



Comment on "The Spatial Extent of 20th-Century Warmth in the Context of the Past 1200 Years"

Gerd Bürger, *et al.*
Science **316**, 1844a (2007);
DOI: 10.1126/science.1140982

The following resources related to this article are available online at www.sciencemag.org (this information is current as of June 29, 2007):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/316/5833/1844a>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/316/5833/1844a/DC1>

This article **cites 5 articles**, 1 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/316/5833/1844a#otherarticles>

This article appears in the following **subject collections**:

Atmospheric Science

<http://www.sciencemag.org/cgi/collection/atmos>

Technical Comments

http://www.sciencemag.org/cgi/collection/tech_comment

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

Comment on “The Spatial Extent of 20th-Century Warmth in the Context of the Past 1200 Years”

Gerd Bürger

Osborn and Briffa (Reports, 10 February 2006, p. 841) identified anomalous periods of warmth or cold in the Northern Hemisphere that were synchronous across 14 temperature-sensitive proxies. However, their finding that the spatial extent of 20th-century warming is exceptional ignores the effect of proxy screening on the corresponding significance levels. After appropriate correction, the significance of the 20th-century warming anomaly disappears.

Osborn and Briffa (1) determined exceedance counts from 14 smoothed and normalized Northern Hemisphere temperature proxies from 800 to 1995 and concluded that “the 20th century is the most anomalous interval in the entire analysis period, with highly significant occurrences of positive anomalies and positive extremes in the proxy records.” The proxies were pooled from three previous studies (2–4) and screened for positive correlations to the associated grid cell temperature of the Climatic Research Unit’s CRUTEM2v data set (5). Significance levels of the exceedance counts were determined by resampling. However, given the large number of candidate proxies and the relatively short temporal overlap with instrumental temperature records, statistical testing of the reported correlations is mandatory. Moreover, the reported anomalous warmth of the 20th century is at least partly based on a circularity of the method, and similar results could be obtained for any proxies, even random-based proxies. This is not reflected in the reported significance levels.

Because of a limited theoretical understanding of proxy temperature relations, especially on longer time scales, temperature-sensitive proxies suitable for climate reconstructions are largely determined empirically. If not already sampled in previous studies, they are culled from a wealth of candidate proxies, normally by inspecting correlations to the local (or global) instrumental temperature record. Osborn and Briffa (1) pooled proxy records from (2–4), selecting only those that showed positive correlations with their local temperature observations (nine proxies based on annual correlations, and another five based on decadal correlations). This method of selecting proxies by screening a potentially large number of candidates for positive correlations runs the danger of choosing a proxy by chance. This is aggravated if the time series show persistence, which reduces the degrees of freedom for calculating correlations

(6) and, accordingly, enhances random fluctuations of the estimates. Persistence, in the form of strong trends, is seen in almost all temperature and many proxy time series of the instrumental period. Therefore, there is a considerable likelihood of a type I error, that is, of incorrectly accepting a proxy as being temperature sensitive. Osborn and Briffa’s exceedance fraction is composed of 14 independent selections of that kind. Hence, the significance of that signal’s being a “clean” temperature proxy—one that is composed solely of temperature-sensitive proxies—is strongly reduced (SOM text) (7). This effect can only be avoided, or at least

mitigated, if the proxies undergo stringent significance testing before selection. Osborn and Briffa did not apply such criteria.

A second and potentially more serious problem in the Osborn and Briffa analysis (1) is also related to the selection process. The red and blue shadings in figure 3 in (1) indicate years when the differences in exceedance fractions were anomalously high or low, exceeding the 5th or 95th percentile band, respectively. That band reflects the likelihood of obtaining such differences by pure chance, that is, from independent time series whose values follow the standard normal distribution (so that the band obviously depends only on the number of available proxies for each year). The 20th century is well above even the 99% percentile and appears unique in that respect for the entire 1200-year period examined. However, this approach does not account for the effects of screening. The majority of those random series would not even have been considered, having failed the initial screening for positive temperature correlations. Taking this effect into account, the independence of the series shrinks for the instrumental period. Moreover, the series are synchronized because, for longer time scales, local temperatures themselves tend to be positively correlated, as evidenced by figure 3D in (1). Thus, for the instrumental period, any set of selected proxies evolves coherently on longer time scales, leading to enhanced exceedance counts. The results described by Osborn and

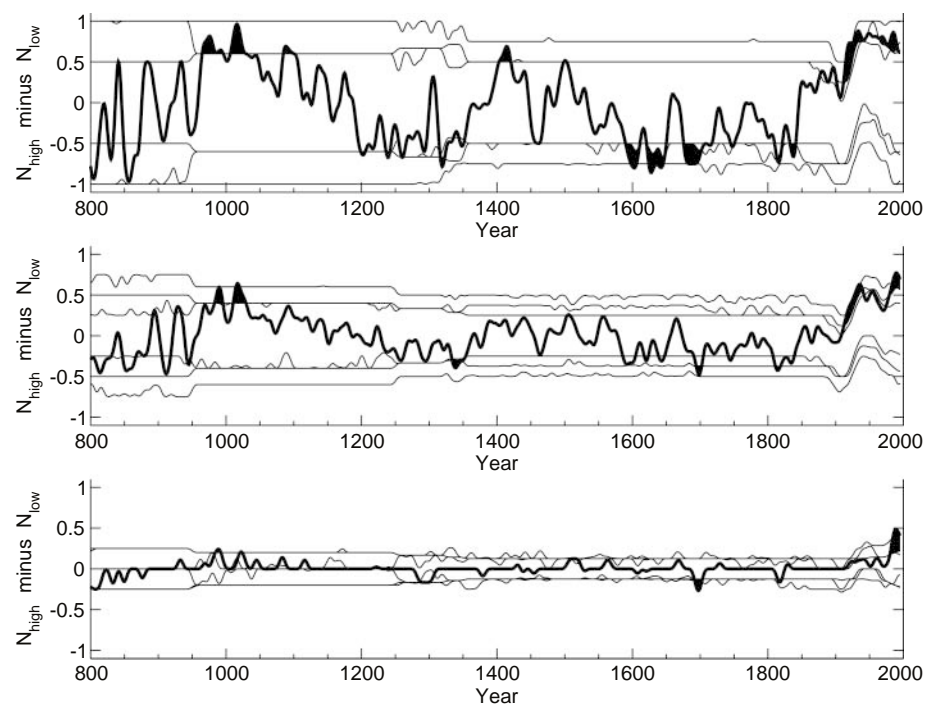


Fig. 1. Difference between the fraction of records that have smoothed and normalized proxy anomalies >0 and <0, >1 and <-1, and >2 and <-2, from top to bottom, using only proxies that are significantly temperature sensitive (bold), along with the (10th, 90th), (5th, 95th), and (1st, 99th) percentile bands (thin) from 1000 random-based series exceeding the critical correlation. The area of the signal outside the (5th, 95th) percentile band is filled black.

Freie Universität-Berlin, Institut für Meteorologie, Carl-Heinrich-Becker-Weg 6-10, D-12165 Berlin, Germany. E-mail: gerd.buerger@met.fu-berlin.de

Briffa are therefore at least partly an effect of the screening, and the significance levels depicted in figure 3 in (I) have to be adjusted accordingly.

I repeated the analysis in (I) with the appropriate adjustments. I first determined which of the proxies was significantly temperature sensitive by applying a *t* test with a 1% significance level (SOM text). Using this criterion, only 8 of the original 14 proxies used in (I) proved temperature sensitive (table S1). After normalizing and smoothing as in (I), for these eight proxies I calculated, for a given year and threshold θ , the difference between the fraction of anomalies $>\theta$ and $<-\theta$, resulting in a new 1200-year time series of such differences. I repeated this procedure 1000 times, with random series in place for each of the eight proxies. The series are fractionally integrated white noise, with memory parameters estimated from the corresponding proxy and with local correlations that exceed the 1% significance level (7). For each year, this gives a distribution of 1000

differences of exceedance counts, along with corresponding percentiles.

As in figure 3 in (I), Fig. 1 shows, for exceedance thresholds of $\theta = 0, 1,$ and $2,$ the differences between corresponding fractions that result from the eight significant temperature-sensitive proxies. Also shown are the (1st, 99th), (5th, 95th), and (10th, 90th) percentile bands of the random distributions. They vary only slowly (with some superimposed sampling noise) in the preinstrumental period, reflecting the changing availability of the eight proxies through the years, as in (I). In the instrumental period, the described synchronizing of signal and significance levels is clearly visible. Only after about 1970 does the signal rise again without a comparable level response. However, that period is very uncertain because it is based on only three to five proxies. As a result, the “highly significant” occurrences of positive anomalies during the 20th century disappear. The 99th percentile is almost never exceeded, except for the very last years

for $\theta = 1, 2.$ The 95th percentile is exceeded mostly in the early 20th century, but also about the year 1000.

References and Notes

1. T. J. Osborn, K. R. Briffa, *Science* **311**, 841 (2006).
2. J. Esper, E. R. Cook, F. H. Schweingruber, *Science* **295**, 2250 (2002).
3. M. Mann *et al.*, *Eos* **84**, 256 (2003).
4. M. E. Mann, P. D. Jones, *Geophys. Res. Lett.* **30**, 1 (2003).
5. P. D. Jones, A. Moberg, *J. Clim.* **16**, 206 (2003).
6. R. Livezey, in *Analysis of Climate Variability: Applications of Statistical Techniques*, H. von Storch, A. Navarra, Eds. (Springer Verlag, Berlin, 1995), pp. 159–176.
7. Materials and methods are available as supporting material on *Science* Online.

Supporting Online Material

www.sciencemag.org/cgi/content/full/316/5833/1844a/DC1

Materials and Methods

SOM Text

Fig. S1

Tables S1 and S2

References

7 February 2007; accepted 30 May 2007

10.1126/science.1140982