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

Supporting Online Material for

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

Gerd Bürger

*To whom correspondence should be addressed. E-mail: gerd.buerger@met.fu-berlin.de

Published 29 June 2007, *Science* **316**, 1844a (2007)
DOI: 10.1126/science.1140982

This PDF file includes:

Materials and Methods
SOM Text
Fig. S1
Tables S1 and S2
References

Materials and methods

Proxy selection. - Table S1 shows the correlations and corresponding critical levels, using 1% significance, of a t-test between each of the 14 proxies and the associated CRUTEM2v grid cell temperature, on an annual and decadal basis. The decadal series are obtained by smoothing, following (1) (we assume here that (2) use a similar method). All series are in rough agreement to the correlations reported (or calculated) by Osborn and Briffa (OB); remaining deviances have no effect on the selection. The critical correlation levels and degrees of freedom are obtained by taking persistence into account (3).

Monte Carlo procedure. - For each selected proxy we independently generated 1000 random series by fractionally integrating white noise. The memory parameters, ranging between 0.1 and 0.7, were estimated from the full proxy series using the local Whittle estimator (4). In this process, a series was rejected until its correlation to local temperature exceeded the critical 1% significance level (screening).

If a proxy is considered a priori genuine (representing temperature) the procedure should be modified by re-replacing the corresponding random series with 1000 copies of the original proxy. In this case, however, the fluctuations of the proxy are imprinted on the significance levels and render them somewhat noisy. This was done for proxy #7, which is a documentary record, but the main conclusions remained unaltered (not shown).

Supporting text

Significance and power of testing temperature sensitivity. - Suppose that after testing for positive correlations to local temperature one believes to have found a temperature sensitive proxy. Then the likelihood of a type I error (the proxy was in fact temperature insensitive) is given by the significance level α of the test, e. g. α=0.05 (=5%). But if the proxy was selected only after screening a set of *n* candidate proxies, i. e. after performing *n* single tests, that likelihood is increased to $\alpha' = 1 - (1-\alpha)^n$. And if a series is composed of *k* such proxies found at different locations the corresponding likelihood of a type I error is $\alpha' = 1 - (1-\alpha)^{nk}$. This new significance level now strongly depends on the initial α, as illustrated by an example with *k=5* and *n=2*: Testing simply for positive correlations as OB do, that is, with α=0.5, gives a new α'=0.99; testing instead with α=0.05 gives α'=0.4, and for α=0.01 one obtains α'=0.1. Hence, to maintain at least some significance of the composed series the components themselves have to be *highly* significant, at least at the 1% level (α=0.01) in settings comparable to the example above. With less significance such as α=0.05 or even α=0.5, as in OB, practically no significance is left. (It is of course false to conclude that the resulting composed series is temperature *insensitive* for that case.)

Choosing α too small, on the other hand, entails the danger to miss some of the actually temperature sensitive proxies (type II error). This relates to the *power* of the test, and depends on the general ability to distinguish temperature sensitive proxies via local correlations. While not much is known about that power, it is likely not very large due to the limited sample size, and choosing α always represents a difficult compromise between significance and power. To illustrate how a relaxed proxy selection affects the final result (and defying the above warning) another set of Monte Carlo experiments was conducted, using an initial significance level of α=5%. The result is shown in Table S2 and Fig. S1. Now 11 proxies pass the test, and their composed signal (differences of exceedance counts) is above the 95% or 99% percentile slightly more often than in the former experiment, mainly in the early 20th century. (Note that these percentiles are unrelated to the significance levels of the proxy selection). The 20th century, however, remains far from being as “highly significant” as it was in OB.

Supporting figures

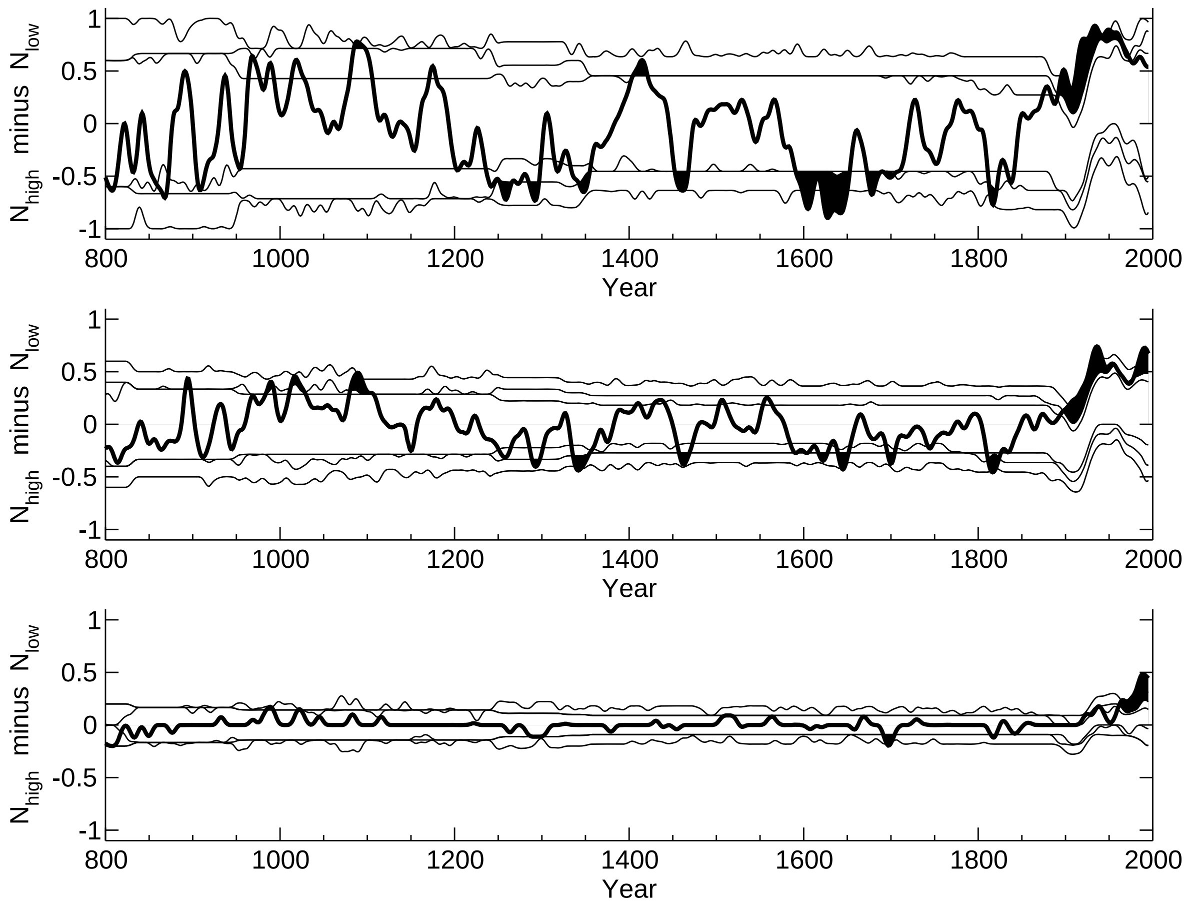

Fig. S1. As Fig. 1, but with proxies selected at the 5% significance level.

Supporting tables

Table S1. Annual (subscript a) and decadal (subscript d) correlations ρ between proxies and associated CRUTEM2v grid cell, and corresponding α=1% significance levels ($\tilde{\rho}$) and degrees of freedom (df) of the t-test. Insignificant values are red, indicating that the proxies were rejected for this study. The proxies from (2) are shaded.

proxy	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
ρ_a	0.13	0.67	0.18	0.38	0.16	0.56	0.67	0.45	0.47	0.38	0.24	0.41	0.28	0.16
$\tilde{\rho}_a$	0.30	0.29	0.30	0.27	0.28	0.32	0.33	0.31	0.31	0.27	0.27	0.27	0.48	0.62
df_a	57.0	62.0	58.0	74.8	66.8	49.3	48.6	54.7	52.7	73.1	70.8	72.6	21.1	11.6
ρ_d	0.65	0.60	0.57	0.12	0.23	0.75	0.71	-0.06	0.69	0.45	-0.07	0.57	0.52	0.22
$\tilde{\rho}_d$	0.72	0.71	0.70	0.64	0.56	0.70	0.56	0.55	0.77	0.57	0.62	0.63	0.77	0.82
df_d	7.8	8.3	8.4	10.8	14.8	8.5	14.6	15.8	6.6	14.4	11.9	11.0	6.5	5.3

Table S2. As Table S1, for α=5%.

proxy	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
ρ_a	0.13	0.67	0.18	0.38	0.16	0.56	0.67	0.45	0.47	0.38	0.24	0.41	0.28	0.16
$\tilde{\rho}_a$	0.22	0.21	0.21	0.19	0.20	0.23	0.23	0.22	0.22	0.19	0.19	0.19	0.35	0.47
df_a	57.0	62.0	58.0	74.8	66.8	49.3	48.6	54.7	52.7	73.1	70.8	72.6	21.1	11.6
ρ_d	0.65	0.60	0.57	0.12	0.23	0.75	0.71	-0.06	0.69	0.45	-0.07	0.57	0.52	0.22
$\tilde{\rho}_d$	0.56	0.54	0.54	0.48	0.41	0.54	0.42	0.40	0.60	0.42	0.46	0.48	0.60	0.66
df_d	7.8	8.3	8.4	10.8	14.8	8.5	14.6	15.8	6.6	14.4	11.9	11.0	6.5	5.3

Supporting references

S1. P. D. Jones, M. E. Mann, *Rev. Geoph.*, **42**, pp. RG2002 (2004).

S2. M. E. Mann, P. D. Jones, *Geophysical Research Letters*, **30**, pp. 1 (2003).

S3. R. Livezey, in *Analysis Of Climate Variability: Applications Of Statistical Techniques*, von Storch, H., Navarra, A. (Ed.) (Springer Verlag, 1995). pp. 159-176.

S4. P. M. Robinson, *Annals of Statistics*, **23**, pp. 1630-1661 (1995).