How do we “know” that 1998 was the warmest year of the millennium?

Stephen McIntyre

Presentation at
Ohio State University
May 16, 2008

Thank you for the invitation to speak here. I've spent most of my life in business, mostly on the stock market side of mining exploration deals - not the most obvious introduction to climate science, but one which is perhaps more relevant than people may think.

In a prep for a radio interview coming here, the radio host commented that it was impossible for members of the public to personally investigate the science and thus, at some point, it was necessary to simply have faith in the scientists. But something similar could be said in all walks of life, where the need for faith is tempered by external due diligence. If you’re offering securities to the public, there are complicated and expensive processes of due diligence, involving audits of financial statements, independent engineering reports, opinions from securities lawyers and so on. There are laws requiring the disclosure of adverse results. These precautions obviously don’t eliminate financial misconduct, but they are serious attempts to protect the public and make markets work more effectively.

There is far more *independent* due diligence on the smallest prospectus offering securities to the public than on a Nature article that might end up having a tremendous impact on policy. At this time, I am not saying that journal peer review processes should be overhauled, only that policy-makers should bear in mind that journal peer review is a very limited form of due diligence. Under any circumstances, radically improving requirements for the archiving of data and methods would be a simple and cost-efficient measure for improving quality control and I urge this policy whenever I get a chance.

Today I will only discuss one particular aspect of the debate – 1000-year temperature reconstructions. I don't claim that the results here invalidate all of climate science or that policy decisions should be deferred because of these problems. On the other hand, the exigencies of policy should not prevent proper consideration of individual smaller issues, even if these ultimately prove only of academic interest. Good coaches and good teams look after the details.

Some people have argued that if the Hockey Stick is not correct, then the situation is worse than we thought. My reaction is: well, then I shouldn't be the only one examining the validity of these reconstructions.

Overview

As a quick overview, I'll introduce you to the infamous Hockey Stick, then its more recent incarnation in what we can call spaghetti graphs. I'll show that these supposedly "independent" studies are nothing of the sort, but rely on the re-cycling of a very small number of stereotyped series to achieve the Hockey Stick effect.

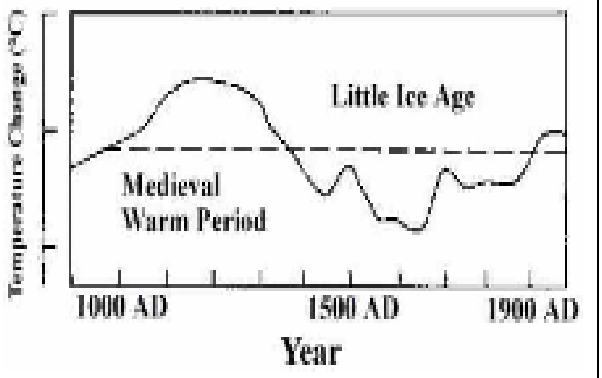
I'll discuss several of these key proxies, identifying problems with each one. Finally, I'll briefly discuss whether another view of the matter can be rationally held.

The Hockey Stick

I'm pretty sure that the first time I ever thought about climate change was in late 2002 when the Canadian Government was promoting acceptance of the Kyoto Protocol. The slogan for their campaign was that the 20th century was the warmest century, the 1990s the warmest decade and 1998 the warmest year in the past millennium - a slogan that got repeated in speech after speech and presentation after presentation.

> *The past decade was the world's warmest decade of the century. And that century was the warmest of the past millennium. Without action, the long-term consequences will be devastating. – David Anderson, Minister of the Environment (Canada) Oct. 27, 2001*
>
> *The 20th century was the warmest in the Northern Hemisphere in the past 1000 years. The 1990s was the warmest decade on record and 1998 was the warmest year - in Canada and internationally. - David Anderson, April 5, 2002.*
>
> *The 20th century was the warmest in the Northern Hemisphere for the past 1000 years and the 1990s the warmest decade on record... The science of climate change has been subjected to international scrutiny, open to all qualified experts, peer review, atmospheric modeling and process studies. – Liberal Party of Canada Caucus, Aug. 22, 2002*

In Canadian grade school, you learn about the Vikings in the Middle Ages - about the colonies in Greenland and their discovery of North America long before Columbus, an explorer presumably well recognized at Ohio State. Had one sought an interpretation of 1000-year climate history as long ago as, say, 1996, one would have been shown a diagram with a pronounced Medieval Warm Period in the early part of the millennium and a cold Little Ice Age from the 17th to 19th centuries.



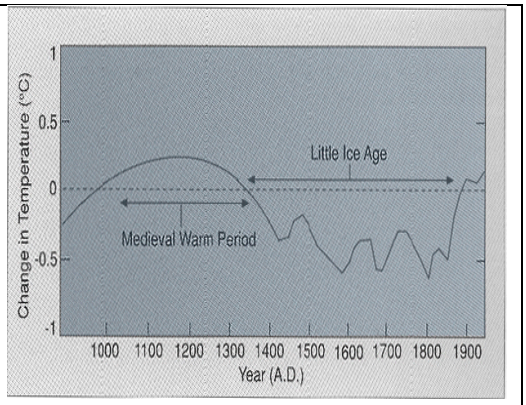


Figure 1. Left - IPCC 1990 Fig 7.1c; Right – From (Crowley 1996). Also see http://www.climateaudit.org/?p=233. This was based on (Lamb 1965) , which is criticized in IPCC AR4 (Box 6.4) as not being based on "any formal statistical calibration".

But in 2002, the Canadian government based its pronouncements on the 2001 International Panel on Climate Change Third Assessment Report (International Panel on Climate Change 2001), which prominently displayed a graphic from then very recent studies by Mann et al in 1998 and 1999 (M. E. Mann, Bradley & Hughes 1998), which announced that 1998 was the warmest year of the millennium. The graphic from the Mann study was re-drawn by IPCC with considerable graphic expertise – indeed, the graphic expertise caught my eye. It was used no fewer than 6 times and occurred as a backdrop in the Working Group 1 press conference. So it was hardly an incidental graphic in the IPCC report; it could almost be termed their logo.

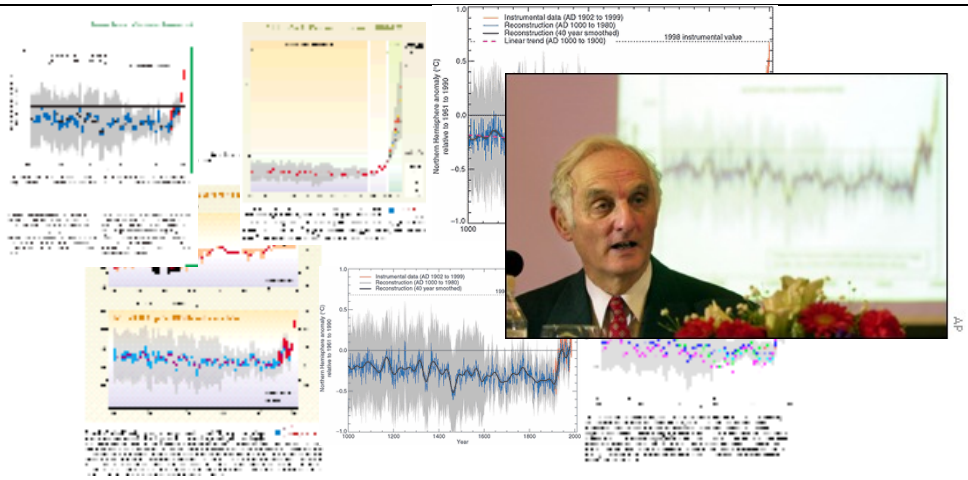


Figure 2. IPCC Third Assessment Report use of Hockey Stick graphic. Insert – Sir John Houghton, Chairman IPCC Working Group 1 at 2001 press conference releasing WG1 Report.

The graphic continues in use to this day. It occurs prominently in Al Gore's Inconvenient

Truth (Gore 2006), where it is called Dr Thompson's thermometer. Gore criticizes the "fierce attack" of skeptics. I guess this "fierce attack" would be me.

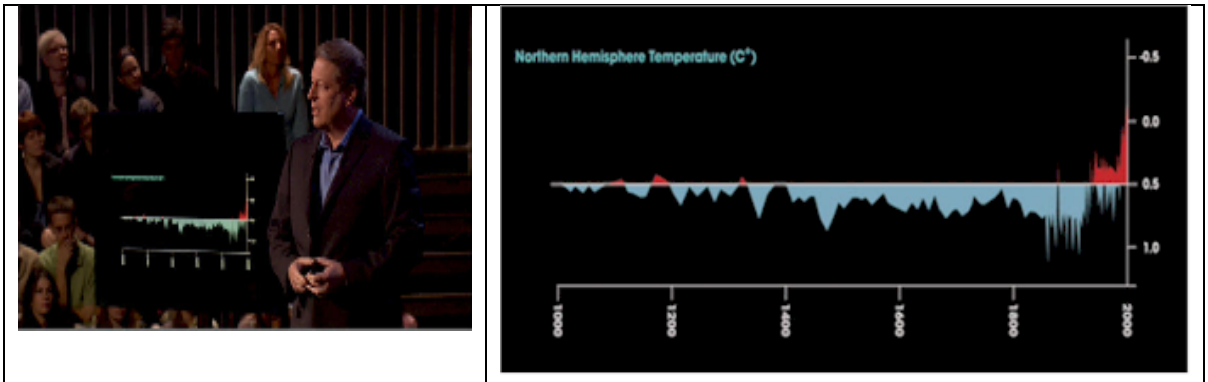


Figure 3. Al Gore presents Hockey Stick in Inconvenient Truth. Gore comments:

> **as Dr Thompson's thermometer shows, the vaunted Medieval Warm Period (the third little red blip from the left below) was tiny in comparison to the enormous increases in temperature in the last half-century - the red peaks at the far right of the graph. These global-warming skeptics - a group diminishing almost as rapidly as the mountain glaciers - launched a fierce attack against another measurement of the 1000 year correlation between CO2 and temperature known as the "hockey stick", a graphic image representing the research of climate scientist Michael Mann and his colleagues.**

The most recent IPCC Assessment Report published in 2007 (International Panel on Climate Change 2007a) substitutes a spaghetti graph of reconstructions for the Mann reconstruction. In addition to the spaghetti graph, they also produced a color-scaled version that rather artfully maintains a sort of iconographic continuity with the Mann hockey stick.

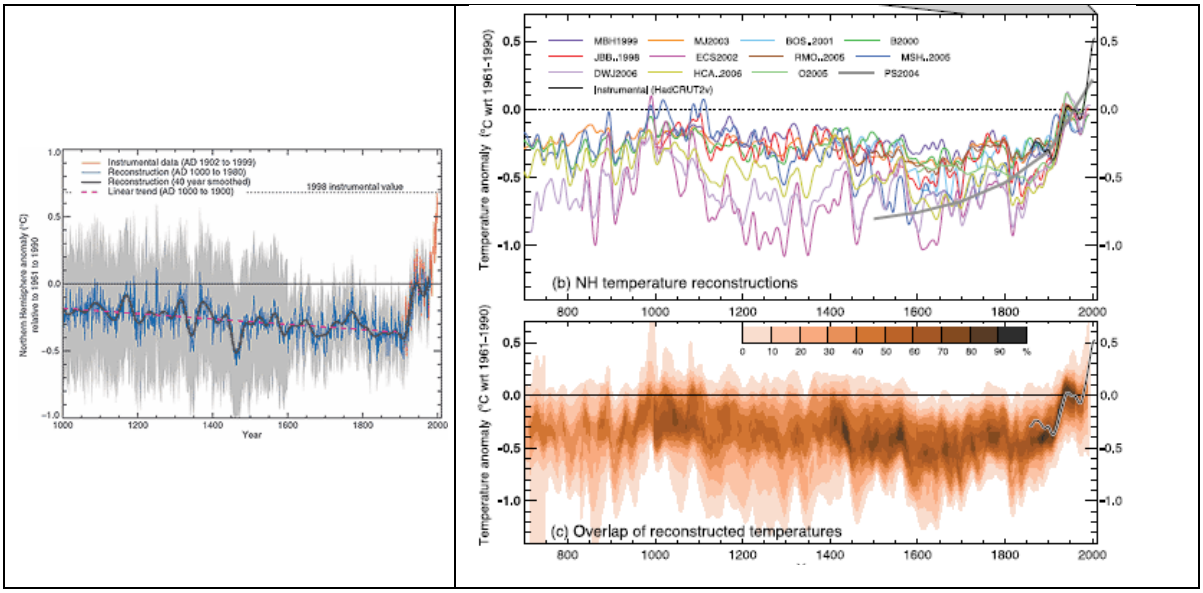


Figure 4. IPCC Northern Hemisphere Temperature History. Left – Fig 2.21, IPCC Third Assessment Report, 2001. Right, IPCC Fourth Assessment Report 2007, Box 6.4

In the most recent report, they change their slogan to say that the last half of the 20th century is the warmest 50-year period of the past 1300 years.

But how do they know either this or that 1998 was the warmest year of the millennium?

I wondered about this in early 2003 in the most casual possible fashion. I thought that it would be interesting to look at the underlying data, rather as I might look at drill data from a mining promotion. Business was slow and I browsed the internet for a due diligence package. I could not locate such a due diligence package nor the underlying proxy data for MBH98. Out of the blue (I was then a Canadian businessman unknown to climate scientists), I emailed Michael Mann, the primary author, inquiring as to the location of the MBH98 proxy data.

To my astonishment, Mann replied that he had "forgotten" the exact location, but that an associate would locate it for me. The associate said that the data did not exist in any one location, but that he would get it together for me. I was dumbfounded. Here was a study that had been on the front page of the IPCC study, used in brochures sent to every household in Canada and there was no due diligence package.

> *Dear Dr. Mann, I have been studying MBH98 and 99. I located datasets for the 13 series used in MBH99 ... and was interested in locating similar information on the 112 proxies referred to in MBH98 ... Thank you for your attention. Yours truly, Stephen McIntyre, Toronto, Canada*
>
> *Dear Mr. McIntyre, These data are available on an anonymous ftp site we have set up.* ***I've forgotten the exact location****, but I've asked my Colleague Dr. Scott Rutherford if he can provide you with that information. best regards, Mike Mann*
>
> ***Steve, The proxies aren't actually all in one ftp site (at least not to my knowledge). I can get them together if you give me a few days.*** *Do you want the raw 300+ proxies or the 112 that were used in the MBH98 reconstruction? Scott*

I realized that this study had never been audited, as I understood the process. Since no one else had done so, I thought that it would be interesting to do so – sort of like doing a big crossword puzzle. I had never written an academic paper nor had I any plans of doing so. Anyway, this led to a very unexpected and unusual introduction to the science community. I associated myself with Ross McKitrick of the University of Guelph, who I've become close friends with along the way. We published several articles, first in 2003 (McIntyre & McKitrick 2003) and then in 2005 (McIntyre & McKitrick 2005).

Our 2005 papers in particular attracted wide attention, being reported on in Nature

(Schiermeier 2005), Science (Kerr 2005) and innumerable newspaper and magazine articles. The controversy was covered on the front page of the Wall Street Journal (Regalado 2005) – something that was more validating in my community than an article in GRL or even Nature.

Figure 5. Wall Street Journal, Feb. 2005; Top right – at a hearing of the Investigations Subcommittee of the House Energy and Commerce Committee, July 2005. Left to right – Mann, Ralph Cicerone, President of the National Academy of Sciences, me, Jay Gulledge, Pew Center, Edward Wegman.

Mann was also interviewed by the Wall Street Journal and was incautious enough to tell the reporter that he would not disclose his algorithm as that would be giving in to the "intimidation tactics" that McKitrick and I were supposedly subjecting him to. This attracted the attention of the House Energy and Commerce Committee who asked Mann to produce the source code that he had refused to provide (House Energy and Commerce Committee 2005a). The academic community, which was not offended at Mann's refusal to disclose his algorithm, was outraged and various institutions protested this inquiry (AAAS 2005) (National Academy of Sciences 2005). Matters then played out a fairly large stage, with two blue-chip panels investigating the matter, leading to hearings held by the House Energy and Commerce Committee (2005b).

Meanwhile, in response originally to being attacked by Mann's blog, realclimate, I started my own blog, www.climateaudit.org, which I enjoy writing. In 2007, it won an internet vote as Best Science Blog.

So why was the Hockey Stick so influential? First, it appeared to provide a much more sophisticated statistical analysis than earlier efforts. It claimed to have "statistical skill", reporting highly significant verification RE and r2 statistics. It claimed to be robust to the presence or absence of tree ring proxies, about which there was then considerable specialist caution. It used seemingly sophisticated principal components methods to handle a much larger data set than had been considered in prior studies.

- ***Skill:*** *"significant skill in independent cross-validation tests"*
- ***Robustness to tree ring problems:*** *"possible low-frequency bias due to non-climatic influences on tree-ring indicators is not problematic ... Whether we use all data, exclude tree rings, or base a reconstruction only on tree rings, has no significant effect on the form of the reconstruction..."*
- ***Size of network:*** *Unprecedentedly large (415 proxies), some of which summarized by principal components*
- ***Geographical Balance***
- ***Qualified Proxies:*** *all had been "analysed by palaeoclimate researchers"*

But we were unable to replicate these claims. Our calculations showed that the verification r^2 statistic in the AD1400 step, the first step in MBH98, was only 0.02 – completely insignificant. Other standard statistics failed as well.

The claimed robustness to presence/absence of tree rings was also untrue. Sensitivity analysis showed that the reconstruction was not only highly sensitive to the presence/absence of bristlecone pines, but indeed the shape of the early part of the reconstruction was entirely dependent on bristlecones. We also observed that they had modified the principal components calculation so that it intentionally or unintentionally mined for hockey stick shaped series. It was so powerful in this respect that I could even produce a HS from random red noise.

This last observation has received much publicity. However, we did not and do not argue that this is the **only** way that a HS series can be obtained from red noise: there is the old fashioned method - manually select series with a hockey stick shape and then average.

- ***MBH statistical "skill" claims were untrue (verification r2 was ~0).***
- ***Robustness claims were untrue.***
- ***Worse, results specifically depended on bristlecone pine data, said by specialists not to be a temperature proxy.***
- ***No benefit from very large network as they had introduced a biased principal components methodology that ended up "mining" for hockey stick shaped series.***

References: McIntyre and McKitrick 2003(EE), 2005a, 2005c, 2005d (GRL), 2005b (EE)

These criticisms caused a visceral reaction in the climate science community. Mann and associates made a series of critical posts at their new blog (www.realclimate.org), some

of which were posted after our 2005 articles were accepted, but before they were published, in a seeming attempt to pre-empt opinion – an effort which was largely successful in the academic climate science community, which by and large was content to accept the re-assurance of Mann and associates that nothing was wrong.

Their principal academic response were two papers from two of Mann's close associates, Wahl and Ammann, one of which was rejected by GRL and one of which was eventually published in 2007 by Climatic Change (Wahl & Ammann 2007). In May 2005, the University Corporation for Atmospheric Research, a multimillion dollar corporation, issued a press release (UCAR 2005) that our claims were "unfounded", a claim that was widely repeated.

> **realclimate:** False Claims by McIntyre and McKitrick regarding the Mann et al. (1998) reconstruction; Myth vs. Fact Regarding the "Hockey Stick"; On Yet Another False Claim by McIntyre and McKitrick; Dummies guide to the latest "Hockey Stick" controversy .
>
> Mann This claim by MM is just another in a series of disingenuous (off the record: plainly dishonest) allegations by them about our work. Our reconstruction passes both RE and R^2 verification statistics if calculated correctly. … I hope you are not fooled by any of the "myths" about the hockey stick that are perpetuated by contrarians, right-wing think tanks, and fossil fuel industry disinformation. (Michael Mann 2005)
>
> UCAR "the highly publicized criticisms of the MBH graph are unfounded" (UCAR 2005)
>
> Houghton **at Senate, July 2005:** "the assertions by McIntyre and McKitrick have been shown to be largely false in the context of the actual data used by Mann and co-workers." (Houghton 2005)

Unfortunately, the UCAR press release was also untrue. To take the simple example of verification statistics, we had reported that the verification r^2 was 0.02; Wahl and Ammann Table 1S showed that it was 0.018. This table was included very unwillingly. It was not in the first draft, which I happened to review. As a reviewer, I asked that this information be included. Wahl and Ammann refused to provide it; the editor supported them and I was terminated as a reviewer. I met with Ammann and told him that he had an obligation to provide these results; he still refused. I filed a complaint with UCAR and while the complaint was rejected the results were reported, obviously showing that this particular claim was not "unfounded." This is far from the only such case.

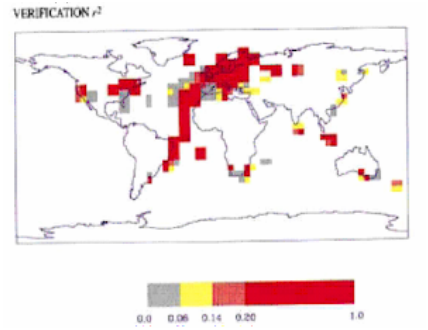


Table 3. Verification Period Statistics for AD1400 Step of MBH98 Reconstruction[a]

	AD1400 Step Results	
	Emulation	**MBH98 Reported**
RE (β)	0.46	0.51
R^2	0.02	NA
CE	−0.26	NA
Sign Test	22	NA
Product Mean Test	1.54	NA

[a]From emulation and as reported by MBH98.

Table 1S Pearson's r^2 and CE scores for MBH reconstruction emulations

Proxy network MBH – periods	NH mean r^2 Calibration-period	NH mean r^2 Verification-period	NH mean CE Verification-period
1400–1449	0.414	0.018	−0.215
1450–1499	0.483	0.010	−0.314
1500–1599	0.487	0.006	−0.253
1600–1699	0.643	0.004	−0.259
1700–1729	0.688	0.00003	−0.161
1730–1749	0.691	0.013	−0.063
1750–1759	0.714	0.156	0.077
1760–1779	0.734	0.050	−0.070
1780–1799	0.750	0.122	0.040
1800–1819	0.752	0.154	0.069
1820–1980	0.759	0.189	0.103

Figure 6. MBH98 Figure 3 showing verification r2 statistics for individual gridcells for the AD1820 step; top right, McIntyre and McKitrick 2005 (GRL) Table 3 showing 0.02 verification r^2 for AD1400 step; bottom right – Wahl and Ammann Table 1S showing verification r^2 of 0.018 for AD1400 step and failed values for other steps.

Wahl and Ammann purported to excuse this statistical failure by arguing that verification r2 was not a relevant statistic for assessing reconstructions – an argument that I find completely unconvincing, but, in any case, should have been made in the original publication together with disclosure of the statistical failure.

As a result of the questions from the House Energy and Commerce Committee in 2005, two senior academic panels were formed to consider the dispute – one consisting of statisticians led by Edward Wegman, an eminent statistician; the other consisting primarily climate scientists, chaired by Gerald North, also very eminent.

The findings of these panels have been spun mercilessly. Several key findings of the NAS panel (National Research Council 2006) clearly supported us. They completely endorsed our criticism of Mann's principal components method; they agreed that the sole use of the RE statistic was insufficient statistically and, notably, they acknowledged the MBH dependence on bristlecones and stated that strip bark trees should be avoided in temperature reconstructions.

Specific NAS Panel Findings

- the MBH principal components method was flawed.
- Passing an RE statistic was insufficient for claiming statistical significance
- MBH results did depend on bristlecones
- Such strip bark forms should be "avoided" in reconstructions

In making this last recommendation, they cited authorities that we had used, including Graybill and Idso (1993) who stated:

> "trends of the magnitude observed in 20th century ringwidth growth conspicuously lacking in all climatic time series …. lack of strong, consistent temperature responses in the subalpine chronologies is a perplexing problem … an ecophysiological study of bristlecone pine reported no marked or direct correlation of summer temperature with ringwidth growth

and Hughes and Funkhouser (2003) who stated:

> the challenge of unraveling the influences of precipitation and temperature at the highest elevations depends on a resolution of the mystery of the century-scale growth increase in certain trees since the mid-19th century

The Wegman Report (E. J. Wegman, Scott & Said 2006) was even more forcefully critical of the Mann study. In subsequent questioning (Edward Wegman 2006), Wegman additionally rejected the Wahl and Ammann defences in categorical terms.

> While the work of Michael Mann and colleagues presents what appears to be compelling evidence of global temperature change, the criticisms of McIntyre and McKitrick, as well as those of other authors mentioned are indeed valid.
>
> I am baffled by the claim that the incorrect method doesn't matter because the answer is correct anyway. Method Wrong + Answer Correct = Bad Science.
>
> It is our understanding that when using the same proxies as and the same methodology as MM, Wahl and Ammann essentially reproduce the MM curves. Thus, far from disproving the MM work, they reinforce the MM work. The debate then is over the proxies and the exact algorithms as it always has been.
>
> Wahl and Ammann adjust the MBH methodology to include the PC4 bristlecone/foxtail pine effects are significant reasons we believe that the Wahl and Amman paper does not convincingly demonstrate the validity of the MBH methodology.

At the House Committee hearings (House Energy and Commerce Committee 2005b), NAS Panel chairman Gerald North and panelist Peter Bloomfield were specifically asked whether they disagreed with the severe criticisms of the Wegman Report and both of them stated that they agreed with his findings. Eduardo Zorita, a prominent climate scientist familiar with the dispute, stated at the time that he thought that the criticisms were as severe as could be possibly expected in the context of the forthcoming IPCC report.

> CHAIRMAN BARTON. Dr. North, do you dispute the conclusions or the methodology of Dr. Wegman's report?
>
> DR. NORTH. No, we don't. We don't disagree with their criticism. In fact, pretty much the same thing is said in our report.
>
> DR. BLOOMFIELD. Our committee reviewed the methodology used by Dr. Mann and his coworkers and we felt that some of the choices they made were inappropriate. We had much the same misgivings about his work that was documented at much greater length by Dr. Wegman.
>
> WALLACE: "the two reports were complementary, and to the extent that they overlapped, the conclusions were quite consistent." (Am Stat Assoc.)

The news coverage told a different story. Nature, the New York Times and Boston Globe announced that the hockey stick had been totally vindicated, while Mann excoriated Wegman at his blog for "parroting" the views of two Canadians – Canadian being a term of disapproval in this context.

> **New York Times:** "Science Panel Backs Study on Warming Climate"
>
> Boston Globe: "National panel supports '98 global warming evidence"
>
> **Nature:** "Academy affirms hockey-stick graph"
>
> **Mann:** "[Wegman] uncritically parrots claims by two Canadians that … were specifically dismissed by the National Academy in their report just weeks ago."

The Spaghetti Graphs

The main line of defence was the argument that other people had got somewhat similar results using somewhat different methods – an observation which supposedly meant that our criticisms were therefore "wrong". Obviously the fact that other studies got similar results in no way invalidated our criticisms of the Mann study. But it was obviously still an open question whether these other studies had been more successful statistically or whether they had committed similar errors. The preferred iconographic style for

representing the results of these other studies now became what can be called a "spaghetti graph" – on the left is one from the NAS panel, with increasing darkness of the background representing the increasing "murkiness" of our knowledge' the one on the right is from IPCC AR4.

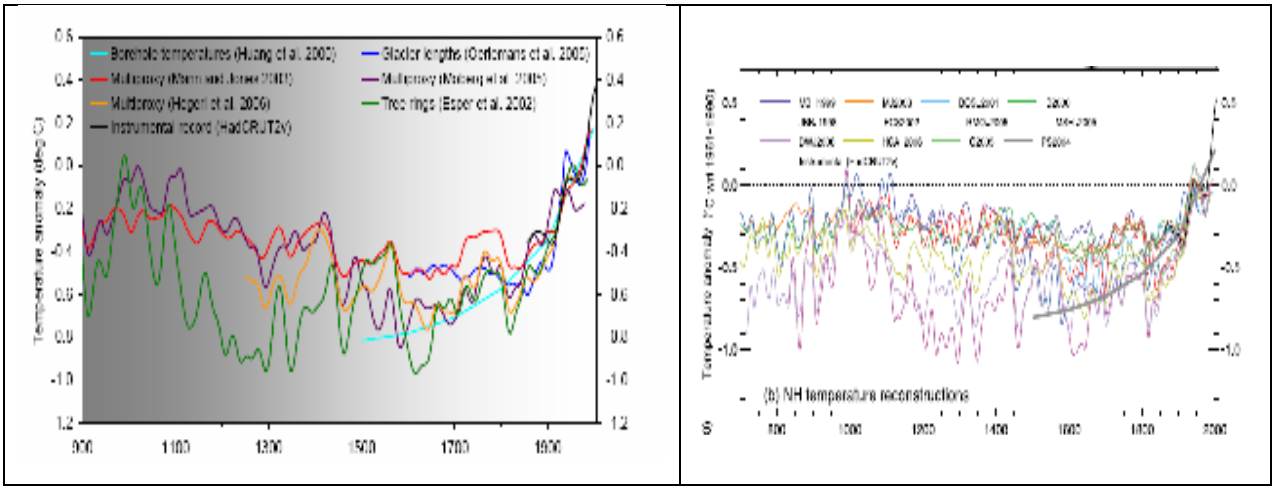


Figure 6. Left – NAS panel spaghetti graph; right – IPCC spaghetti graph. Similar spaghetti graph at Wikipedia.

One of the obvious defects of these supposedly independent studies is that the same proxies are re-cycled over and over. Here is a re-plot of a figure from Wegman (who drew on information at Climate Audit) showing that Polar Urals and Tornetrask proxies were used in every reconstruction, while bristlecones or foxtails were used in all but a few.

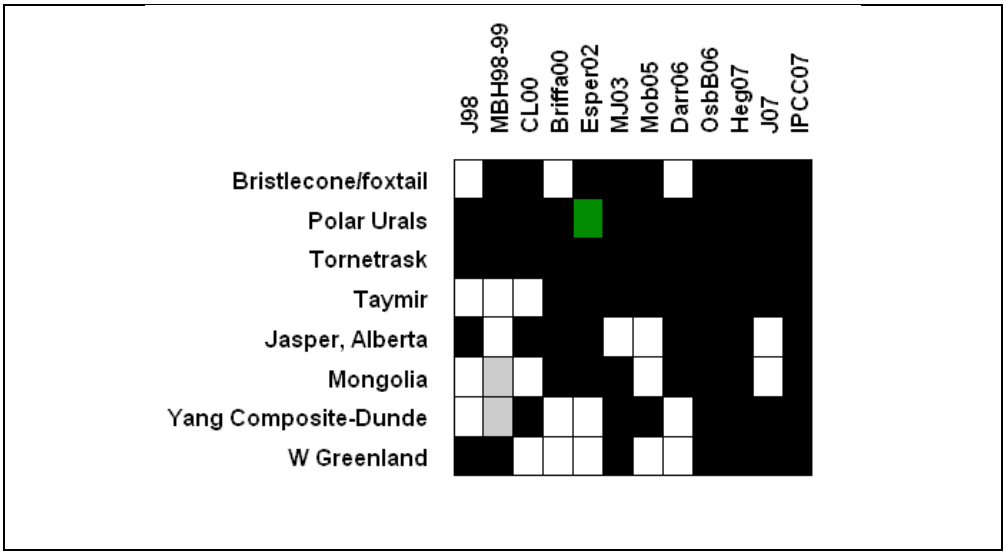


Figure 7. Repetition of Proxies in IPCC Multiproxy Reconstructions. Compare to Wegman et al 2006 Figure 5.8.

Thus any problems with these three proxies would have a dramatic knock-on effect on multiple reconstructions. The NAS panel had already said that strip bark should be avoided, but all 4 series in their spaghetti graph used strip bark. How did this happen? It's as though engineers said that Grade C concrete should not be used in bridge construction and then submitted 4 designs using Grade C concrete.

One day, Gerry North held an online colloquy and I got an opportunity to ask him on the record (Chronicle of Higher Education 2006) (also see CA, 812), whether they had carried out any due diligence to assess whether strip bark had been used in these other studies. He said that they had not carried out any such due diligence.

> **McIntyre Question:** The NRC Panel stated that strip-bark tree forms, such as found in bristlecones and foxtails, should be avoided in temperature reconstructions and that these proxies were used by Mann et al. Did the Panel carry out any due diligence to determine whether these proxies were used in any of the other studies illustrated in the NRC spaghetti graph?
>
> **North:** We did not dissect each and every study in the report to see which trees were used.

North is a very decent guy and, by and large, he's been nice to me, but some of his attitudes demonstrate a lack of understanding by academics on due diligence obligations. In an online seminar (North 2006. minute 55 or so), he says that their panel "just kind of winged it… that's what you do in that kind of expert panel". National Academy of Sciences panels are deemed to expert panels for US Supreme Court evidence; surely such a casual attitude is inconsistent with such responsibility.

We also observed similar statistical defects in the other studies in the spaghetti graphs, a fact reported at AGU in 2006 (McIntyre 2006) All but one fails a Durbin-Watson test in the calibration period, and all fail verification r2 tests, most dismally. Thus the statistical calibration skill used to reject the older literature proves to be a mirage.

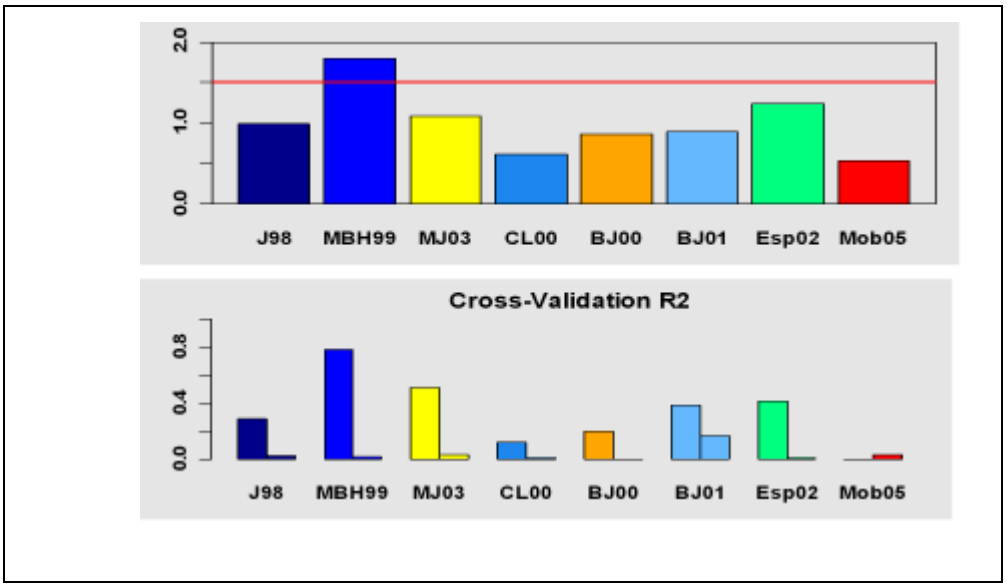


Figure 8. Top – Durbin-Watson statistics in calibration period; bottom – verification r2 statistics. Source: McIntyre (AGU 2006)

These statistical patterns have all the earmarks of a problem long familiar in econometrics – spurious regression. As long ago as 1926, Yule (1926) presented the conundrum of the relationship between proportion of Church of England marriages and mortality – a relationship that in climate science terms was 99,9% significant. This relationship also passes the RE test used as a talisman by climate scientists. The same is true of Hendry's (Hendry 1980) relationship between the UK consumer price index and cumulative rainfall.

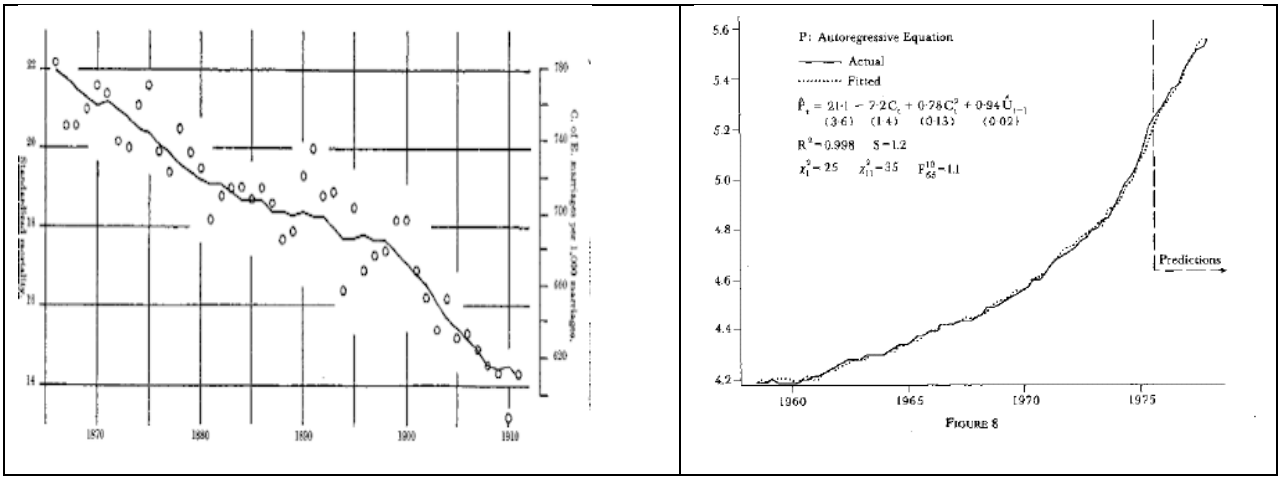


Figure 9. Left: Yule 1926 (J Roy Stat Soc): Mortality per 1000 (points) and proportion of Church of England marriages per 1000 marriages (line); Right – Hendry 1980 (Economica) UK consumer price index versus cumulative rainfall. Both relationships "pass" the RE test.

Ferson et al (2003) had observed that the problem of spurious regression is exacerbated by data mining – something that should be of profound concern in this field, given the

proven recycling of the same proxies over and over.

> Data mining for predictor variables [proxies] interacts with spurious regression bias. The two effects reinforce each other because more highly persistent series are more likely to be found significant in the search for predictor variables. Our simulations suggest that many of the regressions in the literature, based on individual predictor variables, may be spurious…
>
> The pattern of evidence in the instruments in the literature is similar to what is expected under a spurious mining process with an underlying persistent expected return. In this case, we would expect instruments to arise, then fail to work out of sample.

While there are no magic statistical calculations that can, after the fact, separate spurious regressions, given that many of the proxies were originally collected in the 1970s or 1980s, as observed by Greene (2000) in the context of economics, one way to guard against data snooping was simply to wait 30 years and test the hypothesis with new data. This opportunity exists right now for bristlecones and other series: with the warm 1990s and 2000s, ring widths should be at record levels.

> Any test of a theory or model is corrupted if the test is conducted using data which overlaps that of any previous empirical study used to suggest that theory or model…
>
> Testing in un-mined data sets is a difficult standard to meet only to the extent one is impatient. There is a simple and honest way to avoid invalid testing. **Suppose in 1980 one surveys the literature on money demand and decides the models could be improved. File the proposed improvement away until 2010 and test the new model over data with a starting date of 1981**... Only new data represents a new experiment. I do not consider this a pessimistic outlook. Much can be learned from exploring a sample. Patience and slow methodical progress are virtuous.

The "Proxies"

Why haven't the proxies been brought up to date? Michael Mann explained that it was impossible to replicate the heroic collections of the 1970s because of the heavy equipment involved.

> paleoclimatologists are attempting to update many important proxy records to the present, this is a costly, and labor-intensive activity, often requiring expensive field campaigns that involve traveling with heavy equipment to difficult-to-reach locations (such as high-elevation or

remote polar sites). For historical reasons, many of the important records were obtained in the 1970s and 1980s and have yet to be updated.
http://www.realclimate.org/index.php?p=11

The "heavy" equipment involved in tree ring coring is illustrated here.

Figure 10. "Heavy equipment" used in dendrochronology.

In the case of bristlecones, it seems like elementary due diligence to update the proxies. The Mann reconstruction is not merely dependent on bristlecones; it is dependent on the chronologies collected by one researcher, Donald Graybill, back in the 1980s. At left, I show the contribution of bristlecones to the Mann reconstruction; other colors show the contribution of other proxies. Graybill's sites are not particularly remote – one of his sites is within 30 miles of Colorado Springs.

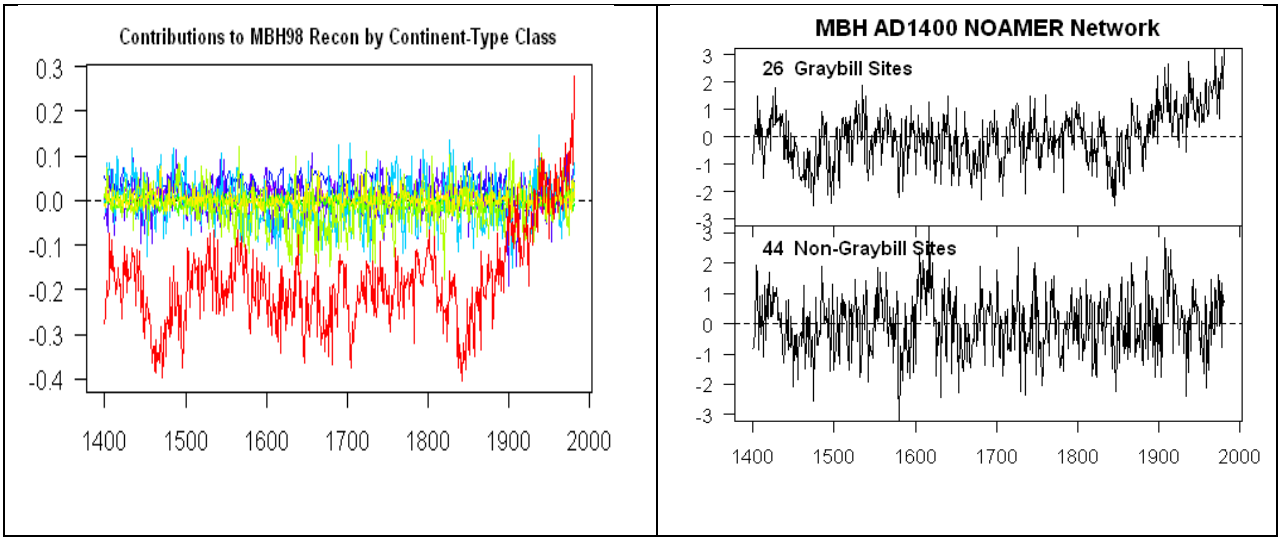


Figure 11. Left- contribution of bristlecones to MBH98; Comparison of Graybill and non-Graybill averages in AD1400 MBH network.

1) Bristlecones

a) Almagre CO

One of my sisters lives in Colorado Springs. My wife and I visited her last summer and we decided to test what I termed the Starbucks Hypothesis on my blog – could you have a Starbucks latte in the morning and still complete a full day's sampling?

We obtained a permit from the Forest Service and, if nothing else, we proved the Starbucks Hypothesis. The real leaders of the expedition were Pete Holzmann, a Climate Audit reader, and his wife and the sampling is primarily due to their efforts. We ended up with 64 cores, which were measured and analyzed at the University of Guelph, with myself and Climate Audit readers paying for the analysis without subsidy. I placed all the measurements online in Oct 2007 as soon as I received them and within a few months of taking the sample. Almagre is (for now) the highest millennium-length tree ring chronology in the world!

Figure 12. Testing the Starbucks Hypothesis at Mt Almagre, Colorado. See http://picasaweb.google.com/Almagre.Bristlecones.2007/

One interesting accomplishment was that we actually located many of the very trees that Graybill sampled. This was partly good luck and partly perseverance. There was no location map or other record of where the samples were taken. We guessed that they were taken near the road up Almagre, but didn't know where. Pete recognized one of the vistas in an old photograph that was online and, from that foothold, we were able to locate the exact tree that Graybill is shown sampling here in 1983.

Figure 13. **Left- Graybill in 1983; right - Pete Holzmann in 2007, both on same tree.**

The NAS panel warned about strip bark but the shape of these trees has to be seen to be believed. These are not circular trees, but are contorted and asymmetric. The bark may only be a few inches wide. If you're looking at one of these trees from the side, it's impossible to guess where the pith is located. One core, which is all that Graybill typically archived, has little chance of establishing the geometry of the tree.

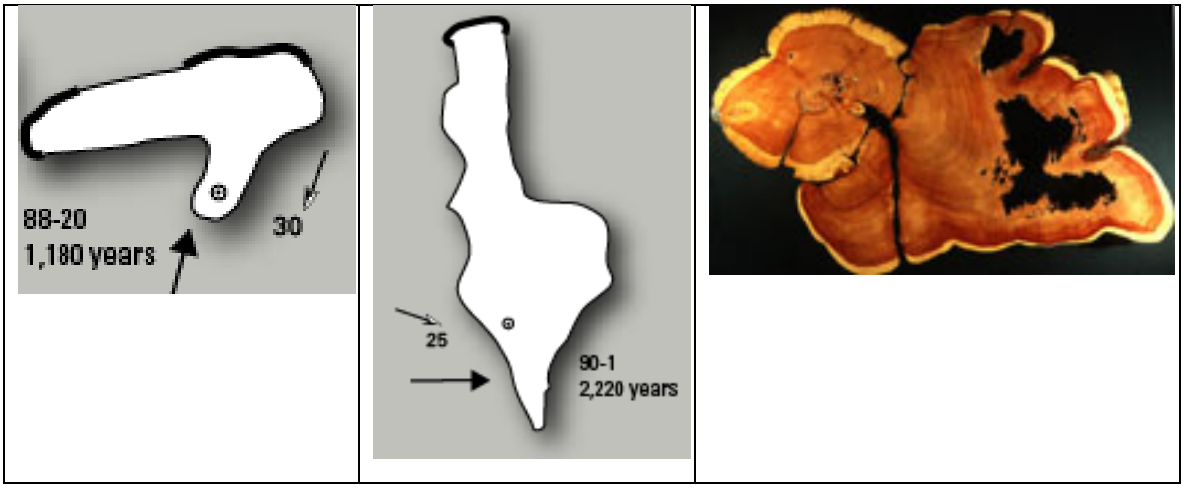


Figure 14. Left and middle – two cross sections of Colorado bristlecones from Brunstein (Brunstein 2006); right – strip bark juniper from Karakorum, Pakistan used in Jan Esper chronology.

One of our findings, which to our knowledge is unreported in the scientific literature, is the fantastic difference in ring width chronologies from cores as little as 6 inches from one another, At top right below is a whole bark tree with a relatively regular decline in ring width with age, modeled by tree ring specialists as a negative exponential. The top right shows two adjacent cores, with only one going to 6 or 7 sigma levels. You can readily see how, in a small collection of 20 or so cores, even a few such 6 or 7 sigma growth pulses could have a huge impact on the chronology. It's hard to say how one would even begin to go about constructing an error model when you're dealing with 6 or 7 sigma growth pulses lasting over a century, and, it seems with an eventual precipitous

decline. In this case, we were able to match this tree to Graybill's data; in this case, he only archived one core, one which had an extreme growth pulse. Graybill *said* that he selected for strip bark; is it possible that this biased his chronologies?

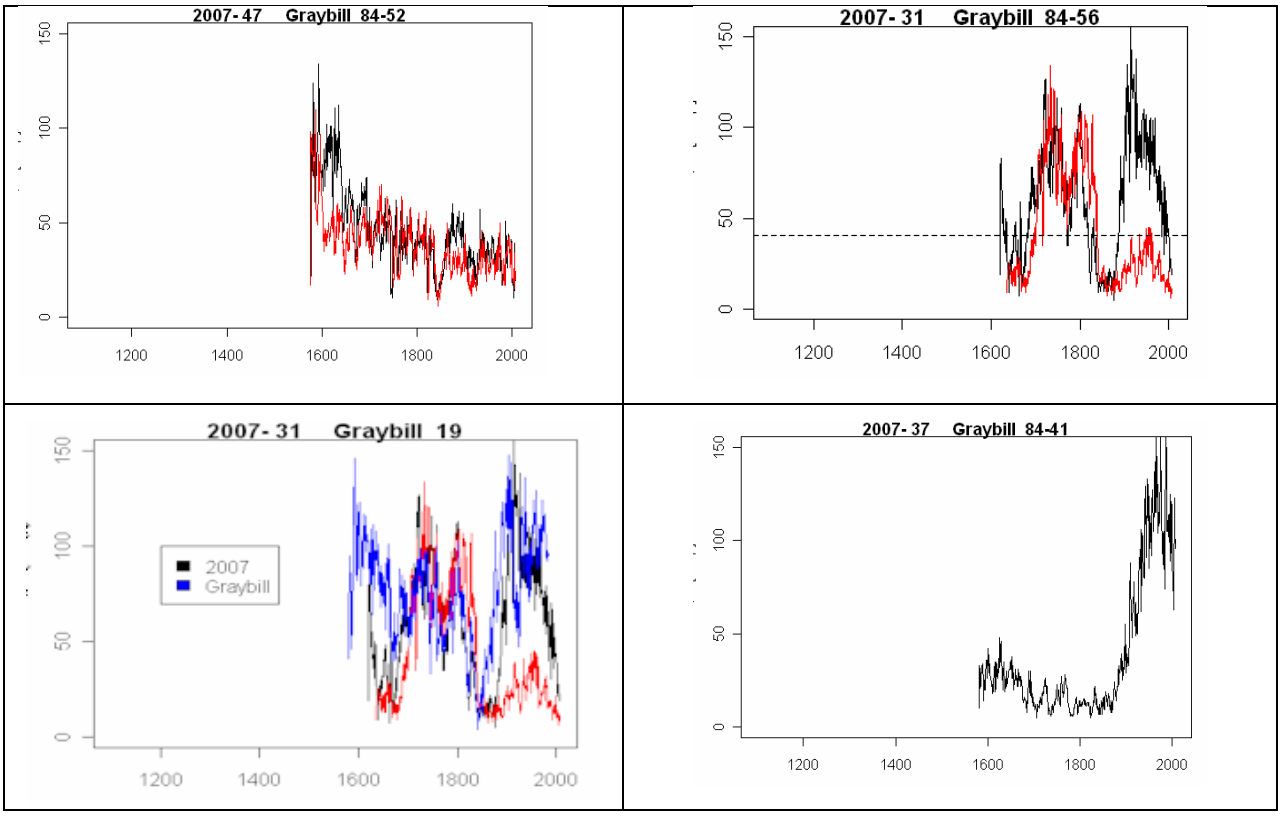


Figure 15. Instead of the expected negative exponential growth trend ("conservative"), strip bark trees had huge 6-7 σ growth pulses, which can persist for a century and then wear off. Cores taken a few inches apart can have very different results (top right). Graybill archive (bottom right) only had one core - with high recent growth. Composites of 15-25 cores will be affected by proportion with growth pulse,

I don't think that this is what the NAS panel had in mind when they made their recommendation, but you can see how relevant the recommendation is.

We also found that growth in the 1990s and 2000s for the majority of trees was not at record levels – far from it, recent growth tended to be low.

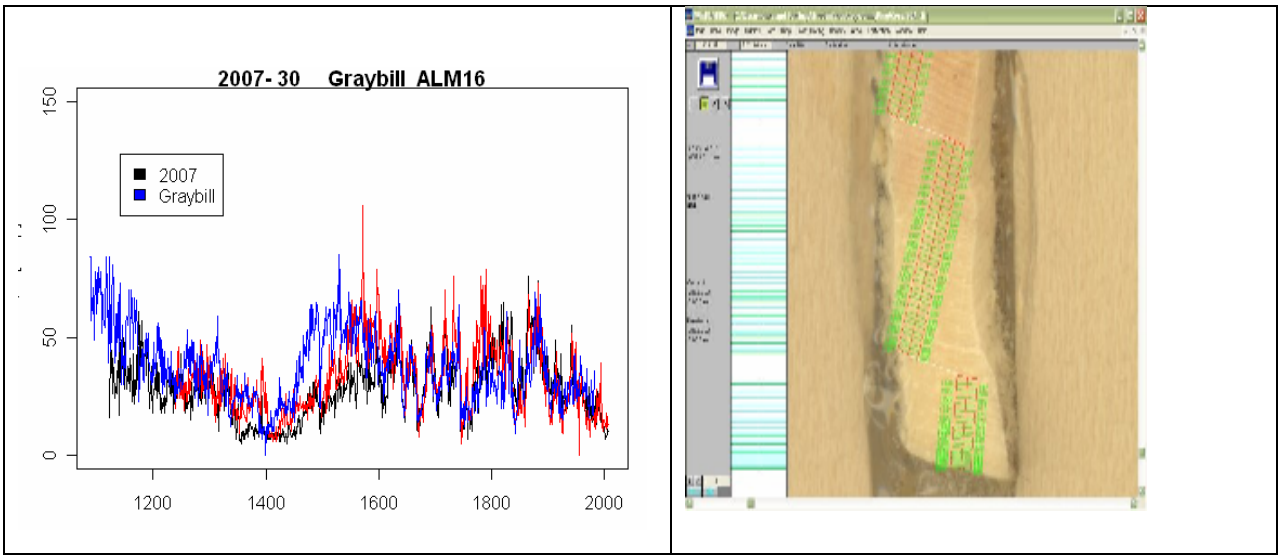


Figure 16. Left- Ring widths for Almagre Tree 30 (black, red – 2007 measurements; blue – Graybill); right – image of recent portion of core showing narrow recent widths at bottom relative to wider widths in 19th century for this core.

Without necessarily arriving at a view as to what meaning a chronology had for this data, I calculated a chronology using both our data and the Graybill data, bringing the results up to date. The hypothesis of a linear relationship between ring widths and world temperature doesn't hold up.

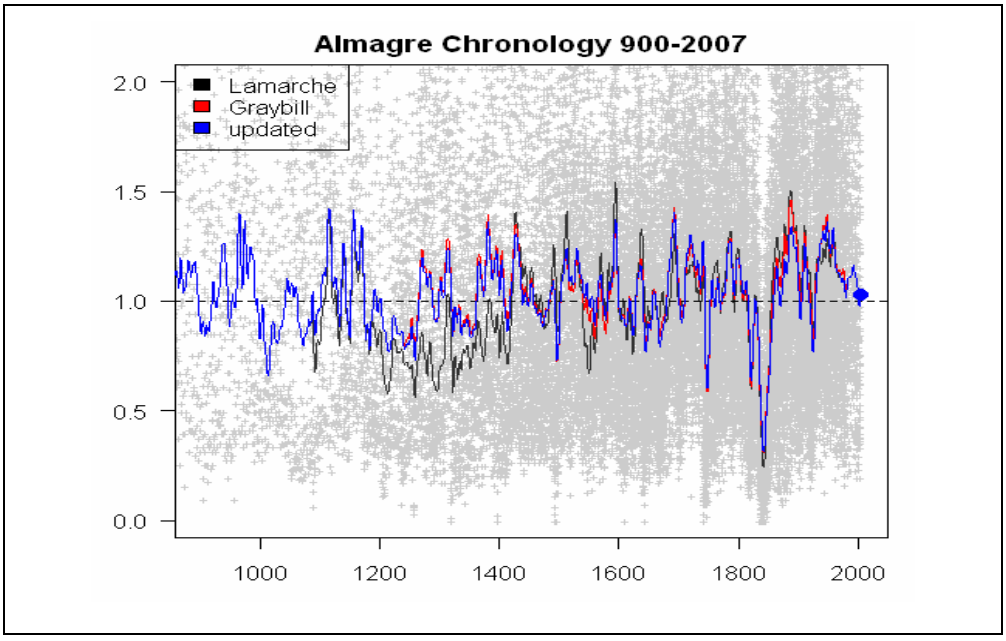


Figure 17. Almagre Chronologies: black – Lamarche (co071); Graybill (co524); blue- incorporating 2007 samples.

In passing, I mention an important archiving problem. Pete Holzmann identified actual tags from the Graybill program. We found that 50% of the data had not been archived. Was this selective or not? No one knows. Graybill died quite young. His

notes were notoriously incomplete. Worse, when the Tree Ring Laboratory moved a few years ago, apparently they forgot to arrange for old samples to be protected. Their former quarters were destroyed. Some of the records were apparently recovered from the trash by one scientist but others are permanently lost.

Pete Holzmann discovered 21 actual tags from Graybill program (of which only 9 were re-sampled as 12 tags were located after end of permit and not sampled). The tag numbers did not match ITRDB archive numbers, but, with the aid of the University of Arizona, a concordance was established using unpublished logbooks. Tags and logs indicate that Graybill sampled 42 trees (84-36 to 84-77). Only 21 trees (23 cores) are archived at ITRDB (although 36 have been measured and cross-dated.)

Figure 18. Graybill tag discovered by Pete Holzmann at Mt Almagre.

b) Sheep Mountain CA

A second important site, Sheep Mt, was updated by a University of Arizona grad student, Linah Ababneh in 2002 with her thesis accepted in 2006(Ababneh 2006) and an article published in 2007(Ababneh 2007). She had a much larger sample than Graybill. Her results failed to replicate the anomalous 20th century previously reported by Graybill (Graybill & Idso 1993). This dramatic inconsistency casts serious doubt on the usability of Graybill chronologies without a complete reconciliation.

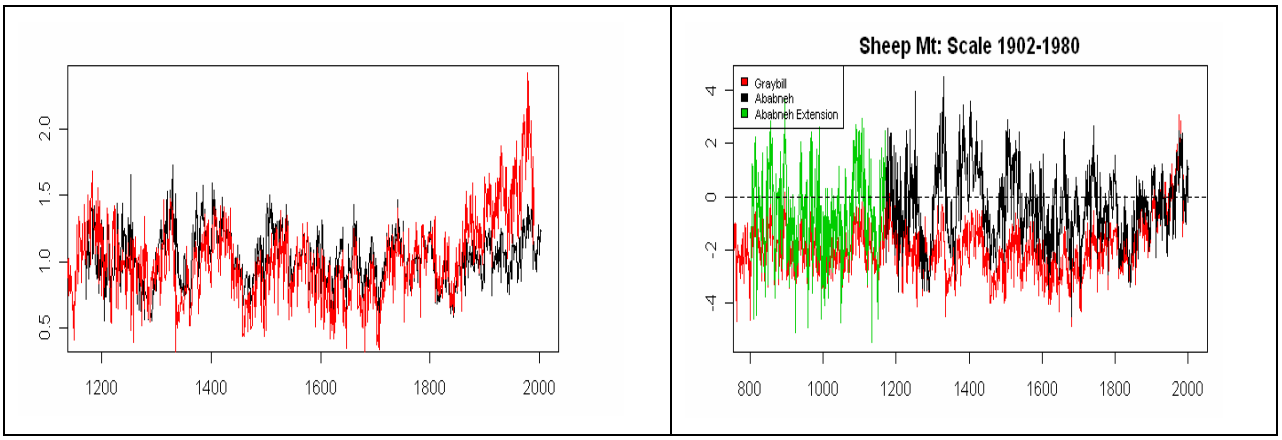


Figure 19. Left - Sheep Mt bristlecone chronologies (dimensionless c*hronology units)* Red-Graybill 1987; black – digitized from Ababneh 2006. Right- same chronologies scaled on 1902-1980. See CA, 2682.

Despite the fact that the measurements were done in 2002 and her thesis accepted in

2006, studies in 2007, including the IPCC Fourth Assessment Report (International Panel on Climate Change 2007a), continue to use the obsolete Graybill version. In addition, the ring widths that she obtained also diverge from temperature in the latter part of the 20th century.

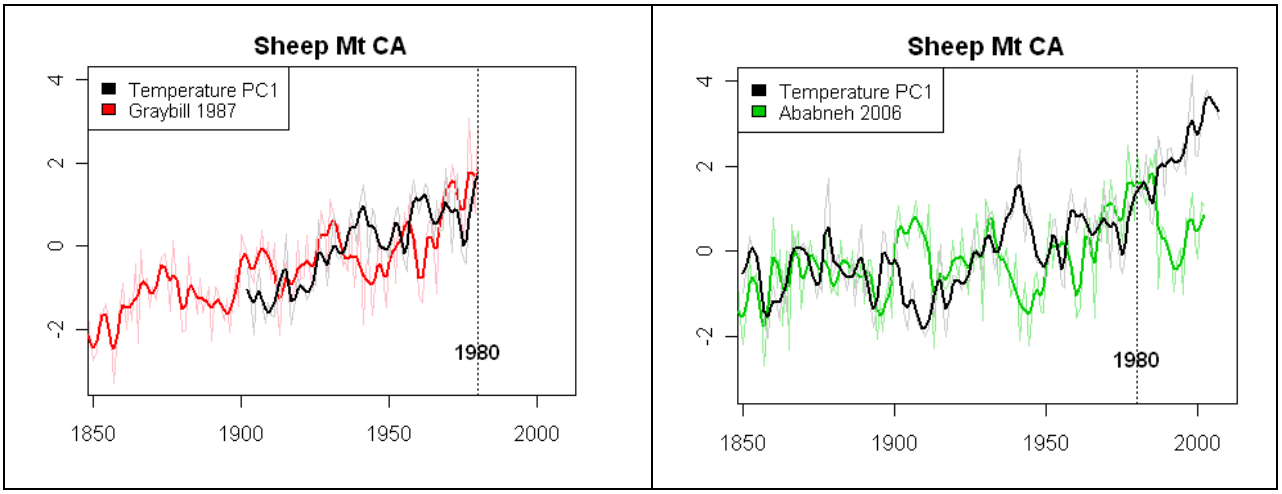


Figure 20. Left: Graybill chronology against MBH PC1; right – Ababneh's chronology against updated temperature PC1. See CA, 2682.

Recent studies in California (Millar et al. 2006) analyzed ecological combinations of trees growing at high altitudes in the MWP, where subfossil trees can be found above the present treeline. Their conclusion was that the MWP was 3.2 deg C warmer than present in California.

> Deadwood tree stems scattered above treeline on tephra-covered slopes of Whitewing Mtn (3051 m) and San Joaquin Ridge (3122 m) show evidence of being killed in an eruption from adjacent Glass Creek Vent, Inyo Craters. We dated deadwood to 815-1350 CE... Using contemporary distributions of the species, **we modeled paleoclimate during [the MWP] to be significantly warmer (+3.2 deg C annual minimum temperature) and slightly drier (-24 mm annual precipitation) than present.**

As a reviewer, I asked IPCC to include this information in AR4; they refused. (International Panel on Climate Change 2007b, 6-1127) The pictures below show medieval subfossil trunks above the present treeline in two different locations.

Figure 21. Left- Whitewing Mt (C Millar); right – Boreal Plateau (Andrew Bunn)

2) Polar Urals

The second key IPCC proxy series – one of three versions being used in all IPCC studies- is from the Polar Urals in Siberia.

In 1995, Keith Briffa, who was the lead author in AR4 (succeeding Mann who was the author in AR3), reported in Nature (Briffa et al. 1995) that the 11th century was cold and 1032 was the coldest year of the millennium – a claim that had a lot of traction in countering apparent MWP warmth in the Viking floruit around 1000.

The Briffa reconstruction was based on only 3-4 short cores in the 11th century – much fewer than QC standards in the field. In the late 1990s, new measurements of medieval trees were made, resulting in a dramatic change. If the new measurements were incorporated, the 11th century was no longer cold, but emerged as one of the warmest periods.

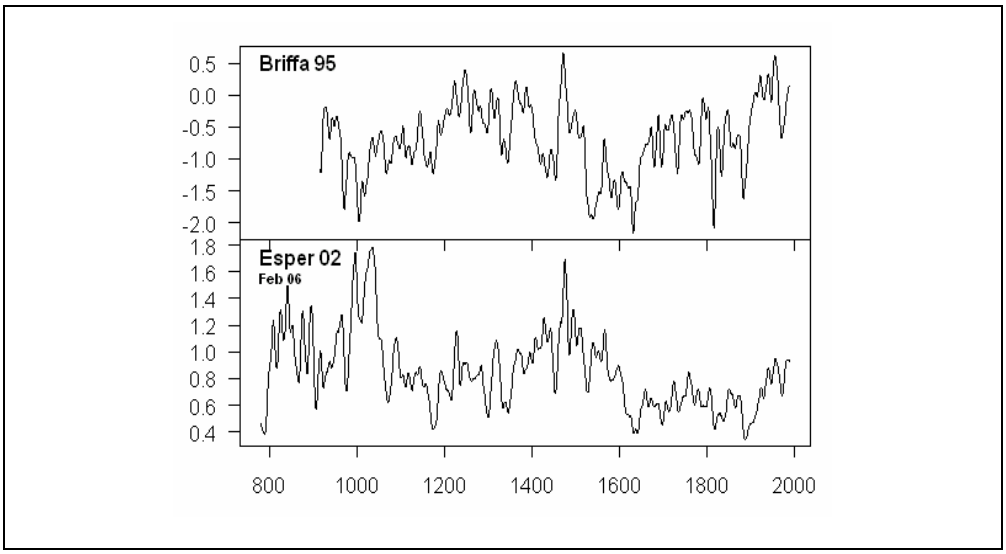


Figure 22.Top – Briffa et al 1995 (Nature) reported that 1032 was the "coldest year of the millennium" and that 20th century was "warmest" century in millennium. This data set important in multiproxy reconstructions (e.g. (Jones et al. 1998). Bottom - New

measurement data from 1998 yielded very different results = a very warm 11th century, especially 1032 (bottom left, repeated top right). Briffa did not report this nor any article published. New data used in only one multiproxy study (Esper, Cook & Schweingruber 2002)) and this one data set causes a different result. *(In sd units) See http://www.climateaudit.org/?p=547*

If these updated results were incorporated in, say, the (Briffa 2000) reconstruction, it would have led to a MWP period warmer than the modern period. Instead of reporting these results, Briffa (Briffa 2000) replaced the Polar Urals site with another site 100 miles away (Yamal) at which he calculated 20th century ring widths that were substantially elevated relative to the MWP, especially towards the end.

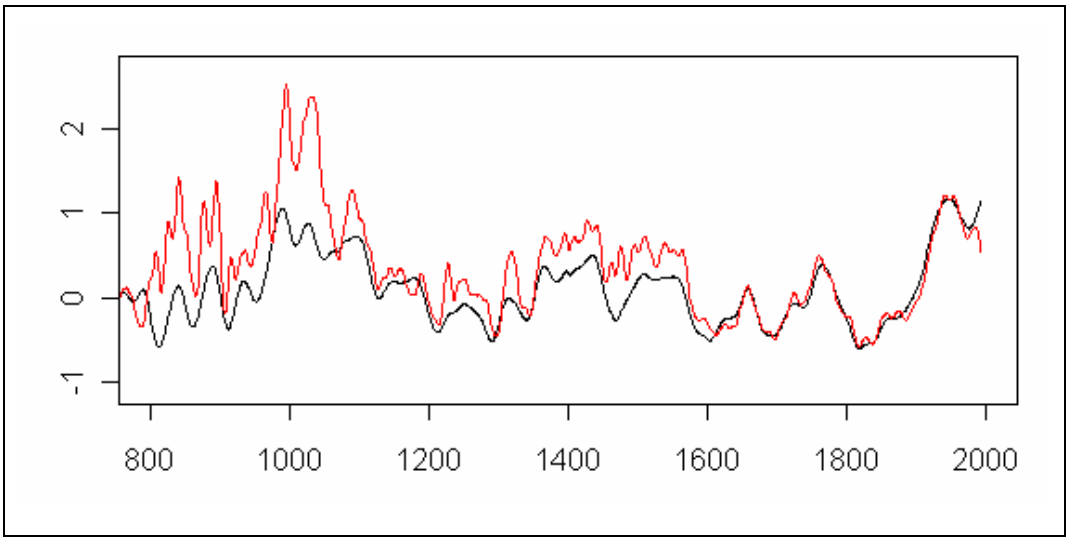


Figure 23. Black – Briffa 2000 reconstruction using Yamal chronology; red – impact using Polar Urals update. Impact on D'Arrigo et al 2006 will be similar since 5 of 6 series overlap. See CA, 693.

Until there is an adequate reconciliation of such inconsistent results, I don't see how any of these results can be incorporated into climate reconstructions.

In Siberia, as in California, there is ecological evidence of a warm MWP. Naurzbaev and MBH co-author Hughes estimated in 2004 (Naurzbaev, Hughes & Vaganov 2004) that the MWP was 1.5 to 3 deg C warmer than at present.

> "trees that lived at the upper (elevational) tree limit during the Medieval Warm Epoch (900 to 1200) show **annual and summer temperature warmer by 1.5 and 2.3 deg C, respectively**, approximately one standard deviation of modern temperature. Note that these trees grew 150-200 m higher (1-1.2 deg C cooler) … implying that this may be an underestimate of the actual temperature difference.

As long ago as 1995, at the same time as Briffa was asserting medieval cold, Shiyatov (Shiyatov 1995) reported that growth at the Polar Urals site in the MWP was the most intense in the entire record:

> The 12th and 13th centuries were most favorable for larch growth … the timberline was the highest, stand density the biggest, longevity of trees the longest, size of trees the largest, increment in diameter and height the most intensive

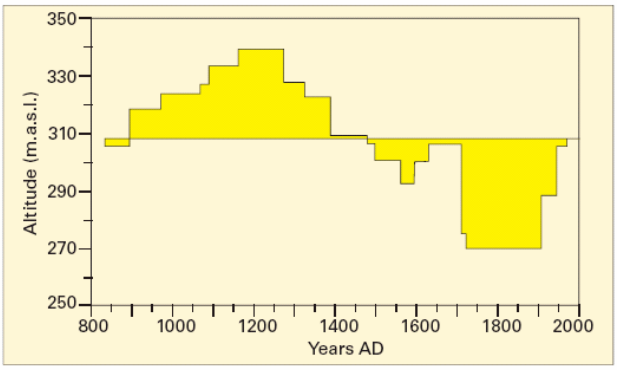


Fig. 1. Altitudinal displacement of the upper treeline in the Polar Ural Mountains during the last 1150 years.

Figure 24. Left – Photograph of Polar Urals treeline, advancing modern treeline in background, medieval treeline in foreground (Jan Esper); right – Shiyatov 1995 figure showing treeline elevation at Polar Urals with high medieval treeline.

3) Tornetrask

Serious problems also occur at Tornetrask, the third staple of every multiproxy reconstruction. In this case, the versions used in the multiproxy studies once again derive from Briffa (Briffa et al. 1992).

In this case, Briffa's reconstruction is based on density (MXD) which happens to trend down in the 20th century. Briffa decided that this didn't make sense and so Briffa et al 1992 adjusted the trend of the density series upwards after 1750 – an adjustment that has been incorporated without comment in subsequent multiproxy reconstructions. In 2005, I reported on this at Climate Audit showing the elevated MWP that resulted from available data without this arbitrary and forced adjustment.

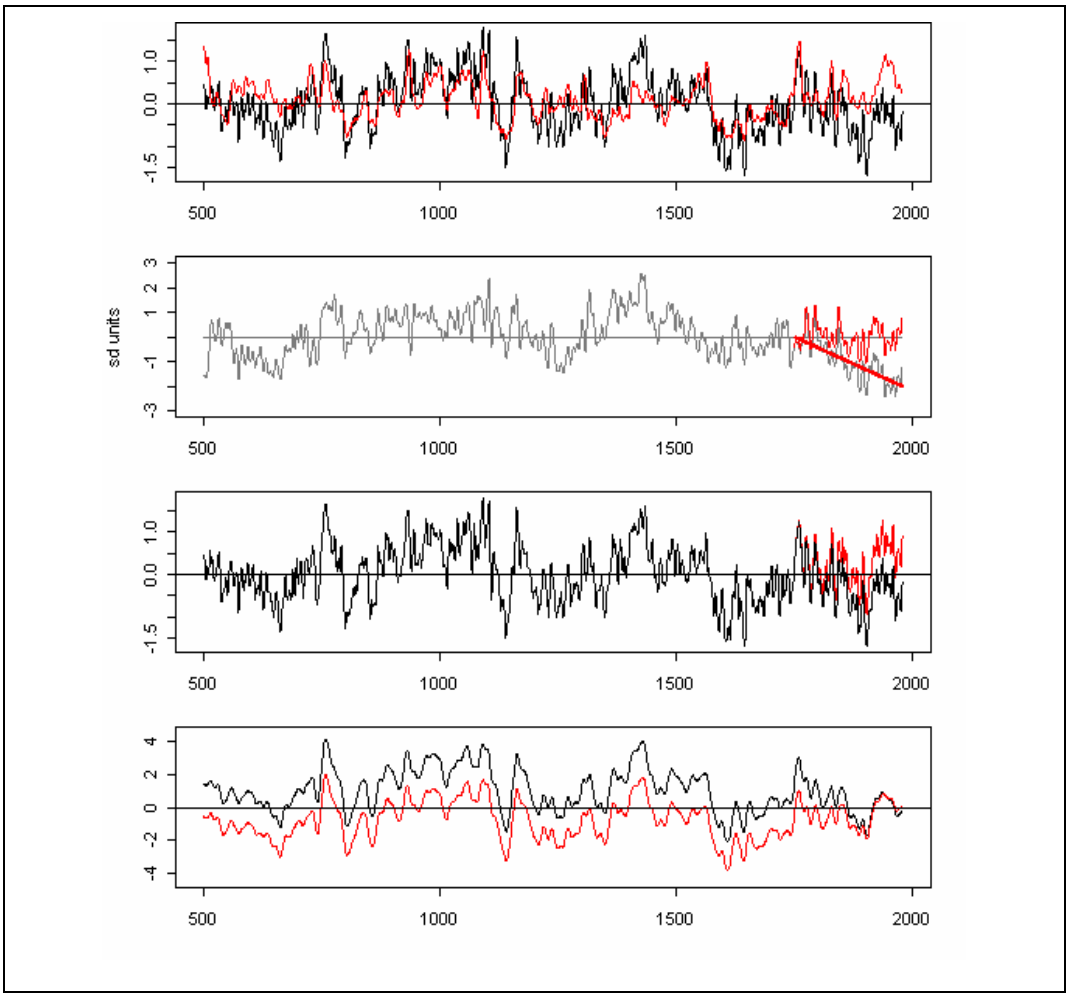


Figure 25. Top panel: RW (red) and MXD (black) chronologies. Second panel: (grey) residuals of MXD against RW after fit on 500-1750; red line - "trend" of residuals; red - "adjusted" residuals. Third panel - "adjusted" chronology after adding back "adjusted" residuals. Bottom panel – MXD chronologies centered on 1902-1980. Black - unadjusted; red - "adjusted". Source: Briffa et al 1992 (Clim Dyn) Fig 7. See www.climateaudit.org/?p=150 (March 2005)

In 2008, a Swedish specialist (Grudd 2008) did a major re-consideration of the Tornetrask data, which resulted in a chronology with a high MWP looking very much like the one that I had calculated in 2005.

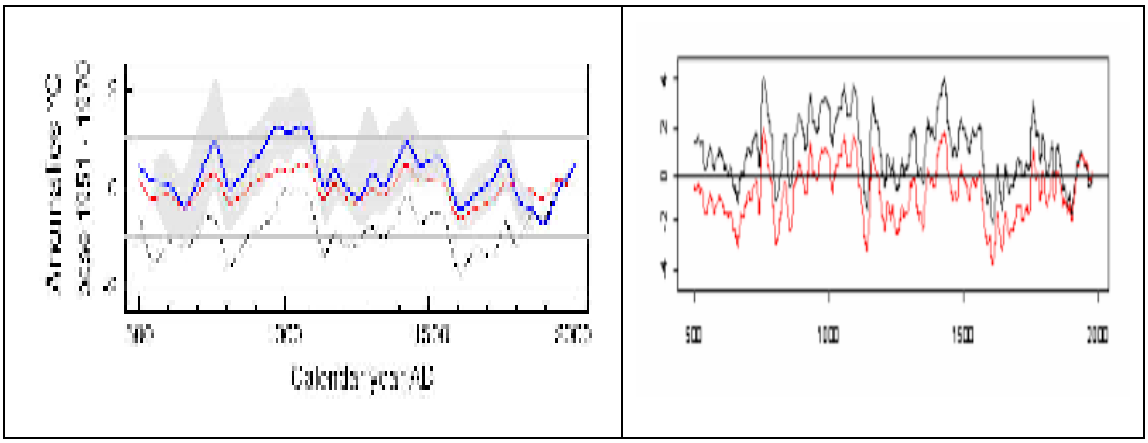


Figure 26. *Left - Grudd* 2008 Fig. 12 The thick blue curve is the new Tornetrask MXD low-frequency reconstruction of April–August temperatures, with a 95% confidence interval (grey shading). The thin red curve is from Briffa et al. (1992); the hatched curve is from Grudd et al. (2002) and based on TRW. Right –Black - "Unadjusted" Briffa version (red-adjusted) shown in bottom panel of above figure.

So for each of these critical data sets – bristlecones, Polar Urals, Tornetrask – each used over and over in virtually every study, there are major problems. And in each case, the direction of the reconsideration is toward re-establishing a MWP and undercutting the evidence relied on by IPCC for asserting relative 20th century warmth.

4) The "Divergence Problem"

At the presentation day to the NAS panel, one of the main issues was the divergence problem (an issue essentially passed over in the final report). In the only large population sample of tree ring chronologies from sites selected in advance to be "temperature sensitive", a collection of 387 sites collected by Schweingruber, some of which has been published by Briffa, there is a pronounced decline in both ring widths and density in the last half of the century – called the divergence problem at the NAS panel hearings. Briffa's density reconstruction, which does not go back to the MWP, has a sharp decline after 1960 to levels comparable to the early 19th century.

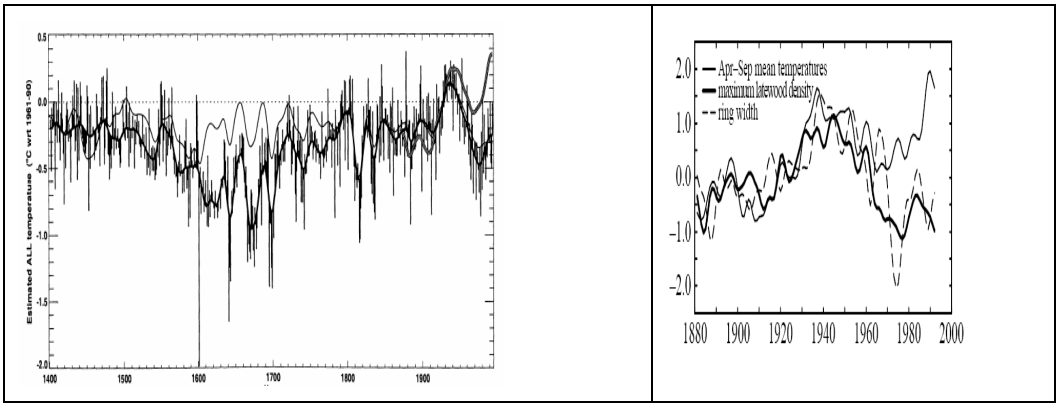


Figure 27. Left - Briffa et al 2001 reconstruction (left) from 387 temperature-sensitive sites; right – from Briffa et al 1998 (Nature). Heavy solid – MXD (used in Briffa et al 2001); dashed – RW; thin solid – temperature.

In commentary at Climate Audit, I've described Briffa's explanation for the late 20th century decline as nothing better than a "cargo cult" explanation.

> In the absence of a substantiated explanation for the decline, we make the assumption that it is likely to be a response to some kind of recent anthropogenic forcing. On the basis of this assumption, the pre-twentieth century part of the reconstructions can be considered to be free from similar events and thus accurately represent past temperature variability.

At the AGU 2007 Fall Meeting, there was a session on the "Divergence Problem" and the younger dendrochronologists realized that this sort of explanation was not acceptable. However the discussion in the NAS Panel report and IPCC AR4 maintains the unscientific approach used here by Briffa, unsurprising in the case of IPCC given that Briffa was lead author of this section. This cargo cult explanation is then used to justify ignore the "divergence problem" – results which would be interpreted by econometricians as refuting the hypothesis of a linear relationship between ring widths and increased temperature.

In the IPCC Third Assessment Report, they did worse than simply ignoring the problem. They deleted the declining portion after 1960, thereby giving a false sense of coherence between the proxies.

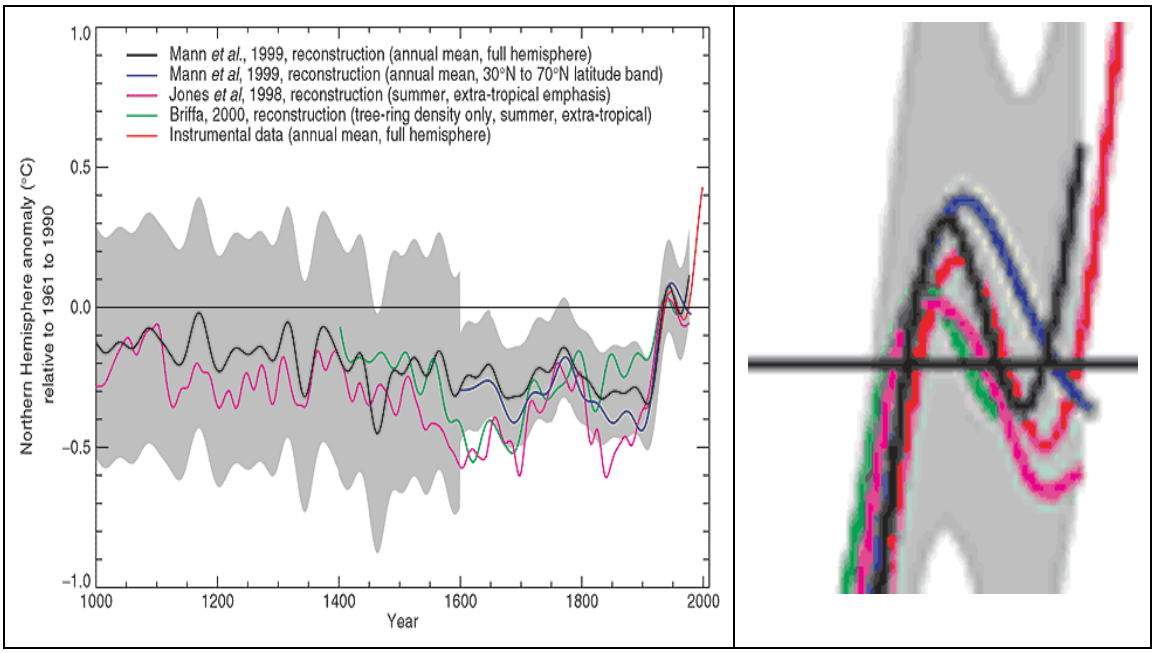


Figure 28. Left – IPCC TAR Spaghetti Graph. Right- blow-up of right hand portion. The divergent portion (after 1960) of the Briffa reconstruction (green) was deleted in IPCC TAR (green) 1960 and thus no visible "divergence". Similar truncation in AR4.

In AR4, as a reviewer, I asked them to restore the deleted portion. They refused saying that showing this information would be "inappropriate" (See IPCC WG1 chapter 6

Review Comments) and the downward late 20th century portion of the Briffa et al 2001 reconstruction was once again deleted in IPCC AR4.

5) Thompson's Tropical Ice Cores

Now to Ohio State's own hockey stick, courtesy of Lonnie Thompson, a celebrated and adventurous scientist. While some of my remarks may seem critical, I don't intend them to be seen as diminishing Thompson's accomplishments; indeed, I think that the suggestions that I make here would enhance his legacy.

First as an ironic point, what is called Dr Thompson's thermometer in the Gore movie isn't really. It's just the Michael Mann hockey stick spliced together with the CRU temperature series. So one reason that it seems to support the Mann hockey stick is that it is the Mann hockey stick.

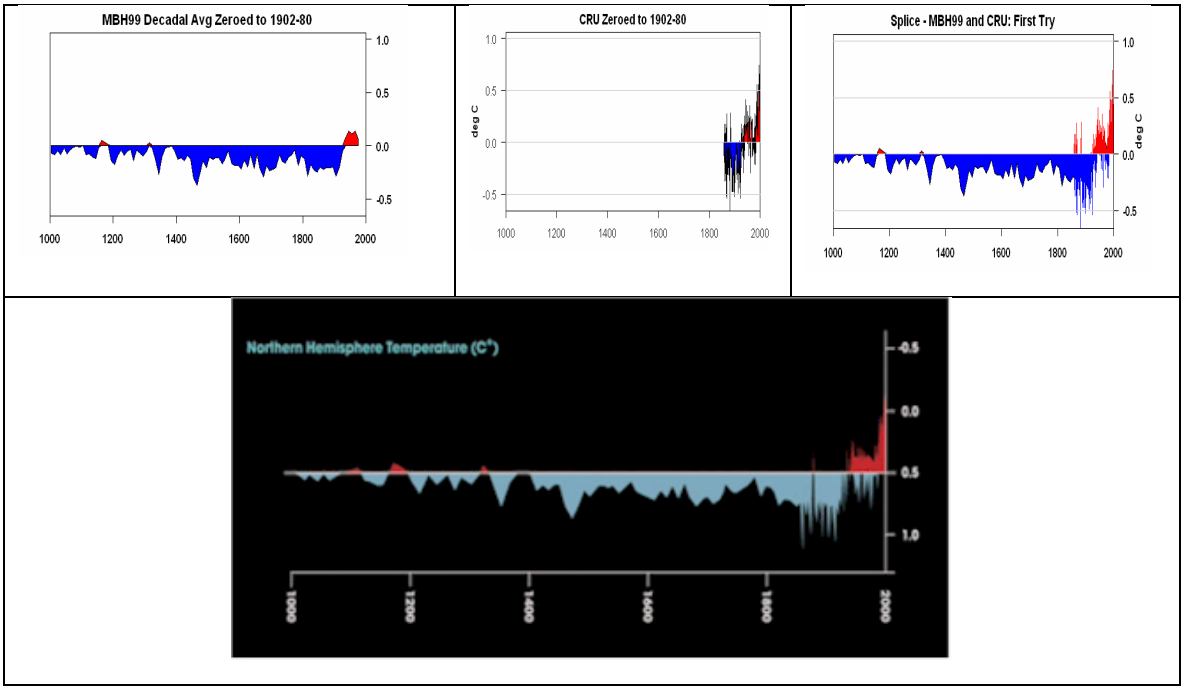


Figure 29. "Dr Thompson's Thermometer" in Inconvenient Truth ironically has nothing to do with ice cores, but is simply the Mann reconstruction spliced (decadally averaged –left) with the CRU instrumental record (middle), spliced as the right panel. Also see Thompson et al 2003 Fig 7 (Clim Chg) and CA.

Having said that, Thompson has produced his own hockey stick, which is shown in the middle panel here. It does have a HS shape, seemingly adding support to the HS reconstructions. A big problem for me here is that other δO18 series that seem just as relevant – here I show one from Mt Logan AK and another from Law Dome Antarctica - don't have a HS shape. The Mt Logan series surprisingly goes down rather sharply at the

start of the 20th century, while the Law Dome series has a noticeable MWP. These results have been attributed to regional circulation changes, but, if that explains the high altitude ice cores where the δO18 effect should be more evident, then how can this effect be excluded in consideration of tropical cores?

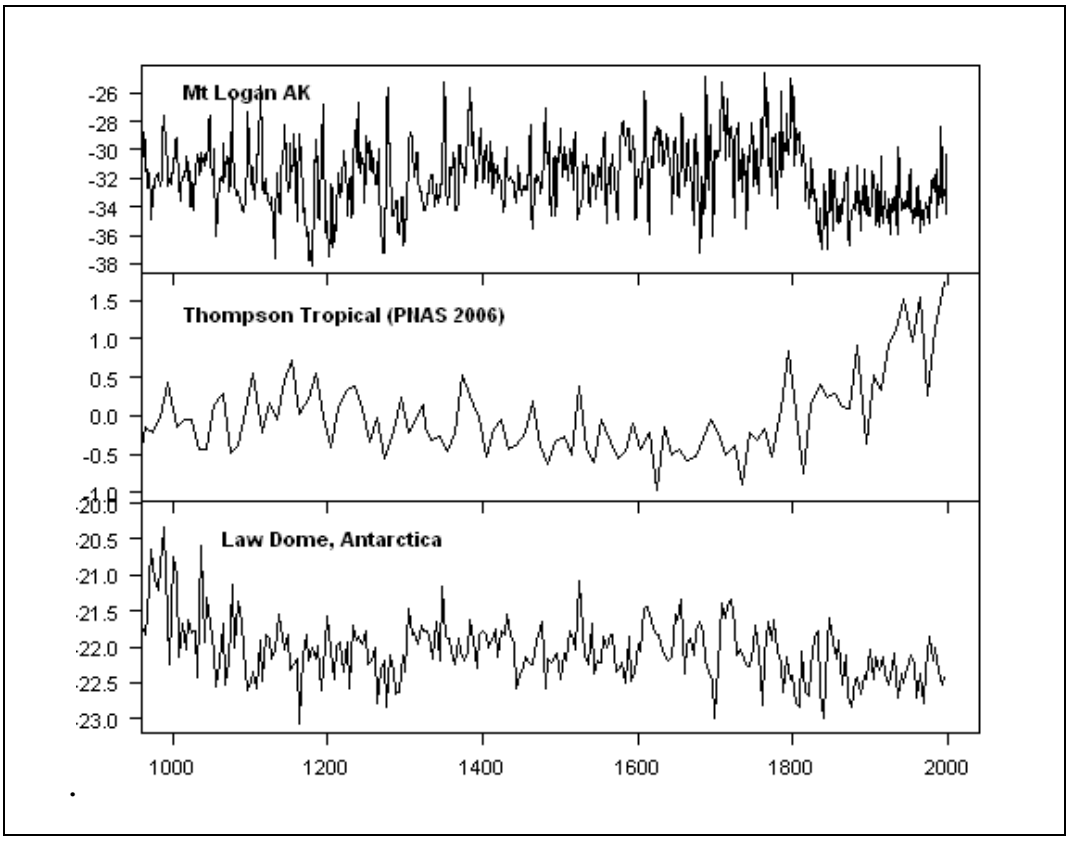


Figure 30. Middle – Dr Thompson's δO18 "Hockey Stick"(Thompson et al. 2003). Top - Mount Logan AK (Fisher et al. 2004); bottom - Law Dome, Antarctica δO18 (Jones & M. E. Mann 2004 from van Ommen data). If tropical δO18 shows global warming, then what do the high-latitude cores show?

In addition, there is an important "conundrum" associated with these tropical cores – which is debated in the specialist literature. In the tropics, the highest δO18 values occur in winter and the most depleted values in summer.

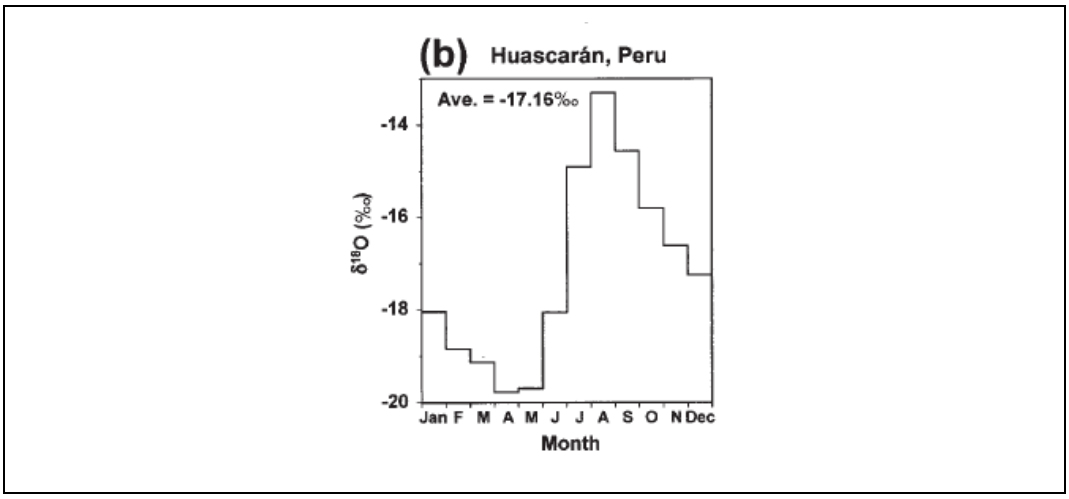


Figure 31. Seasonal δO18 at Huascaran from Figure 7b, (Thompson 2000). In the Andes and Himalayas, δO18 is high (enriched) in winter and low (depleted) in summer – the reverse of polar δO18.

This is the reverse of the polar situation where δO18 is interpreted as a temperature proxy. Important academic opinion has been that tropical δO18 is measuring the amount of summer precipitation. Hoffman et al (2003):

> At low latitudes, the relation between isotopes and temperature breaks down and the amount of precipitation becomes dominant.

Hardy et al (2003) stated:

> A strong relationship exists between Sajama δ18O and precipitation; both snowfall and net accumulation explain nearly half of the isotopic variance. In contrast, no significant association exists between air temperature and δ 18O.

Thompson has contested this interpretation, arguing that the tropical δO18 can still be interpreted as a temperature proxy. A third party is obviously not in a position to determine which side of the debate is correct. However, there is clearly no consensus on this, with, as far as I can tell, more opinion against Thompson's interpretation than for it. Until this issue is clearly resolved, I don't see how this can be introduced as evidence in the debate.

In this case, the problem is particularly intense because the largest contributor to the HS-ness of the Thompson average, Dasuopu, is the series most identified with an amount effect. Thompson's original publication of Dasuopu (Thompson et al. 2000) was entitled "A High-Resolution Millennial Record of the Asian Monsoon", interpreted Dasuopu δ O18 as an index of monsoon activity...

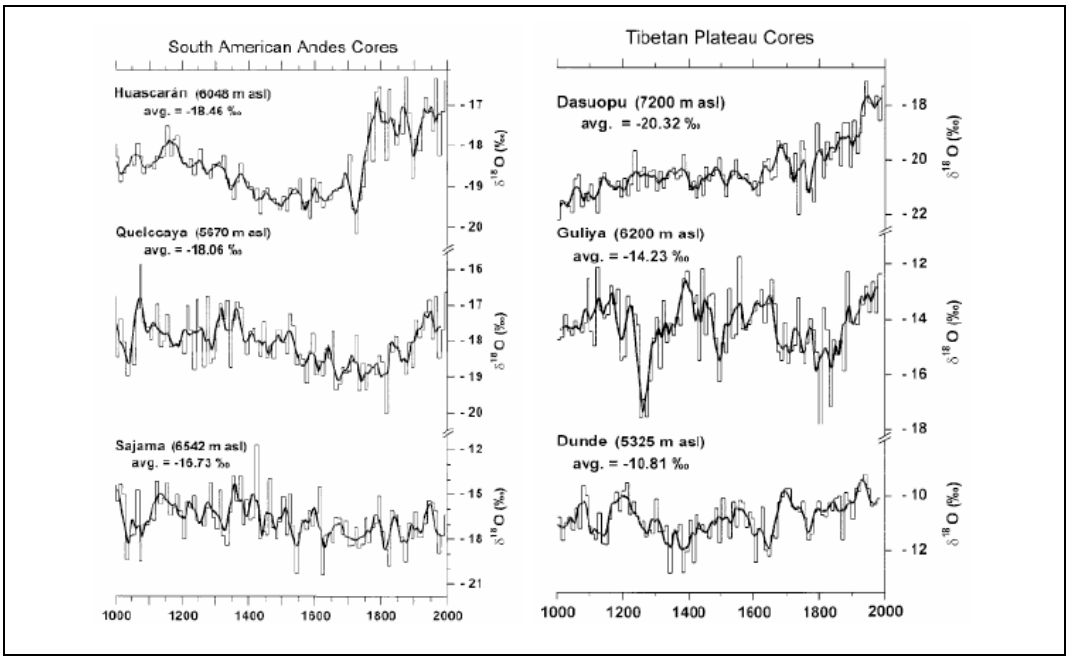


Figure 32. Excerpt from Thompson et al 2003. Dasuopu (top right) is the most HS-shaped series,

Interpretation of Thompson's data by third parties is further complicated by multiple inconsistent versions of key series appearing in peer reviewed literature. This graphic illustrates various versions of the Dunde ice core that have appeared in one publication or another. Doubtless there is an explanation for this and reconciliation is possible. But in order for that reconciliation to be verifiable and replicable, I see no alternative other than a complete archiving of every individual sample from the Dunde core – in this case results from about 10,000 samples.

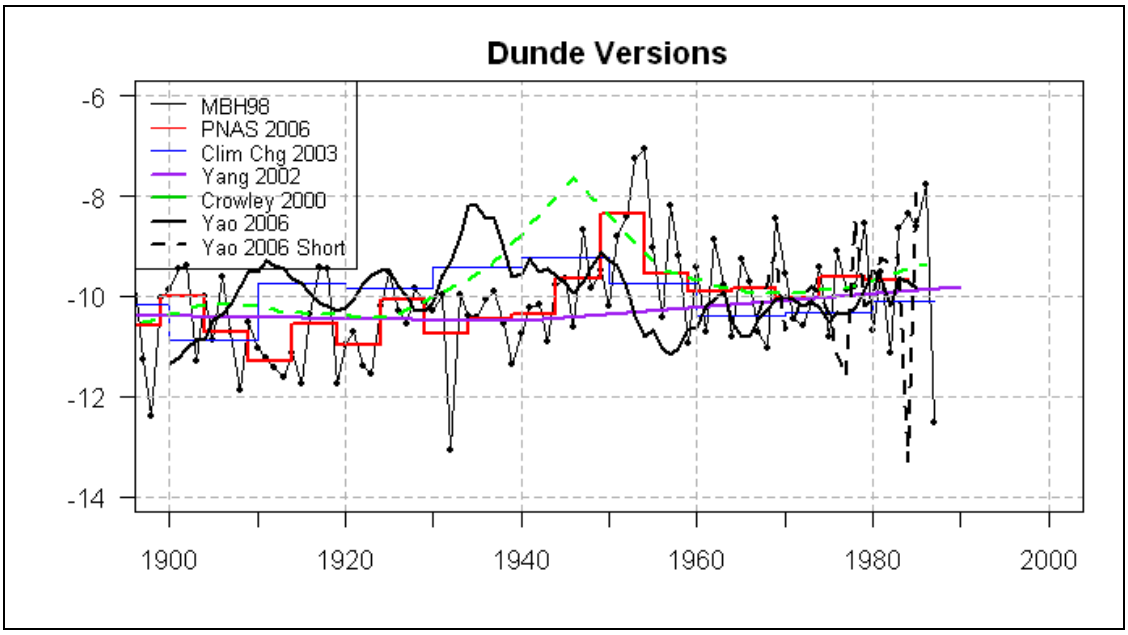


Figure 33. Different versions of Thompson's Dunde data from various peer reviewed studies (1- M. E. Mann, Bradley & Hughes 1998) (2- Thompson et al. 2006) (3- Thompson et al. 2003); (4- Yang et al. 2002); (5- Crowley & Lowery 2000); (6,7- Yao et al. 2006)

During the course of his career, Thompson has collected over 70,000 samples from Dunde, Guliya, Dasuopu, Huascaran, Sajama, Quelccaya, Kilimanjaro, Puruogangri and Bona-Churchill. A standard suite of sample measurements for ice core data includes: δ18O, δD, 3 dust sizes (coarse, middlings, fines), F, Cl, NO3, SO4, Na, NH4, K, Mg, Ca, conductivity plus annual accumulation. Some samples appear to have had only a subset of measurements. By contrast, the data archive for Thompson's data at http://www.ncdc.noaa.gov/paleo/ or journals is an inconsequential fraction of the population. For Dunde, the archive is merely 99 decadally averaged δO18 measurements, seemingly less than 1% of the potential archive.

I think that Thompson's accomplishments in collecting all this data both require and deserve a complete archive. A complete digital archive can be interpreted as the best way of ensuring that Thompson's legacy as a scientist be preserved. It would be tragic if this data were lost through some misadventure.

6) Glacier Recession

I would like to touch briefly on glacier recession, since this is much in the news. Obviously I can't make a full survey, but wish to draw two interesting articles to your attention. Schlüchter and Joerin have analysed organic material disgorged from retreating glaciers and concluded that glaciers in the past had sometimes receded considerably from present locations. In one of their articles, they provide a vivid rendering of Green Alps in Roman times, a time when archaeological evidence suggests traffic through passes that have been closed through the 20th century.

Figure 34. Left - modern view of a pass in the Alps (with glacier lines of 1922 and 1856); right - reconstructed view in Roman times of 2000 years ago Schlüchter and Jörin (2004). Also see http://www.climateaudit.org/?p=772

In the Venezuelan Alps, the changing content of sediments has been interpreted as evidence of the presence/absence of glaciers. In this case, the authors concluded that the glaciers did not exist in the MWP and that temperatures in the Venezuelan Alps declined 2.6 -4.3 deg C between the MWP and the Little Ice Age.

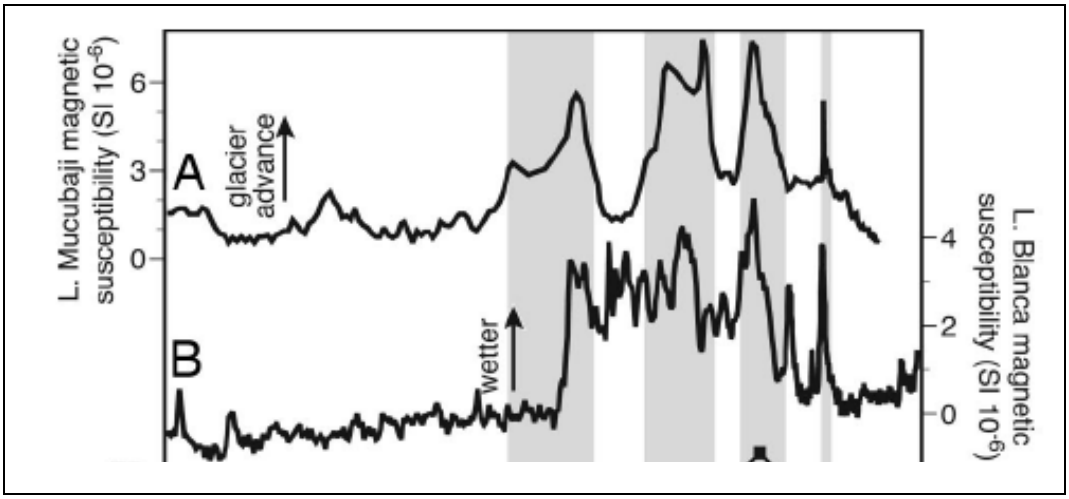


Figure 35. Polissar et al (2006) conclude that sediments evidence the absence of a glacier in the watershed before approximately A.D. 1100. Estimated LIA temperature decline of 2.6-4.3°C, greater than ~2°C inferred for Caribbean sea-surface temperatures. http://www.climateaudit.org/?p=704

7) Ocean Sediments

One of the promising lines of new evidence comes from ocean sediments. In the past, most of these cores have been very difficult to date with considerable variability from radiocarbon dating, making decadal analyses tenuous at best. In a recent analysis of sediments offshore Iceland, individual tephras have been attributed to well-dated volcanic eruptions enabling relatively close control of the dating. In this case, δO18 estimates of temperature show a dramatic decline from the MWP to the LIA, supporting the more informal concepts of the early authors.

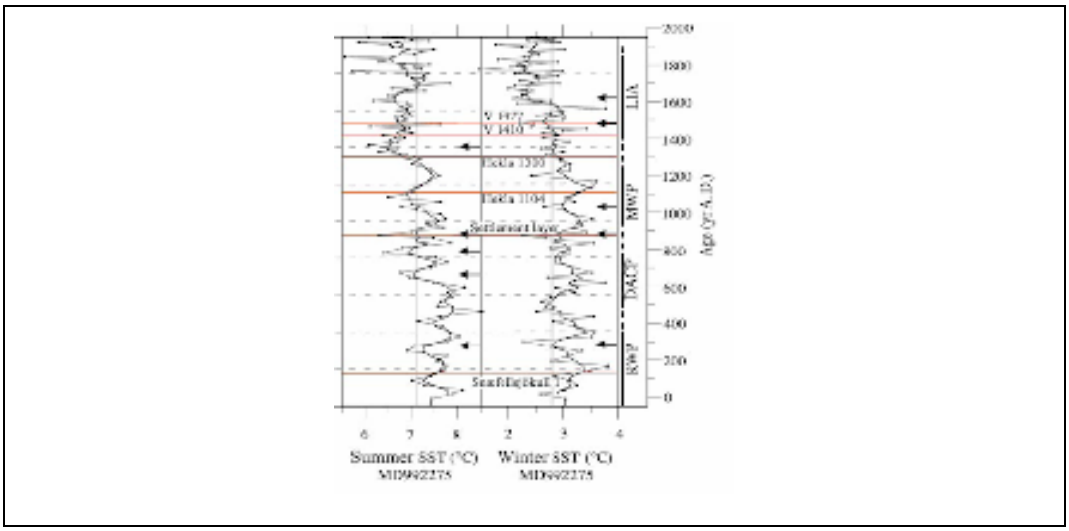


Figure 36. From Jiang et al (2005). Sediment is well-dated by tephra and show a noticeable MWP and LIA (concepts applied by the authors).

There are an increasing number of high-resolution ocean sediment records. Even in the tropical Pacific, far from northern Europe, there is evidence of a MWP, even in Antarctica. Here are 6 relatively high-resolution ocean sediment series that I showed at the AGU 2006 Fall Meeting (McIntyre 2006). Since then, interesting new series from Cariaco and Pigmy Basin, Gulf of Mexico have been published, also with a noticeable MWP - see CA, 2300, 2306 discussing Black et al (2007) and Richey et al (2007). Of the series shown below, only one was selected in Juckes et al (Juckes et al. 2007) – the one that doesn't have a MWP.

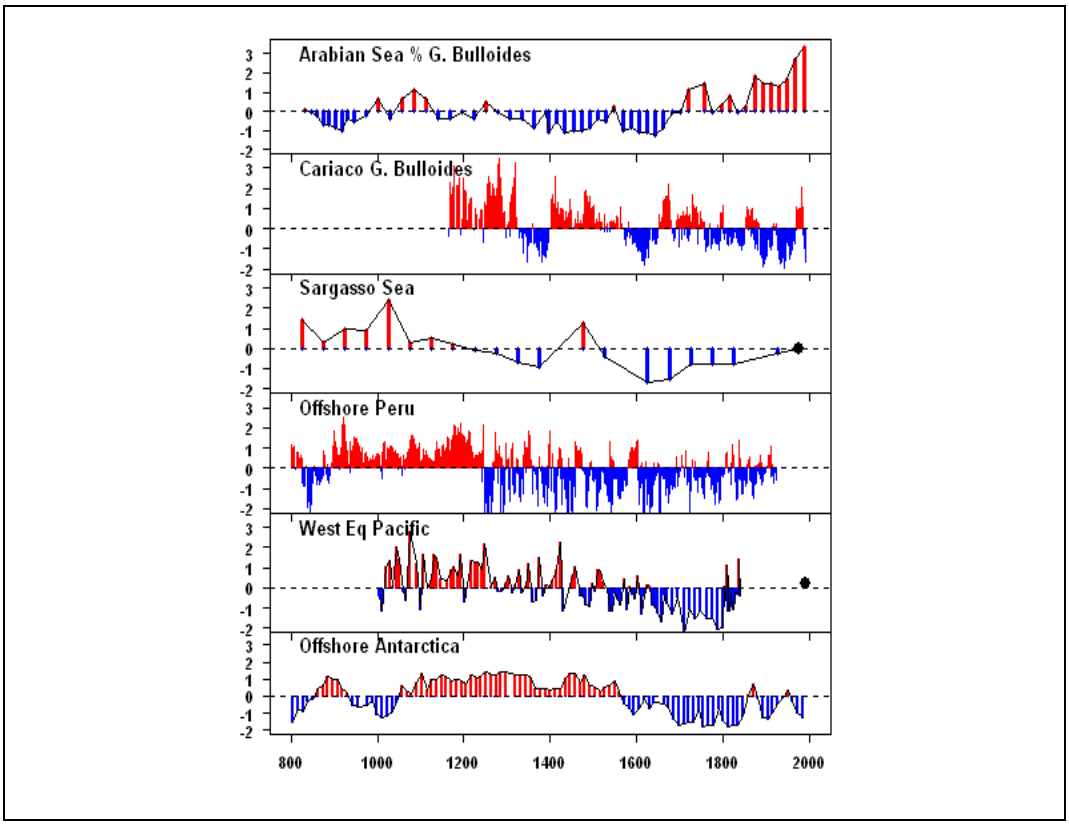


Figure 37. Six ocean sediment series. 1 - Arabian Sea G. Bulloides percentage (Gupta et al (2003); (Anderson et al. 2002); b- Cariaco G Bulloides (Black et al. 1999); c- Sargasso Sea SST (Keigwin 1996) ; d- offshore Peru lithics (Rein et al 2004); e- Pacific Warm Pool SST (Newton et al 2006); Bramsfield Basin, Antarctica from Khim et al (2002).

Alternative Spaghetti Graphs?

But can these proxies be put together in a comparable way to the studies in the spaghetti graphs? Sure they can. At the NAS panel in 2006, we presented the results of what we called apple-picking, showing that we could pick series that yielded a MWP period with just as good (or bad) statistical performance as the canonical studies ((McIntyre & McKitrick 2006)

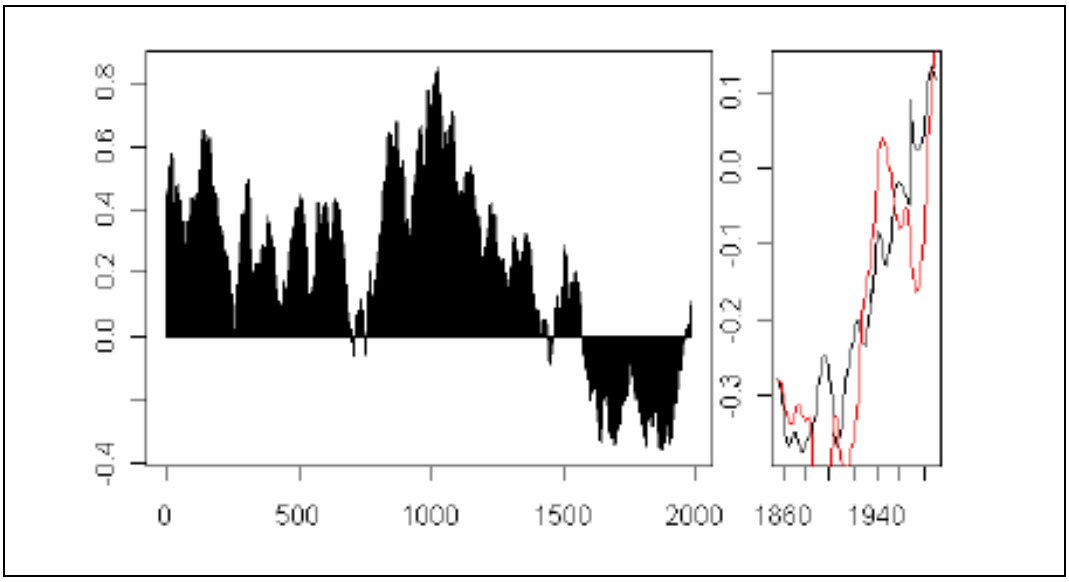


Figure 38. Pseudo-reconstruction submitted to NAS Panel in 2006, taking alternate selection of 8 proxies.

I wasn't completely sold on the idea that this meant anything – just as I wasn't sold on the opposite. Hu McCulloch and Crag Loehle have re-visited this problem in an interesting way (Loehle & McCulloch 2008). They collected non-tree ring series that authors had explicitly calibrated in temperature units. Remarkably, to my knowledge, this was the first such study ever done. All the other studies to a greater or lesser degree used series that had never been calibrated to temperature, but were merely said to be proxies, standardized and averaged – with calibration to temperature taking place after the averaging rather than before. Loehle and McCulloch neatly circumvent this by totally relying on calibration done by the primary authors. To some extent this is what people thought that Mann had done, but he hadn't.

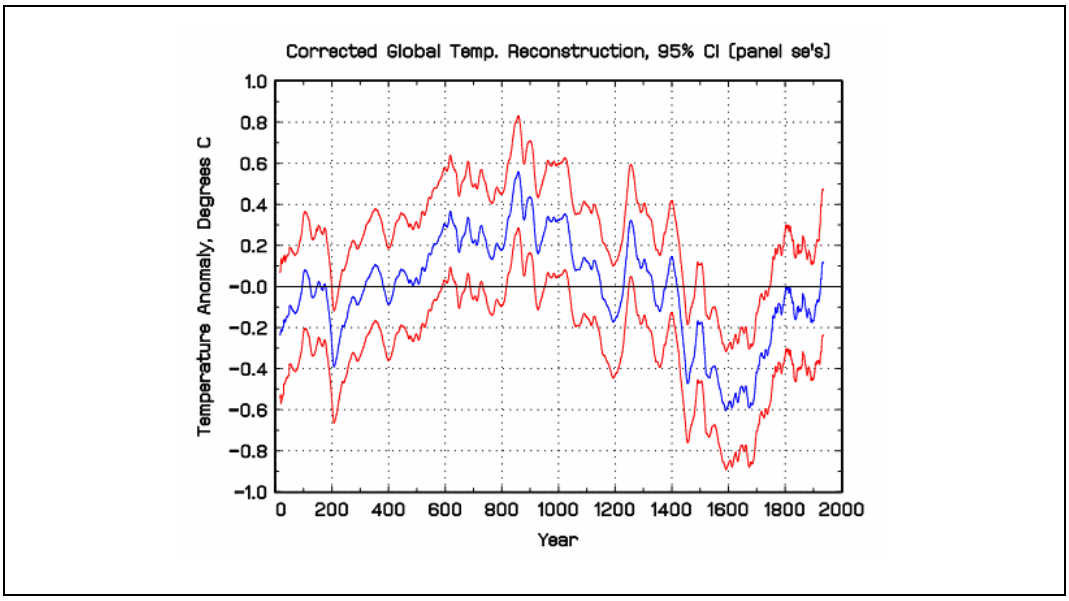


Figure 39. From Loehle and McCulloch 2008.

By doing this, they end up with a rather elevated MWP. The availability of their series tails off towards the present, as all too often ocean sediment series lack the top portion due to mechanical exigencies. However, in this limited case – where the proxy is already calibrated to temperature through a hypothesized physical relationship (and not through mere correlation), there may be an argument for splicing the instrumental record – a view that I verified with William Curry, one of the most eminent oceanographers.

But even if one doesn't fully subscribe to this approach, one is still left with the situation where there is no objective way to distinguish between one class of reconstructions with a warm medieval period and another class of reconstructions with a warm modern period, such as the comparison shown below in my 2006 AGU presentation (McIntyre 2006). If anything, the ones with the warmer modern period are more reliant on proxies known to be problematic – such as the Graybill bristlecones.

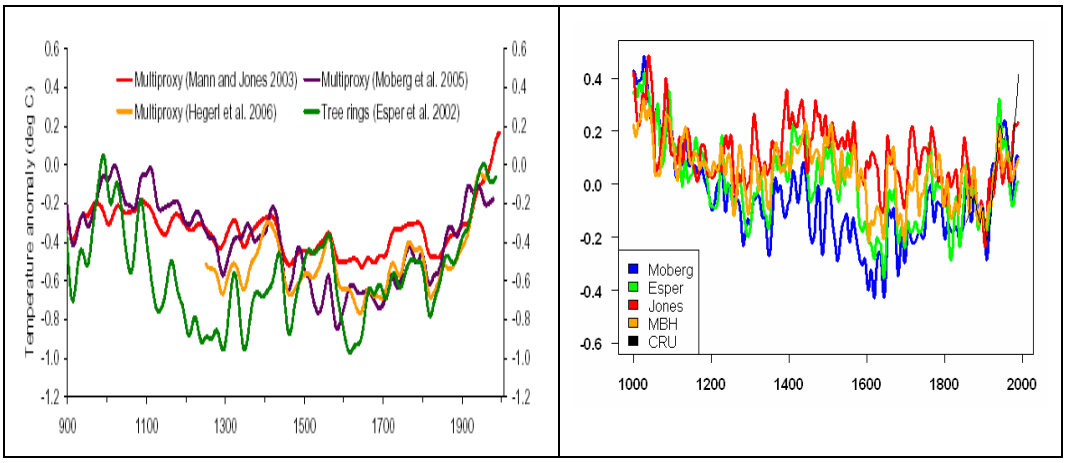


Figure 40. Left: NRC panel. Right: Variations of standard reconstructions using Polar Urals update instead of Yamal and Sargasso Sea SST instead of G Bulloides wind speed proxy and Yakutia instead of problematic bristlecones/foxtails

Conclusions

So where does that leave us?

In my opinion, there are serious and probably fatal problems with the main proxies used as supposed evidence against a warm MWP – the Graybill strip bark chronologies, Briffa's adjustment to the Tornetrask series, the inconsistency between Briffa's Yamal substitution and the updated Polar Urals series and so on. For every proxy that supposedly shows a MWP cooler than the present, there seems to be one that is just as good or better evidencing the opposite. For the California and Urals proxies so fundamental to the Hockey Stick, the ecological evidence is further evidence against the Graybill and Briffa chronologies being interpretable as temperature proxies.

The selection of proxies in studies displayed by IPCC seems to me to be biased against proxies with a warm MWP. IPCC itself does not carry out any independent due diligence of the type that might be expected in a prospectus. Further, in 2007, as in 2001, the authors involved in preparing the paleoclimate section were active parties in controversies and, in the end, IPCC Fourth Assessment Report strongly reflects their partisan point of view.

Is there a wider lesson here for engineers? We are often told that the "Science is settled". But engineers, of all people, know that, even if the "science is settled", the engineering work may have just begun. One would hardly derive the parameters for a chemical process from an article in Nature without an engineering feasibility study.

The most critical question in climate is the estimation of a parameter – is the sensitivity of climate to doubled CO2 1.5, 2.5 or 3.5 deg C? Or could it be 6 deg C or 0.6 deg C?

In some ways, the estimation of such parameters through the development of complicated computer models is reminiscent of activities carried out by engineers. One important difference is that climate scientists typically report their results in highly summarized form in journals like Nature, rather than in the 1000-page or 2000-page engineering studies that an aerospace engineering enterprise would produce.

Viewed from this perspective, a remarkable aspect of the climate debate has been the seeming inability of the climate science community to narrow confidence intervals on this estimate. In 1979, the Charney Report (National Research Council 1979) estimated the impact at 3 deg C with a 1.5 degree range either way. In 2007, IPCC AR4 estimates are virtually unchanged. With all the improvements in scientific knowledge and all the efforts of climate scientists over the years, why has the improvement of these confidence intervals proved so resistant? I don't know, but it's worth thinking about.

REFERENCES

AAAS, 2005. Letter to House Energy and Commerce Committee. Available at: http://www.aaas.org/news/releases/2005/0714letter.pdf.

Ababneh, L., 2006. Analysis of Radial Growth Patterns of Strip-Bark and Whole-Bark Bristlecone Pine Trees in the White Mountains of California: Implications iin Paleoclimatology and Archaeology of the Greeat Basin. *University of Arizona Ph.D. Thesis*.

Ababneh, L., 2007. Bristlecone pine (BCP) paleoclimatic model for archaeological patterns in the White Mountain of California. *Quaternary International*.

Anderson, D.M., Overpeck, J.T. & Gupta, A.K., 2002. Increase in the Asian southwest monsoon during the past four centuries. *Science*, 297(5581), 596-9.

Black, D.E. et al., 2007. An 8-century tropical Atlantic SST record from the Cariaco Basin: Baseline variability, twentieth-century warming, and Atlantic hurricane frequency. *Paleoceanography*, 22(4).

Black, D.E. et al., 1999. Eight Centuries of North Atlantic Ocean Atmosphere Variability. *Science*, 286(5445), 1709.

Briffa, K.R., 2000. Annual climate variability in the Holocene: interpreting the message of ancient trees. *Quaternary Science Reviews*, 19(1-5), 87-105.

Briffa, K.R. et al., 1992. Fennoscandian summers from ad 500: temperature changes on short and long timescales. *Climate Dynamics*, 7(3), 111-119.

Briffa, K.R. et al., 1995. Unusual twentieth-century summer warmth in a 1, 000-year temperature record from Siberia. *Nature*, 376(6536), 156-159.

Brunstein, C., 2006. *Growth-Form Characteristics of Ancient Rocky Mountain Bristlecone Pines (Pinus aristata), Colorado*,

Chronicle of Higher Education, 2006. A Scientific Graph Stands Trial. Available at: http://chronicle.com/live/2006/09/hockey_stick/.

Crowley, T.J., 1996. Remembrance of Things Past: Greenhouse Lessons from the Geologic Record. *Consequences*, 2(1), 3-12.

Crowley, T.J. & Lowery, T.S., 2000. How Warm Was the Medieval Warm Period? *AMBIO: A Journal of the Human Environment*, 29(1), 51-54.

Esper, J., Cook, E.R. & Schweingruber, F.H., 2002. Low-Frequency Signals in Long Tree-Ring Chronologies for Reconstructing Past Temperature Variability. *Science*, 295(5563), 2250-2253.

Ferson, W.E., Sarkissian, S. & Simin, T.T., 2003. Spurious Regressions in Financial Economics? *The Journal of Finance*, 58(4), 1393-1414.

Fisher, D.A. et al., 2004. Stable Isotope Records from Mount Logan, Eclipse Ice Cores and Nearby Jellybean Lake. Water Cycle of the North Pacific Over 2000 Years and Over Five Vertical Kilometres: Sudden Shifts and Tropical Connections. *Géographie physique et Quaternaire*, 58(2-3), 337-352.

Gore, A., 2006. *An Inconvenient Truth. The Planetary Emergency of Global Warming and What We Can Do About It*, Rodale Press, Emmaus, Pennsylvania (United States).

Graybill, D.A. & Idso, S.B., 1993. Detecting the aerial fertilization effect of atmospheric CO 2 enrichment in tree-ring chronologies. *Global Biogeochemical Cycles*, 7(1), 81-95.

Greene, C.A., 2000. I am not, nor have I ever been a member of a data-mining discipline. *Journal of Economic Methodology*, 7(2), 217-230.

Grudd, H., 2008. Torneträsk tree-ring width and density AD 500–2004: a test of climatic sensitivity and a new 1500-year reconstruction of north Fennoscandian summers. *Climate Dynamics*, (DOI 10.1007/s00382-007-0358-2).

Gupta, A.K., Anderson, D.M. & Overpeck, J.T., 2003. Abrupt changes in the Asian southwest monsoon during the Holocene and their links to the North Atlantic Ocean. *Nature*, 421(6921), 354-7.

Hardy, D.R., Vuille, M. & Bradley, R.S., 2003. Variability of snow accumulation and isotopic composition on Nevado Sajama, Bolivia. *J. Geophys. Res*, 108, 4693.

Hendry, D.F., 1980. Econometrics: Alchemy or Science? *Economica*, 47, 387-406.

Hoffmann, G. et al., 2003. Coherent isotope history of Andean ice cores over the last century. *Geophys. Res. Lett*, 30(4), 1179.

Houghton, J., 2005. Sir John Houghton testifying before the Senate Energy and Natural Resources Committee. Available at: http://www.creationcare.org/files/houghton_qa.pdf.

House Energy and Commerce Committee, 2005a. Letter to Michael Mann. Available at: http://energycommerce.house.gov/reparchives/108/Letters/062305_Mann.pdf.

House Energy and Commerce Committee, 2005b. *Questions Surrounding the 'Hockey Stick' Temperature Studies: Implications for Climate Change Assessments*, Available at:

http://energycommerce.house.gov/reparchives/108/Hearings/07192006hearing1987/hearing.htm.

Hughes, M.K. & Funkhouser, G., 2003. Frequency-Dependent Climate Signal in Upper and Lower Forest Border Tree Rings in the Mountains of the Great Basin. *Climatic Change*, 59(1), 233-244.

International Panel on Climate Change, 2001. Climate Change 2001: The Scientific Basis.

International Panel on Climate Change, 2007a. *Climate Change 2007: The Physical Basis*,

International Panel on Climate Change, 2007b. *Climate Change 2007: The Physical Basis. Review Comments and Responses*,

Jiang, H. et al., 2005. Evidence for solar forcing of sea-surface temperature on the North Icelandic Shelf during the late Holocene. *Geology*, 33(1), 73-76.

Jones, P.D. et al., 1998. High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures. *The Holocene*, 8(4), 455.

Jones, P.D. & Mann, M.E., 2004. Climate over past millennia. *Rev. Geophys*, 42(2).

Juckes, M.N. et al., 2007. Millennial temperature reconstruction intercomparison and evaluation, Clim. *Climate of the Past*, 3, 591-609.

Keigwin, L.D., 1996. The Little Ice Age and Medieval Warm Period in the Sargasso Sea. *Science*, 274(5292), 1503.

Kerr, R.A., 2005. *GLOBAL WARMING: Millennium's Hottest Decade Retains Its Title, for Now*,

Khim, B.K. et al., 2002. Unstable Climate Oscillations during the Late Holocene in the Eastern Bransfield Basin, Antarctic Peninsula. *Quaternary Research*, 58(3), 234-245.

Lamb, H., 1965. The Early Medieval Warm Epoch and Its Sequel. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 13-37.

Loehle, C. & McCulloch, J.H., 2008. Correction to: A 2000-Year Global Temperature Reconstruction Based on Non-Tree Ring Proxies. *Energy & Environment*, 19(1), 93-100.

Mann, M.E., Bradley, R.S. & Hughes, M.K., 1998. Global-scale temperature patterns and

climate forcing over the past six centuries. *Nature*, 392, 779-787.

Mann, M., 2005. Het antwoord van Mann. Available at: http://www.natutech.nl/00/nt/nl/49/nieuws/2299/index.html.

McIntyre, S., 2006. The Impact of NRC Recommendations on Climate Reconstructions. *American Geophysical Union, Fall Meeting 2006, abstract# U11B-05*. Available at: http://www.climateaudit.org/pdf/agu06.ppt.

McIntyre, S. & McKitrick, R., 2003. Corrections to the Mann et al.(1998) proxy data base and northern hemispheric average temperature series. *Energy and Environment*, 14(6), 751–771.

McIntyre, S. & McKitrick, R., 2005. Hockey sticks, principal components, and spurious significance. *Geophysical research letters*, 32(2), 3710-3710.

McIntyre, S. & McKitrick, R., 2006. Presentation to the National Academy of Sciences Expert Panel, "Surface Temperature Reconstructions for the Past 1,000-2,000 Years." : Supplementary Comments. Available at: http://www.climateaudit.org/pdf/NAS-followup-M&M.pdf.

Millar, C.I. et al., 2006. Late Holocene forest dynamics, volcanism, and climate change at Whitewing Mountain and San Joaquin Ridge, Mono County, Sierra Nevada, CA, USA. *Quaternary Research*, 66(2), 273-287.

National Academy of Sciences, 2005. Letter to House Energy and Commerce Committee. Available at: http://www.realclimate.org/Cicerone_to_Barton.pdf.

National Research Council, 1979. *Carbon Dioxide and Climate: A Scientific Assessment* , Available at: http://www.atmos.ucla.edu/%7Ebrianpm/download/charney_report.pdf.

National Research Council, 2006. *Surface Temperature Reconstructions for the Last 2,000 Years* , Available at: http://books.nap.edu/openbook.php?isbn=0309102251.

Naurzbaev, M.M., Hughes, M.K. & Vaganov, E.A., 2004. Tree-ring growth curves as sources of climatic information. *Quaternary Research*, 62(2), 126-133.

Newton, A., Thunell, R. & Stott, L., 2006. Climate and hydrographic variability in the Indo-Pacific Warm Pool during the last millennium. *Geophys. Res. Lett*, 33.

North, G., 2006. Presentation by Gerald North to Dessler Seminar, Texas A&M. Available at: http://www.met.tamu.edu/people/faculty/dessler/NorthH264.mp4.

Polissar, P.J. et al., 2006. Solar modulation of Little Ice Age climate in the tropical Andes. *Proceedings of the National Academy of Sciences*, 103(24), 8937-8942.

Regalado, A., 2005. In Climate Debate, The 'Hockey Stick' Leads to a Face-Off. *Wall Street Journal*.

Rein, B., Lückge, A. & Sirocko, F., 2004. A major Holocene ENSO anomaly during the Medieval period. *Geophys. Res. Lett*, 31.

Richey, J.N. et al., 2007. 1400 yr multiproxy record of climate variability from the northern Gulf of Mexico. *Geology*, 35(5), 423-426.

Schiermeier , Q., 2005. Past climate comes into focus but warm forecast stays put. *Nature*, 433, 562.

Schlüchter, C. & Jörin, U., 2004. Holz- und Torffunde als Klimaindikatoren ALPEN OHNE. *DIE ALPEN*, 6/2004, 34-47.

Shiyatov, S.G., 1995. Reconstruction of climate and the upper timberline dynamics since AD 745 by tree-ring data in the Polar Ural Mountains. *International Conference on Past, Present and Future Climate. Helsinki: Publication of the Academy of Finland*, 144-147.

Thompson, L.G., 2000. Ice core evidence for climate change in the Tropics: implications for our future. *Quaternary Science Reviews*, 19(1-5), 19-35.

Thompson, L.G. et al., 2006. Abrupt tropical climate change: Past and present. *Proceedings of the National Academy of Sciences*, 103(28), 10536.

Thompson, L.G. et al., 2003. Tropical Glacier and Ice Core Evidence of Climate Change on Annual to Millennial Time Scales. *Climatic Change*, 59(1), 137-155.

Thompson, L.G. et al., 2000. A High-Resolution Millennial Record of the South Asian Monsoon from Himalayan Ice Cores. *Science*, 289(5486), 1916.

UCAR, 2005. Media Advisory: The Hockey Stick Controversy. New Analysis Reproduces Graph of Late 20th Century Temperature Rise. Available at: http://www.ucar.edu/news/releases/2005/ammann.shtml.

Wahl, E.R. & Ammann, C.M., 2007. Robustness of the Mann, Bradley, Hughes reconstruction of Northern Hemisphere surface temperatures: Examination of criticisms based on the nature and processing of proxy climate evidence. *Climatic Change*, 85(1), 33-69.

Wegman, E.J., Scott, D.W. & Said, Y.H., 2006. Ad Hoc Committee Report on the« Hockey Stick» Global Climate Reconstruction. *A Report to Chairman Barton, House Committee on Energy and Commerce and to Chairman Whitfield, House Subcommittee on Oversight and Investigations: Paleoclimate Reconstruction*.

Wegman, E., 2006. Response of Dr. Edward Wegman to Questions Posed by the Honorable Bart Stupak . Available at: http://www.uoguelph.ca/~rmckitri/research/StupakResponse.pdf.

Yang, B. et al., 2002. General characteristics of temperature variation in China during the last two millennia. *Geophysical Research Letters*, 29(9), 38-1.

Yao, T. et al., 2006. 18O records from Tibetan ice cores reveal differences in climatic changes. *Annals of Glaciology*, 43(1), 1-7.

Yule, G.U., 1926. Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series. *Journal of the Royal Statistical Society*, 89(1), 1-63.