

Interactive comment on “Millennial temperature reconstruction intercomparison and evaluation” by M. N. Juckes et al.

S. McIntyre

stephen.mcintyre@utoronto.ca

Received and published: 19 December 2006

Section 2 of Juckes et al is less comprehensive than and adds nothing to the corresponding review of the NRC Panel.

Section 4 presents a reconstruction (the Union Reconstruction) whose proxies differ little from those in other recent literature; the statistical analysis of the reconstruction is very deficient, with the reconstruction even failing an elementary statistical significance test recommended by the NRC Panel.

Wegman et al 2006 criticized the overlap of proxies in supposedly “independent” studies. Despite this criticism published prior to the submission of Juckes et al, the “Union” reconstruction uses virtually the same collection of proxies as Osborn and Briffa 2006

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

and Hegerl et al 2006. Each consists of small collections (12-18 series). However, all three studies use two or more bristlecone/foxtail series, Tornetrask (twice in Juckes et al), Yamal, Taimyr, the Yang Composite and Fisher's West Greenland. See <http://www.climateaudit.org/?p=967>.

This repetitive use of the same proxies compromises any claim of “independence” between studies - a problem also noted by the NRC Panel. Because of this repetitive use of the same data, important premises of significance testing are violated, an issue discussed in economics literature. For example, Greene [Journal of Economic Methodology 2000] observed that standard distributions cannot be used with re-cycled data:

"Because the existing data series is no longer free of pre-testing or specification search and so cannot yield test statistics with known distributions. An attempt to re-use the original data implies the actual distribution of any test statistic differs from standard distributions in an unknown manner."

In addition to this problem, there is inadequate testing against the possibility of “spurious” or “nonsense” correlations between unrelated series [Yule 1926; Granger and Newbold 1974; Hendry 1980; Phillips 1986 and a large economics literature]. Yule's classic example of spurious correlation was between alcoholism and Church of England marriages. Hendry showed a spurious correlation between rainfall and inflation. The simulations performed in Juckes et al have virtually no “power” (in the statistical sense) as a test against possible spurious correlation between the Union reconstruction and temperature. For this purpose, a common, and not especially demanding, test is the Durbin-Watson test [Granger and Newbold 1974], whose use was encouraged by the NRC Panel (p. 87). According to my calculations, the Union Reconstruction failed even this test, contradicting the claims of Juckes et al to “99.98% significance”. (See <http://www.climateaudit.org/?p=945>).

Calibration-verification is a standard methodology in multiproxy studies and was rec-

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

ommended by the NRC Panel (88ff). In MM2005a-b, we observed that the 15th century MBH reconstruction failed the verification r2 test (that was said in MBH98 to have been considered) and, in MM2005b, we criticized the failure of Mann et al to report these adverse verification r2 results. Our finding of negligible verification r2 (and CE) values was confirmed by Wahl and Ammann. These findings were specifically noted by the NRC panel, in their decision to withdraw confidence intervals from the early portion of the MBH reconstruction. Juckes et al conspicuously did not reported calibration-verification results. My calculations indicate an extremely low verification r2 (or CE) values for the Union reconstruction. Verification r2 and CE results for the Union reconstruction should be reported; if the reconstruction fails verification r2 or CE tests, the authors should attempt to account for the failure if they can.

Reconstructions that are slightly varied from the Juckes reconstruction (but with different medieval-modern relationships) are also “99.98% significant” by the criterion of Juckes et al. Obviously the two different reconstructions cannot both be “99.98% significant” - evidence that neither reconstruction is "99.98% significant". See <http://www.climateaudit.org/?p=903>

Juckes et al failed to provide any statistical references for the results in their Appendix 1, nor any proof of the claimed optimality (or a reference of the fact). They assert a noise model, but do not show that they carried out any tests to demonstrate that the noise model in Appendix 1 was applicable to the actual proxy network. Inspection of the residuals in the individual series strongly indicates that the noise model of their Appendix 1 is not valid - see <http://www.climateaudit.org/?p=938>

I was able to replicate some of Juckes’ CVM calculations, but not all of them. In the Union reconstruction, there is an unreported flipping of the Chesapeake Mg-Ca series, the procedure for which is not described. The mbhpc reconstruction appears not to have carried out a flipping of PC series said to have been carried out. MBH99 said that bristlecones should be corrected for CO2 fertilization. We disagree that MBH99 carried out a relevant correction, but Juckes et al appear to have use PC series without any

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

effort whatever to apply such a correction - see <http://www.climateaudit.org/?p=930>

Juckes et al have put source code online (good), but the source code contains virtually no relevant comments and seems to be a grudging accommodation, rather than an earnest effort to illuminate methodology for subsequent readers.

Juckes' SI Figure 1 used rms normalization without any disclosure or explicit justification. Rms normalization is not used elsewhere in the study or, to my knowledge, in the relevant paleoclimate literature. It has the effect of minimizing the difference between MBH and other PC studies. I see no purpose whatever in permitting its use in this figure - especially without any disclosure of the methodology. See <http://www.climateaudit.org/?p=897>

I have tested some of Juckes' CVM reconstruction, finding that trivial variations can yield different medieval-modern relations e.g. Esper CVM without foxtails; <http://www.climateaudit.org/?p=885> ; Moberg CVM using Sargasso Sea SST instead of Arabian Sea G Bulloides wind speed and Polar Urals update instead of Yamal - see <http://www.climateaudit.org/?p=903> and <http://www.climateaudit.org/?p=887> Juckes' justification for not using Sargasso Sea SST is not convincing <http://www.climateaudit.org/?p=898> , nor is the exclusion of the Indigirka River series of Moberg et al 2005, which is an extension of the Yakutia series used in MBH98 - see <http://www.climateaudit.org/?p=901>

Juckes et al Table 1 contains numerous geographical mislocations. Table 1 shows lists the Tornetrask site 4 times under different alter egos, using 3 different coordinates, none of which are correct. The two "independent" foxtail sites are only about 30 km apart (the coordinates being inaccurately reported in Juckes et al.) The Union reconstruction used two different versions of the Tornetrask site (which are obviously not "independent") and neither justified this duplicate use nor the similar duplication of foxtail and bristlecone sites.

Juckes failed to evaluate the validity of individual Union proxies in light of criticisms by

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

the NRC panel and others. The use of percentage *G. Bulloides* as a temperature proxy was criticized by David Black, author of a *G. Bulloides* series from Cariaco. Without addressing such criticisms, Juckes et al used a percentage *G. Bulloides* series from the Arabian Sea in the Union reconstruction - see <http://www.climateaudit.org/?p=957>. The NRC panel specifically said that strip-bark bristlecones and foxtails should be “avoided” in temperature reconstructions. Without addressing this criticism, out of only 18 proxy series in the Union reconstruction, Juckes et al used no fewer than 4 bristlecone and foxtail series from one gridcell.

There has been extensive discussion of various aspects of Juckes et al at www.climateaudit.org - see <http://www.climateaudit.org/?cat=36>.

Interactive comment on Clim. Past Discuss., 2, 1001, 2006.

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)