cycle lasts ~11 years, there are very few data points available for the modern era, e.g., we can see from Fig. 7 that the satellite era has so far only comprised ~3.5 solar cycles. Another problem is that you can obtain slightly different values depending on the method (and/or smoothing choice if applied) you use for calculating the lengths, e.g., compare the two curves in Fig. 4(e) — one calculated using solar cycle minima and the other using solar cycle maxima.

In an attempt to reduce these problems, Friis-Christensen and Lassen (1991) averaged together the values calculated from the solar cycle minima with those from the solar cycle maxima and applied a 5-point smoothing (1,2,2,2,1). However, this smoothing meant that their analysis would have finished in the late-1960s. So, to extend their analysis, they also included the most recent data points unsmoothed. Unfortunately, this artificially increased the apparent strength of the correlation to Northern Hemisphere surface air temperatures, and the apparent correlation became much less compelling when the data was updated and alternative smoothing methods were applied (e.g., Thejll and Lassen, 2000; Laut, 2003; Damon and Peristykh, 2005; Stauning, 2011).

Still, the fact that the initially impressive correlation identified by Friis-Christensen and Lassen (1991) no longer seems as compelling does not in itself mean that solar cycle lengths are a poor solar proxy.

Moreover, Hoyt and Schatten (1993) developed an annually resolved version which overcomes some of the above problems. By comparing the annual sunspot numbers to the maximum sunspot numbers for that cycle, and then tracking the changes in these ratios, they were able to estimate the effective number of cycles per year for that year. They used the 23 year running mean of this measure as one component in their Total Solar Irradiance reconstruction, which we will discuss later.

We have plotted a digitized version of Hoyt and Schatten (1993)'s annually resolved solar cycle lengths in Fig. 4(f). The 23 year running means of this solar proxy shares a lot of similarities to the 23 year running means of the raw sunspot numbers, and to a lesser extent with the sunspot envelope and average sunspot groups/cycle, particularly until the mid-20th century — compare Fig. 4(d) and (f), (g) and (h). This suggests that the annually resolved solar cycle lengths are capturing some of the same aspects of solar variability as the sunspot numbers. However, there are also some key differences in the trends, suggesting that the two types of proxies may also be capturing some slightly different aspects of solar variability.

In particular, while the International and Group sunspot numbers imply solar activity peaked around 1958 and remained fairly constant until the end of the 20th century, the solar cycle lengths imply an earlier peak in the 1940s and a decrease in solar activity until the 1970s. Interestingly, we note that Clette et al. (2014)'s recent update to the sunspot number records seems to have *slightly* reduced the apparent differences between the two types of solar proxy, particularly for the smoothed versions — compare Fig. 4(d) and (f).

Since faculae are the bright components and sunspots are the dark components of the active regions (see Table 1), the increase in Total Solar Irradiance during periods of high sunspot activity is mostly a consequence of the increase in facular emission, and not the increase in sunspots (Foukal et al., 2006). Therefore, we might expect facular areas to be a more relevant proxy for Total Solar Irradiance than sunspot numbers. Unfortunately, the available data for facular areas is still very limited.

The Royal Greenwich Observatory did maintain a record of facular areas for just over a century, starting in 1874 (Foukal, 1993). However, this record finished in 1976, i.e., just before the start of the satellite era. We plot this record in Fig. 5(b) (digitized from Foukal, 1993).

In general, the facular areas implied by this record were closely correlated to the sunspot numbers and also the Royal Greenwich Observatory's parallel measurements of sunspot areas. However, unlike the sunspot records which implied a maximum of solar activity in the late 1950s, the facular areas peaked in the 1940s and implied a decline in solar activity during the 1950s and 1960s (Foukal, 1993). Interestingly, this is somewhat reminiscent of the trends for the solar cycle lengths discussed above.

If this apparent breakdown in the relationship between sunspots and faculae is genuine, then this poses problems for relying on sunspot numbers as a solar proxy. Some researchers have argued that there are problems with the Royal Greenwich Observatory facular area records (e.g., Foukal, 1993; Chapman et al., 2011). However, if we are to discard this pre-satellite data, then our understanding of the facular-sunspot relationship is limited to a few solar cycles. In other words, it is still possible that the secular trends in faculae (and hence, Total Solar Irradiance) are different from those in sunspots.

At the San Fernando Observatory, Chapman et al. have been monitoring the relationships between the faculae, network and sunspot areas since 1988 (e.g., Chapman et al., 2001, 2011; Preminger et al., 2002; de Toma et al., 2013). Chapman et al. (2011) claim that the relationship between sunspots and faculae is reasonably linear, if the facular areas are calculated from singly-ionized calcium K-line images, instead of the white light images used by Royal Greenwich Observatory. However, there are some potential problems with this claim:

- In an earlier study, Chapman et al. (2001) found the facular-to-spot area ratio varied significantly over their period of analysis with the ratio increasing during cycle 22 and decreasing in the first part of cycle 23.
- While Chapman et al. (2011) were able to calculate a reasonably linear relationship over their 22 year period of analysis (1988–2009) using annually averaged data, this represented less than two solar cycles, i.e., it is unclear whether this relationship would hold over longer timescales.
- Indeed, using the same dataset, de Toma et al. (2013a) argued that the reason why the apparent reduction in Total Solar Irradiance during cycle 23 was less than implied by the reduction in sunspot numbers was because the total facular area had decreased by about the same amount as the total sunspot area. However, the reason why Total Solar Irradiance seems to be correlated to sunspot area is that the reduction in irradiance from the dark sunspots is normally outweighed by the increase in irradiance by the bright faculae (Foukal et al., 2006). So, de Toma et al. (2013a)'s results actually imply a change in the facular-sunspot relationship during cycle 23.

We note that the *exact* mechanisms as to how and why sunspots, faculae and the network occur are still poorly understood (e.g., Parker, 2009; Harvey, 2013). So, even though several groups have proposed models relating these magnetic features to the Total Solar Irradiance (e.g., Lean et al., 1995; Lean, 2000; Unruh et al., 1999; Fligge et al., 2000; Wang et al., 2005; Krivova et al., 2007; Vieira et al., 2011, etc.), the available data from which these models were derived is still surprisingly limited.

For this reason, the collection of more data on the relationships between sunspots, faculae and the network should still be a high priority for improving our understanding. In this sense, the detailed analysis of faculae and network during periods of high solar activity by Ortiz et al. (2002) and Yeo et al. (2013), and the San Fernando Observatory measurements (e.g., Chapman et al., 2011) are important programmes. It is also possible that more faculae data could still be found and digitized from the historic archives. For instance, Muñoz-Jaramillo et al. (2012) recently standardized, validated and calibrated a compilation of a century's worth of *polar* faculae (white light) measurements made by the Mount Wilson Observatory.

In the meantime, even if we stick to using only the data on sunspots, there are other aspects of solar variability (aside from the raw sunspot numbers) which *may* provide insight into the long term Total Solar Irradiance trends in the pre-satellite era.

As well as recording facular areas (and sunspot numbers), the Royal Greenwich Observatory also recorded the total sunspot areas over the period 1874–1976. Balmaceda et al. (2009) have combined these measurements with more recent measurements by other observatories to construct a sunspot area database covering more than 130 years. The relationship between sunspot area and sunspot numbers seems to be roughly linear, although not exactly (Balmaceda et al., 2009). Also, as for sunspot numbers, different observers had different measuring systems, and so caution is required when comparing sunspot area measurements by different observatories (Balmaceda et al., 2009). We note that, like the sunspot record which may be retrospectively revised and updated as early historical records are found, digitized and archived (e.g., Aparicio et al., 2014; Clette et al., 2014), some researchers are actively searching for additional historical sunspot area records which could be added to the available databanks. For instance, Aparicio et al. (2014) recently retrieved and digitized several sets of sunspot number and area measurements made at the Madrid Astronomical Observatory over the 1876–1986 period.

One aspect of solar variability related to sunspots which has received considerable attention lately is to do with the sizes of sunspots (e.g., Kilcik et al., 2011, 2014; Lefèvre and Clette, 2011; Clette and Lefèvre, 2012; Nagovitsyn et al., 2012; de Toma et al., 2013; Muñoz-Jaramillo et al., 2015). Unfortunately, depending on how the sunspot sizes are classified and the data used, researchers can

reach different conclusions. For instance, comparing the numbers of "large sunspots" between Solar Cycles 22 and 23, de Toma et al. (2013) found a decrease; Kilcik et al. (2011) found an increase; while Lefèvre and Clette (2011) found no change.

Another aspect of observed solar variability which may be relevant is the actual structure of sunspots. Most sunspots are very dark at the centre (the "*umbra*"), but are a bit brighter at the outer edges of the sunspot (the "*penumbra*") — see Fig. 6 and Table 1. However, the ratio of the umbral area to the penumbral area varies from spot to spot. Indeed, some spots have almost no umbra (Hoyt and Schatten, 1993), while others have almost no penumbra (McIntosh, 1990; Penn and Livingston, 2006).

Several researchers have noted that the average umbral/penumbral ratio can change from cycle to cycle (Hoyt, 1979; Hoyt and Schatten, 1993; Hathaway, 2013; Bludova et al., 2014) — see Fig. 5(d) and (f). However, most reconstructions of solar activity have effectively assumed that this ratio has remained unchanged over time (Hathaway, 2013). So, if this ratio has significance for Total Solar Irradiance trends, this could pose a problem. Hoyt (1979) noted a quite strong correlation ($R = 0.57$) between the umbral/penumbral ratio and Northern Hemisphere surface temperature trends. If this correlation is genuine then it suggests that the umbral/penumbral ratio is indeed significant.

A useful solar proxy for which relatively long records exist (1947–present) is the F10.7 cm radio flux (often called $F_{10.7}$ for short) — see Fig. 5(h). Although the exact sources of this flux are still being debated (see Henney et al., 2012 and Schonfeld et al., 2015 for a brief review), it seems to mostly originate from the coronal plasma in the upper atmosphere of the Sun.

Interestingly, it seems to have been closely correlated to the sunspot numbers for the first four sunspot cycles since 1947, i.e., Solar Cycles 19–22 (Svalgaard and Hudson, 2010; Le Mouël et al., 2012). This suggests that both solar proxies are capturing much of the same aspects of solar variability — despite originating in different parts of the solar atmosphere. However, Svalgaard and Hudson (2010) noted that the relationship which held reasonably well for those four cycles seemed to break down for the most recent cycle, i.e., Solar Cycle 23.

Specifically, when the observed F10.7 cm flux was used to estimate the sunspot numbers using the relationship derived from Solar Cycles 19–22, it predicted that more sunspots should have occurred than were actually observed (Svalgaard and Hudson, 2010). Livingston et al. (2012) suggested that the reason for this change in the relationship could be that the average magnetic field strength had significantly decreased during Solar Cycle 23 (Penn and Livingston, 2006). They argue that sunspots do not form unless the magnetic field strength exceeds about 1500 G and that during Solar Cycle 23 this minimum threshold was not reached as often as it was in previous cycles (Penn and Livingston, 2006; Livingston et al., 2012).

Livingston et al. (2012) suggest that once this minimum threshold is taken into account this fully explains the change in the relationship between F10.7 cm flux and sunspot numbers. We agree that this is a plausible explanation, but we do caution that while the correlation between F10.7 cm flux and sunspot numbers was quite strong for Solar Cycles 19–22, it was not perfect. So, it is possible that other changes in the relationship may occur during future cycles.

Another solar proxy which seems to have a close relationship to sunspots, faculae and the magnetic network is the solar emission at the calcium II (Ca II) "*K-line*" (i.e., at 393.4 nm) — see Fig. 5(g). This spectral line is believed to mostly originate from the chromosphere, i.e., the region of the Sun's atmosphere just above the photosphere where the sunspots, faculae and network occur. In particular, it seems to be generated with the bright "*plages*" which tend to be associated with the faculae and network. As we will discuss in Section 2.2.3, this is a particularly useful solar proxy because it can be compared to stellar emissions by other stars. Foukal et al. (2009) have compiled together more than a century's worth of Ca II measurements including the Mount Wilson Observatory measurements shown in Fig. 5(g).

2.2.2. Geomagnetic and cosmic ray-based solar proxies

Most of the proxies for solar variability are understandably derived from solar observations, e.g., sunspots, faculae, changes in the solar emission spectra. However, some widely-used solar proxies are actually based on the variability in certain terrestrial phenomena which are believed to themselves act as proxies for aspects of solar variability, e.g., the so-called "*solar wind*".

The solar wind is a stream of high energy plasma particles (e.g., electrons, protons and alpha particles) which flows outwards from the Sun's corona (i.e., upper atmosphere) throughout the Solar System. During periods of high solar activity, the strength of this solar wind seems to increase. Therefore, changes in terrestrial phenomena which are influenced by the strength of the solar wind can sometimes be used as an indirect proxy for solar activity. Several of these indirect solar proxies have been proposed.

When the solar wind nears the Earth, it interacts with the Earth's magnetic field (i.e., the "*geomagnetic field*"). Therefore, several researchers have proposed using various indices of geomagnetic activity as an indirect proxy for solar activity — specifically as a proxy for variations in the solar wind (e.g., Cliver et al., 1998; Lockwood and Stamper, 1999; Le Mouël et al., 2012).

One geomagnetic activity index which has been particularly popular among researchers is the *aa* index shown in Fig. 5(a). This is an index that is derived from roughly simultaneous measurements of the geomagnetic field which have been made since 1868 at two stations located at nearly opposite sides of the Earth, i.e., England and Australia.

Because of its relatively long and complete record, several groups have used the *aa* index for generating estimates of the long term trends in solar activity (e.g., Cliver et al., 1998; Lockwood and Stamper, 1999). Love (2011) has noted that there may be some long-term biases in the *aa* index which may have altered the magnitude of the long-term trends. Although this does not seem to have altered the shape of the curves, and they suggest it may still be *qualitatively* reliable.

At any rate, Le Mouël et al. (2012) note that there is a wide range of different geomagnetic indices which could be used as solar proxies. Indeed, Le Mouël et al. (2010) have even suggested that records of the slight variations (on the order of milliseconds) in the length of day could act as a useful proxy for solar wind variations. Many of these geomagnetic indices are closely correlated to each other and also to the sunspot records. However, the relationships are not exact, and it is still unclear which (if any) of them are best at capturing the true trends in solar activity — see Le Mouël et al. (2012) for a useful discussion and comparison of some of these indices.

In particular, while the geomagnetic indices show a lot of similarity to the sunspot records (and other solar proxies), the ~11 year cycles which are so prominent in the sunspot records are less pronounced in the geomagnetic indices. Some of the difference between the geomagnetic indices and the sunspot records may result from the fact that they are both capturing different aspects of solar variability (i.e., solar wind strength vs. sunspot formation). However, some of the difference may be due to the fact that the geomagnetic indices are only an indirect solar proxy.

Another set of indirect solar proxies are related to the numbers of galactic cosmic rays reaching the Earth. Most of the cosmic rays reaching the Earth are believed to have been generated outside the Solar System, e.g., from exploding supernovae (Svensmark, 2012; Shaviv et al., 2014). However, the strength of the solar wind appears to influence the number of cosmic rays which enter the Earth's atmosphere. Specifically, when the solar wind is strong, the number of incoming cosmic rays is reduced. As a result, the number of galactic cosmic rays entering the Earth's atmosphere seems to be inversely proportional to solar activity.

As we will discuss in Section 2.3, some solar-climate theories propose that cosmic rays have a direct influence on the Earth's climate, and so there is considerable interest in the trends in cosmic ray intensities. The annual mean measurements from one neutron monitoring station (Moscow) are shown in Fig. 5(i). However, this observation is also

useful for generating long-term solar proxies. When high energy cosmic rays reach the Earth, they occasionally interact with the atoms in the atmosphere (or rocks and soil) to generate relatively rare isotopes, such as beryllium-10 (^{10}Be) and carbon-14 (^{14}C). These *cosmogenic isotopes* behave almost identically to their regular counterparts and a result can end up being incorporated into various natural long-term records, e.g., ice cores and tree rings, respectively.

Therefore, if we assume that the average concentration of cosmogenic isotopes present in the atmosphere at any given time is proportional to the intensity of incoming cosmic rays, which is in turn inversely proportional to the strength of the solar wind, then we can use the relative concentrations of these cosmogenic isotopes as an indirect proxy for solar activity. Although this is a rather indirect type of solar proxy, the relatively long length of various cosmogenic isotope records – chiefly ice cores and tree ring chronologies – makes them one of our best solar proxies for studying long-term trends (e.g., Bard et al., 2000; Muscheler et al., 2007a; Steinhilber et al., 2009, 2012). Recently, a study involving one of us (WS) was able to identify and confirm co-variations of three solar activity proxies (namely, ^{14}C, ^{10}Be and nitrates) on timescales ranging from centuries to millennia (Soon et al., 2014).

The fact that several of these records pre-date the sunspot records makes them particularly intriguing. The post-1600 portion of two of the solar proxy records derived from cosmogenic isotopes are shown in Fig. 4, i.e., those by Bard et al. (2000) and Muscheler et al. (2007a).

There are some key differences between the long-term trends of individual cosmogenic isotope records. As a result, there has been some controversy over which ones are most representative of solar trends (e.g., Bard et al., 2000, 2007; Muscheler et al., 2007a,b; Steinhilber et al., 2009, 2012). Steinhilber et al. (2012) argue that the best approach to overcoming these discrepancies is probably to identify the common trends between multiple records. Also, it has been suggested that changes in climate could *slightly* alter the relative concentrations of cosmogenic isotopes in ice cores (Heikkilä et al., 2008; Field et al., 2009).

Nonetheless, it can be seen from Fig. 4 that the long-term solar activity trends implied by these cosmogenic isotope records are broadly similar to those implied by other solar proxies, e.g., sunspot numbers and solar cycle lengths. In particular, they suggest that the relatively low *sunspot* activity during the Maunder Minimum and Dalton Minimum did indeed correspond to periods of low solar activity. Moreover, they imply that similar periods of low solar activity have occurred several times over the centuries, e.g., the Spörer minimum during the late-15th/early-16th centuries (Eddy, 1976).

2.2.3. Astronomical observations of "Sun-like" stars

As we saw in the discussion above, at present, the currently available data from historical observations of the Sun is still very limited (Harvey, 2013). This poses a serious challenge in trying to understand how the solar variability we have observed in the relatively short available records compares to trends over timescales of centuries to millennia. An intriguing shortcut to estimating how recent solar activity compares to solar activity in the pre-instrumental period is to compare the behaviour of the Sun to astronomical observations of other "*Sun-like*" stars (sometimes called "*solar analogs*"), i.e., stars that closely match our Sun in mass and colour/effective temperature.

That is, by comparing the observed solar variability to the stellar variability of Sun-like stars, we may be able to place the recent solar trends in the context of other similar stars. This should give us an indication of how normal or unusual the recent solar variability is, and thereby how wide a range of solar variability we should expect there to have been in the pre-instrumental era.

Obviously, we cannot study another star in as much detail as we can for the Sun, since we are located in the Solar System. Instead, the chief advantage of stellar observation programmes is that you can study a large sample of stars simultaneously. This means that you can collect a large amount of data in a relatively short time. For instance, by monitoring ~30 Sun-like stars for ~20 years you would collect ~600 stellar years worth of information, which is longer than the current ~400 years of sunspot records. This information could also be useful to the wider astronomy community.

For this reason, there has been considerable interest from the solar physics community in the results of stellar observation programmes such as those by the Mount Wilson Observatory and Lowell observatory (e.g., Baliunas and Jastrow, 1990; White et al., 1992; Zhang et al., 1994; Baliunas et al., 1995; Baliunas and Soon, 1995; Radick et al., 1998; Wright, 2004; Hall and Lockwood, 2004; Judge and Saar, 2007; Lockwood et al., 2007; Hall et al., 2009; Saar and Testa, 2011; Judge et al., 2012; Shapiro et al., 2013; Basri et al., 2013 and others). One of us (WS) has been involved in some of this research (e.g., Zhang et al., 1994; Baliunas and Soon, 1995; Baliunas et al., 1995; Lockwood et al., 2007).

In 1966, Olin Wilson began a programme at the Mount Wilson Observatory of monitoring the emission in the calcium II K (393.4 nm) and H (396.8 nm) lines (relative to two nearby continuum passbands) of 91 lower main-sequence stars thought to be Sun-like. By continuing and expanding on this programme, Baliunas et al. were able to collect some data for 74 Sun-like stars with more detailed and continuous measurements for 13 of the stars (Baliunas and Jastrow, 1990; Zhang et al., 1994; Baliunas and Soon, 1995; Baliunas et al., 1995; Radick et al., 1998; Lockwood et al., 2007). As we discussed in Section 2.2.1, emission at the calcium K line is a common proxy for solar activity (Fig. 5(g)). Indeed, White et al. (1992) noted that it was possible to convert the available calcium II K solar observations into equivalent "*Sun as a star*" measurements, allowing a reasonably direct comparison with the Mount Wilson Observatory measurements.

Analysis of this data revealed that, like the solar cycles of the Sun, the majority of the observed stars also seem to go through "*stellar cycles*". However, a noticeable fraction of the observed stars seemed to be in a "*non-cycling state*" (Baliunas and Jastrow, 1990; White et al., 1992; Zhang et al., 1994; Baliunas et al., 1995). Baliunas and Jastrow (1990) suggested that these non-cycling states could be analogous to the state that the Sun might have been in during the Maunder Minimum.

Baliunas and Jastrow (1990) also noted that of the 13 stars with the most data, four seemed to have been non-cycling ("*flat*"), and that all four of these seemed to be less bright. This suggested that the ratio of the brightness for cycling and non-cycling Sun-like stars could be used as a rough estimate of the relative difference in solar activity between the Maunder Minimum and present (Baliunas and Jastrow, 1990; White et al., 1992; Zhang et al., 1994; Baliunas et al., 1995). Unfortunately, while these preliminary results are intriguing, the sample size of 13 was probably too small for drawing definitive conclusions on the Maunder Minimum to present ratio.

More recently, similar stellar observation programmes at Lowell Observatory and Fairborn Observatory have been combined to create a larger and more up-to-date database for 32 Sun-like stars (Lockwood et al., 2007; Hall et al., 2009). 27 of these stars seem to be in cycling states, suggestion that the fraction of non-cycling stars at any stage is about 15% (Lockwood et al., 2007) — about half the ~30% originally implied by Baliunas and Jastrow (1990).

Using this larger database, Hall and Lockwood (2004) failed to replicate Baliunas and Jastrow (1990)'s second result, since some of their non-cycling stars seemed to have a relatively high stellar brightness. However, it now seems that identifying appropriate analogues for Sun-like stars in the Maunder Minimum is more challenging than originally envisaged (Hall and Lockwood, 2004; Wright, 2004; Judge and Saar, 2007; Hall et al., 2009; Saar and Testa, 2011).

Interestingly, this larger database implies that for some cycling stars, the stellar brightness (equivalent to TSI) seems to decrease during the maxima of the stellar cycles (equivalent to sunspot cycles), i.e., the opposite of what has been observed for the Sun during the satellite era (Lockwood et al., 2007; Hall et al., 2009). Lockwood et al. (2007) suggest that, for younger active stars, the stellar cycles are "*spot-dominated*" and it

is only for older, less active stars (like the Sun), that the stellar cycles are "*faculae-dominated*". However, it also suggests to us further evidence that the faculae:sunspot ratio is not as constant as is often assumed on the basis of the satellite observations during Solar Cycles 21 and 22 (see Section 2.2.1).

By splitting the Shapiro et al. (2011) Total Solar Irradiance reconstruction (see Section 2.2.4) into multiple 15 year segments, Judge et al. (2012) were able to compare the secular trends during these segments to the 10–20 year trends of the Sun-like stars in the database. Although they found that the relatively high amplitudes of the solar variability implied by the Shapiro et al. (2011) reconstruction were *probably* overestimated by a factor of two, they found that the secular trends themselves were consistent with the stellar observations. However, they cautioned that to properly assess the reliability of the 30–50 year secular trends implied by the Shapiro et al. (2011) reconstruction and others, it would be highly preferable to have at least 30–50 years of stellar observations, rather than the ~15 years they had. Hence, they called for the stellar observation programmes to be continued for at least another few decades. Shapiro et al. (2013) echoed this recommendation, as do we.

With this in mind, it is disappointing that the Mount Wilson Observatory programme has been discontinued. However, we note that the recent *Kepler* mission to try and identify exoplanets (i.e., planets orbiting other stars) gathered short-term information for a very large number of Sun-like stars (Basri et al., 2013). Although this data is far too short-term (i.e., of the order of 30 min to three months) for studying secular trends (i.e., greater than 10–20 years), the data can still provide us with more information on how the Sun compares with other stars (Basri et al., 2013). It also could be used for identifying other potential Sun-like stars if future researchers wish to expand the number of stars to be considered for stellar observation.

2.2.4. Solar proxy-based Total Solar Irradiance reconstructions

From the discussion above, it should be apparent that depending on which solar proxy dataset(s) are used, different estimates of the trends and fluctuations in solar activity since the 19th century (and earlier) can be obtained. Additionally, in order to convert these datasets into estimates of Total Solar Irradiance, researchers often calibrate the datasets using one of the satellite composites discussed in Section 2.1. We saw that each of the three composites implies slightly different trends over the satellite era (Fig. 3), so the choice of which composite to use for calibration can significantly influence the longer term trends. Hence, several different reconstructions of the Total Solar Irradiance trends have been proposed. Fig. 8 shows the trends and fluctuations of eight of these reconstructions since the 19th century which have recently been proposed (or updated).

Although each of the reconstructions uses a different method and/or model for converting the solar proxies into Total Solar Irradiance values and incorporates some information from more than one type of proxy, four of the estimates are *primarily* based on the sunspot numbers and related proxies (e.g., sunspot areas), i.e., Lean et al. (1995); Wang et al. (2005); Krivova et al. (2007; updated Krivova et al. 2010) and Vieira et al. (2011). A fifth reconstruction, Shapiro et al. (2011), also used sunspot numbers to calculate the high frequency variability of their estimate.

On the other hand, while Hoyt and Schatten (1993; updated by Scafetta & Willson, 2014) also used sunspot numbers as one component of their reconstruction, their reconstruction was actually a composite of several different solar proxies. The exact proxies they used is a little unclear from the text, although there is some further information in Hoyt and Schatten (1997). However, their composite seems to have been constructed from the following:

- 1700–1992: The international sunspot numbers — Fig. 4(a) — rescaled to match the Nimbus 7 satellite Total Solar Irradiance measurements, i.e., Fig. 2.

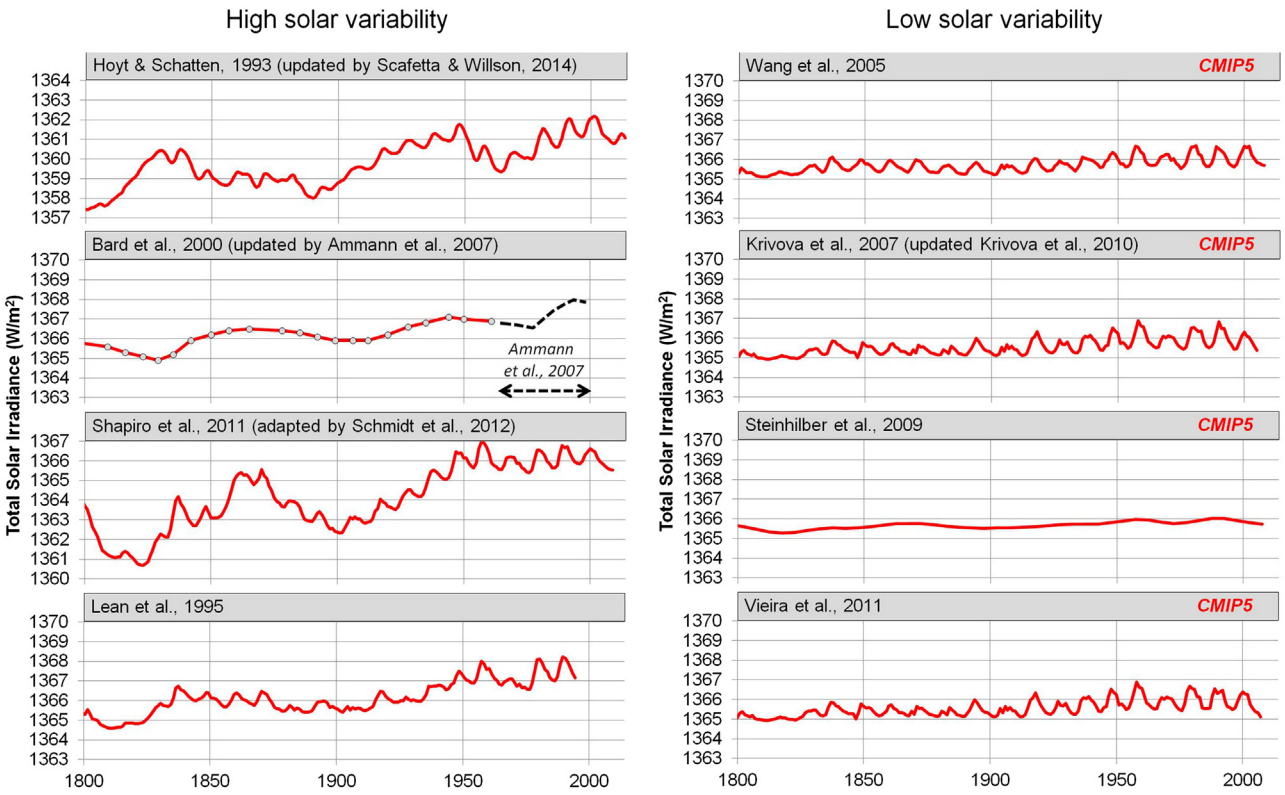

Fig. 8. Some of the proposed Total Solar Irradiance (TSI) reconstructions (shown since 1800). The four reconstructions on the right are ones that were used as "natural forcings" for the CMIP5 Global Climate hindcasts in the Intergovernmental Panel on Climate Change (IPCC)'s 5th Assessment Report (5AR) All eight reconstructions are shown with the same scale, but the absolute values of the y-axes have been varied to fit.

- 1700–1978: The "*mean level of solar activity*" — we suspect this is the 23 year running mean of the sunspot numbers, i.e., Fig. 4(d).
- 1700–1978: The mean solar cycle lengths, i.e., Fig. 4(e).
- 1700–1978: The "*normalized rate of solar cycle decay*", i.e., the annually-resolved solar cycle lengths in Fig. 4(f).
- 1874–1990: The fraction of spots with no umbra, i.e., Fig. 5(d) and, when available, the umbral fraction, i.e., Fig. 5(f).
- 1879–1979: The solar equatorial rotation rate, i.e., Fig. 5(c).
- 1979–1992: Nimbus 7 satellite Total Solar Irradiance measurements, i.e., top left dataset in Fig. 2.

Scafetta and Willson (2014) used the post-1980 ACRIM composite in Fig. 3 instead of the Nimbus 7 satellite data for the updated version used in this paper (1700–2013).

The remaining two estimates (Steinhilber et al., 2009; Bard et al., 2000; updated by Ammann et al., 2007) were derived from cosmogenic isotope records, as was the pre-1640 portion (not shown here) of the Vieira et al. (2011) estimate. Shapiro et al. (2011) also used cosmogenic ^{10}Be isotope records to calculate the low frequency variability of their estimate.

As mentioned earlier, much of the debate over the various long-term solar activity reconstructions has centred over how recent solar activity compares to solar activity during the Maunder Minimum. In this regard, there are currently two main groups of reconstructions — those arguing that there has been a substantial increase in solar activity since the Maunder Minimum ("high solar variability"), and those arguing that the increase has been fairly modest ("low solar variability") — see Fig. 8. Some of the different estimates which have been made of the approximate difference in Total Solar Irradiance between the Maunder Minimum and present are listed in Table 2.

Interestingly, Lean et al. have repeatedly reduced the variability of their reconstruction from their original 1995 estimate (Lean et al., 1995) to their 2000 estimate (Lean, 2000) to the current Wang et al. (2005) estimate. We can see from Fig. 8 that the Lean et al. (1995) reconstruction implied a relatively high solar variability. However, the more recent update of this reconstruction by Wang et al. (2005) is a low solar variability reconstruction. With this in mind, in Fig. 9, we compare all three of these Lean et al. reconstructions.

Lean et al. (1995) consisted of two components — both of which were derived from the group sunspot numbers. One component described the ~11 year solar cycle and was derived by rescaling the sunspot numbers to match satellite-based estimates of Total Solar Irradiance in the satellite era, and combining this with information from sunspot area measurements from the Royal Greenwich Observatory for the post-1874 period. The other component was to describe the long-term background trends, and was based on the average number of sunspot groups in each cycle, i.e., Fig. 4(h). This background component was scaled to match the estimates of the difference between the Maunder Minimum and the present implied by the stellar and solar

Table 2
Various estimates of the difference in Total Solar Irradiance between the Maunder Minimum and present.

Study	Maunder Minimum vs. present difference
Shapiro et al. (2011)	3–9 W/m^2
Bard et al. (2000)	3.4–8.9 W/m^2
Zhang et al. (1994)	2.7–8.2 W/m^2 (i.e., 0.2–0.6%)
Cliver et al. (1998)	~7.4 W/m^2 (i.e., 0.54%)
Lean et al. (1995)	~3.3 W/m^2
Lean (2000)	~2.8 W/m^2
Wang et al. (2005)	~1 W/m^2
Steinhilber et al. (2009)	0.5–1.3 W/m^2
Vieira et al. (2011)	~1.3 W/m^2
Krivova et al. (2007)	0.9–1.5 W/m^2

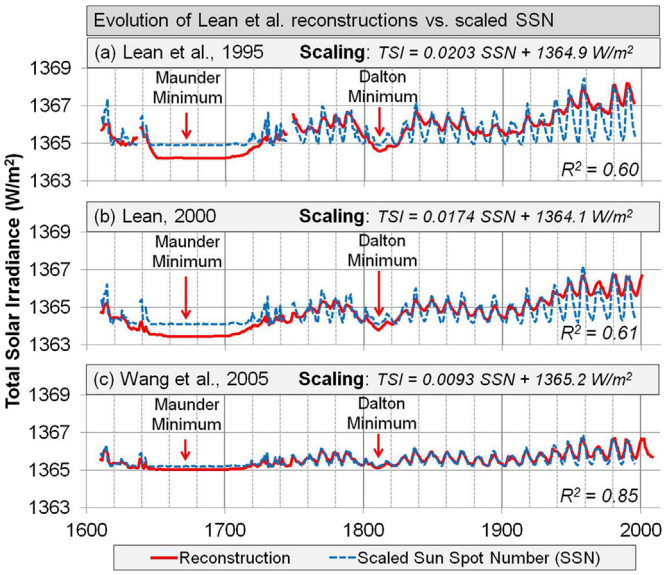

Fig. 9. Plots showing how the three Lean et al. Total Solar Irradiance (TSI) reconstructions compare to a simple linear scaling of the group sunspot numbers (SSN). The linear scaling functions and associated R^2 values for each reconstruction are shown in the corresponding panels.

observations of White et al. (1992) and Baliunas and Jastrow (1990). We can see from Fig. 9(a) that the Lean et al., 1995 reconstruction is heavily derived from the sunspot number record, but implies a relatively high solar variability due to this background component.

The Lean (2000) reconstruction was similarly split into two components. For the ~11 year cycle component, Lean used a series of different solar proxies for different portions of the reconstruction. For the background component, she used a 15 year running mean of the group sunspot numbers, instead of using the average group counts per cycle. Both of these components yielded broadly similar results to the Lean et al., 1995 versions. However, when scaling the background component, she reduced the magnitude of the difference between the Maunder Minimum and present. As a result, the solar variability of the Lean, 2000 reconstruction was lower than the earlier version.

Wang et al. (2005) developed a rather complex theoretical model for their reconstruction which was based on the results of magnetic flux transport simulations. However, despite the complex nature of their theoretical model, it can be seen from Fig. 9(c) that their reconstruction is actually not very different from a simple rescaling of the sunspot number records.

Climate modellers contributing to the Coupled Model Intercomparison Project Phase 5 (CMIP5) that was used for the latest Intergovernmental Panel on Climate Change (IPCC) reports were all specifically recommended by the organizers to use the Wang et al., 2005 dataset (http://solarisheppa.geomar.de/cmip5), so not surprisingly all of the hindcasts gave similar results. However, since the Wang et al., 2005 dataset only begins in 1610, some modellers that were also contributing to the related Paleoclimate Model Intercomparison Project Phase III (PMIP3) decided to use slightly different datasets so that they could extend their hindcasts further into the past. HadCM3 used Steinhilber et al. (2009) and several others used Vieira et al. (2011) for the pre-1850 part of their hindcast, while MPI-COSMOS used one by Krivova et al. (2007; updated by Krivova et al., 2010) for their entire reconstruction, following the recommendation of Jungclaus et al. (2010)(see Schurer et al., 2013; Schmidt et al., 2011, 2012 for more details).

Although these exceptions meant that the CMIP5 modellers were not all using exactly the same solar dataset, it can be seen from Fig. 8 that Wang et al., 2005 and these three additional datasets (Krivova

et al., 2007; Steinhilber et al., 2009; Vieira et al., 2011) all are "low solar variability" datasets. That is, they imply there has been very little solar variability since 1800 when compared with some of the other datasets on the left hand side, e.g., Hoyt and Schatten, 1993 (Hoyt and Schatten, 1993; updated by Scafetta and Willson, 2014), Bard et al., 2000 (Bard et al., 2000; updated by Ammann et al., 2007) or Shapiro et al., 2011 (Shapiro et al., 2011; adapted by Schmidt et al., 2012).

As we will discuss in Section 4.5, this could partially explain why the Intergovernmental Panel on Climate Change (IPCC) failed to find much of a role for natural climate change in recent global temperature trends in their latest reports (Bindoff et al., 2013). Partly for this reason, in Section 5, we will explicitly consider one of the high solar variability datasets not considered by CMIP5, i.e., the updated version of the Hoyt and Schatten, 1993 dataset (Hoyt and Schatten, 1993; updated by Scafetta and Willson, 2014).

2.3. Possible "amplification" mechanisms for solar-climate links

In the previous sections, we considered various estimates of the trends in Total Solar Irradiance, since this seems to be the solar variability factor which should have the most obvious effect on surface air temperature trends. However, several researchers have suggested that there could be other subtle and/or indirect mechanisms by which solar variability could significantly influence the Earth's climate.

If these hypothesised mechanisms are valid then it is possible that solar variability could have a pronounced effect on the Earth's climate over a given time period even if the corresponding changes in Total Solar Irradiance were modest. In that case, the effects of solar variability on climate change would appear to have been "*amplified*" relative to the smaller changes in Total Solar Irradiance.

A detailed discussion of these solar-climate amplification mechanisms is beyond the scope of this review. However, the controversy over these mechanisms has comprised a substantial fraction of the debate over possible solar-climate links. So, here, we will provide a brief overview of the main proposed amplification mechanisms which have been considered in the literature.

Many of these mechanisms are based on the possible influence of galactic cosmic rays on the Earth's climate—see Kirkby (2007) for a review. However, so far, the evidence for and against such mechanisms has been quite ambiguous and controversial.

One proposed cosmic ray-climate mechanism which has received considerable attention is the suggestion by Svensmark et al. that cosmic rays could play an important role in the formation of clouds (Svensmark and Friis-Christensen, 1997; Marsh and Svensmark, 2000, 2003; Svensmark, 2007; Svensmark and Calder, 2007; Svensmark et al., 2009, 2013). If the presence of cosmic rays increases the probability of cloud formation then it is plausible that periods of low solar activity could lead to generally cloudier conditions (since more cosmic rays enter the Earth's atmosphere when the solar wind is weaker). Cloudier conditions could increase the Earth's planetary albedo (Stephens et al., 2015), reducing the amount of incoming solar radiation absorbed by the Earth system, and thereby cool the planet (Svensmark, 2007). Svensmark and Calder (2007) provide an accessible introduction to this theory.

Svensmark and Friis-Christensen (1997) found an apparently strong correlation between global cloud cover and observed cosmic rays by using a satellite-based cloud cover dataset which covered the period 1980–1995. This initially seemed to offer strong support for this mechanism. However, the correlation does not seem to have held for updated versions of the dataset (Soon et al., 2000; Laut, 2003). Marsh and Svensmark (2000) modified the original theory, by suggesting that the proposed effect was most pronounced on low-lying clouds, and that there was an apparently strong correlation between cosmic rays and low-lying clouds. Yu (2002) offered further support for modifying the theory to take into account the altitudes of clouds. However, when alternative versions of the low cloud cover and cosmic ray datasets were used, this correlation was not apparent (Laut, 2003).

Although some more recent studies have offered support to the theory (e.g., Marsh and Svensmark, 2003; Harrison and Stephenson, 2006; Svensmark et al., 2009, 2013; Yu and Luo, 2014), others have criticised the theory (e.g., Laken et al., 2012; Sloan and Wolfendale, 2013; Tsonis et al., 2015). Other studies have been more equivocal and/or called for more data to be collected (e.g., Kirkby et al., 2011; Agee et al., 2012; Magee and Kavic, 2012; Voiculescu and Usoskin, 2012).

Still, if the proposed link between cosmic rays and cloud cover transpires to be insignificant, this does not preclude the possibility of other mechanisms by which cosmic rays could significantly influence the climate (Kirkby, 2007). For instance, Tinsley et al. have suggested that cosmic rays could influence the climate through interactions with the global atmospheric electric circuit (e.g., Tinsley, 2012; Voiculescu et al., 2013 and references therein). Specifically, Tinsley (2012) suggests that incoming cosmic rays can electrically induce changes in the ionosphere-earth current density (J_z), which in turn can modulate cloud microphysics and storm vorticity (see e.g., Hebert et al., 2012). This could, for instance, affect winter atmospheric circulation patterns such as the North Atlantic Oscillation and the Arctic Oscillation (e.g., Zhou et al., 2014).

If some substantial link between cosmic rays and the Earth's climate exists, then this *also* offers an additional mechanism for a very long-term astronomical influence on climate change, aside from solar variability. If the interstellar concentration of cosmic rays changes significantly as the Solar System moves through the galaxies, then this could alter the relative magnitude of this solar-climate effect over timescales of hundreds of millions of years, and could partly explain the Earth's apparent alternation between ice ages and non-ice ages which is suggested by geological records (Svensmark, 2012; Shaviv et al., 2014).

Other proposed amplification mechanisms are based on the differences in the solar irradiance trends at different wavelengths. Although satellite measurements suggest that the changes in Total Solar Irradiance are relatively small over a single solar cycle (~0.1%), the corresponding changes for some of the ultraviolet (U.V.) frequencies can be of the order of 5–10% (Lean et al., 1995; Lean, 2000).

Most of the incoming Total Solar Irradiance in the ultraviolet frequencies is absorbed by the ozone layer in the stratosphere. As a result, it is often assumed that this higher solar variability would not *directly* affect tropospheric temperatures. However, several researchers have proposed several mechanisms by which a possible link between ultraviolet solar variability and the stratosphere could *indirectly* influence the climate in the troposphere (e.g., Haigh et al., 2010; Lockwood et al., 2010; Lean and DeLand, 2012).

If this is the case then significant changes in the ultraviolet component of the Total Solar Irradiance could offer a possible amplification mechanism for a solar-climate link. However, a major problem for studying these possible links is that our understanding of the variability in the ultraviolet component is still mostly limited to the satellite era measurements. Some researchers have tried to develop models for estimating the variability of the solar spectrum (e.g., Lean et al., 1995; Lean, 2000; Krivova et al., 2007; Fontenla et al., 2011; Vieira et al., 2011), but probably we will need a longer period of direct monitoring of ultraviolet variability before we can draw robust conclusions about how the long term trends in the ultraviolet frequencies compare to those for Total Solar Irradiance.

As we will discuss in Section 5.3, two of us (MC & RC) have argued elsewhere that the stratosphere and troposphere are much more closely linked than has been previously assumed (Connolly and Connolly, 2014f,g,h). If this is the case, then a solar influence on stratospheric temperatures would also influence tropospheric temperatures. Hence, the effects of a possible solar-stratosphere link on tropospheric temperatures would be less indirect than was assumed by Haigh et al. (Haigh et al., 2010; Lockwood et al., 2010; Lean and DeLand, 2012). However, this would not change the fact that if there are significant trends in

the ultraviolet component of the Total Solar Irradiance, this might provide an additional mechanism for a solar-climate link which might not be apparent from the non-partitioned Total Solar Irradiance trends.

Reid (1991) found an apparently strong correlation between the envelope of sunspot activity (see Fig. 4(g)) and sea surface temperature trends but the correlation was not as strong with land temperature trends. As we will discuss in Section 4.2, the latest estimates of sea surface temperature trends are more compatible with land temperature trends (at least for the Northern Hemisphere). However, this early study highlighted the possibility that solar activity might have a more pronounced influence on sea surface temperatures than those on land, due to the large heat capacity of the oceans (Reid, 1991; Gray et al., 2010; Soon and Legates, 2013).

One of us (WS) has recently suggested that, if changes in solar activity strongly influenced sea surface temperatures, this could alter the oceanic circulation patterns, which could in turn influence land temperature trends (Soon and Legates, 2013).

2.4. Summary of the current debates

Main points of general agreement:

- Solar input is probably one of the most important drivers of Earth's climate. So, any substantial changes in solar input that might have occurred in the past should in theory have resulted in significant climate change.
- During the satellite era, there are considerable uncertainties in accurately estimating solar activity, especially during the so-called "ACRIM gap".
- The average Total Solar Irradiance at 1 Astronomical Unit (AU) is currently somewhere in the range 1360–1380 W/m^2.
- All the different solar proxies are just "proxies", and so, before the satellite era, they can only be used to suggest approximate TSI trends.
- The Total Solar Irradiance was probably lower during the Maunder Minimum than it is today by some value in the range 0.5–8.9 W/m^2 — see Table 2.
- There was a general increase in Total Solar Irradiance during the early-to-mid-20th century, and Total Solar Irradiance was probably on average lower in the 19th century than it is today.
- There has been a slight reduction in Total Solar Irradiance since the end of the 20th century, although there is debate over whether this reduction began earlier or not.

Main points of debate:

- During the satellite era, what trends in solar activity (if any) have occurred between solar cycles? There is considerable disagreement on this with one group claiming there has been a decrease in solar activity (PMOD), one claiming solar activity has remained fairly constant (RMIB) and the other claiming solar activity increased during the 1980s and 1990s (ACRIM).
- What is the exact value of the current Total Solar Irradiance (at 1 AU)?
- In terms of solar proxies for the pre-satellite era, which aspects of solar variability are most relevant for reconstructing past Total Solar Irradiance?
- In particular, how much of the trends in Total Solar Irradiance are/are not captured by the widely-used sunspot number records (and similar proxies)?
- Did the mid-20th century solar peak occur in the 1930s, 1940s, 1950s or 1960s? Also how much did solar activity decrease from this peak until the satellite era?
- What is the magnitude of the difference in Total Solar Irradiance between Maunder Minimum and now?
- Is studying Total Solar Irradiance trends sufficient for studying possible solar-climate links, or are there "solar-climate amplification mechanisms" which also need to be accounted for?

3. Surface air temperature data: compilation of regional trends

As two of us (RC & MC) discussed in Connolly and Connolly (2014c), the main dataset used by most of the current weather station-based estimates of global temperature trends is the Global Historical Climatology Network (GHCN) monthly dataset. Until recently, this dataset was maintained by the NOAA National Climatic Data Center, and at the time of writing could still be accessed from: https://www.ncdc.noaa.gov/ghcnm/v3.php. However, in April 2015, the National Climatic Data Center was merged with the new NOAA National Centers for Environmental Information (NCEI), and may need to be accessed via the relevant link on the new website instead, in the future.

We have used the GHCN dataset for our analysis below. However, as discussed in Connolly and Connolly (2014c), there are several serious limitations with this dataset, which impose the following restrictions on our analysis:

- Aside from the contiguous U.S. component of the dataset (called the USHCN dataset), there is no accompanying station history information. Instead, all that is provided for each station is a monthly temperature record, and some brief metadata to describe the station's current location and properties, i.e., station co-ordinates, whether or not it is an airport station, etc.
- There is a dramatic decrease in station coverage after 1990. In some cases, this is due to station closures, e.g., some former Soviet Union stations. However, for many of the stations, this is because the monthly temperature records were *not* collected directly from the station, but rather were taken from some previous data compilation, and some of these compilations have not been updated since the 1990s.
- Aside from the USHCN, there are very few fully rural stations with relatively long & complete station records. E.g., Connolly and Connolly (2014c) showed that only 8 of the (non-USHCN) fully rural stations had data for at least 95 of the last 100 years.

It is well known that station moves, changes in instrumentation, changes in the station's environment and other changes can introduce non-climatic biases into a station's record (e.g., Mitchell, 1953). For this reason, many weather station observers will maintain detailed station histories, documenting any notable changes associated with the station. These station histories are often incomplete, and do not necessarily provide any information on gradual non-climatic biases which may occur, such as urbanization bias. However, if available, they could be useful for accurately identifying and accounting for some of the main non-climatic biases in individual station records.

Since NOAA also organizes the U.S. Cooperative Observer Program (COOP) which is used for constructing the U.S. component of the dataset (called the USHCN), NOAA had access to the station histories for this component (Menne et al., 2009). In particular, they were able to use the recorded changes in observer time to calculate the non-climatic "time of observation biases" associated with the records (Karl et al., 1986; Vose et al., 2003).

As we will discuss in Section 3.2, we are able to use these corrections for time of observation bias for our analysis of U.S. temperature trends. However, other than the USHCN component, no station histories are provided with the Global Historical Climatology Network dataset.

In the case of our regional estimate for Ireland (Section 3.3), we have obtained the relevant station histories (and parallel measurements) directly from the Irish meteorological organization. However, we do not have station histories for our other two components: China (Section 3.1) and the Arctic (Section 3.4). According to Ren et al. (2008), some station history information (including the times of station moves) is available for several hundred Chinese stations at least. It is likely that station histories could also be obtained for many of the Arctic stations.

Therefore, a potential problem for the analysis in this paper is our lack of station histories when constructing our Chinese and Arctic regional estimates. We suggest that a high research priority for improving the reliability of the available global temperature trend estimates should be the organized collection, digitization and compilation of detailed station histories for individual station records.

Since the Global Historical Climatology Network dataset has no station history information (aside from the U.S. component), in order to attempt to correct the station records for non-climatic biases, NOAA have applied an automated homogenization algorithm to one version of their dataset. This algorithm was developed by Menne and Williams (2009) and systematically compares each station record with the records from 40 of its nearest neighbours. If a station record shows an abrupt jump (or trend) relative to some of its neighbours, the algorithm may identify this as a "non-climatic bias", and apply a step change adjustment to the record (Menne and Williams, 2009). Believing that these adjustments are accurate and reliable, many researchers have assumed that the homogenized version of the Global Historical Climatology Network dataset is more reliable than the original unadjusted version (e.g., Menne et al., 2009; Lawrimore et al., 2011; Jones et al., 2012).

It is true that the Menne and Williams, 2009 algorithm has performed very well at identifying (and correcting) artificially introduced biases in simulated data (Venema et al., 2012; Williams et al., 2012). However, two of us (RC & MC) have shown that the algorithm performs poorly when a substantial fraction of neighbouring stations are similarly affected by comparable non-climatic biases (Connolly and Connolly, 2014c,d). In such cases, the algorithm can often lead to a blending of the non-climatic biases between stations, rather than actually removing the biases.

In particular, when a substantial fraction of neighbouring stations have all been affected by urbanization bias (as is the case for many urban areas), the Menne & Williams algorithm can lead to "urban blending". That is, the urbanization bias of heavily urbanized stations is only reduced to match the average urbanization bias of the neighbours, and if there are only a few rural neighbours, they may have "urbanization" bias artificially introduced into their records (Connolly and Connolly, 2014c). Similarly, if the majority of stations in an area are affected by similar biases, as seems to be the case for siting biases in the U.S., the algorithm can actually introduce biases into the records of non-biased stations (Connolly and Connolly, 2014d).

For these reasons, in this study, we will treat the homogenized version of the Global Historical Climatology Network with caution, and will predominantly base our analysis on the unadjusted version. However, where appropriate, we will also discuss the homogenization adjustments generated by the Menne & Williams algorithm.

Recently, considerable work has been carried out into developing more comprehensive temperature datasets than the Global Historical Climatology Network, e.g., the Berkeley Earth project (Rohde et al., 2013) and the International Surface Temperature Initiative (ISTI) project (Rennie et al., 2014). Although, these new projects do not currently seem to involve the collection of station histories either, they do seem to have significantly increased the available data from station records. So, this could be a useful dataset for future research. However, for this review we will focus on the Global Historical Climatology Network (version 3), since this is the dataset which has been predominantly used (directly or indirectly) since the late 1990s for most temperature trend studies — see Connolly and Connolly (2014c).

For calculating the annual mean gridded temperature trends for each of the regions, we adopt the same approach used in Connolly and Connolly (2014c), i.e.,

1. All stations meeting the required characteristics for a particular subset are identified.
2. Each station's monthly temperature record is converted into an annual temperature record, by calculating the mean temperature of all 12 months for a calendar year. If data for one or more months is missing for a year, we do not calculate the mean for that year.
3. The annual records for each station are then converted into "temperature anomaly records", by subtracting the mean annual temperature over the 1961–1990 period for that station from each annual value. This means that if the temperature anomaly for a given year is negative, then it was colder than the 1961–1990 average for that year, and if it is positive, then it was warmer.
4. Stations are assigned into 5° latitude × 5° longitude boxes. The annual mean temperature anomalies for a given grid box are then calculated as the mean of all of the available anomaly records for each year, in that grid.
5. Gridded regional temperature anomalies are then calculated for the subset by averaging together the anomalies of all grid boxes with data for a given year, weighting by the surface area of the grid boxes, which is approximated as being proportional to the cosine of the mid-latitude of the box.
6. To calculate our smoothed values, we use an 11 pt binomial weighting, i.e., we apply the following weights to the 11 year period starting 5 years before a given year and ending 5 years after that given year: [1, 10, 45, 120, 210, 252, 210, 120, 45, 10, 1]. These weights are the 11th power binomial coefficients, i.e., the coefficients of $(1 + x)^{11}$. By using this weighting (as opposed to e.g., an 11 point boxcar running mean), the contributions of values far from the given year are quite small, but the annual variability is still substantially dampened, making the plots seem less "*noisy*".

Although the National Climatic Data Center did *not* include any station *histories* for the stations in the Global Historical Climatology Network dataset, they did provide some basic station *metadata*, i.e., data describing the station and its environment. This metadata includes two different estimates of how urbanized each of the stations is. For each station, they included rough estimates of the population associated with any neighbouring towns or cities in the vicinity, provided the estimated population was greater than 10,000. On the basis of this estimate they defined each station as being "rural" (population < 10,000), "small town" (10,000 ≤ population < 50,000) or "urban" (population ≥ 50,000). They also considered an additional measure based on the brightness of the night-lights in the area, as determined from satellite measurements.[4] As for the population estimates, there are three possibilities: rural, small town or urban.

As discussed in Connolly and Connolly (2014a), both of these metrics are relatively crude, and a station which is identified as "urban" by one of the metrics might be considered "rural" by the other, or vice versa. However, since both metrics are completely independent, if a station is considered "urban" by both metrics, it is likely to be a highly urbanized station. Similarly, if a station is considered "rural" by both metrics, it is unlikely to be majorly affected by urbanization. Therefore, for this analysis, we combine these two urbanization metrics together to create a more rigorous threshold, as in Connolly and Connolly (2014c). That is, we define stations as being "fully rural" if they are rural according to both estimates, "fully urban" if they are urban according to both estimates and otherwise "intermediate".

3.1. Rural China

The first of our four regions which we will consider is China. Several attempts have already been made to estimate the long-term (100 years+) regional surface temperature trends of China, e.g., see Tang et al. (2010) for a review.

Unfortunately, not many Chinese weather stations were active before the People's Republic was founded in 1949. Instead, most of the available Chinese station records began in 1953 or later. From Table 3,

[4] Areas which have been highly urbanized and have access to electricity tend to be brighter at night because of street lights, building lights, etc.

Table 3
Maximum numbers of Chinese stations in the Global Historical Climatology Network for different periods.

Period	Fully rural		Intermediate		Fully urban		All
1841–1906	0	*0%*	0	*0%*	11	*100%*	**11**
1907–1950	13	*12%*	34	*32%*	60	*56%*	**107**
1951–1990	126	*30%*	183	*44%*	108	*26%*	**417**
1991–2014	17	*23%*	18	*25%*	38	*52%*	**73**
All time	126	*30%*	183	*44%*	108	*26%*	**417**

it can also be seen that there is a sharp drop in the number of available stations in the Global Historical Climatology Network after 1990. Several researchers using alternative datasets do not seem to have had this post-1990 drop off problem (e.g., Li et al., 2004, 2013; Wang et al., 2004; Tang and Ren, 2005; Hua et al., 2008; Ren et al., 2008; Yang et al., 2011; Wang and Ge, 2012). So, this seems to mostly be a limitation of the Global Historical Climatology Network dataset, as opposed to a genuine large-scale reduction in station numbers. However, since we are relying on this dataset for the analysis in this paper, the reliability of our post-1990 regional estimates for China will be limited by this restriction.

Despite the limitations in the available data, most long-term estimates which consider pre-1953 temperatures find evidence of an early 20th century warm period (1920s–40s) which has some similarity to the recent warm period, e.g., see Fig. 1 of Tang et al. (2010). Uncertainty exists over the *relative* warmth of the modern & early warm periods, i.e., was one warmer than the other, and if so, by how much?

In a previous study, one of us (WS) considered the influence of solar activity on regional temperature trends for China (Soon et al., 2011), by using one of these long-term estimates, i.e., the Wang et al. (2001) regional temperature estimate. However, although Soon et al. assumed that the homogenization techniques used by Wang et al. had accounted for urbanization bias (and other non-climatic biases), Tang et al., 2010 noted that none of the current *long-term* estimates have explicitly dealt with the urbanization bias problem.

In the last few decades, China has seen a dramatic increase in urbanization. For this reason, it is likely that urbanization bias is a problem for many station records. The acceleration in urbanization has been so rapid that most of the formerly "rural" stations have now become urbanized (Wang and Ge, 2012; Yang et al., 2013).

This is a particular problem for the Global Historical Climatology Network, as can be seen from Table 3 — although there are 126 fully rural stations with data during the 1951–1990 period, there are only 13 and 17 fully rural stations during the 1907–1950 and 1991–2014 periods respectively, and none of the stations with records before 1907 are still fully rural.

Despite the challenges of the urbanization bias problem for Chinese records, there have been several attempts to quantify the bias in recent years (e.g., Li et al., 2004, 2013; Zhou et al., 2004; Hua et al., 2008; Ren et al., 2008; Ren and Ren, 2011; Yang et al., 2011, 2013; Wang and Ge, 2012). For individual urban stations and areas in China, the temperature records can be very strongly affected by urbanization bias. For instance, Ren et al. (2007) found that between 65–80% of the apparent warming trend over the 1961–2000 period for the Beijing and Wuhan station records was probably due to increasing urban heat islands. However, there is quite a bit of debate over the net effect of these biases on regional trends, e.g., see Table 1 of Yang et al. (2011). These estimates range from very small (0.012 °C/decade) (Li et al., 2004) to quite substantial (0.09 °C/decade) (Wang and Ge, 2012).

From Table 3, we can see that for the 1951–1990 period, there is a relatively large number of fully rural (126), intermediate (183) and fully urban (108) stations for China. Therefore, by comparing the trends for these subsets, we can get a reasonable estimate of the net urbanization bias over this 40 year period. We can see from Fig. 10 that the more urbanized a subset is, the greater the warming trend.

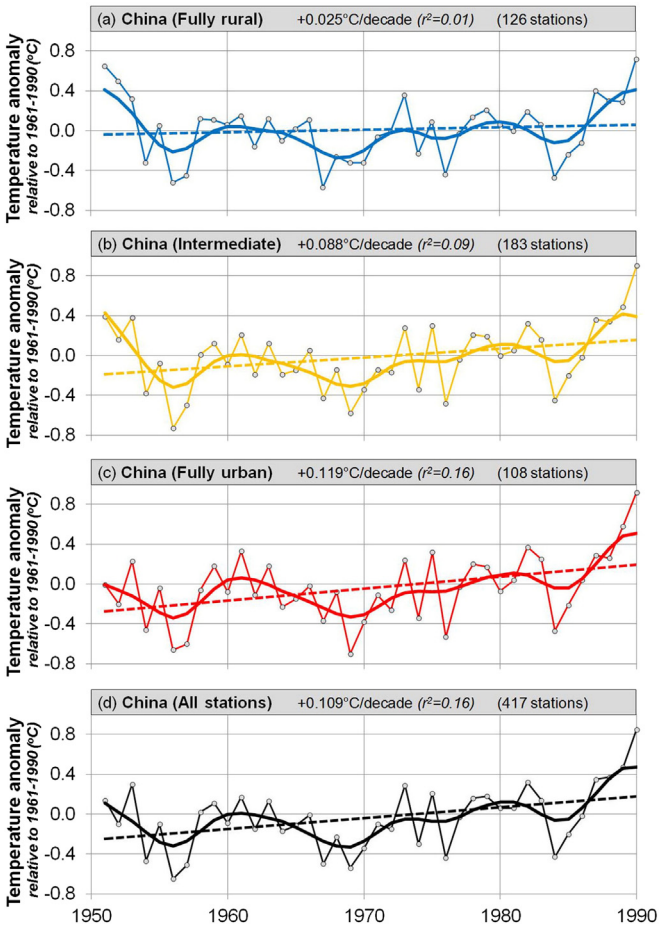

Fig. 10. Urbanization bias study for China during period of maximum station coverage (1951–1990).

For all subsets, the temperature trends over this period are quite non-linear (as can be seen from the relatively low r^2 values for the linear fits). However, if we approximate the trends as linear, the temperature trends increase from +0.025 °C/decade (fully rural) to +0.088 °C/decade (intermediate) to +0.119 °C/decade (fully urban). The trends for all stations are +0.109 °C/decade. If we assume that the fully rural stations are unaffected by urbanization bias, while the other subsets are, then we can estimate the extent of urbanization bias in the "all stations" trends by subtracting the fully rural trends. This gives us an estimate of +0.094 °C/decade urbanization bias over the 1951–1990 period — similar to Wang and Ge (2012)'s +0.09 °C/decade estimate.

This suggests that as much as 86% of the warming trend in the "all stations" estimate for the 1951–1990 period may be urbanization bias. Indeed, if we look at the fully rural subset in Fig. 10(a), the temperatures at the beginning of the period were comparable to those at the end of the period. Admittedly, if we look outside this period (see Supplementary Information), the 1970s–1990 warming continued for all subsets until at least the end of the 1990s, and so stopping our analysis in 1990 could be a bit misleading. However, the 1970s–1990s warming followed a period of cooling which began in the late 1940s. So, starting our analysis in 1951 could *also* be a bit misleading. In other words, we should be wary in extrapolating our analysis in Fig. 10 far beyond the 1951–1990 period. Nonetheless, the substantial difference between each of the subsets over this period confirms that urbanization bias is a serious problem for the Chinese stations in the Global Historical Climatology Network.

Some researchers have argued that the Menne and Williams (2009) homogenization algorithm which NOAA apply to generate the *homogenized* version of the Global Historical Climatology Network should have

removed most of the non-climatic biases in the data, including urbanization bias (e.g., Menne et al., 2009; Lawrimore et al., 2011; Hausfather et al., 2013). For this reason, in Figs. 11 and 12, we compare the homogenized and non-homogenized trends, and the net effects of the homogenization adjustments to each of the subsets.

In light of our analysis above, the first point to note is that there are almost no *net* adjustments to the trends of any of the subsets over the 1951–1990 period. If the adjustments were successfully removing the urbanization bias from the subsets, then we would expect to see the warming trends of the "fully urban", "intermediate" and "all stations" subsets substantially reduced over this period. The fact that this does *not* happen confirms the analysis by two of us (RC & MC) in Connolly and Connolly (2014c), where we concluded that the Menne and Williams (2009) homogenization algorithm is unreliable for removing urbanization bias.

Of course, there are many other types of station changes which could introduce non-climatic biases into the station records, e.g., station moves (Mitchell, 1953) or changes in times of observation (Karl et al., 1986). If we had access to the station histories for the Chinese stations, then we might be able to manually assess the possible non-climatic biases introduced by such changes into individual station records. However, since station histories are not available with the Global Historical Climatology Network dataset, such an analysis is beyond the scope of this study.

On the other hand, Menne and Williams (2009) have argued that their algorithm should be able to identify and remove most of the main non-climatic biases in the station records. So, it is worth considering whether or not their adjustments are still useful for removing the non-urbanization-related biases from the Chinese station records.

While the times of observation since 1953 seem to have been well regulated in China, this does not seem to have been the case for earlier records (Tang et al., 2010). From Figs. 11 and 12, we can see that most of the Menne & Williams adjustments occur in the pre-1953 portions of the records. So, it is plausible that some of the net adjustments are a result of correcting for time of observation bias. However, we note that Tang and Ren (2005) attempted to partially account for time of observation bias by using a combination of maximum and minimum temperature records (Tang and Ren, 2005; Tang et al., 2010), yet their

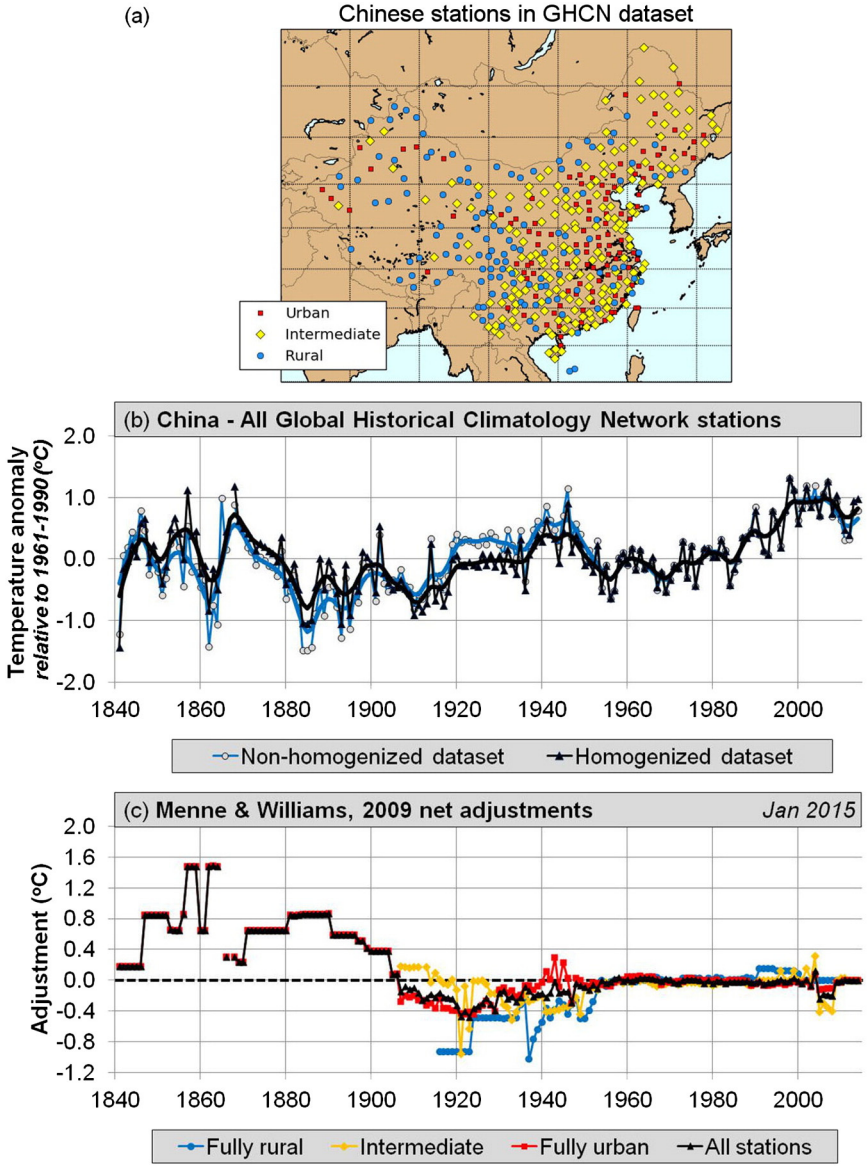

Fig. 11. (a) Location of all 417 Chinese stations in the Global Historical Climatology Network (GHCN) datasets. "Urban" and "Rural" correspond to fully urban and fully rural stations respectively. (b) Gridded mean temperature trends for all Chinese stations in the non-homogenized (light blue) and homogenized (black) GHCN datasets. (c) Net differences between the gridded means of the non-homogenized and homogenized versions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

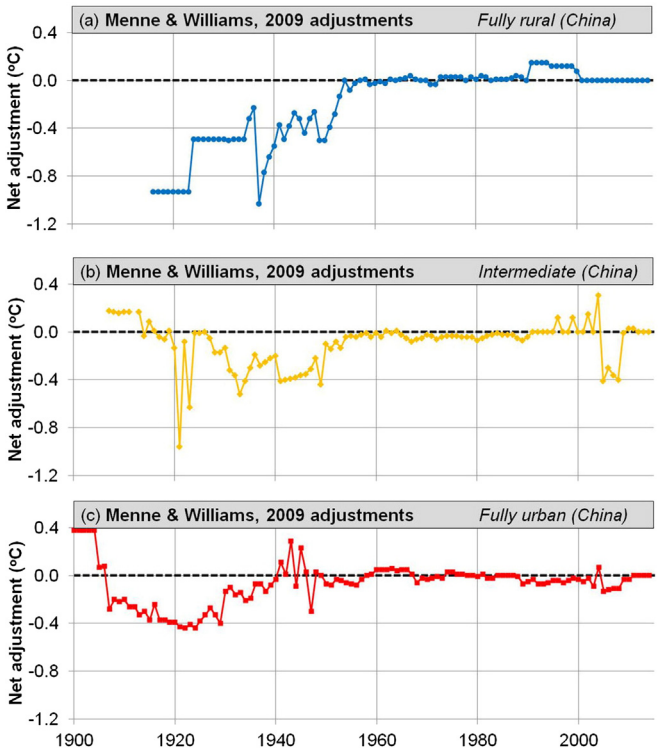

Fig. 12. Net effect of the Menne and Williams, 2009 homogenization adjustments on the three Chinese subsets for the January 2015 version of the Global Historical Climatology Network (GHCN) dataset.

reconstruction was quite similar to the gridded mean average when using the non-homogenized dataset, i.e. compare the light blue curve in Fig. 11(b) to Tang and Ren (2005)'s Fig. 2. This suggests that the *net* biases from changes in time of observation are quite modest for the Chinese network.

Still, this does not rule out the possibility that the relatively large pre-1950s Menne & Williams adjustments may be correctly removing some other non-climatic biases, such as those introduced by station moves.

On the other hand, as discussed earlier, when the number of fully rural stations in an area is low (i.e., pre-1951), this algorithm & other similar ones such as the Easterling & Peterson 1995 algorithm used by Li et al. (2004) will lead to urban blending, meaning that the urbanization bias will tend to be distributed among all stations (urban and rural) to generate a uniform "homogenous" blend. This could explain why Li et al., 2004 failed to identify much urbanization bias in their analysis.

Indeed, from Fig. 12, we can see that the magnitude of the adjustments for the pre-1951 period is actually greatest for the fully rural subset, and the homogenization algorithm reduces the warmth of the mid-20th century warm period, effectively introducing a warming trend into the fully rural subset. This is consistent with urban blending. If the homogenization algorithm has indeed introduced urban blending, then this would have reduced the reliability of the station records, rather than improved it.

For these reasons, to construct our Chinese regional estimate we will use the non-homogenized dataset. However, if detailed station histories for these stations can be obtained, future work could involve using this information to correct our estimate for non-climatic biases other than urbanization bias.

Ideally, in order to minimise the amount of urbanization bias in our regional temperature estimate, we would only use fully rural stations. However, we can see from Table 3 that outside of the peak 1951–1990 period, there are very few fully rural stations in the dataset. Therefore, outside this period, as a compromise to maintain reasonable station coverage, we merely remove the fully urban stations. That is, we include both fully rural and intermediate stations for the post-1990 and 1907–1950 periods.

For the pre-1907 period, the only stations available are fully urban stations. However, if we assume that *most* of the urbanization bias introduced into these station records has occurred in recent decades, then perhaps the urbanization bias in the pre-1907 period for the fully urban stations is relatively small in magnitude. With this in mind, we extend our composite back to 1841 by including all available stations for the 1841–1906, even though all of these stations are now fully urban (Fig. 13).

Since the long term trends for each of these subsets is different (see Supplementary Information), it is important to renormalize each of the segments before combining them together. Otherwise, there would appear to be a sharp temperature jump between segments. To overcome this, when adding the 1907–1950 segment to our reconstruction, we renormalize the "no fully urban" subset so that it has the same mean temperature anomaly as the fully rural subset over the following 15 year period, i.e., 1951–1975. Similarly, when adding the 1841–1906 segment, we renormalize the fully urban subset so that it has the same mean temperature anomaly as our (renormalized) "no fully urban" subset over the 1907–1922 period. However, we do not need to renormalize the 1991–2014 segment, since both the 1951–1990 and 1991–2014 segments are already calculated relative to the same 1961–1990 period.

3.2. Rural U.S.

Recently, two of us (RC & MC) carried out a detailed analysis of the regional temperature trends for the contiguous U.S., i.e., all United States except Hawai'i and Alaska (Connolly and Connolly, 2014c,d) (Fig. 14). As part of this analysis, we identified that a relatively large

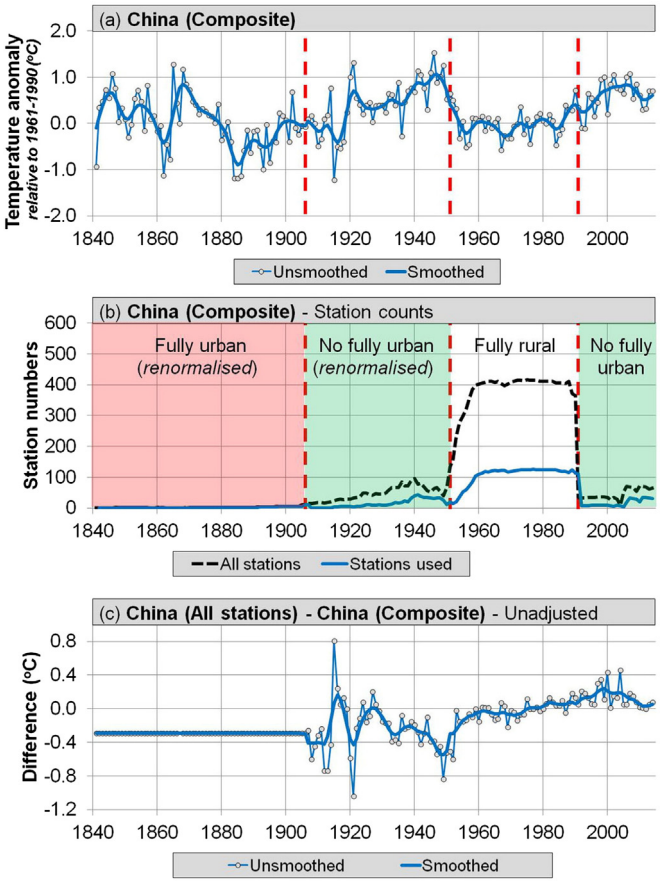

Fig. 13. Our mostly rural Chinese regional temperature composite.

number of fully rural stations were available for the period 1895–present. This was partly, as a result of the U.S. Cooperative Observer Network (COOP) program and partly because Karl et al. (1988) had intentionally selected mostly rural stations from this network when they were constructing the USHCN dataset (i.e., the U.S. component of the Global Historical Climatology Network) to minimise the effects of urbanization bias.

In total, 272 of the 1218 stations (22.3%) in the USHCN dataset were fully rural, and had enough data for our analysis, and the locations of these stations are fairly evenly distributed across the contiguous U.S. — see Fig. 14(e). The unadjusted gridded mean temperature trends of these stations are shown in Fig. 14(a).

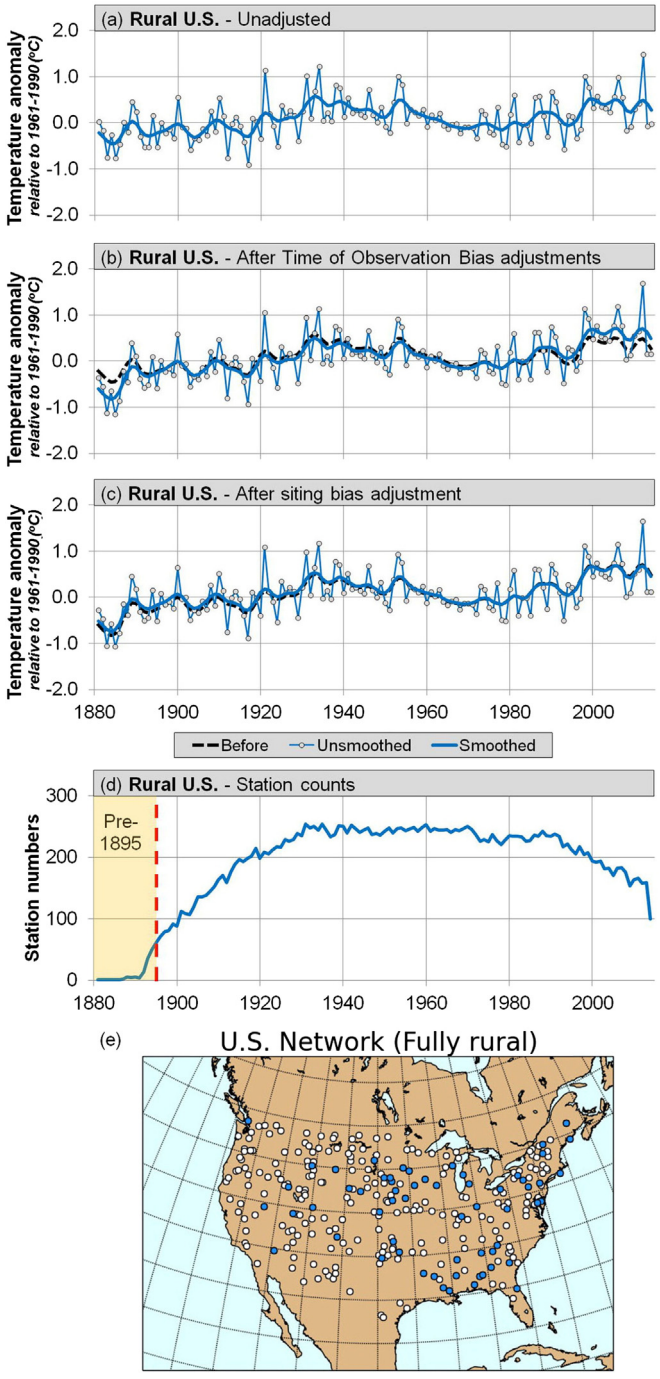

Fig. 14. (a–c) Rural U.S. temperature trends before (black dashed lines) and after various adjustments to account for non-climatic biases.(d) Stations used for each year.(e) Locations of fully rural U.S. Historical Climatology Network (USHCN) stations used. Filled blue circles indicate the locations of those stations with data before 1895. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The COOP program only began providing large amounts of data from 1895 on. So, in the earlier Connolly and Connolly (2014c, 2014d) analysis, we only considered U.S. trends for the 1895–2011 period. However, from Fig. 14(d) and (e), we can see that there are also some fully rural stations with data for the 1881–1894 period. Hence, for this study, we have extended our original analysis to cover the 1881–2014 period, but we would recommend treating the pre-1895 period with some caution.

Karl et al. (1986) noted that there had been systemic changes in the average times during which COOP observers had made their measurements over the course of the 20th century. Karl et al. argued that these changes had introduced "time of observation biases" into many of the COOP station records. For this reason, they developed a series of statistically and empirically derived adjustments which could be applied to the COOP station records to account for these biases, provided that the times of observations were known.

With this in mind, since NOAA have access to the station histories for the USHCN stations, they have applied Karl et al. (1986)'s "Time of Observation Bias" adjustments to each of their USHCN datasets except for the unadjusted version. Hence, for this study, we will use the Time of Observation Bias adjusted dataset for our U.S. regional estimates — Fig. 14(b).

It can be seen from Fig. 14 that these adjustments introduce a slight warming trend into the temperature estimates. Specifically, two of us (RC & MC) have shown elsewhere that these Time of Observation Bias adjustments introduce a warming trend of roughly +0.20 °C/century to U.S. temperature trends relative to the unadjusted data (Connolly and Connolly, 2014d).

Balling and Idso (2002) were sceptical of the reliability of these adjustments and noted that the unadjusted U.S. trends showed a better match to U.S. temperature trends derived from upper air (weather balloon and satellite) measurements than the adjusted data. However, Vose et al. (2003) reconfirmed that the Time-of-observation bias adjustments were necessary due to the systemic change in U.S. observer behaviour over the 20th century.

Connolly and Connolly (2014c) found that, after applying Time-of-Observation bias adjustments, the 10% of U.S. stations that were "fully urban" showed an urbanization bias of approximately +0.50 °C/century relative to the 23% of stations that were "fully rural". If, as a crude approximation we assume the remaining 67% of the U.S. stations (i.e., "intermediate" urbanization) had an urbanization bias of roughly half that, then the average urbanization bias in the Time of Observation bias adjusted dataset is about +0.22 °C/century, i.e., 10% × 0.50 + 67% × 0.25.

In that case, the urbanization bias in the U.S. Historical Climatology Network would be *slightly* greater than the time-of-observation bias — although of opposite sign. In other words, the warming trend introduced by the time-of-observation adjustments should probably be counteracted by a similar cooling trend to account for urbanization bias. For this paper, we do not need to do this since our U.S. estimates are calculated from only the "fully rural" stations. However, we note that this could explain the apparent contradiction between the Balling and Idso (2002) and Vose et al. (2003) analyses for the full U.S. Historical Climatology Network.

Having said that, although urbanization bias is unlikely to be a major problem for our U.S. rural subset, the recent Surface Stations survey has revealed that about 70% of the U.S. Historical Climatology Network stations are currently sited in poorly or badly exposed locations (Menne et al., 2010; Fall et al., 2011; Connolly and Connolly, 2014d). The local environment within a few hundred metres of a thermometer station can lead to unusual "micro-climates", which are unrepresentative of the climate in the surrounding area. As a result, it is quite likely that a reduction in the siting quality of the stations over time has introduced a "siting bias" into the station trends for both rural and urban stations.

The magnitude (and sign) of the average siting bias has been controversial — see Connolly and Connolly (2014d) for a review and summary by two of us (RC & MC). However, Connolly and Connolly (2014d) found that poor station siting seems to have introduced a warming bias of roughly +0.09 °C/century to U.S. temperature trends in the Time-of-Observation adjusted dataset. Unfortunately, because only 8% of the U.S. Historical Climatology Network stations are currently well-sited according to the Surface Stations results, there were not enough well-sited stations to determine the exact biases with much accuracy. Also, the siting biases did not seem to be linear in nature. So, the +0.09 °C/century calculation is only a very crude estimate.

According to Watts (Watts, 2014 — personal communication with RC & MC), the Surface Stations team have recently adopted a new classification system which they believe is able to more accurately distinguish between good and poor station siting, and under this new system, about 20% of the U.S. Historical Climatology Network stations are well-sited. So, when these new ratings have been published it is possible that a more accurate estimate of the siting biases can be determined.

In the meantime, to correct for siting bias, we will apply the above +0.09 °C/century adjustment to our rural U.S. regional temperature estimate Fig. 14(c). However, by comparing Fig. 14(b) and (c), it can be seen that this adjustment does not majorly alter the long term trends. So, our analysis in this paper should not be substantially different if we used a slightly larger adjustment *or* included no adjustment for siting biases.

3.3. Rural Ireland

In Connolly and Connolly (2014c), two of us (RC & MC) showed that there were only eight non-US fully rural stations in the Global Historical Climatology Network dataset with data for at least 95 of the last 100 years. In other words, there is a serious shortage of fully rural stations with long, up-to-date and relatively complete records.

Although five of these stations were in Europe, there were several inconsistencies between the trends for each record. It is quite likely that some/all of these inconsistencies are a consequence of non-climatic biases. Indeed, a preliminary internet search identified a number of station changes which could potentially have introduced non-climatic biases.

Unfortunately, because there were so few fully rural stations with long & complete records in the region, it is difficult to establish the true climatic trends for these stations without a station inspection and/or a detailed station history. However, because two of us (RC & MC) are only a few hours' drive from one of these eight stations (Valentia Observatory, Ireland), we have visited the station and in consultation with Met Éireann (the Irish National Meteorological Service), we have been able to carry out a detailed assessment of the possible non-climatic biases which might have affected the station record.

According to Met Éireann's station history for Valentia Observatory, there have only been four significant changes since the start of the record in 1869:

1. A station move in March 1892 (from Valentia Island to the mainland);
2. A change of ownership in September 1937 (from the British Meteorological Office to Met Éireann);
3. A station move of ~350 m in February 2001;
4. A switch from a manual weather station to an automatic weather station on 1 April 2012,

The 4 dates of these changes are highlighted in Fig. 15 along with the annual mean temperature trends for its entire record (1869–2014). For some reason, the Global Historical Climatology Network was missing monthly values for June 2004 and July 2012 for Valentia Observatory. In order to provide a continuous record for the entire record, we obtained these two missing values directly from Met Éireann's monthly weather summaries for those months (http://www.met.ie/climate/monthly-weather-reports.asp).

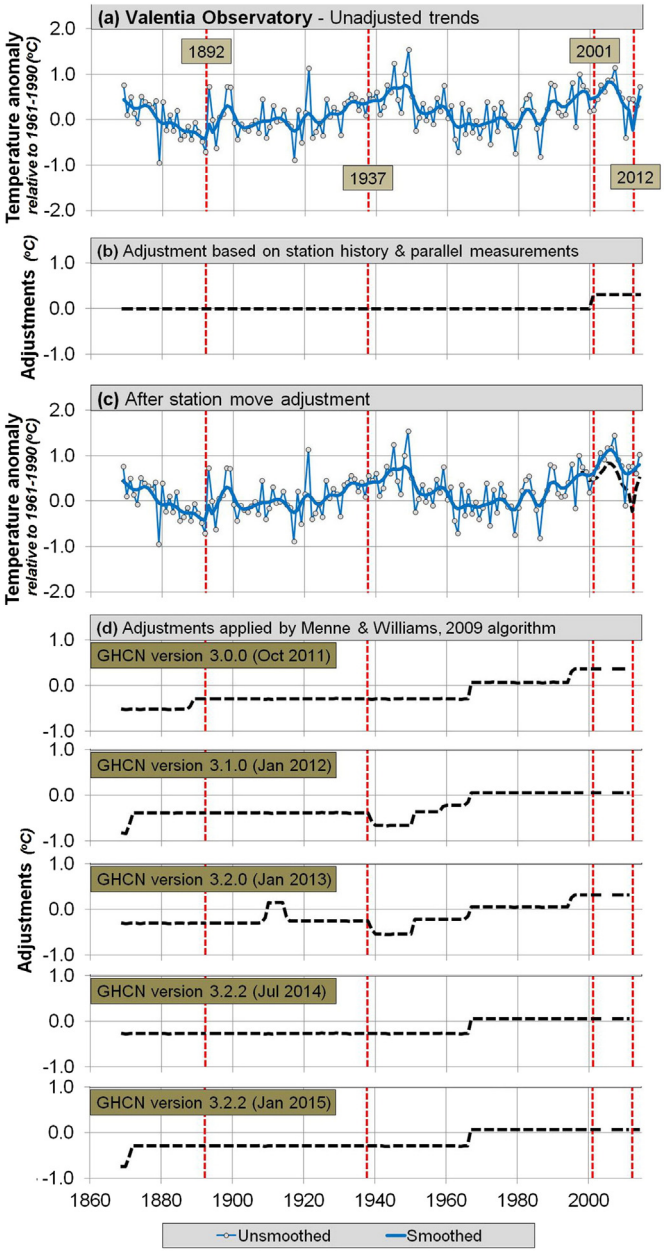

Fig. 15. (a) Valentia Observatory record before adjustment. (b) Record adjustment to account for 2001 station move. (c) Record after adjustment. (d) Adjustments generated by the Menne and Williams, 2009 algorithm. Red dashed lines indicate station history events. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We did not visit the old station location on Valentia Island. However, from our site inspection, the two station locations before and after the 2001 move are both well-sited, with a rating of 2 according to the Leroy (1999) classification[5] used in our Connolly and Connolly (2014d) study. That is, both station locations were more than 30 m from any buildings, concrete surfaces or parking lots, and the slope of the ground was less than the recommended maximum slope of 19°. Therefore, the station record is unlikely to have been affected by siting biases.

There is a small town nearby, Cahirciveen (population 1401 in 2013), which has been expanding over the years. So, it is possible that, in the future, expansion of Cahirciveen might start to introduce some urbanization bias into the station record. But, for now, Valentia

[5] The Leroy, 1999 siting ratings range from 1 (excellent) to 5 (very bad).

Observatory remains outside the town, and the town is still relatively small. So, urbanization bias is unlikely to be a problem for the current analysis.

Now, let us consider what non-climatic biases (if any) the four changes recorded in the station history might have introduced.

It is plausible that the station move in 1892 introduced a step change into the record. However, unfortunately, parallel measurements do not seem to have been made for this station move. Also, none of the other fully rural Irish stations in the Global Historical Climatology Network have data for the 19th century.

As a result, it is unknown what the exact sign and magnitude of any such bias was. The mean temperature for the five year period after the station move (1892–1896) of 10.4 °C was slightly higher than the mean for the previous five years (1887–1891) of 10.2 °C. So, it is possible that the station move could have introduced a slight warming bias. However, it can be seen from Fig. 15 that the year-to-year variability of the record is actually greater than this apparent warming, e.g., the standard deviation over the 1887–1896 period was 0.4 °C. Therefore, it is plausible that much or all of this apparent warming was genuine. Indeed, the mean temperature for the 23 years up to the station move (1869–1891) of 10.5 °C was actually the same as the mean temperature of the following 23 years (1892–1914). The station move could even have introduced a cooling bias, in which case the apparent warming might have been underestimated.

Since it is unclear what bias (if any) was introduced by the 1892 station move, we do not apply any correction for this station change. However, we would recommend treating comparisons between the pre-1892 and post-1892 record cautiously. Perhaps if a weather station was reinstalled near the pre-1892 site, and parallel measurements were carried out, the possible bias introduced by the station move could be estimated.

At any rate, for the more recent station history events, more information is available. The change of ownership in 1937 was a political change corresponding to the founding of the Republic of Ireland, but according to Met Éireann, it did not involve any actual changes in the day-to-day running of the station. Therefore, we do not need to apply any correction for the 1937 ownership change.

The February 2001 station move was from one part of the observatory grounds to another. There is a gentle slope of ~3.3° on the grounds, and as a result the new enclosure was ~20 m higher than the old enclosure (old elevation = 9.1 m; new elevation = ~30 m) making it slightly cooler. Also, the old enclosure was only ~100 m from the ocean, while the new enclosure is ~450 m from the ocean. Either of these factors could have introduced a non-climatic bias.

Before the February 2001 station move, Met Eireann carried out a 24 month comparison between the old and new enclosures, which revealed that the new enclosure was on average 0.3 °C colder than the older enclosure (−0.4 °C for maximum temperatures and −0.2 °C for minimum temperatures) (Met Eireann, 2014 — personal communication with RC & MC). Therefore, we apply a +0.3 °C correction to the Valentia Observatory record to all the months after January 2001 — see Fig. 15(b).

Met Éireann also provided us with two years (2011–12) of parallel temperature measurements for the automatic station and the older system, which allows us to estimate any non-climatic biases introduced by the switch to the automated system in April 2012. According to the new system, the mean annual temperatures for 2011 and 2012 were 10.91 ± 0.04 °C and 11.03 ± 0.04 °C respectively, while for the older system, the corresponding means were 10.89 ± 0.04 °C and 11.02 ± 0.04 °C. Therefore, since the annual mean differences between the two systems were less than the error bars, and less than 0.1 °C, no correction is necessary for the 2012 switch.

The corrections we apply to the Valentia Observatory record for non-climatic biases are summarised by Fig. 15(b). Essentially, the only correction we apply in this case is to increase the post January 2001 temperatures by 0.3 °C. It is interesting to compare these corrections (which are based on the actual station history and parallel measurements) to the automated Menne and Williams (2009) data adjustments used to generate the homogenized version of the Global Historical Climatology Network.

Fig. 15(c) shows the different homogenization adjustments applied to the Valentia Observatory record for five different updates of the Global Historical Climatology Network. There are several important points to note:

- If the Menne and Williams (2009) algorithm was even moderately effective in correctly identifying and correcting for any non-climatic biases in the Valentia Observatory record, then we would expect *some* consistency in the adjustments it applies between updates. However, all five of the updates yielded different estimates of the "non-climatic biases" allegedly in the record, even though the raw data had not changed aside from the latest data being added to the end of the record each month [*For reference*, *in*Connolly and Connolly (2014c), *our analysis of the Valentia Observatory record was just based on the January 2013 update.*].
- None of the adjustments applied in any of the updates corresponded to actual station changes in the recorded station history. Hence, it seems that, in this case, all of the "biases" detected by the automated algorithm were false positives. It could be argued that one of the adjustments for the January 2012 and January 2013 updates occurred near the 1937 change, which might indicate a true positive. However, we saw above that the 1937 change was only an administrative one. Also, none of the other three updates made an adjustment then.
- None of the updates identified the non-climatic bias introduced by the 2001 station move or any bias for the 1892 station move (which we hypothesised might have introduced a bias). So, the algorithm also generated false negatives.

All of the above highlights the dangers of over-relying on automated homogenization algorithms, such as Menne and Williams (2009) to remove non-climatic biases from station records. They may be a useful tool in identifying possible biases which might otherwise be overlooked, particularly when the available station history is limited, or even absent, as is the case for the non-US component of the Global Historical Climatology Network. However, the above example illustrates the inconsistency and high incidence of false positives and false negatives which can occur.

For these reasons, we strongly recommend making the collection of detailed station histories and other relevant information for individual stations a high research priority. In our case, by obtaining a detailed station history and useful parallel measurements directly from Met Éireann, we were able to *manually* construct a more plausible, consistent and empirically justified set of adjustments for the Valentia Observatory record.

Let us now compare our adjusted record to the other available station records for the region. In total, in the Global Historical Climatology Network, there are 15 stations for the island of Ireland (one station, Belfast, is in Northern Ireland and the rest are in the Republic of Ireland). However, only five of these stations meet the fully rural requirements.

In Fig. 16, we compare our corrected Valentia Observatory record to the trends of the other four fully rural Irish stations, i.e., Belmullet, Malin Head, Birr and Clones. Although none of these records have data before 1950, it can be seen that all five fully rural stations have fairly similar trends during the period of overlap (1950–2014).

For these reasons, we believe that our corrected Valentia Observatory record is fairly representative of Irish rural trends since 1869. With its location at the western tip of an island on the edge of western Europe, right beside the Atlantic Ocean, it is probably also the land station most representative of long term North Atlantic temperature trends (at least on the eastern side of the Atlantic). If similar station histories and/or parallel measurements were located for some of the other long and complete fully rural station records discussed in Connolly and

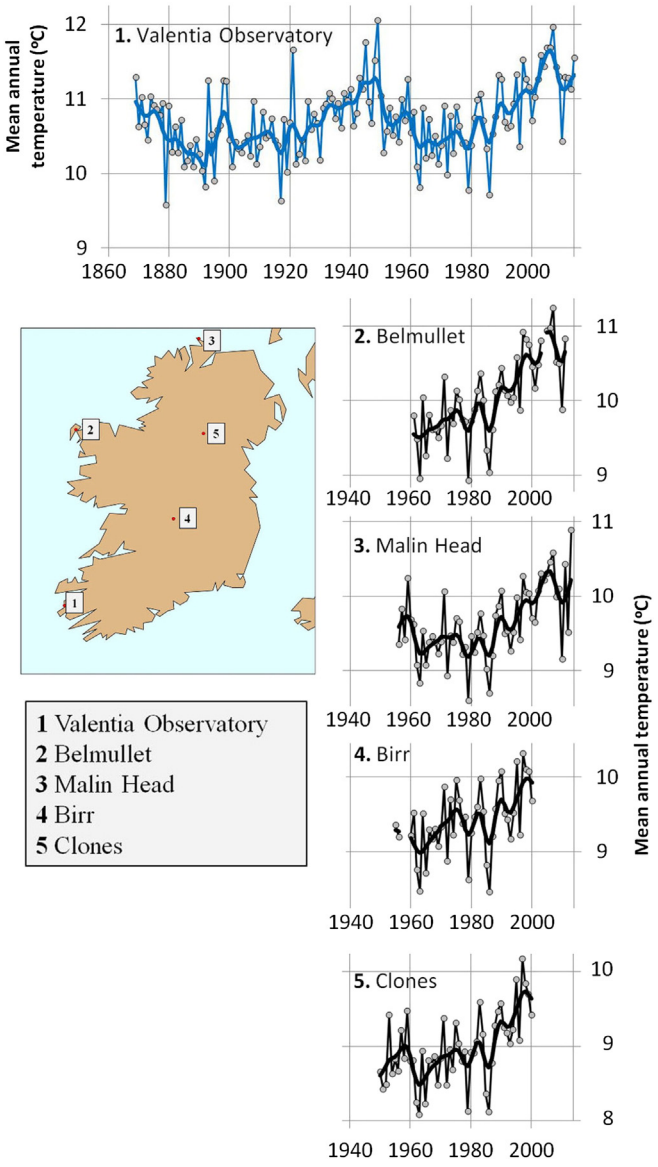

Fig. 16. Annual mean temperature trends of all five "fully rural" Global Historical Climatology Network (GHCN) stations for Ireland, after applying our 2001 station move correction to the Valentia Observatory record.

Connolly (2014c), then our analysis could also be extended to include mainland Europe.

We note that the regional temperature trends Le Mouël et al. (2008, 2009) calculated for mainland Europe are quite different from the above rural Ireland trends. Specifically, Le Mouël et al. found that European temperatures remained fairly constant for most of the 20th century, followed by a sharp stepwise increase in temperatures around 1986. However, when Le Mouël et al. were selecting stations for their estimate, they did not distinguish between urban and rural stations (their main priority was that the stations would have relatively long and complete records). As a result, many of the stations they used were urban stations and these records may be affected by urbanization bias. Also, as Legras et al. (2010) pointed out, because Le Mouël et al. had (intentionally) used raw unadjusted station records, it is possible that non-climatic biases in the raw records could have introduced significant biases into their regional estimates. On the other hand, as we discussed above, the automated pairwise homogenization algorithms favoured by Legras et al. for removing non-climatic biases can also be problematic. Therefore, while it is plausible that the regional temperature trends for mainland Europe were different than those for rural Ireland, it is unclear how much (if any) of the trends reported by Le Mouël et al. are due to non-climatic biases. For this reason, in this paper, we limit our European analysis to rural Ireland, for which we have detailed information. If, in the future, this analysis is to be extended to mainland Europe, we recommend taking an approach similar to our rural Ireland analysis, i.e., focusing on rural trends and collecting and using station histories (and parallel measurements where available) to try and account for non-climatic biases.

3.4. Arctic Circle

For the purposes of this study, we define "the Arctic" as the region of the Northern Hemisphere that is north of the Arctic Circle, i.e., 66° 33′N. As two of us (RC & MC) discussed elsewhere (Connolly and Connolly, 2014c), there are 77 Global Historical Climatology Network stations in this region. The gridded mean temperature trends of these stations are shown in Fig. 17a.

We can see evidence of a regional warming from the early 1890s to 1930s followed by a regional cooling until the 1970s, followed by a second regional warming in recent decades. This is broadly similar to other estimates of Arctic temperature trends, e.g., Kuzmina et al. (2008), or the estimate used for the Soon (2005) study, i.e., Polyakov et al. (2003).

From Fig. 17a and b, we can see that the main net effect of the Menne and Williams (2009) homogenization on the homogenized version of the Global Historical Climatology Network is to slightly reduce the warmth of pre-1958 and post 2005 temperatures relative to the 1958–2005 period. As discussed above, we have concerns over this automated homogenization algorithm, and so for this study, we have preferred to use the estimate based on the unhomogenized data. However, since the net adjustments to the regional trends are relatively small (e.g., compare the dark blue homogenized and light blue unhomogenized curves in Fig. 17a), our analysis would have been fairly similar even if we had used the homogenized version. In either case, as for the U.S., substantial non-climatic trends may still exist in *both* versions. So, it is worth briefly discussing the possibility of non-climatic biases in these estimates.

As a region, the Arctic is still relatively unaffected by urbanization, and so one might naïvely assume urbanization bias to be a negligible problem. However, while the Arctic is a region that is mostly unoccupied by humans, weather stations tend to be located in those areas which are near, or in, human settlements. Klein Goldewijk et al. (2010) has calculated that, even though the total urban area in 2000 for Greenland is only 24 km^2, the percentage of people living in urban areas in Greenland has increased from 20% in 1900 to 82% in 2000. The world average urban fraction for 2000 was only 47%. As a result, urbanization can still be a problem for Arctic stations.

Moreover, modern settlements in harsh climatic permafrost regions often make the most of technological advances when they become available, e.g., snow ploughs, building insulation and heating. In particular, during the polar winter, anthropogenic heat can often be the main source of thermal energy in or near human settlements (Hinkel and Nelson, 2007; Konstantinov et al., 2014). So, even towns with a modest population could have quite a large urban heat island in such regions.

For instance, even though Barrow, Alaska (USA) has a relatively small population (4600 in 2000), Hinkel and Nelson, 2007 found it has an average urban heat island of ~2 °C for the four winter months (December-March). Although this high latitude urban heat island seems to be mostly a winter phenomenon, this would still introduce a warming bias of ~0.5–1.0 °C to annual temperatures over the entire Barrow record. Similarly, Magee et al. (1999) have shown that Fairbanks, Alaska (USA) developed an average annual urban heat island of 0.4 °C (1.0 °C during the winter) over the 49 year period from 1949–1997, even though its population was still less than 35,000.

Only one of the 77 Global Historical Climatology Network stations in the Arctic Circle is "fully urban" — Murmansk, Russian Federation (68.97°N, 33.05°E; 2010 population of 307,257). However, it can be

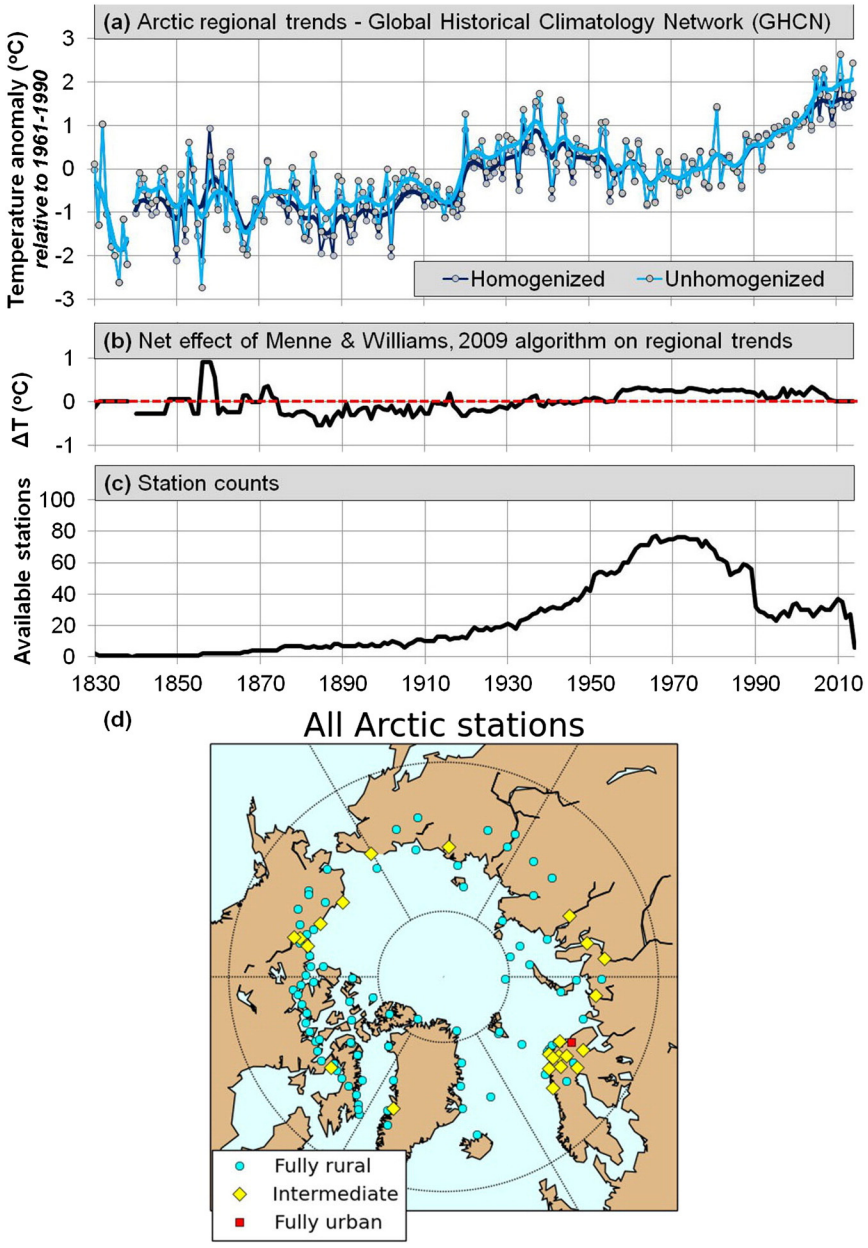

Fig. 17. Gridded mean regional temperature trends for the Arctic circle according to the Global Historical Climatology Network stations. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

seen from Table 1 that a substantial fraction of the stations are "Intermediate", i.e., the same urban rating as Barrow and Fairbanks.

In addition, a lot of the "fully rural" Arctic stations in the Global Historical Climatology Network have fairly short station records, and mostly only cover the period of 1951–1989. As a result, even though 77% of the Arctic stations are fully rural, the *long-term* trends of the Arctic estimates in Fig. 17a are heavily influenced by stations that are currently somewhat urbanized (i.e., "intermediate", or in the case of Murmansk, "fully urban"). In particular, in Connolly and Connolly (2014c), two of us (RC & MC) noted that only one of the six Arctic stations with data for 75 of the last 80 years was "fully rural", i.e., Sodankylä, Finland.

Aside from the Barrow and Fairbanks studies, there has been very ittle research into quantifying the extent of urban heat islands in the Arctic, although Konstantinov et al. (2014) recently presented some preliminary research finding urban heat islands for four Arctic cities (including Murmansk). So, at present, it is difficult to accurately quantify the magnitude of the net urbanization bias effect on the Arctic trends in Fig. 17a.

For this reason, we will *not* attempt to correct our Arctic estimate for urbanization bias. However, it is likely that urbanization bias has led to a slight overestimation of the recent warming period.

In addition, although a survey like the Surface Stations project has not yet being carried out for the Arctic, it is likely that changes in station microclimates could have introduced significant siting biases into the Arctic station records, as well as the other possible non-climatic biases such as changes in time-of-observation and station moves.

In particular, with regards to siting bias, it can be seen from Table 4 that nearly half of the Arctic stations (49%) are currently airport stations. Instanes and Mjureke (2005) have noted that there has been considerable effort in recent decades to better insulate permafrost-based runways such as Svalbard airport. The motivation is to keep the ground under the runway frozen for longer, and extend the length of the season for which the runway is usable. However, on an energetic basis, this belowground "cooling" could have also introduced a relatively warm microclimate above the runway. Since Arctic airport weather stations are often located beside the runway, it is plausible that this slightly

Table 4
Mean number of Arctic Circle stations in the Global Historical Climatology Network for different periods. "Rural" and "urban" categories refer to the "fully rural" and "fully urban" stations.

Period	Rural		Intermed.	Urban	All	Airport	
1830–1874	0	*0%*	4	0	**4**	1	*25%*
1875–1918	6	*46%*	7	0	**13**	5	*38%*
1919–1950	31	*70%*	13	1	**45**	19	*42%*
1951–1989	59	*77%*	17	1	**77**	38	*49%*
1990–2014	24	*65%*	13	1	**38**	20	*53%*
All time	59	*77%*	17	1	**77**	38	*49%*

warmer microclimate could have introduced a warming bias in recent decades to the records of several Arctic airport station records. However, as for the urbanization bias problem, since this has not been quantified yet for the Arctic, we will not apply any adjustments for these biases for this study.

For these reasons, it is plausible that non-climatic biases have altered the apparent relative magnitudes of the warm (1920s–40s; 1990s–2000s) and cool periods (late 19th century; 1960s–70s). Still, as two of us (RC & MC) discussed in Connolly and Connolly (2014c), there is considerable evidence to *qualitatively* confirm the signs of the trends, i.e., warming from early 1900s to 1940s; cooling from 1940s to 1970s; warming from 1970s to 2000s. For this reason, in Section 4, we will use the above Arctic regional estimate as part of our Northern Hemisphere composite, with the caveat that uncertainties remain over the relative magnitudes of the peaks and troughs.

As an aside, we note that Cowtan and Way (2014) have argued that the HadCRUT4 global temperature trend dataset has underestimated the magnitude of recent warming trends due to a lack of coverage for the Arctic in recent years, particularly since 2005. From Table 4, we can see that the number of available Arctic stations before the peak period of 1951–1989 was also quite low. So, if Cowtan and Way (2014) are correct, then presumably this would also mean that HadCRUT4 also underestimated the magnitude of the early 20th century warm period during the 1930s–40s.

4. Northern Hemisphere composite

Fig. 18 compares each of the four rural regional temperature trend estimates described in Section 3. It also includes a Northern Hemisphere composite derived from all four estimates during the period of overlap for all four estimates, i.e., 1881–2014. Details on the stations used and relative weight of each component in the Northern Hemisphere composite are provided in Table 5.

Although each of the regional estimates was derived in different manners and with different numbers of stations, it was assumed that each estimate was representative of the temperature trends for that region. However, it is known that regional temperature trends tend to be greater in magnitude at higher latitudes — a phenomenon known as "polar amplification" (e.g., Soon and Legates, 2013; Cowtan and Way, 2014). Also, the total land area in a given latitudinal band decreases as the cosine of the latitude. Therefore, if all latitudes were weighted equally, this would bias the hemispheric estimates towards the higher latitudes. For this reason, the relative weights for each component were determined by calculating the cosine of the mean latitude associated with each region.

There are some differences between the trends for each of the regions. However, there are several clear similarities between them:

- A warming trend from the late 19th century until the mid-to-late 1940s.
- A cooling trend from the 1950s until the 1970s.
- A warming trend from the 1980s until the 2000s.

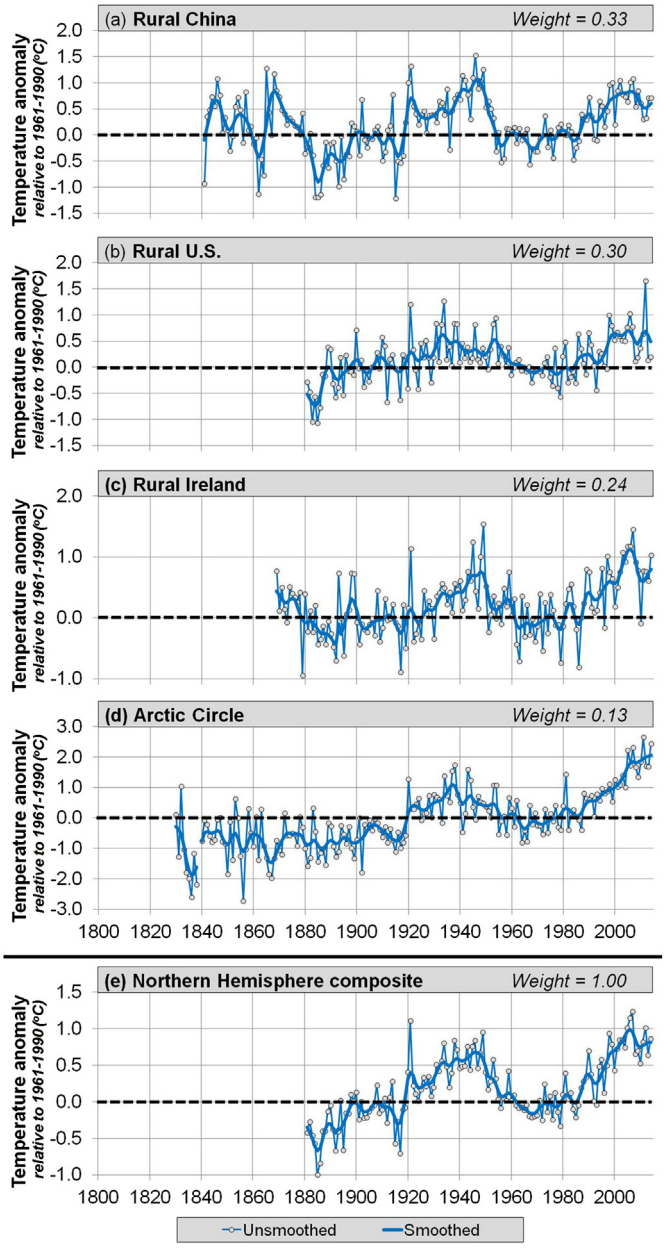

Fig. 18. Construction of our Northern Hemisphere temperature trend estimate (bottom panel) from the four components discussed in text.

Hence, these three common characteristics are also apparent in our Northern Hemisphere composite (Fig. 18e).

Fig. 19 shows the locations of the regions used for our Northern Hemisphere composite. We can see that all four regions are fairly evenly distributed across the Northern Hemisphere. Admittedly, aside from the southern parts of China, our composite does not contain many tropical (i.e., <23°N) stations. This is because there is a shortage of rural tropical stations with long records in the Global Historical Climatology Network dataset. Moreover, there still has not been much research into quantifying the extent of urbanization bias in these regions — see Roth (2007) for a review. However, perhaps in future work, some suitable tropical stations could be identified in one of the new larger temperature record datasets, e.g., the Berkeley Earth (Rohde et al., 2013) or ISTI datasets (Rennie et al., 2014).

Two of us (RC & MC) have noted elsewhere (Connolly and Connolly, 2014c) that aside from Ireland's Valentia Observatory, there are several

Table 5
Station statistics of our regional and Northern Hemisphere (N.H.) composites. For Ireland's "Mean latitude", we have used the latitude of the Valentia Observatory station (Section 3.3), while for the other regions we used the mean latitude of all stations used.

	China	U.S.	Ireland	Arctic	N. H.
Period covered	1841–2014	1881–2014	1869–2014	1830–2014	1881–2014
Stations used	180	272	1	77	530
% fully rural	70.0%	100.0%	100.0%	76.6%	86.8%
Avail. stations	417	1218	10	77	4402
% used	43.2%	22.3%	10.0%	100.0%	12.2%
Latitude range	16.53–53.47°N	24.56–49.00°N	51.80–55.37°N	66.55–82.50°N	16.53–82.50°N
Mean latitude	33.98°N	40.34°N	51.93°N	70.82°N	–
Cos (Latitude)	0.829	0.762	0.617	0.329	–
Rel. weight	0.33	0.30	0.24	0.13	1.00

other fully rural European stations with relatively long and complete records in the Global Historical Climatology Network. So, if detailed station histories for some (or all) of these stations could be obtained, it may be possible to carry out a similar analysis to our Valentia Observatory assessment for them. In that case, it might be possible to expand our rural Ireland estimate to include some of mainland Europe.

4.1. Comparison with other land-based Northern Hemisphere estimates

Fig. 20(a) shows the two different estimates of Northern Hemisphere temperature trends if we apply our gridding method to *all* of the Northern Hemisphere station records (rural and urban) in the non-homogenized (blue) and homogenized (red) versions of the Global Historical Climatology Network datasets. The net effect of the Menne and Williams (2009) algorithm used for generating the homogenized version is to introduce an almost linear +0.020 °C/decade warming trend (linear fit of $R^2 = 0.91$).

Fig. 20(b) compares the homogenized dataset estimate (henceforth, "All GHCN stations") with the Northern Hemisphere (land-only) estimates of:

- NOAA National Climatic Data Center[6] (NOAA NCDC) (Lawrimore et al., 2011)
- Climate Research Unit's CRUTEM4 (Jones et al., 2012)
- NASA Goddard Institute for Space Studies' GISTEMP (Hansen et al., 2010)
- Berkeley Earth (Rohde et al., 2013)
- Lugina et al. (2006)'s update to "the Russian group's" estimate (Vinnikov et al., 1990).

Although there are some slight year-to-year differences between the estimates, all of them show basically identical trends. This has already been noted by most of the above groups (Hansen et al., 2010; Jones et al., 2012; Rohde et al., 2013) and others (e.g., Hartmann et al., 2013; Xu et al., 2014). However, as two of us (RC & MC) have discussed elsewhere (Connolly and Connolly, 2014c), this is not too surprising, since most of these estimates are heavily derived from the Global Historical Climatology Network dataset and/or similar datasets. These latest estimates also apparently imply essentially the same trends (Le Treut et al., 2007; Hawkins and Jones, 2013) as earlier estimates (e.g., Callendar, 1961; Mitchell, 1961). At any rate, it seems reasonable to treat our "All GHCN stations" estimate as representative of all of the previous weather station-based estimates of Northern Hemisphere temperature trends since the 19th century.

Fig. 21 compares this "All GHCN stations" estimate to our new Northern Hemisphere composite derived from mostly rural stations. Since both estimates were constructed relative to the 1961–1990 anomaly baseline period, the convergence between both plots is greatest during that period — see Fig. 21(a). However, by rescaling the estimates relative to the 1881–1910 base period, the long-term divergence between the two estimates is more apparent — see Fig. 21(b).

Both estimates have a lot of similarities, e.g., a warming trend from the late 19th century until the mid-1940s and a second warming trend from the 1970s until the 2000s. However, in our new composite, there was a pronounced cooling trend between the two warming trends (1950s–1970s), while in the "All GHCN stations" estimate, temperatures remained fairly static during this period. Moreover, in our new composite, the relative warmth of the mid-20th century warm period is considerably warmer than in the "All GHCN stations" estimate.

As a result of these nuanced differences, our new composite provides a quite different view of Northern Hemisphere temperature trends since the 1880s than the previous estimates. Both estimates are consistent with the widely-cited observations that there has been "global warming" since the 1970s, e.g., see Hartmann et al. (2013) for an extensive review. However, in our new composite, this recent global warming no longer appears to be particularly unusual. That is, the mid-20th century warm period seems to have been quite comparable to the recent warm period, and temperature trends seem to have alternated between warming and cooling periods since at least the 1880s.

This is consistent with the prediction by two of us (RC & MC) in Connolly and Connolly (2014a) that urbanization bias had led to an underestimation of the relative warmth of the mid-20th century warm period and the magnitude of 1950s–1970s "global cooling" trends.

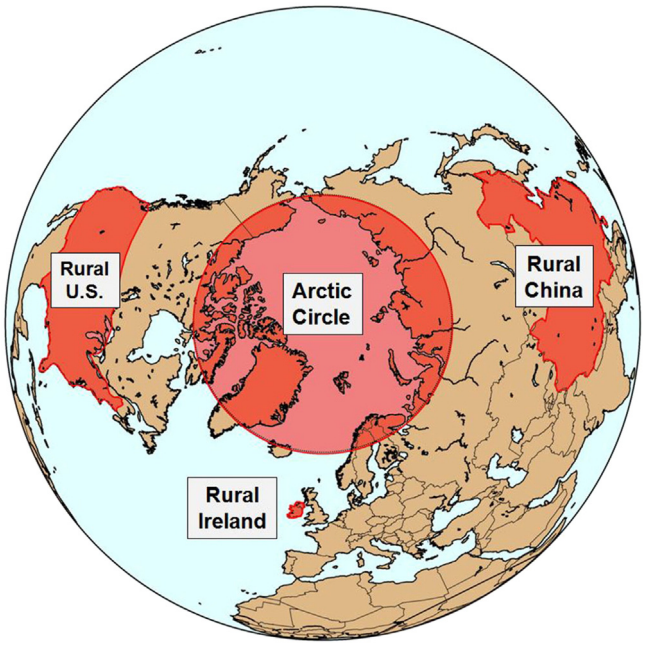

Fig. 19. Regions in the Northern Hemisphere considered for this study.

[6] In April 2015, NOAA NCDC was replaced by NOAA's National Centers for Environmental Information (NCEI).

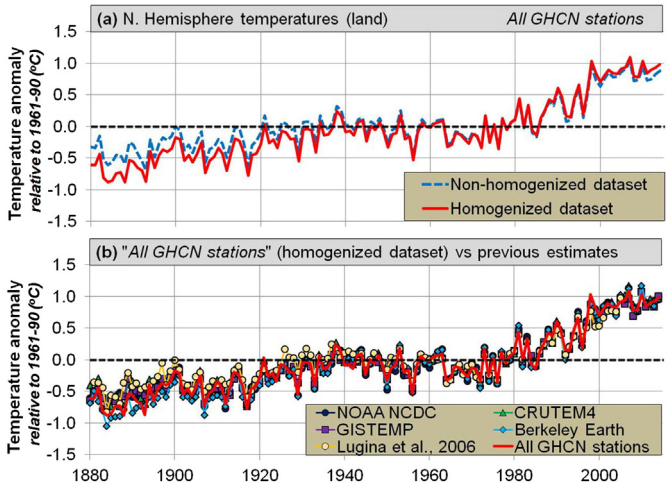

Fig. 20. (a) Comparison of gridded mean temperature trends for all Northern Hemisphere stations in the Global Historical Climatology Network (GHCN) homogenized and non-homogenized datasets. (b) Comparison of the homogenized version to five other recent estimates, which were downloaded from their respective websites on 18th March 2015: CRUTEM4 (http://www.cru.uea.ac.uk/cru/data/temperature/); NOAA NCDC (http://www.ncdc.noaa.gov/monitoring-references/faq/anomalies.php); GISTEMP (http://data.giss.nasa.gov/gistemp/); Lugina et al., 2006 (http://cdiac.ornl.gov/trends/temp/lugina/lugina.html); Berkeley Earth (http://berkeleyearth.lbl.gov/regions/northern-hemisphere). (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

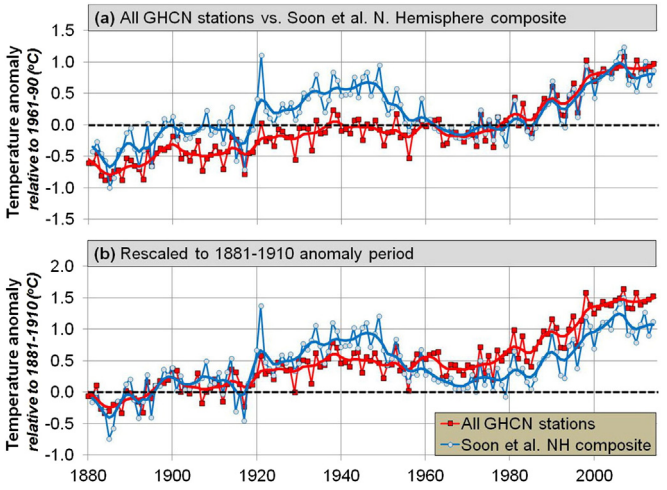

Fig. 21. Comparison of our Northern Hemisphere composite to the equivalent estimate using all Northern Hemisphere stations in the Global Historical Climatology Network (GHCN) homogenized dataset. Temperatures relative to (a) 1961–1990 and (b) 1881–1910.

Because we only used 530 out of the 4402 Northern Hemisphere stations in the Global Historical Climatology Network (i.e., 12.0%), it might be argued that our new composite does not include enough station records, and is therefore less reliable than the "All GHCN stations" composite, and the other estimates in Fig. 20(b). However, this does not necessarily follow.

Although we only considered four regions in the Northern Hemisphere, these regions were selected on the basis that they had a relatively high amount of rural data. Hence, while only 24.7% of the stations in the "All GHCN stations" estimates are fully rural, our new composite is mostly comprised from fully rural stations (86.4% of stations). Therefore, our new composite is far less likely to be affected by urbanization bias than the previous estimates.

As a result, it can be seen from Table 6 that, even though our new composite only uses 12.0% of the available stations, this actually accounts for 42.2% of the available fully rural stations (458 out of 1086). Indeed, many of the fully rural stations in the Global Historical Climatology Network only have data during the 1950–1990 period, and it can be seen from the Supplementary Information that outside this period our composite uses more than half of the available fully rural stations. For the early-20th century, our composite uses nearly 2/3 of the available fully rural stations (see Supplementary Information).

We can also see from Table 6 that several of the previous estimates actually used fewer stations than our new composite (i.e., Callendar, 1961; Mitchell, 1961; Lugina et al., 2006). Yet, these low station number estimates implied essentially the same trends as the other estimates in Fig. 20(b) (Le Treut et al., 2007; Hawkins and Jones, 2013). Indeed, Jones et al. (1997) calculated that about 50 well-distributed stations should be adequate for determining annual global temperature trends (Jones et al., 1997; Hawkins and Jones, 2013). As our composite is just for the Northern Hemisphere, even fewer stations should be needed. Therefore, the differences between our new composite and previous estimates are *not* a result of the smaller number of stations used.

Another potential concern is that our composite only considered four regions, i.e., China, U.S., Ireland and the Arctic. For this reason, it might not be representative of trends for the entire Northern Hemisphere. However, all four of the regions considered are reasonably well distributed throughout the Northern Hemisphere (Fig. 19). Moreover, as seen in Fig. 18, the trends of all four of these widely-spaced regions are broadly similar. Therefore, it appears that our new composite is indeed reasonably representative of Northern Hemisphere trends.

Nonetheless, due to the fact that all of the previous estimates provided basically the same description of temperature trends since the 1880s (Fig. 20), while our new composite provides a different one (Fig. 21), it is worth checking it further against other estimates of Northern Hemisphere temperature trends. With this in mind, in the following subsections, we will compare our new composite to several other independent estimates.

4.2. Comparison with sea surface temperatures

Fig. 22(a) compares our Northern Hemisphere composite to the combined uncertainty ranges of the ERSST and HadSST Northern Hemisphere Sea Surface Temperature (SST) estimates, as in Fig. 5 of Kennedy (2014).

There are several qualitative similarities between the two estimates, i.e., both estimates imply a warming trend during the 1880s–1940s; a cooling trend during the 1950s–1970s; and a warming trend during the 1980s–2000s. However, the magnitudes of all three of these trends are greater for our land-based composite than for the ocean-based composite.

This is not surprising since the oceans have a greater heat capacity than the land and we would therefore expect the ocean temperature trends to be slightly dampened relative to the land temperature trends. Indeed, while the standard deviation of the mean SST estimate is only 0.24 °C over the 1881–2012 period (i.e., the period of overlap between the two datasets), for our land-based estimate it is 0.44 °C. With this in mind, for Fig. 22(b), we have rescaled our land-based estimate by a factor of 0.53, so that both estimates have the same standard deviation over the 1881–2012 period.

Table 6
Numbers of stations used for our composite, and some other Northern Hemisphere estimates.

Estimate	Northern Hemisphere stations
Our composite	530 (458 fully rural, i.e., 86.4%)
"All GHCN stations"	4402 (1086 fully rural, i.e., 24.7%)
Lugina et al. (2006)	384
Callendar (1961)	~350
Mitchell (1961)	119

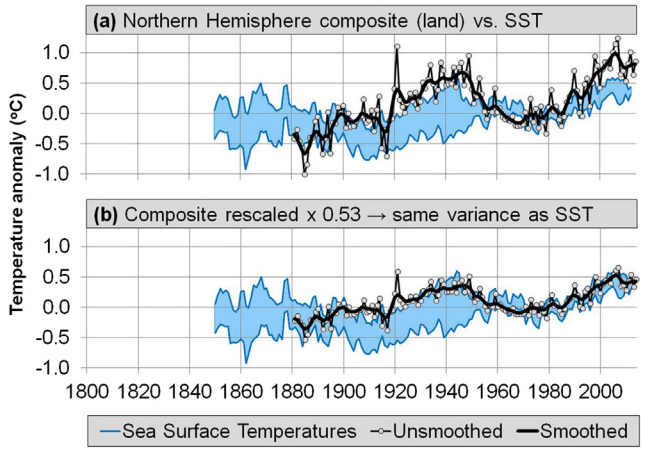

Fig. 22. Comparison of Northern Hemisphere composite (black lines) and the combined uncertainty ranges of the ERSST and HadSST Northern Hemisphere sea surface temperature trend estimates (blue envelope) — digitized from Fig. 5 of Kennedy, 2014. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

When we compare our rescaled composite to the SST estimates (Fig. 22(b)), we find that our early 20th century warming trends are near the upper bound for the SST estimates. However, over most of the records, our rescaled composite is within the uncertainty bands associated with the SST estimates. So, while the previous land-based Northern Hemisphere estimates shown in Fig. 20 have been shown to lie within the uncertainty bands (Kennedy, 2014), after rescaling, our new land-based composite is *also* quite consistent with the SST estimates.

In the Supplementary Information, we provide both the non-rescaled version of our Northern Hemisphere composite, and the version which has been rescaled to match the reduced variability of the ocean temperatures, i.e., Fig. 22(b). Also, since the Northern Hemisphere surface is approximately 40% land and 60% oceans, we provide a third "land + oceans" version that has been rescaled accordingly. That is, since the temperature variability of the ocean component is only 0.53 times the land component, the temperature variability of the combined land plus oceans composite has been rescaled to 72% of the land component ($0.53 \times 60\% + 1.00 \times 40\% = 72\%$). We will use this "land + oceans" version at several points in later sections.

4.3. Comparison with temperature proxies: glacier length-derived estimates

In Fig. 23, we compare our composite to the upper and lower bounds of Leclercq and Oerlemans (2012)'s Northern Hemisphere temperature reconstruction derived from glacier length records. The yellow envelope corresponds to the 95% confidence intervals of their reconstruction.

Leclerq & Oerlemans also provided global and Southern Hemisphere reconstructions, and all three of their reconstructions covered the period 1600–2000. However, although Leclercq and Oerlemans (2012) were able to locate four records (from the European Alps) with data beginning in the 1500s, the vast majority of their records begin in the 19th century or later. Specifically, only 24 of their 247 Northern Hemisphere records (less than 10%) have data for before 1800 (locations highlighted with blue circles in Fig. 23). Even for the early 19th century, they only had an extra 15 records that started before 1850.

For this reason, the earlier portions of the Leclercq and Oerlemans (2012) reconstruction should be treated cautiously. Having said that, 92 of their records (37%) began in 1881 or earlier and the average length of time between data points is much lower in the later parts of the reconstruction. So, the post-1880 portion of their reconstruction which overlaps with our composite is probably relatively robust.

It is therefore worth noting that, even though the mean values of their estimates differ from our composite, the trends of our Northern Hemisphere composite mostly lie within the confidence intervals of the Leclercq & Oerlemans reconstruction. That is, the earlier part of our composite (1880s–1940s) is consistent with the upper bound of their reconstruction while the more recent part of our composite (1980s–2000s) is consistent with the lower bound of their reconstruction.

It should be noted that the two y-axes of Fig. 23(a) cover slightly different ranges, i.e., the modelled temperature variability of their reconstruction is slightly reduced compared to the observed temperature variability of our composite. However, this is to be expected since the response of glaciers to temperature changes tends to be somewhat dampened and lagged (Leclercq and Oerlemans, 2012), even though (i) year-to-year temperature fluctuations are also expected to play an important role in the annual glacier lengths and (ii) glacier behaviour depends on many factors such as precipitation, the underlying topology of the ground, etc. (e.g., Roe and O'Neal, 2009; Roe and Baker, 2014).

Therefore, our Northern Hemisphere composite temperature index seems to be consistent with temperature trends derived from glacier length records.

4.4. Comparison with temperature proxies: tree ring-derived estimates

Several researchers using tree ring chronologies as proxies to extend Northern Hemisphere temperature reconstructions into the pre-instrumental era have noted an apparent "divergence problem" (e.g., Wilson et al., 2007 and references therein). Apparently, although a reasonable correlation can often be found over much of the instrumental record, the tree ring proxies fail to show the strong warming for recent decades which is implied by the weather station-based estimates.

Two of us (RC & MC) have argued in Connolly and Connolly (2014e) that one factor which could explain this would be the presence of non-climatic biases such as urbanization bias in the instrumental record. However, although Wilson et al. (2007) briefly considered the possibility of urbanization bias, on the basis of some of the studies referred to in Section 1, they assumed it would only be a minor problem.

With this in mind, it is worth comparing our new mostly-rural composite to one of these tree ring-derived temperature reconstructions, e.g., the Wilson et al. (2007) proxy reconstruction.[7] We rescaled the Wilson et al. (2007) reconstruction to convert the proxy values into temperature anomalies relative to 1961–90 (°C) by fitting the proxy data to our Northern Hemisphere composite using linear least squares fitting.[8]

In Fig. 24, we compare this rescaled reconstruction to our composite. Although the trends of the tree ring-based reconstruction seem somewhat dampened relative to our weather station-derived composite, they both have some qualitative similarities. Specifically, both estimates imply:

a. A warming trend during the 1920s–1940s period
b. A cooling trend during the 1950s and 1960s
c. A warming trend from the 1970s until at least the end of the 20th century.

The main differences seem to be in the magnitudes of these warming and cooling trends. That is, the temperature changes implied by the tree ring proxies are smaller than those implied by our weather station-derived composite.

We should not necessarily expect an exact linear relationship between a proxy derived from biological growth (i.e., tree rings) and measured instrumental temperature (Loehle, 2009). For instance, if the

[7] Note that the Wilson et al. tree ring-based reconstruction is considered a proxy for Northern Hemisphere *summer* temperature trends, while our composite is for *annual* temperature trends. For this discussion, we will assume they are comparable.

[8] Fitting equation: $y = 0.2381x - 0.0240$ ($r^2 = 0.28$), where y = proxy values and x = temperature anomalies relative to 1961–90 (°C).

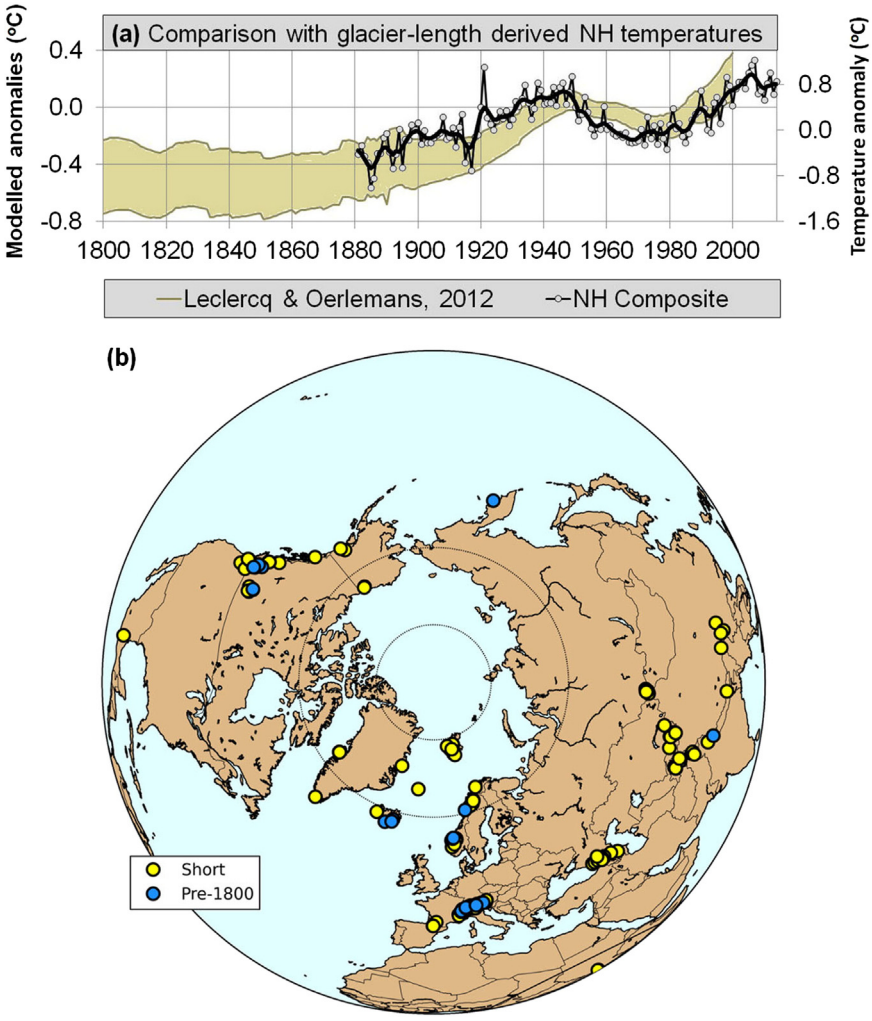

Fig. 23. (a) Comparison with Leclercq and Oerlemans, 2012's glacier length record derived Northern Hemisphere temperature reconstruction. The y-axis on the left corresponds to the modelled temperature anomalies of the reconstruction and the y-axis on the right corresponds to our composite. (b) Location of glaciers used by Leclercq and Oerlemans, 2012 for their Northern Hemisphere glacier length record derived temperature reconstruction. Blue circles indicate the locations of the 24 glaciers (out of 247) with data before 1800. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

growth rate of a particular tree is temperature-limited (e.g., a boreal or alpine tree near the tree-line), then an increase in the summer temperature should lead to an increase in growth rate, and hence wider tree rings for that year. However, if the average temperatures continue to increase, the temperature limiting factor may start to become less important for tree ring growth, and other limiting factors (e.g., available nutrients, water or sunlight) may start to become more important. In that case, we would expect the relationship between tree rings and summer temperatures to become somewhat dampened.

Even if the temperature response is dampened, such a relationship would still be very useful for *qualitative* comparisons of the relative warmth of different periods, e.g., as in a 2003 series of studies by one of us (WS) (i.e., Soon and Baliunas, 2003; Soon et al., 2003a, 2003b), although we note that those studies were quite controversial and generated considerable debate (Mann et al., 2003a,b; Soon et al., 2003b).

At any rate, the qualitative agreement of the trends of the Wilson et al. series with our Northern Hemisphere composite over the common period of overlap (1881–2000) offers further support to the reliability of our composite.

Moreover, although Wilson et al. (2007) claimed to still have a divergence problem with CRUTEM3 for recent decades, this does not seem to be a problem for our mostly-rural composite. This suggests that much (if not all) of the apparent divergence problem was a consequence of urbanization bias in the CRUTEM3 dataset after all.

We do have some concern over the reliability of the Wilson et al. reconstruction, however. As two of us (RC & MC) discussed in Connolly and Connolly (2014e), when carrying out *post hoc* selection of data based on how well individual series compare to a particular parameter, there is a risk that the apparent agreement of the data with that parameter (in this case, instrumental records) will be artificially inflated. This could have led to some "overfitting" (Babyak, 2004; Simmons et al., 2011), whereby the apparently strong relationship between the data and parameter is reduced outside of the calibration (or "training") period.

It is unclear from the text of Wilson et al. (2007) exactly how many series were discarded for not having a high enough correlation to the local gridded instrumental records. It is possible that only a few series were discarded, in which case these effects would be relatively minor. However, if a large number of series were discarded then this could have significant implications for the reliability of the reconstruction in the pre-1850 period.

Fortunately, there is cause for some optimism in this case. Wilson et al. (2007) used the CRUTEM3 gridded temperature dataset for carrying out their data selection, which is an earlier version of the CRUTEM4 dataset discussed in Section 4.1. However, we saw from Figs. 20 and 21 that our new Northern Hemisphere composite implies quite different trends from the Climate Research Unit's Northern Hemisphere temperature estimates. Therefore, since they were discarding series to improve

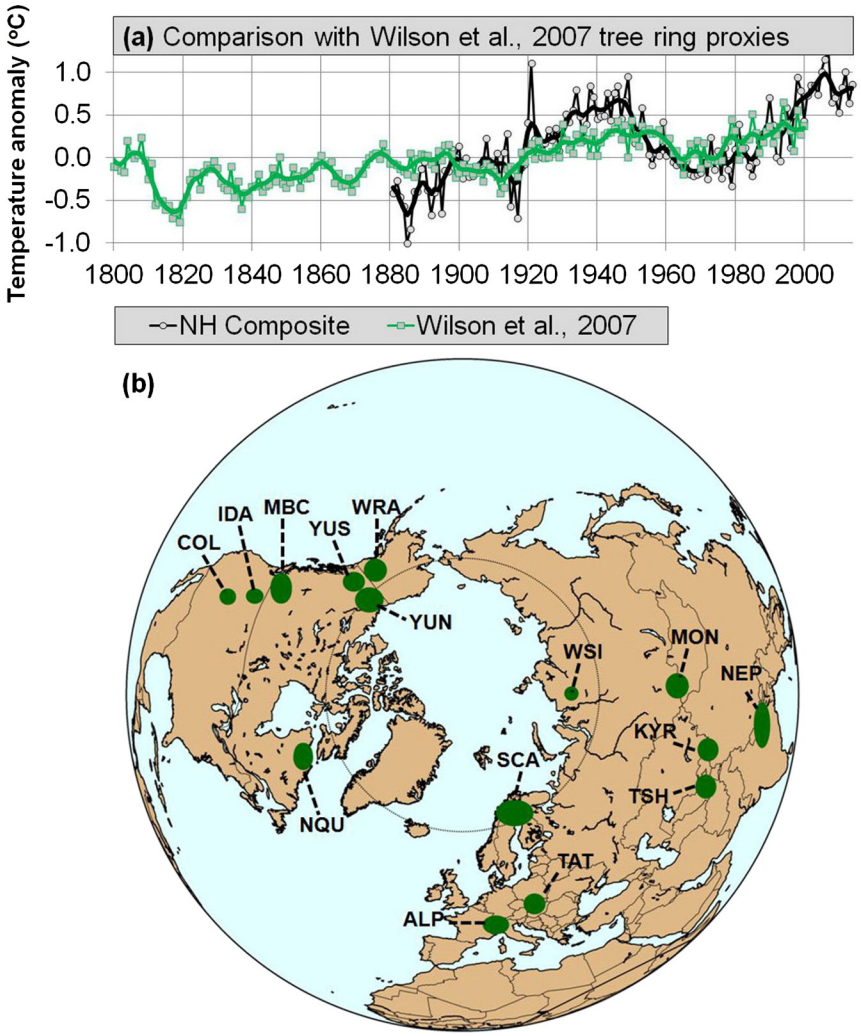

Fig. 24. (a) Comparison with Wilson et al. (2007)'s extra-tropical Northern Hemisphere tree-ring-based temperature proxy reconstruction. (b) Approximate location of the 15 tree ring chronologies used by Wilson et al. (2007) for their extra-tropical temperature proxy reconstruction (adapted from their Fig. 2). For explanations of the labels, see Wilson et al. (2007).

the fit with CRUTEM3, yet their reconstruction actually ended up matching quite well with our composite, the potential overfitting problems introduced by their data discarding were probably quite minor.

4.5. Comparison with Global Climate Model hindcasts (CMIP5)

For the latest Intergovernmental Panel on Climate Change (IPCC) Assessment Reports, a series of Global Climate Model (GCM) simulations were carried out by most of the main climate modelling groups. These simulations were organized and compiled by the Coupled Model Intercomparison Project Phase 5 (CMIP5).[9] One of the sets of simulations carried out for CMIP5 were hindcasts by each of the models of how climate should have behaved since the mid-19th century.

The CMIP5 models account for four main types of climatic "forcings" (two natural and two anthropogenic):

1. *Natural*: volcanic cooling — relatively sudden, yet short-lasting (2–3 year), cooling trends after large volcanic eruptions which transported large quantities of aerosols (mainly sulphates) into the stratosphere, e.g., the 1991 Mount Pinatubo eruption.
2. *Natural*: solar variability — in the CMIP5 models, increases or decreases in the annual Total Solar Irradiance lead to a corresponding response in global surface air temperatures.
3. *Anthropogenic*: greenhouse warming — in the CMIP5 models, increases in the average atmospheric composition of infrared-active "greenhouse gases" due to human activities — chiefly, carbon dioxide (CO_2) — are expected to have introduced an exponentially increasing "global warming" trend since the 19th century.
4. *Anthropogenic*: aerosol cooling — in the CMIP5 models, anthropogenic emissions of sulphates and other aerosols are expected to have introduced a "global cooling" trend due to an enhanced albedo effect.

Bindoff et al. (2013) compared the CMIP5 multi-model mean hindcasts to the global (land & ocean) temperature trend estimates of HadCRUT4, NOAA NCDC and NASA GISS over the period 1860–2012. These trends are comparable to the trends of the various Northern Hemisphere (land-only) estimates shown in Fig. 20 as discussed in Section 4.1.

Bindoff et al. found a strong correlation between the hindcasts and the three global temperature trend estimates over the entire period of overlap once the hindcasts accounted for all four of the above forcings, i.e., "natural + anthropogenic forcings". However, when the hindcasts were carried out without including the anthropogenic forcings (i.e., "natural forcings only") the hindcasts only matched the temperature trend estimates until about 1950. After about 1950, the three global temperature trend estimates implied a pronounced global warming trend, but the hindcasts did not.

[9] http://cmip-pcmdi.llnl.gov/cmip5/.

On this basis, the IPCC WG1's 5th Assessment Report (2013) concluded that:

> "*More than half of the observed increase in global mean surface temperature (GMST) from 1951 to 2010 is very likely due to the observed anthropogenic increase in greenhouse gas (GHG) concentration*" — Bindoff et al. (2013).

However, as can be seen from Fig. 25, when we compare the multi-model mean hindcasts to our new Northern Hemisphere composite, the fits are no longer as compelling. Importantly, *none* of the hindcasts are able to satisfactorily reproduce the early-to-mid 20th century warm period.

It is true that there seems to be an apparently reasonable fit to the "Natural + Anthropogenic forcings" hindcasts (Fig. 25a) from the late 1960s onwards. However, up until then the fit is rather poor. Indeed, arguably, up until the late 1980s, the fits are at least as good (if not better) for the "Natural forcings only" hindcasts as for the "Natural + Anthropogenic forcings". For this reason, if we assume that the CMIP5 models are reliable, then the starting point at which the IPCC attributed surface temperature trends to being mostly "anthropogenic" in nature should probably be changed from 1951 to at least the late 1980s.

Moreover, as discussed in Section 2, the solar variability estimates used by the CMIP5 models were all ones that implied relatively little solar variability since the 19th century, e.g., those on the right hand side of Fig. 8. So, it is possible that the "Natural forcings only" hindcasts might have yielded better fits, if one of the other solar variability estimates had been considered, e.g., the Ammann et al. (2007) update to Bard et al. (2000) or the Scafetta and Willson (2014) update to Hoyt and Schatten (1993) — see left hand side of Fig. 8. In that case, it would weaken the IPCC's conclusion even further.

With this in mind, in Section 5, we will compare our new Northern Hemisphere composite to one of these higher variability estimates of Total Solar Irradiance, i.e., the updated version of the Hoyt & Schatten reconstruction (Hoyt and Schatten, 1993; updated by Scafetta and Willson, 2014).

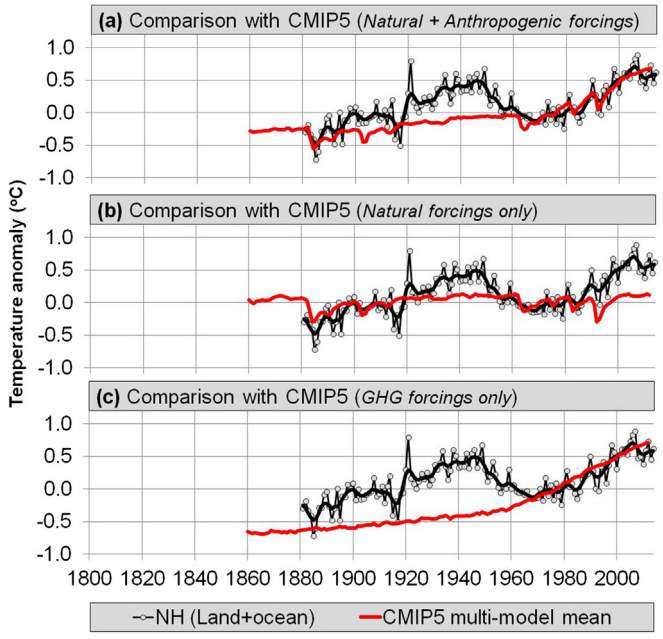

Fig. 25. Comparison with multi model means of CMIP5 hindcasts. CMIP5 means digitized from Fig. 10.1of IPCC WG1 5th Assessment Report (Bindoff et al., 2013) and rescaled to anomalies relative to 1961–1990. NH composite was rescaled by 0.72 to emulate land and ocean trends.

4.6. Comparison with stratospheric volcanic eruptions

If the force of a volcanic eruption is strong enough to reach into the stratosphere, several studies have suggested that this can induce a large-scale cooling effect which may last for a few years (e.g., Sato et al., 1993; Ammann et al., 2003; Arfeuille et al., 2014 and references therein). The theoretical basis for this is that the presence of extra stratospheric aerosols (mainly sulphates) is believed to temporarily increase the planetary albedo, thereby reducing the incoming solar irradiance until such time as the aerosol particles fall out of atmospheric circulation.

Currently, it is believed that a strong stratospheric volcanic eruption like Pinatubo in 1991 will lead to two or three years of global cooling, starting a few months after the eruption (Sato et al., 1993). However, because these particularly strong eruptions ("stratospheric volcanic eruptions") are relatively rare, there have so far not been many examples against which these theories can be tested. As can be seen from Fig. 26, there have only been about 10 major (documented) stratospheric eruptions since the late 19th century, and only two since the start of the satellite era — El Chichón (1982) and Pinatubo (1991).

As a result, several different model estimates of the relative impacts of these stratospheric volcanic eruptions have been made, and there is still debate over which of these models are most accurate (e.g., Sato et al., 1993; Ammann et al., 2003; Arfeuille et al., 2014). Nonetheless, as we discussed in Section 4.5, the climatic forcings implied by these models are one of the main "natural forcings" considered by the current climate models, and a popular test of the reliability of a Global Climate Model is its ability to accurately hindcast the climatic impacts from stratospheric volcanic eruptions. See Driscoll et al. (2012) for an assessment of the modelling of the climatic effects of the stratospheric volcanic eruptions by the CMIP5 hindcasts.

With that in mind, we now test for the presence of any indication of volcanic cooling associated with stratospheric eruptions in our new Northern Hemisphere composite. In Fig. 26, we compare our composite (black line with circles) against the 10 known stratospheric eruptions (vertical lines), and also with the CMIP5 multi-model hindcast (red solid line).

Following several of the eruptions the CMIP5 hindcasts model a temporary, yet pronounced, global cooling which lasts 2–3 years, as can be seen by the short, downward "blips". The magnitude of these cooling periods varies from eruption to eruption, depending on the modelled eruptions in the datasets used by the CMIP5 hindcasts. In particular, six of these volcanic eruptions are believed to have introduced a substantial mass of sulphates and other aerosols into the stratosphere:

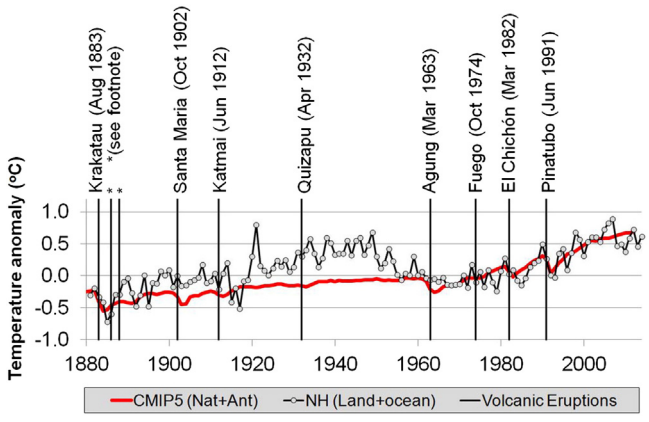

Fig. 26. Timing of the 10 stratospheric volcanic eruptions since the start of our Northern Hemisphere composite in 1881. *Several large volcanic eruptions occurred in the late 1880s, including Tarawera (Jun 1886) and Bandai (Jul 1888). (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

Krakatau (1883), Santa Maria (1902), Katmai (1912), Agung (1963), El Chichón (1982) and Pinatubo (1991).

When we compare our Northern Hemisphere composite to these six eruptions, fairly pronounced cooling periods are apparent after both the Krakatau and Pinatubo eruptions, which were two of the largest eruptions over the period. So, for these two eruptions, our composite provides quite a good match with the CMIP5 hindcasts.

There also seems to have been a few relatively cool years following the El Chichón eruption, but in this case, the cooling seems to have lasted longer and been more pronounced than the hindcasts. So, it is possible that some (or all) of this cooling has nothing to do with El Chichón.

While the hindcasts predicted that there should also have been a noticeable cooling after the Santa Maria and Agung eruptions, these coolings are not reflected in our composite temperature record. Having said that, both of those eruptions are believed to have mostly affected the Southern Hemisphere, e.g. (Arfeuille et al., 2014). So, it is possible that the Northern Hemisphere was relatively unaffected by either eruption.

There does seem to have been a pronounced cool period of several years that occurred after the Katmai (1912) eruption. However, this cool period only began in 1915 and so is probably not related to Katmai.

So, in summary, several cooling events apparent from our Northern Hemisphere composite seem to be broadly consistent with the modelled cooling of the CMIP5 hindcasts. However, given the relatively large inter-annual variability of the average Northern Hemisphere temperatures, plus the fact that there were only a few major stratospheric volcanoes over the period 1881–2014, it is hard to conclusively attribute a particular cool period to volcanic cooling. We conclude, therefore, that the evidence for and against the modelled climatic responses to stratospheric volcanic eruptions by the CMIP5 hindcasts is equivocal.

5. Comparison between Northern Hemisphere temperature and solar activity trends

In Fig. 27, we compare the Hoyt & Schatten reconstruction (Hoyt and Schatten, 1993; updated by Scafetta and Willson, 2014) of Total Solar Irradiance trends (Fig. 8) to each of the four regional temperature trend estimates of Section 3, and our Northern Hemisphere composite, i.e., the plots in Fig. 18.

In all cases, the general agreement between the temperature and solar activity trends is striking. This is especially so when we consider the comparatively poor agreement between our composite and the various CMIP5 hindcasts discussed in Section 4.5 (Fig. 25).

Therefore, if the Hoyt & Schatten reconstruction is a reasonably accurate estimate of solar activity trends (see Section 2.2), and our mostly-rural temperature trend estimates are reasonably accurate (as seems to be the case from Section 4), then this suggests that Northern Hemisphere surface air temperature trends have been heavily influenced by changes in Total Solar Irradiance since at least 1881. This result directly contradicts the conclusions of the latest IPCC reports (Bindoff et al., 2013), in which it is argued that surface air temperature trends have been dominated by changes in atmospheric greenhouse gas concentrations since at least the 1950s.

As discussed in Sections 2.2 and 4.5, the CMIP5 hindcasts were based on either the Wang et al. (2005) Total Solar Irradiance reconstruction or one of the other "low solar variability" reconstructions in Fig. 8. So, they did not evaluate the Hoyt & Schatten reconstruction used in Fig. 27. Moreover, Bindoff et al. (2013) relied on the HadCRUT4, NOAA NCDC and NASA GISS estimates of global temperature trends, which as we discussed in Sections 4.1 and 4.5 imply different trends than our new mostly-rural Northern Hemisphere composite (Fig. 21). It appears that these are the main reasons why Bindoff et al. (2013) reached a different conclusion from us on the relative importance of solar variability on temperature trends.

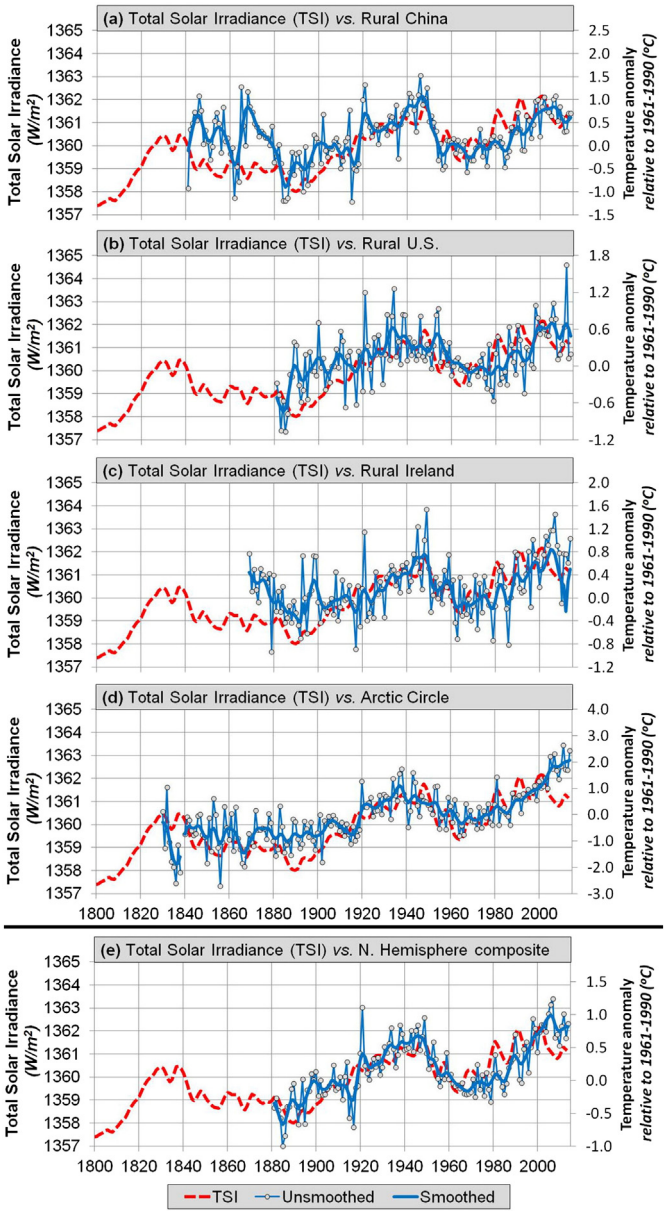

Fig. 27. Comparison between the updated Hoyt & Schatten Total Solar Irradiance trends (red dashed curve) and the various Northern Hemisphere regions/gradients discussed above. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The fact that the Hoyt & Schatten reconstruction seems to imply a relatively large role for solar variability on surface air temperatures has been noted before. In particular, one of us (WS) has already noted a strong correlation between the Hoyt & Schatten reconstruction and regional temperature trends for the Arctic (Soon, 2005, 2009); China (Soon et al., 2011) and the U.S. (Robinson et al., 2007). However, for those earlier Chinese and U.S. studies, no explicit attempt was made to account for urbanization bias or other non-climatic biases, as has been done in this paper.

In this new analysis, all of our temperature trend estimates have been extended up to the latest complete year (i.e., 2014) and the Hoyt & Schatten reconstruction has been updated to 2013 (Scafetta and Willson, 2014; personal communication). We have also combined the regional estimates for China, U.S. and the Arctic with a new regional estimate for rural Ireland to create a new mostly-rural temperature trend estimate for the Northern Hemisphere. So, the analysis in Fig. 27 represents a more comprehensive and up-to-date assessment than the

earlier analyses of Soon et al. (Soon, 2005, 2009; Robinson et al., 2007; Soon et al., 2011).

Soon and Legates (2013) have also identified a strong correlation between the Hoyt & Schatten reconstruction and the so-called "Equator-to-Pole Temperature Gradient" (EPTG) in the Northern Hemisphere. It was found that increases in Total Solar Irradiance led to a reduction in the temperature gradients between the Equator and the Arctic. In other words, the temperature responses to changes in Total Solar Irradiance were greater at higher latitudes, a phenomenon sometimes referred to as "polar amplification".

When Scafetta & Willson updated the Hoyt & Schatten reconstruction using the ACRIM composite (Scafetta and Willson, 2014 — see their Appendix B), they noted an apparently strong correlation with Manley's Central England Temperature (CET) dataset (Manley, 1974), as updated by Parker and Horton (2005). This dataset is not part of the Global Historical Climatology Network, and so was not considered in our analysis. However, since England is also part of the British Isles, we would expect the trends for central England to be similar to those we found for Ireland (Section 3.3).

Having said that, and although Manley did apply slight and somewhat ad hoc adjustments to the post-1960 data for one of his records to account for urbanization bias (Manley, 1974), and Parker & Horton expanded on and increased these adjustments for their update (Parker and Horton, 2005), it may be that urbanization bias is still a problem for the Central England Temperature dataset. For instance, two of us (RC & MC) have noted elsewhere (Connolly and Connolly, 2014a) that the grounds of the Rothamsted station (one of the four stations used for the recent portion of the Central England composite) are surrounded by the town of Harpendon, Hertfordshire (current population ~30,000). Also, one of the stations which was used for the 1974–2004 period (Ringway) was located at a busy international airport (i.e., Manchester Airport).

So, it is possible that the urbanization bias adjustments which have been applied, and which Parker and Horton (2005) conceded were "*still somewhat arbitrary*" are inadequate, and that some residual urbanization bias remains in the Central England Temperature dataset. Indeed, Scafetta & Willson noticed a divergence between the Hoyt & Schatten reconstruction and the Central England Temperatures in the last few decades, with the Central England Temperatures showing more warming (see Fig. 16 of Scafetta and Willson, 2014), while this does not seem to be the case for rural Ireland (Fig. 27c).

When Hoyt and Schatten (1993) originally presented their reconstruction, they compared it to an early version of the GISTEMP Northern Hemisphere temperature trend estimate shown in Fig. 20(b), i.e., the Hansen and Lebedeff (1987) dataset. They found a reasonable correlation over most of the record which covered the period 1880–1990, and also for the 1700–1879 period, which they estimated using the Groveman and Landsberg (1979) reconstruction. However, they noted that temperature trends seemed to be increasing faster than solar activity for the last few decades up until 1990. Hoyt and Schatten (1993) suggested that this could be due to anthropogenic global warming from an increase in atmospheric greenhouse gases, although they also considered the possibility that non-climatic biases such as urbanization bias could be a factor.

Unlike the analysis in Hoyt and Schatten (1993) and the CMIP5 hindcasts discussed in Section 4.5, the changes in Total Solar Irradiance seem to be strongly correlated to our new mostly rural Northern Hemisphere composite over almost the entire record, i.e., 1881–2014 — see Fig. 27e. This suggests that much of the apparent divergence between the Hoyt & Schatten reconstruction and the Hansen and Lebedeff (1987) dataset noted by Hoyt and Schatten (1993) for the late 20th century was due to urbanization bias. It also suggests that much of the observed global warming in recent decades may have been a result of changes in solar activity, contradicting Bindoff et al. (2013)'s claim that the recent global warming is mostly due to changes in anthropogenic greenhouse gas concentrations (chiefly, carbon dioxide).

In light of these new findings, in Section 5.1, we will re-assess the relative roles solar variability (as implied by the Hoyt & Schatten reconstruction) and atmospheric greenhouse gases have played in determining Northern Hemisphere temperature trends since the 19th century. As we discussed in Section 4.5, changes in solar activity and greenhouse gas concentrations are two of the four main climatic forcings considered by the CMIP5 hindcasts. The other two forcings (stratospheric volcanic eruptions and anthropogenic aerosols) are both believed to lead to cooling, and so should not be responsible for any of the *warming* trends in Fig. 27e. Moreover, we have already considered the possible influence of stratospheric volcanic eruptions on Northern Hemisphere temperature trends in Section 4.6.

5.1. Fitting of Northern Hemisphere temperatures to changes in solar activity and atmospheric carbon dioxide (CO_2)

If we assume a linear response between Total Solar Irradiance (TSI) and Northern Hemisphere surface air temperatures then we should be able to derive a theoretical model of Northern Hemisphere surface air temperatures from Hoyt & Schatten's Total Solar Irradiance reconstruction using linear least squares fitting. See Fig. 28.

We note this model is able to describe our new Northern Hemisphere composite fairly well, and that the residuals are relatively small. That is, the model seems to describe most of the temperature variability. For this analysis we are using the rescaled "land + ocean" version described in Section 4.2.

This would imply that a change in Total Solar Irradiance of 1 W/m^2 should (on average) correspond to a change of ~0.211 °C for Northern Hemisphere, i.e., the value of the slope of the line in Fig. 28(a).

If we assume that the average Total Solar Irradiance is ~1361 W/m^2 (Section 2.1), then a 1 W/m^2 change in Total Solar Irradiance would amount to a 0.0735% change in Total Solar Irradiance (i.e., 1/1361). If we approximate the average Northern Hemisphere surface air temperature as 288 K and assume that it would be less than 4 K in the absence

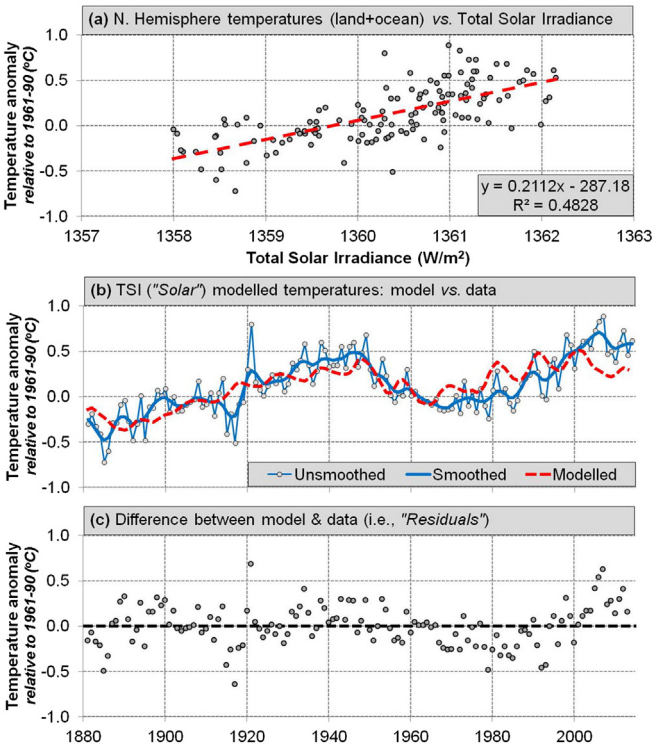

Fig. 28. Derivation of a model for our Northern Hemisphere composite (land & ocean version) based on Total Solar Irradiance trends.

of the Sun, then this implies Northern Hemisphere surface air temperatures are ~284 K warmer than they would be in the absence of the Sun. Therefore, a 0.0735% change in the solar-induced Northern Hemisphere surface air temperature would be ~0.209 K, i.e. 0.209 °C. This is remarkably similar to the ~0.211 °C response implied by our model. That is, it seems that a 0.0735% change in Total Solar Irradiance results in a nearly equivalent change in Northern Hemisphere surface air temperatures.

Although the variations in Total Solar Irradiance are relevant when considering solar variability, when we are considering the effects of solar variability on the Earth's climate, we are usually more interested in the incoming solar irradiance that is actually absorbed by the Earth system (i.e., surface + atmosphere). For this reason, it can be helpful to convert the Total Solar Irradiance values at 1 AU into Radiative Forcing (RF) values, where the solar radiative forcing is the globally averaged solar irradiance absorbed by the Earth system (W/m^2).

The Earth is essentially a sphere and at any given moment, only one side of this sphere is facing the Sun. Therefore, to convert the Total Solar Irradiance into Radiative Forcing, we must first divide our values by four, i.e., the surface area of a sphere divided by the surface area of the irradiated disc. Hence, a Total Solar Irradiance value of 1361 W/m2 would correspond to a Radiative Forcing of 340 W/m^2.

We also saw from Fig. 1 that there is a strong seasonal variation in the Total Solar Irradiance reaching the Earth due to orbital variability. So, if we were considering timescales of less than one year, we would need to consider these seasonal variations. However, since our analysis is based on annual averages, we can neglect this seasonal variability and use the mean annual values at 1 AU.

We must also consider the incoming solar irradiance which is reflected back to space without being absorbed, i.e., the planetary albedo. Currently, it is estimated that about 29% of the solar irradiance reaching the Earth is reflected back to space through albedo (Stephens et al., 2015). Averaged over the year and both hemispheres, this albedo factor of 0.29 can be roughly broken down into three components[10]: clouds (~47%); surface reflection (~13%); and atmospheric reflection (~40%).

If there are secular changes in any of these albedo components, e.g., through large-scale changes in cloud cover or ice cover, then these climate changes could alter the fraction of the solar irradiance absorbed — independently of the true solar variability. Moreover, we discussed in Section 2.3 how some researchers have argued that solar variability could induce climate change by influencing cloud formation (e.g., Svensmark, 2007; Svensmark et al., 2009). However, if we assume that the planetary albedo remains relatively constant at 0.29, then we can calculate the Radiative Forcing (RF) from the Total Solar Irradiance (TSI) as follows:

$$RF = \frac{TSI(1 - 0.29)}{4}.$$

Therefore, a change in TSI of 1 W/m^2 would cause a change in RF of 0.178 W/m^2. Hence, according to our model derived from Fig. 28, a solar Radiative Forcing of 1 W/m^2 on average causes a change of 1.18 °C. This is more than twice the value of 0.5 °C/W m^{-2} assumed by Gray et al. (2010).

If the above model is accurate, then we can determine the maximum possible influence of Greenhouse Gases (GHG) on Northern Hemisphere by fitting the residuals to the changes since 1881 in either (a) the calculated Radiative Forcing from "well-mixed greenhouse gases" dataset used by CMIP5 simulations or (b) carbon dioxide (CO_2) concentrations. We do this in Fig. 29, using datasets taken from the Supplementary Materials accompanying Schmidt et al. (2011), although since those datasets finish in 2000, we extended the CO_2 dataset using Mauna Loa measurements.

[10] We have calculated this breakdown in the albedo components using the figures in Section 4 of Stephens et al. (2015).

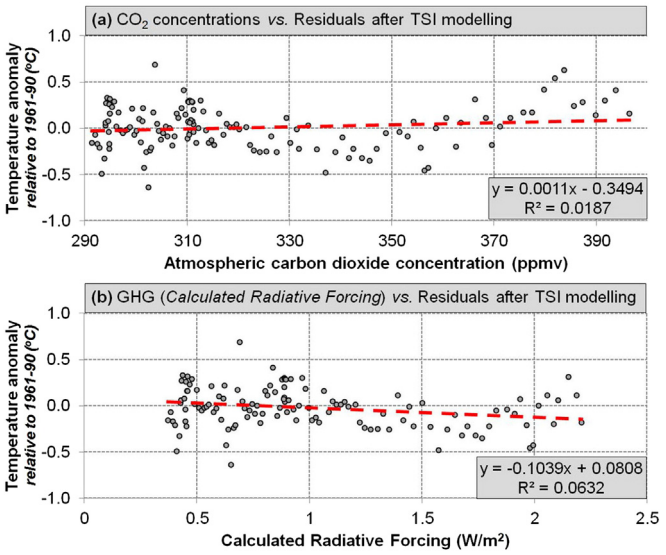

Fig. 29. Attempts to fit the residuals from the solar model using changes in greenhouse gas concentrations.

There does not appear to be much of a relationship between the residuals and either of the datasets, suggesting that greenhouse gases do not appear to have much influence on Northern Hemisphere surface air temperatures, after the above solar relationship has been accounted for.

There did seem to be a slight positive relationship between the residuals and carbon dioxide (CO_2) concentrations, but fitting using the calculated greenhouse gas (GHG) radiative forcing values actually gave a slightly *negative* relationship. However, in both cases, the R^2 values from linear least squares fitting were almost zero, so the fits are probably not significant.

Nonetheless, let us ignore the negative relationship with greenhouse gas (GHG) radiative forcing, and assume the carbon dioxide (CO_2) relationship is valid. If atmospheric carbon dioxide concentrations have risen by ~110 ppmv since 1881 (i.e., 290 → 400 ppmv), this would imply that carbon dioxide (CO_2) is responsible for a warming of at most 0.0011 × 110 = 0.12 °C over the 1881–2014 period, where 0.0011 is the slope of the line in Fig. 29(a).

We can use this relationship to calculate the so-called "climate sensitivity" to carbon dioxide, i.e., the temperature response to a doubling of atmospheric carbon dioxide. According to this model, if atmospheric carbon dioxide concentrations were to increase by ~400 ppmv, this would contribute to at most 0.0011 × 400 = 0.44 °C warming. That is, the climate sensitivity to atmospheric carbon dioxide is at most 0.44 °C. We note that this estimate of the climate sensitivity is substantially less than predicted by current Global Climate Models, and disagrees with the conclusions of the most recent IPCC reports (Bindoff et al., 2013). However, it is quite similar to the 0.4 °C estimate of Idso (1998). It is also worth noting that if we define the climate sensitivity in terms of a doubling of atmospheric carbon dioxide from pre-industrial concentrations, i.e., ~280 ppmv (as is sometimes done), then the climate sensitivity is further reduced to 0.31 °C (i.e., 0.0011 × 280). This illustrates that, despite its popularity in the scientific literature (e.g., Bindoff et al., 2013), "climate sensitivity" can be a rather ambiguous metric (see e.g., Essex, 2011).

If we instead assume that changes in atmospheric carbon dioxide concentrations are predominantly responsible for the changes in Northern Hemisphere surface air temperature since 1881, then we can reverse the process and apply linear least squares fitting directly to the Northern Hemisphere composite — see Fig. 30.

In this case, the attempted fit is considerably worse than the solar model in Fig. 28. Statistically, this is apparent from the fact that the

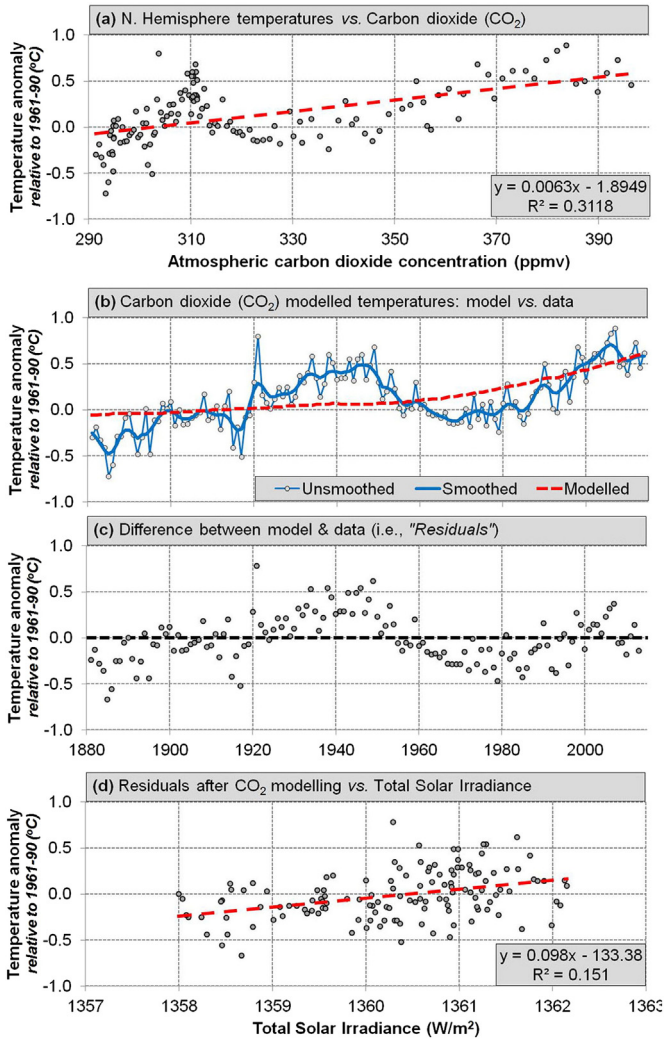

Fig. 30. As for Fig. 28, but fitting using measured changes in atmospheric carbon dioxide (CO_2).

residuals (Fig. 30c) contain noticeable "structure", i.e., the residuals are not completely randomly distributed, but still have trends. In the Supplementary Information, we carry out the equivalent analysis using the calculated greenhouse gas (GHG) radiative forcing values, and obtain similar results.

Moreover, the trends in the residuals from these fits show some similarity to the Hoyt & Schatten Total Solar Irradiance reconstruction. So, in Fig. 30d, we fitted the residuals to Hoyt & Schatten's Total Solar Irradiance reconstruction using linear least squares fitting. Comparing Figs. 28(a) and 30(d), we can see that the expected temperature response to a change in Total Solar Irradiance has more than halved in this second model, i.e., the slope of the linear fit has been reduced from ~0.211 °C/W m^{-2} to ~0.098 °C/W m^{-2}.

Converting this result into the expected response to a solar Radiative Forcing of 1 W/m^2, as before, implies a temperature response of ~0.55 °C/W m^{-2}, which is quite similar to the value of 0.5 °C/W m^{-2} assumed by Gray et al. (2010). Our model in Fig. 30 also implies a larger temperature response to changes in atmospheric carbon dioxide than our model in Figs. 28 & 29: It implies that carbon dioxide (CO_2) is responsible for a warming trend of 0.0063 × 110 = 0.69 °C over the 1881–2014 period, where 0.0063 is the slope of the line in Fig. 30(a). Moreover, according to this model, if atmospheric carbon dioxide concentrations were to increase by ~400 ppmv, this would contribute to 0.0063 × 400 = 2.52 °C warming. This is within the 1.5 °C–4.5 °C "likely" range of equilibrium climate sensitivities predicted by the CMIP5 Global Climate Models (Bindoff et al., 2013), albeit in the lower half of the range. It is worth noting that, as before, if we define the climate sensitivity in terms of a doubling of atmospheric carbon dioxide *from pre-industrial concentrations*, this reduces the climate sensitivity to 0.0063 × 280 = 1.76 °C warming. This reduced value would still be within the 1.5 °C–4.5 °C "likely" CMIP5 range, but at the lower end.

It follows that it is important to establish which of our fitted models is the most reliable. In Fig. 31, we compare our Northern Hemisphere composite to our four different models, i.e.,

1. Fitting our Northern Hemisphere temperature composite only using atmospheric carbon dioxide (CO_2) concentrations
2. First fitting our Northern Hemisphere temperature composite using atmospheric carbon dioxide (CO_2) concentrations, and then fitting the residuals to changes in Total Solar Irradiance
3. Fitting our Northern Hemisphere temperature composite only using changes in Total Solar Irradiance
4. First fitting our Northern Hemisphere temperature composite using changes in Total Solar Irradiance, and then fitting the residuals to atmospheric carbon dioxide (CO_2) concentrations.

The first two models describe our Northern Hemisphere composite rather poorly, while the latter two models describe the composite surprisingly well. The better fits of the latter models can be confirmed by comparing the R^2 correlation coefficients for the fits which are shown for each of the models in Fig. 31, where R^2 is a statistic which varies from 0 (no correlation) to 1 (perfect correlation).

This suggests that, since at least 1881, Northern Hemisphere temperature trends have been *primarily* influenced by changes in

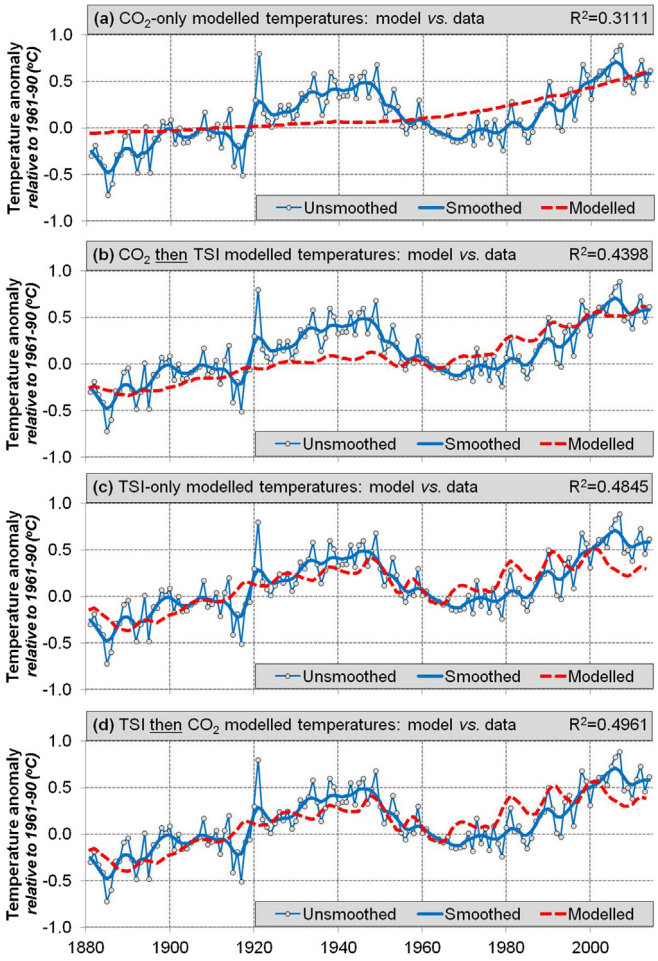

Fig. 31. Comparison of the four different models to our Northern Hemisphere composite (land & ocean rescaled version).

Total Solar Irradiance, as opposed to atmospheric carbon dioxide concentrations. Note, however, that this result does *not* rule out a secondary contribution from atmospheric carbon dioxide. Indeed, the correlation coefficient for Model 4 is slightly better than Model 3 (i.e., ~0.50 vs. ~0.48). However, as we discussed above, this model (Model 4) suggests that changes in atmospheric carbon dioxide are responsible for a warming of at most ~0.12 °C over the 1881–2014 period, i.e., it has so far only had at most a modest influence on Northern Hemisphere temperature trends.

From the above discussion, it appears that there is a quite strong correlation between the Hoyt & Schatten Total Solar Irradiance reconstruction and Northern Hemisphere surface air temperature trends once urbanization bias has been accounted for. However, as we discussed in Section 1.1, correlation does not imply causation. Therefore, it is important to consider which of our four types of correlations apply.

As we discussed in Section 1.1, if the correlation is *either* causal or commensal, then it should not make too much difference to our conclusions which of the two types it is. That is, if the changes in Total Solar Irradiance are caused by some other aspect of solar variability which commensally also has a strong influence on Northern Hemisphere temperature trends, then we can still conclude that Northern Hemisphere temperature trends are strongly influenced by "solar variability", even if we do not know the exact form of solar variability that is involved.

We saw in Section 2 that there are quite a few subjective decisions which need to be made when constructing an estimate of the long term changes in solar activity, e.g., which of the satellite composites is most reliable (Section 2.1) and which (if any) of the solar proxies in Section 2.2 (e.g., those in Figs. 4 and 5) best capture the true trends in solar variability. Depending on these decisions, there are many different possible solar trend reconstructions, e.g., compare those in Fig. 8. For these reasons, it is possible that the strong correlations between the Hoyt & Schatten reconstruction and our Northern Hemisphere composite might just be a coincidence.

In Section 3 we saw that there were also some approximations and decisions made in the construction of the Northern Hemisphere dataset which involve some subjectivity. For instance, although the adjustments we applied to our contiguous U.S. for siting biases were quite small (Fig. 14), if it turns out that larger adjustments are necessary, this could slightly alter the trends for our U.S. estimate. Similarly, we noted in Section 3.4 that there are significant uncertainties over the relative magnitudes of the warm and cool periods for our Arctic regional temperature trend estimates. If these, or any other approximations or decisions were inappropriate, this could have left non-climatic biases in one or more of our regional estimates. Depending on the net signs and magnitude of these potential biases, it is possible this could have coincidentally made the correlations in Fig. 27 appear stronger than they actually are.

Still, the fact that the correlations exists for multiple, widely separated, regions within the Northern Hemisphere (i.e., contiguous U.S., China, Arctic and Ireland) suggests that it is *probably* not just a coincidental correlation. This is reinforced by the fact that the correlation seems quite strong over the entire length of the record (1881–2013), and outside that period, it continues to hold for the longer components, e.g., the Arctic component covers the period 1830–2013 (Fig. 27). Moreover, the Hoyt & Schatten dataset was originally constructed in 1993 (Hoyt and Schatten, 1993). So, the fact that the correlation still seems to hold for the updated dataset with more than 20 years extra data (Scafetta and Willson, 2014) supports the idea that it is not just a coincidental correlation.

There is still the possibility that constructional correlations are involved, however. When Hoyt and Schatten (1993) initially constructed their dataset, *part* of their justification for using some of their solar proxies was that they had apparently shown some agreement with Northern Hemisphere temperature trends (Hoyt, 1979; Friis-Christensen and Lassen, 1991; Hoyt and Schatten, 1993). Therefore, it is plausible that this may have prematurely biased Hoyt & Schatten into believing their chosen proxies were more relevant than they were.

On the other hand, other researchers have used similar arguments to justify the use of other solar proxies which imply different trends from the proxies used by Hoyt and Schatten (1993). For instance, Reid (1991) argued that the solar cycle envelope derived from sunspot numbers is a better solar proxy, partly due to an apparent correlation to sea surface temperature trends, and Cliver et al. (1998) favoured the *aa* geomagnetic index partly on the basis of an apparent correlation to global temperature trends.

Second, the main justifications that Hoyt & Schatten based their choices on were physical grounds, i.e., they chose their various proxies because there are plausible links between the solar variability aspects they describe and changes in the long term Total Solar Irradiance (e.g., through changes in convective patterns). Also, by using multiple different solar proxies – each capturing a different aspect of solar variability – it is likely that the composite of them all has captured some of the key relevant underlying trends in solar activity.

Third, as discussed at the start of Section 5, while Hoyt and Schatten (1993) did note that their solar reconstruction seemed very well correlated to Northern Hemisphere temperature trends for most of their reconstruction, they found a divergence in the last few decades.

In addition, from Fig. 8, it can be seen that other, independent, Total Solar Irradiance estimates imply similar trends to the Hoyt & Schatten dataset, e.g., Ammann et al. (2007)'s update to the Bard et al. (2000) estimate which was derived from cosmogenic nuclides.

Finally, from reading Hoyt and Schatten (1997) which describes Hoyt & Schatten's approach in more detail, it appears that these authors were acutely aware of the dangers of constructional correlation, and made a conscious effort to try and avoid confirmation bias.

With this in mind, it is worth providing some discussion of the proposed mechanisms by which solar variability could be influencing Northern Hemisphere surface air temperatures. As this is a collaborative paper, the authors have slightly different explanations to explain the above results. So, we will provide two separate discussions below. However, we believe that both discussions may be complementary.

5.2. Soon's proposed mechanisms: dynamical considerations

As discussed earlier, one of us (WS) has already identified strong correlations between the updated Hoyt & Schatten solar reconstruction and temperature trends for three of the regions considered in Section 3, i.e., the Arctic (Soon, 2005, 2009), China (Soon et al., 2011) and the U.S. (Robinson et al., 2007). However, the results shown in Fig. 27 now reveal that these correlations are *not* just regional, but apply to the entire Northern Hemisphere — once urbanization bias is accounted for.

With this in mind, it is worth investigating the mechanisms by which this strong solar influence on climate could be maintained across the Northern Hemisphere. We saw in Section 2.3 that much of the recent research into sun-climate relationships has focused on possible "solar amplification" mechanisms, due to the perception that there has not been enough solar variability to explain recent temperature trends. However, from the discussion in Section 5.1, we can see that this is not actually a problem — if the updated Hoyt & Schatten reconstruction is accurate.

Instead, an important issue which has yet to be resolved is the relative roles played by *direct* solar irradiance and *indirect* dynamical interactions between different climatic regions of the Earth. Additionally, if indirect mechanisms are a significant component, a better understanding of the dynamical interactions between different regions of the Earth could provide us with useful further insights into how and why the climate changes.

As explained in Section 5.1, changes in the average Total Solar Irradiance should directly influence surface air temperatures across the entire Northern Hemisphere simply as a result of the changes in direct solar irradiance. It seems likely that this direct solar influence is a major

component for the correlations in Fig. 27 (as MC & RC argue in Section 5.3). However, there is considerable evidence to suggest that indirect solar influences through various climatic feedbacks and/or interactions between different climatic "centres of climatic action" (Soon, 2009) could also be important components in the correlations, e.g.,

- A close inspection of the y-axes in Fig. 27 reveals that the magnitude of the temperature trends is generally larger at higher latitudes, i.e., there is "polar amplification". This implies a non-uniform solar influence on different climatic regions, suggesting that various feedback mechanisms (positive or negative) may be important.
- Several recent paleoclimatic studies have provided empirical evidence for time delays between solar activity forcing and certain climatic variables (e.g., Hormes et al., 2006; Eichler et al., 2009; Cruz-Rico et al., in press). This suggests an important role for feedbacks which provide lagged indirect secondary climatic influences — see also Soon (2009).
- Wyatt and Curry (2014) have noted that several important climatic indices seem to fluctuate at roughly the same quasi-periodic 50–80 year tempo, but that these indices are often out of phase with each other. They suggest that a sequence of feedbacks allow climatic signals to slowly propagate through the climate system over multidecadal timescales. Although, it must be noted that this "stadium wave hypothesis" is being debated — e.g., Mann et al. (2014) vs. Kravtsov et al. (2014).

Several studies have indicated that the Arctic and North Atlantic are particularly important centres of climatic action as a result of the various interconnected interactions between sea level pressures, the associated wind fields, Arctic sea ice patterns and the Atlantic Ocean circulation (e.g., Dima and Lohmann, 2007; van Loon et al., 2012; Courtillot et al., 2013; Klöwer et al., 2014; Wyatt and Curry, 2014; McCarthy et al., 2015). Soon (2009) describes a framework by which these multiple factors can interact to amplify the direct solar influence of changing Total Solar Irradiance on the Arctic, through several important feedback mechanisms.

Soon and Legates (2013) noted that changes in the so-called Equator to Pole Temperature Gradient (EPTG) are also strongly correlated to the updated Hoyt & Schatten Total Solar Irradiance reconstruction. Although the poles are obviously colder than the Equator, the exact magnitudes of the differences between temperatures at different latitudes varies over time, and the EPTG is a useful metric for describing these variations. Several researchers have proposed that small changes in the EPTG could induce quite substantial changes on the atmospheric and oceanic circulation, and thereby factors that influence the EPTG can be a major indirect driver of global climate change (e.g., Lindzen, 1994; Davis and Brewer, 2011).

Specifically, if changes in the Total Solar Irradiance alter the EPTG, this will alter the atmospheric heat and moisture transport from the tropics to the Arctic. This will in turn influence the export of Arctic sea ice into the North Atlantic, and thereby alter both the thermal and saline components of the North Atlantic thermohaline circulation, including the Atlantic Meridional Overturning Circulation. This provides an indirect north–south link between solar-induced changes in the Arctic and the lower latitudes.

Direct monitoring of the Atlantic Meridional Overturning Circulation only began in 2004. So, studying the long-term variations in this circulation system generally involves estimating trends from *related* variables for which longer records are available, e.g., Klöwer et al. (2014). However, climate model simulations of an idealised Earth subjected to a sinusoidally-varying solar input (periods of either 60 or 100 years) by Park and Latif (2012) suggest that solar variability could induce a "*rather complex and nonlinear*" response on the Atlantic Meridional Overturning Circulation.

Soon et al. (2011) noted that the circumglobal teleconnection pattern described by Ding and Wang (2005) also provides an important east–west link between the various climatic regions across the Eurasian continent, via the jet streams. This could provide an important mechanism for linking the Northern Hemisphere climatic systems longitudinally. Ding and Wang (2005) suggest that this circumglobal teleconnection also interacts with the El Niño-Southern Oscillation (ENSO), via the Indian summer monsoon, thereby providing additional links between the Northern and Southern Hemispheres, although that is beyond the scope of this paper.

5.3. Connolly & Connolly's proposed mechanisms

Two of us (MC & RC) have recently written a series of three companion papers providing new insights into the physics of the Earth's atmosphere (Connolly and Connolly, 2014f,g,h). Although the findings in these papers are not exclusively related to understanding possible solar-terrestrial climate links, they may provide plausible mechanisms by which such links could occur.

In Paper 1 (Connolly and Connolly, 2014f), we identified a previously-overlooked phase change associated with the troposphere/tropopause transition. When this phase change is taken into account, it is found that most of the atmospheric temperature profile from the ground to at least the mid-stratosphere can be explained in terms of the properties of the bulk gases (N_2, O_2 and Ar) and water vapour (H_2O). This implies that the troposphere and stratosphere are roughly in Thermodynamic Equilibrium with each other, as opposed to being only in Local Thermodynamic Equilibrium, as implicitly assumed by the current climate models.

Although the variability in the ultraviolet (UV) region of the Total Solar Irradiance seems to be quite substantial relative to the lower frequencies, most of the ultraviolet irradiance is absorbed in the stratosphere. As a result, it had been assumed that this ultraviolet variability would not *directly* influence tropospheric temperatures (e.g., Gray et al., 2010; Haigh et al., 2010; Lean and DeLand, 2012) — see Section 2.3. However, if the troposphere and stratosphere are in Thermodynamic Equilibrium with each other, then this relatively high variability in the ultraviolet regions *could* directly influence tropospheric climate.

In Paper 2 (Connolly and Connolly, 2014g), we argue that this phase change is associated with the formation of multimers of oxygen and/or nitrogen above the troposphere. If this is correct, this has several implications:

a) Our calculations suggest that the phase conditions corresponding to the thermodynamic stability of multimers could be influenced by both solar variability and variability in the geomagnetic field. Therefore, it is plausible that changes in either of these factors could alter the region of the atmosphere over which multimers occur.
b) In Paper 1, we showed that the altitude at which the phase change occurs directly affects the air temperature at ground level. Therefore, if solar variability influences the phase change conditions, this could directly change the average Northern Hemisphere surface air temperatures.
c) In Paper 2, we argue that rapid changes in the altitude at which the phase change occurs at a given location are responsible for the direction and path of the jet streams. The path and direction of the jet streams has been found to indirectly influence regional climate in the lower troposphere, e.g., regional surface air temperatures. It has also been suggested that the jet streams are influenced by solar variability (e.g., Haigh, 1996).

The conventional mechanisms for thermal energy transmission (i.e., conduction, enthalpic convection and radiation) and mechanical energy transmission (i.e., kinetic convection) do not appear to be rapid enough to keep the troposphere and stratosphere in Thermodynamic Equilibrium with each other. This is why it was implicitly *assumed* they were only in Local Thermodynamic Equilibrium. However,

in Paper 3 (Connolly and Connolly, 2014h), we note that the current atmospheric models fail to account for an additional "*through-mass*" mechanical energy transmission mechanism, which is distinct from the "*with-mass*" mechanism of kinetic convection. Within the conventional *thermal* energy transmission mechanisms, conduction is a "*through-mass*" mechanism, i.e., the thermal energy is transmitted through the mass without the mass itself having to move, while enthalpic convection is a "*with-mass*" mechanism, i.e., the thermal energy is transported with the mass (in other words, through the movement of the air). Our laboratory measurements of the relative rates of these energy transmission mechanisms in air indicate that, although the thermal "*through-mass*" mechanism (i.e., conduction) only plays a negligible role in the atmosphere (since air is a poor conductor), "*through-mass*" mechanical energy transmission is actually a very rapid mechanism for transmitting energy throughout the atmosphere (and probably throughout other fluids, e.g., the oceans). This could explain why the troposphere and stratosphere appear to be in Thermodynamic Equilibrium with each other.

6. Conclusions

We have constructed a new estimate of Northern Hemisphere surface air temperature trends derived from mostly rural stations — thereby minimizing the problems introduced to previous estimates by urbanization bias. Similar to previous estimates, our composite implies warming trends during the periods 1880s–1940s and 1980s–2000s. However, this new estimate implies a more pronounced cooling trend during the 1950s–1970s. As a result, the relative warmth of the mid-20th century warm period is comparable to the recent warm period — a different conclusion to previous estimates.

Although our new composite implies different trends from previous estimates, we note that it is compatible with Northern Hemisphere temperature trends derived from (a) sea surface temperatures; (b) glacier length records; (c) tree ring widths. However, the recent multi model means of the CMIP5 Global Climate Model hindcasts failed to adequately reproduce the temperature trends implied by our composite, even when they included both "*anthropogenic and natural forcings*".

One reason why the hindcasts might have failed to accurately reproduce the temperature trends is that the solar forcings they used all implied relatively little solar variability. However, in this paper, we carried out a detailed review of the debate over solar variability, and revealed that considerable uncertainty remains over exactly how the Total Solar Irradiance has varied since the 19th century.

When we compared our new composite to one of the high solar variability reconstructions of Total Solar Irradiance which was not considered by the CMIP5 hindcasts (i.e., the Hoyt & Schatten reconstruction), we found a remarkably close fit. If the Hoyt & Schatten reconstruction and our new Northern Hemisphere temperature trend estimates are accurate, then it seems that *most* of the temperature trends since at least 1881 can be explained in terms of solar variability, with atmospheric greenhouse gas concentrations providing at most a minor contribution. This contradicts the claim by the latest Intergovernmental Panel on Climate Change (IPCC) reports that most of the temperature trends since the 1950s are due to changes in atmospheric greenhouse gas concentrations (Bindoff et al., 2013).

Acknowledgements

We thank Noelle Gillespie of Met Éireann for information on the station history and the corresponding parallel measurements for Valentia Observatory. We thank Nicola Scafetta of Università degli Studi di Napoli Federico II for the updated composite HS93 and ACRIM3 TSI record. We also thank Eugene Avrett, Ray Bates, Robert Carter, Ole Humlum, Jan-Erik Solheim, Don Zieman, the editor and the two reviewers for useful comments and feedback.

W.S. would like to thank Eugene Avrett, Sallie Baliunas, Robert Carter, Dave and Steve Fettig, Joseph Kunc, Eugene Parker, Geoff Smith and many friends and family members for their many years of unwavering encouragements and support for carrying out his studies of the sun-climate connections at the Harvard-Smithsonian Center for Astrophysics to the fullest extent possible without fear or prejudice.

All of the research in this collaborative paper was carried out during the authors' free time and at their own expense, and none of the authors received any funding for this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.earscirev.2015.08.010.

References

Agee, E.M., Kiefer, K., Cornett, E., 2012. Relationship of lower-troposphere cloud cover and cosmic rays: an updated perspective. J. Clim. 25, 1057–1060. http://dx.doi.org/10.1175/JCLI-D-11-00169.1.

Ammann, C.M., Meehl, G.A., Washington, W.M., Zender, C.S., 2003. A monthly and latitudinally varying volcanic forcing dataset in simulations of 20th century climate. Geophys. Res. Lett. 30, 1657. http://dx.doi.org/10.1029/2003GL016875.

Ammann, C.M., Joos, F., Schimel, D.S., Otto-Bliesner, B.L., Tomas, R.A., 2007. Solar influence on climate during the past millennium: results from transient simulations with the NCAR Climate System Model. Proc. Natl. Acad. Sci. 104, 3713–3718. http://dx.doi.org/10.1073/pnas.0605064103.

Aparicio, A.J.P., Vaquero, J.M., Carrasco, V.M.S., Gallego, M.C., 2014. Sunspot numbers and areas from the Madrid Astronomical Observatory (1876–1986). Sol. Phys. 289, 4335–4349. http://dx.doi.org/10.1007/s11207-014-0567-x.

Arfeuille, F., Weisenstein, D., Mack, H., Rozanov, E., Peter, T., Brönnimann, S., 2014. Volcanic forcing for climate modeling: a new microphysics-based data set covering years 1600–present. Clim. Past 10, 359–375. http://dx.doi.org/10.5194/cp-10-359-2014.

Arlt, R., Leussu, R., Giese, N., et al., 2013. Sunspot positions and sizes for 1825–1867 from the observations by Samuel Heinrich Schwabe. Mon. Not. R. Astron. Soc. 433, 3165–3172. http://dx.doi.org/10.1093/mnras/stt961.

Babyak, M.A., 2004. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom. Med. 66, 411–421. http://dx.doi.org/10.1097/01.psy.0000127692.23278.a9.

Baliunas, S., Jastrow, R., 1990. Evidence for long-term brightness changes of solar-type stars. Nature 348, 520–523. http://dx.doi.org/10.1038/348520a0.

Baliunas, S., Soon, W., 1995. Are variations in the length of the activity cycle related to changes in brightness in solar-type stars? Astrophys. J. 450, 896–901. http://dx.doi.org/10.1086/176193.

Baliunas, S.L., Donahue, R.A., Soon, W.H., et al., 1995. Chromospheric variations in main-sequence stars. II. Astrophys. J. 438, 269–287. http://dx.doi.org/10.1086/175072.

Balling Jr., R.C., Idso, S.B., 1989. Historical temperature trends in the United States and the effect of urban population growth. J. Geophys. Res. 94 (D3), 3359–3363. http://dx.doi.org/10.1029/JD094iD03p03359.

Balling Jr., R.C., Idso, C., 2002. Analysis of adjustments to the United States Historical Climatology Network (USHCN) temperature database. Geophys. Res. Lett. 29, 1387. http://dx.doi.org/10.1029/2002GL014825.

Balling Jr., R.C., Roy, S.S., 2005. Analysis of spatial patterns underlying the linkage between solar irradiance and near-surface air temperatures. Geophys. Res. Lett. 32, L11702. http://dx.doi.org/10.1029/2005GL022444.

Balmaceda, L.A., Solanki, S.K., Krivova, N.A., Foster, S., 2009. A homogeneous database of sunspot areas covering more than 130 years. J. Geophys. Res. 114, A07104. http://dx.doi.org/10.1029/2009JA014299.

Bard, E., Raisbeck, G., Yiou, F., Jouzel, J., 2000. Solar irradiance during the last 1200 years based on cosmogenic nuclides. Tellus 52B, 985–992. http://dx.doi.org/10.3402/tellusb.v52i3.17080.

Bard, E., Raisbeck, G.M., Jouzel, J., 2007. Comment on "Solar activity during the last 1000 yr inferred from radionuclide records" by Muscheler et al. (2007). Quat. Sci. Rev. 26, 2301–2304. http://dx.doi.org/10.1016/j.quascirev.2007.06.002.

Basri, G., Walkowicz, L.M., Batalha, N., et al., 2011. Photometric variability in Kepler Target Stars. II. An overview of amplitude, periodicity, and roation in first quarter data. Astron. J. 141 (20). http://dx.doi.org/10.1088/0004-6256/141/1/20.

Basri, G., Walkowicz, L.M., Reiners, A., 2013. Comparison of KEPLER photometric variability with the Sun on different timescales. Astrophys. J. 769, 37. http://dx.doi.org/10.1088/0004-637X/769/1/37.

BenMoussa, A., Gissot, S., Schuhle, U., et al., 2013. On-orbit degradation of solar instruments. Sol. Phys. 288, 389–434. http://dx.doi.org/10.1007/s1207-013-0290-z.

Berger, A., 1988. Milankovitch theory and climate. Rev. Geophys. 26, 624–657. http://dx.doi.org/10.1029/RG026i004p00624.

Berger, A., Loutre, M.-F., Tricot, C., 1993. Insolation and Earth's orbital periods. J. Geophys. Res. 98D, 10341–10362. http://dx.doi.org/10.1029/93JD00222.

Bindoff, N.L., Stott, P.A., et al., 2013. Detection and attribution of climate change: from global to regional. In: Stocker, T.F., Qin, D., et al. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group 1 to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, USA.

Bludova, N.G., Obridko, V.N., Badalyan, O.G., 2014. The relative umbral area in spot groups as an index of cyclic variation of solar activity. Sol. Phys. 289, 1013–1028. http://dx.doi.org/10.1007/s11207-013-0370-0.

Callendar, G.S., 1961. Temperature fluctuations and trends over the earth. Q. J. R. Meteorol. Soc. 87, 1–12. http://dx.doi.org/10.1002/qj.49708737102.

Chapman, G.A., Cookson, A.M., Dobias, J.J., Walton, S.R., 2001. An improved determination of the area ratio of faculae to sunspots. Astrophys. J. 555, 462–465. http://dx.doi.org/10.1086/321466.

Chapman, G.A., Dobias, J.J., Arias, T., 2011. Facular and sunspot areas during Solar Cycles 22 and 23. Astrophys. J. 728, 150. http://dx.doi.org/10.1088/0004-637X/728/2/150.

Chapman, G.A., Cookson, A.M., Preminger, D.G., 2013. Modeling Total Solar Irradiance with San Fernando Observatory ground-based photometry: comparison with ACRIM, PMOD, and RMIB composites. Sol. Phys. 283, 295–305. http://dx.doi.org/10.1007/s11207-013-0233-8.

Clette, F., Lefèvre, L., 2012. Are the sunspots really vanishing? J. Space Weather Space Clim. 2, A06. http://dx.doi.org/10.1051/swsc/2012007.

Clette, F., Svalgaard, L., Vaquero, J.M., Cliver, E.W., 2014. Revisiting the sunspot number: a 400-year perspective on the solar cycle. Space Sci. Rev. 186, 35–103. http://dx.doi.org/10.1007/s11214-014-0074-2.

Cliver, E.W., Boriakoff, V., Feynman, J., 1998. Solar variability and climate change: geomagnetic aa index and global surface temperature. Geophys. Res. Lett. 25, 1035–1038. http://dx.doi.org/10.1029/98GL00499.

Connolly, R., Connolly, M., 2014a. Urbanization bias I. Is it a negligible problem for global temperature estimates? Open Peer Rev. J. 28 ((Clim Sci) ver 0.2. http://oprj.net/articles/climate-science/28. [Accessed 3 February 2015]).

Connolly, R., Connolly, M., 2014b. Urbanization bias II. An assessment of the NASA GISS urbanization adjustment method. Open Peer Rev. J. 31 ((Clim Sci) ver 0.1. http://oprj.net/articles/climate-science/31. [Accessed 27 December 2014]/).

Connolly, R., Connolly, M., 2014c. Urbanization bias III. Estimating the extent of bias in the Historical Climatology Network datasets. Open Peer Rev. J. 34 ((Clim Sci) ver 0.1. http://oprj.net/articles/climate-science/34. [Accessed 27 December 2014]).

Connolly, R., Connolly, M., 2014d. Has poor station quality biased U.S. temperature trend estimates? Open Peer Rev. J. 11 ((Clim Sci) ver 0.1. http://oprj.net/articles/climate-science/11. [Accessed 27 December 2014]).

Connolly, R., Connolly, M., 2014e. Global temperature changes of the last millennium. Open Peer Rev. J. 16 ((Clim Sci) ver 1.0. http://oprj.net/articles/climate-science/16. [accessed 27 December 2014]).

Connolly, M., Connolly, R., 2014f. The physics of the Earth's atmosphere I. Phase change associated with tropopause. Open Peer Rev. J. 19 ((Atm Sci) ver 0.1. http://oprj.net/articles/climate-science/19. [Accessed 27 December 2014]).

Connolly, M., Connolly, R., 2014g. The physics of the Earth's atmosphere II. Multimerization of atmospheric gases above the troposphere. Open Peer Rev. J. 22 ((Atm Sci) ver 0.1. http://oprj.net/articles/climate-science/22. [Accessed 27 December 2014]).

Connolly, M., Connolly, R., 2014h. The physics of the Earth's atmosphere III. Pervective power. Open Peer Rev. J. 25 ((Atm Sci) ver 0.1. http://oprj.net/articles/climate-science/25. [Accessed 27 December 2014]).

Courtillot, V., Le Mouël, J.-L., Kossobokov, V., et al., 2013. Multi-decadal trends of global surface temperature: a broken line with alternating ~30 yr linear segments? Atmos. Clim. Sci. 3, 364–371. http://dx.doi.org/10.4236/acs.2013.33038.

Cowtan, K., Way, R.G., 2014. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. Q. J. R. Meteorol. Soc. 140, 1935–1944. http://dx.doi.org/10.1002/qj.2297.

Crommelynck, D., Fichot, A., Lee III, R.B., Romero, J., 1995. First realisation of the Space Absolute Radiometric Reference (SARR) during the ATLAS 2 flight period. Adv. Space Res. 16, 17–23. http://dx.doi.org/10.1016/0273-1177(95)00261-C.

Cruz-Rico, J., Rivas, D., Tejeda-Martinez, A., 2014. Variability of surface air temperature in Tampico, northeastern Mexico. Int. J. Climatol. http://dx.doi.org/10.1002/joc.4200 (in press).

Damon, P.E., Peristykh, A.N., 2005. Solar forcing of global temperature change since AD 1400. Clim. Change 68, 101–111. http://dx.doi.org/10.1007/s10584-005-1152-y.

Davis, B.A.S., Brewer, S., 2011. A unified approach to orbital, solar, and lunar forcing based on the Earth's latitudinal insolation/temperature gradient. Quat. Sci. Rev. 30, 1861–1874. http://dx.doi.org/10.1016/j.quascirev.2011.04.016.

de Toma, G., Chapman, G.A., Preminger, D.G., Cookson, A.M., 2013. Analysis of sunspot area over two cycles. Astrophys. J. 770, 89. http://dx.doi.org/10.1088/0004-637X/770/2/89.

Dima, M., Lohmann, G., 2007. A hemispheric mechanism for the Atlantic multidecadal oscillation. J. Clim. 20, 2706–2719. http://dx.doi.org/10.1175/JCLI4174.1.

Ding, Q.H., Wang, B., 2005. Circumglobal teleconnection in the Northern Hemisphere summer. J. Clim. 18, 3483–3505. http://dx.doi.org/10.1175/JCLI3473.1.

Driscoll, S., Bozzo, A., Gray, L.J., et al., 2012. Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. J. Geophys. Res. 117, D17105. http://dx.doi.org/10.1029/2012JD017607.

Eddy, J.A., 1976. The Maunder Minimum. Science 192, 1189–1202. http://dx.doi.org/10.1126/science.192.4245.1189.

Eichler, A., Olivier, S., Henderson, K., et al., 2009. Temperature response in the Altai region lags solar forcing. Geophys. Res. Lett. 36, L01808. http://dx.doi.org/10.1029/2008GL035930.

Essex, C., 2011. Climate theory versus a theory for climate. Int. J. Bifurcation Chaos 21, 3477–3487. http://dx.doi.org/10.1142/S0218127411030672.

Fall, S., Watts, A., Nielsen-Gammon, J., et al., 2011. Analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends. J. Geophys. Res. 116, D14120. http://dx.doi.org/10.1029/2010JD015146.

Field, C.V., Schmidt, G.A., Shindell, D.T., 2009. Interpreting 10Be changes during the Maunder Minimum. J. Geophys. Res. 114, D02113. http://dx.doi.org/10.1029/2008JD010578.

Fligge, M., Solanki, S.K., Unruh, Y.C., 2000. Modelling irradiance variations from the surface distribution of the solar magnetic field. Astron. Astrophys. 353, 380–388.

Fontenla, J.M., Harder, J., Livingston, W., Snow, M., Woods, T., 2011. High-resolution solar spectral irradiance from extreme ultraviolet to far infrared. J. Geophys. Res. 116, D20108. http://dx.doi.org/10.1029/2011JD016032.

Foukal, P., 1993. The curious case of the Greenwich faculae. Sol. Phys. 148, 219–232. http://dx.doi.org/10.1007/BF00645087.

Foukal, P., Fröhlich, C., Spruit, H., Wigley, T.M.L., 2006. Variations in solar luminosity and their effect on the Earth's climate. Nature 443, 161–166. http://dx.doi.org/10.1038/nature05072.

Foukal, P., Bertello, L., Livingston, W.C., et al., 2009. A century of solar Ca II measurements and their implication for solar UV driving of climate. Sol. Phys. 255, 229–238. http://dx.doi.org/10.1007/s11207-009-9330-0.

Friis-Christensen, E., Lassen, K., 1991. Length of the solar cycle: an indicator of solar activity closely associated with climate. Science 254, 698–700. http://dx.doi.org/10.1126/science.254.5032.698.

Fröhlich, C., 2006. Solar irradiance variability since 1978 — revision of the PMOD composite during Solar Cycle 21. Space Sci. Rev. 125, 53–65. http://dx.doi.org/10.1007/s11214-006-9046-5.

Fröhlich, C., 2012. Total solar irradiance observations. Surv. Geophys. 33, 453–473. http://dx.doi.org/10.1007/s10712-011-9168-5.

Fröhlich, C., 2013. Total solar irradiance: what have we learned from the last three cycles and the recent minimum? Space Sci. Rev. 176, 237–252. http://dx.doi.org/10.1007/s11214-011-9780-1.

Galindo, S., Saladino, A., 2008. An early comment on the sunspot-climate connection. Rev. Mex. Fis. E 54, 234–239 (url:http://www.scielo.org.mx/scielo.php?pid=S1870-35422008000200018&script=sci_abstract&tlng=en [accessed May 2015]).

Gray, L.J., Beer, J., Geller, M., et al., 2010. Solar influences on climate. Rev. Geophys. 48, RG4001. http://dx.doi.org/10.1029/2009RG000282.

Groveman, B.S., Landsberg, H.E., 1979. Simulated northern hemisphere temperature departures 1579–1880. Geophys. Res. Lett. 6, 767–769. http://dx.doi.org/10.1029/GL006i010p00767.

Haigh, J.D., 1996. The impact of solar variability on climate. Science 272, 981–984. http://dx.doi.org/10.1126/science.272.5264.981.

Haigh, J.D., Winning, A.R., Toumi, R., Harder, J.W., 2010. An influence of solar spectral variations on radiative forcing of climate. Nature 467, 696–699. http://dx.doi.org/10.1038/nature09426.

Hall, J.C., Lockwood, G.W., 2004. The chromospheric activity and variability of cycling and flat activity solar-analog stars. Astrophys. J. 614, 942–946. http://dx.doi.org/10.1086/423926.

Hall, J.C., Henry, G.W., Lockwood, G.W., Skiff, B.A., Saar, S.H., 2009. The activity and variability of the Sun and sun-like stars. II. Contemporaneous photometry and spectroscopy of bright solar analogs. Astron. J. 138, 312–322. http://dx.doi.org/10.1088/0004-6256/138/1/312.

Hansen, J., Lebedeff, S., 1987. Global trends of measured surface air temperature. J. Geophys. Res. 92D, 13345–13372. http://dx.doi.org/10.1029/JD092iD11p13345.

Hansen, J., Ruedy, R., Sato, M., Lo, K., 2010. Global surface temperature change. Rev. Geophys. 48, RG4004. http://dx.doi.org/10.1029/2010RG000345.

Harrison, R.G., Stephenson, D.B., 2006. Empirical evidence for a nonlinear effect of galactic cosmic rays on clouds. Proc. R. Soc. A 462, 1221–1233. http://dx.doi.org/10.1098/rspa.2005.1628.

Hartmann, D.L., Klein Tank, A.M.G., Rusticucci, M., et al., 2013. Observations: atmosphere and surface. In: Stocker, T.F., Qin, D., et al. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group 1 to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, USA.

Harvey, J.W., 2013. The Sun in time. Space Sci. Rev. 176, 47–58. http://dx.doi.org/10.1007/s11214-010-9726-z.

Hathaway, D.H., 2013. A curious history of sunspot penumbrae. Sol. Phys. 286, 347–356. http://dx.doi.org/10.1007/s11207-013-0291-y.

Hausfather, Z., Menne, M.J., Williams, C.N., et al., 2013. Quantifying the effect of urbanization on U.S. Historical Climatology Network temperature records. J. Geophys. Res. Atmos. 118, 481–494. http://dx.doi.org/10.1029/2012JD018509.

Hawkins, E., Jones, P.D., 2013. On increasing global temperatures: 75 years after Callendar. Q. J. R. Meteorol. Soc. 139, 1961–1963. http://dx.doi.org/10.1002/qj.2178.

Hebert III, L., Tinsley, B.A., Zhou, L., 2012. Global electric circuit modulation of winter cyclone vorticity in the northern high latitudes. Adv. Space Res. 50, 806–818. http://dx.doi.org/10.1016/j.asr.2012.03.002.

Heikkilä, U., Beer, J., Feichter, J., 2008. Modeling cosmogenic radionuclides 10Be and 7Be during the Maunder Minimum using the ECHAM5-HAM general circulation model. Atmos. Chem. Phys. 8, 2797–2809. http://dx.doi.org/10.5194/acp-8-2797-2008.

Henney, C.J., Toussaint, W.A., White, S.M., Arge, C.N., 2012. Forecasting F10.7 with solar magnetic flux transport modeling. Space Weather 10, S02011. http://dx.doi.org/10.1029/2011SW000748.

Herschel, W., 1801a. Observations tending to investigate the nature of the Sun, in order to find the causes or symptoms of its variable emission of light and heat; with remarks on the use that may possibly be drawn from solar observations. Philos. Trans. R. Soc. Lond. 91, 265–318. http://dx.doi.org/10.1098/rstl.1801.0015.

Herschel, W., 1801b. Additional observations tending to investigate the symptoms of the variable emission of the light and heat of the Sun; with trials to set aside darkening glasses, by transmitting the solar rays through liquids; and a few remarks to remove objections that might be made against some of the arguments contained in the former paper. Philos. Trans. R. Soc. Lond. 91, 354–363.

Hinkel, K.M., Nelson, F.E., 2007. Anthropogenic heat island at Barrow, Alaskam during winter: 2001–2005. J. Geophys. Res. 112, D06118. http://dx.doi.org/10.1029/2006JD007837.

Hormes, A., Beer, J., Schluchter, C., 2006. A geochronological approach to understanding the role of solar activity on Holocene glacier length variability in the Swiss Alps. Geogr. Ann. 88A, 281–294. http://dx.doi.org/10.1111/j.0435-3676.2006.00301.x.

Hoyt, D.V., 1979. Variations in sunspot structure and climate. Clim. Change 2, 79–92. http://dx.doi.org/10.1007/BF00138229.

Hoyt, D.V., Schatten, K.H., 1993. A discussion of plausible solar irradiance variations, 1700–1992. J. Geophys. Res. 98, 18895–18906. http://dx.doi.org/10.1029/93JA01944.
Hoyt, D.V., Schatten, K.H., 1997. The role of the sun in climate change. Oxford University Press, New York, USA (279 pp.).
Hoyt, D.V., Schatten, K.H., 1998. Group sunspot numbers: a new solar activity reconstruction. Sol. Phys. 179, 189–219. http://dx.doi.org/10.1023/A:1005007527816.
Hua, L.J., Ma, Z.G., Guo, W.D., 2008. The impact of urbanization on air temperature across China. Theor. Appl. Climatol. 93, 179–194. http://dx.doi.org/10.1007/s00704-007-0339-8.
Idso, S.B., 1998. CO_2-induced global warming: a skeptic's view of potential climate change. Clim. Res. 10, 69–82. http://dx.doi.org/10.3354/cr010069.
Instanes, A., Mjureke, D., 2005. Svalbard airport runway. Performance during a climate-warming scenario. Seventh international conference on the bearing capacity of roads, railways and airfields (27–29 June 2005, Trondheim, Norway. http://www.ivt.ntnu.no/bat/vs/bcra/workshops/workshop1/W1_2%20Instanes_BCRA2005BCRA05paper00241.PDF. Accessed 27 December 2014).
Jones, P.D., Osborn, T.J., Briffa, K.R., 1997. Estimating sampling errors in large-scale temperature averages. J. Clim. 10, 2548–2568. http://dx.doi.org/10.1175/1520-0442(1997)010<2548:ESEILS>2.0.CO;2.
Jones, P.D., Lister, D.G., Osborn, T.J., et al., 2012. Hemispheric and large-scale land-surface air temperature variations: an extensive revision and an update to 2010. J. Geophys. Res. 117, D05127. http://dx.doi.org/10.1029/2011JD017139.
Judge, P.G., Saar, S.H., 2007. The outer solar atmosphere during the Maunder Minimum: a stellar perspective. Astrophys. J. 663, 643–656. http://dx.doi.org/10.1086/513004.
Judge, P.G., Lockwood, G.W., Radick, R.R., et al., 2012. Confronting a solar irradiance reconstruction with solar and stellar data. Astron. Astrophys. 544, A88. http://dx.doi.org/10.1051/0004-6361/201218903.
Jungclaus, J.H., Lorenz, S.J., Timmreck, C., et al., 2010. Climate and carbon-cycle variability over the last millennium. Clim. Past 6, 723–737. http://dx.doi.org/10.5194/cp-6-723-2010.
Karl, T.R., Williams Jr., C.N., Young, P.J., 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. J. Clim. Appl. Meteorol. 25, 145–160. http://dx.doi.org/10.1175/1520-0450(1986)025<0145:AMTETT>2.0.CO;2.
Karl, T.R., Diaz, H.F., Kukla, G., 1988. Urbanization: its detection and effect in the United States climate record. J. Clim. 1, 1099–1123. http://dx.doi.org/10.1175/1520-0442(1988)001<1099:UIDAEI>2.0.CO;2.
Kennedy, J.J., 2014. A review of uncertainty in in situ measurements and data sets of sea surface temperature. Rev. Geophys. 52, 1–32. http://dx.doi.org/10.1002/2013RG000434.
Kilcik, A., Yurchyshyn, V.B., Abramenko, V., et al., 2011. Time distributions of large and small sunspot groups over four solar cycles. Astrophys. J. 731, 30. http://dx.doi.org/10.1088/0004-637X/731/1/30.
Kilcik, A., Yurchyshyn, V.B., Ozguc, A., et al., 2014. Solar cycle 24: curious changes in the relative numbers of sunspot group types. Astrophys. J. 794, L2. http://dx.doi.org/10.1088/2041-8205/794/1/L2.
Kirkby, J., 2007. Cosmic rays and climate. Surv. Geophys. 28, 333–375. http://dx.doi.org/10.1007/s10712-008-9030-6.
Kirkby, J., Curtius, J., Almeida, E., et al., 2011. Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation. Nature 476, 429–433. http://dx.doi.org/10.1038/nature10343.
Klein Goldewijk, K., Beusen, A., Janssen, P., 2010. Long-term dynamic modelling of global population and built-up area in a spatially explicit way: HYDE 3.1. The Holocene 20, 565–573. http://dx.doi.org/10.1177/0959683609356587.
Klöwer, M., Latif, M., Ding, H., et al., 2014. Atlantic meridional overturning circulation and the prediction of North Atlantic sea surface temperature. Earth Planet. Sci. Lett. 406, 1–6. http://dx.doi.org/10.1016/j.epsl.2014.09.001.
Konstantinov, P., Baklanov, A., Varentsov, M., et al., 2014. Experimental urban heat island research of four biggest polar cities in Northern Hemisphere. EGU General Assembly 2014, held 27 April–2 May, 2014 in Vienna, Austria (url:http://adsabs.harvard.edu/abs/2014EGUGA.1610699K. [Accessed 4 January 2015]).
Kopp, G., Lean, J.L., 2011. A new, lower value of total solar irradiance: evidence and climate significance. Geophys. Res. Lett. 38, L01706. http://dx.doi.org/10.1029/2010GL045777.
Kravtsov, S., Wyatt, M.G., Curry, J.A., et al., 2014. Two contrasting views of multidecadal climate variability in the twentieth century. Geophys. Res. Lett. 41, 6881–6888. http://dx.doi.org/10.1002/2014GL061416.
Krivova, N.A., Balmaceda, L., Solanki, S.K., 2007. Reconstruction of solar total irradiance since 1700 from the surface magnetic flux. Astron. Astrophys. 467, 335–346. http://dx.doi.org/10.1051/0004-6361:20066725.
Krivova, N.A., Vieira, L.E.A., Solanki, S.K., 2010. Reconstruction of solar spectral irradiance since the Maunder minimum. J. Geophys. Res. 115, A12112. http://dx.doi.org/10.1029/2010JA015431.
Kuzmina, S.I., Johannessen, O.M., Bengtsson, L., et al., 2008. High northern latitude surface air temperature: comparison of existing data and creation of a new gridded data set 1900–2000. Tellus A 60, 289–304. http://dx.doi.org/10.3402/tellusa.v60i2.15260.
Laken, B.A., Pallé, E., Čalogović, J., Dunne, E.M., 2012. A cosmic ray-climate link and cloud observations. J. Space Weather Space Clim. 2, A18. http://dx.doi.org/10.1051/swsc/2012018.
Langley, S.P., 1888. The new astronomy. Ticknor and Company, Boston, USA (260 pp.).
Laut, P., 2003. Solar activity and terrestrial climate: an analysis of some purported correlations. J. Atmos. Sol. Terr. Phys. 65, 801–812. http://dx.doi.org/10.1016/S1364-6826(03)00041-5.
Lawrimore, J.H., Menne, M.J., Gleason, B.E., et al., 2011. An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. J. Geophys. Res. 116, D19121. http://dx.doi.org/10.1029/2011JD016187.
Le Mouël, J.-L., Courtillot, V., Blanter, E., Shnirman, M., 2008. Evidence for a solar signature in 20th-century temperature data from the USA and Europe. C. R. Geosci. 340, 421–430. http://dx.doi.org/10.1016/j.crte.2008.06.001.
Le Mouël, J.-L., Blanter, E., Shnirman, M., Courtillot, V., 2009. Evidence for solar forcing in variability of temperatures and pressures in Europe. J. Atmos. Sol. Terr. Phys. 71, 1309–1321. http://dx.doi.org/10.1016/j.jastp.2009.05.006.
Le Mouël, J.-L., Blanter, E., Shnirman, M., Courtillot, V., 2010. Solar forcing of the semi-annual variation of length-of-day. Geophys. Res. Lett. 37, L15307. http://dx.doi.org/10.1029/2010GL043185.
Le Mouël, J.-L., Blanter, E., Courtillot, V., et al., 2011. A note on comments on papers published in Journal of Atmospheric and Solar-Terrestrial Physics and our responses. J. Atmos. Sol. Terr. Phys. 73, 2042. http://dx.doi.org/10.1016/j.jastp.2011.04.001.
Le Mouël, J.-L., Blanter, E., Shnirman, M., Courtillot, V., 2012. On secular changes of correlation between geomagnetic indices and variations in solar activity. J. Geophys. Res. 117, A09103. http://dx.doi.org/10.1029/2012JA017643.
Le Treut, H., Somerville, R., Cubasch, U., et al., 2007. Historical overview of climate change. In: Solomon, S., Qin, D., et al. (Eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, USA.
Lean, J., 2000. Evolution of the Sun's spectral irradiance since the Maunder Minimum. Geophys. Res. Lett. 27, 2425–2428. http://dx.doi.org/10.1029/2000GL000043.
Lean, J.L., DeLand, M.T., 2012. How does the Sun's spectrum vary? J. Clim. 25, 2555–2560. http://dx.doi.org/10.1175/JCLI-D-11-00571.1.
Lean, J., Beer, J., Bradley, R., 1995. Reconstruction of solar irradiance since 1610: implications for climate change. Geophys. Res. Lett. 22, 3195–3198. http://dx.doi.org/10.1029/95GL03093.
Leclercq, P.W., Oerlemans, J., 2012. Global and hemispheric temperature reconstruction from glacier length fluctuations. Clim. Dyn. 38, 1065–1079. http://dx.doi.org/10.1007/s00382-011-1145-7.
Lefèvre, L., Clette, F., 2011. A global small sunspot deficit at the base of the index anomalies of solar cycle 23. Astron. Astrophys. 536, L11. http://dx.doi.org/10.1051/0004-6361/201118034.
Legras, B., Mestre, O., Bard, E., Yiou, P., 2010. A critical look at solar-climate relationships from long temperature series. Clim. Past 6, 745–758. http://dx.doi.org/10.5194/cp-6-745-2010.
Leroy, M., 1999. Classification d'un site. Note Technique #35. Direction des Systemes d'Observation, Meteo-France, Trappes, France (url: http://www.ccrom.org/ccrom/IMG/pdf/note_2499.technique35-2.pdf).
Li, Q., Zhang, H., Liu, X., et al., 2004. Urban heat island effect on annual mean temperature during the last 50 years in China. Theor. Appl. Climatol. 79, 165–174. http://dx.doi.org/10.1007/s00704-004-0065-4.
Li, Y., Zhu, L.J., Zhao, X.Y., et al., 2013. Urbanization impact on temperature change in China with emphasis on land cover change and human activity. J. Clim. 26, 8765–8780. http://dx.doi.org/10.1175/JCLI-D-12-00698.1.
Lindsey, R., 2003. Under a variable Sun. NASA Earth Observatory (url: http://earthobservatory.nasa.gov/Features/VariableSun/printall.php [accessed May 2015]).
Lindzen, R.S., 1994. Climate dynamics and global change. Annu. Rev. Fluid Mech. 26, 353–378. http://dx.doi.org/10.1146/annurev.fl.26.010194.002033.
Livingston, W., Penn, M.J., Svalgaard, L., 2012. Decreasing sunspot magnetic fields explain unique 10.7 cm radio flux. Astrophys. J. Lett. 757, L8. http://dx.doi.org/10.1088/2041-8205/757/1/L8.
Lockwood, M., Fröhlich, C., 2008. Recent oppositely directed trends in solar climate forcings and the global mean surface air temperature. II. Different reconstructions of the total solar irradiance variation and dependence on response time scale. Proc. R. Soc. A 464, 1367–1385. http://dx.doi.org/10.1098/rspa.2007.0347.
Lockwood, M., Stamper, R., 1999. Long-term drift of the coronal source magnetic flux and the total solar irradiance. Geophys. Res. Lett. 26, 2461–2464. http://dx.doi.org/10.1029/1999GL900485.
Lockwood, G.W., Skiff, B.A., Henry, G.W., et al., 2007. Patterns of photometric and chromospheric variation among sun-like stars: a 20 year perspective. Astrophys. J. Suppl. 171, 260–303. http://dx.doi.org/10.1086/516752.
Lockwood, M., Bell, C., Woollings, T., et al., 2010. Top-down solar modulation of climate: evidence for centennial-scale change. Environ. Res. Lett. 5, 034008. http://dx.doi.org/10.1088/1748-9326/5/3/034008.
Loehle, C., 2009. A mathematical analysis of the divergence problem in dendroclimatology. Clim. Change 94, 233–245. http://dx.doi.org/10.1007/s10584-008-9488-8.
Love, J.J., 2011. Long-term biases in geomagntic K and aa indices. Ann. Geophys. 29, 1365–1375 http://dx.doi.org/105194/angeo-29-1365-2011.
Love, J.J., 2013. On the insignificance of Herschel's sunspot correlation. Geophys. Res. Lett. 40, 4171–4176. http://dx.doi.org/10.1002/grl.50846.
Lugina, K.M., PYa, Groisman, Ya, Vinnikov K., et al., 2006. Monthly surface air temperature time series area-averaged over the 30-degree latitudinal belts of the globe, 1881–2005. Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. http://dx.doi.org/10.3334/CDIAC/cli.003.
Magee, N., Kavic, M., 2012. Probing the climatological impact of a cosmic ray-cloud connection through low-frequency radio observations. J. Atmos. Sol. Terr. Phys. 74, 224–231. http://dx.doi.org/10.1016/j.jastp.2011.10.003.
Magee, N., Curtis, J., Wendler, G., 1999. The urban heat island effect at Fairbanks, Alaska. Theor. Appl. Climatol. 64, 39–47. http://dx.doi.org/10.1007/s007040050109.
Manley, G., 1974. Central England temperatures: monthly means 1659 to 1973. Q. J. R. Meteorol. Soc. 100, 389–405. http://dx.doi.org/10.1002/qj.49710042511.
Mann, M., Amman, C., Bradley, R., et al., 2003a. On past temperatures and anomalous late-20th century warmth. Eos Trans. 84, 256–257. http://dx.doi.org/10.1029/2003E02700039.

Mann, M., Amman, C., Bradley, R., et al., 2003b. Reply to comment on "On past temperatures and anomalous late-20th century warmth". Eos Trans. 84, 473–476. http://dx.doi.org/10.1029/2003E0440008.

Mann, M.E., Steinman, B.A., Miller, S.K., 2014. On forced temperature changes, internal variability, and the AMO. Geophys. Res. Lett. 41, 3211–3219. http://dx.doi.org/10.1002/2014GL059233.

Marsh, N., Svensmark, H., 2000. Low cloud properties influenced by cosmic rays. Phys. Rev. Lett. 85, 5004–5007. http://dx.doi.org/10.1103/PhysRevLett.85.5004.

Marsh, N., Svensmark, H., 2003. Solar influence on Earth's climate. Space Sci. Rev. 107, 317–325. http://dx.doi.org/10.1023/A:1025573117134.

McCarthy, G.D., Gleeson, E., Walsh, S., 2015. The influence of ocean variations on the climate of Ireland. Weather 70, 242–245. http://dx.doi.org/10.1002/wea.2543.

McIntosh, P.S., 1990. The classification of sunspot groups. Sol. Phys. 125, 251–267. http://dx.doi.org/10.1007/BF00158405.

Mekaoui, S., Dewitte, S., 2008. Total solar irradiance measurement and modelling during Cycle 23. Sol. Phys. 247, 203–216. http://dx.doi.org/10.1007/s11207-007-9070-y.

Mekaoui, S., Dewitte, S., Conscience, C., Chevalier, A., 2010. Total solar irradiance absolute level from DIARAD/SOVIM on the international space station. Adv. Space Res. 45, 1393–1406. http://dx.doi.org/10.1016/j.asr.2010.02.014.

Menne, M.J., Williams Jr., C.N., 2009. Homogenization of temperature series via pairwise comparisons. J. Clim. 22, 1700–1717. http://dx.doi.org/10.1175/2008JCLI2263.1.

Menne, M.J., Williams Jr., C.N., Vose, R.S., 2009. The U.S. Historical Climatology Network monthly temperature data, version 2. Bull. Am. Meteorol. Soc. 90, 993–1007. http://dx.doi.org/10.1175/2008BAMS2613.1.

Menne, M.J., Williams Jr., C.N., Palecki, M.A., 2010. On the reliability of the U.S. surface temperature record. J. Geophys. Res. 115, D11108. http://dx.doi.org/10.1029/2009JD013094.

Menzel, D.H., 1959. Our Sun. Harvard University Press, Cambridge MA (350 pp.).

Mitchell, J.M., 1953. On the causes of instrumentally observed secular temperature trends. J. Meteorol. 10, 244–261. http://dx.doi.org/10.1175/1520-0469(1953)010<0244:OTCOIO>2.0.CO;2.

Mitchell, J.M., 1961. Recent secular changes of global temperature. Ann. N. Y. Acad. Sci. 95, 235–250. http://dx.doi.org/10.1111/j.1749-6632.1961.tb50036.x.

Monin, A.S., Shishkov, Yu.A., 2000. Climate as a problem of physics. Physics-Uspekhi 43 (4), 381–406.

Muñoz-Jaramillo, A., Sheeley Jr., N.R., Zhang, J., deLuca, E.E., 2012. Calibrating 100 years of polar faculae measurements: implications for the evolution of the heliospheric magnetic field. Astrophys. J. 753, 146. http://dx.doi.org/10.1088/0004-637X/753/2/146.

Muñoz-Jaramillo, A., Senkpeil, R.R., Longcope, D.W., et al., 2015. The minimum of solar cycle 23: as deep as it could be? Astrophys. J. 804, 68. http://dx.doi.org/10.1088/0004-637X/804/1/68.

Muscheler, R., Joos, F., Beer, J., et al., 2007a. Solar activity during the last 1000 yr inferred from radionuclide records. Quat. Sci. Rev. 26, 82–91. http://dx.doi.org/10.1016/j.quascirev.2006.07.012.

Muscheler, R., Snowball, I., Joos, F., et al., 2007b. Reply to the comment by Bard et al. on "Solar activity during the last 1000 yr inferred from radionuclide records". Quat. Sci. Rev. 26, 2304–2308. http://dx.doi.org/10.1016/j.quascirev.2007.06.014.

Nagovitsyn, Y.A., Pevtsov, A.A., Livingston, W.C., 2012. On a possible explanation of the long-term decrease in sunspot field strength. Astrophys. J. 758, L20. http://dx.doi.org/10.1088/2041-8205/758/1/L20.

Neff, U., Burns, S.J., Mangini, A., et al., 2001. Strong coherence between solar variability and the monsoon in Oman between 9 and 6 kyr ago. Nature 411, 290–293. http://dx.doi.org/10.1038/350770048.

Nickerson, R.S., 1998. Confirmation bias: a ubiquitous phenomenon in many guises. Rev. Gen. Psychol. 2, 175–220. http://dx.doi.org/10.1037/1089-2680.2.2.175.

Ogurtsov, M.G., Oinonen, M., 2014. Evidence of the solar Gleissberg cycle in the nitrate concentration in polar ice. J. Atmos. Sol. Terr. Phys. 109, 37–42. http://dx.doi.org/10.1016/j.jastp.2013.12.017.

Ortiz, A., Solanki, S.K., Domingo, V., et al., 2002. On the intensity contrast of solar photospheric faculae and network elements. Astron. Astrophys. 388, 1036–1047. http://dx.doi.org/10.1051/0004-6361:20020500.

Park, W., Latif, M., 2012. Atlantic meridional overturning circulation response to idealized external forcing. Clim. Dyn. 39, 1709–1726. http://dx.doi.org/10.1007/s00382-011-1212-0.

Parker, D.E., 2006. A demonstration that large-scale warming is not urban. J. Clim. 19, 2882–2895. http://dx.doi.org/10.1175/JCLI3730.1.

Parker, E.N., 2009. Solar magnetism: the state of our knowledge and ignorance. Space Sci. Rev. 144, 15–24. http://dx.doi.org/10.1007/s11214-008-9445-x.

Parker, D., Horton, B., 2005. Uncertainties in central England temperature 1878–2003 and some improvements to the maximum and minimum series. Int. J. Climatol. 25, 1173–1188. http://dx.doi.org/10.1002/joc.1190.

Penn, M.J., Livingston, W., 2006. Temporal changes in sunspot umbral magnetic fields and temperatures. Astrophys. J. 649, L45–L48. http://dx.doi.org/10.1086/508345.

Peterson, T.C., Gallo, K.P., Lawrimore, J., et al., 1999. Global rural temperature trends. Geophys. Res. Lett. 26, 329–332. http://dx.doi.org/10.1029/1998GL900322.

Polyakov, I.V., Bekryaev, R.V., Alekseev, G.V., et al., 2003. Variability and trends of air temperature and pressure in the maritime Arctic 1875–2000. J. Clim. 16, 2067–2077. http://dx.doi.org/10.1175/1520-0442(2003)016<2067:VATOAT>2.0.CO;2.

Preminger, D., Walton, S.R., Chapman, G.A., 2002. Photometric quantities for solar irradiance modeling. J. Geophys. Res. 107, A11. http://dx.doi.org/10.1029/2001JA009169.

Pustil'nik, L.A., Yom Din, G., 2004. Influence of solar activity on the state of the wheat market in medieval England. Sol. Phys. 223, 335–356. http://dx.doi.org/10.1007/s11207-004-5356-5.

Radick, R.R., Lockwood, G.W., Skiff, B.A., Baliunas, S.L., 1998. Patterns of variation among Sun-like stars. Astrophys. J. Suppl. 118, 239–258. http://dx.doi.org/10.1086/313135.

Reid, G.C., 1991. Solar total irradiance variations and the global sea surface temperature record. J. Geophys. Res. 96D, 2835–2844. http://dx.doi.org/10.1029/90JD02274.

Ren, Y.Y., Ren, G.Y., 2011. A remote-sensing method of selecting reference stations for evaluating urbanization effect on surface air temperature trends. J. Clim. 24, 3179–3189. http://dx.doi.org/10.1175/2010JCLI3658.1.

Ren, G.Y., Chu, Z.Y., Chen, Z.H., Ren, Y.Y., 2007. Implications of temporal change in urban heat island intensity observed at Beijing and Wuhan stations. Geophys. Res. Lett. 34, L05711. http://dx.doi.org/10.1029/2006GL027927.

Ren, G.Y., Zhou, Y.Q., Chu, Z.Y., et al., 2008. Urbanization effects on observed surface air temperature trends in north China. J. Clim. 21, 1333–1348. http://dx.doi.org/10.1175/2007JCLI1348.1.

Rennie, J.J., Lawrimore, J.H., Gleason, B.E., et al., 2014. The international surface temperature initiative global land surface databank: monthly temperature data release description and methods. Geosci. Data J. 1, 75–102. http://dx.doi.org/10.1002/gdj3.8.

Robinson, A.B., Robinson, N.E., Soon, W., 2007. Environmental effects of increased atmospheric carbon dioxide. J. Am. Phys. Surg. 12 (79:90. url:http://www.jpands.org/jpands1203.htm [accessed May 2015]).

Roe, G., 2006. In defense of Milankovitch. Geophys. Res. Lett. 33, L24703. http://dx.doi.org/10.1029/2006GL027817.

Roe, G.H., Baker, M.B., 2014. Glacier response to climate perturbations: an accurate linear geometric model. J. Glaciol. 60, 670–684. http://dx.doi.org/10.3189/2014JoG14J016.

Roe, G.H., O'Neal, M.A., 2009. The response of glaciers to intrinsic climate variability: observations and models of late-Holocene variations in the Pacific Northwest. J. Glaciol. 55, 839–854. http://dx.doi.org/10.3189/002214309790152438.

Rohde, R., Muller, R.A., Jacobsen, R., et al., 2013. A new estimate of the average earth surface land temperature spanning 1753 to 2011. Geoinfor. Geostat. Overview 1, 1. http://dx.doi.org/10.4172/2327-4581.1000101.

Roth, M., 2007. Review of urban climate research in (sub)tropical regions. Int. J. Climatol. 27, 1859–1873. http://dx.doi.org/10.1002/joc.1591.

Saar, S.H., Testa, P., 2011. Stars in magnetic grand minima: where are they and what are they like? In: Mandrini, C.H., Webb, D.F. (Eds.), Proceedings IAU Symposium No. 286 http://dx.doi.org/10.1017/S1743921312005066

Sato, M., Hansen, J.E., McCormick, M.P., Pollack, J.B., 1993. Stratospheric aerosol optical depths, 1850–1990. J. Geophys. Res. 98 (D12), 22987–22994. http://dx.doi.org/10.1029/93JD02553.

Scafetta, N., 2012. A shared frequency set between the historical mid-latitude aurora records and the global surface temperature. J. Atmos. Sol. Terr. Phys. 74, 145–163. http://dx.doi.org/10.1016/j.jastp.2011.10.013.

Scafetta, N., Willson, R.C., 2014. ACRIM total solar irradiance satellite composite validation versus TSI proxy models. Astrophys. Space Sci. 350, 421–442. http://dx.doi.org/10.1007/s10509-013-1775-9.

Schmidt, G.A., Jungclaus, J.H., Ammann, C.M., et al., 2011. Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0). Geosci. Model Dev. 4, 33–45. http://dx.doi.org/10.5194/gmd-4-33-2011.

Schmidt, G.A., Jungclaus, J.H., Ammann, C.M., et al., 2012. Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.1). Geosci. Model Dev. 5, 185–191. http://dx.doi.org/10.5194/gmd-5-185-2012.

Schonfeld, S.J., White, S.M., Henney, C.J., et al., 2015. Coronal sources of the solar F10.7 radio flux. Astrophys. J. 808. http://dx.doi.org/10.1088/0004-637X/808/1/29.

Schurer, A.P., Hegerl, G.C., Mann, M.E., et al., 2013. Separating forced from chaotic climate variability over the past millennium. J. Clim. 26, 6954–6973. http://dx.doi.org/10.1175/JCLI-D-12-00826.1.

Shapiro, A.I., Schmutz, W., Rozanov, E., et al., 2011. A new approach to the long-term reconstruction of the solar irradiance leads to a large historical solar forcing. Astron. Astrophys. 529, A67. http://dx.doi.org/10.1051/0004-6361/201016173.

Shapiro, A.I., Schmutz, W., Cessateur, G., Rozanov, E., 2013. The place of the Sun among the Sun-like stars. Astron. Astrophys. 552, A114. http://dx.doi.org/10.1051/0004-6361/201220512.

Shaviv, N.J., Prokoph, A., Veizer, J., 2014. Is the Solar System's galactic motion imprinted in the Phanerozoic climate? Sci. Rep. 4, 6150. http://dx.doi.org/10.1038/srep06150.

Simmons, J.P., Nelson, L.D., Simonsohn, U., 2011. Psychol. Sci. 20, 1–8. http://dx.doi.org/10.1177/0956797611417632.

Sloan, T., Wolfendale, A.W., 2013. Cosmic rays, solar activity and the climate. Environ. Res. Lett. 8, 045022. http://dx.doi.org/10.1088/1748-9326/8/4/045022.

Solanki, S.K., Fligge, M., 1999. A reconstruction of total solar irradiance since 1700. Geophys. Res. Lett. 26, 2465–2468. http://dx.doi.org/10.1029/1999GL900370.

Solanki, S.K., Fligge, M., 2000. Reconstruction of past solar irradiance. Space Sci. Rev. 94, 127–138. http://dx.doi.org/10.1023/A:1026754803423.

Solanki, S.K., Krivova, N.A., 2003. Can solar variability explain global warming since 1970? J. Geophys. Res. 108 (A5), 1200. http://dx.doi.org/10.1029/2002JA009753.

Solheim, J.-E., Stordahl, K., Humlum, O., 2012. The long sunspot cycle 23 predicts a significant temperature decrease in cycle 24. J. Atmos. Sol. Terr. Phys. 80, 267–284. http://dx.doi.org/10.1016/j.jastp.2012.02.008.

Soon, W.W.H., 2005. Variable solar irradiance as a plausible agent for multidecadal variations in the Arctic-wide surface air temperature record of the past 130 years. Geophys. Res. Lett. 32, L16712. http://dx.doi.org/10.1029/2005GL023429.

Soon, W.W.H., 2009. Solar Arctic-mediated climate variation on multidecadal to centennial timescales: empirical evidence, mechanistic explanation and testable consequences. Phys. Geogr. 30, 144–184. http://dx.doi.org/10.2747/0272-3646.30.2.144.

Soon, W., 2014. Sun shunned. Climate Change—The Facts 2014. Institute of Public Affairs, Melbourne, Australia, pp. 57–66.

Soon, W., Baliunas, S., 2003. Proxy climatic and environmental changes of the past 1000 years. Clim. Res. 23, 89–110. http://dx.doi.org/10.3354/cr023089.

Soon, W., Legates, D.R., 2013. Solar irradiance modulation of equator-to-pole (Arctic) temperature gradients: empirical evidence for climate variation on multi-decadal

timescales. J. Atmos. Sol. Terr. Phys. 93, 45–56. http://dx.doi.org/10.1016/j.jastp.2012.11.015.
Soon, W.W.H., Yaskell, S.H., 2003. The Maunder Minimum and the Variable Sun–Earth Connection. World Scientific Publishing Company, Singapore (278 pp.).
Soon, W., Baliunas, S., Posmentier, E.S., et al., 2000. Variations of solar coronal hole area and terrestrial lower tropospheric air temperature from 1979 to mid-1998: astronomical forcings of change in Earth's climate? New Astron. 4, 563–579. http://dx.doi.org/10.1016/S1384-1076(00)00002-6.
Soon, W., Baliunas, S., Idso, C., et al., 2003a. Reconstructing climatic and environmental changes of the past 1000 years. Energy Environ. 14, 233–296. http://dx.doi.org/10.1260/095830503765184619.
Soon, W., Baliunas, S., Legates, D., 2003b. Comment on "On past temperatures and anomalous late-20th century warmth". Eos Trans. 84, 473–476. http://dx.doi.org/10.1029/2003E0440007.
Soon, W., Dutta, K., Legates, D.R., et al., 2011. Variation in surface air temperature of China during the 20th century. J. Atmos. Sol. Terr. Phys. 73, 2331–2344. http://dx.doi.org/10.1016/j.jastp.2011.07.007.
Soon, W., Velasco Herrera, V.M., Selvaraj, K., et al., 2014. A review of Holocene solar-linked climatic variation on centennial to millennial timescales: physical processes, interpretative frameworks and a new multiple cross-wavelet transform algorithm. Earth-Sci. Rev. 134, 1–15. http://dx.doi.org/10.1016/j.earscirev.2014.03.003.
Stauning, P., 2011. Solar activity-climate relations: a different approach. J. Atmos. Sol. Terr. Phys. 73, 1999–2012. http://dx.doi.org/10.1016/j.jastp.2011.06.011.
Steinhilber, F., Beer, J., Fröhlich, C., 2009. Total solar irradiance during the Holocene. Geophys. Res. Lett. 36, L19704. http://dx.doi.org/10.1029/2009GL040142.
Steinhilber, F., Abreu, J.A., Beer, J., et al., 2012. 9,400 years of cosmic radiation and solar activity from ice cores and tree rings. Proc. Natl. Acad. Sci. 109, 5967–5971. http://dx.doi.org/10.1073/pnas.1118965109.
Stephens, G.L., O'Brien, D., Webster, P.J., et al., 2015. The albedo of Earth. Rev. Geophys. 53, 141–163. http://dx.doi.org/10.1002/2014RG000449.
Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bull. Am. Meteorol. Soc. 93, 1879–1900. http://dx.doi.org/10.1111/j.1541-0064.2000.tb00709.x.
Suzuki, M., 2014. On the long-term modulation of solar differential rotation. Sol. Phys. 289, 4021–4029. http://dx.doi.org/10.1007/s11207-014-0576-9.
Svalgaard, L., Hudson, H.S., 2010. The solar microwave flux and the sunspot number. SOHO23 Conference Proceedings, Astronomical Society of the Pacific (arXiv: 1003.4281v1 [astro-ph.SR]).
Svensmark, H., 2007. Cosmoclimatology: a new theory emerges. Astron. Geophys. 48, 1.18–1.24. http://dx.doi.org/10.1111/j.1468-4004.2007.48118.x.
Svensmark, H., 2012. Evidence of nearby supernovae affecting life on Earth. Mon. Not. R. Astron. Soc. 423, 1234–1253. http://dx.doi.org/10.1111/j.1365-2966.2012.20953.x.
Svensmark, H., Calder, N., 2007. The Chilling Stars. Icon Books Ltd, Cambridge, UK (246 pp.).
Svensmark, H., Friis-Christensen, E., 1997. Variation of cosmic ray flux and global cloud coverage—a missing link in solar-climate relationships. J. Atmos. Terr. Phys. 59, 1225–1232. http://dx.doi.org/10.1016/S1364-6826(97)00001-1.
Svensmark, H., Bondo, T., Svensmark, J., 2009. Cosmic ray decreases affect atmospheric aerosols and clouds. Geophys. Res. Lett. 36, L15101. http://dx.doi.org/10.1029/2009GL038429.
Svensmark, H., Enghoff, M.B., Pedersen, J.O.P., 2013. Response of cloud condensation nuclei (>50 nm) to changes in ion-nucleation. Phys. Lett. A 377, 2343–2347. http://dx.doi.org/10.1016/j.physleta.2013.07.004.
Symmons, P., 2014. The 'heat island' effect and estimates of the temperature of Britain. Weather 69, 40–41. http://dx.doi.org/10.1002/wea.2246.
Tang, G.L., Ren, G.Y., 2005. Reanalysis of surface air temperature change of the last 100 years over China. Clim. Environ. Res. 10, 791–798 (in Chinese).
Tang, G.L., Ding, Y.H., Wang, S.W., et al., 2010. Comparative analysis of China surface air temperature series for the past 100 years. Adv. Clim. Chang. Res. 1, 11–19. http://dx.doi.org/10.3724/SP.J.1248.2010.00011.
Thejll, P., Lassen, K., 2000. Solar forcing of the Northern hemisphere land air temperature: new data. J. Atmos. Sol. Terr. Phys. 62, 1207–1213. http://dx.doi.org/10.1016/S1364-6826(00)00104-8 (Erratum. 64:105. doi:10.1016/S1364-6826(01)00106-7).
Tinsley, B., 2012. A working hypothesis for connections between electrically-induced changes in cloud microphysics and storm vorticity, with possible effects on circulation. Adv. Space Res. 50, 791–805. http://dx.doi.org/10.1016/j.asr.2012.04.008.
Tsonis, A.A., Deyle, E.R., May, R.M., et al., 2015. Dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. Proc. Natl. Acad. Sci. 112, 3253–3256. http://dx.doi.org/10.1073/pnas.1420291112.
Unruh, Y.C., Solanki, S.K., Fligge, M., 1999. The spectral dependence of facular contrast and solar irradiance variations. Astron. Astrophys. 345, 635–642.
Usoskin, I.G., Rainer, A., Asvestari, E., et al., 2015. The Maunder minimum (1645–1715) was indeed a Grand minimum: a reassessment of multiple datasets. Astron. Astrophys. http://dx.doi.org/10.1051/0004-6361/201526652 (in press).
Vahrenholt, F., Lüning, S., 2013. The Neglected Sun. Stacey International, London, UK (412 pp.).
Van Loon, H., Brown, J., Milliff, R.F., 2012. Trends in sunspots and North Atlantic sea level pressure. J. Geophys. Res. 117, D07106. http://dx.doi.org/10.1029/2012JD017502.
Venema, V.K.C., Mestre, O., Aguilar, E., et al., 2012. Benchmarking homogenization algorithms for monthly data. Clim. Past 8, 89–115. http://dx.doi.org/10.5194/cp-8-89-2012.
Vieira, L.E.A., Solanki, S.K., Krivova, N.A., Usoskin, I., 2011. Evolution of the solar irradiance during the Holocene. Astron. Astrophys. 531, A6. http://dx.doi.org/10.1051/0004-6361/201015843.
Vinnikov, K.Y., Groisman, P.Y., Lugina, K.M., 1990. Empirical data on contemporary global climate changes (temperature and precipitation). J. Clim. 3, 662–677. http://dx.doi.org/10.1175/1520-0442(1990)003<0662:EDOCGC>2.0.CO;2.
Voiculescu, M., Usoskin, I., 2012. Persistent solar signatures in cloud cover: spatial and temporal analysis. Environ. Res. Lett. 7, 044004. http://dx.doi.org/10.1088/1748-9326/7/4/044004.
Voiculescu, M., Usoskin, I., Condurache-Bota, S., 2013. Clouds blown by the solar wind. Environ. Res. Lett. 8, 045032. http://dx.doi.org/10.1088/1748-9326/8/4/045032.
Vose, R.S., Williams Jr., C.N., Peterson, T.C., et al., 2003. An evaluation of the time of observation bias adjustment in the U.S. historical climatology network. Geophys. Res. Lett. 30, 2046. http://dx.doi.org/10.1029/2003GL018111.
Wang, F., Ge, Q.S., 2012. Estimation of urbanization bias in observed surface temperature change in China from 1980 to 2009 using satellite land-use data. Chin. Sci. Bull. 57, 1708–1715. http://dx.doi.org/10.1007/s11434-012-4999-0.
Wang, S.W., Gong, D.Y., Zhu, J.H., 2001. Twentieth-century climatic warming in China in the context of the Holocene. The Holocene 11, 313–321.
Wang, S.W., Zhu, J.H., Cai, J.N., 2004. Interdecadal variability of temperature and precipitation in China since 1880. Adv. Atmos. Sci. 21, 307–313. http://dx.doi.org/10.1007/BF02915560.
Wang, Y.-M., Lean, J.L., Sheeley Jr., N.R., 2005. Modeling the Sun's magnetic field and irradiance since 1713. Astrophys. J. 625, 522–538. http://dx.doi.org/10.1085/429689.
White, O.R., Skumanich, A., Lean, J., et al., 1992. The Sun in a noncycling state. Publ. Astron. Soc. Pac. 104, 1139–1143. http://dx.doi.org/10.1086/133100.
Wickham, C., Rohde, R., Muller, R.A., et al., 2013. Influence of urban heating on the global temperature land average using rural sites identified from MODIS classifications. Geoinfor. Geostat. 1, 2. http://dx.doi.org/10.4172/gigs.1000104.
Williams, C.N., Menne, M.J., Thorne, P.W., 2012. Benchmarking the performance of pairwise homogenization of surface temperatures in the United States. J. Geophys. Res. 117, D05116. http://dx.doi.org/10.1029/2011JD016761.
Willson, R.C., 2014. ACRIM3 and the total solar irradiance database. Astrophys. Space Sci. 352, 341–352. http://dx.doi.org/10.1007/s10509-014-1961-4.
Wilson, R., D'Arrigo, R., Buckley, B., et al., 2007. A matter of divergence: tracking recent warming at hemispheric scales using tree ring data. J. Geophys. Res. 112, D17103. http://dx.doi.org/10.1029/2006JD008318.
Wright, J.T., 2004. Do we know of any Maunder Minimum stars? Astron. J. 128, 1273–1278. http://dx.doi.org/10.1086/423221.
Wyatt, M.G., Curry, J.A., 2014. Role for Eurasian Arctic shelf sea ice in a secularly varying hemispheric climate signal during the 20th century. Clim. Dyn. 42, 2763–2782. http://dx.doi.org/10.1007/s00382-013-1950-2.
Xu, W.-H., Li, Q.-X., Yang, S., Xu, Y., 2014. Overview of global monthly surface temperature data in the past century and preliminary integration. Adv. Clim. Chang. Res. 5, 111–117. http://dx.doi.org/10.1016/j.accre.2014.11.003.
Yan, H., Wei, W., Soon, W., et al., 2015. Dynamics of the intertropical convergence zone over the western Pacific during the Little Ice Age. Nat. Geosci. 8, 315–320. http://dx.doi.org/10.1038/ngeo2375.
Yang, X.C., Hou, Y.L., Chen, B.D., 2011. Observed surface warming induced by urbanization in east China. J. Geophys. Res. 116, D14113. http://dx.doi.org/10.1029/2010JD015452.
Yang, Y.J., Wu, B.W., Shi, C.E., et al., 2013. Impacts of urbanization and station-relocation on surface air temperature series in Anhui province, China. Pure Appl. Geophys. 170, 1969–1983. http://dx.doi.org/10.1007/s0024-012-0619-9.
Yeo, K.L., Solanki, S.K., Krivova, N.A., 2013. Intensity contrast of solar network and faculae. Astron. Astrophys. 550, A95. http://dx.doi.org/10.1051/0004-6361/201220682.
Yu, F., 2002. Altitude variations of cosmic ray induced production of aerosols: implications for global cloudiness and climate. J. Geophys. Res. 107, A7. http://dx.doi.org/10.1029/2001JA000248.
Yu, F., Luo, G., 2014. Effect of solar variations on particle formation and cloud condensation nuclei. Environ. Res. Lett. 9, 0450004. http://dx.doi.org/10.1748-9326/9/4/045004.
Zacharias, P., 2014. An independent review of existing total solar irradiance records. Surv. Geophys. 35, 897–912. http://dx.doi.org/10.1007/s10712-014-9294-y.
Zhang, Q., Soon, W.H., Baliunas, S.L., et al., 1994. A method of determining brightness variations of the Sun in past centuries from observations of solar-type stars. Astrophys. J. 427, L111–L114. http://dx.doi.org/10.1086/187377.
Zhou, L., Dickinson, R.E., Tian, Y.H., et al., 2004. Evidence for a significant urbanization effect on climate in China. Proc. Natl. Acad. Sci. 101, 9540–9544. http://dx.doi.org/10.1073/pnas.0400357101.
Zhou, L., Tinsley, B.A., Huang, J., 2014. Effects on winter circulation of short and long term solar wind changes. Adv. Space Res. 54, 2478–2490. http://dx.doi.org/10.1016/j.asr.2013.09.017.