

Zombie-Wissenschaft: das schlimmste Beispiel der Manipulation von Klimadaten – jemals: der zurückgezogene Beitrag Gergis 2012 wird zu Gergis 2016

geschrieben von Steve McIntyre | 11. August 2016

...

Das Wiedererscheinen des Artikels von Gergis im Journal of Climate wurde begleitet von einer unwahren Darstellung zum Zurückziehen der Version aus dem Jahr 2012. Gergis' Phantastereien und Fehlinterpretationen riefen übermäßiges Lob von Akademikern und anderen Kommentatoren hervor. Gergis zitierte mich persönlich mit der Aussage aus dem Jahr 2012, dass es „fundamentale Mängel“ in dem Artikel gab – eine Behauptung, die sie (fälschlich) als „unrichtig“ bezeichnete und die ihrer Ansicht nach vermutlich „eine Schmierenkampagne koordinieren sollte, um (ihre) Wissenschaft zu diskreditieren“. Ihre folgenden Schwierigkeiten, den Artikel erneut zu veröffentlichen, wofür sie über vier Jahre brauchten, scheint mir eine so eloquente Bestätigung meiner ursprünglichen Diagnose zu sein wie man nur erwarten kann.

Ich habe einige längere Anmerkungen zu Gergis' falschen Statements zu dem Vorfall entworfen, im Besonderen über falsche Behauptungen von Gergis und Karoly, dass die ursprünglichen Autoren den ursprünglichen Fehler unabhängig voneinander entdeckten „zwei Tage“ bevor er bei Climate Audit diagnostiziert worden ist. Diese Behauptungen wurden schon vor vielen Jahren widerlegt in Gestalt von E-Mails, die im Zuge einer Anforderung im Rahmen des FOI herausgegeben werden mussten. Gergis charakterisierte die Anforderungen aus dem FOI als „einen Versuch, Wissenschaftler einzuschüchtern und unsere Bemühungen zu torpedieren, unsere Arbeit zu machen“, aber zu diesen Anforderungen war es nur gekommen infolge der implausiblen Behauptungen von Gergis und Karoly bzgl. Climate Audit.

Obwohl in Gergis et al. 2016 nicht ausdrücklich darauf hingewiesen worden war (um es milde auszudrücken), stellte sich deren Arbeit als **identisch** mit den australasiatischen Rekonstruktionen in PAGES2K (Nature 2013) heraus, während die Rekonstruktionen selbst nahezu identisch sind. PAGES2K wurde im April 2013 veröffentlicht, und man kann nicht anders als sich zu fragen, warum es über drei Jahre gedauert und es neun Überarbeitungen bedurft hatte, um etwas zu veröffentlichen, das sich so ähnlich ist.

Außerdem war es eine der Erwartungen hinsichtlich des PAGES2K-Programms dass es verfügbare Proxy-Daten identifizieren und ausweiten würde, die die vergangenen zwei Jahrtausende abdecken. In dieser Hinsicht sind Gergis und die AUS2K-Arbeitsgruppe kläglich gescheitert. Das Fehlen jedweder Fortschritte von der AUS2K-Arbeitsgruppe ist sowohl erstaunlich als auch erbärmlich, ein Scheitern, dass in Gergis et al. 2016 nicht erwähnt wird, wo doch in Anspruch genommen wurde, „die regionale Konsolidierung australasiatischer Temperatur-Proxys der AUS2K-Arbeitsgruppe zu evaluieren“.

Trendbereinigtes und nicht trendbereinigtes Sichten [screening]

Die folgende Diskussion über die Datenmanipulation bei Gergis et al. 2016 baut auf meiner ähnlichen Kritik an Datenmanipulationen in PAGES2K auf (hier).

In einer Erwiderung auf damals gerade erfolgte Skandale in Sozialpsychologie hat Wagenmakers (2011 pdf, 2012 pdf) die Skandale in den Zusammenhang mit Akademikern gestellt, die ihre Analyse so frisieren, dass ein „erwünschtes Ergebnis“ herauskommt. Er klassifizierte dies als eine Form von „Datenmanipulation“ [data torture]:

*Wir sprechen über eine unbequeme Tatsache, die das Herz der Forschung zu Psychologie an Akademien bedroht: Fast ohne Ausnahme verpflichten sich Psychologen nicht zu einem Verfahren der Datenanalyse, bevor sie die tatsächlichen Daten auf den Tisch bekommen. **Die Versuchung ist dann groß, die Analyse einer Feinabstimmung zu unterziehen, um ein erwünschtes Ergebnis zu erhalten – ein Verfahren, dass die Interpretation der allgemeinen statistischen Tests hinfällig macht. Das Ausmaß der Feinabstimmung variiert erheblich bei Experimenten und Experimentierern, aber es ist praktisch unmöglich für Begutachter und Leser, dies aufzudecken...***

*Einige Forscher erliegen dieser Versuchung leichter als andere, und **aus der präsentierten Arbeit geht in keiner Weise hervor, bis zu welchem Ausmaß die Daten manipuliert worden sind, um die berichtete Bestätigung zu erhalten.***

Wie ich weiter unten zeigen werde, ist es schwierig, ein noch besseres Beispiel von Datenmanipulation zu finden, wie es von Wagenmakers beschrieben worden ist, als Gergis et al. 2016.

Die Kontroverse um Gergis et al. 2012 entzündete sich um das *ex post screening** von Daten, ein bei IPCC-Klimawissenschaftlern sehr populäres Verfahren, aber eines, dass ich seit Jahren scharf kritisiere. Jeff Id und Lucia haben auch etwas Deutliches zu diesem Thema geschrieben (z. B. Lucia hier und im Zusammenhang mit Gergis et al. hier). Ich selbst habe meinen ersten Beitrag zu Gergis et al. 2012 am 31. Mai 2012 geschrieben. Eng damit verbundene statistische Dinge tauchen in anderen Bereichen mit anderer Terminologie auf, z. B. Verzerrungen beim Sammeln von

Stichproben, Nachbehandlung von Variablen, endogener Bias bei der Auswahl. Das Potential von *ex post screening* scheint absurd trivial, falls man das Beispiel eines Drogenprozesses betrachtet. Aber aus irgendwelchen Gründen leugnen die IPCC-Klimawissenschaftler den Bias stumpfsinnig immer weiter. (Eine Schwäche ist, dass statistischer Bias von *ex post screening* nicht automatisch gegenteilige Ergebnisse beweist. Mein Punkt lautet einfach, dass verzerrte Verfahren statistisch nicht durchgeführt werden sollten).

[*Der Terminus ‚*ex post screening*‘ taucht im Folgenden noch öfter auf. Da mir keine vernünftige Übersetzung einfällt, belasse ich es beim Original. Kann jemand einen Vorschlag machen, was gemeint ist? Anm. d. Übers.]

Trotz der öffentlichen Beschränktheit von Klimawissenschaftlern hinsichtlich des Verfahrens hat Karoly kurz nach einer ursprünglichen Kritik an Gergis et al. 2012 *privat* den Bias erkannt, der verbunden ist mit dem *ex post screening*, wie aus einer E-Mail an Neukom vom 7. Juni 2012 hervorgeht (FOI K,58):

Falls die Auswahl der Proxys ohne Trendbereinigung erfolgt, d. h. der gesamten Proxy-Aufzeichnungen über das 20. Jahrhundert, dann wird man Aufzeichnungen mit starken Trends wählen, was effektiv zu einem Hockeyschläger-Ergebnis führt. Dann ist die Kritik von Steve McIntyre valid. Ich denke, dass es wirklich wichtig ist, trendbereinigte Proxydaten für die Auswahl heranzuziehen und dann Proxys auszuwählen, die über eine Stichprobenmenge für Korrelationen während des Kalibrierungs-Zeitraumes hinausgehen, entweder für die jährliche oder dekadische Variabilität der trendbereinigten Daten ... Die Kritik, dass das Auswahlverfahren zu einem Hockeyschläger-Ergebnis führt, wird valid sein, falls der Trend bei der Auswahl der Proxys nicht außen vor gelassen wird.

Gergis et al. 2012 hatten so getan, als ob sie diesen Bias umgangen hätten mit der Betrachtung **trendbereinigter** Daten. Sie haben dieses Verfahren sogar beworben als eines, das „die Aufblähung des Korrelations-Koeffizienten umgeht“:

Bei der Auswahl der Prädiktoren wurden Proxy-Klima-Daten als auch instrumentelle Daten linear trendbereinigt über den Zeitraum von 1921 bis 1990, um das Aufblähen des Korrelations-Koeffizienten zu vermeiden, zu der es kommt infolge der Präsenz des Signals der globalen Erwärmung, der sich in den gemessenen Temperaturaufzeichnungen zeigt. Nur Aufzeichnungen, die signifikant ($p < 0,05$) mit dem trendbereinigten instrumentellen Ziel korrelierten, wurden für die Analyse herangezogen. Dieses Verfahren identifizierte 27 temperatur-sensitive Prädiktoren der warmen Saison von September bis Februar.

Wie inzwischen allgemein bekannt ist, haben sie die behauptete Berechnung tatsächlich nicht durchgeführt. Stattdessen haben sie

Korrelations-Koeffizienten mit nicht trendbereinigten Daten durchgeführt. Dieser Fehler wurde zuerst von dem Kommentator Jean S am 5. Juni 2012 bei Climate Audit (CA) angesprochen (hier). Zwei Stunden später (um 2 Uhr nachts Schweizer Zeit) hat Mitautor Raphi Neukom Gergis und Karoly über diesen Fehler informiert (FOI 2G, Seite 77). Obwohl Karoly später (fälschlich) behauptet hatte, dass seine Mitautoren sich der Bedrohung von Climate Audit nicht bewusst waren, zeigen E-Mails, die im Zuge des FOI herausgegeben werden mussten, dass Gergis an seine Mitautoren eine E-Mail gesandt hatte (FOI 2G, Seite 17), in der er auf die Bedrohung bei CA hinwies, dass Karoly selbst an Myles Allen geschrieben hatte (FOI 2K, Seite 11) über Kommentare, die ihm bei dieser Bedrohung zugeordnet wurden und dass Climate Audit und/oder ich selbst in vielen anderen E-Mails erwähnt werden (FOI 2G).

Wenn Korrelations-Koeffizienten nach der beschriebenen Methode neu berechnet werden, kam nur eine Handvoll davon durch die Sichtung, ein Punkt, der von Jean S am 5. Juni bei Climate Audit angesprochen und von mir in einem Beitrag am 6. Juni beschrieben worden war. Meinen Berechnungen zufolge haben nur 6 der 27 Proxys im G12-Netzwerk die trendbereinigte Sichtung passiert. Am 8. Juni (FOI 2G, Seite 112) schrieb Neukom an Karoly und Gergis, dass acht Proxys die trendbereinigte Sichtung passiert hätten (wobei die Differenz zwischen seinen und meinen Ergebnissen vermutlich Unterschieden bei den Algorithmen geschuldet ist). Er sandte ihnen eine Abbildung (gegenwärtig nicht verfügbar), in der die berichtete Rekonstruktion mit der Rekonstruktion nach dem beschriebenen Verfahren verglichen wird:

Die gestrichelte Rekonstruktion verwendet nur 8 von den Proxys, die die trendbereinigte Sichtung passiert hatten. Die durchgezogene Linie ist die ursprüngliche Rekonstruktion.

Leider war diese Abbildung nicht Teil der FOI-Erwiderung. Sie wäre extrem interessant gewesen.

Da immer mehr Menschen online den Fehler bemerkten, beschloss Leitautor Karoly, dass sie das Journal of Nature informieren müssten. Gergis meldete dem Journal am 8. Juni einen „Fehler beim Daten-Processing“, und der Herausgeber hob die Akzeptanz der Studie am nächsten Tag sofort auf, und zwar mit der Feststellung, dass er verstehe, dass sie die Analyse noch einmal durchführen würden, um mit ihrem beschriebenen Verfahren konform zu gehen:

*Nach Gesprächen mit dem Chefredakteur über unsere Lage lautet meine Entscheidung, die Akzeptanz Ihres Manuskriptes für die Veröffentlichung aufzuheben. So wie ich es verstehe, werden Sie die Analyse erneut durchführen, **um mit ihrer Original-Beschreibung der Auswahl der Prädiktoren konform zu gehen**. In diesem Fälle könnten Sie zu einem anderen Ergebnis kommen als im Original-Manuskript. Angesichts dieser Umstände fordere ich Sie auf, das Manuskript zurückzuziehen.*

Im Gegensatz zu ihrer kürzlich erschienenen Story bei Conservation versuchte Gergis, die erneute Durchführung der Analyse zu vermeiden. Stattdessen versuchte sie den Herausgeber zu überreden, dass der Fehler rein semantischer Natur war („Irrtum der Worte“) anstatt eines Programmierfehlers, und sie suchte nach Unterstützung der nicht trendbereinigten Sichtung von Michael Mann, der Gergis hinter den Kulissen anfeuerte:

Nur zur Klarstellung – es gab einen Irrtum in den Worten, mit denen das Auswahlverfahren der Proxys beschrieben wurde, und keine Fehler in der Gesamtanalyse, wie es von Amateuren klimaskeptischer Blogger behauptet wurde ... Man hat argumentiert, dass trendbereinigte Aufzeichnungen von Proxys, wenn man die Temperatur rekonstruiert, tatsächlich unerwünscht sind.

Die Herausgeber des Journal of Climate ließen sich nicht überreden und forderten Gergis spitz auf, die Differenz zwischen ihrer ersten E-Mail, in der der Fehler als Programmierfehler beschrieben worden war, und ihrer zweiten E-Mail, in der der Fehler als semantisch bezeichnet worden war, aufzuklären:

Ihre jüngste E-Mail an John charakterisiert den Fehler in Ihrem Manuskript als einen der Wortwahl. Aber dies unterscheidet sich von der Charakterisierung in Ihrer ersten E-Mail, in der Sie den Fehler ansprachen. In jener E-Mail (vom 7. Juni) beschrieben Sie es als „einen unglücklichen Fehler beim Daten-Processing“, was Ihre Absicht bekundete, die Daten vom Trend zu bereinigen. Dies würde bedeuten, dass das Problem nicht die Wortwahl war, sondern in der Durchführung des beabsichtigten Verfahrens lag. Würden Sie bitte erklären, warum Ihre beiden E-Mails unterschiedliche Eindrücke von der Natur des Fehlers vermitteln?

Gergis versuchte, von der Frage abzulenken. Sie fuhr fort, das Journal of Climate zu überreden, die Änderung ihrer Beschreibung des Verfahrens zu akzeptieren, war sie doch gegen die Neudurchführung der Analyse nach dem beschriebenen Verfahren. Sie bot lediglich an, die Differenzen in einer Fußnote kurz zu beschreiben:

Die Mail vom 8. Juni war eine rasche Erwiderung, als wir bemerkten, dass es eine Inkonsistenz gab zwischen dem in der Studie beschriebenen Verfahren der Auswahl von Proxys und dem tatsächlichen Verfahren. Die E-Mail wurde überhastet verschickt, da wir Sie so schnell wie möglich informieren wollten, bevor die Studie zum Druck vorbereitet würde. Inzwischen hatten wir jedoch mehr Zeit, uns ausführlich mit Kollegen auszutauschen und die bestehende Literatur zu dem Thema durchzugehen. Es gibt Gründe, warum die Trendbereinigung vor der Proxy-Auswahl ungeeignet sein könnte. Die Unterschiede zwischen den beiden Verfahren werden im Begleitmaterial beschrieben, wie ich in meiner E-Mail vom 14. Juni erklärt habe. Als solche werden die Änderungen im Manuskript vermutlich gering sein, wobei Details der alternativen Proxy-Auswahl im Begleitmaterial umrissen werden.

...

Der Herausgeber des Journal of Climate widersetzte sich dem, räumte aber Gergis widerstrebend ein kurzes Zeitfenster ein (bis Juli 2012), den Artikel zu überarbeiten, verlangte aber, dass sie direkt die Sensitivität der Proxy-Auswahl bzgl. der Rekonstruktion ansprechen und die „Robustheit ihrer Schlussfolgerungen zu demonstrieren“ sollte:

In der Überarbeitung fordere ich strikt, dass das Thema Sensitivität der Klima-Rekonstruktion hinsichtlich der Auswahl des Verfahrens zur Proxy-Aufzeichnungen (trendbereinigt oder nicht) angesprochen wird. So wie ich es verstehe, ist dies genau das, was Sie vorhaben, und dies ist eine gute Gelegenheit, die Robustheit Ihrer Schlussfolgerungen zu belegen.

Unter den Umständen war Chiangs Angebot sehr generös. Gergis griff nach dieser Gelegenheit und versprach, am 27. Juli einen überarbeiteten Artikel zu übermitteln, der den Einfluss dieser Entscheidung auf die sich ergebenden Rekonstruktionen zeigte:

Unser Team wäre sehr erfreut, ein überarbeitetes Manuskript einzureichen am oder vor dem 27. Juli 2012 zur Begutachtung. Wie Sie unten gefordert haben, werden wir ausführlich die Proxy-Auswahl ansprechen zu trendbereinigten oder nicht trendbereinigten Daten sowie den Einfluss auf die resultierenden Rekonstruktionen.

...

Manipulation und Verwässerung der Daten

In der zweiten Hälfte des Jahres 2012 verschrieben sich Gergis und ihre Mitautoren einem bemerkenswerten Programm zur Datenmanipulation mit dem Ziel, ein Netzwerk von 27 Proxys zu retten, während sie immer noch angeblich eine „trendbereinigte“ Sichtung verwendeten. Ihr vermutliches Vorgehen beim *ex post screening* hatte keine Ähnlichkeit mit der grob vereinfachenden Sichtung von Mann und Jones 2003.

Eines ihrer Kernpunkte bei der Datenmanipulation war es, Proxydaten-Korrelationen nicht einfach mit den Temperaturen im gleichen Jahr zu vergleichen, sondern mit Temperaturen im Vorjahr und im folgenden Jahr.

Um Proxys mit jahreszeitlichem Bezug Rechnung zu tragen zu anderen Jahreszeiten als September bis Februar (d. h. Mittelwerte über das Kalenderjahr) wurden die Vergleiche durchgeführt mittels Verschiebungen von -1, 0 und +1 Jahr für jedes einzelne Proxy.

Hauptsächlich waren hier Baumring-Proxys gemeint. In ihrem Verfahren bedeutete eine Verschiebung um -1 Jahr, dass eine Baumring-Serie ein Jahr früher zugewiesen wird als die Chronologie (+1 Jahr wird ein Jahr später zugeordnet). Für eine Reihe mit einer Verschiebung von -1 Jahr soll also die Ringdicke z. B. im Sommer 1989-90 mit den Sommertemperaturen des Vorjahres korrelieren. Es gibt Präzedenzfälle

bzgl. einer Korrelation mit Temperaturen des Vorjahres in speziellen Studien [specialist studies]. Beispielsweise sagen Brookhouse et al. 2008 (Abstract hier), dass die Baumringdaten von *Baw Baw* (ein Gergis-Proxy) positiv mit Frühjahrstemperaturen des Vorjahres korrelieren. In diesem Falle jedoch ordnete Gergis dieser Reihe eine Null-Verschiebung zu, ebenso wie eine negative Orientierung.

Die Verschiebung um +1 Jahr, wie sie 5 Orten zugeordnet worden ist, ist physikalisch nur sehr schwierig zu interpretieren. Eine solche Verschiebung erfordert (zum Beispiel) Mangawhera-Ringdicken, die dem Sommer 1989-90 zugeordnet sind, mit den Temperaturen des Folgesommers (1990-91) korrelieren – was bedeutet, dass die Dicke der Ringe als ein Prädiktor der Temperatur im nächsten Jahr fungiert. Gergis' vermeintliche Rechtfertigung im Text war nichts weiter als Schaumschlägerei, aber den Juroren schien das egal zu sein.

Von den 19 Baumring-Reihen im 51 Reihen umfassenden G16-Netzwerk wurde fünf Reihen eine (unphysikalische) Verschiebung von +1 Jahr, zwei Reihen eine Verschiebung von -1 Jahr und sieben Reihen gar keine Verschiebung zugeordnet. Fünf Reihen wurden aussortiert. Von den sieben Reihen ohne Verschiebung hatten zwei eine inverse Orientierung im PAGES2K. Im Einzelnen gibt es kaum Konsistenz bei Bäumen und Stellen mit der gleichen Spezies. Beispiel: Neuseeland LIBI-Komposit-1 hatte eine +1-Verschiebung, während Neuseeland LIBI-Komposit-2 keine Verschiebung aufwies. Eine andere LIBI-Reihe (Urewara) wird eine inverse Orientierung in der (identischen) PAGES2K-Reihe zugeordnet und folglich vermutlich auch in der CPS-Version von G16. Zwei LIBI-Reihen (Takapari und Flanagan's Hut) werden in G16 ausgesucht, obwohl Takapari in G12 enthalten war. Weil die Zuordnung von Verschiebungen nichts weiter ist als ein Versuch, das Netzwerk zu retten, ist es unmöglich, den Ergebnissen irgendeine Bedeutung zuzumessen.

[Die Übersetzung des vorstehenden Absatzes war für den Übersetzer ein reiner Blindflug. Weil Fehler in einem solchen Fall kaum zu vermeiden sind, folgt er hier im Original:

Of the 19 tree ring series in the 51-series G16 network, an (unphysical) +1 lag was assigned to five series, a -1 lag to two series and a 0 lag to seven series, with five series being screened out. Of the seven series with 0 lag, two had inverse orientation in the PAGES2K. In detail, there is little consistency for trees and sites of the same species. For example, New Zealand LIBI composite-1 had a +1 lag, while New Zealand LIBI composite-2 had 0 lag. Another LIBI series (Urewara) is assigned an inverse orientation in the (identical) PAGES2K and thus presumably in the CPS version of G16. Two LIBI series (Takapari and Flanagan's Hut) are screened out in G16, though Takapari was included in G12. Because the assignment of lags is nothing more than an ad hoc after-the-fact attempt to rescue the network, it is impossible to assign meaning to the results.

Ende Original.]

Zusätzlich hat sich Gergis auch bei einem Daten-Manipulations-Verfahren von Michael Mann bedient. Mann et al. 2008 waren unzufrieden mit der Anzahl der Proxys, die einen Sichtungstest bestanden hatten, der auf der Grundlage einer lokalen Gitterzelle erfolgt war, ein allgemein verwendetes Kriterium (z. B. Mann and Jones 2003). Stattdessen verglich Mann also Ergebnisse mit den zwei „am nächsten gelegenen Gitterzellen“, pickte die höchste der beiden Korrelationen heraus, jedoch ohne den Signifikanz-Test zu modifizieren, um das „Wähle-Zwei“-Verfahren zu reflektieren. (Eine Diskussion dazu gibt es hier). Anstatt lediglich mit den zwei am nächsten gelegenen Gitterzellen zu vergleichen, weitete Gergis den Vergleich auf alle Gitterzellen aus „innerhalb eines Radius‘ von 500 km um den Ort der Proxys“, ein Verfahren mit erlaubten Vergleichen mit 2 bis 6 Gitterzellen abhängig sowohl von der geogr. Breite und der Nähe des Proxys zur Kante der Gitterzelle:

*Wie in Anhang A detailliert ausgeführt, wurden für die weitere Analyse über den Zeitraum 1931 bis 1990 nur Aufzeichnungen ausgewählt, die signifikant mit Temperaturvariationen korrelierten ($p < 0,05$) in **mindestens einer Gitterzelle innerhalb von 500 km um den Ort des Proxys.***

Wie im Artikel beschrieben wurden beide Faktoren in den G16-Vergleichen vermischt. Mit der Multiplikation von drei Verschiebungen mit 2 bis 6 Gitterzellen scheint Gergis 6 bis 18 *trendbereinigte* Vergleiche angestellt zu haben, wobei sie jene Proxys beibehielt, für die eine „statistisch signifikante“ Korrelation bestand. *It doesn't appear that any allowance was made in the benchmark for the multiplicity of tests.** In jedem Falle haben sie es mit diesem „trendbereinigten“ Vergleich fertig gebracht, mit einem Netzwerk von 28 Proxys aufzuwarten, einem mehr als im Netzwerk von Gergis et al. 2012. Die meisten der längeren Proxys sind in beiden Netzwerken die Gleichen mit einer Verschachtelung von etwa sieben kürzeren Proxys. Keine Eisbohrkern-Daten sind im überarbeiteten Netzwerk enthalten und nur eine kurze speläologische Proxy. Es besteht fast ausschließlich aus Baumring- und Korallen-Daten.

[*Ich kapituliere hier. Fachleute wissen sicher, was gemeint ist. Anm. d. Übers.]

Offensichtlich enthielt die ursprüngliche Datenanalyse von Gergis et al. kein verschnörkeltes [baroque] Sichtungs-Verfahren. Es ist klar, dass sie dieses bizarre Sichtungs-Verfahren ausgeheckt haben, um eine Rekonstruktion zu erhalten, die wie die ursprüngliche Rekonstruktion aussah anstatt einer divergenten Version, die sie nicht erwähnt haben. Wer weiß, wie viele Permutationen und Kombinationen und Iterationen getestet worden sind, bevor man zu einem finalen Sichtungs-Verfahren gekommen war.

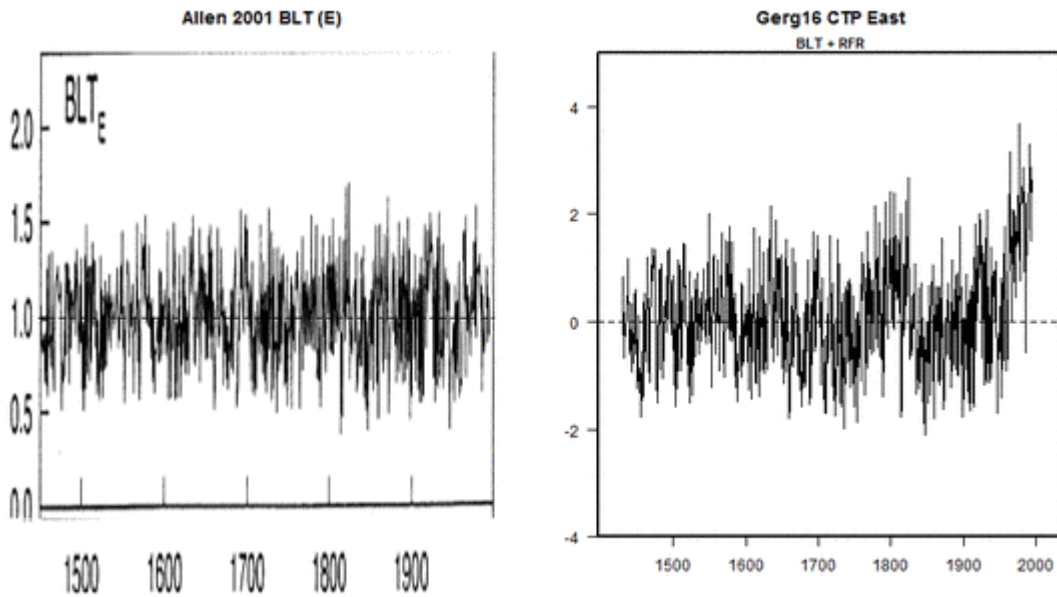
Es ist unmöglich, ein noch klareres Beispiel von „Datenmanipulation“ zu erhalten (selbst Mann et al. 2008).

Auch werden hierdurch nicht alle Elemente der Datenmanipulation in der Studie offensichtlich, da Manipulations-Verfahren aus Gergis et al. 2012 auch in Gergis et al. 2016 zur Anwendung kamen. Mittels originaler und (noch) nicht archivierter Daten hatten Gergis et al. 2012 alle Baumring-Chronologien erneut berechnet mit Ausnahme von zweien. Dabei wendeten sie ein undurchsichtiges, an der University of East Anglia entwickeltes Verfahren an. Die beiden Ausnahmen waren die beiden langen Baumring-Chronologien, die bis ins Mittelalter zurückreichen:

*Alle Baumring-Chronologien wurden auf der Grundlage von Roh-Messungen entwickelt mittels des signalfreien Trendbereinigungs-Verfahrens (Melvin et al. 2007; Melvin and Briffa 2008) ... Die einzige Ausnahme dieses signalfreien Trendbereinigungs-Verfahrens von Baumringen war das Silver Pine-Baumring-Komposit aus Neuseeland (Oroko Swamp und Ahaura), welches nach 1957 durch Abholzen gestört wurde (D'Arrigo et al., 1998; Cook et al., 2002a; Cook et al., 2006) und die Mount Read Huon Pine-Chronologie aus Tasmanien, welches eine komplexe Ansammlung von Material ist, welches von lebenden Bäumen und sub-fossilem Material stammt. Der Konsistenz mit veröffentlichten Ergebnissen halber verwenden wir die finalen Temperatur-Rekonstruktionen, die von den ursprünglichen Autoren zur Verfügung gestellt worden waren und die störungs-korrigierte Daten für die Silver Pine-Aufzeichnung und die Regional Curve-Standardisierung enthalten für die komplexe Altersstruktur des Holzes, das zur Entwicklung der Mount read-Temperatur-Rekonstruktion verwendet wurde.**

[*Wieder ein Blindflug. Ich verweise auf das Original! Anm. d. Übers.]

Damit erhebt sich die offensichtliche Frage, warum „Konsistenz mit veröffentlichten Ergebnissen“ von so großem Belang ist für Mt Read und Oroko, aber nicht für die anderen Reihen, von denen ebenfalls veröffentlichte Ergebnisse vorliegen. Zum Beispiel zeigen Allen et al. 2001 die Chronologie links in Blue Tier, während Gergis et al. 2016 die Chronologie rechts verwendeten als Kombination von Blue Tier und einer benachbarten Stelle. Mittels der Verfahren von East Anglia zeigt die Chronologie eine scharfe Zunahme im 20. Jahrhundert, und „Konsistenz“ mit den Ergebnissen von Allen et al. 2001 machte den Autoren keine Bedenken. Man vermutet, dass Gergis et al. ähnliche Berechnungen für Mount Read und Oroko durchgeführt haben, sich aber entschlossen haben, diese nicht zu verwenden. Man kommt kaum umhin sich zu fragen, ob die aussortierten Berechnungen die erwünschte Story vielleicht nicht gestützt haben.



Dies ist auch nicht die einzige Ad-Hoc-Auswahl, welche diese beiden wichtigen Proxys involviert. Gergis et al. sagten, dass ihr Proxy-Bestand eine Untermenge von 62 Reihen war, entnommen dem Bestand von Neukom und Gergis 2011. (Es war mir nicht möglich, diese Zahl abzugleichen, und Gergis et al. 2016 geben keine solche Liste). Dann sortierten sie Reihen aus, die „zur Zeit der Analyse noch im Entwicklungsstadium waren“ (obwohl sie an einer anderen Stelle sagen, dass der Datensatz ab Juli 2011 eingefroren war wegen der „Komplexität der extensiven mehrdimensionalen Analyse“).

*Von den sich daraus ergebenden 62 Reihen sortieren wir außerdem Reihen aus, die sich zur Zeit der Analyse noch in der Entwicklung befanden ... und Reihen mit **einem in der Literatur identifizierten Mangel** oder durch persönliche Kommunikation.*

Allerdings war die Anwendung dieses Kriteriums inkonsistent. Gergis et al. räumen ein, dass die Oroko-Stelle nach 1957 von einer „Störung durch Holzeinschlag“ betroffen war – ein eindeutiges Beispiel für einen „in der Literatur identifizierten Mangel“. Die Daten wurden aber trotzdem verwendet. In einigen populären Oroko-Versionen (bei Climate Audit gibt es dazu eine Diskussion hier) wurden Proxy-Daten nach 1957 sogar ersetzt durch instrumentelle Daten. Gergis et al. 2016 fügten eine Diskussion dieses Problems hinzu, taten jedoch die Einfügung von instrumentellen Daten in die Proxy-Reihe als belanglos ab:

Man beachte, dass die instrumentellen Daten, die den von Störungen betroffenen Zeitraum seit 1957 in der Baumring-Aufzeichnung von Silver Pine (Oroko) ersetzen, die Sichtung der Proxys beeinflussen haben können, ebenso wie die Kalibrierungs-Verfahren in diesem Zeitraum. Angesichts dessen jedoch, dass unsere Rekonstruktionen zu Beginn des Verifikations-Intervall Qualität [skill] zeigen, also außerhalb des Zeitraumes mit den Störungen, und dass unsere Unsicherheits-Schätzungen das Proxy-

Resampling einschließen, argumentieren wir, dass diese Unregelmäßigkeit in der Silver Pine-Reihe unsere Schlussfolgerungen nicht verzerrt.

Es gibt eine Art von Blindkuh-Spiel in Gergis' Analyse an dieser Stelle. Es sieht für mich so aus, als ob G16 eine Oroko-Version enthalten könnte, die nicht mit instrumentellen Daten versetzt ist. Weil jedoch niemals Messdaten für Oroko archiviert worden waren und eine Schlüsselversion nur als Bestandteil eine Klimagate-E-Mail bekannt geworden ist, ist es schwierig, solche Details zu erkennen.

...

Schlussfolgerungen

Gergis ist viel vertrauensseliges Lob von Akademikern bei Conversation zuteil geworden, aber keiner von ihnen scheint sich die Mühe gemacht zu haben, den Artikel vor der Belobigung zu evaluieren. Anstatt dass die Version 2016 eine Bestätigung oder Verbesserung der Version 2012 ist, zeigt sie ein Beispiel von Datenmanipulation so klar wie man es sich nur wünschen kann. Wir kennen den *ex ante*-Plan zur Datenanalyse von Gergis, weil er in Gergis et al. 2012 beschrieben worden ist. Leider machen sie in ihrem Computerskript einen Fehler und waren nicht in der Lage, ihre Ergebnisse zu wiederholen mittels des Sichtungs-Verfahrens, das in Gergis et al. 2012 beschrieben worden ist.

...

Man fragt sich, ob die Herausgeber und Begutachter des Journal of Climate die extreme Datenmanipulation überhaupt in vollem Umfang verstanden haben, um deren Genehmigung sie aufgefordert waren. Eindeutig scheint es einigen Widerstand von Herausgebern und Begutachtern gegeben zu haben – anderenfalls hätte es nicht neun Runden der Überarbeitung und 21 Begutachtungen gegeben. Da die zahlreichen Begutachtungen das Netzwerk unverändert ließen, ohne auch nur ein Iota vom Netzwerk der Rekonstruktion in PAGES2K abzuweichen (April 2013), kann man nur vermuten, dass Gergis et al. schließlich über ein zögerliches Journal of Climate triumphierten, nach vier Jahren Einreichung und Neu-Einreichung, so dass der Artikel schließlich angenommen wurde.

Wie oben erwähnt, hat Wagenmakers Datenmanipulation definiert als „der Versuchung erliegen, eine Analyse mit den Daten abzustimmen, um ein erwünschtes Ergebnis zu bekommen“. Er diagnostizierte das Phänomen als besonders wahrscheinlich, wenn sich die Autoren nicht selbst „eines Verfahrens der Datenanalyse verpflichtet haben, bevor sie die tatsächlichen Daten gesehen haben“. In diesem Falle hatten Gergis et al. ironischerweise sich selbst einer Methode der Datenanalyse verschrieben nicht privat, sondern im Zusammenhang mit einem angenommenen Artikel, aber offensichtlich haben ihnen die Ergebnisse nicht gefallen.

Es ist verständlich, wie erleichtert Gergis war, nachdem ein derartig manipuliertes Manuskript angenommen worden war. Aber gleichzeitig waren

die Probleme vollständig hausgemacht. Gergis nahm besonderen Anstoß an meiner ursprüngliche Behauptung, dass es „fundamentale Probleme“ bei Gergis et al. 2012 gebe, was sie „unrichtig“ nannte. Aber da ist nichts „Unrichtiges“ an der tatsächlichen Kritik (hier):

Eines der der Analyse im Gergis-Stil zugrunde liegenden Mysterien ist, dass scheinbar äquivalente Proxys „signifikant“ sein können, andere dagegen nicht. Leider werden diese fundamentalen Dinge in der „begutachteten Literatur“ niemals angesprochen.

Dieser Kommentar ist heute noch genauso treffend wie im Jahre 2012.

In ihrem Conversion-Beitrag behauptete Gergis, dass ihr „Team“ die Fehler in Gergis et al. 2012 unabhängig voneinander entdeckt hätten und „zwei Tage“ bevor die Fehler bei Climate Audit beschrieben wurden. Diese Behauptungen sind falsch. Sie haben die Fehler nicht „unabhängig“ von oder vor Climate Audit entdeckt.

Anmerkungen von Anthony Watts:

Es ist wieder und immer wieder die Mann'sche Leugnung von Aufsplittung und Zuordnung, nur um den „Hockeyschläger“ zu erhalten.



Das ist keine Wissenschaft, sondern Verteidigung von „The Cause“.

Der ganze Beitrag steht hier.

Es ist verblüffend, dass Derartiges immer weiter geht, und ich bewundere Steve McIntyre für seine Geduld und seine mikroskopische Detail-Verliebtheit, die er an den Tag legen musste, um durch diese Art der – zweimal – total gescheiterten Begutachtung zu waten. Das Journal of Climate sollte Gergis 2016 zurückziehen.

Link:

<https://wattsupwiththat.com/2016/08/08/zombie-science-the-worst-example-of-climate-data-torturing-ever-withdrawn-gergis-2012-becomes-gergis-2016/>

Übersetzt von Chris Frey EIKE