Better defining (or altogether avoiding) the term “pre-industrial” would aid interpretation of internationally agreed global temperature limits and estimation of the required constraints to avoid reaching those limits.

The basis for international negotiations on climate change has been to “prevent dangerous anthropogenic interference with the climate system,” (p. 9) using the words in Article 2 of the United Nations Framework Convention on Climate Change (UNFCCC; United Nations 1992). The 2015 Paris COP21 Agreement (United Nations 2015) aims to maintain global average temperature “well below 2°C above pre-industrial levels and pursuing efforts to limit the temperature increase to 1.5°C above pre-industrial levels” (p. 3). However, there is no formal definition of what is meant by “pre-industrial” in the UNFCCC or the Paris Agreement. Neither did the Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) use the term when discussing when global average temperature might cross various levels because of the lack of a robust definition (Kirtman et al. 2013).

Ideally, a preindustrial period should represent the mean climate state just before human activities...

AFFILIATIONS: Hawkins, Ortega, and Suckling—NCAS-Climate, and Department of Meteorology, University of Reading, Reading, United Kingdom; Schurer and Hegerl—School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom; Jones—Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom, and Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia; Joshi and Osborn—Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom; Masson-Delmotte—Laboratoire des Sciences du Climat et de l’Environnement (CEA-CNRS-UVSQ, Université Paris Saclay), Institut Pierre Simon Laplace, Gif-sur-Yvette, France; Mignot—Climate and Environmental Physics, Physics Institute, and Oeschger Center for Climate Change Research, University of Bern, Bern, Switzerland, and Sorbonne Universités (UPMC-CNRS-IRD-MNHN), and LOCEAN/IPSL, Paris, France; Thorne—Department of Geography, Maynooth University, Maynooth, Ireland; van Oldenborgh—Koninklijk Nederlands Meteorologisch Instituut, De Bilt, Netherlands

CORRESPONDING AUTHOR: Ed Hawkins, e.hawkins@reading.ac.uk

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-D-16-0007.1

In final form 20 December 2016 ©2017 American Meteorological Society For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy.

This article is licensed under a Creative Commons Attribution 4.0 license.
started to demonstrably change the climate through combustion of fossil fuels. Here we discuss which time period might be most suitable, considering various factors such as radiative forcings, availability of observations, and uncertainties in our knowledge.

We will focus on global temperatures, specifically for informing discussions on future temperature limits, and make an assessment of how much global average temperature has already warmed since our defined preindustrial period using a range of approaches. We will also provide recommendations for i) how future international climate reports and agreements might use this assessment and ii) how the assessment itself may be improved in the future, particularly regarding the use of instrumental data, proxy evidence, and simulations of past climate.

RELEVANCE OF THE PREINDUSTRIAL PERIOD FOR CROSSING GLOBAL TEMPERATURE THRESHOLDS. In the absence of a formal definition for preindustrial, the IPCC AR5 made a pragmatic choice to reference global temperature to the mean of 1850–1900 when assessing the time at which particular temperature levels would be crossed (Kirtman et al. 2013). In the final draft, 1850–1900 was referred to as preindustrial, but at the IPCC AR5 plenary approval session, “a contact group developed a proposal, in which reference to ‘pre-industrial’ is deleted, and this was adopted [by the governments]” (IISD 2013). However, the term preindustrial was used in AR5, often inconsistently, in other contexts—for example, when discussing atmospheric composition, radiative forcing (the year 1750 is used as a zero-forcing baseline), sea level rise, and paleoclimate information. These discussions highlight the importance of defining preindustrial consistently and more precisely.

In AR5, the observed increase in global temperature was calculated as the mean of 1986–2005 minus the mean of 1850–1900 in the HadCRUT4 dataset (0.61°C; Morice et al. 2012), which was the only combined global land and ocean temperature dataset available back to 1850 at the time. The 1986–2005 modern period was chosen because the design of the CMIP5 simulations required a recent reference baseline for the projections of future climate [discussed further in Hawkins and Sutton (2016)]. Note that the warming between 1850–1900 and the most recent decade covered (2003–12) was given by AR5 as 0.78° ± 0.03°C (IPCC 2013).

The choice of 1850–1900 as the historical reference period benefits from relatively widespread, but still sparse, temperature observations, and quantified uncertainties in the estimates of global temperature. Since the AR5, two further datasets have been produced that allow a comparison for the 1850–1900 period. In the Cowtan and Way (2014, hereafter CW14) dataset, which is based on interpolating the spatial gaps in HadCRUT4, the difference from 1850–1900 to 1986–2005 is 0.65°C and in the Berkeley Earth global land and sea data (BEST-GL; berkeleyearth.org), it is 0.71°C, suggesting that the AR5 value may be slightly too low. Also, Cowtan et al. (2015) presented GCM-based evidence that sparse observation-based datasets may have significantly underestimated the changes in global surface air temperature due to slower warming regions being preferentially sampled in the past. However, infilling the gaps in the early period is especially problematic owing to the sparse observations and may accentuate the dominant observed anomaly.

However, some anthropogenic warming is estimated to have already occurred by 1850 (Hegerl et al. 2007; Schurer et al. 2013; Abram et al. 2016) as greenhouse gas concentrations had started increasing around a century earlier (Fig. 1). On the other hand, the 1880s and 1890s were cooler than the preceding decades because of the radiative impact of aerosols from several volcanic eruptions (Fig. 1), which may have compensated for the earlier anthropogenic influence. It is therefore plausible that a “true” preindustrial temperature could be warmer or cooler than 1850–1900, depending on the balance of these two factors. A key question which we will consider is how representative the 1850–1900 period is for preindustrial global average temperature.

DEFINING A SUITABLE PREINDUSTRIAL PERIOD USING RADIATIVE FORCING ESTIMATES. Anthropogenic climate change is occurring on top of i) internal climate variability, such as ENSO, the Pacific decadal oscillation (PDO), Atlantic multidecadal variability (AMV), and possibly longer time scales [see Deser et al. (2010)]
for a review] and ii) multidecadal scale variations in natural radiative forcings, such as solar activity, changes in Earth’s orbit, and the frequency of large volcanic eruptions.

A preindustrial climate should therefore be defined as a period that is close to the present but before the “industrial age,” with small anthropogenic forcings. Ideally, levels of natural forcings would also be similar to present and widespread direct or indirect observations would be available. The better part of a century would appear to be required to average over the longer-time-scale internal variations.

Unfortunately, such a perfect time period does not exist so compromises have to be made. In particular, there are very few instrumental temperature records before 1850, which limits our ability to determine pre-1850 global temperatures. Changes in land use and other human activities (e.g., biomass burning, deforestation) may have altered the composition of the atmosphere several millennia ago (Ruddiman 2003; Ruddiman et al. 2016). There are also variations in greenhouse gas concentrations (of a few ppm) before 1700 (Bauska et al. 2015). However, we assume that these early influences are not relevant for defining a preindustrial period for use by policymakers.

Bradley et al. (2016) identified the period 725–1025 as a “medieval quiet period,” without major tropical eruptions or solar variations, and that might represent a reference climate state. However, proxy evidence suggests a slow decline of global temperatures, surface ocean temperatures, and reductions in sea level over the last two millennia, which has been attributed to orbital forcing (Kaufman et al. 2009) or to increasing volcanic activity (McGregor et al. 2015; Stoffel et al. 2015; Kopp et al. 2016). Given this multimillennial trend, whatever its cause, it makes sense to choose a reference period as close to the present as possible.
An important moment at the start of the industrial age was when James Watt patented the steam engine condenser in 1769, dramatically improving Thomas Newcomen’s 1712 steam engine design. Various agricultural revolutions also began around the same time. However, there was probably only a small climate effect of these developments for several decades at least. For these reasons, historical anthropogenic radiative forcings are often considered relative to 1750 levels (IPCC 2007; Meinshausen et al. 2011).

It is also important to ensure that the natural forcings in any chosen period are not unusual, compared to the present (Fig. 1). The period before 1720, often called the Little Ice Age (Mann et al. 2009), was influenced by several large tropical volcanic eruptions in the 1600s (Briffa et al. 1998; Crowley et al. 2008; Gao et al. 2008; Sigl et al. 2013) and the Maunder Minimum in solar activity, which finished in the early 1700s (Steinhilber et al. 2009; Lockwood et al. 2014; Usoskin et al. 2015). The period after 1800 is influenced by the Dalton Minimum in solar activity and the large eruptions of an unlocated volcano in 1808/09, Tambora (1815; Raible et al. 2016), and several others in the 1820s and 1830s. In addition, greenhouse gas concentrations had already increased slightly by this time (Fig. 1).

In contrast, between 1720 and 1800 the evidence suggests that natural radiative forcings are closer to modern levels, with only very weak anthropogenic forcings. It could be argued that this period has slightly anomalously low volcanic activity, including one relatively small tropical eruption (Makian, Indonesia, in 1761) and one long-lasting northern extratropical eruption (Laki, Iceland, in 1783). This issue is returned to later.

There is also no evidence for unusual AMV/PDO variability during the 1720–1800 period (e.g., Gray et al. 2004; MacDonald and Case 2005), suggesting that these modes of variability are not expected to significantly affect the multidecadal temperature average.

We, therefore, suggest that 1720–1800 is the most suitable period to be called preindustrial for assessing global temperature levels in terms of the radiative forcings and we concentrate on this period in the analysis that follows. Different choices may be made if considering changes in other variables (Knutti et al. 2015), such as regional temperatures, rainfall, sea level, carbon storage, or glacier extents, but assessing those is beyond the scope of this study.

Using three different approaches, we now address two related questions, based on the reference periods used in IPCC AR5: i) what is the global temperature change from our preindustrial choice to a recent baseline (1986–2005) and ii) is 1850–1900 a reasonable pragmatic surrogate for the preindustrial period? We also consider the precision to which such questions can be answered.

**Approach 1: Using Radiative Forcing s.** Our first approach uses radiative forcings to estimate changes in global temperature before the available observations. Phase 5 of the Coupled Model Intercomparison Project (CMIP5) provides estimated historical radiative forcings for 1765–2005, referenced to 1750, and for a range of representative concentration pathways (RCPs) after 2005 (Meinshausen et al. 2011). We use RCP4.5 for the period 2006–15 but this makes little difference.

We adopt a weighted least squares multiple linear regression approach, using the radiative forcings (provided in W m−2), multiplied by individual scaling factors, to best fit the observed global mean surface temperature (GMST):

\[
\text{GMST}(t) = \left[ \sum_{j=1}^{4} \alpha_j F_j(t) \right] + \gamma E(t - \tau) - \beta. \tag{1}
\]

We consider four radiative forcings (\(F_j\), with scalings \(\alpha_j\)): greenhouse gases, other anthropogenic effects (mainly aerosols, land use, and ozone), solar, and volcanic activity. Annual means are used everywhere. We also use an ENSO index (\(E\), scaled by \(\gamma\)) as a “forcing” to remove the effects of the leading mode of interannual variability from the observations. This \(E\) index is defined as the linearly detrended Niño-3.4 anomaly from 1857 to 2015 (Kaplan et al. 1998) and zero before 1857, with a lag (\(\tau\)) of 4 months to maximize the variance explained (i.e., the annual mean is a September to August average). A similar approach to fitting global temperatures was taken by Lean and Rind (2009) and Suckling et al. (2016). All global temperature data are referenced to 1986–2005 to match the analysis in IPCC AR5 (Kirtman et al. 2013) and \(\beta\) is a constant offset to account for this reference period.

We perform the analysis separately for five global temperature datasets to represent the uncertainty in temperature reconstructions, although this is an underestimate of the true uncertainty because they are all based on similar observations. For HadCRUT4, BEST-GL, and CW14, the multiple linear regression is performed over the period 1850–2015. The NOAA GlobalTemp (Karl et al. 2015) and NASA Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) (Hansen et al. 2010) datasets are fitted over the full extent of their available data.
We use the HadCRUT4 uncertainties in the weighted regression (except for BEST-GL and NOAA GlobalTemp, which have their own uncertainty estimates), so that the older (and more uncertain) data have less weight.

Figure 2 (top) shows one estimate of GMST (HadCRUT4) and the scaled forcings for the full 1765–2015 period, using the regression parameters derived over 1850–2015. The correlation between the scaled forcings (including ENSO) and observed temperatures is 0.94 for each of the global datasets.

There are two ways to estimate a change in temperature using this approach. First, we can average the scaled forcings over 1765–1800 to produce an estimate of the preindustrial global temperature for each dataset with associated uncertainties, accounting for the covariance in derived $c$'s. Note that this is the longest period available using the CMIP5 forcings in the 1720–1800 period. The Paleoclimate Modeling Intercomparison Project (PMIP) protocol does not currently provide consistent forcing estimates in this way for the 850–1850 period (Schmidt et al. 2012). For the five temperature datasets, the best estimates are found to range from 0.64°C to 0.76°C with uncertainties of around ±0.05°C. Alternatively, the value of the regression constant ($\beta$) is an estimate of the temperature change from a state of zero forcing (in this case 1750) to 1986–2005. For the five temperature datasets, $\beta$ ranges from 0.69°C to 0.82°C (with uncertainties of ±0.02°C), which is around 0.06°C larger than using the 1765–1800 average. This difference is consistent with the small increase in greenhouse gas forcing and the relatively weak volcanic forcing after 1765. Overall, these results suggest that preindustrial was slightly cooler than the 1850–1900 period.

Also, the derived estimates for the warming are all larger than the value used in IPCC AR5 (0.61°C), with the HadCRUT4-based estimates being the smallest and GISTEMP the largest. The differences between estimates from the various datasets are larger than the stated uncertainties and are dominated by the uncertainty in global change since 1850, partly related to the way missing data are treated. The CW14 dataset, which interpolates between the gaps in HadCRUT4, finds slightly larger warming, consistent with Cowtan et al. (2015) who show a similar effect when examining simulated data to determine the effects of incomplete spatial coverage. The NOAA and GISTEMP datasets also use slightly different interpolation techniques. These various infilling approaches may reduce the bias from poor spatial sampling, especially for fast warming regions such as the Arctic, but may simply accentuate the dominant anomaly and add uncertainty. These inconsistencies merit further investigation elsewhere.

This approach does not account for nonlinearities in the temperature response to forcings, or uncertainties in the assumed CMIP5 forcing history itself, which are likely to be particularly large for aerosols

---

3 These estimates are largely insensitive to whether a lag is introduced in the greenhouse gas forcing [as done in Lean and Rind (2009)], or if only the 1900–2015 period is used for fitting, or if the anthropogenic forcings are combined before fitting.
(e.g., Carslaw et al. 2013; Stevens 2013) and ozone (Marenco et al. 1994). However, this approach does allow for varying sensitivities ($\alpha$) to the different assumed forcings (or “efficacies”) (Hansen et al. 2005; Shindell 2014). Another approach would be to use a simple energy balance model, tuned to the observational record (e.g., Osborn et al. 2006; Aldrin et al. 2012) and this could be examined in future work.

**APPROACH 2: USING LAST MILLENNIUM SIMULATIONS.** An alternative approach to considering the forcings alone is to use “last millennium” ensembles (LMEs), which use global climate models (GCMs) to simulate global climate from 850 to 2005 using the PMIP3 estimates of greenhouse gas concentrations, solar variations, and volcanic eruptions detailed by Schmidt et al. (2012). Here we consider three ensembles with different GCMs: Goddard Institute for Space Studies Model E2, coupled with the Russell ocean model (GISS-E2-R; 3 members; Schmidt et al. 2014), Community Earth System Model, version 1 (CESM1; 10 members; Otto-Bliesner et al. 2016), and Max Planck Institute Earth System Model (MPI-ESM; 3 members; Jungclaus et al. 2014). These are the only models to have made continuous simulations available for the whole time period using all radiative forcings\(^4\) and multiple ensemble members (Fig. 2, bottom).

In the GCM simulations, 1720–1800 is 0.00°–0.06°C cooler than 1850–1900 (using ensemble means), which is slightly smaller than the result using approach 1. However, the three GCMs produce very different estimates for the warming from 1720 to 1800 until 1986–2005 (0.51° ± 0.08°C for CESM1, 1.04° ± 0.07°C for GISS E2-R, and 0.91° ± 0.04°C for MPI-ESM).\(^5\) These differences are not what would be expected as a result of climate sensitivity alone as CESM1 has the largest transient climate response (TCR; 2.2 K) and GISS E2-R the smallest (1.5 K). It is more likely that the differences are due to a combination of several factors, including climate sensitivity, different amplitude responses to anthropogenic aerosols and volcanic eruptions (Stoffel et al. 2015), different assumed forcings (e.g., the size of the 1761 eruption), and different implementations of the forcings. In addition, the global temperature response to volcanic eruptions appears to be larger in the GCMs than the real world (e.g., Schurer et al. 2013), although Stoffel et al. (2015) suggest this effect is much reduced with an improved representation of the aerosol microphysics.

Given the diversity in global temperature response, a robust estimate of change in global temperature since preindustrial using these simulations should consider scaling the responses to the observations or using detection and attribution techniques on the range of simulations available (Schurer et al. 2013; Otto-Bliesner et al. 2016). In addition, the comparison with observations is not necessarily like-with-like given sparse observations and different use of air or sea temperatures (Cowtan et al. 2015; Richardson et al. 2016).

However, an additional use for the LMEs is to examine uncertainty in the estimate of preindustrial temperatures due to internal variability alone. This can be done by considering the spread of estimated change using the 10 CESM1 ensemble members ($\sigma = 0.05$ K), which suggests an uncertainty of around ±0.1°C. Note that this range is similar to the uncertainty ranges from long instrumental records discussed below. The other ensembles are too small to reliably estimate this range. We also use the CESM1 simulations to consider issues of differential seasonal warming in the appendix.

**APPROACH 3: USING LONG INSTRUMENTAL RECORDS.** The above two approaches have considered the response to estimated radiative forcings. An alternative approach to estimate GMST further back in time is to use direct observations from long instrumental records and calibrate them against each of the five global mean temperature datasets.

For example, central England temperature (HadCET, herein referred to as CET; Manley 1974; Parker et al. 1992) is available for 1659–present. CET covers just 0.005% of Earth’s surface but is highly correlated with GMST on multidecadal time scales (Sutton et al. 2015). Here, we utilize this correlation and scale GMST to CET:

$$\text{CET} = \delta \text{GMST} + \epsilon,$$

using the overlapping periods (1850–2015), and adopt the same parameters to scale CET back to 1659.

\(^4\) Note that the GISS-E2-R simulations used a different aerosol forcing over the historical period than the CMIP5 historical simulations performed with the same GCM. The PMIP3 simulations warm by about 0.3 K more than the CMIP5 simulations (not shown).

\(^5\) We also tested approach 1 using the global temperatures from the PMIP simulations. This produced compatible values for the warming (0.45° ± 0.09°, 1.09° ± 0.04°, and 0.90° ± 0.06°C, respectively), building confidence in that approach.
as an estimate of GMST (Fig. 3, top). When using HadCRUT4 as GMST, $\delta = 1.20 \pm 0.23$, although other global temperature datasets give lower values (e.g., for BEST-GL, $\delta = 1.06 \pm 0.21$). The major caveats to this approach are that we assume the historical temperature estimates are unbiased and that the relationship between GMST and CET is the same whatever forcing is dominant, neither of which may be true (Zanchettin et al. 2013; Haarsma et al. 2013; see appendix).

We take the mean of the scaled CET over two periods: i) 1765–1800 (for consistency with approach 1) and ii) 1720–1800 (the full period identified from the radiative forcing history). An additional issue that arises from scaling a local record to global temperatures is the possible regional effect of external forcing. In particular, the eruption of Laki (located in Iceland) in 1783 likely only had a small global effect, but it certainly influenced western Europe (Thordarson and Self 2003). Therefore the years 1783 and 1784 are removed from the averages owing to the eruption of Laki to avoid biasing the estimated temperature change. However, this does not change the results significantly.

These two periods produce consistent estimates for the warming to 1986–2005: 0.75° ± 0.10°C (for 1765–1800) and 0.64° ± 0.08°C (for 1720–1800) when using HadCRUT4 for GMST. The other global temperature datasets give larger values for the warming to 1986–2005, by up to 0.09°C (Fig. 3, top). The quoted uncertainty ranges account for the uncertainties in the regression parameters and assume the uncertainty in each CET annual mean from 1720 to 1800 is independent and equal to 0.2°C [based on Parker (2010)].

The difference between the two averaging periods is mainly because the 1720s and 1730s were unusually warm in the CET record. Internal climate variability and a recovery from the negative forcings of the previous decades are possible explanations, although this warmth was less pronounced in some other European instrumental records (e.g., Berlin) (Jones and Briffa 2006).

Figure 3 (bottom) repeats this analysis with the Berkeley global land temperature (BEST-Land; Rohde et al. 2013), which starts in 1753. A similar approach was adopted by Mann (2014). Using BEST-Land produces a consistent but slightly lower warming than derived with CET. Using the scaled temperatures over the 1753–1800 period, the estimates of the warming to 1986–2005 range from 0.62° ± 0.10°C for HadCRUT4 to 0.71° ± 0.12°C for GISTEMP.

It may seem surprising that the error bars are not smaller for the BEST-Land dataset than for CET. The regression uncertainty is indeed much larger for the local example; however, the error in representing the whole global land area with sparse data are larger than in representing central England with a small number of stations. These two sources of uncertainty combine to give similar overall ranges. Note that BEST-Land looks very similar to the long European records and the variability increases further back in time (also for CET), highlighting that fewer and fewer (mostly European) stations are used in the reconstruction.

We also consider a long temperature series from the Netherlands, referenced to De Bilt, which starts in 1706...
(Van Engelen and Nellestijn 1990) and a central Europe instrumental series from Dobrovolný et al. (2010) that starts in 1760, which are also both well correlated with GMST in the overlapping period. These results are summarized in Fig. 4, which shows that the central Europe series consistently produces slightly lower estimates of the warming than CET or BEST-Land.

**OVERALL ASSESSMENT.** We consider that approaches based on the radiative forcings and scaled instrumental observations currently produce more reliable estimates of the global temperature change since preindustrial than the last millennium GCM simulations. This weighting of methods could change in the future with additional evidence, analysis, and model development (see implications discussed below). Furthermore, the estimates using radiative forcings are generally larger than when using the observational datasets, as summarized in Fig. 4. Much of the uncertainty in the assessment derives from the range of global temperature change estimates available since 1850. For example, the uninterpolated HadCRUT4 dataset produces lower values than the other infilled records.

Our overall assessment is that the change in global average temperature from preindustrial to 1986–2005 is “likely” between 0.55° and 0.80°C.

This range reflects the authors’ aggregated assessment of the three approaches and contains virtually all of the best estimates using the various combinations of regional and global temperature datasets and scaled radiative forcing estimates. Note that there are potentially important uncertainties in each approach that we cannot quantify. As in IPCC AR5 we consider that likely refers to greater than 66% probability, although this is not a formal uncertainty quantification.

It is also helpful to assess a lower bound and we suggest that the warming from preindustrial until the 1986–2005 period is likely greater than 0.60°C, implying that the value used by IPCC AR5 for the warming since 1850–1900 (0.61°C) was probably smaller than the true change since preindustrial. Such differences matter more when considering the chance of crossing lower temperature levels such as 1.5°C than when considering higher values.

Using this lower bound, 2015 was the first year to be more than 1°C above preindustrial levels in each global temperature dataset (Fig. 5). The year 2016 was warmer than 2015, but future years could still be cooler than 2015 owing to internal variability, such as a La Niña event.

The available proxy-based evidence is consistent with our assessment, but currently too uncertain to make more precise estimates, partly because of different seasonal signals (see appendix). However, defining a preindustrial period offers a target for proxy reconstructions to aid future assessments.

**CONCLUSIONS AND IMPLICATIONS.** We have examined estimates of historical radiative forcings to determine which period might be most suitable to be termed preindustrial and used several approaches to estimate a change in global temperature since this preindustrial reference period. The main conclusions are as follows:

1) The 1720–1800 period is most suitable to be defined as preindustrial in physical terms, although we have incomplete information about the
radiative forcings and very few direct observations during this time. However, this definition offers a target period for future analysis and data collection to inform this issue.

2) The 1850–1900 period is a reasonable pragmatic surrogate for preindustrial global mean temperature. The available evidence suggests it was slightly warmer than 1720–1800 by around 0.05°C, but this is not statistically significant.

3) We assess a likely range of 0.55°–0.80°C for the change in global average temperature from preindustrial to 1986–2005.

4) We also consider a likely lower bound on warming from preindustrial to 1986–2005 of 0.60°C, implying that the AR5 estimate of warming was probably too small and that 2015 was the first year to be more than 1°C above preindustrial levels.

We have assumed in the motivation for this discussion and choice of reference periods that the UNFCCC agreements on temperature limits refer to anthropogenic increases only, but this is not explicitly stated. We have not attempted to attribute the observed increase in global temperatures (but see Schurer et al. 2013; Otto et al. 2015); nonanthropogenic factors (including internal variability) may have either offset or contributed to the warming. We have attempted to minimize issues of varying natural forcing and internal variability, but this effect cannot be removed entirely.

Our chosen preindustrial period likely has slightly weaker volcanic activity than a typical period and the modern reference period (1986–2005) includes the large Pinatubo eruption. These effects would bias our estimated change in temperature to be slightly too low, highlighting the value of assessing a lower bound in the warming since preindustrial. We also note that future climate projections do not usually include volcanic eruptions, so choosing a relatively weak volcanic baseline is perhaps appropriate. The recent period has a slightly positive PDO index that would act as a small positive bias for some of our estimates, but this modern reference period will likely be updated for the next IPCC assessment.

There are a number of ways that this assessment could be improved. Better understanding of historical radiative forcings, particularly of volcanic eruptions, solar activity, and anthropogenic aerosols, would help narrow the uncertainties in past global and regional temperature change. We did not include the estimates for preindustrial temperature from the last millennium simulations in this assessment because of the diverse derived values, which are due to differences in both the forcings used and climate sensitivity (Fernández-Donado et al. 2013). Future work might consider scaling the simulations (Schurer et al. 2013) or use of simple energy balance models (EBMs).

However, we may not necessarily expect simulated and observed values to agree, even in the case of perfect knowledge of radiative forcings and climate sensitivity. This is because the global observations are a sparse blend of sea surface temperatures over the ocean and air temperatures over the land, whereas virtually all analyses of GCM simulations use air temperatures with complete global coverage. Cowtan et al. (2015) and Richardson et al. (2016) used GCM simulations to suggest that if we had complete coverage of air temperature, the observed change from 1850 to present would be 24% ± 15% larger than currently estimated in HadCRUT4. The use of infilled temperature datasets only partly overcomes this issue.

This creates a dilemma—are the temperature limits adopted by the UNFCCC designed to use
observationally based estimates of global temperature change (as generally used here) or on what those observations mean for a “true” global mean air temperature change (as used in most climate impact assessments)? The available evidence suggests that the latter is larger. If such findings are borne out by further research, and if the true change is what is desired by UNFCCC, then our assessed temperature change since preindustrial is too small and should probably be increased by 0.05°–0.10°C.

It is possible to obtain additional data for the historical period. Recovery of additional instrumental observations of temperature and sea level pressure from undigitized handwritten logbooks from ships and in currently data-sparse regions could significantly aid similar future assessments. Some such efforts are ongoing (e.g., the Atmospheric Circulation Reconstructions over the Earth (ACRE) and OldWeather.org initiatives; Allan et al. 2011) but these could be expanded. The available observations can also be combined with data assimilation techniques to allow longer atmospheric reanalyses to be produced (Widmann et al. 2010; Compo et al. 2011; Matsikaris et al. 2016; Brohan et al. 2016). Additional seasonal proxy information would be of great value for informing this discussion, especially for winter (see appendix) and for the tropics and Southern Hemisphere (e.g., Jones et al. 2016), although the temporal resolution and continuity of proxies into the modern period is also a potential issue. Also note that a suitable preindustrial period may be different for other climate variables (e.g., sea level) or for carbon cycle considerations.

Two specific recommendations for future GCM-based analyses and simulations are i) to use blended observation-like estimates of global mean temperature when comparing observations and simulations and ii) use 1750 forcings to perform preindustrial control simulations and to start historical transient simulations, rather than 1850. Adopting these recommendations would allow an ensemble of transient historical simulations to better quantify the role of natural variability and the impacts of the total radiative forcing changes since the preindustrial period, especially the potentially long-term impact of the large volcanic eruptions in the early 1800s (Raible et al. 2016). We recognize, however, that this increases the computational demand in producing historical simulations. In addition, increased usage of tracers (e.g., water stable isotopes) and proxy models within GCMs would allow more direct comparisons between simulations and proxy observations, including GCM simulations nudged to atmospheric reanalyses (e.g., Jouzel et al. 2000; LeGrande and Schmidt 2009; Evans et al. 2013).

Finally, these findings have a number of implications for policy-relevant issues. For example, the date at which future temperature thresholds are expected to be crossed may be shifted slightly earlier than estimated in IPCC AR5 (see Joshi et al. 2011; Kirtman et al. 2013; Hawkins and Sutton 2016). In addition, the cumulative emissions allowed to avoid reaching a particular temperature threshold (Meinshausen et al. 2009; Allen et al. 2009) may need to be reassessed, although any difference would likely be well within the current uncertainty ranges. Moving the baseline may also affect how historical responsibility for emissions needs to be accounted for (Knutti et al. 2015).

More specifically, given the uncertainty in the global mean temperature change since preindustrial, the UNFCCC might consider alternative equivalent baselines and limits to global temperature change. For example, “well below 2°C above pre-industrial” (p. 3) might be translated to “well below X°C above 1986–2005.” Using a recent baseline is possibly more relevant for defining some impacts of climatic changes, with the value of X (and choice of baseline period) being decided by the UNFCCC. Given the uncertainty in defining the temperature change since preindustrial, such a framing would allow a more precise assessment of when such levels might be reached in the future, given our much improved recent observational coverage and availability of atmospheric reanalyses for the modern period (e.g., Dee et al. 2011; Simmons et al. 2016). It would also remove the need to precisely assess inherently uncertain changes since the preindustrial period.

ACKNOWLEDGMENTS. We thank John Fasullo and Johann Jungclaus for providing the CESM1 LME and MPI-ESM data, respectively. EH is funded by the U.K. National Centre for Atmospheric Science (NCAS) and a Natural Environment Research Council (NERC) fellowship (Grant NE/I020792/1). TJO and GH were supported by the NERC (Grant NE/N006348/1, SMURPHS). ES was supported by the European Union Seventh Framework Programme (FP7/2007–2013) under the SPECS project (Grant Agreement 308378) and by the UK-China Research and Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund. AS and GH are supported by the ERC funded project TITAN (EC-320691). GH was further supported by NCAS and the Wolfson Foundation and the Royal Society as a Royal Society Wolfson Research Merit Award (WM130060) holder. VMD acknowledges support from Agence Nationale de la Recherche, project ANR-15-
APPENDIX: COMPARISON WITH PROXY RECONSTRUCTIONS. There are numerous efforts to reconstruct past climate using different proxies and archives that could be used to aid an assessment of change since the preindustrial period. For temperature, these include ice cores, glaciers, tree rings, pollen, corals, and sediment cores.

For example, Leclercq and Oerlemans (2012) suggest a global land warming of $0.94\pm0.31^\circ$C between 1830 and 2000 using glacier reconstructions, although the mid-1700s is around $0.25^\circ$C warmer than 1830 in their estimates. Holloway and Smelhod (2004) suggest that global land temperatures in the mid-1700s were around $0.65^\circ$–$0.90^\circ$C below the year 2000 using borehole proxies. Mann et al. (2008) perform a multiproxy analysis and report that global average temperature was around $0.3^\circ$C below 1961–90 in the mid-1700s, with large uncertainties. This is equivalent to around $0.6^\circ$C below 1986–2005, consistent with the recent PAGES2k global reconstruction (PAGES 2k Consortium 2013) and this study.

Overall, these proxy reconstruction estimates for preindustrial temperature are consistent with the approaches adopted above, but the uncertainties are currently too large to make more precise statements. Defining a preindustrial period (1720–1800) will hopefully provide a target for future reconstructions using the proxy data available. Certain long proxy series could also be used in approach 3. However, it is important that such efforts focus on all seasons, as we discuss next.

Seasonal effects in proxies, observations, and simulations. There are likely some seasonal differences in the rates of temperature change that are important to consider (e.g., Hegerl et al. 2011; Jones et al. 2014). For example, different proxies are sensitive to climate in certain seasons. In general, summer is more widely represented because many proxies rely on biological activity, which tends to occur in the extended summer season. This is a potential issue for using proxies to reconstruct past temperatures if winter and summer change at different rates (Jones et al. 2003). In that case, the different seasonal proxies may not agree and/or produce biased estimates of an annual average. Some reconstructions (e.g., Van Engelen et al. 2001; Luterbacher et al. 2004; Vinther et al. 2010) for Holland, Europe, and Greenland, respectively, do show seasonal warming differences. However, the restricted availability of winter proxies limits the scope of such a comparison.

To investigate how representative of annual-mean changes the seasonal data are, we repeated the instrumental analysis (approach 3) using extended seasons (April to September and October to March) for the regional data, while retaining the annual global data as the reference. Figure A1a shows how the derived warming since the 1753–1800 period depends on the choice of season for the instrumental series—the extended winter season warms much faster than the extended summer season.

However, if this seasonal difference in the rate of change over Europe was constant with time it should be scaled out. This suggests that there is i) a seasonal bias in the observed temperatures in certain periods (e.g., before standardized measurements) and/or ii) a different seasonal response to different radiative forcings.

For example, there is evidence that some historical observations may be biased, especially in summer, where warm biases due to nonoptimal ob-

**Fig. A1.** Seasonal differences in warming rates. (a) Derived scaled warming from 1753–1800 to 1986–2005 (using approach 3) for annual means (black) and for the extended seasons (Apr to Sep—AMJJAS, red; and Oct to Mar—ONDJFM, blue) for the different regional time series, all using annual-mean HadCRUT4 as the reference dataset. (b) Seasonal warming derived from the CESM1 LME simulations for the global mean (crosses, with black lines linking the same ensemble members in each season) and for the ensemble mean of simulated CET (circles).
observation techniques in the past have been identified (Parker 1994; Böhm et al. 2010; Jones 2016), which fits the pattern seen in Fig. 1a. Dobrovolný et al. (2010) note that their documentary temperature data agree best with their instrumental data during winter, adding credence to this hypothesis. In addition, the cooling due to tropospheric aerosols in the twentieth century may be seasonally dependent (Hunter et al. 1993; Krishnan and Ramanathan 2002), there is a trend in westerly wind characteristics in winter (Haarsma et al. 2013), and many of the observations are located in the northern extratropics and therefore influenced by Arctic amplification, which is observed and simulated to be larger in winter than in summer (Serreze et al. 2009; Pithan and Mauritsen 2014).

We can also examine whether this seasonal warming difference is present in the last millennium model simulations. Figure 1b highlights that the CESM1 LME simulations do not show a strong global mean warming seasonal difference since the preindustrial period and only a very small seasonal effect when considering the central England location. The complex nature of these different seasonal features merits further analysis in a range of observations and simulations.

REFERENCES


IISD, 2013: Earth Negotiations Bulletin: Summary of the 12th session of Working Group I of the Intergovernmental Panel on Climate Change (IPCC) and thirty-sixth session of the IPCC. [Available online at www.iisd.ca/vol12/enb12581e.html.]


American Meteorological Society


http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf]