

Evidence

Submitted to the Independent Climate Change Email Review (ICCER)
Sir Muir Russell, Chairman

Ross McKittrick, Ph.D.
Professor of Economics,
University of Guelph,
Guelph Ontario N1G 2W1
Canada

February 26, 2010

INTRODUCTION	2
TERMS OF REFERENCE QUESTION 1.1	6
1. THE ALLEGATION OF IGNORING POTENTIAL PROBLEMS IN DEDUCING PALAEOTEMPERATURES FROM TREE RING DATA THAT MIGHT UNDERMINE THE VALIDITY OF THE SO-CALLED “HOCKEY-STICK” CURVE.....	6
2. THE ALLEGATION THAT CRU HAS COLLUDED IN ATTEMPTING TO DIMINISH THE SIGNIFICANCE OF DATA THAT MIGHT APPEAR TO CONFLICT WITH THE 20TH CENTURY GLOBAL WARMING HYPOTHESIS.....	20
3. IT IS ALLEGED THAT PROXY TEMPERATURE DEDUCTIONS AND INSTRUMENTAL TEMPERATURE DATA HAVE BEEN IMPROPERLY COMBINED TO CONCEAL MISMATCH BETWEEN THE TWO DATA SERIES.....	25
4. IT IS ALLEGED THAT THERE HAS BEEN AN IMPROPER BIAS IN SELECTING AND ADJUSTING DATA SO AS TO FAVOUR THE ANTHROPOGENIC GLOBAL WARMING HYPOTHESIS AND DETAILS OF SITES AND THE DATA ADJUSTMENTS HAVE NOT BEEN MADE ADEQUATELY AVAILABLE	26
TERMS OF REFERENCE QUESTION 1.2.	43
5. IT IS ALLEGED THAT THERE HAVE BEEN IMPROPER ATTEMPTS TO INFLUENCE THE PEER REVIEW SYSTEM AND A VIOLATION OF IPCC PROCEDURES IN ATTEMPTING TO PREVENT THE PUBLICATION OF OPPOSING IDEAS.	43
6. THE SCRUTINY AND RE-ANALYSIS OF DATA BY OTHER SCIENTISTS IS A VITAL PROCESS IF HYPOTHESES ARE TO RIGOROUSLY TESTED AND IMPROVED. IT IS ALLEGED THAT THERE HAS BEEN A FAILURE TO MAKE IMPORTANT DATA AVAILABLE OR THE PROCEDURES USED TO ADJUST AND ANALYSE THAT DATA, THEREBY SUBVERTING A CRUCIAL SCIENTIFIC PROCESS.	54
7. THE KEEPING OF ACCURATE RECORDS OF DATASETS, ALGORITHMS AND SOFTWARE USED IN THE ANALYSIS OF CLIMATE DATA.	56
TERMS OF REFERENCE QUESTION 1.3.	58
8. RESPONSE TO FREEDOM OF INFORMATION REQUESTS.	58
APPENDIX A: REFERENCES.....	60
APPENDIX B: SUPPORTING PAPER	62

Introduction

Personal details

[1] I am a Professor of Economics at the University of Guelph in Canada where I specialize in environmental economics. In addition to academic publications in the field of economics I have published numerous articles in climatology journals. These are mainly related to statistical methods in paleoclimatic research and the analysis of trends in surface temperatures.

Concerns regarding the composition of the ICCER

[2] All of the ICCER members initially named have sound professional credentials and qualifications. Yet two of the members turned out to have made statements indicative of prejudicial views on the subjects at issue. One panelist (Dr. Campbell) resigned when his statements came to light. Another (Dr. Boulton) has remained on the panel. I list herewith the concerns that I believe are unresolved at this stage.

- [3] Dr. Boulton is a signatory to a petition circulated by the UK Met Office in December (<http://www.metoffice.gov.uk/climatechange/news/latest/uk-science-statement.html>). The petitioners declare “the utmost confidence in the observational evidence for global warming and the scientific basis for concluding that it is due primarily to human activities,” they assert their belief that the scientists who have done the research “adhere to the highest levels of professional integrity” and that the material in question “has been subject to peer review and publication, providing traceability of the evidence and support for the scientific method.” *Yet these are precisely the points under investigation:* whether the observational evidence has been compromised, whether key scientists have acted with less than the utmost integrity, whether the peer review process has been obstructed

and whether evidence actually is traceable. By signing the petition Dr. Boulton has advocated for conclusions that are supposed to be under review.

- [4] The Inquiry claims that none of its members have any links to the CRU (<http://www.cce-review.org/about.php>). Dr. Boulton's CV indicates that he was employed at the University of East Anglia in the School of Environmental Sciences from 1968 until 1986, a fact not revealed on the Inquiry website. It stretches credibility to claim that he could have been at the UEA, in the Environmental Sciences area, for 18 years, without interaction with the CRU. At the very least his long employment at the UEA creates the appearance of a lack of independence.
- [5] The Inquiry has emphasized that its members are not from the climate change field. At a press conference in mid-February Professor Boulton stated [sic] "I am not involved in recent and the issues of recent and current climate nor am I part of that community." He is described on the Inquiry web site as having expertise "in fields related to climate change and is therefore aware of the scientific approach, though not in the climate change field itself." Yet his CV, which his university distributed to Xiamen University (<http://spa.xmu.edu.cn/edit/UploadFile/2007101883249846.doc>), states "His research is in the field of climatic and environmental change and energy, and is an advisor to the UK Government and European Commission on climate change. He leads the Global Change Research Group in the University of Edinburgh, the largest major research group in the University's School of Geosciences." In a 2005 address to the Royal Academy of Engineering, Dr. Boulton said of himself "I am also still a practicing scientist, working on issues such as climate change and nuclear waste disposal..." (http://www.raeng.org.uk/news/publications/list/reports/Ethics_transcripts.pdf). In a January 2008 speech to the Glasgow Centre for Population Health (http://www.gcph.co.uk/component/option,com_docman/task,doc_download/gid,385/) he was introduced with the following comments:

He also heads up the Global Change Research Group which is hosted in Edinburgh and he has just told Carol and I that he has recently arrived back from China where he has been having discussions there with governmental and NGO representatives around global climate change and the role that China and its industrialisation will be playing in that.

He did not gainsay that description, and the talk he gave was a detailed presentation on the subject of climate change. In a speech to the Royal Society of Edinburgh in February 2008 he is reported (<http://www.ma.hw.ac.uk/RSE/events/reports/2007-2008/ecrr.pdf>) as having focused on climate change, saying “I believe that we can currently say that the probability of severe climate change with massive impacts is uncomfortably high.” In a contribution to a report from the David Hume Institute in October 2008 (<http://tinyurl.com/yjok56a>) Professor Boulton wrote a fictional retrospective from 2050 on the subject of climate change, elaborating a pessimistic scenario in which extreme damages from greenhouse gas emissions played out around the world. Other examples can be given of detailed public presentations on climate change, which frequently focus on extreme risks and high-end warming scenarios, and of his public representation as an expert in the field of climate change. Thus it strains credibility for the Inquiry to maintain that Professor Boulton is not “in the field of climate change itself” and for Professor Boulton to say that he is not involved in these issues.

[6] In light of the above, it is reasonable to take the view that Professor Boulton, his impressive credentials notwithstanding, is insufficiently independent of the climate change community in general, and the Climate Research Unit in particular, nor are his stated views on the subject matter sufficiently neutral, to avoid the appearance of bias.

[7] Thus two of the five panelists brought onto the ICCER can reasonably be described as not being impartial. It is somewhat improbable that an Inquiry operating with the utmost neutrality would recruit five members and two of them would turn out to have demonstrated biases in the same direction.

[8] Therefore, I am making this submission accompanied by the objection that the actions of the Inquiry to date have not provided convincing evidence of good faith and neutrality. I understand the enormous responsibility and difficulty of the task confronting members of the ICCER. I will lay out detailed evidence that I believe cannot be ignored in your investigations, even though it may lead you towards conclusions you would strongly prefer not to have to make. Your willingness to confront all the evidence will ultimately determine the credibility of the Inquiry's work.

[9] My submission is organized using the Terms of Reference and "Cross-Examination" document released by the Inquiry at <http://www.cce-review.org/Workplan.php>. Text from the Inquiry is quoted in gray Arial Font.

Terms of Reference Question 1.1

1.1 Examine the hacked e-mail exchanges, other relevant e-mail exchanges and any other information held at CRU to determine whether there is any evidence of the manipulation or suppression of data which is at odds with acceptable scientific practice and may therefore call into question any of the research outcomes.

1. The allegation of ignoring potential problems in deducing palaeotemperatures from tree ring data that might undermine the validity of the so-called “hockey-stick” curve.

In the late 20th century, the correlation between the tree ring record and instrumental record of temperature change diverges from that for the earlier period. The cause of this divergence does not appear to be understood. If the method used to deduce temperatures from tree ring proxy metrics for the earlier tree ring record is applied to the late 20th century tree ring series, then declining temperatures would be deduced for the late 20th century. It is alleged that if the cause of divergence between the tree ring and instrumental temperature record is unknown, it may have existed in earlier periods. Therefore if tree rings had similarly failed to reflect the warming of the early Middle Ages, they may significantly under-estimate the warming during the Medieval Warm Period, thus falsely enhancing the contrast between the recent warming and that earlier period. (It is this contrast that has led to statements that the late 20th century warming is unprecedented during at least the last 1000 years.)

QUESTIONS TO ADDRESS:

What method do you use to deduce palaeotemperatures from tree ring data?

General comments on paleoclimate statistical methods and uncertainty

[10] There are many *ad hoc* methods in use, all of which involve a statistical calibration of temperature and proxy data together. In ordinary regression modeling, a dependent variable is regressed on one or more independent variables, and out-of-sample observations of the independent variables are used to forecast the out-of-sample values of the dependent variable. The challenge in paleoclimate work is that proxies are (in principle) the dependent variable and temperatures are independent, and we seek forecasts of the temperature data rather than the proxy data; in other words forecasting the independent variable given observations of the dependent variable. Hence the paleoclimate calibration problem is an inverse calibration—intuitively the problem involves estimating a confidence interval around the reciprocal of a slope coefficient. In this case, weak correlations between dependent and independent variables greatly amplify the width of confidence intervals, as do conflicting trends among the proxy variables (Brown and Sundberg 1987).

[11] *Ad hoc* methods can conceal the magnitude of these uncertainties, either by simply omitting confidence ellipsoids, which is common, or by generating them using undisclosed and non-standard procedures (such as Mann et al. 1998, 1999). One of the conclusions of the National Research Council Report (North et al. 2006), specifically attributed to the technical contributions I and Stephen McIntyre had made to the panel, was that “uncertainties of the published reconstructions have been underestimated.” (p. 121)

Does not the problem of divergence for the late 20th century record invalidate the deduction of tree ring palaeotemperatures for the period prior to the instrumental record?

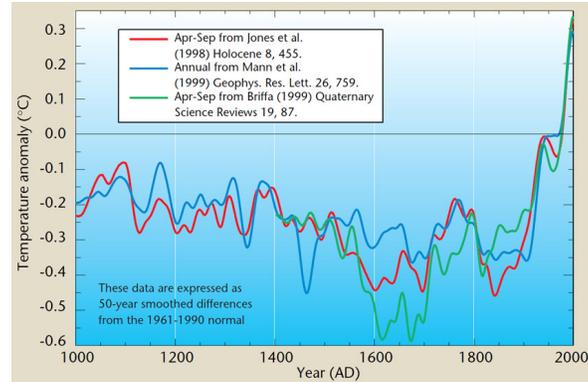
The Divergence Problem

[12] The break-down of the correlation between temperatures and tree rings in the late 20th century leads to widening of the coefficient confidence ellipsoids, and in practice the coefficients become statistically insignificant, i.e. the confidence intervals encompass zero. As a result the confidence intervals around the temperature reconstruction, properly calculated, can become infinitely large. This problem also arises when proxies are inconsistent with one another, not merely with the temperature record. Consequently, merely expanding the proxy sample to include some that do not diverge from temperatures will not solve the problem of indeterminacy if the proxies are inconsistent. McIntyre and McKittrick (2009) raised this point in a comment on Mann et al. (2008), who in their reply did not rebut the point.

How open have you been about this issue?

[13] The divergence problem was well-known during the preparation of the IPCC Report. It was brought up during the meetings at the US National Academy of Sciences for the NRC (2006) report at which I was present. Both Keith Briffa, in his capacity as IPCC Lead Author, and Phil Jones, in his capacity as author of a 1999 World Meteorological Organization Report, have supervised the presentation of graphical data in which the divergence problem is present in the underlying data. The forms those graphs took should be a matter of close scrutiny in the process of answering the questions in your remit.

[14] Jones produced the following diagram for the 1999 WMO Statement on the Status of the Global Climate <http://www.wmo.ch/pages/prog/wcp/wcdmp/statemnt/wmo913.pdf>. The graph appears to show three different proxy series all converging in an impressive unity to reveal a rapid modern warming trend.

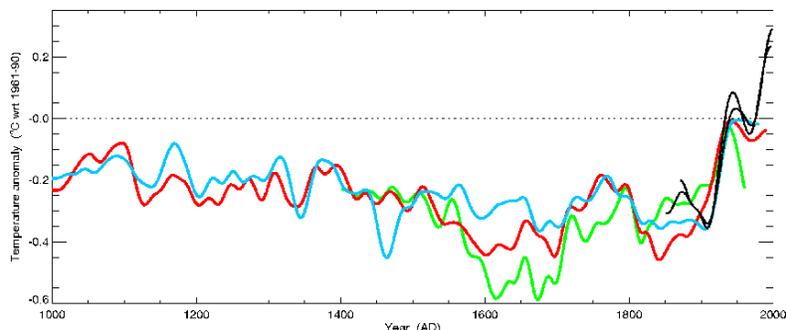


[15] This graph is also disseminated on the UEA website at http://www.uea.ac.uk/polopoly_fs/1.138392!imageManager/1009061939.jpg.

[16] The apparent agreement between the proxy records and the temperature records was achieved by the undisclosed step of replacing the ending 2-4 decades of the proxy records with the CRU temperature series and heavily smoothing over the splice. This is the “trick” referred to in email 942777075.txt wherein Jones says

“I've just completed Mike's Nature trick of adding in the real temps to each series for the last 20 years (ie from 1981 onwards) and from 1961 for Keith's to hide the decline.”

Without this step the diagram would have looked something like this (the black line is the instrumental record):



<http://climateaudit.org/2009/11/20/mike's-nature-trick/>

[17] The problem of Briffa's divergent proxy series was also raised during the preparation of the 3rd IPCC report, as reported in an email from Michael Mann (0938018124.txt), dated September 22 1999.

Keith's series... differs in large part in exactly the opposite direction that Phil's does from ours. This is the problem we all picked up on (everyone in the room at IPCC was in agreement that this was a problem and a potential distraction/detraction from the reasonably consensus viewpoint we'd like to show w/ the Jones et al and Mann et al series.

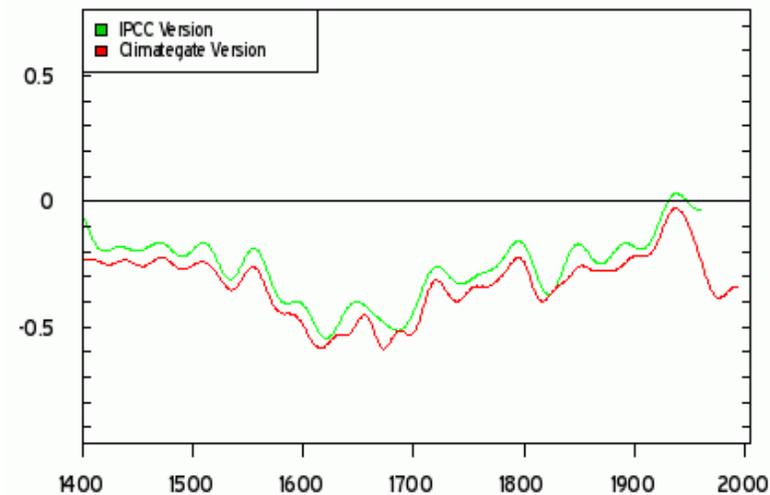
[18] Elsewhere in that email Mann makes it clear that they have no explanation for why Briffa's series diverges from the others, and yet they consider it a priority to present a coherent message so as not to give "fodder" to skeptics.

[19] Briffa, in an email (0938031546.txt) also dated September 22 1999, voiced doubts about the "nice tidy story" that they were pressured to produce, and says

I believe that the recent warmth was probably matched about 1000 years ago. I do not believe that global mean annual temperatures

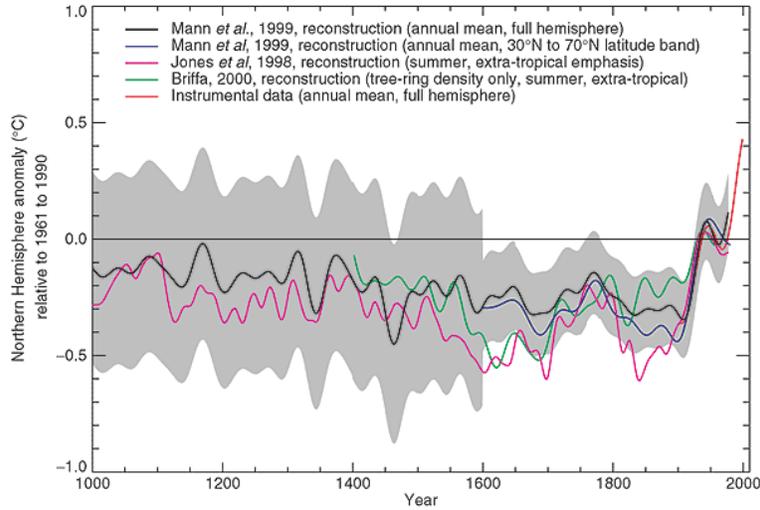
have simply cooled progressively over thousands of years as Mike [Mann] appears to and I contend that that there is strong evidence for major changes in climate over the Holocene (not Milankovich) that require explanation and that could represent part of the current or future background variability of our climate.

[20] These doubts were not reflected in either the text or the graphics of the IPCC Reports. In both the 3rd Assessment Report and the 2007 Fourth Assessment Report, Briffa's graph was truncated at 1960 and no text reflecting the above privately-expressed views went into the Report. The end-result of the above exchange was that Mann agreed to include Briffa's data in his diagram. But he repositioned it so that the first half of the 20th century lined up with other series and the post-1960 portion was deleted. The change is illustrated as follows (original in red, modified in green):



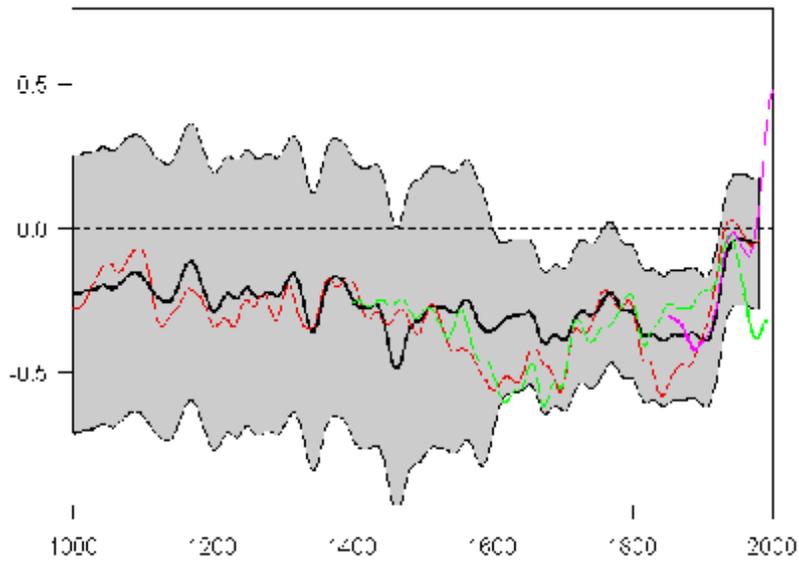
<http://climateaudit.org/2009/12/10/ipcc-and-the-trick/>

[21] The published IPCC image looked like this:



http://www.grida.no/climate/ipcc_tar/wg1/images/fig2-21.gif

[22] Without the deletion of the Briffa data it would have looked something like this:



<http://climateaudit.org/2007/06/26/ipcc-and-the-briffa-deletions/>

[23] I am unaware of any record of Briffa objecting to these manipulations. Such emails may exist, but the same manipulations were employed in the IPCC Fourth Assessment Report when Briffa himself was Lead Author.

[24] In 2006 the draft IPCC report again included a version of the above graph with the post-1960 proxy data deleted. Briffa was the Lead Author for the section. Stephen McIntyre was a chapter reviewer and submitted the following comment:

Show the Briffa et al reconstruction through to its end; don't stop in 1960. Then comment and deal with the "divergence problem" if you need to. Don't cover up the divergence by truncating this graphic. This was done in IPCC TAR; this was misleading. (Reviewer's comment ID #: 309-18)]

[25] Briffa's response was:

Rejected — though note divergence' issue will be discussed, still considered inappropriate to show recent section of Briffa et al. series.

[26] The divergence problem was not discussed in the drafts shown to reviewers, and thus any explanatory text in the final report was inserted without expert review. An email dated March 9 2006 (1141930111.txt) from Eystein Jansen to Phil Jones and Keith Briffa noted that the IPCC text as of the close of the scientific review period (i.e. the Second Order Draft or SOD) still did not deal adequately with the problem of "bad proxies."

One side effect of being stranded and in horizontal working mode is more time to browse the net, thus I have monitored the Climate Audit page. Looking at the discussions after the NAS panel meeting we should expect focus now to be sidetracked from PC-analyses and over to the issue of bad proxies and divergence from

temperature in the last 50 years. Thus this last aspect needs to be tackled more candidly in AR4 than in the SOD, and we need to discuss how to do this, soon.

[27] At some point in July 2006, Briffa sent IPCC chapter materials, including Reviewer Comments, to Eugene Wahl of Alfred College, in apparent contravention of IPCC rules, seeking advice on how to respond to review comments that rebutted his dismissive treatment of the McIntyre-McKittrick critique of the hockey stick. The divergence issue emerged in an interesting way in this exchange. Wahl, along with Mann's former student Caspar Ammann, were coauthors of a paper that defended Mann's interpretation of his data. Wahl was also (as he said in the email thread 1155402164.txt) involved in coaching Congressional witnesses who were going to defend the hockey stick in hearings that summer. As a side note, the emails in this thread show that Briffa was aware of the complexities of this topic as well as his own potential bias, yet instead of seeking guidance from a qualified, neutral party, he instead sought assistance from a partisan on Mann's side.

[28] In an email thread (1155402164.txt) spanning July 21 to August 12 2006, Wahl supplied Briffa with unpublished material that had not gone through the IPCC review process. One of the noteworthy points in that thread is that Briffa had apparently, though perhaps inadvertently, conceded the seriousness of the divergence problem. Wahl urged him to rewrite his response so as not to leave that impression:

I question the way the response to the comment there is currently worded, as it seems to imply that the divergence issue really does invalidate any dendro-based reconstructions before about 1850--which I imagine is not what you would like to say. I give a series of arguments against this as a general conclusion. Maybe I got over-bold in doing so, as in my point (1) I'm examining

issues that are at the very core of your expertise! Excuse me that one, but I decided to jump in anyway. Let me know if I got it wrong in any way!

[29] On July 31 2006 Briffa responded

I do give an implied endorsement of the sense of the whole comment. This is not, of course what I intended. I simply meant to agree that some reference to the "divergence" issue was necessitated . I will revise the reply to say briefly that I do not agree with the interpretation of the reviewer. I am attaching what I have done (see blue highlighting) to the section in response to comments (including the addition of the needed extra section on the "tree-ring issues" called for by several people). I have had no feedback yet on this as it has not been generally circulated , but thought you might like to see it.

[30] Note that by this point Expert Review had been closed for several weeks and nothing Briffa wrote subsequently would be seen by the scientific review group until after publication. In sum: the deletion of the divergent data was done over the objection of IPCC reviewers, the text inserted to rationalize the divergence problem was not shown to expert reviewers, and the response to review comments was revised on the advice of someone outside the IPCC review process whose concern was to downplay the apparent seriousness of the issue.

What attempts have you made to resolve it?

[31] As I have shown above, CRU researchers Keith Briffa and Phil Jones have not “resolved” the problem, despite their direct involvement in the publication of graphs that effectively concealed it. They may believe that there is a good explanation—and indeed there may well be a good explanation. But from the time the problem was noted in 1999, during the preparation of the TAR, to late 2006 during the preparation of AR4, no such explanation had emerged. The 2006 NRC Report (North et al. 2006 pp. 48—52) pointed to a few conjectures, including regional precipitation changes and stratospheric ozone depletion, but they did not report any resolution of the problem, and it cannot be assumed that there is one. The Memorandum of the University of East Anglia to the Parliamentary Inquiry (page 4, paragraph 3.5.4) attempts to mitigate the problem by stating:

The use of the term “hiding the decline” referred to the method of combining the tree-ring evidence and instrumental temperatures, removing the post-1960 tree-ring data to avoid giving a false impression of declining temperatures.

But this begs the question. If the tree rings really are temperature-sensitive, then their decline cannot be assumed to be a “false impression” *unless specific evidence has been shown to account for it*. If they are not temperature sensitive, then their pre-19th century values should not be used as temperature data. Far from avoiding “a false impression,” the removal of the post-1960 tree-ring data created a false impression of certainty on a topic subject to great and ongoing uncertainty.

What is the evidence that the amplitude of warming during the Medieval Warm Period (MWP) is not underestimated by tree ring evidence?

[32] The necessary evidence on this point would consist of an identifiable third factor mediating the temperature-proxy relationship that could be quantified and put into a calibration regression model. Upon

controlling for its influence, the parameterized relationship between the proxy and temperature measures should become positive and significant in the recent era, justifying the assumption that such a relationship holds in eras when the third factor is low or non-existent. Qualitative conjectures are no substitute for empirical evidence. It is not sufficient for paleoclimate researchers simply to guess that some unspecified variable accounts for the divergence, and without any statistical proof, simply delete the divergent data portions. I do not believe this can be considered sound scientific practice.

How does the tree ring evidence of the MWP compare with other proxy data? Have you showed how data from different sources compare or have you conflated them? If the latter, what is the justification?

[33] This is a large issue subject to ongoing debate. Loehle and McCullough (2008, a corrected version of Loehle's 2007 paper) published a reconstruction derived solely from non-tree ring proxies, which shows an elevated MWP.

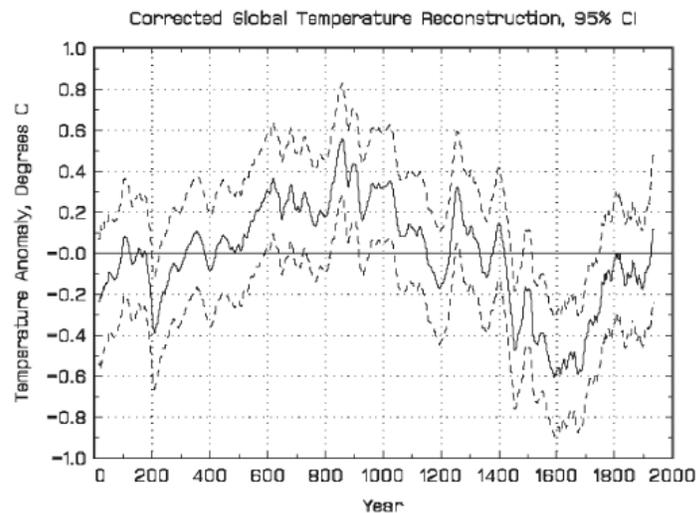


Figure 2. Corrected reconstruction with 95% confidence intervals.
Data for this graph is online at <<http://www.econ.ohio-state.edu/fhm/AGW/Loehle/>>

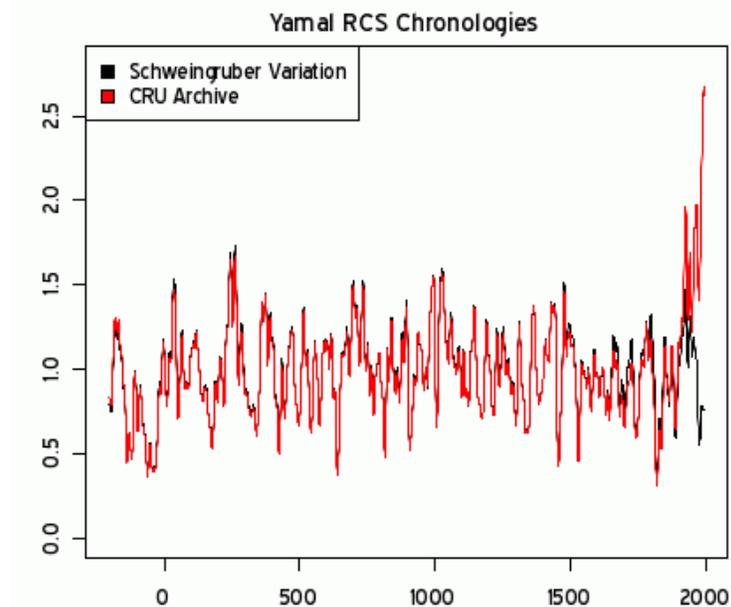
If tree ring proxies are removed from reconstructions, what evidence remains of the MWP?

[34] In the case of the Mann et al. hockey stick graph, removal of the small group of bristlecone pine proxies is sufficient to eliminate the hockey stick shape under all variations and permutations of methodology. This point has been acknowledged by all parties, including McIntyre and McKittrick (2005), Wahl and Ammann (2007), the NRC report (2006), etc.

The Yamal Chronology

Have you been selective in utilizing tree ring evidence from Yamal in Siberia; and if so, what is the justification for selectivity and does the selection influence the deduced pattern of hemispheric climate change during the last millennium?

[35] Briffa's Yamal series was presented in Briffa (2000), which did not provide "core counts" – the number of cores contributing to the chronology. The core counts by decade were not made available until late 2009, and they showed that the sample fell from over 300 in early years to 10 in 1990, then 5 in 1995, well below replication standards. Yet in the meantime the Yamal chronology had been used as an input to several published climate reconstructions, including ones cited by Briffa in his capacity as IPCC Lead Author.



<http://climateaudit.org/2009/09/27/yamal-a-divergence-problem>.

Red: Briffa's Yamal data;

Black: Same with Schweingruber data replacing the CRU archive data after 1800.

[36] The release of the data used in the Yamal chronology showed that the sample size drops rapidly in the 20th century and collapses right at the point (~1990) where the graph's most remarkable behaviour emerges, namely the sharp ending trajectory that creates the hockey stick shape. Hence the remarkable behaviour is not a robust feature of the full data set but coincides with the point where the sample size collapses. A larger nearby sample (Khadyta River) by Schweingruber trends downward over this interval (black line above). There may be a legitimate reason for limiting the input series to the one site and not using the Schweingruber series from nearby to maintain the sample size. But the rapid drop in the sample size ought to have been reported to readers.

2. The allegation that CRU has colluded in attempting to diminish the significance of data that might appear to conflict with the 20th century global warming hypothesis

The CRU group, in consultation with Professor Michael Mann, is alleged to have systematically attempted to diminish the significance of the Medieval Warm Period, evidenced by an email from Mann 4th June 2003: “I think that trying to adopt a timeframe of 2K, rather than the usual 1K, addresses a good earlier point that Peck made w/ regard to the memo, that it would be nice to try to “contain” the putative “MWP”, even if we don't yet have a hemispheric mean reconstruction available that far back [Phil and I have one in review--not sure it is kosher to show that yet though--I've put in an inquiry to Judy Jacobs at AGU about this].” The use of the words “contain” and “putative” are alleged to imply an improper intention to diminish the magnitude and significance of the MWP so as to emphasise the late 20th century warming.

QUESTIONS TO ADDRESS

What does the word “contain” mean in this context?

[37] Interpreting “contain” to imply an attempt to diminish the perceived magnitude of the MWP in comparison to the modern era, is the obvious, prima facie meaning. It is also the reading that makes this email consistent with the other discussions quoted above, especially in paragraphs [14] to [23], which deal with the desire to present a “tidy” story that shows modern warming unusually high compared to the MWP.

What involvement have you had in “containing” the MWP?

[38] See previous section.

How important is the assertion of “unprecedented late 20th century warming” in the argument for anthropogenic forcing of climate?

The MWP Question

[39] The answer to this question has been provided in part by Professor Jones in an interview with the BBC on February 13 2010, in which he said, in part, ““Of course, if the MWP was shown to be global in extent and as warm or warmer than today, then obviously the late 20th Century warmth would not be unprecedented.” This is, at one level, a mere truism. But the larger point is that, since the overall effect of greenhouse gases on the Earth’s temperature field cannot be derived from first principles (owing to the involvement of inscrutable processes such as convection, cloud feedbacks, etc), empirical evidence must be used. This could take the form of showing that the climate has recently moved out of the bounds of natural variability in comparison to the past one or two millennia. The importance the IPCC attached to showing that the modern era is unusually warm is revealed by the conspicuous efforts the IPCC has made to highlight studies, such as the Mann hockey stick, that assert the view that modern warming is unprecedented, the many efforts made in the last IPCC report to denigrate research critical of that view, and the determination to use the Wahl and Ammann work, even to the point of redefining the cut-off date for using in-press literature (see submission to this Inquiry by David Holland) and allowing Wahl to supply unreviewed information to Briffa for use in Chapter 6 through backchannels (see paragraphs [27]-[30]).

[40] If it were shown that the MWP was globally warmer than the present, even though CO₂ levels were likely much lower, it might call into question whether CO₂ is a primary climate driver on century time-

scales, whether natural variability may be larger than is currently thought, and whether the climate sensitivity to CO₂ can be as high as is generally assumed, if increased levels had not brought about any more warming than was experienced naturally in the past. For that reason it is notable that the IPCC 2007 Summary for Policymakers claimed:

Paleoclimate information supports the interpretation that the warmth of the last half century is unusual in at least the previous 1300 years

[41] Briffa had privately expressed doubts about such a view at the time of the previous IPCC report (see paragraph [19]). An email exchange (1051638938.txt) between Ed Cook and Briffa in April 2003 revealed their continuing doubts on the question, even after the 2001 IPCC Report had displayed the hockey stick graph so prominently. Cook to Briffa:

[Ray] Bradley still regards the MWP as "mysterious" and "very incoherent" (his latest pronouncement to me) based on the available data. Of course he and other members of the [Mann-Bradley-Hughes] MBH camp have a fundamental dislike for the very concept of the MWP, so I tend to view their evaluations as starting out from a somewhat biased perspective, i.e. the cup is not only "half-empty"; it is demonstrably "broken". I come more from the "cup half-full" camp when it comes to the MWP, maybe yes, maybe no, but it is too early to say what it is. Being a natural skeptic, I guess you might lean more towards the MBH camp, which is fine as long as one is honest and open about evaluating the evidence (I have my doubts about the MBH

camp). We can always politely(?) disagree given the same admittedly equivocal evidence.

[42] Briffa replied:

Can I just say that I am not in the MBH camp - if that be characterized by an unshakable "belief" one way or the other , regarding the absolute magnitude of the global MWP. I certainly believe the " medieval" period was warmer than the 18th century - the equivalence of the warmth in the post 1900 period, and the post 1980s , compared to the circa Medieval times is very much still an area for much better resolution. I think that the geographic /seasonal biases and dating/response time issues still cloud the picture of when and how warm the Medieval period was . On present evidence , even with such uncertainties I would still come out favouring the "likely unprecedented recent warmth" opinion - but our motivation is to further explore the degree of certainty in this belief - based on the realistic interpretation of available data.

[43] This discussion took place in 2003, and by common agreement there has been no significant warming since then, so the lack of clarity about the relative ranking of the medieval/modern climatic state would not have changed by the time the IPCC Report was being prepared in 2004/2005.

[44] In his recent BBC interview Jones states that the basis for concluding the MWP was warmer than the present is still heavily disputed, and cannot be settled with current data.

There is much debate over whether the Medieval Warm Period was global in extent or not. The MWP is most clearly expressed in parts of North America, the North Atlantic and Europe and parts of Asia. For it to be global in extent the MWP would need to be seen clearly in more records from the tropical regions and the Southern Hemisphere. There are very few palaeoclimatic records for these latter two regions.

<http://news.bbc.co.uk/2/hi/science/nature/8511670.stm>

[45] I make no criticism of scientists comparing notes and expressing doubts on complex research questions. The issue here is that *in private* and *among themselves* CRU scientists expressed relatively high levels of uncertainty about the MWP/modern comparison compared to what they were saying through IPCC and WMO Reports. The suppression of legitimate, known uncertainties for the purpose of sharpening up communication to policymakers is a form of activism. In the Iraq war intelligence reporting issue it was referred to as “sexing up” a report. It was also fingered as a key contributing factor in the 1986 Challenger disaster.

3. It is alleged that proxy temperature deductions and instrumental temperature data have been improperly combined to conceal mismatch between the two data series

An attempt to hide the difficulty of combining these two data series and to mislead is alleged to be revealed in the following sentence in a November 1999 email from Professor Phillip Jones which is alleged to imply a conscious attempt to mislead: “I’ve just completed Mike’s Nature trick of adding in the real temps to each series for the last 20 years (i.e. from 1981 onwards) and from 1961 for Keith’s to hide the decline”.

QUESTIONS TO ADDRESS

What is the meaning of the quotation from the 1999 email?

[46] See paragraph [16]. The word “trick” is not the issue. The quotation and the actions to which it refers would be equally objectionable had the word “procedure” been used instead. The problematic wording is “hide the decline.” It reveals that Jones knew the proxy data showed a decline, and he employed a technique that concealed the fact and showed a uniform increase instead. There is no ambiguity on this point and the context does not change anything about the meaning.

How do you justify combining proxy and instrumental data in a single plotted line?

[46] As an academic matter, scientists combine different types of data all the time for the purpose of extracting information and constructing statistical models. As long as the methods are clearly explained and the reader is given the information necessary to evaluate the quality of the calibration/fitting process,

there is nothing wrong with this, and indeed it is often the path to important discoveries and progress. But in the case of the preparation of the WMO and IPCC diagrams, the problem is that readers were not told about the way different data sets were being trimmed and/or combined, hence materially adverse information was withheld from readers, thus exaggerating the quality of the statistical model.

What method do you use?

[47] Many methods are used, see paragraph [10].

4. It is alleged that there has been an improper bias in selecting and adjusting data so as to favour the anthropogenic global warming hypothesis and details of sites and the data adjustments have not been made adequately available

It is alleged that instrumental data has been selected preferentially to include data from warmer, urban in contrast to rural sites; that the rationale for the choice of high/low latitude sites is poor; and that the processes by which data has been corrected, accepted and rejected are complex and unclear.

QUESTIONS TO ADDRESS

What is the rationale for the choice of data stations worldwide?

Jones' disclosure of CRU Input data

[48] Any answer you receive to this question from CRU cannot be independently verified since Jones has not clarified the list of stations he used in the CRUTEM3 compilation. Up to around 2003 Jones had been

forthcoming and courteous about explaining the inputs used to produce the CRU data. The 1985 technical reports to the US Department of Energy are, indeed exhaustive (<http://www.cru.uea.ac.uk/st/>). But they refer to data sets that have since been superseded, so they are not adequate for understanding the post-1980 CRUTEM series. At <http://climateaudit.org/2009/08/06/a-2002-request-to-cru/> Stephen McIntyre relates that in 2002 he had asked Jones for a list of stations used for an earlier CRU dataset. Jones promptly replied with a list of stations and their data, but he cautioned that those data were out of date. He pointed to the forthcoming CRUTEM2 edition and said:

Once the paper comes out in the Journal of Climate, I will be putting the station temperature and all the gridded databases onto our web site.

[49] McIntyre notes at the above web page that the promised disclosure never took place, but by 2003 he had moved onto looking at the hockey stick issue and did not pursue his request for the station data.

[50] In July 2004 Jones received a request from Warwick Hughes for the location of the CRUTEM2 station data (see <http://climateaudit.files.wordpress.com/2008/05/cru.correspondence.pdf>). Jones replied by referring Hughes to a division of the WMO for the information.

[51] To my knowledge, from then until February 23 2005 Jones received no subsequent requests for the data from individuals outside his own circle of collaborators, and in particular, he received none from me or Stephen McIntyre. On February 2 2005, Jones emailed to Mann (1107454306.txt)

Just sent loads of station data to Scott. Make sure he documents everything better this time ! And don't leave

stuff lying around on ftp sites - you never know who is trawling them. The two MMs have been after the CRU station data for years. If they ever hear there is a Freedom of Information Act now in the UK, I think I'll delete the file rather than send to anyone.

[50] 'MM' refers to McIntyre and McKittrick, as is clear since the surrounding conversation refers to data connected to the hockey stick. The above email is important for several reasons.

- [51] There have been suggestions that Jones was under siege with requests for data and became uncooperative out of sheer frustration. As superficially plausible as this sounds, the timeline shows it is untrue. Jones' remark to Mann that he would delete data rather than share it, was made *before* he had received data requests. At the point in time that Jones wrote the above email he had received one request for a list of station identifiers from McIntyre back in 2002, in reply to which he had promised to post the information (but did not do so), and one request from Warwick Hughes the summer of 2004 for a list of stations, in response to which he had referred Hughes to the WMO. I had never contacted Jones asking for his station data, and apart from his 2002 request neither had McIntyre, nor had we any given any indication of planning to do so during this interval. Any campaign by McIntyre and me to get the station data was in Jones' imagination.
- [52] Even if we had made such requests, it is unclear why that should be seen as a vexation since he had already promised to McIntyre that he would post the data on the CRU website, and it is only fitting that a data product as prominent as CRUTEM should be as transparent as possible.

- [53] The reference to sending “loads of station data” to Scott, presumably meaning Mann’s associated Scott Rutherford, indicates that Jones was willing and able to send the station data when so inclined. It appears that his inclination was influenced by a preference for uncritical recipients. This is borne out by the fact that Warwick Hughes emailed Jones later that same month (February 18 2005) telling him that the WMO had not replied to his emails, and asking for another contact person. Jones replied that he was traveling and would reply soon. Before doing so, on February 21, Jones wrote (1109021312.txt) to Mann, Bradley and (Malcolm) Hughes

I'm getting hassled by a couple of people to release the CRU station temperature data. Don't any of you three tell anybody that the UK has a Freedom of Information Act !

Jones then responded to Warwick Hughes on February 23, saying in part:

Even if WMO agrees, I will still not pass on the data. 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.

This email is not in the East Anglia compilation, but it is available online at <http://climateaudit.files.wordpress.com/2008/05/cru.correspondence.pdf>. It is hard to imagine a sentiment more antithetical to good science. Once again, it was not sent by someone who was being “besieged” with unreasonable requests for data, it was sent by someone who had, to that point, only received two requests over the previous three years, for data he had already promised to release, and who had readily shared “loads of” the data with a trusted colleague.

[54] Following the publication of the CRUTEM3 data series (Brohan et al. 2006), it was not possible to discern from information on the CRU website, or in accompanying publications, which locations and weather stations had been used to produce the gridcell anomalies. On September 28 2006 Willis Eschenbach and Douglas Keenan filed an FOIA request for the list of meteorological stations used for the CRUTEM3 data product. This request was rejected by David Palmer of the University of East Anglia on February 10 2007 on the alleged grounds that CRU input data was already published on websites at the Global Historical Climatology Center (GHCN) and the National Center for Atmospheric Research (NCAR), both in the US. CRU also claimed that they had sent all their data to the GHCN and it was thus publicly available. However these archives contain large collections of station series, only some of which are used by CRU. Without knowing which ones are selected, it is not possible to back out the CRU data set from the GHCN and NCAR archives.

[55] Eschenbach and Keenan appealed the decision on the grounds that without the station identifiers it was impossible to know which data series had been used in the CRUTEM3 series, even if the full archives are on the internet. On April 12 the UEA again rejected the request, pointing again to the GHCN and NCAR archives and saying that “more than 98% of the CRU data are on these sites.” Eschenbach appealed again, reiterating that without the WMO identifiers it would be impossible to tell which of the thousands of GHCN and NCAR data series had been used by CRU. He specified that he was only looking for a list of station IDs and locations. On April 23 2007 a document was created at the CRU (<http://www.climate-gate.org/cru/documents/jones-foiathoughts.doc>) comparing 3 response options, one of which was simply to send the data. The other two involved deleting portions of the data in ways that “would annoy them.” Four days later the FOIA request was refused outright, on April 27 2007, on the basis of the claim that CRU no longer had a list of the stations it used:

We cannot produce a simple list with this format and with the information you described in your note of 14 April. Firstly, we do not have a list consisting solely of the sites we currently use. Our list is larger, as it includes data not used due to incomplete reference periods, for example. Additionally, even if we were able to create such a list we would not be able to link the sites with sources of data. The station database has evolved over time and the Climate Research Unit was not able to keep multiple versions of it as stations were added, amended and deleted. This was a consequence of a lack of data storage in the 1980s and early 1990s compared to what we have at our disposal currently. It is also likely that quite a few stations consist of a mixture of sources. (<http://climateaudit.files.wordpress.com/2008/05/cru.correspondence.pdf>)

[56] The above statement is a striking contrast to the exhaustive disclosure in the 1985 DOE Reports (<http://www.cru.uea.ac.uk/st/>) Eschenbach immediately appealed this decision to the Information Commissioner, who worked out an agreement that a list of stations with WMO identifiers would be released, but it would not indicate which stations were used at which points in time, nor would it indicate which stations are currently in use, nor the sources of the data, nor any of the data adjustment code. Again, this was a remarkable departure from the generous disclosure of the 1985 reports. As noted in the April FOIA refusal, the CRU now claimed only to have a large and inaccurate list of stations, some of which had not actually been used. That file was eventually posted in October 2007 at <http://www.cru.uea.ac.uk/cru/data/landstations>. Thereafter Eschenbach abandoned his inquiries for data to the CRU. In 2007 CRU received some FOIA requests regarding unpublished materials used in the 2007 IPCC Report. As far as I am aware, further inquiries about CRUTEM data did not come until June 2009. Meanwhile, on June 19 2007, Jones wrote to two colleagues (Wang and Peterson, 1182255717.txt)

Think I've managed to persuade UEA to ignore all further FOIA requests if the people have anything to do with Climate Audit.

[57] In May 2007 Doug Keenan wrote to Jones, referring to Jones' claim that there were non-disclosure agreements preventing release of station data. Keenan asked which countries were covered by these agreements. In Jones' reply he listed Germany, Bahrain, Oman, Algeria, Japan, Slovakia, Syria, Mali, India, Pakistan, Poland, Indonesia, Democratic Republic of the Congo, Sudan and "some Caribbean Islands."

[58] Two years later, in May 2009, Stephen McIntyre observed that there was a note on the Hadley Centre Website (<http://hadobs.metoffice.com/indicators/index.html>) saying

To obtain the archive of raw land surface temperature observations used to create CRUTEM3, you will need to contact Phil Jones at the Climate Research Unit at the University of East Anglia. Recently archived station reports used to update CRUTEM3 and HadCRUT3 are available from the CRUTEM3 data download page.

(this has since been deleted). McIntyre wrote the Hadley Centre asking for a copy of the data that they had received from the CRU. When this request was refused on the grounds that Jones had forbidden them to pass it on, McIntyre submitted an FOIA request to the Hadley Centre for the archive. (See <http://climateaudit.org/2009/06/04/the-uk-hadley-center-refuses-crutem-data>). This was refused with the claim that the Hadley Centre did not receive station data from CRU (despite earlier saying they had it but were not allowed to share it), only the gridded (or "value-added") CRUTEM data. McIntyre then requested the Met Office supply him with the source data and "*documents that you hold describing the*

procedures under which the data has been quality controlled and where deemed appropriate, adjusted to account for apparent non-climatic influences.” This request was rejected on the grounds that

The Met Office received the data information from Professor Jones at the University of East Anglia on the strict understanding by the data providers that this station data must not be publicly released. If any of this information were released, scientists could be reluctant to share information and participate in scientific projects with the public sector organisations based in the UK in future. It would also damage the trust that scientists have in those scientists who happen to be employed in the public sector and could show the Met Office ignored the confidentiality in which the data information was provided. (<http://climateaudit.org/2009/07/23/uk-met-offices-refuses-to-disclose-station-data-once-again/>)

[59] Note that the above statement contradicts the reason given to Eschenbach and Keenan for refusing their 2007 request (paragraph [55]), namely that there is no need to comply with the FOIA request because 98% of the CRU station data is already in the public GHCN/NCAR archives and the CRU had supplied all its station data to the public GHCN archive.

[60] On June 25 2009, Peter Webster of Georgia Tech told McIntyre that he had asked for station data from Jones and it had been sent to him. McIntyre sent an FOIA request for the data supplied to Webster, but it was rejected by David Palmer on the grounds that “the information requested was received by the University on terms that prevent further transmission to non-academics.” (<http://climateaudit.org/2009/07/24/cru-refuses-data-once-again/>) On July 24 I submitted the request, pointing out that I am an academic working in the field of the assessment of surface temperature data quality (see correspondence at <http://sites.google.com/site/rossmckittrick/CRUdata.pdf>). This was rejected on August 13 2009 on two grounds: there is no need to supply the data I requested because it is

already released on the internet at the GHCN archive, and it is not possible to supply the data because its release by the CRU is prevented under the terms of agreements with persons and agencies in other countries who had supplied the data. The contradiction between these two claims is obvious, as is the contradiction with the claim to Eschenbach and Keenan that the CRU data had sent all its data to the GHCN (paragraph [54]). Confronted with incoherent and implausible reasons for not releasing the data, the readers of ClimateAudit decided to ask to see the non-disclosure agreements. Since a pattern of apparent stonewalling was by now established, we decided to request the texts of agreements on a country-by-country basis.

[61] Claims that the CRU was besieged by a flood of FOIA requests in summer 2009 (such as paragraph 3.7.4 in the UEA submission to the UK Parliamentary Inquiry) pertain to this final phase in the process. Some allege that the CRU endured a long campaign of multiple, frivolous FOIA requests for data they were not allowed to release, and out of understandable frustration decided to stonewall them. This is untrue. Instead, the CRU had been asked for data they had previously said they were at liberty to share, and that they had already shared with international colleagues. After getting frustrated by the CRU's increasingly implausible refusals, researchers resorted to the FOIA process to ascertain the nature of the alleged non-disclosure agreements. Also the dozens of FOIA requests in July 2009 were not for data, they were pro forma requests for the texts of the non-disclosure agreements that the CRU cited as grounds for not releasing its data. As it turned out, responding to each such request was not onerous since each response was identical. It referred to a web page at the CRU (<http://www.cru.uea.ac.uk/cru/data/availability/>) that listed a few such agreements, and of the rest it said "We know that there were others, but cannot locate them, possibly as we've moved offices several times during the 1980s."

[62] The UEA Memorandum to the Parliamentary Inquiry also disputes reports (paragraph 3.7.1) that the CRU lost or discarded its raw data. But the CRU itself makes this very claim on its web page <http://www.cru.uea.ac.uk/cru/data/availability>:

“Data storage availability in the 1980s meant that we were not able to keep the multiple sources for some sites, only the station series after adjustment for homogeneity issues. We, therefore, do not hold the original raw data but only the value-added (i.e. quality controlled and homogenized) data.”

However this claim is still misleading. The CRU clearly holds station data in some form (i.e. that which was shared with Scott Rutherford and Peter Webster). These series are used to produce the CRUTEM gridded products, and these are the series that were sought by Warwick Hughes in 2005, Willis Eschenbach in 2007 and Stephen McIntyre in 2009. Even if they are not the absolute raw data, they are still the basic inputs to the CRUTEM gridcells, and as such these are the records we need to see if we are to understand how the CRUTEM products are generated. The CRU refused (and still refuses) to release these records. The comment that these records themselves reflect some initial processing, and the absolute raw data are no longer available, is irrelevant. What appears to be most relevant in this whole episode is Jones’ original rebuff to Warwick Hughes in 2005, before the requests for data had even begun, when he said

Why should I make the data available to you, when your aim is to try and find something wrong with it.

How has this choice been tested as appropriate in generating a global or hemispheric mean temperature (both instrumental and proxy data)?

Documentation of the inhomogeneity corrections

[63] In the next few paragraphs I will argue that the choices of station inputs and the adjustment algorithms have not been adequately tested by Jones and other CRU scientists, and where others outside the CRU have done the testing the results have revealed strong evidence of contaminating non-climatic influences in the CRU data. Much of the discussion on this point has been misplaced to the extent it asks *whether* the data have been adjusted and whether the adjustments have been documented. I will show that not only is documentation of the adjustments inadequate, but also that independent testing of the adjustments has shown important problems likely remain in the data.

[64] Everyone who works on climatic research knows that assembling a global surface temperature data set is a large task, and a debt of gratitude is owed to those who have undertaken it. But gratitude does not warrant an exemption from critical scrutiny. My research has focused the post-1979 interval which displays a strong upward global trend. Climatic data are not simply temperature records. It is universally acknowledged that temperatures at land-based observational sites can be affected by changes in the local land surface due to deforestation, introduction of agriculture, road-building, urbanization, changes in monitoring equipment, measurement discontinuities, and so forth; as well as by local emissions of particulates and other air pollutants. These are *non-climatic* influences, since they are driven by local, and in principle reversible, changes, rather than global climatic forcing. Hence the raw temperature record must be adjusted, if possible, to reveal the *climatic* record. An ideal record of surface climatic changes would require a monitoring site untouched by human development, the equipment for which was consistent and perfectly maintained over the entire measurement interval. However the actual data used to produce climate data sets almost never satisfies these ideals. Consequently, data sets published as “climate” records are not simply observations: they are the outputs of models that take weather records as inputs, apply adjustments aimed at removing non-climatic influences, group the resulting records into

regional grids and then translate the data into deviations from local averages, yielding what are called gridded climate “anomalies”.

[65] The problems with raw temperature data are acknowledged by the CRU. The CRU web page (<http://www.cru.uea.ac.uk/cru/data/hrg/>) references data compilations called CRU TS 1.x, 2.x and 3.x which are not subject to adjustments for non-climatic influences. Users are explicitly cautioned not to use the TS data for measuring or analyzing climate change in the ways applicable to IPCC reports. The 1.2 release of this product provided a list of FAQ’s related to time series analysis (see <http://www.cru.uea.ac.uk/cru/data/hrg/timm/grid/ts-advice.html>). The first question, and its answer, are reproduced (in part) below.

Question One

Q1. Is it legitimate to use CRU TS 2.0 to 'detect anthropogenic climate change' (IPCC language)?

A1. No. CRU TS 2.0 is specifically *not* designed for climate change detection or attribution in the classic IPCC sense. The classic IPCC detection issue deals with the distinctly anthropogenic climate changes we are already experiencing. Therefore it is necessary, for IPCC detection to work, to remove all influences of urban development or land use change on the station data....If you want to examine the detection of anthropogenic climate change, we recommend that you use the [Jones](#) temperature data-set. This is on a coarser (5 degree) grid, but it is optimised for the reliable detection of anthropogenic trends.

[66] The implication is that the Jones data has been adjusted “for the reliable detection of anthropogenic trends.” Readers are referred to some academic papers for explanation of the adjustments. The first is Brohan et al. (2006). It does not explain how the data are adjusted, instead it focuses on defending the claim that the potential biases are very small, for which two references are cited in support. One is by US

scientist Thomas Peterson, which refers to the contiguous US only. Another is by David Parker of the Hadley Centre, whose argument relied on an apparent similarity between trends on windy and calm nights. None of the published literature critical of Parker's methods are cited. Section 2.3.3 of Brohan et al. states that to properly adjust the data would require a global comparison of urban versus rural records, but classifying records in this way is not possible since "no such complete meta-data are available" (p. 11), so the authors instead impose the assumption that the bias is no larger than 0.006 degrees per century. This assumption reappears in the 2007 IPCC Summary for Policymakers as a research finding (see paragraph [86]).

[67] Brohan et al. refer to a 2003 paper in *Journal of Climate* by Jones and Moberg, explaining the CRUTEM version 2 data product. This paper also has little information about the data adjustments. Reference is made to combining multiple site records into a single series, but not to removing non-climatic contamination. Moreover, the article points out (page 208) that it is difficult to say what homogeneity adjustments have been applied since the original data sources do not always include this information.

[68] The other reference on the CRU website is to a 1999 *Reviews of Geophysics* paper by Jones, New, Parker et. al. This paper emphasizes that non-climatic influences (therein referred to as "inhomogeneities") must be corrected (Section 2, p. 37) for the data to be useful for climatic research. The only part of the paper that provides information on the adjustments is Section 2.1, consisting of only 3 paragraphs, none of which explains the CRU procedures. The only explanatory statement is (page 174):

"All 2000+ station time series used have been assessed for homogeneity by subjective interstation comparisons performed on a local basis. Many stations were adjusted and some

omitted because of anomalous warming trends and/or numerous nonclimatic jumps (complete details are given by Jones et al. [1985, 1986c]).”

[69] Jones et al. [1985, 1986c] are technical reports that were submitted to the US Department of Energy, and are posted at <http://www.cru.uea.ac.uk/st/>. They only cover data sets ending in the early 1980s, whereas the data currently under dispute is the post-1979 interval. Those documents caution (page 3) that even with station-by-station examination, correction of all the problems is not possible due to insufficient detail in the site records to calculate correction factors. Even if the adjustments were adequate in the pre-1980 interval it is likely impossible to have estimated empirical adjustments in the early 1980s that would apply to changes in socioeconomic patterns that did not occur until the 1990s and after. Also, Jones had told McIntyre in 2002 that data sets published prior to that point are “out of date” <http://climateaudit.org/2009/08/06/a-2002-request-to-cru/>. Hence these reports do not provide the disclosure necessary for understanding how CRUTEM3 was assembled.

[70] In sum, the CRU cautions that its unadjusted temperature data products (TS 2.x etc.) are inappropriate for the IPCC’s purpose, and for detection and attribution analysis more generally. The CRU refers users instead to the CRUTEM products. Yet the accompanying documentation does not appear to explain the adjustments made or the grounds for claiming the data products are reliable for climate research purposes.

Independence from NASA and NOAA Temperature Products

[71] The papers provide tables of sources for the CRUTEM input data, from which it can be inferred that a substantial portion are from the Global Historical Climatology Network (GHCN) maintained by NOAA. CRU has elsewhere stated that 98% of their data come from GHCN (paragraph [55] above). The GHCN data are also used as inputs for the NASA and NOAA global temperature series. Hence the three

global climate data series are not independent. However CRU also claims to depend so heavily on unpublished data from other countries covered by non-disclosure agreements, that it cannot release its input data. So the extent of actual overlap cannot be determined without knowing exactly which GHCN series are used for the CRU data set, which was one of the points subject to Freedom of Information requests described above. In addition, without provision of the non-GHCN source data, and a clear description of the adjustments applied to all input data, it is likely impossible to determine the overall independence between the CRU, GISS and NOAA series.

Testing the adequacy of the corrections and adjustments

[72] Jones made a strong claim about the quality of his surface temperature data in email 1141930111.txt, dated March 9 2006. This was during the final review phase of the 4th IPCC Report on which Jones was a Chapter 3 Coordinating Lead Author, which dealt among other things with the quality of the surface temperature record. Jones was, by then, already in possession of IPCC review comments pointing to published evidence from two independent groups calling into question the quality of the CRU data (see paragraph [73] below). The email was sent to IPCC colleagues Jansen, Overpeck and Briffa in response to news of a forthcoming US government report that would present a partial reconciliation of satellite and surface data series. Though the recipient list is short, it is indicative of the attitude that Jones took towards his data products, namely a categorical dismissal of the possibility of problems.

I can say for certain (100% - not any probable word that IPCC would use) is that the surface temperature data are correct.

[73] I have spent several years devising and implementing statistical models to test the claim that the adjustments to CRU data are adequate. I have argued that an indication of inadequate adjustments would

be a significant correlation between the spatial pattern of warming trends in climate data and the spatial pattern of industrialization/socioeconomic development. McKittrick and Michaels (2004), published in *Climate Research*, showed that such correlations are large and statistically significant, implying that the adjustments are likely inadequate. Our follow-up paper in the *Journal of Geophysical Research* in 2007 re-established these results on a new and larger global data base. Meanwhile a pair of Dutch meteorologists (de Laat and Maurellis) also published peer-reviewed research in 2004 and 2006 showing that temperature trends in gridded climate data sets appear to be correlated with the spatial pattern of industrialization, adjustments notwithstanding. De Laat and Maurellis used different methodologies, and we worked independently. Indeed we knew nothing of each other's work prior to its publication. The uniform conclusion across these four papers—published in *Climate Research*, *Geophysical Research Letters*, *International Journal of Climatology* and *Journal of Geophysical Research*—was that spatial patterns of warming are strongly correlated with spatial patterns of industrialization in ways that strongly imply inadequate adjustments for non-climatic effects, and which likely create an overall warm bias in the global record. Hence peer-reviewed research by two independent teams, working independently of the CRU and the Hadley Centre to test the CRUTEM products had, by 2007, showed ample evidence of problems in the CRU data. Beginning at paragraph [78] below I will explain how this issue was kept out of the 2007 IPCC Report.

Describe as clearly as possible the protocols you have followed in selecting, correcting and rejecting data and stations.

[74] As described above, notwithstanding the 1985 DOE reports, incomplete information has been published on this topic, and attempts by other researchers to ascertain exactly which stations are used in each grid cell have been thwarted. The CRU will not even disclose the stations it currently uses.

Has this been an orderly and objective process applied to all datasets?

[75] Claims to this effect from the CRU should be considered in light of the fact that no independent verification has been possible.

To what extent have different procedures for data of different vintages and different sources been unified?

[76] All that has been disclosed is a list of stations, with no indication of which stations were used at which points in time. So this question is not answerable on the basis of publicly-disclosed information.

What means do you use to test the coherence of the datasets?

[77] It is not clear what is meant by “coherence.” The publication record does not indicate that any tests are applied by CRU to ensure station records are adjusted to remove non-climatic biases. The quotation from paragraph [68] only says the CRU uses “subjective interstation comparisons performed on a local basis” where possible.

Terms of Reference Question 1.2.

Review CRU's policies and practices for acquiring, assembling, subjecting to peer review and disseminating data and research findings, and their compliance or otherwise with best scientific practice.

ISSUES ARISING ON Para 1.2 OF THE TERMS OF REFERENCE

5. It is alleged that there have been improper attempts to influence the peer review system and a violation of IPCC procedures in attempting to prevent the publication of opposing ideas.

It is alleged that there has been an attempt to subvert the peer review process and exclude publication of scientific articles that do not support the Jones-Mann position on global climate change. A paper by Soon & Balunias was published in the Journal *Climate Research* arguing that the 20th century was not abnormally warm. An email from Professor Michael Mann on 11th March 2003 contained the following:

"I think we have to stop considering Climate Research as a legitimate peer-reviewed journal. Perhaps we should encourage our colleagues in the climate research community to no longer submit to, or cite papers in, this journal."

The allegation is that journals might be pressured to reject submitted articles that do not support a particular view of climate change.

In an email to a fellow researcher in June 2003, Briffa wrote: “Confidentially I now need a hard and if required extensive case for rejecting (an unnamed paper) to support Dave Stahle’s and really as soon as you can.”

In an email to Mann on 8th July 2004, Jones wrote:

“The other paper by MM is just garbage. [...] I can’t see either of these papers being in the next IPCC report. Kevin and I will keep them out somehow — even if we have to redefine what the peer-review literature is!”

The allegation is of an attempt to prevent ideas being published and the author being prepared to subvert the peer review process for a journal and to undermine the IPCC principle of accounting properly for contradictory views.

QUESTIONS TO ADDRESS

Give full accounts of the issue in relation to the journal Climate Research, the June 2003 email, and the March 2004 email to Mann (“recently rejected two papers (one for Journal of Geophysical Research & one for Geophysical Research Letters) from people saying CRU has it wrong over Siberia. Went to town over both reviews, hopefully successfully. If either appears I will be very surprised”.

Suppression of Information in the IPCC Report: Surface Data Contamination

[78] The affair over the Soon and Baliunas paper is, in my view, a sad indicator of the intolerant environment in the climatology community, especially since the paper in question was giving evidence of

uncertainties that we now know were privately shared by others in the field. However the main instigators of the campaign to harass and discredit *Climate Research* were not apparently at the CRU so it is not directly relevant to the Inquiry. Instead I will address the email of July 8 2004, since it refers to a paper of which I was a coauthor (McKittrick and Michaels 2004). I note that in a UK Guardian article of February 2, 2010, Trenberth is quoted as strongly disavowing the statement by Jones (<http://www.guardian.co.uk/environment/2010/feb/02/hacked-climate-emails-flaws-peer-review>).

[79] It has been suggested (for example by the UEA Memorandum to the UK Parliamentary Inquiry, paragraph 3.8.2) that the email is mitigated by the fact that the paper in question was actually cited in the IPCC report. This is misleading. As I will show, the citation was kept out of the drafts shown to expert reviewers, and the text that appeared in the published IPCC report relied on invented evidence.

[80] The UEA Memorandum (paragraph 3.8.3) also attempts to mitigate the comment by saying that papers were published by Benestad (2004) and Schmidt (2009) criticizing our methods, thus apparently vindicating Jones' views. This is unconvincing in several respects.

- First, the UEA document did not explain that the Benestad paper was a short comment on McKittrick and Michaels, it was printed by *Climate Research* without being subject to peer review, and in any event the IPCC did not use the Benestad argument to criticise McKittrick and Michael's findings.
- Second, the UEA Memorandum failed to cite the reply of McKittrick and Michaels (2004b) which rebutted Benestad's criticism.
- Third, the UEA Memorandum failed to note that our 2004 results were replicated and reinforced by the findings in our 2007 paper on a new and larger data set, as well as being independently supported by the de Laat and Maurellis findings.

- Fourth, the UEA Memorandum failed to disclose that the Schmidt article was peer-reviewed for the journal by Phil Jones himself (see http://www.climate-gate.org/cru/documents/review_schmidt.doc), so the fact of its publication cannot be offered as independent support for Phil Jones' views.
- Fifth, the UEA Memorandum failed to note that Schmidt's paper was published long after the IPCC Report came out, so its content was irrelevant to the deliberations at the time of the IPCC Report's preparation.
- Sixth, the UEA Memorandum failed to point out that in his capacity as IPCC Lead Author, Jones disputed the findings of McKittrick and Michaels and de Laat and Maurellis on grounds unrelated to Schmidt's comment. As I will explain below, the specific claim made in the IPCC text relied on the apparently fabricated claim that if the effects of atmospheric circulations are taken into account, our results become statistically insignificant. This has since been refuted in a peer-reviewed article (McKittrick 2010, included as Appendix B), a copy of which the UEA Memorandum authors could easily have obtained had they looked into the matter.
- Finally the UEA Memorandum ought to have noted that the Schmidt comment was published by a journal that did not ask for review comments from nor seek a reply from either McKittrick and Michaels or de Laat and Maurellis, so the issues raised therein have not been resolved in the literature, however McKittrick (2010) does rebut the main argument in Schmidt (2009), namely that spatial autocorrelation undermines the conclusions of the McKittrick and Michaels papers (see Appendix B).

[81] The IPCC released the First Order Draft in August 2005. Since this was over a year after Jones' email to Mann it is clear he was aware of my study (it is not clear what is the second paper to which he refers, but it might have been one by de Laat and Maurellis, and I assume that it was). The relevant section of the IPCC Draft was Chapter 3, pages 3-9 to 3-10. Consistent with the intent expressed in the

email there was no mention of either McKittrick and Michaels or the de Laat and Maurellis work. IPCC Expert Reviewer Vincent Gray criticized the omission as follows:

The "corrections" to the surface temperature record have always been based on very poor evidence. The many references to studies on individual or regional stations which find the need for much higher corrections than are currently applied, are ignored. Now you have ignored the persuasive evidence of McKittrick and Michaels 2004 Climate Research 26 156-173 who have shown a significant influence on your "corrected" figures of a series of socioeconomic factors. You cannot just ignore this paper.
[Vincent Gray]

(<http://pds.lib.harvard.edu/pds/view/7795947?n=7&imagesize=1200&jp2Res=.25>) My expert review comments also criticized the omission.

[82] The IPCC Second Order Draft was released in March 2006. Again consistent with the intent revealed in Jones' email to Mann, and despite reviewer demands, there was still no mention of our findings or those of deLaat and Maurellis. I provided lengthy feedback objecting to this omission. In June 2006 the expert review period closed.

[83] The final, published IPCC report in May 2007 included a new paragraph in Chapter 3, on page 244, that had not been included in either of the drafts shown to reviewers. I surmise that Professor Jones, as Coordinating Lead Author for Chapter 3, wrote the paragraph alone or in consultation with Trenberth, and bears responsibility for its inclusion in the published report. It read (emphasis added):

McKittrick and Michaels (2004) and De Laat and Maurellis (2006) attempted to demonstrate that geographical patterns of warming trends over land are strongly correlated with geographical patterns of industrial and socioeconomic development, implying that urbanisation and related land surface changes have caused much of the observed warming. However, the locations of greatest socioeconomic development are

also those that have been most warmed by atmospheric circulation changes (Sections 3.2.2.7 and 3.6.4), which exhibit large-scale coherence. **Hence, the correlation of warming with industrial and socioeconomic development ceases to be statistically significant.** In addition, observed warming has been, and transient greenhouse-induced warming is expected to be, greater over land than over the oceans (Chapter 10), owing to the smaller thermal capacity of the land.

[84] The concept of “statistical insignificance” has a specific numerical interpretation: it implies that an empirical test has been done of a null hypothesis yielding a p value greater than 0.1. The effects reported in McKittrick and Michaels (2004) had p values on the order of 0.002 or 0.2%, indicating statistical significance of the effects. The claim that our results were statistically insignificant is false and was made without any supporting evidence. To my knowledge no study showing such a thing exists, and I have included a forthcoming paper in a peer-reviewed statistics journal (Appendix B) countering the specific claim that accounting for atmospheric circulation effects renders our results insignificant.

[85] No supporting evidence is provided for the highlighted portion of the inserted paragraph, hence it appears to reflect a fabricated conclusion. It was not shown to expert reviewers during the IPCC Report preparation. Moreover, the references to sections 3.2.2.7 and 3.6.4 of the IPCC Report are misleading since neither section presents evidence that warming due to atmospheric circulation changes occurs in the regions of greatest socioeconomic development. Neither section even mentions industrialization, socioeconomic development, urbanization or any related term. The final sentence in the quoted paragraph is irrelevant to the present discussion since the debate only concerns data over land: there is obviously no economic development over the open ocean.

[86] Evidence of quality problems in CRU data had immediate implications for some of the main conclusions in the published version of IPCC Working Group I Report. Global temperature trends are presented in Table 3.2 on page 243 of the IPCC Report. The accompanying text (page 242) states that the CRU data uncertainties “take into account” biases due to urbanization. The Executive Summary to the chapter (page 237) asserts that “Urban heat island effects are real but local, and have not biased the large-scale trends...the very real but local effects are avoided or accounted for in the data sets used.” The influential Summary for Policymakers stated:

“Urban heat island effects are real but local, and have a negligible influence (less than 0.006°C per decade over land and zero over the oceans) on these values.”

[87] The supporting citation was to Section 3.2, which relied on the unsubstantiated material on page 244. IPCC Chapter 9 provides the summary of evidence attributing warming to greenhouse gases. The problem of CRU surface data contamination is set aside as follows (p. 693):

Systematic instrumental errors, such as changes in measurement practices or urbanisation, could be more important, especially earlier in the record (Chapter 3), although these errors are calculated to be relatively small at large spatial scales. Urbanisation effects appear to have negligible effects on continental and hemispheric average temperatures (Chapter 3).

[88] Again, the rationale for ignoring the issue of CRU data quality problems relies on a citation to Chapter 3, which in turn relied upon the apparently unsubstantiated evidence on page 244.

[89] I submit that evidence sufficient to disprove a claim of fabrication would consist of the p value supporting the claim of statistical insignificance made on page 244 of IPCC Working Group I, the peer-reviewed journal article in which it was presented, and the page number where the study is cited in the IPCC Report. **I request that the Inquiry ask Dr. Jones to produce these things.** An inability on his part to do so would, I submit, establish that the insertion of the paragraph quoted above at paragraph [83] amounted to fabrication of evidence, with the effect of concealing problems in the CRU temperature data upon which some of the core conclusions of the IPCC were based.

Are the first two instances evidence of attempts to subvert the peer review process?

[90] The peer-review process can be said to be subverted when information is withheld from reviewers and evidence is fabricated.

In relation to the third, where do you draw the line between rejecting a paper on grounds of bad science etc, and attempting to suppress contrary views?

To what extent is your attitude to reviewing conditioned by the extent that a paper will set back the case for anthropogenic global warming and the political action that may be needed to mitigate it?

What is the justification for an apparent attempt to exclude contrary views from the IPCC process?

[91] CRU staff may attempt to argue that their duties as IPCC Lead Authors required them to weigh conflicting evidence, not merely to report all findings that appear in the literature. However in this case the facts suggest that bias was at work, not impartial scholarship.

- [92] Jones' email to Mann was sent in July 2004, a year prior to the release of the first IPCC draft. It proves that Jones was aware of the paper. His email contains no discussion of the content of the paper, no indication that he had identified any actual flaws in it and indeed no indication that he had even read it. Even though the paper provided statistical evidence of quality problems in the CRU data, Jones did not submit a reply or comment to the journal and has never addressed the evidence in print. His email expresses a derisive attitude and an intention to use his status in the IPCC to suppress discussion of it. The evidence shows that he kept it out of the review drafts and then inserted unsubstantiated rebuttal material without subjecting his own conjectures to the peer review process.
- [93] Jones' email of March 2006 (quoted at paragraph [72]) indicates he held an unrealistic view of the quality of CRU data and that he was unreceptive to criticism of his data.
- [94] Jones' responses to the IPCC review comments was incoherent. His response to the Gray comment quoted at [80] stated

The locations of socioeconomic development happen to have coincided with maximum warming, not for the reasons given by McKittrick and Michaels (2004) but because of the strengthening of the Arctic Oscillation and the greater sensitivity of land than ocean to greenhouse forcing owing to the smaller thermal capacity of land.

<http://pds.lib.harvard.edu/pds/view/7795947?n=7&imagesize=1200&jp2Res=.25>

This *ad hoc* reasoning was unsupported by any evidence. The McKittrick and Michaels paper examined data from all over the Northern and Southern Hemispheres. The idea that the Arctic Oscillation controls warming trends in all those places is an exceedingly implausible invention. The

IPCC did not even attribute *Arctic* warming to the Arctic Oscillation, much less warming throughout Africa and South America. The statement gives the impression that Jones had no credible reason to exclude the McKittrick and Michaels evidence, but he was determined to do so nevertheless.

- [95] Jones' review of the Schmidt paper wholly endorses Schmidt's hypothesis that spatial autocorrelation explains the results in both McKittrick and Michaels and de Laat and Maurellis (http://www.climate-gate.org/cru/documents/review_schmidt.doc), even though this contradicts his own hypothesis that the Arctic Oscillation explains the results. He even emphasizes that "it is all down to the calculation of spatial degrees of freedom." His willingness to abandon his own hypothesis suggests that he did not believe it himself, it was simply inserted without evidence to create an appearance of scientific support for what was in reality a foregone conclusion.

Suppression of Information in the IPCC Report: Long Term Persistence

[96] Additional evidence that suppressing material from the IPCC Report was motivated not by impartial scholarship but by bias is found by looking at the treatment of the topic of Long Term Persistence in Chapter 3 of the IPCC Report. In this case, text was introduced into the Second Order Draft of the report, based on expert reviewer comments on the First Draft and supported by citations to peer-reviewed literature, that expressed caution about the statistical significance of warming trends in climate data. Despite the fact that the text had been agreed-to during the review phase, it was then deleted after the close of scientific review and prior to final publication.

[97] The underlying issues are technical. While it is relatively straightforward to estimate a linear trend through time series data, it is much more difficult to determine if it is statistically significant when the data exhibits a strong form of autocorrelation called "persistence." The statistical literature discusses a related family of concepts called, variously, long memory, long term persistence (LTP),

Persistency/Antipersistency (P/AP), autoregressive integrated moving averages (ARIMA), fractional integration (ARFIMA), and so forth. A large literature exists on each of these topics and many papers have been published applying the estimation methods to climatology data sets. Indeed many of the foundational works come out of the hydrology field, where LTP models were developed by pioneers like Hurst and Mandelbrot to provide more physically-realistic characterizations of long data sets (see Koutsoyiannis 2002).

[98] One of the established results of the LTP literature is that taking it into account tends to reduce the apparent significance of trends in long data sets. This has been shown in temperature analysis as well as in analysis of other data sets (e.g. Cohn and Lins 2005).

[99] The First Draft of the IPCC report Chapter 3 contained no discussion of this topic and also made some strong claims about trend significance based on unpublished calculations done at the CRU. I was one of the reviewers who called attention to the issue and requested insertion of some cautionary text dealing with the LTP issue. Chapter 3 was revised so that the Second Draft now included the following paragraph on page 3-9:

Determining the statistical significance of a trend line in geophysical data is difficult, and many oversimplified techniques will tend to overstate the significance. Zheng and Basher (1999), Cohn and Lins (2005) and others have used time series methods to show that failure to properly treat the pervasive forms of long-term persistence and autocorrelation in trend residuals can make erroneous detection of trends a typical outcome in climatic data analysis.

[100] Similar text was also included in the Chapter 3 Appendix, but was supplemented with a disputatious and incorrect claim that LTP models lacked physical realism. I criticized the addition of that

gloss, but other than that there were no second round review comments opposing the insertion of the new text.

[101] After the close of Expert Review the above paragraph was deleted and does not appear in the published IPCC Report, yet the disputatious text in the Appendix was retained. The sections in question were under the control of Jones and Trenberth, who were Coordinating Lead Authors. It is difficult to see how this exclusion of contradictory evidence regarding the significance of warming trends can be justified. The science in question was in good quality peer-reviewed journal articles, the chapter authors had agreed to its inclusion during the review process and there were no reviewer objections to its inclusion.

6. The scrutiny and re-analysis of data by other scientists is a vital process if hypotheses are to rigorously tested and improved. It is alleged that there has been a failure to make important data available or the procedures used to adjust and analyse that data, thereby subverting a crucial scientific process.

It is alleged that there has been a systematic policy of denying access to data that has been used in publications, referring to an email from Jones to Mann on 2nd February 2005 which contains the following:

“And don't leave stuff lying around on ftp sites - you never know who is trawling them. The two MMs have been after the CRU station data for years. If they ever hear there is a Freedom of Information Act now in the UK, I think I'll delete the file rather than send to anyone. Does your similar act in the US force you to respond to enquiries within 20

days?—our does! The UK works on precedents, so the first request will test it. We also have a data protection act, which I will hide behind”.

QUESTIONS TO ADDRESS

Do you agree that releasing data for others to use and to test hypotheses is an important principle?

If so, do you agree that this principle has been abused?

[102] As explained above, it is important to note that the quoted email was sent on February 2 2005, before Jones had received the data requests in 2005, 2007 and 2009. At no time had I ever requested Jones’ station data, despite his claim. The email reveals that a determination not to release data to those who might question his work pre-dated receiving the requests.

If so, should not data be released for use by those with the intention to undermine your case, or is there a distinction you would wish to make between legitimate and illegitimate use?

[103] When I publish a paper I release all my data and code with it. I do not control who can access my data, and indeed I am quite aware that some of my strongest critics take my data and code and try to find something wrong with it. This led, in one embarrassing case, to the discovery of a programming error that I had to correct. It is not a fun process but it is essential to good scientific practice. I would think this is especially the case for a high-profile statistical product like the CRUTEM data set which is so heavily relied upon by researchers around the world and upon which massive public policy decisions now rest.

If not, do others have reasonable access to the data at all levels and to the description of processing steps, in order to be able to carry out such a re-analysis?

[104] I have shown above that this is not the case.

Can you describe clearly the data-sets and relevant meta-data that have been released; what has not been released and to what extent is it in useable form? Where has it been released?

[105] This kind of disclosure was done in 1985 for a previous edition of the land data set, but that data is out of date and subsequent products, especially CRUTEM3, are largely undocumented.

Where access is limited, or not possible, or not meaningful, for legitimate reasons please explain why?

[106] The CRU's claim that non-disclosure agreements forbid releasing the data needs to be reconciled to the CRU's other claim that there is no need to release its data because it has already been published at the GCHN (see paragraphs [54-60]).

7. The keeping of accurate records of datasets, algorithms and software used in the analysis of climate data.

A key concern expressed by a number of correspondents and commentators has been as to whether datasets, and analyses based thereon, were deleted.

QUESTIONS TO ADDRESS

Were formal “data dictionaries” kept of the data sets acquired by the CRU at various times from other bodies such as the UK Meteorological Office Hadley Centre and its equivalents around the World?

Were comprehensive records kept of the way these various data sets were used, the statistical and other algorithms used in processing them, and the various software programmes and modules used to carry out that processing?

Does a formal library of algorithms and software used by the CRU exist?

What quality control measures were used to test the various algorithms and software modules developed by the CRU?

What techniques did members of the CRU employ to ensure the integrity of the various applications used to process climate data?

What policies are in place to ensure the formal archiving of data sets and resultant analyses for future use and review.

[107] The evidence I have submitted above should be sufficient to answer most of these questions.

Terms of Reference Question 1.3.

Review CRU's compliance or otherwise with the University's policies and practices regarding requests under the Freedom of Information Act ("the FOIA") and the Environmental Information Regulations ("the EIR") for the release of data.

8. Response to Freedom of Information requests.

A number correspondents and commentators assert that requests under the Freedom of Information Act (FOIA) and the Environmental Information Regulations (EIR) were incorrectly denied by the University of East Anglia on advice from the CRU. This is the subject of a separate inquiry by the Data Protection Commissioner, but does fall within the terms of reference of the Review Team.

QUESTIONS TO ADDRESS

What formal processes were in place both centrally and within the CRU to ensure fair and impartial assessment of FOIA requests?

Were there any processes in place centrally to review recommendations from the CRU that information should not be released?

Over the five years to November 2009:

- how many requests were received?*
- how many were rejected, and on what grounds?*

- *how many received full release of information?*
- *how many received partial release of information?*

[108] The evidence I have submitted above should be sufficient to answer most of these questions.

Appendix A: References

- Benestad R.E. (2004) Are temperature trends affected by economic activity? Comment on McKittrick & Michaels. (2004). *Climate Research* 27:171–173
- Briffa, K.R., 2000: Annual variability in the Holocene: interpreting the message of ancient trees. *Quaternary* 43 *Science Reviews*, **19**, 87-105.
- Brohan, P., Kennedy, J., Harris, I., Tett, S.F.B. and Jones, P.D., 2006: Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *J. Geophys. Res.* **111**, D12106, doi:10.1029/2005JD006548.
- Brown PJ, Sundberg R (1987) Confidence and conflict in multivariate calibration. *Journal of the Royal Statistical Society. Ser. B* 49: 46-57.
- Cohn, T.A., and H.F. Lins (2005). Nature's Style: Naturally Trendy. *Geophysical Research Letters* 32, DOI: 10.1029/2005GL024476.
- De Laat, A.T.J., and A.N. Maurellis (2004) Industrial CO₂ emissions as a proxy for anthropogenic influence on lower tropospheric temperature trends. *Geophysical Research Letters* Vol. 31, L05204, doi:10.1029/2003GL019024.
- De Laat, A.T.J., and A.N. Maurellis (2006), Evidence for influence of anthropogenic surface processes on lower tropospheric and surface temperature trends. *International Journal of Climatology*. 26:897—913.
- IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.).. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Jones, P.D. and A. Moberg (2003). “Hemispheric and Large-Scale Surface Air Temperature Variations: An Extensive Revision and an Update to 2001.” *Journal of Climate* 16 203—223.
- Jones, P.D., M. New, D. E. Parker, S. Martin, and I. G. Rigor, (1999) Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics.*, **37**, 173–199.
- Koutsoyiannis, D. (2002). The Hurst Phenomenon and Fractional Gaussian Noise Made Easy. *Hydrological Sciences Journal* 47, 4: 573–95.
- Loehle, CL and McCulloch, JH 2008. Correction to: A 2000-year global temperature reconstruction based on non-tree ring proxies. *Energy and Environment*, v. 19, no. 1, p. 93-100
- Mann ME, et al (2008) Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences* 105: 13252-13257.
- Mann, M.E., Bradley, R.S. and Hughes, M.K., 1998. Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries, *Nature*, 392, 779-787.
- Mann, M.E., Bradley, R.S. and Hughes, M.K., Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations, *Geophysical Research Letters*, 26, 759-762, 1999.
- McIntyre, Stephen and Ross R. McKittrick (2009) “Proxy inconsistency and other problems in millennial paleoclimate reconstructions.” *Proceedings of the National Academy of Sciences* February 2, 2009. 106:E10; doi:10.1073/pnas.0812509106
- McKittrick, Ross R. and Stephen McIntyre (2005) “The M & M Critique of the MBH98 Northern Hemisphere Climate Index: Update and Implications. “ *Energy and Environment* 16(1) pp. 69-100.
- McKittrick, Ross R. “Atmospheric Oscillations do not Explain the Temperature-Industrialization Correlation.” Forthcoming, *Statistics, Politics and Policy*.
- McKittrick, Ross R. and P.J. Michaels (2007) “Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data.” *Journal of Geophysical Research*,

APPENDIX A: References

- 112, D24S09, doi:10.1029/2007JD008465.
- McKittrick, Ross R. and Patrick J. Michaels (2004) “Are Temperature Trends Affected by Economic Activity? Reply to Benestad (2004b)” *Climate Research* 27(2) pp. 175-176.
- McKittrick, Ross R. and Patrick J. Michaels, (April 2004a) “A Test of Corrections for Extraneous Signals in Gridded Surface Temperature Data.” *Climate Research* 26(2) 159-173, April 2004.
- North, G. et al. (National Research Council, NRC) (2006). *Surface Temperature Reconstructions for the Last 2,000 Years*. Washington: National Academies Press.
- Schmidt, M. (2009) Spurious correlations between recent warming and indices of local economic activity. *International Journal of Climatology* DOI:10.1002/joc.1831.
- Wahl, Eugene R. and Caspar M. Ammann, 2006. “Robustness of the Mann, Bradley, Hughes Reconstruction of Surface Temperatures: Examination of Criticisms Based on the Nature and Processing of Proxy Climate Evidence.” *Climatic Change*

Appendix B: Supporting Paper

Atmospheric Circulations do not Explain the Temperature-Industrialization Correlation*

Ross McKittrick
University of Guelph, rmckitri@uoguelph.ca

Abstract:

Gridded land surface temperature data products are used in climatology on the assumption that contaminating effects from urbanization, land-use change and related socioeconomic processes have been identified and filtered out, leaving behind a “pure” record of climatic change. But several studies have shown a correlation between the spatial pattern of warming trends in climatic data products and the spatial pattern of industrialization, indicating that local non-climatic effects may still be present. This, in turn, could bias measurements of the amount of global warming and its attribution to greenhouse gases. The 2007 report of the Intergovernmental Panel on Climate Change (IPCC) set aside those concerns with the claim that the temperature-industrialization correlation becomes statistically insignificant if certain atmospheric circulation patterns, also called oscillations, are taken into account. But this claim has never been tested and the IPCC provided no evidence for its assertion. I estimate two spatial models that simultaneously control for the major atmospheric oscillations and the distribution of socioeconomic activity. The correlations between warming patterns and patterns of socioeconomic development remain large and significant in the presence of controls for atmospheric oscillations, contradicting the IPCC claim. Tests for outlier influence, spatial autocorrelation, endogeneity bias, residual nonlinearity and other problems are discussed.

Key words: global warming, data quality, industrialization, spatial regression

*Acknowledgments: I thank Werner Antweiler and Glen Waddell for assistance in the econometric programming. Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

1 INTRODUCTION

1.1 Overview

This paper examines a claim in the most recent report of the Intergovernmental Panel on Climate Change (IPCC) concerning temperature data quality. At issue is whether trends in climate data sets over land are solely due to global climatic change, or are to some extent measuring local non-climatic effects such as urbanization, land surface modification, instrument changes, etc. The latter factors are supposed to have been filtered out of climate data products. Inadequate filtering would imply a form of data contamination. In its most recent report, the IPCC claimed that residual non-climatic effects are negligible in the data upon which they base their main conclusions. But published evidence has contradicted this on the grounds that, in the climate data relied upon by the IPCC, the spatial pattern of warming over land correlates strongly with the spatial pattern of industrialization and economic growth, a pattern not predicted as a feature of general climatic warming. The IPCC dismissed this as a spurious effect attributable to large-scale atmospheric circulation systems, which, they claim, renders the temperature-industrialization correlations statistically insignificant. But they provided no evidence for this position. The purpose of this paper is to evaluate their assertion in a statistical framework.

1.2 Background

Numerous studies have shown that land-use change, such as urbanization, removal of forest cover and introduction of irrigated agriculture, introduce warming biases into local surface temperature data records (e.g. Jones et al. 2009, Christy et al. 2006, Pielke Sr. et al. 2002, Mahmood et al. 2010, etc., see review in McKittrick and Michaels 2007). Since these local changes are not related to global atmospheric climate changes, they need to be filtered out of climate data sets. But de Laat and Maurellis (2004, 2006, collectively denoted DM) and McKittrick and Michaels and (2004a, 2007, collectively denoted MM) showed that the spatial pattern of temperature trends in gridded surface climate data products is strongly correlated with indicators of industrial and socioeconomic development, which are broadly called anthropogenic surface processes. MM additionally test for and establish significant effects from measures of data inhomogeneity, or variations in measurement quality, on data sets that ostensibly have been adjusted to remove such effects. The aggregate effects of these influences are not small: DM and MM07 each estimate non climate-related effects in post-1980 surface temperature

data amounting to between one-third and one-half of the observed warming trend over the global land surface.

The presence of non-climate-related trends in climatic data sets is a form of contamination that may be overstating atmospheric temperature trends (Klotzbach et al. 2009) and leading to misattribution of temperature changes to greenhouse gas effects. Pielke Sr. et al. (2002) found land surface changes produce regional climatic modifications that are not accounted for in the standard radiative forcing metric, further supporting the possibility of misattribution of spatiotemporal variability in gridded surface data.

Benestad (2004) and Schmidt (2009) both argued that the evidence of data contamination can be dismissed as artifacts of spatial autocorrelation, but neither provided a formal test. The spatial autocorrelation issue is examined in McKittrick and Nierenberg (2009) and will be discussed below. Benestad (2004) and McKittrick and Michaels (2004c) debated the extent to which global-scale patterns should be replicated in subsamples. McKittrick and Michaels (2007) performed random subsampling experiments and found consistently strong replication (see Section 3.5 below). Schmidt (2009) showed that the results for individual coefficients are weaker when using the satellite reanalysis product of Mears et al. (2001) rather than the Spencer and Christy (1990) data. McKittrick and Nierenberg (2009) showed that the joint significance tests on which the main conclusions are based remain significant regardless of which satellite data product is used.

Empirical papers in climatology rely strongly on the assumption that climate data products are free of effects from surface processes and measurement inhomogeneity. To take one example, in a comparison of warming trends in climate models and climatic data, Jun et al. (2008, p. 935), state:

Inhomogeneities in the data arise mainly due to changes in instruments, exposure, station location (elevation, position), ship height, observation time, urbanization effects, and the method used to calculate averages. However, these effects are all well understood and taken into account in the construction of the data set.

Later, after observing discrepancies between model-generated and observed trends (which they denote D_i), when explaining why they do not attribute them to data contamination but rather assume they are all attributable to climate model biases, they state: "...climate scientists have fairly strong confidence in the quality of their observational data compared with the climate model biases. Therefore, we assume that the effect of observational errors to D_i is negligible." (Jun et al. 937)

The same assumption is made in the most recent report of the Intergovernmental Panel on Climate Change (IPCC 2007), and indeed is

fundamental to their interpretation of the surface temperature data. Confining the data contamination question only to urban heat island (UHI) effects, though the underlying issue is in fact broader, the IPCC states (p. 244)

In summary, although some individual sites may be affected, including some small rural locations, the UHI effect is not pervasive, as all global-scale studies indicate it is a very small component of large-scale averages.

The quoted statement is misleading since studies of UHI effects are inherently local, whereas the global-scale studies of DM and MM looked at more general issue of surface processes and data inhomogeneities, and did find large effects.

This paper focuses on the treatment of the contamination problem by the IPCC in its 2007 Fourth Assessment Report, where the issue was raised but dismissed as follows.

McKittrick and Michaels (2004) and De Laat and Maurellis (2006) attempted to demonstrate that geographical patterns of warming trends over land are strongly correlated with geographical patterns of industrial and socioeconomic development, implying that urbanisation and related land surface changes have caused much of the observed warming. However, the locations of greatest socioeconomic development are also those that have been most warmed by atmospheric circulation changes (Sections 3.2.2.7 and 3.6.4), which exhibit large-scale coherence. **Hence, the correlation of warming with industrial and socioeconomic development ceases to be statistically significant.** In addition, observed warming has been, and transient greenhouse-induced warming is expected to be, greater over land than over the oceans (Chapter 10), owing to the smaller thermal capacity of the land.

(IPCC 2007 Chapter 3 page 244, emphasis added). The emphasized sentence makes a specific statistical claim: temperature-industrialization correlations cease to be statistically significant once account is taken of atmospheric circulation effects. Numerically, a result ceases to be significant if its P value rises above 0.05, and loses marginal significant when P goes above 0.1. The IPCC did not cite any published P values, or any published evidence of any kind, in support of their claim, and indeed none exists; nor were any new statistical calculations presented in the IPCC report itself. It is also noteworthy that the claim was not subject to the IPCC's peer review process since the paragraph in question did not appear in

either of the two drafts that were circulated for expert review. It appeared for the first time in the final published version.¹

Since the quality of the land surface temperature data is integral to so many reports and studies on climate change it is important to assess the IPCC's claim of statistical insignificance by means of a proper testing. In this paper I examine whether the results in MM04 and MM07 become insignificant once the effects of atmospheric oscillations on the surface temperature field are entered into the models in a reasonable way. I show that after augmenting the models in McKittrick and Michaels (2004, 2007) with four major atmospheric circulation indexes, the correlations in question remain highly significant and continue to indicate that urbanisation, related land surface changes and data inhomogeneities can account for much of the post-1980 warming over land. In order to establish the robustness of these findings I test the augmented McKittrick and Michaels (2007, hereinafter "MM07") model for spatial autocorrelation, endogeneity bias, error misspecification, outlier effects, and overfitting. No evidence for any of these problems emerges, supporting an overall conclusion that the IPCC conjecture was not only presented without support, but is also untrue, and that evidence pointing to significant contamination of climate data over land should therefore not have been dismissed.

2 EMPIRICAL TESTING OF CIRCULATION PATTERN EFFECTS

2.1 MM 2004 Model

Most of the temperature data used in this paper is grouped into 5 degree-by-5 degree grid cells on the Earth's surface. The unit of observation is a linear trend (degrees C per decade) through monthly "anomalies," or deviations from local averages. Hence the regressions are cross-sectional, and seek to explain the spatial pattern of warming and cooling trends over land.

By including the spatial pattern of socioeconomic variables such as population growth and GDP growth, as well as the spatial pattern of geographical and climatological factors, McKittrick and Michaels (2004a) sought to test the hypothesis that the spatial pattern of warming is independent of the socioeconomic influences that climatologists claim have been identified and removed from the data set. Their first regression equation was:

¹ IPCC drafts and review comments are available at <http://www.ipcc-wg3.de/publications/assessment-reports/ar4/forth-assessment-review-comments>

$$\begin{aligned}
 STREND_i = & \alpha + \gamma_1 PRESS_i + \gamma_2 WATER_i + \gamma_3 COSABLAT_i + \beta_1 POP_i \\
 & + \beta_2 SCALE79_i + \beta_3 COAL80_i + \beta_4 COALGROW_i + \beta_5 INC79_i \\
 & + \beta_6 GDPGROW_i + \theta_1 SOVIET_i + \theta_2 SURFMISS_i + \theta_3 LIT79_i + \varepsilon_i \quad [1]
 \end{aligned}$$

where $STREND_i$ is the 1979-2000 trend in weather station data from the Goddard Institute of Space Studies (GISS, http://data.giss.nasa.gov/gistemp/station_data), $PRESS_i$ is mean air pressure, $COSABLAT_i$ is cosine of absolute latitude, $WATER_i$ is a dummy (0,1) variable representing proximity to an ocean coast or a large body of water, POP_i is the GISS estimate of the local population, $SCALE79_i$ is the product of 1979 local per capita income and local population (i.e. a measure of the total scale of measured economic activity in the station's vicinity income of the station location), $COAL80_i$ is the 1980 total national coal consumption in million short tons, $COALGROW_i$ is the average (compound) annual increase in total coal consumption from 1980 to 1998, $INC79_i$ is 1979 real per capita income, $GDPGROW_i$ is average growth in annual real national Gross Domestic Product (GDP) from 1979 to 2000, $SOVIET_i$ is a dummy for membership in the former Soviet Union, $SURFMISS_i$ is the number of months between 1979:1 and 2000:12 in which the observation is missing and $LIT79_i$ is the 1979 average literacy rate for the country.

The use of weather station data was intended to provide a benchmark for the use of climatic data. Local weather station data are not adjusted for the contaminating influences of nearby socioeconomic activity, so it was expected that the various socioeconomic coefficients would be significant, as indeed they were. Then [1] was re-estimated using a different variable, $GTREND_i$, as the dependent variable. This is the linear trend over 1979-2000 in the Climatic Research Unit (CRU) grid cell temperature anomaly data (Brohan et al. 2006) for the same locations as each station in equation [1], where $i = 1, \dots, 218$ is the location index. The CRU data are used by the IPCC and others on the assumption that the socioeconomic effects have been removed through various ad hoc adjustments described in Brohan et al (2006). Further details on all data sources are in McKittrick and Michaels (2004a). Equation [1] was estimated using Generalized Least Squares applying a correction for heteroskedastic errors. Application of an additional correction for clustering of error terms is discussed below. The working assumption of climate data users implies that the socioeconomic coefficients should vanish when $GTREND_i$ is the dependent variable. But MM found that the while the coefficients had somewhat smaller magnitude they were still highly significant. On this basis they concluded that the corrections for non-climatic biases in the data were inadequate. MM extended their data and testing framework in a 2007 paper, as described the next section.

Variable	MM04s	MM04sc	MM04g	MM04gc
press	0.016 <i>1.49</i>	0.025 <i>2.22</i>	0.011 <i>2.88</i>	0.016 <i>3.45</i>
water	0.098 <i>1.80</i>	0.100 <i>1.85</i>	0.003 <i>0.10</i>	0.009 <i>0.29</i>
cosablat	-0.381 <i>-1.61</i>	-0.555 <i>-2.32</i>	-0.602 <i>-5.95</i>	-0.660 <i>-6.23</i>
pop	0.002 <i>0.56</i>	0.004 <i>1.14</i>	-0.002 <i>-1.07</i>	-0.001 <i>-0.75</i>
scale79	-0.001 <i>-0.31</i>	-0.003 <i>-0.87</i>	0.001 <i>0.47</i>	0.001 <i>0.25</i>
coal80	-0.454 <i>-2.73</i>	-0.507 <i>-2.69</i>	-0.309 <i>-3.98</i>	-0.278 <i>-3.33</i>
coalgrow	-0.005 <i>-1.28</i>	-0.006 <i>-1.32</i>	0.001 <i>0.23</i>	0.001 <i>0.34</i>
inc79	0.040 <i>3.56</i>	0.037 <i>3.05</i>	0.018 <i>3.50</i>	0.014 <i>2.76</i>
gdpgrow	0.085 <i>4.12</i>	0.080 <i>3.83</i>	0.026 <i>2.50</i>	0.019 <i>1.80</i>
soviet	0.489 <i>3.90</i>	0.431 <i>3.13</i>	0.129 <i>1.88</i>	0.080 <i>1.22</i>
surfmiss	0.000 <i>0.19</i>	0.000 <i>0.29</i>	-0.003 <i>-0.48</i>	-0.002 <i>-0.23</i>
lit79	-0.005 <i>-3.18</i>	-0.006 <i>-3.22</i>	-0.003 <i>-3.09</i>	-0.002 <i>-2.77</i>
ao		1.023 <i>2.94</i>		0.343 <i>1.61</i>
nao		-0.946 <i>-2.96</i>		-0.218 <i>-1.12</i>
pdo		0.148 <i>1.04</i>		0.150 <i>1.59</i>
so		-0.125 <i>-1.06</i>		0.021 <i>0.24</i>
Constant	-15.392 <i>-1.43</i>	-24.357 <i>-2.15</i>	-10.536 <i>-2.69</i>	-14.995 <i>-3.28</i>
P(X = 0)	0.0021	0.0050	0.0004	0.0219
P(Circ = 0)		0.0046		0.1061
N	218	218	205	205
r ²	0.26	0.30	0.38	0.41

TABLE 1. Results from McKittrick and Michaels (2004b,c) re-done introducing atmospheric circulation measures AO, NAO PDO and SO. Number in italics is *t*-statistic for coefficient immediately above. **Bold** denotes coefficients that are significant at 5%. First column (MM04s): reproduces original results, dependent variable is station trends, estimator is GLS with heteroskedasticity correction. Second column (MM04sc) introduces circulation indexes as correlation coefficients (variables ao—so). Third column (MM04g) reproduces MM04 results using grid cell trend as dependent variable. Fourth column (MM04gc) introduces circulation indexes as correlation coefficients (variables ao—so). P(X=0) is test that non-climatic effects (*Pop* through *Lit79*) are jointly zero. P(Circ=0) is test that atmospheric circulation effects (AO, NAO PDO, SO) are jointly zero. *N* is sample size, *r*² is coefficient of determination.

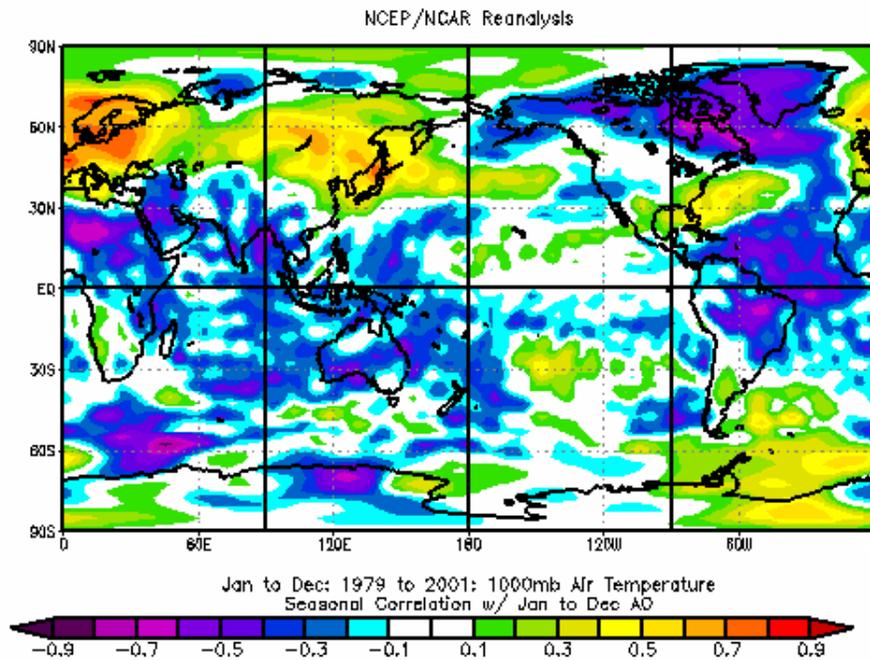
In order to assess the IPCC claim that the MM (2004a) result is spurious due to atmospheric circulation changes, I first obtained a set of correlation fields using the National Oceanic and Atmospheric Administration web site (NOAA, <http://www.cdc.noaa.gov/Correlation>) to generate relevant terms for addition to the regression model. Various atmospheric circulation patterns are discussed in the IPCC Report, chiefly the Arctic Oscillation (AO), North Atlantic Oscillation (NAO), Pacific Decadal Oscillation (PDO) and the Southern Oscillation Index (SO), also known as the El Niño cycle. Each one is an oscillating pattern in air pressure or ocean-atmosphere interactions operating over long (multidecadal) timescales which are known to have strong influences on prevailing weather patterns over land. Indexes measuring the state of each of these oscillations are constructed, typically using pressure gradients across fixed measurement points, or principal components of pressure data over the region of interest.

The index values themselves cannot be used since there is only one monthly value for the whole world. What we are instead interested in is the influence of each oscillation pattern on ground temperatures around the world. The relationship at the gridcell level between surface air temperatures over the 1979-2001 interval and, respectively, the indexes of AO, NAO, PDO and the SO were obtained. These are the most appropriate measures to use for testing the IPCC claim, since they represent the component of temperature changes within a grid cell that are most directly associated with changes in the standard index of the state of the oscillation pattern.

Associations can be measured using a slope from a regression of each gridcell temperature series on the global index, or a Pearson correlation coefficient between the same measures. Both types of coefficient were tried and the one most favorable to the IPCC hypothesis was selected for each model. All results, data and code are available in Supplementary Information. Figure 1 illustrates the correlation field values for the AO.

Equation [1] was re-estimated on both station and gridded data after augmenting with AO_i , NAO_i , PDO_i and SO_i , each of which denotes the correlation in grid cell i between the indicated oscillation pattern and grid cell temperatures. The results are in Table 1. In the station data sample (columns 1—2) the socioeconomic coefficients retain their approximate size and significance levels and an F test shows they remain jointly significant ($P = 0.005$) after the circulation indexes are introduced (note that all joint parameter tests herein use F tests of linear restrictions). The AO and NAO terms are individually significant and the four circulation indexes are jointly significant (joint $P = 0.005$).

With gridded trends as the dependent variable (columns 3—4), there is no support for introducing circulation indexes into the model as they fail to achieve individual or joint significance ($P = 0.106$). If they are included anyway, three of



NOAA/ESRL Physical Sciences Division

Figure 1. Partial correlation between surface air temperature and the AO index 1979-2001. Source: <http://www.cdc.noaa.gov/Correlation>.

the four significant socioeconomic coefficients remain significant and the fourth (GDP growth) falls to marginal significance, while the group remain jointly significant ($P = 0.022$).

Applying a correction for clustered standard errors (Moulton 1990) increases the error variances slightly. Atmospheric circulation indexes remain individually and jointly insignificant (joint $P = 0.124$) and thus their inclusion in the model is not supported. Other inferences remain the same in the station data. In the gridded data, GDP growth falls to insignificance but the economic variables (POP through GDP_{grow}) retain joint significance ($P = 0.035$) and all the socioeconomic variables together remain jointly marginally significant ($P=0.077$).

Note that McKittrick and Michaels (2004) included a suite of model specification tests, including sensitivity to removal of influential outliers, out-of-sample predictive ability, robustness to re-specification of the dependent variable as surface-troposphere trend differences, and insensitivity to stepwise inclusion of independent variables. Since the tests reported herein show that atmospheric circulation indexes are not supported in the model these test results remain valid.

MM07 undertook a wider set of model specification tests, and the effects of including atmospheric circulation indexes will be discussed in Section 3.

2.2 MM 2007 Model

MM07 developed a new and larger data set to re-test their 2004 results. They obtained temperature data for all 469 land-based grid cells with temperature data over 1979:1 to 2002:12 in the CRU ‘crutem2v’ edition data set, and corresponding socioeconomic data, and estimated the regression

$$\begin{aligned} \theta_i = & \beta_0 + \beta_1 TROP_i + \beta_2 PRESS_i + \beta_3 DRY_i + \beta_4 DSLP_i + \beta_5 WATER_i \\ & + \beta_6 ABSLAT_i + \beta_7 p_i + \beta_8 m_i + \beta_9 y_i + \beta_{10} c_i + \beta_{11} e_i + \beta_{12} g_i \\ & + \beta_{13} x_i + u_i \end{aligned} \quad [2]$$

where θ_i is the linear (Ordinary Least Squares) trend through monthly temperature anomalies in 5x5 degree grid cells (data obtained from <http://ipcc-ddc.cru.uea.ac.uk>), $TROP_i$ is the time trend of Microwave Sounding Unit (MSU)-derived temperatures in the lower troposphere in the same grid cell as θ_i over the same time interval (Spencer and Christy 1990), obtained from <http://vortex.nsstc.uah.edu/data/msu/t2lt>, $PRESS_i$ is as above, DRY_i is a dummy variable denoting when a grid cell is characterized by predominantly dry conditions (which is indicated by the mean dewpoint being below 0 °C), $DSLP_i$ is $DRY_i \times PRESS_i$, $WATER_i$ is a dummy variable indicating the grid cell contains a major coastline, $ABSLAT_i$ denotes the absolute latitude of the grid cell, p_i is 1979-1999 population growth, m_i is the 1979 to 1999 percent change in real GDP per capita, y_i is the corresponding percent change in national GDP, c_i is the corresponding growth of national coal consumption, g_i is the GDP density (GDP per square km) as of 1979 and e_i is the average level of educational attainment as late in the interval as possible; each in the country where gridcell i is located; and x_i is the number of missing months in the observed temperature series for gridcell i over the interval 1979—2002. Complete details on data sources and model derivation are in MM07. Educational attainment is included as a measure of the difficulty of recruiting qualified technical staff to operate the meteorological monitoring network.

Equation [2] was estimated on the subsample ($n = 440$) excluding Antarctica and gridcells with too many missing values. They used Generalized Least Squares, controlling for heteroskedasticity and error clustering. Table 2 reports the MM07 results, and also shows the results from augmenting the model with indicators of atmospheric circulation patterns as above. Only the PDO is individually significant, and the four together are jointly significant ($P = 0.0002$).

Variable	MM07	MM07circ-r
trop	0.8631 <i>8.62</i>	0.8279 <i>9.16</i>
slp	0.0044 <i>1.02</i>	0.0060 <i>1.34</i>
dry	0.5704 <i>0.10</i>	-0.1276 <i>-0.02</i>
dslp	-0.0005 <i>-0.09</i>	0.0002 <i>0.04</i>
water	-0.0289 <i>-1.37</i>	-0.0301 <i>-1.48</i>
abslat	0.0006 <i>0.51</i>	0.0008 <i>0.60</i>
g	0.0432 <i>3.36</i>	0.0520 <i>4.17</i>
e	-0.0027 <i>-5.14</i>	-0.0025 <i>-5.57</i>
x	0.0041 <i>1.66</i>	0.0041 <i>1.63</i>
p	0.3839 <i>2.72</i>	0.4376 <i>3.20</i>
m	0.4093 <i>2.39</i>	0.4463 <i>2.80</i>
y	-0.3047 <i>-2.22</i>	-0.3231 <i>-2.54</i>
c	0.0062 <i>3.45</i>	0.0057 <i>3.51</i>
ao		0.1319 <i>1.03</i>
nao		-0.1101 <i>-1.32</i>
pdo		0.1845 <i>2.31</i>
so		0.1510 <i>1.25</i>
constant	-4.2081 <i>-0.96</i>	-5.8078 <i>-1.29</i>
P(X=0)	7.1E-14	1.2E-15
P(Circ=0)		0.0002
N	440	440
r ²	0.53	0.54

TABLE 2.

Results from McKittrick and Michaels (2007) re-done introducing atmospheric circulation measures AO, NAO PDO and SOI. Number in italics is *t*-statistic for coefficient immediately above. **Bold** denotes coefficients that are significant at 5%. First column (MM07): reproduces main original results, dependent variable is grid cell trends, estimator is GLS with heteroskedasticity correction and error clustering. Second column (MM07circ-r) uses same dependent variable and estimator, and introduces atmospheric circulation variables measured using regression coefficients (ao—so). P(X=0) is test that non-climatic effects (g through c) are jointly zero. P(Circ=0) is test that atmospheric circulation effects (AO, NAO, PDO, SO) are jointly zero. *N* is sample size, *r*² is coefficient of determination.

The remaining coefficients are quite robust to their inclusion. Five of the six significant socioeconomic indicators gain size and/or significance, and the joint significance *P* value falls to the 10⁻¹⁵ scale. The estimated average surface

warming trend after removing contaminating effects falls from 0.30 C decade⁻¹ to 0.17 in MM07 and to 0.16 C decade⁻¹ in this case (see MM07 for filtering methodology).

3 FURTHER SPECIFICATION TESTS

The calculations in Section 2 falsify the assertion in IPCC (2007) that the strong correlations between indicators of industrialization and surface temperature trends become statistically insignificant upon controlling for the influence of atmospheric circulation patterns on temperatures. In this section I apply a battery of tests to the results from the augmented MM07 regressions, to check for the overall robustness of the findings and to rule out various other attempts to dismiss the findings as spurious.

3.1 Spatial Autocorrelation

Schmidt (2009) and Benestad (2004) argued that spatial autocorrelation (SAC) of the climate trend field might lead to exaggerated significance in the above regressions. The fact that a dependent variable exhibits SAC is not a problem for hypothesis testing if the model on the right hand side explains it and leaves an uncorrelated residual. Rewrite the regression model [2] in matrix notation as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + u \quad [3]$$

where \mathbf{y} is the linear trend in the temperature series for each of 440 surface grid cells, \mathbf{X} is the matrix of climatic and socioeconomic covariates, $\boldsymbol{\beta}$ is the vector of least-squares slope coefficients and u is the residual vector. SAC in the residual vector can be treated using

$$u = \lambda \mathbf{W}u + e \quad [4]$$

where λ is the autocorrelation coefficient, \mathbf{W} is a symmetric $n \times n$ matrix of weights that measure the influence of each location on the other, and e is a vector of homoskedastic Gaussian disturbances (Pisati 2001). A test of $H_0: \lambda=0$ measures whether the error term in [2] is spatially independent. Anselin et al. (1996) shows that, if the alternative model allows for possible spatial dependence of the \mathbf{y} variables, i.e.

$$\mathbf{y} = \phi \mathbf{Zy} + \mathbf{X}\beta + e \quad [5]$$

where \mathbf{Z} is a matrix of spatial weights for \mathbf{y} and may not be identical to \mathbf{W} , standard adaptations of Wald and Lagrange Multiplier (LM) formulae yield tests that are severely biased towards over-rejection of the null. Anselin et al. (1996) propose a $\chi^2(1)$ Lagrange Multiplier (LM) test of $\lambda=0$ robust to possibly nonzero ϕ in [5], which has substantially superior performance in Monte Carlo evaluations compared to the non-robust LM test.

Hypothesis tests, and any subsequent parameter estimations, are conditional on the assumed form of the spatial weights matrix \mathbf{W} in [4]. I consider three possibilities. Denote the great circle distance between the grid cell centers from which observation i and observation j are drawn as g_{ij} . Weighting matrix 1 (W1) is computed such that each element is $1/g_{ij}$ and the rows are standardized to sum to one. Weighting matrix 2 (W2) is computed such that each element is $1/\sqrt{g_{ij}}$ and the rows are standardized to sum to one. Weighting matrix 3 (W3) is computed such that each element is $1/g_{ij}^2$ and the rows are standardized to sum to one. Matrix W1 assumes the influence of adjacent cells diminishes at a hyperbolic rate. Matrix W2 assumes the inter-cell influence declines more slowly with distance while W3 assumes it declines more rapidly with distance.

Table 3 shows the LM test values (robust and non-robust) for weighting matrices W1—W3 applied to equation [2], with and without the atmospheric circulation terms. For the robust LM statistic (Panel *a*) in none of the three cases is there evidence of significant SAC in the residuals of [2]. The W2 weighting rule, which allows for the slowest decline in the influence of adjacent grid cells as the distance increases, shows the largest test score, though it is still insignificant. These results reverse for the non-robust test (Panel *b*) where W2 yields the smallest test scores. The other two weighting schemes reject the null in the MM07 case, though with the addition of the atmospheric circulation terms the score in W1 becomes insignificant.

While the evidence thus supports not treating for SAC, as a precaution we can nonetheless re-estimate [2] augmented with atmospheric circulation terms and applying [4] as the error term model, to make sure that important conclusions do not hinge on the decision about SAC. The parameter estimates change very little and the filtering method still causes the estimated average trend over land to fall from ~ 0.30 C/decade to ~ 0.18 C/decade (~ 0.27 C/decade to ~ 0.14 C/decade if gridcells are cosine-weighted). The joint socioeconomic effects remain highly significant (see Panel *c*).

<i>a</i>		
Robust LM Score ($\chi^2(1)$, <i>P</i> value)		
Weighting Matrix	MM07	MM07+circulations
W1 (Inverse-linear)	0.032 (0.858)	0.041 (0.840)
W2 (Inverse-square root)	2.564 (0.109)	2.796 (0.095)
W3 (Inverse-squared)	0.094 (0.759)	0.669 (0.413)

<i>b</i>		
Non-Robust LM Score ($\chi^2(1)$, <i>P</i> value)		
Weighting Matrix	MM07	MM07+circulations
W1 (Inverse-linear)	4.240 (0.039)	2.69 (0.101)
W2 (Inverse-square root)	0.030 (0.862)	0.003 (0.954)
W3 (Inverse-squared)	16.032 (0.000)	10.824 (0.001)

<i>c</i>		
Joint Tests in Augmented MM07 Model Applying Controls for Spatial Autocorrelation		
Weighting Matrix	Socioecon ($\chi^2(7)$, <i>P</i> value)	Circulation ($\chi^2(4)$, <i>P</i> value)
W1 (Inverse-linear)	48.05 (0.000)	10.13 (0.038)
W2 (Inverse-square root)	68.88 (0.000)	11.44 (0.022)
W3 (Inverse-squared)	23.64 (0.001)	6.49 (0.166)

Table 3. Hypothesis tests for spatial autocorrelation in model [2] of surface temperature trends and inhomogeneity-anthropogenic surface process biases. **Bold** denotes coefficients that are significant at 5%. Panel *a*: Robust LM test on null hypothesis of no spatial dependence in the model residuals for MM07 model and MM07 augmented with atmospheric circulation variables. Panel *b*: Non-robust LM test on null hypothesis of no spatial dependence in the model residuals for MM07 model and MM07 augmented with atmospheric circulation variables. Panel *c*: Linear restrictions test on null hypothesis of no joint significance of, respectively, socioeconomic and circulation index variables in MM07 augmented with atmospheric circulation variables, controlling for spatial autocorrelation.

Interestingly, if the results from the non-robust LM test under W3 are invoked to recommend adding the controls for SAC, the results in Panel *c* show that while the surface process and inhomogeneity effects remain highly significant, the atmospheric circulation effects become jointly insignificant under that specification. Likewise under W1, the non-robust LM hints at SAC, but not if circulation controls are added (Panel *b*). Hence there is no configuration of tests and models in Table 3 that indicates both spatial autocorrelation and significant atmospheric circulation effects; moreover the conclusions of MM07 concerning the significance of the non-climatic effects are unaffected by the decision to control for spatial autocorrelation or not.

3.2 Influence of Outliers

As in MM07, the results of the model with circulation indexes added were checked to make sure that outliers are not driving the conclusions. Observations were removed if the corresponding diagonal element of the OLS hat matrix exceeded twice the mean of the diagonal elements (Kmenta 1986, pp. 424-426). This resulted in removal of 21 observations, leaving a sample size of 419. The coefficients of the model without outliers were quite similar to those in MM07 Table 2, and the tests of contaminating influences remained highly significant. The Hausman chi-squared statistic was used to test for systematic change in the model parameters, yielding a $\chi^2(18)$ score of 16.39, which is insignificant ($P = 0.565$). Consequently there is no evidence that the conclusions are dependent on outlier observations.

3.3 Overfitting and Collinearity

Overfitting refers to the fact that if there are n observations and k independent variables, as k approaches n the model converges to a perfect fit even if none of the independent variables actually have any explanatory power. The indication that overfitting may be a problem is a significant joint F statistic for all independent variables even though none of them are individually significant. This can also arise if two or more independent variables are highly correlated, or collinear. The results reported herein clearly do not exhibit this problem, since many independent variables are individually significant. MM07 noted that correlations among the model variables were low, and variance inflation factors (VIF) indicated that the explanatory variables were nearly independent of one another. Augmenting the model with atmospheric circulation indexes leaves these results unchanged: the socioeconomic variables exhibit VIFs below 10, indeed most were below 3.

3.4 Regression Error Specification (RESET) and Endogeneity Tests

The RESET score was insignificant in MM07 (Sct. 4.3); likewise with the circulation indexes added it remains so ($P = 0.564$) indicating no evidence of untreated nonlinearity in the model structure. The Hausman endogeneity score also remained insignificant ($P = 0.97$) providing evidence that the model findings are not spurious effects of reverse causality (see discussion in MM07 Sct. 4.4).

3.5 Out-of-sample Prediction Test

A good check against spurious results is the ability of a model to predict the values of a portion of the sample withheld during estimation. MM07 (Sct. 4.5) applied a test as follows. Thirty percent of the data were randomly selected and removed, then the model was re-estimated on the remaining 70%. The model was then used to predict the values of the withheld sample. If the model is perfectly accurate then a regression of the observations on the predicted values would yield a 45 degree line (constant = 0, slope = 1). An *F*-test of these coefficient restrictions is thus an exact test of whether the model systematically fails to predict out-of-sample.

In MM07 this procedure was repeated 500 times. The same number of repetitions were applied after the model was augmented with the atmospheric circulation terms. The mean value of the constant was 0.016, the mean slope coefficient was 0.946, the mean R^2 was 0.498 and the mean *P* value on the test $H_0:(\text{constant} = 0, \text{slope} = 1)$ was 0.373. These scores were almost identical to those in MM07, indicating that the additional terms neither added nor detracted from the model's stability for out-of-sample prediction.

3.6 Tropospheric Pattern

One of the tests in MM07 (Section 4.6) involved an alternative estimation in which the UAH-derived tropospheric trends were removed from the right hand side and used as the dependent variable instead of the surface trends. Had the socioeconomic coefficients retained their size and significance it would suggest that the surface results were spurious, since we would not expect the surface processes to have much effect at the height measured by the satellites (approximately 5 to 15 km aloft). The results in MM07 showed that the surface process variables did indeed lose size and significance in the troposphere, in line with expectations. However, that specification leaves a strong SAC pattern in the residuals. Augmenting the regression with the oscillation variables and applying an SAC correction (using the likelihood-maximizing inverse square weights) yields the expected reduction in magnitude of all the socioeconomic coefficients. The surface process variables become individually and jointly insignificant. Two of the variables that measure surface data quality become much smaller in size but remain (or become) significant, indicating that they are acting as a proxy for some regional effect, since the underlying variables cannot affect the satellite trends. The missing observation count becomes significant, but has a positive value only in 5% of the sample (mainly in the tropics) and is insignificant in the full model anyway. The educational attainment coefficient falls by three-quarters and

changes sign, yet remains significant. In this case the educational attainment variations apparently overlap some aspect of the spatial trend pattern in the tropospheric record. But since the tropospheric trends are included in the full model anyway, this portion of the variance in the education variable is conditioned out in the full model, as indicated by the sign change. Finally, the SAC coefficient in this regression is extremely large (>0.994) indicating that a single spatial lag for this test regression is likely inadequate, and that the variances are likely overstated a bit. Hence the apparent significance levels may simply be due to underestimated variances.

4 CONCLUSIONS

Regional patterns of industrialization, land-use change and variations in the quality of temperature monitoring have been shown by several groups of authors to leave significant imprints on climate data, adding up to a widespread net warming bias that may account for as much as half the post-1980 warming over land. The Fourth Assessment Report of the IPCC dismissed this evidence with the claim that “the correlation of warming with industrial and socioeconomic development ceases to be statistically significant” upon controlling for atmospheric circulation patterns. This claim was presented without any supporting statistical evidence. The models in this paper implement a reasonable way of augmenting the original regressions with the relevant oscillation data, and the results contradict the IPCC claim. The temperature-industrialization correlations in question are quite robust to the inclusion of standard measures of the effects of atmospheric circulation patterns on temperatures, confirming the presence of significant extraneous signals in surface climate data on a scale that may account for about half the observed upward trend over land since 1980.

As discussed in the underlying papers by deLaat and Maurellis and McKittrick and Michaels, socioeconomic activity can lead to purely local atmospheric modifications (such as changes in water vapour and fine particle levels), which, along with other land-surface modifications and data inhomogeneities, can cause apparent trends in temperature data that are not attributable to general climatic changes. As was noted half a century ago by J. Murray Mitchell Jr., referring to the use of temperature observations for measuring climatic trends, “The problem remains one of determining what part of a given temperature trend is climatically real and what part the result of observational difficulties and of artificial modification of the local environment.” (Mitchell Jr., 1953). The results herein show that this concern is still valid, and the conjecture invoked by the IPCC to dismiss it is not supported by the data. A

substantial fraction of the post-1980 trends in gridded climate data over land are likely not “climatically real” but arise from measurement quality problems and local environmental modifications.

REFERENCES

- Anselin, L, A.K. Bera, R. Florax and M.J. Yoon (1996) Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26: 77—104.
- Benestad R.E. (2004) Are temperature trends affected by economic activity? Comment on McKittrick & Michaels. (2004). *Climate Research* 27:171–173
- Brohan, P., J.J. Kennedy, I. Harris, S.F.B. Tett and P.D. Jones (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *Journal of Geophysical Research* **111**, D12106, doi:10.1029/2005JD006548
- Christy, J.R., W.B. Norris, K. Redmond, and K.P. Gallo (2006) Methodology and Results of Calculating Central California Surface Temperature Trends: Evidence of Human-Induced Climate Change? *Journal of Climate* Volume 19, Issue 4 (February 2006) DOI: 10.1175/JCLI3627.1
- De Laat, A.T.J., and A.N. Maurellis (2004) Industrial CO₂ emissions as a proxy for anthropogenic influence on lower tropospheric temperature trends. *Geophysical Research Letters* Vol. 31, L05204, doi:10.1029/2003GL019024.
- De Laat, A.T.J., and A.N. Maurellis (2006), Evidence for influence of anthropogenic surface processes on lower tropospheric and surface temperature trends. *International Journal of Climatology*. 26:897—913.
- IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds).. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Jun, Mikyoung, Reto Knutti and Douglas W. Nychka (2008) Spatial analysis to quantify numerical model bias and dependence: How many climate models are there? *Journal of the American Statistical Association* 108 No. 483 934—947 DOI 10.1198/016214507000001265.
- Jones, P. D., D. H. Lister, and Q. Li (2008) Urbanization effects in large-scale temperature records, with an emphasis on China. *Journal of Geophysical Research* 113, D16122, doi:10.1029/2008JD009916.

- Mahmood, R., R.A. Pielke Sr., K.G. Hubbard, D. Niyogi, G. Bonan, P. Lawrence, B. Baker, R. McNider, C. McAlpine, A. Etter, S. Gameda, B. Qian, A. Carleton, A. Beltran-Przekurat, T. Chase, A.I. Quintanar, J.O. Adegoke, S. Vezhapparambu, G. Conner, S. Asefi, E. Sertel, D.R. Legates, Y. Wu, R. Hale, O.W. Frauenfeld, A. Watts, M. Shepherd, C. Mitra, V.G. Anantharaj, S. Fall, R. Lund, A. Nordfelt, P. Blanken, J. Du, H.-I. Chang, R. Leeper, U.S. Nair, S. Dobler, R. Deo, and J. Syktus, (2010) Impacts of land use land cover change on climate and future research priorities. *Bulletin of the American Meteorological Society* in press
- McKittrick, R.R. and P.J. Michaels (2004a) A test of corrections for extraneous signals in gridded surface temperature data. *Climate Research* 26(2) pp. 159-173
- McKittrick, R.R. and P. J. Michaels (2004b) Erratum, *Climate Research*. 27(3) 265—268.
- McKittrick, R.R. and P. J. Michaels (2004c) Are Temperature Trends Affected by Economic Activity? Reply to Benestad (2004) *Climate Research* 27:175-176.
- McKittrick, R.R. and P.J. Michaels (2007) Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data. *Journal of Geophysical Research*, 112, D24S09, doi: 10.1029/2007JD008465.
- McKittrick, R.R. and N. Nierenberg (2009) Correlations between Surface Temperature Trends and Socioeconomic Activity: Toward a Causal Interpretation. submitted to *International Journal of Climatology*.
- Mitchell Jr., J.M. (1953) On the causes of instrumentally-observed secular temperature trends. *Journal of Meteorology* 10, pp. 244-261.
- Moulton, B.R. (1990) An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics* 72(2) 334-338.
- Pielke R.A. Sr., G. Marland, R.A. Betts, T.N. Chase, J.L. Eastman, J.O. Niles, D.D.S. Niyogi and S.W. Running (2002) The influence of land-use change and landscape dynamics on the climate system: Relevance to climate-change policy beyond the radiative effect of greenhouse gases. *Philosophical Transactions of the Royal Society of London Series A*. A360:1705-1719
- Pisati, M. (2001) Tools for spatial data analysis. *Stata Technical Bulletin* STB-60, March 2001, 21—37.
- Schmidt, M. (2009) Spurious correlations between recent warming and indices of local economic activity. *International Journal of Climatology* DOI:10.1002/joc.1831.
- Spencer, R.W. and J.C. Christy (1990) Precise monitoring of global temperature trends from satellites. *Science* 247:1558—1562.