WASHINGTON ROUNDTABLE
ON SCIENCE & PUBLIC POLICY

The Hockey Stick Debate: Lessons in Disclosure And Due Diligence

By Stephen McIntyre and Ross McKitrick

GEORGE C. Marshall INSTITUTE

George C. Marshall Institute

The George C. Marshall Institute, a nonprofit research group founded in 1984, is dedicated to fostering and preserving the integrity of science in the policy process. The Institute conducts technical assessments of scientific developments with a major impact on public policy and communicates the results of its analyses to the press, Congress and the public in clear, readily understandable language. The Institute differs from other think tanks in its exclusive focus on areas of scientific importance, as well as a Board whose composition reflects a high level of scientific credibility and technical expertise. Its emphasis is public policy and national security issues primarily involving the physical sciences, in particular the areas of missile defense and global climate change.

The Washington Roundtable on Science and Public Policy

The Washington Roundtable on Science and Public policy is a program of the George C. Marshall Institute. The Roundtable examines scientific questions that have a significant impact on public policy and seeks to enhance the quality of the debate on the growing number of policy decisions that look to science for their resolution.

The opinions expressed during Roundtable discussions do not necessarily represent those of the Marshall Institute or its Board of Directors. Additional copies of this transcript may be ordered by sending $7.00 postage paid to:

The George Marshall Institute
1625 K Street, NW Suite 1050
Washington, D.C. 20006
Phone: 202/296-9655
Fax: 202/296-9714
E-mail: info @marshall.org
Website: www.marshall.org

The Hockey Stick Debate: Lessons in Disclosure and Due Diligence

by

Stephen McIntyre and Ross McKitrick

The George Marshall Institute
Washington, D.C.

Stephen McIntyre has worked in mineral exploration for 30 years, much of that time as an officer or director of several public mineral exploration companies. He has also been a policy analyst at both the governments of Ontario and of Canada.

Ross McKitrick is an Associate Professor in the Economics Department at the University of Guelph, Ontario, and a Senior Fellow of the Fraser Institute in Vancouver, B.C. He specializes in the application of economic analysis to environmental policy design and climate change.

The Hockey Stick Debate: Lessons in Disclosure and Due Diligence*

Stephen McIntyre and Ross McKitrick
May 11, 2005

Jeff Kueter: Good afternoon everyone and welcome to this afternoon's discussion on climate science and policy. I am Jeff Kueter, the President of the George Marshall Institute. It is my great pleasure to co-host this luncheon talk today with two of our good colleagues from north of the border who have done just tremendous work on exposing the process that goes into scientific publications. If you take nothing else away from what they tell you today, I hope that you have a greater appreciation for how that process works and its weaknesses as well as its strengths. It is my great pleasure now to introduce our co-host for today, our good colleague Myron Ebell from the Competitive Enterprise Institute.

Myron Ebell: Thanks, Jeff. I am pleased again to work with the Marshall Institute to put together one of these briefings. The Cooler Heads Coalition, as most of you know, has put together a number of briefings over the years on the science, economics and politics of global warming. In fact, I now believe Ross McKitrick has done more of these than any other individual on a wider variety of topics. So I am very pleased that Ross and Steve McIntyre can be here today. This is their second Cooler Heads and Marshall presentation on the hockey stick.

Let me just briefly say that Steve McIntyre is an independent person on these issues, because he is in the minerals exploration and investing business. He has a degree from the University of Toronto and another from Oxford and he will explain how he got interested in the hockey stick from the viewpoint of a person who has to scrutinize very carefully the claims of people who want you to invest your money in their project. He found out that you have to be very careful. Ross McKitrick is associate professor of economics at the University of Guelph in Ontario and his degrees are from the University of British Columbia. He has written a number of interesting things on environmental economics, particularly related to global warming, the last few years and is the author, with Christopher Essex, of *Taken by Storm*, an excellent book on global warming which also won the Canadian Donner Prize as the "Notable Book of the Year" a couple of years ago. I urge

* The views expressed by the authors are solely those of the authors and may not represent those of any institution with which they are affiliated.

you to look at their website and also to look at the *Taken by Storm* website. Please join me in welcoming Steve and Ross.

Ross McKitrick: Thank you all for coming. I should warn you before I start that Steve and I have about six hours worth of material to present. We have whittled it down as best we can, but some sections we are going to skip pretty quickly over some material that some of you may want to know in all the technical details. We are more than happy to fill in afterwards any technical details that we skipped over, but we have tried to tailor this for the general audience and address the several levels that catch people's interest on this topic.

So in summary for today, we will give an explanation and an update of the hockey stick debate. We want to widen the coverage a little bit to include the whole hockey team, and I will explain what that means. (With the hockey strike, Canadians are trying to find substitutes.) An underlying theme all the way through is that science is being used to make public policy decisions. We want people to think about the need for higher standards of disclosure and due diligence when science is driving public policy, so there is sort of a policy hook in the background of all of this.

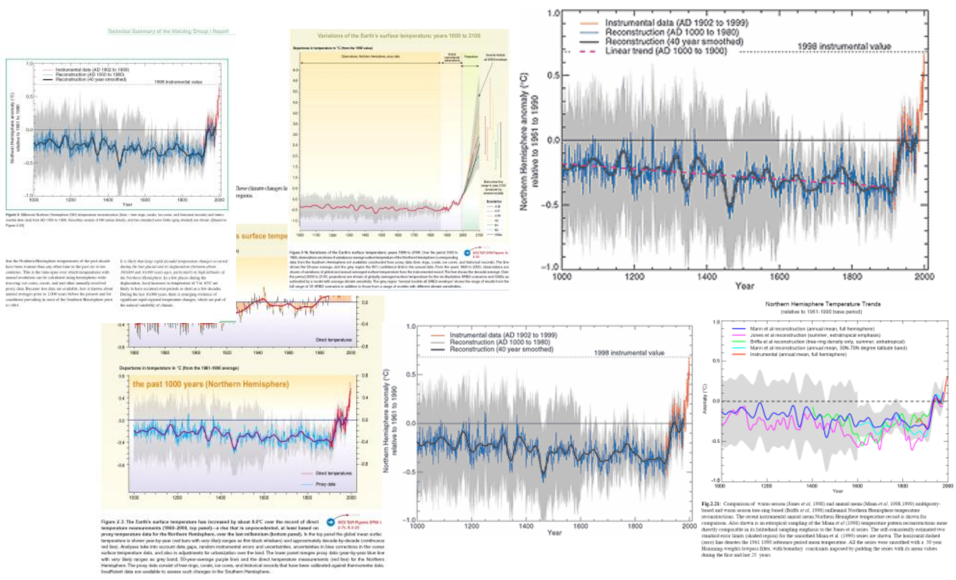


Figure 1

Now I suspect I don't really need to explain what the hockey stick is, but it is a graph that was essential to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). The Third Assessment Report concluded, among other things, that humans are causing climate change of a magnitude exceeding that observed in the past thousand years. The people that put the report together obviously felt the hockey stick was quite important for the purpose of conveying their message because it appears and reappears many times throughout the report and through related summary documents that were produced (Figure 1). It is very prominent in the Summary for Policymakers, in the Technical Summary, in Chapter 2 twice (figures 2.20 and 2.21), and in the Synthesis Report where again it appears twice. It is the basis for the claim that temperatures in the latter half of the twentieth century were unprecedented.

I am taking some pains to emphasize how important the hockey stick graph was to the IPCC because one of the responses that Steve and I have gotten since our papers have come out is, “Why are you so hung up on the hockey stick? It was never that important. It was sort of peripheral to the whole process and there is lots of evidence in the report.” There is indeed lots of evidence in the report, but not everything was selected for so much emphasis. To give you an example, not only was the hockey stick used repeatedly, but it is also very visually prominent.

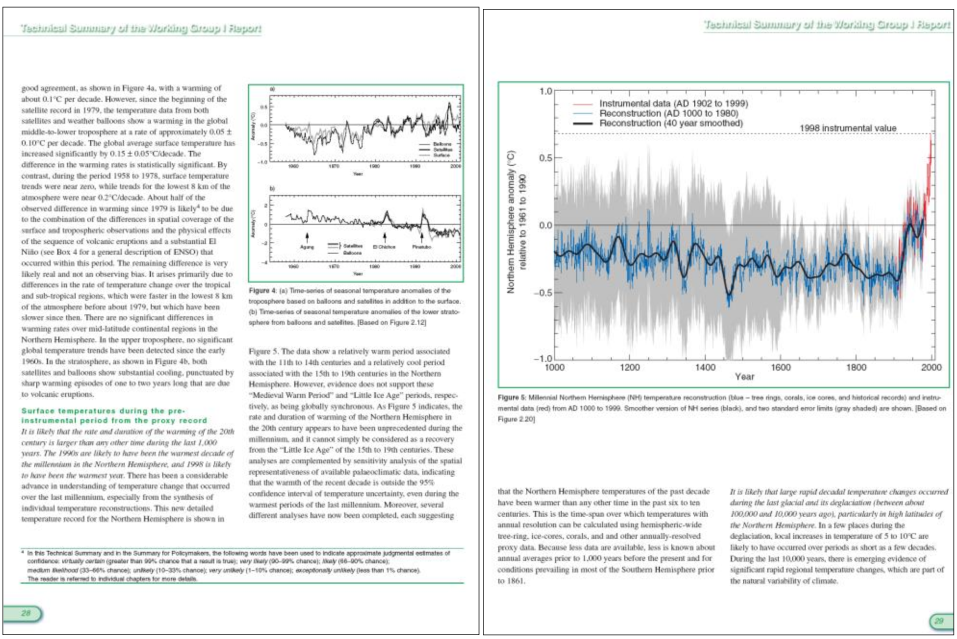


Figure 2

The left page shown on Figure 2 is an extract from the Technical Summary. If you look closely there is a data series in the top panel that is very important for the current debates over global warming. It is a temperature series that is derived from weather satellite readings of microwave emissions in the troposphere. It has attracted a lot of attention; there is a huge amount of literature about it and some very pertinent debates are going on to this day over what the data series means. It is there in the Technical Summary, but if you didn't know how to find it, it wouldn't really catch your eye and it is overlaid there with the surface temperature series as well as a balloon record. But nevertheless, it is there in the Technical Summary. But notice the facing page. If you are going through this quickly, only one of those diagrams is really going to catch your eye. What strikes me is not only how often the hockey stick appears, but when it appears, it is very large, it is in full color, and it is prominently displayed. That is where the deliberate editorial decisions really play a role; this is an issue of emphasis and that is why I think it is important that people spend a lot of time thinking about the hockey stick and what it means and how much we can trust it.

As far as the message goes, I can say that in Canada the point was not lost. When the Third Assessment Report came out, Henry Hengeveld, Canada's recently retired Chief Climate Science Advisor to the federal government, was reported in the papers saying,

> "This gives a fairly clear signal that this isn't just a future issue, it's happening now." ***"Among the strongest evidence is the fact that the past century has likely been the warmest in the Northern Hemisphere in the past millennium," he said. "Not only that, the 1990s ranked as the warmest decade of the millennium, and 1998 was the warmest year of the millennium*** in the Northern Hemisphere, which is where most of the data have been acquired."[1]

So again, notice the choice of emphasis here on the hockey stick graph.

In terms of why that should have been played up so much, it is helpful to go back to the earlier IPCC report and look at what we might call the Medieval Warm Period problem. Of course, it is not a problem, but I will call it that for now. To give you an idea of what we are talking about, Figure 3 shows a graph showing the location, the altitude of the upper tree line in the polar Ural Mountain range during the last 1,150 years.

[1] *Globe and Mail* January 22, 2001 (emphasis added).

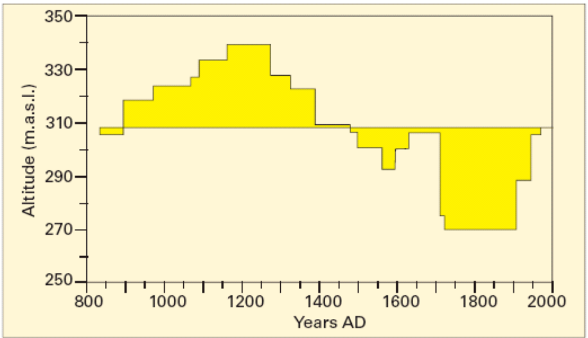


Fig. 1. Altitudinal displacement of the upper treeline in the Polar Ural Mountains during the last 1150 years.

Figure 3

The tree line migrated up to quite a high altitude compared to where it is today, then in the centuries that followed fell quite a bit compared to today, and now we have got halfway back. So if this is a proxy for temperature and for the state of the climate, that would indicate there must have been pretty sustained warm conditions in that region. The picture on the right of Figure 3 shows a medieval tree stump in California that is currently in an area where no trees grow, because of the climate conditions. The evidence from these two sites suggests that a thousand years ago, conditions were rather warm compared to what they are today.

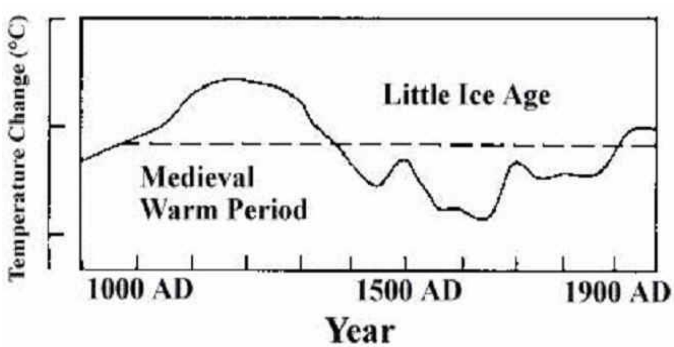


Figure 4

This was reflected in the IPCC 1995 report which has a schematic illustration of the state of the climate, or at least the temperature of the climate, showing a long Medieval Warm Period, then a Little Ice Age and then a recovery to the present (Figure 4). But we have only come part way back up to where we were in the medieval era.

So if you want to sell the story that we are now in uncharted territory as far as the climate goes and that we are experiencing unusually rapid and unprecedented warming conditions, it is very hard to do that if you have the Medieval Warm Period sitting in the background suggesting that this isn't at all unprecedented.

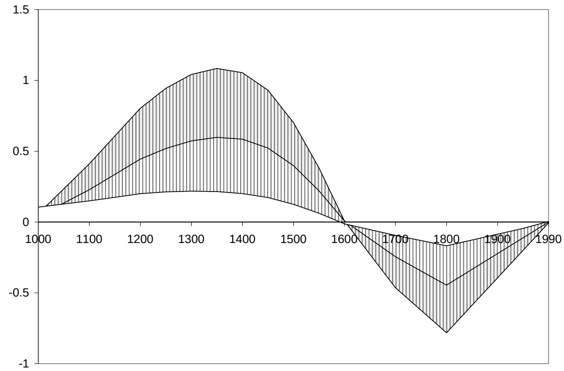


Figure 5

Now other evidence came out after the 1995 report, for instance, a paper in 1997 in the *Geophysical Research Letters* (GRL) using a global sample of borehole thermometry from Huang, Pollock and Shin. Figure 5 is just an expansion of the final portion of their graph. Their data set goes back 20,000 years. It too shows this strong medieval warm period, based on a global sample of borehole temperatures.

Not too long ago, another borehole researcher published an essay describing some things that happened to him after he published a paper on this in 1995. He published a paper in *Science* reconstructing climatic conditions in North America based on borehole record and concluded in the paper that present conditions still appeared to be within the range of natural variability. In his essay he comments,

> "With the publication of the article in Science [in 1995], I gained significant credibility in the community of scientists working on climate change. They thought I was one of them, someone who would pervert science in the service of social and political causes. So one of them let his guard down. A major person working in the area of climate change and global warming sent me an astonishing email that said, "**We have to get rid of the Medieval Warm Period.**"
>
> – D. Denning, Science 1995.

So that is why I call it the Medieval Warm Period problem, because it is something that some people want to get rid of. Shortly after that, the problem appeared to be solved with the release of the hockey stick graph in 1999. This is from the press release that the authors put out:

> **March 4, 1999.** Researchers at the Universities of Massachusetts and Arizona who study global warming have released a report strongly suggesting that **the 1990s were the warmest decade of the millennium, with 1998 the warmest year so far**… The latest reconstruction supports earlier theories that temperatures in **medieval times** were relatively warm, but "even the warmer intervals in the reconstruction pale in comparison with mid-to-late 20th-century temperatures," said Hughes.

Now this was based on the 1999 *Geophysical Research Letters* paper, but the methodology and data were introduced a year earlier in 1998 in *Nature* and there is a very tight timeline here. From the *Nature* paper in April 1998 to the *GRL* letter in March 1999 to the draft of the IPCC report for government and expert review in April 2000, the hockey stick was the one palaeoclimate graph chosen for emphasis throughout. It came to prominence very quickly in that process.

Now in the 1998 paper which we called MBH98 (for Mann, Bradley and Hughes 1998), there are 112 proxy series used. Of these, seventy-one are site records from individual sites; the tree line altitude record, for instance, would be an individual site record. Thirty-one are weighted averages of a very large set of underlying tree ring proxy data and these weighted averages are called principal components. We decided to leave out all the math on the principal components, but the gist of it is that if you have a large matrix, you can reduce it to a few key series that summarize all the activity in the larger matrix. Principal component (PC) analysis gives you an ordered list of these vectors. The first principal component is the dominant pattern of variability in the matrix and then the second principal component captures the second most important pattern, the third principal component, the third most important and so on down the line. In many cases, if the data is at all correlated, principal component analysis lets you take potentially hundreds of series and replace them with a few principal component averages and capture most of what you are interested in. But the first principal component is really the key series coming out of a principal component analysis.

Now, going back a couple of years, we published a paper in the fall of 2003 in the British journal *Energy and Environment* where we discussed problems that we had found as we had inspected the MBH98 data set closely. From what we could tell, we were the first people to really inspect it closely. We found

- 35 series listed as being used were not used
- Truncation of sources
- Obsolete data
- Duplication of series
- Series in incorrect geographic locations
- Problems in the PC calculations

Most of these things (and they are listed in the paper in lots of detail) don't have a huge influence on the outcome, but the last one really did matter. There were problems in how the principal components were calculated and we just couldn't replicate the original principal components. We didn't know what the problem was, but we couldn't replicate it and we couldn't get any information on what was going wrong here. However by removing these errors and recalculating the principal components, our conjecture was that a more authentic rendering of the information in the data was the red line, rather than the blue line (Figure 6), which is obviously quite a different story than the blue line in terms of the climate history.

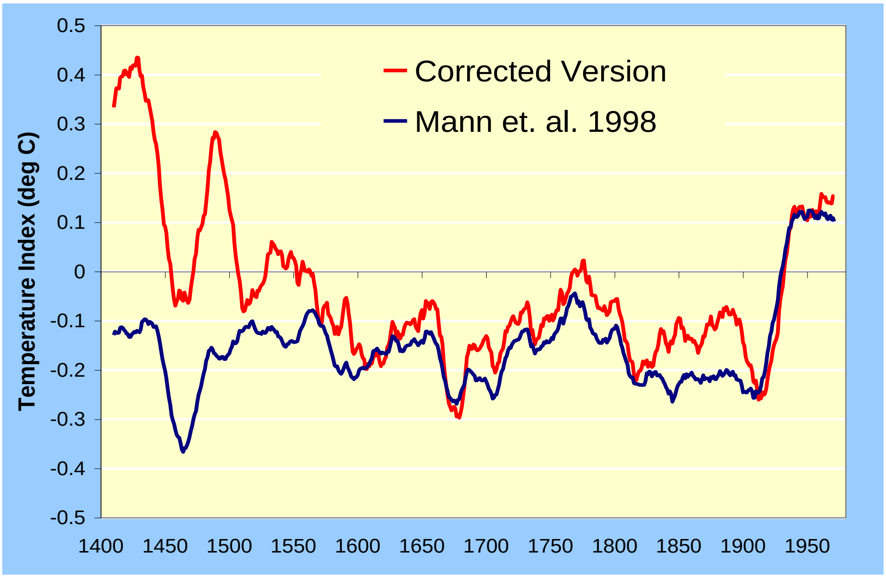


Figure 6

That all came out in November 2003 and following from that there was over a year of very vigorous exchanges back and forth, new debates including the release of a new FTP site and later a new archive at *Nature* by

Mann and his co-authors, which clarified some ambiguities about what the data was. They also revealed for the first time just a small portion of the computer code that they used to generate their results, most of which we still don't have access to. But we did get access to the portion of the code that calculates the principal components. Through analyzing this Fortran program, we came to understand why their principal components looked different. It is a very subtle difference between the way they do it and the way a standard statistical package would do it. It has to do with the way the data is standardized or centered only against the ending portion of the data series, rather than against the whole length of the data series. This has the effect of strongly favoring data series that trend upward in the 20th century.

To show how strong this effect was, we wrote a paper published not long ago in *Geophysical Research Letters* where we fed a type of random numbers called "red noise" into the program and we showed that it reliably produces a hockey stick-shaped first principal component, even when there is no trend in the underlying process. It is just a curious thing.

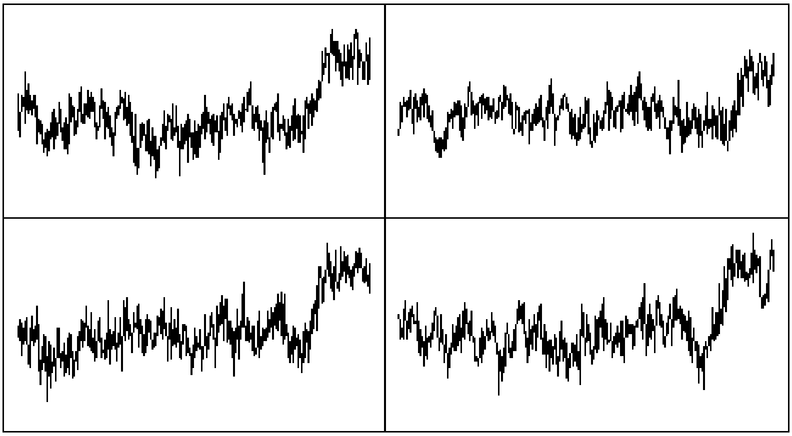


Figure 7

Figure 7 shows four graphs. Three of them are examples of "red noise" yielding hockey sticks from the Mann algorithm and one of them is the actual proxy data portion of their temperature reconstruction. If you have trouble telling which is which, that is the trouble.

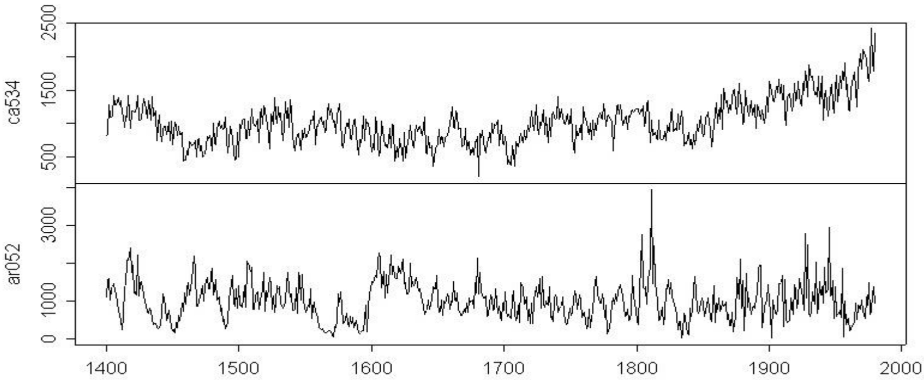
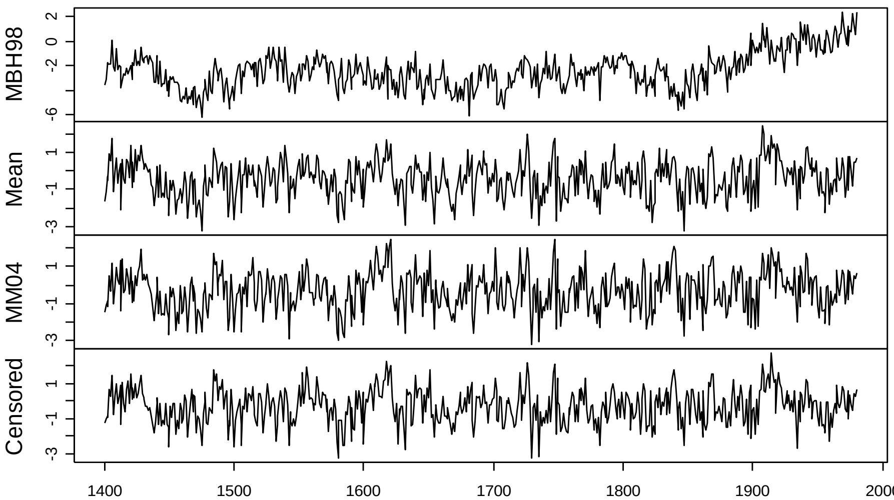


Figure 8

Now to give you another example, Figure 8 is two individual tree ring records from the MBH98 data set going back to 1400. The top is a bristlecone pine series from Sheet Mountain in California. The bottom is from Mayberry Flue in Arkansas. They are both full-length series and they would both be considered qualified for use in the study. The top one, you can see, has an upward trend in the 20th century and the bottom one does not. In the first principal component, the top one, because of the way the data is handled in this program, gets 390 times the weight of the bottom series in the first principal component.

As we looked at the top-weighted series in the first principal component, it turned out that they all have a couple of things in common. One is that they have an upward trend in the 20th century, but they also all turned out to be of one type of tree ring series. They are all bristlecone pines and they almost all of them could be traced to a single research project back in the 1980s in which Donald Graybill and Sherwood Idso were investigating patterns of CO_2 fertilization in bristlecone pine growth. In the data that they archived and they published, their own comment was that this is not a temperature signal. It is a signal perhaps of the CO_2 content of the air, but it doesn't match local temperature records. Many people in the tree ring field have worked on these series to try to explain this growth spurt, but nobody argues that it is a temperature proxy. There are various theories about why that growth spurt is there, but it is not a good temperature proxy. Unfortunately what happens here is that the data that is the least qualified to represent temperature ends up getting most of the weight and drives the results in the hockey stick graph. As a result, the dominant pattern in that hockey stick graph is non-climatic; it is not a temperature signal.



Top panel: PC1 of the post-1400 NOAMER tree ring network, calculated by MBH98 using short-segment standardization. **Second panel:** simple mean of proxies. **Third panel:** PC1 using standard software without short-segment standardization. **Bottom panel:** Unreported PC1 calculated by MBH after censoring Graybill-Idso high-altitude series. All normalized to 1902-1980.

Figure 9

One more way of explaining this is to show you this four-panel diagram. If we take the North American data set of tree ring records, a very important data set here, the first principal component as computed by Mann, Bradley and Hughes is the top series. It looks like the Sheet Mountain series and has a strong upper trend in the twentieth century, which is vital to their final results. If you take the same data and take the average, the mean of them, you get the second graph. If anything, there is a slight downward trend across the twentieth century but really there is nothing there; it is just static up and down and there is no trend, on average, in that data set. Doing a standard principal component analysis using an ordinary statistics package yields the first principal component that looks a lot like the mean. In other words, it is identifying the dominant pattern of variance as being just the unweighted mean. Again, there is no upper trend in the twentieth century.

Now our conjecture was that the Mann program picks those series that have an upper trend in the twentieth century and loads all the weight on them. If we take those top twenty series out, by and large what is left tends to look like just static, without a strong trend in the twentieth century. In that case, the program shouldn't be fooled or trip up. The bottom panel shows what happens if you take the Mann algorithm and apply it to the data set, but take out those top twenty bristlecone pine series. Then it reverts back to normal behavior and gives you the standard-looking first principal component. So those bristlecone pines are really driving the results. One of the interesting points here is that we didn't actually calculate the series itself; we found it on Mann's FTP site in a folder, but the result hadn't been reported.

This was clear to the authors and in their 1999 paper they say that these bristlecone pines really are driving the show. So that has been part of the debate that has carried on over the past year. People need to understand that almost all the data in this data set is there for show; it pads the list out, but it is really the bristlecone pines that give the result.

There is another important series which has not attracted too much attention. There is a Gaspé cedar series that appears twice in the data set and that is odd enough; in one case, there is a segment of very sparse information at the beginning, including a segment of missing data. The early years were extrapolated to cover that missing data segment and then the start date was listed as 1400 AD rather than, in this case, 1404. Steve looked at what happened if you take the extrapolation out and it ends up having a surprising large effect on a controversial 15th century portion of the study because this Gaspé cedar series has an upper trend in the twentieth century. So it is quite influential as well.

We have published two papers on the topic, one in *Geophysical Research Letters* and a second one in *Environment and Energy*. We pointed out that not only is there the problem of the program itself strongly favoring hockey stick shapes, but there is also a problem of deciding how to judge statistical significance here. If you feed random numbers into a program and it spits out a result that looks like it fits well, then the actual data have to meet a high benchmark in terms of its performance against the temperature series. So we recalculated the benchmark statistics after taking into account this unusual transformation of data and found that the predictive skill of their model is actually statistically insignificant.

Once we had taken account of the way the PCs are done, the particular statistic that they are using, the RE statistic, the benchmark was higher than they compared it to and in fact, they don't exceed the benchmark. The conclusion in statistical terms is that their model is insignificant. That probably would have been noticed earlier because there are some simple related statistics, most notably the R^2 statistic. R^2 is just a standard, very simple statistic that people routinely look at in any kind of application like this because it gives you a rough and ready statement of the explanatory power of the model. They did not report their R^2 statistics and they are still not publicly available, but in the simulation of their results, we conclude that their R^2 is zero. If this had been known, I think people would have wondered why they got an apparently high RE statistic but a zero R^2. The reason, again, is the way the PCs were calculated; that shifts where the RE benchmark is. Overall we conclude that their model actually does not have statistical significance sufficient to use it for the study of historical climates.

There has been a lot of reaction around the world to our papers so I will show you some of the professional reaction. We are including this more or less for vanity here, but also to let you know that serious people have looked at this and are coming around on it.

- Professor Francis Zwiers, a Canadian Climate Centre statistician, quoted in the *Wall Street Journal* said, "Mann's statistical method "preferentially produces hockey sticks when there are none in the data."

- Prof. Hans von Storch, a well-known German climate scientist at the GKSS Research Centre, said our criticism on this point is "entirely valid."

- A Dutch science magazine that wrote a cover article hired Professor Mia Hubert, a statistician at the Catholic University of Leuven to go over all our work and to issue her opinion. She concurred that "Tree rings with a hockey stick shape dominate the principal component analysis with this method."

- Professor Richard Muller, a geophysicist at University of California at Berkeley, wrote about this last fall saying, among other things, the findings "hit me like a bombshell, and I suspect it is having the same effect on many others."

- Dr. Rob van Dorland, an IPCC Lead Author and climate scientist at the Dutch National Meteorological Agency, said "It is strange that the climate reconstruction of Mann passed both peer review rounds of the IPCC without anyone ever really having checked it." It is not so strange if you understand how the IPCC works.

- Not long ago after our first paper came out, we learned that German climatologist Ulrich Cubasch had asked a PhD student to try to replicate Mann's hockey stick curve and they discovered that they could not do it either. There was a TV show about this in Germany (Das Erste, Feb. 16, 2005) which said, "He [Ulrich Cubasch] discussed with his coworkers - and many of his professional colleagues - the objections, and sought to work them through … Bit by bit, it became clear also to his colleagues: **the two Canadians were right**. … With that, the core conclusion, and that also of the IPCC 2001 Report, was completely undermined." With that, the core conclusion and that also of the IPCC 2001 report was completely undermined.

- Not long ago, we got an email from Dr Hendrik Tenekes, who is the retired Director of the Royal Meteorological Institute in the Netherlands. Among other things, he commented, "The IPCC review process is fatally flawed ... The scientific basis for the Kyoto protocol is grossly inadequate." (Feb. 22, 2005).

Now at this point we move into the third period where Mann and his co-authors have vigorously defended their position. I will go on the bench and let Steve continue.

Steve McIntyre: Ross, thank you very much for being here. They are obviously defending their position very intensely. The first argument that they make is that even if they did the principal components calculations the wrong way, it doesn't matter. Intuitively when you can produce these kinds of hockey stick-shaped PCs from "red noise," you feel it probably does matter. They say if we use five PCs instead of two PCs, we can still get a hockey stick shape. Or if we don't use PCs, we can still get a hockey stick shape and there are ten other studies that show the same thing.

Our first argument is that if five PCs are the right way to do it, why didn't they do it right the first time? At this point, they are throwing out reconstructions that are not peer-reviewed. They are up on a website and no one has examined them. If you develop a new method, there maybe a few warts on it and in your attempts to salvage it, you can create new problems.

If we take the bristlecone pines out, we don't get a hockey stick but we will do a full reconstruction. All we will do is do what Graybill and Idso did and calculate a CO_2 effect and at least allow for that as a possibility and see what happens. If we take a CO_2 effect out of the bristlecones, we get a high 15th century, which shows one more time that their reconstruction is not robust to the bristlecone pines. There is a very definite alternative possibility in this hockey stick, suggesting that all we are doing is getting a non-climatic effect on bristlecone pines which is being carried forward into the reconstruction and that the reconstruction itself is not an authentic signal of temperature.

They answer that if we don't include PCs, we still get a hockey stick. But what happens there? First of all, one of the warranties of the original study was that there was relatively even geographical balance, that this was the best the geographical method had ever done and it was testing everything in the world. But in reality, eighty of the ninety-five proxies are American tree rings and twenty of the ninety-five are bristlecone pines, so really all they are doing is getting the bristlecone pines in through the back door. Again, the bristlecone pines are not just any old proxy; these are ones that are specifically identified by specialists as not being temperature related. This is a breach of another warranty; they had said that all their proxies had been carefully examined to determine that each proxy was a valid temperature proxy. So here we are with non-temperature related proxies driving the bus. Further, they had specifically studied the effect of excluding these proxies and had not gotten a hockey stick and not reported that result.

Again, for any of these new salvage reconstructions, our prediction is that if and when they produce the digital versions of the 15th century stuff, they will all fail the R^2 step and every one of these reconstructions will also be statistically insignificant by that measure. They respond that the RE statistic is the "preferred statistic" for climatologists. Perhaps. When you talk to general audiences about differences between R^2 statistics and RE statistics, people's eyes can glaze over. But even if you are not used to these particular statistics, you are all used to working with statistics every day, whether profit or loss. To make it vivid, if you have a statistical model with an R^2 of zero, it is bankrupt. You cannot have a model which purports to be correlated to temperature and then in validation have an R^2 of zero and still have anything useful. It's a test of bankruptcy. This is like saying, well, we are bankrupt, but we have a high growth rate!

I come from a business background and this looked to me like a dot.com promotion. The dot.com promotions thought they had found a new form of economics in which you didn't need profits. In fact, you do. I am not saying that just because you report a high R^2, you necessarily have a valid statistical model. There may be warts on that; it may be spurious. There are many tests you have to do. But if you don't pass statistical significance of an R^2, it isn't any good. Another objection I have is that he won't release the results of the 15th century step and surprisingly, climate scientists have not risen outraged at this. I think it is an absolute scandal, not just for Mann and his associates but for the entire discipline that one particular scientist is allowed to be a prima donna on this. If it were just a matter of scientists in their common rooms, that's fine, but people are making serious public policy on this. Certainly in Canada, we are spending a lot of money based on data that simply can't be replicated and on statistical methods that are, at best, highly questionable.

We have also tried to get hold of the source code. While we have been able to replicate his methods well enough to draw pretty strong conclusions, there are still aspects of it that we haven't been able to replicate. This point has received a lot of attention; in a front page article in the *Wall Street Journal* (Feb. 14, 2005) Mann was quoted as saying "Giving them the algorithm would be giving in to the intimidation tactics that these people are engaged in." At first, when we were began to see discrepant results, we sent a polite email saying, "We don't want to argue about methodological issues; we would like to reconcile methods so that we are comparing apples and apples." He has consistently refused to provide a source code and I think it is a very poor practice to base public policy on studies whose authors don't release their methodologies in complete detail. In our own GRL study, we archived all our source code and data as used at the AGU website at publication.

Another argument is that there are other studies which arrive at the same conclusion. I have two responses to that. Even if these other studies were correct, which I don't think they are, that wouldn't salvage Mann. If the other study has no statistical significance and is merely an imprint of bristlecone pines, that study is no good. Equally it is a fair comment that we haven't written about these other studies and haven't said too much about what we think about them. But there has been no due diligence done on these studies by anybody else, either. Until we started looking at Mann's study, nobody had every looked at it. People also say, "You are not just competing against a single hockey stick, you are playing against the whole team." Well, I think it is probably an even match.

First of all, the studies and proxy data are not independent, as any ordinary person would understand the word independent.

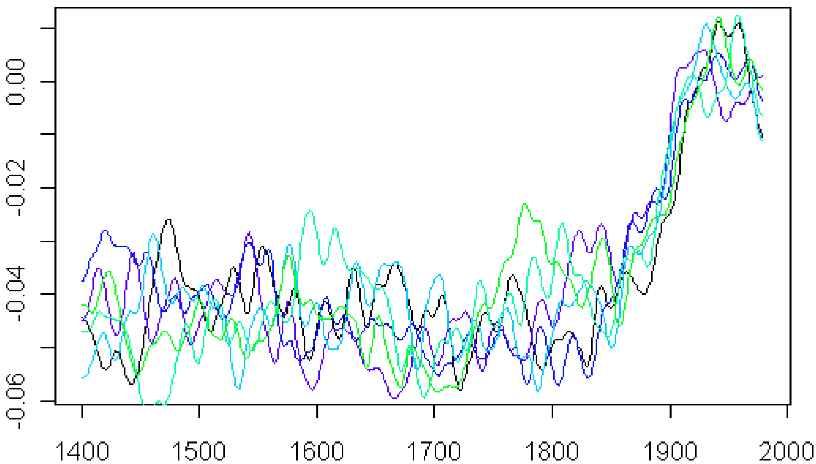


Figure 10

Figure 10 is a "spaghetti graph" that I made from some of our simulated hockey sticks. If we just put them all together and color them different colors, we get a more convincing spaghetti graph from "red noise" than any of the ones that they did. If we have an underlying cherry-picking process, we can produce hockey stick graphs from "red noise," so the statistical task is to find out if we are doing any better than just simple cherry picking.

I also contend that many of the studies are not independent. Here is a listing of some of the most often cited ones.

- Bradley and Jones [1993]
- Hughes and Diaz [1994]
- Mann, Bradley and Hughes [1998, 1999]
- Jones, Briffa and others [1998]
- Briffa [2000]
- Briffa, Jones and others [2001]
- Mann and Jones [2003]
- Bradley, Hughes and Diaz [2003]
- Jones and Mann [2004]
- Rutherford, Mann, Bradley, Hughes, Briffa, Jones and Osborn [2005]

They are all sort of "et al." studies: Jones et al., Briffa et al., and Mann et al. Now here are the MBH co-authors in red and the Jones and Briffa co-authors in blue:

- Bradley and Jones [1993]
- Hughes and Diaz [1994]
- Mann, Bradley and Hughes [1998, 1999]
- Jones, Briffa and others [1998]
- Briffa [2000]
- Briffa, Jones and others [2001]
- Mann and Jones [2003]
- Bradley, Hughes and Diaz [2003]
- Jones and Mann [2004]
- Rutherford, Mann, Bradley, Hughes, Briffa, Jones and Osborn [2005]

The first name on the masthead rotates, so Briffa et al. is not a different set of authors than Jones et al. and Mann et al. is not different than Jones and Mann. In the most recent one where they take a crack at us, in 2005, all six of them join up together.

In terms of the proxies not being independent, if we actually look in detail at what is in these series, we see certain stereotype series occurring again and again. Bristlecone pines don't just affect the Mann studies; bristlecone pines are in Crowley and in Mann and Jones. There are other series like the polar Urals, the tree ring series, and the Tornetrask tree ring series that are in virtually every study. So the question then would be, if there is a problem with bristlecone pines and bristlecone pines are in eight of the other ten studies, is the "active ingredient" of each of these hockey sticks the same thing? Can each of these other papers survive a sensitivity study to bristlecone pines possibly being non-climatic? This sort of non-independence in the proxies was even recognized by Briffa, one of the "hockey team," who pointed out that very few of the series are independent; they are a common input to them all:

> "An uninformed reader would be forgiven for interpreting the similarity between the 1000-year temperature curve of Mann *et al.* and a variety of others also representing either temperature change over the NH as a whole or a large part of it (see the figure) as strong corroboration of their general validity .... **Unfortunately, very few of the series are truly independent: There is a degree of common input to virtually every one**, [emphasis added] because there are still only a small number of long, well-dated, high-resolution proxy records."
>
> Briffa and Osborn, *Science* [1999]

In looking at some of the other studies, we found that even to get a toehold on these things, we had to get access to the data. In many cases they refuse to provide data. For me it is much worse. Before when I was just some aging Canadian businessman, as a member of the public at least I could get return emails. Now if I send an email to one of these guys asking for data, they don't write back even to refuse anymore. But there are many problems: they splice different kinds of series, they use obsolete data, and they cherry-pick. I will illustrate some of these things.

Crowley has a famous study. After twenty-six emails, he sent me a smoothed and transformed version of his underlying data. (This was back when I could still get return emails.) Then I asked to see the actual used, not a smoothed version. He said he couldn't locate it and was not sure if he still had it. I checked some of the smoothed versions and couldn't match them to archived sources and I asked where they came from. He couldn't recall where he got the data. Some of this data he got from Jones, so in fact even Crowley is not fully independent of the Jones-Briffa group. But more surprising, he was astonished at being held responsible. He said, "This is five years old, I did it elsewhere; how can you expect me to have this data?" Yet this study is quoted by IPCC and people are relying on it. He did also mention in a speech about a month ago that he asked for some data from me and I hadn't sent it to him. In fact, I did send that data to him; I said, "Look, if you check your in-box, you will find that the data is there." He said, "I am sorry, I have been very busy with other things and didn't notice."

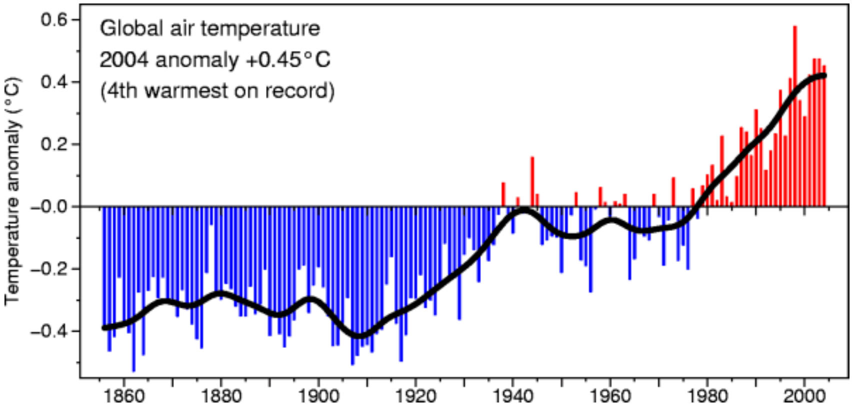


Figure 11

Figure 11 is the famous Jones temperature series and I will show you how this ties in to the proxy one. This is the other part of the promotional diptych with the Mann thousand-year study. A researcher we know said, “We want to see the underlying station data for an archive of that.” Jones said, “We have 25 or so years invested in the work. Why should I make the data available to you when your aim is to try and find something wrong with it?” He is English, but all of this has been funded by the Department of Energy. The underlying data should all be archived and it should not matter whether you’re in the “in crowd” to get the data. It should be there for everybody to look at.

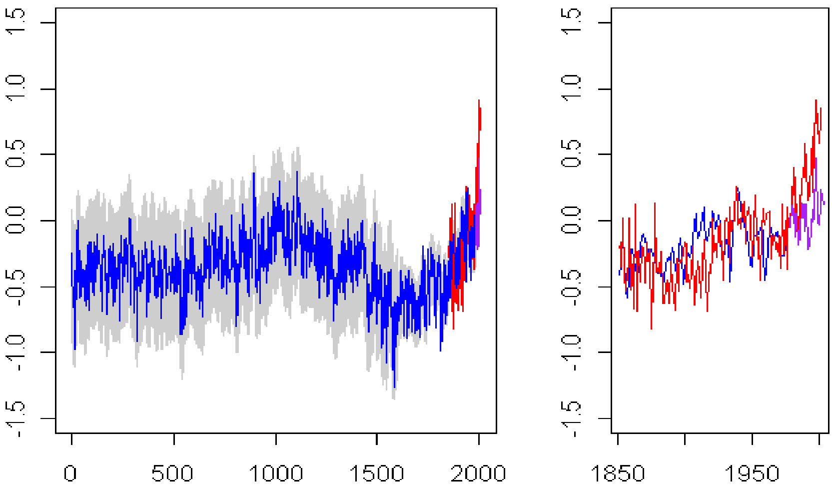


Figure 12

There has been a lot of publicity recently about a new study by Moberg in *Nature*. He is the “new sheriff in town” and people are asking, “What are you doing criticizing Mann?” The day that Moberg’s paper came out, I got an email saying, “The climate community has moved on; so should you.” I actually plotted up Moberg’s line in an IPCC form (Figure 12) and the blue line is what the proxy data shows. When I first plotted it, I said that doesn’t even look that much like a hockey stick anymore. There is a big Medieval Warm Period and the 20th century is a little different because it is coming up to the Medieval Warm Period, but it is not blowing it off the chart. What blows it off the chart is the Jones instrumental series, which is the one for which he won’t let people look at the data. I thought just as a little experiment, what happens if I splice the satellite data on to this instead of the Jones data? The blow-up on the right is on the same scale. The purple line is the satellite data so if you just take the rate of increase, you are not even getting the oddness of the 1990s with satellite data spliced onto the proxy data. This is a little different twist on the satellite debate because everybody is now saying we are looking at tenths of a degree or .01 degree growths. I say if you splice that back into the proxy record, it takes a lot of the steam out of it.

Another thing is that even though Moberg's study is new, he uses some very old data; he uses a bristlecone pine series that ends in 1962. Even Mann used one that went up to 1980. So though it is a newer study and people claim this is newer data, a lot of it isn't. They are still using proxy data that ends in the 1970s. This is important. Why don't they use up-to-date proxy data? If the 1990s are super-warm, the proxy data should be just going off the charts and so we have a wonderful opportunity to benchmark whether these proxies are any good. My personal feeling is that most of the proxies are no good. Just calling something a temperature proxy doesn't make it a temperature proxy.

What is Mann's answer? He says we have to rely on proxies ending in the 1970s and 1980s because getting new proxies is

> "a costly, and labor-intensive activity, often requiring expensive field campaigns that involve traveling with heavy equipment to difficult-to-reach locations. [...] For historical reasons, many of the important records were obtained in the 1970s and 1980s and have yet to be updated."

An awful lot more money has been spent on climate research in the 1990s and the 1970s, so that doesn't seem to be a very good reason. And what is the heavy equipment that is used? A hand-held drill weighing a few pounds, hardly a heavy drill. In terms of remoteness of the sites, bristlecone pines are about twenty miles from Bishop, California. Granted, you would probably have to leave a vehicle to get a sample from them and you would have to stay overnight in a motel that was not five-star. That may seem difficult to a professor at the University of Virginia, but I could find many geologists who would be able to handle that kind of heavy equipment and go to any remote site that you could name. I don't think that is a particularly good reason not to update the proxies.

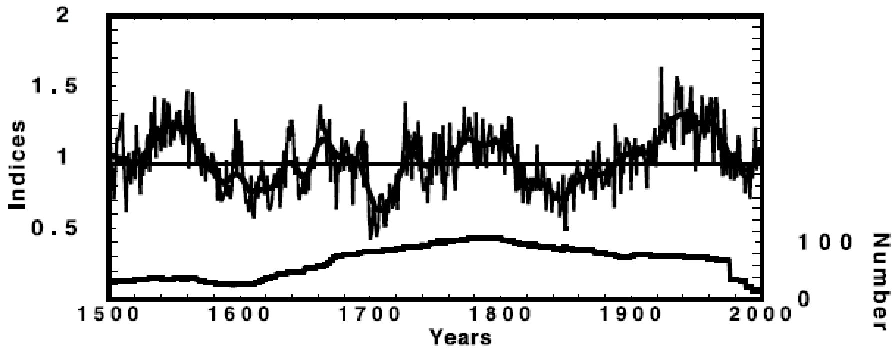


Figure 14

In fact, there is a lot of new proxy data. What does it show? Figure 14 shows data taken by Jacoby on the Twisted Tree Heartrot Hill. This is a classic site that was used because you see a trend going up in the middle of the 19th century and then starting around 1980 it goes down. This data set goes up to 2002 and obviously this particular proxy is not going off the chart in the 1990s. If it did not miss the 1990s, how do we know there wasn't some other warm period in that proxy?

The Gaspé series was really important in the Mann study. Somebody unofficially sent us an updated version of the Gaspé series. Figure 15 shows two versions: the purple one is the one that was used in the Mann study and the navy blue one is the updated data. You see that the updated version does not have a hockey stick shape.

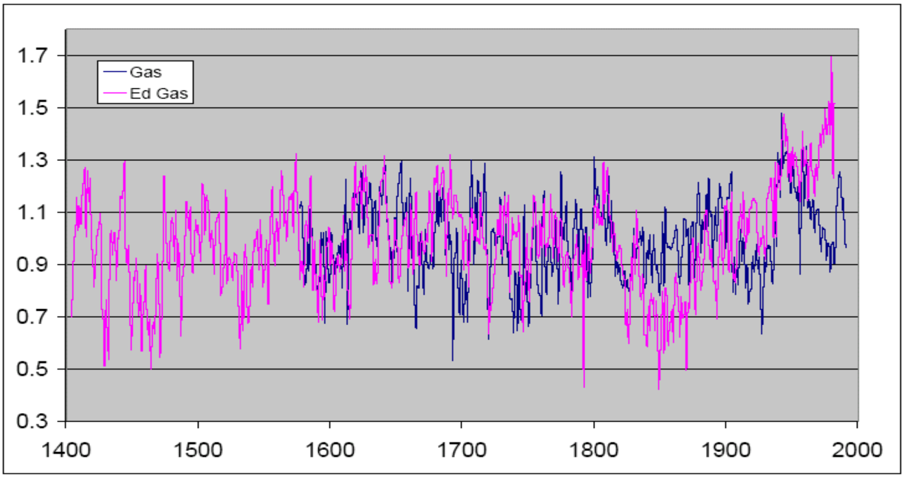


Figure 15

I asked Jacoby and D'Arrigo for the updated version, since I just had this graphic, but did not have the data. First of all, they said the earlier version was superior in giving a temperature signature so I should continue using the earlier version; the later one didn't have a good temperature signature in it. I said I would like to commission my own re-sampling of this site and asked if they would tell me where they took the samples. I worked at that for about six months and then complained to the journal that published it. Then the authors said,

> "There was an attempt to update this record but the original site was not located. The original sampling was prior to GPS locating. Therefore there is no newer data for this particular site. If we implied this is any published paper, we misspoke."

In Jacoby's 1999 paper on the northern tree line series that is often used, he mentioned that he had done thirty-six sites. He had selected the ten most "temperature sensitive" and then averaged those. I did some tests and found that if you do that with "red noise" and all you do is pick out "red noise" series that trend up in the 20th century, you get hockey sticks every time. It is a different method than Mann's PC method and it is a little low-tech, but it is just as effective a way of cherry picking. I wrote to the journal and asked to see the twenty-six series that he didn't use. Jacoby refused to release them and said,

> "Most of our research has been mission-oriented … If we get a good climatic story from a chronology, we write a paper using it. That is our funded mission ... The rejected data are set aside and not archived …. As an ex- marine I refer to the concept of a few good men."

As for being mission-oriented, I would simply say in the exploration business, geologists are mission-oriented: they are trying to find ore bodies and they drill holes. We are obligated to publish the results of bad holes as well as good holes. This is really a symptom of a very general problem. This is also one area where we can quantify the cherry picking. One of the fundamental problems in all the other multiple proxy studies is how were these particular series selected and was the analysis controlled against the cherry-picking process? I will go through a couple quickly.

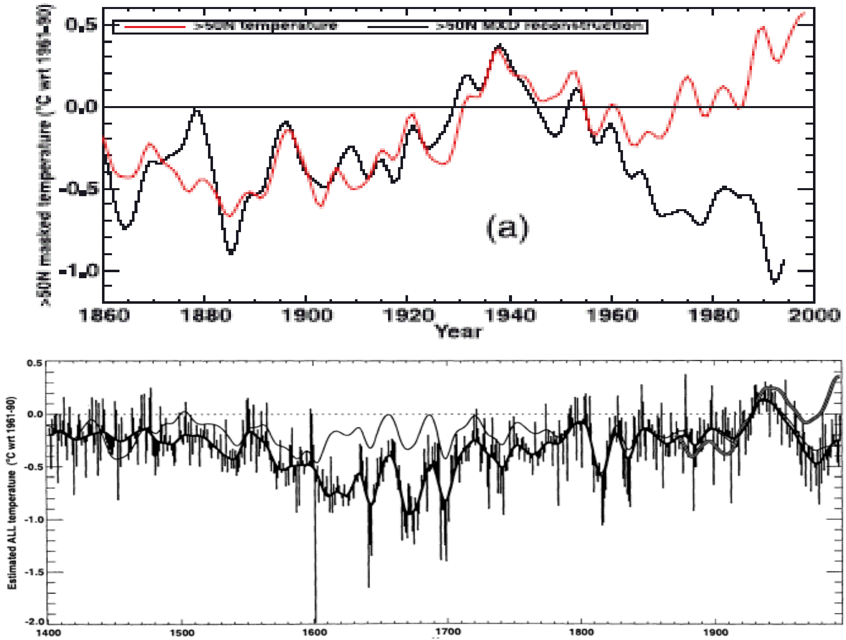


Figure 16

Figure 16 is Briffa's maximum density series, which is up to date. The top is the temperature series and the bottom is the density series and they go right up through the nineties. There is an increase up to about 1960 and then the proxy comes down. How do Briffa et al. explain that? They say there is some unknown anthropogenic factor causing this series to go down. How does the "hockey team" handle this sort of stuff in their reports and still come up with spaghetti graphs with everything going up? Good question. Here is one way of doing it.

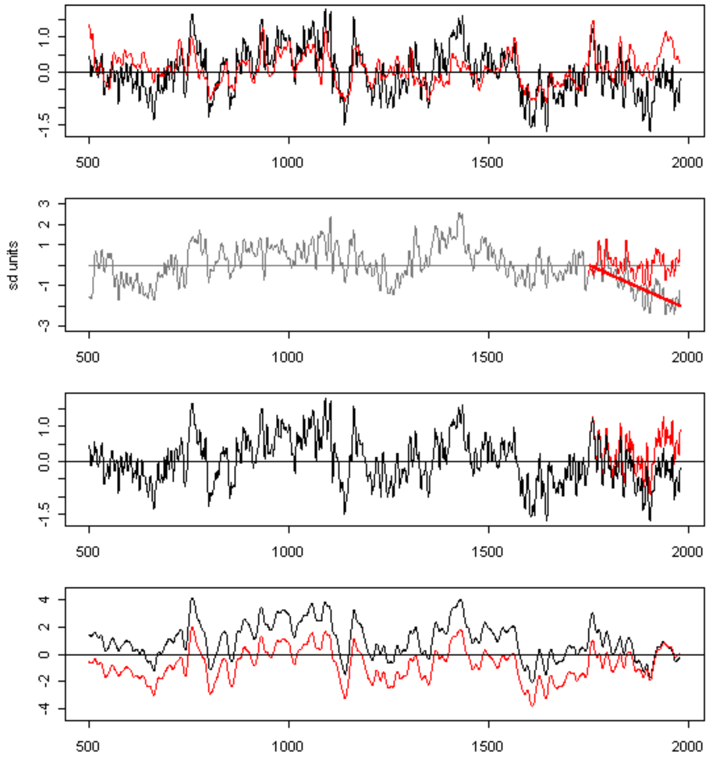


Figure 17

Figure 17 shows the Tornetrask series. This is an emulation of actual diagrams that they did. He saw that after 1750, these residuals are going down, but he felt that they should not go down – there is something wrong with that. So what we will do is just straighten out that downward trend and even them out. Then we change the series to one and instead of a low 20th century, we get a high 20th century. There is a knock-on effect because if you then standard re-normalize these things to a 1902 base, if you hadn't done that, you would get a big Medieval Warm Period from this study. But by doing this adjustment, we get a low medieval period, but nothing fancy. What is the rationale for that? If that is a good thing to do there, is it done in every other series? Of course not. Only for the ones that are going down.

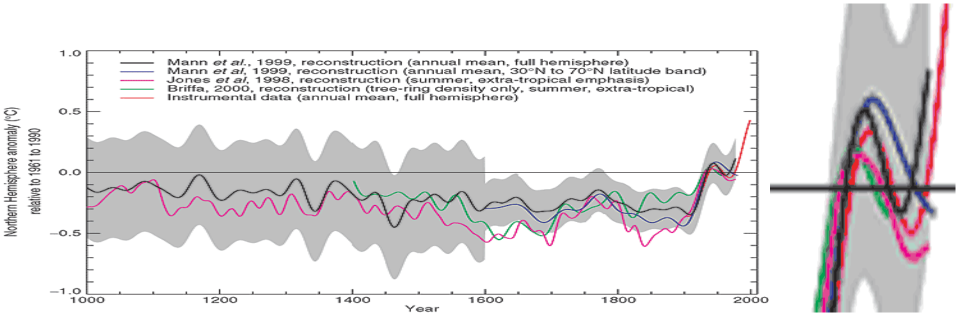


Figure 18

In the graph in Figure 18, which is actually used in the IPCC report, the version of the **MXD** going down. How does that turn up in the IPCC? How do they handle that and still get a spaghetti graph? Well, here is their spaghetti graph, and the Briffa series is the one in green. It looks like it is going up at the end. But if you actually blow it up (right side), what you find is that the green series stop around 1960. When it starts going down, they just delete that portion of the series that goes down.

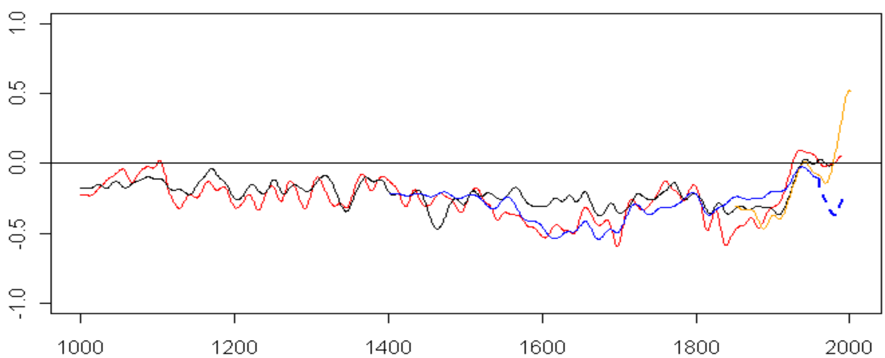


Figure 19

If you take the full study without the deletion and put the deletion back in, you then get this blue series where the Briffa series goes down at the end and you end up with a much less convincing spaghetti graph of everything doing the same thing. So part of the appearance of consensus is created just by deleting an unfavorable portion of a record.

In closing up, I got interested in this because I thought this whole thing looked like a promotion. I started out from the mineral exploration background so I am used to promotions and I am used to fairly wily guys. I never expected anybody would be interested in what I had to say about it; I was just doing it for my own curiosity. My entry point was one day I emailed

Mann and asked him where his data was. I was just going to look at like a field core. He said he had forgotten where the data was, but he would get one of his associates to look at it. His associate said it wasn't in any one place but he would get it together for me. At that point, I got very interested in it and I thought, "Nobody has ever looked at this stuff!" It just seemed impossible to imagine, but in fact that was right – nobody ever had.

There is lots of hair on the story about what happened with the data that was supplied to us, but essentially governments trusted the IPCC and the IPCC relied on journals, their referees and their authors. But who checked it and whose responsibility is it to do due diligence? I have done this on my own nickel; it is sheer madness as far as my family is concerned because I haven't made any money for two years. But literally nobody has done this stuff. So what testing did the IPCC do? Here is Mann's response:

> "It is distinctly against the mission of the IPCC to "carry out independent programs", so the premise of the question is false. However, the IPCC's author team did engage in a lively interchanges about the quality and overall consistency of all of the papers as the chapter was drafted and revised in the course of review."

But the public thinks that this data has been checked and they think that this is an engineering-quality document, not a matter of some guys just talking at a coffee klatch about their studies. When we tried to get code from the National Science Foundation, who funded this stuff, they said,

> "Mann is under no obligations to provide you with his programs. Dr. Mann and his other US colleagues are under no obligation … to provide you with computer programs, codes, etc. His research is published in the peer-reviewed literature which has passed muster with the editors of those journals and other scientists who have reviewed his manuscripts. You are free to your analysis of climate data and he is free to his."

At the journal, I was asked to review a submission by Mann, curiously, on climatic change and in my capacity of reviewer, I asked him for the source code and the journal editor (*Schneider*) said,

> "I have run the issue by the full Climatic Change Editorial Board since a source code request by a reviewer is unprecedented in the 28 years since I founded the journal."

Because of that, they at least adopted a policy that authors had to archive their data. I then asked for Mann's supporting calculations, which he refused to provide the journal. We haven't heard any more about that submission. *Nature* said,

> "we do not take the view that [source codes] are something that in general should automatically be provided on request - the decision of whether or not to do so normally rests with the authors of such codes."

I had some correspondence with *Science* and they have no policies on this right now.

Going back to the original representation of what had been done: the representations made to the government are that this has been rigorously reviewed, that every step along the way has been checked, that we have engineering quality and due diligence. Well, we don't. Thank you.

Addendum

Steve McIntyre's website is [*www.climateaudit.org*](http://www.climateaudit.org)
Ross McKitrick's website is [*www.uoguelph.ca/~rmckitri/research/trc.html*](http://www.uoguelph.ca/~rmckitri/research/trc.html)
Michael Mann's website is [*www.realclimate.org*](http://www.realclimate.org)

Questions and Answers.

Question: I have talked about your synthetic time solution[?] to "red noise." Why is it that they only show ups at the end and not an equal number of downs? Why don't you see similar kinds of ends in the series effects at the beginning of the time series and the end? If you are just putting in "red noise," most analyses should be symmetrical both up and down and left and right.

McKitrick: The two questions there. The up portion, the "red noise" process will generate an even number of up and down portions. For the purpose of the temperature reconstruction, they are going to be regressed against an upward sloping temperature series, so it just sticks a negative coefficient on the downward sloping ones. In fact, the first principal component for the 99 GRL paper for MBH99 was actually downward sloping; it just gets flipped over in the regression step.

McIntyre: In the GRL article, we defined a hockey stick shape as being something where there is greater than one standard deviation difference in the

20^{th} century and the historical series. The graph we illustrated showed that 99% of the time, it is going to be one standard deviation up; half the time it will be a negative hockey stick, and half the time, a positive hockey stick.

Question: What about the beginning of the time series?

McKitrick: The standardization is done at the end of the time series, so for a regular standardization, you divide by the mean and the standard deviation of the whole length. You get a series of mean zero variance of one and then PC algorithm looks for the dominant patterns there and it is not going to be fooled by changes of scale. In this case, Mann is standardizing using only the end portion of the series, so he is just taking the mean at the end of the series. In this case, the algorithm is strongly sensitive to this gap between the mean of the whole series and the mean at the end. Now if you did the standardization using the first eighty years of the series, then you might generate hockey sticks laying the other way. I don't think we have checked that out, but in this case, the reason they come out with that step going up into the last portion is precisely because that is where the standardization is done. There are always departures, but for most of the series, there is enough time for the series to return back to the mean. But you will have series that have a departure in the 20^{th} century that don't have time, because of the autocorrelation, to get back down to the mean. So it is able to pick weights that will take those series that have a departure in the 20^{th} century that doesn't revert back to the mean and construct the hockey stick out of that. That is why it is all at the top end.

McIntyre: I used the spaghetti graphs to illustrate the cherry picking. There are different ways that you can implement a cherry-picking process. One way is through this automated weighting system, so that Mann's PC method is kind of automated cherry picking. But you can actually just do a simple cherry picking by just picking series that slope up at the end, which is what Jacoby does. So in terms of the illustration of the spaghetti graphs, you are quite right, that is why I picked upper ones. Because half the hockey sticks would go down. But part of the premise is that the guys are picking ones that go one direction or they say that if it is going down, that is because it has a negative correlation and they flip it and make it go up. That flipping process is accommodated effortlessly in proxy __. All I am saying is that you can generate hockey sticks in a lot of different ways. If all you are doing is picking data with a 20^{th} century trend out of "red noise," and you have random stuff in the shaft, then the random stuff in the shaft cancels out, so you get a little amplitude in the shaft and the result is a blade of the hockey stick.

Question: Even with your conclusions, Canada is implementing Kyoto, Europe is implementing Kyoto, and in the Senate, there is energy legislation coming up. There is a lot of buzz around town that that might include some climate provisions. What impact do you hope your conclusions will have on the policy debate both here and elsewhere?

McKitrick: Let me correct you on one thing: Canada is not implementing Kyoto. Canada has *ratified* Kyoto and it is introducing a lot of toy programs that have the name Kyoto on them, but in terms of actual CO_2 emissions, Canadian emissions have not departed from the business-as-usual trend and they won't. It is just sort of an odd PR gesture. As for the policy implications, in a way it would be odd to think that this would have policy implications, but then it is odd to think that the original study had policy implications. If there is a policy implication out of all this, my preference would be that it just raises a whole lot of questions about what happens when a scientific study is used to drive policy and everybody is assuming that everybody else checked it out. Everybody is assuming that journal peer review constitutes some kind of a rigorous due diligence process and it turns out that actually not even the journals did the foundation checking. I would hope that the one policy implication that would come out of this is for people to realize that journal peer review is not a basis for setting policy and if studies are going to be used as a basis for expensive decisions, then they should be able to withstand a higher level of scrutiny. You should be able to take them apart from scratch and at least verify the results.

Question: What processes are in place with the journals and the IPCC for peer review? Clearly they are saying that scientists have looked at these studies and have signed off on them, essentially.

McIntyre: People misunderstand what peer review is. From the business point of view, I think it is important to understand that peer review is not an audit. When a company is doing a prospective, we have to use audited financial statements. A lot of it is really due auditing. They check invoices, they check records, and they check computer programs. No peer review for a journal is remotely equivalent to an audit. They are very low level due diligence: some guy going home while his kids are screaming and reading it for an hour and deciding whether it makes sense. At worst. And at best, he will check on it. You can't really expect that; the peer reviewers are not being paid to do it. To replicate a study, to do a full audit, is a big job. Auditors in financial statements are well-paid professionals. They spend a lot of time at it and journals are not equipped to do that kind of work. At a minimum, one way of control on this process is to do a simple thing like provide the source code and data so you don't have to fight the authors for this stuff. Anyone should be able to go to the journal website and pull it down. In empirical economics journals, that has become standard practice. Bruce McCullough has written a working paper for the Federal Reserve Bank of St. Louis on this recently for economics and cites our work as being an interesting example illustrating this process in another area. There is no magic bullet for it, but it is astonishing to me how little due diligence there is in this, compared to the due diligence in a prospectus or any kind of public offering of stock, even by cruddy little mining companies. Peer reviewers merely give advice to the editor as to whether a paper should be published. There is no warranty that the results are correct or that they can be reproduced. The reviewers whom the editor picks say yes, we would like to see this in the journal; that is the start and finish of it.

Question: Does that mean we should not trust any of the studies that suggest that climate change is a problem? Was your article peer reviewed?

McIntyre: The first was obviously a rhetorical question and we are limiting our comments here to the studies that we looked at. I don't think there is anything wrong with the journal peer review process, as long as people understand what it is. It works at weeding out a lot of errors and raising the standard of what gets published and filtering things in terms of which journal they should go to. I certainly don't want to suggest kind of a nihilistic attitude toward scientific publications. Our article was peer-reviewed. But that is not the main line of defense actually for our work. I think the real line of defense is that all the data and the code are archived at GRL and we know a lot of people have downloaded it and reproduced our results. People are trying to find holes in it and that is fine; that is what they should be doing. But everybody agrees on what we did and exactly how we did it.

McKitrick: So far there have been three comments submitted on our article to GRL which are under review and that we have been asked to respond to. So there is quite a large aftermarket for this and it is a fair bit of work servicing the aftermarket. Certainly it strikes me as ironic that the climate community is spending much more effort trying to show that we are wrong than they ever tried to check the original studies. Even the guys who are commenting on us can't even begin to replicate what Mann did; they have no idea what he did. Now in terms of the three comments that we have seen so far, I would say all of them agree that the process used is biased toward producing hockey sticks. That finding, that that method is no good, has been confirmed not simply by the five or six people quoted in various articles, but in the submission.[?] The defense that they are trying to take is that the error doesn't matter. I think they are going to find that a very heavy burden to try to prove that the error doesn't matter. But they are trying hard. The three comments that I have seen so far don't, in my opinion, lay a glove on any of our points.

Question: As I understand it, you are demonstrating that the temperatures in the latter part of the 15th century were warmer, in fact, that in the latter part of the 20th century and that is why the hockey stick formulation is flawed?

McIntyre: No, we are not claiming anything at all. What we are saying is that their data and methods don't enable them to conclude that the 20th century is warmer. We are taking a *reductio ad absurdum* and we are saying that their results are no good. As to what really happened, that is a completely different study and we are not saying what really happened. But none of these studies, as far as we are concerned, are any good for concluding what happened. If you ask me what I *think* happened, I am certainly content that the 20th century was warmer than the 19th century, but I think there is lots of evidence that the 19th century was a particularly cold century.

Question: Is it conceivable, however, in the 21st century that the additional amounts of greenhouse gases that have been poured into atmosphere as a result of human activity may tip that balance?

McIntyre: We are not commenting on that. We are saying that we think that these reports are flawed. We think it is never a good idea for people to try to help the truth along, if that's what it is. I view myself looking at this like a Securities Commission report. This is an offering to the public: does it meet the standards of an offering to the public, is it replicable, is there full disclosure? From the disclosure point of view, one of my biggest beefs about it would be withholding the R^2 statistic. If people had known that this had an R^2 of zero, it would never have got off the ground. In a prospectus, you would have to disclose this and if you didn't disclose it, the SEC would be down on you. What happens if you exclude the bristlecone pines? They knew what would happen if they excluded the bristlecone pines: you don't get a hockey stick. They didn't report that. Again, if they were following SEC standards, they would have had a report and they didn't report it in the core agenda. If you are playing in the big leagues where you are dealing with public policy, you have to have full, true, plain disclosure and that is the standard that I am advocating for this. I am not trying to say what did or didn't happen, but as the public, we are entitled to full, true, plain disclosure. I don't see how you can argue against that.

Question: I have had the pleasure of attending a lot of IPCC meetings in the last year and a half so I have seen how the sausage is made. It is interesting that the argument has always been, in terms of accepting scientific literature, that the IPCC process is much more rigorous as peer review than any journal's peer review process. And now it has suddenly flipped back the other argument saying the IPCC has to accept the journal's peer review process on its face value. What is interesting about this debate is that we realize how curious a process all that can be, irrespective of whether it is related either to the IPCC and this issue or not, what is scientific certainly generally, and how that process evolves over time. As we speak, the Working Group 1 of the IPCC is meeting in Beijing to decide the basic first draft for the Fourth Assessment of the science which we will get to see sometime presumably at the end of September for the first review process. This issue is presumably being debated and discussed: what are they going to do with this? Are they going to fix it somehow or not? Also we know from the broad debate that the Europeans have apparently decided it is a big enough issue so they are going to have their own process, the so-called hockey team approach, and have a conference next year on this issue. Is this going to become a way of saying that the IPCC will not change anything in the Fourth Assessment at this time, pending further studies, etc., or will they accept information that is done from this date on this? Allegedly May of this year is the cut-off period for accepting peer-reviewed literature that will be assessed in the Fourth Assessment, but that is somewhat tongue in cheek, because we know they accept stuff after this date, even if it has not been published but it has been cleared. Having said all that, I would just query you gentlemen if you are involved in this IPCC process now? Have you been invited to any of these kinds of deliberations that are occurring at this time?

McIntyre: We are not involved in it, nor has the Canadian government ever contacted me about our findings. I **actually __mentioned __ written** some scathing article about us so I invited him out to lunch and bought lunch for him and told him the problems with this. He said he didn't care. He said if you knock down that study, there will be ten other ones. The IPCC is aware of our work, but they are trying to work around it.

Question: On April 22, Ira Flatow at NPR had a *Science Friday* program where he had Michael Mann on along with another newspaper writer basically to lay out the opposition to your piece. Were you all invited to appear on that program?

McIntyre: No.

McKitrick: That is the first I have heard of it.

Question: Any contact at all?

McIntyre: No.

Question: You mentioned both the favorable and unfavorable comments you have gotten from scientists in this field. Do the favorable comments fall into any patterns in terms of what disciplines those scientists are in, where their funding comes from, their geographical location?

McKitrick: I don't see a pattern in that myself. We showed you a sample of them. They are North American and European university-based people. It is not like we get a whole lot of feedback, one way or the other. I think most of the people that we have corresponded with are just not well-known people, they are not prominent skeptics or prominent on the other side, they are just people toiling in the field who appreciate the study that we produced. But I don't see a pattern to them one way or the other.

Question: I have been keeping track of some of the reactions and it is interesting to see many reputable scientists now who have quite serious doubts about the hockey stick, but who are defending the IPCC conclusions. Essentially they are saying the hockey stick was never really important. It would be interesting to see how this turns out now in the Fourth Assessment Report of the IPCC. In your opinion, will they stick to the hockey stick, or will they just dump it, like they dumped the business that Santa [?] had in the Second Report; that has been forgotten. In other words, they pick these things up, they use them, then if they don't work out, they drop them.

McKitrick: Consensus can be a very fragile thing. On my way down here, I have been reading a book called *Conspiracy of Fools,* which is the story of the Enron fiasco. In the year 2000, General Electric lost its claim as the best managed US company and the company that took over – by consensus – as the best-managed US company was Enron and Andrew Fastow of Enron was named CEO of the year. So consensus can be fragile. This is probably more common in business than in science. One of the problems with Enron in retrospect was that nobody knew how Enron made any money and the answer was, they did not. Audits are done in business situations not because businessmen are more honest or are willingly submitting to such procedures because they are good citizens, but because people's money is involved. We are betting a great deal of money on premises that are far less well examined than audited statements. An audited case – Enron – does not mean that the statements were not crooked; there are lots of ways that you can trick processes. But the hockey stick has not even had the equivalent of an Enron audit. It doesn't mean that it is wrong; it just means that we are taking an awful lot of things on trust. Another theme in the Enron scandal was that everyone assumed that somebody else had done the due diligence. In our case, it was an epiphany for me – at some point I realized nobody has ever looked at this stuff! At that point, I decided to keep pulling at strings and see what happens. And lots of interesting things have happened.

Question: Do you think that it is fair to suggest that perhaps that even the co-authors, like Ray Bradley and Malcolm Hughes, really don't have the program and don't really understand it?

McIntyre: I can't comment on it and at the end of the day I don't think it matters.

✻ ✻ ✻