

The Yamal implosion

Sep 29, 2009

Bishop Hill in Climate

*There is a great deal of excitement among climate sceptics over Steve McIntyre's recent posting on Yamal. Several people have asked me to do a layman's guide to the story in the manner of **Caspar and the Jesus paper**. Here it is.*

The story of Michael Mann's Hockey Stick reconstruction, its statistical bias and the influence of the bristlecone pines is well known. McIntyre's research into the other reconstructions has received less publicity, however. The story of the Yamal chronology may change that.

The bristlecone pines that created the shape of the Hockey Stick graph are used in nearly every millennial temperature reconstruction around today, but there are also a handful of other tree ring series that are nearly as common and just as influential on the results. Back at the start of McIntyre's research into the area of paleoclimate, one of the most significant of these was called Polar Urals, a chronology first published by Keith Briffa of the Climate Research Unit (CRU) at the University of East Anglia. At the time, it was used in pretty much every temperature reconstruction around. In his paper, Briffa made the startling claim that the coldest year of the millennium was AD 1032, a statement that, if true, would have completely overturned the idea of the Medieval Warm Period. It is not hard to see why paleoclimatologists found the series so alluring.



Keith Briffa Some of McIntyre's research into Polar Urals deserves a story in its own right, but it is one that will have to wait for another day. We can pick up the narrative again in 2005, when McIntyre discovered that an update to the Polar Urals series had been collected in 1999. Through a contact he was able to obtain a copy of the revised series. Remarkably, in the update the eleventh century appeared to be much warmer than in the original - in fact it was higher even than the twentieth century. This must have been a severe blow to paleoclimatologists, a supposition that is borne out by what happened next, or rather what didn't: the update to the Polar Urals was not published, it was not archived and it was almost never seen again.

With Polar Urals now unusable, paleoclimatologists had a pressing need for a hockey stick shaped replacement and a solution appeared in the nick of time in the shape of a series from the nearby location of Yamal.

The Yamal data had been collected by a pair of Russian scientists, Hantemirov and Shiyatov, and was published in 2002. In their version of the data, Yamal had little by way of a twentieth century trend. Strangely though, Briffa's version, which had made it into print before even the Russians', was somewhat different. While it was very similar to the Russians' version for most of the length of the record, Briffa's version had a sharp uptick at the end of the twentieth century -- another hockey stick, made almost to order to meet the requirements of the paleoclimate community. Certainly, after its first appearance in Briffa's 2000 paper in *Quaternary Science Reviews*, this version of Yamal was seized upon by climatologists, appearing again and again in temperature reconstructions; it became virtually ubiquitous in the field: apart from Briffa 2000, it also contributed to the reconstructions in Mann and

Jones 2003, Jones and Mann 2004, Moberg et al 2005, D'Arrigo et al 2006, Osborn and Briffa 2006 and Hegerl et al 2007, among others.

When McIntyre started to look at the Osborn and Briffa paper in 2006, he quickly ran into the problem of the Yamal chronology: he needed to understand exactly how the difference between the Briffa and Hantemirov versions of Yamal had arisen. McIntyre therefore wrote to the Englishman asking for the original tree ring measurements involved. When Briffa refused, McIntyre wrote to *Science*, who had published the new paper, pointing out that, since it was now six years since Briffa had originally published his version of the chronology, there could be no reason for withholding the underlying data. After some deliberation, the editors at *Science* declined the request, deciding that Briffa did not have to publish anything more as he had merely re-used data from an earlier study. McIntyre should, they advised, approach the author of the earlier study, that author being, of course, Briffa himself. Wearily, McIntyre wrote to Briffa again, this time in his capacity as author of the original study in *Quaternary Science Reviews* and he was, as expected, turned down flat.



That was how the the investigation of the Yamal series stood for the next two years until, in July 2008, a new Briffa paper appeared in the pages of the *Philosophical Transactions of the Royal Society B*, the Royal Society's journal for the biological sciences. The new paper discussed five Eurasian tree ring datasets, which, in fairly standard Hockey Team fashion, were unarchived and therefore not susceptible to detailed analysis. Among these five were Yamal and the equally notorious Tornetrask chronology. McIntyre observed that the only series with a strikingly anomolous twentieth century was Yamal. It was frustratingly therefore that he had *still* not managed to obtain Briffa's measurement data. It appeared that he was going to hit another dead end. However, in the comments to his article on the new paper, a possible way forward presented itself. A reader pointed out that the Royal Society had what appeared to be a fairly clear and robust policy on data availability:

As a condition of acceptance authors agree to honour any reasonable request by other researchers for materials, methods, or data necessary to verify the conclusion of the article...Supplementary data up to 10 Mb is placed on the Society's website free of charge and is publicly accessible. Large datasets must be deposited in a recognised public domain database by the author prior to submission. The accession number should be provided for inclusion in the published article.

Having had his requests rejected by every other journal he had approached, McIntyre had no great expectations that the Royal Society would be any different, but there was no harm in trying and he duly sent off an email pointing out that Briffa had failed to meet the Society's requirement of archiving his data prior to submission and that the editors had failed to check that Briffa had done so. The reply, to McIntyre's surprise, was very encouraging:

We take matters like this very seriously and I am sorry that this was not picked up in the publishing process.

Was the Royal Society, in a striking contrast to every other journal in the field, about to enforce its own data availability policy? Had Briffa made a fatal mistake?

Summer gave way to autumn and as October drew to a close, McIntyre had still heard nothing from the Royal Society. However, in response to some further enquiries, the journal sent McIntyre some more encouraging news -- Briffa *would* be producing most of his data, although not immediately. Most of it would be available by the end of the year, with the remainder to follow in early 2009.

The first batch of data appeared on schedule in the dying days of 2008 and it was something of a disappointment. The Yamal data, as might have been expected, was to be archived with the second batch, so there would be a further delay before the real action could start. Meanwhile, however, McIntyre could begin to look at what Briffa had done elsewhere. It was not to be plain sailing. For a start, Briffa had archived data in an obsolete data format, last used in the era of punch-cards. This was inconvenient, and apparently deliberately so, but it was not an insurmountable problem -- with a little work, McIntyre was able to move ahead with his analysis. Briffa had also thrown a rather larger spanner in the works though: while he had archived the tree ring measurements, he had not supplied any metadata to go with it -- in other words there was no information about where the measurements had come from. All there was was a tree number and the measurements that went with it. However, McIntyre was well used to this kind of behaviour from climatologists and he had some techniques at hand for filling in some of the gaps. Climate Audit postings on the findings followed in fairly short order, some of which were quite intriguing. There was, however, no smoking gun.

There followed a long hiatus, with no word from the Royal Society or from Briffa. McIntyre would occasionally visit Briffa's web page at the CRU website to see if anything new had appeared, but to no avail. Eventually, though, Briffa's hand was forced, and in late September 2009, a reader pointed out to McIntyre that the remaining data was now available. It had been quietly posted to Briffa's webpage, without announcement or the courtesy of an email to McIntyre. It was nearly ten years since the initial publication of Yamal and three years since McIntyre had requested the measurement data from Briffa. Now at last some of the questions could be answered.

When McIntyre started to look at the numbers it was clear that there were going to be the usual problems with a lack of metadata, but there was more than just this. In typical climate science fashion, just scratching at the surface of the Briffa archive raised as many questions as it answered. Why did Briffa only have half the number of cores covering the Medieval Warm Period that the Russian had reported? And why were there so few cores in Briffa's twentieth century? By 1988 there were only 12 cores used, an amazingly small number in what should have been the part of the record when it was easiest to obtain data. By 1990 the count was only ten, dropping still further to just five in 1995. Without an explanation of how the selection of this sample of the available data had been performed, the suspicion of 'cherrypicking' would linger over the study, although it is true to say that Hantemirov also had very few cores in the equivalent period, so it is possible that this selection had been due to the Russian and not Briffa.

The lack of twentieth century data was still more remarkable when the Yamal chronology was compared to the Polar Urals series, to which it was now apparently preferred. The ten or twelve cores used in Yamal was around half the number available at Polar Urals, which should presumably therefore have been considered the more reliable. Why then had climatologists almost all preferred to use Yamal?

Could it be because it had a hockey stick shape?

None of these questions was likely to be answered without an answer to the question of which trees came from which locations. Hantemirov had made it clear in his paper that the data had been collected over a wide area - Yamal was an expanse of river valleys rather than a single location. Knowing exactly which trees came from where might well throw some light onto the question of why Briffa's reconstruction had a hockey stick shape but Hantemirov's didn't.

As so often in McIntyre's work, the clue that unlocked the mystery came from a rather unexpected source. At the same time as archiving the Yamal data, Briffa had recorded the numbers for another site discussed in his Royal Society paper: Taimyr. Taimyr had, like Yamal, also emerged in Briffa's *Quaternary Science Reviews* paper in 2000. However, in the Royal Society paper, Briffa had made major changes, merging Taimyr with another site, Bol'shoi Avam, located no less than 400 kilometres away. While the original Taimyr site had something of a divergence problem, with narrowing ring widths implying cooler temperatures, the new composite site of Avam--Taimyr had a rather warmer twentieth century and a cooler Medieval Warm Period. The effect of this curious blending of datasets was therefore, as so often with paleoclimate adjustments, to produce a warming trend. This however, was not what was interesting McIntyre. What was odd about Avam--Taimyr was that the series seemed to have more tree cores recorded than had been reported in the two papers on which it was based. So it looked as if *something else* had been merged in as well. But what?

With no metadata archived for Avam-Taimyr either, McIntyre had another puzzle to occupy him, but in fact the results were quick to emerge. The Avam data was collected in 2003, but Taimyr only had numbers going up to 1996. Similarly, the Taimyr trees were older, with dates going back to the ninth century. It was therefore possible to make a tentative split of the data by dividing the cores into those finishing after 2000 and those finishing before. This was a good first cut, but the approach assigned 107 cores to Avam, which was more than reported in the original paper. This seemed to confirm the impression that there was something else in the dataset.

At the same time, McIntyre's rough cut approach assigned 103 cores to Taimyr, a number which meant that there were still over 100 cores still unallocated. The only way to resolve this conundrum was by a brute force technique of comparing the tree identification numbers in the dataset to tree ring data in the archives. In this way, McIntyre was finally able to work out the provenance of at least some of the data.

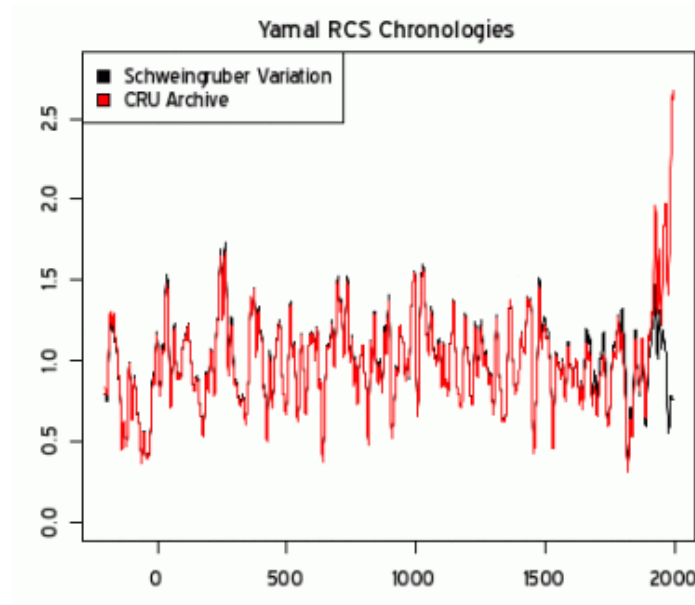
Forty-two of the cores turned out to be from a location called Balschaya Kamenka, some 400 km from Taimyr. The data had been collected by the Swiss researcher, Fritz Schweingruber. The fact that the use of Schweingruber's data had not been reported by Briffa was odd in itself, but what intrigued McIntyre was why Briffa had used Balschaya Kamenka and not any of the other Schweingruber sites in the area. Several of these were much closer to Taimyr -- Aykali River was one example, and another, Novaja Rieja, was almost next door.

By this point then, McIntyre knew that Briffa's version of Yamal was very short of twentieth century data, having used just a selection of the available cores, although the grounds on which this selection had been made was not clear. It was also obvious that there was a great deal of alternative data available from the region, Briffa having been happy to supplement Taimyr with data from other locations such as Avam and Balschaya Kamenka. Why then had he not supplemented *Yamal* in a similar way, in order to bring the number of cores up to an acceptable level?

The reasoning behind Briffa's subsample selection may have been a mystery, but with the other information McIntyre had gleaned, it was still possible to perform some tests on its validity. This could be done by performing a simple sensitivity test, replacing the twelve cores that Briffa had used for the

modern sections of Yamal with some of the other available data. Sure enough, there was a suitable Schweingruber series called Khadyta River close by to Yamal, and with 34 cores, it represented a much more reliable basis for reconstructing temperatures.

McIntyre therefore prepared a revised dataset, replacing Briffa's selected 12 cores with the 34 from Khadyta River. The revised chronology was simply staggering. The sharp uptick in the series at the end of the twentieth century had vanished, leaving a twentieth century apparently without a significant trend. The blade of the Yamal hockey stick, used in so many of those temperature reconstructions that the IPCC said validated Michael Mann's work, was *gone*.



[Updated 30/9/09 to correct minor dating issue. Also removed the reference to KB's illness which is apparently genuine]

Article originally appeared on (<http://www.bishop-hill.net/>).
See website for complete article licensing information.