

Gefälschte Ergebnisse bei Untersuchungen des Proxys „Schichtdicke“

geschrieben von Willis Eschenbach | 22. April 2013

Ihre Behauptung aus dem Abstract lautet folgendermaßen:

Mittels einer hierarchischen Bayesianischen Analyse [?] von Thermometerablesungen, Baumringen, Eisbohrkernen und Sedimenten aus Seen zeigen wir hier, das Magnitude und Frequenz jüngster Warm-Extreme der Temperatur in hohen nördlichen Breiten während der letzten 600 Jahre ohne Beispiel sind. Die Sommer der Jahre 2005, 2007, 2010 und 2011 waren wärmer als alle Sommer zuvor bis zum Jahr 1400 (Wahrscheinlichkeit $P > 0,95$), jedenfalls im Flächenmittel. Der Sommer 2010 war der wärmste Sommer der letzten 600 Jahre in Westrussland ($P > 0,99$) und möglicherweise der wärmste in Westgrönland und der kanadischen Arktis ($P > 0,90$). Diese und andere Extreme aus jüngster Zeit liegen deutlich weit oberhalb der in einem stationären Klima zu erwartenden Extreme, können aber verstanden werden als resultierend aus konstanter Raum-Zeit-Variabilität über eine gestiegene Mitteltemperatur.

[Der letzte Satz lautet im Original: *probably the warmest in western Greenland and the Canadian Arctic as well ($P > 0.90$). These and other recent extremes greatly exceed those expected from a stationary climate, but can be understood as resulting from constant space-time variability about an increased mean temperature.*]

Inzwischen hat Steve McIntyre bei ClimateAudit einige liebenswerte Probleme in ihren Behauptungen gefunden. Ich überlegte mir, einmal einen Blick auf ihre Aufzeichnungen aus Sedimenten in Seen zu werfen. Die Rohdaten vor jedweder Analyse sehen so aus:

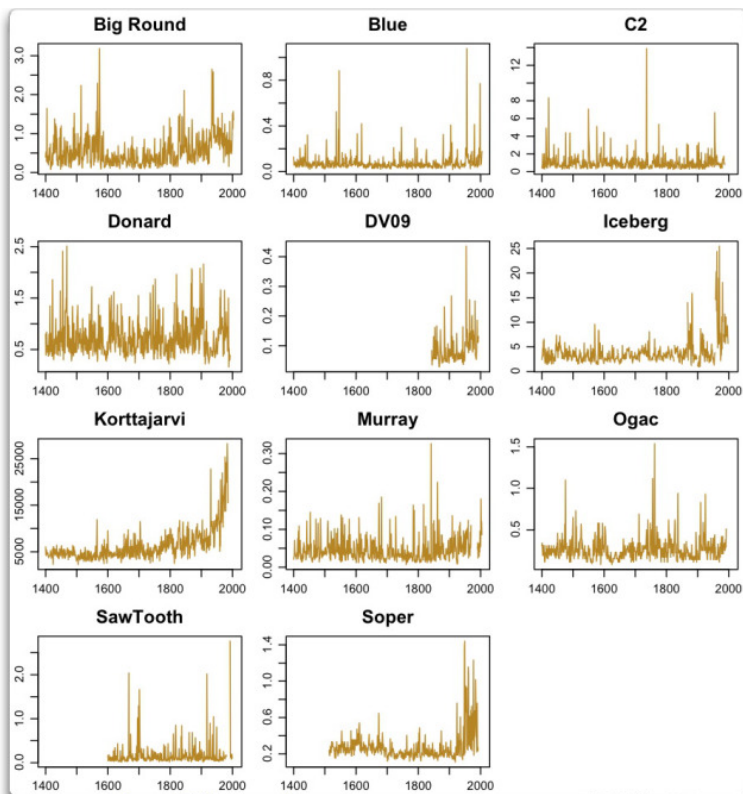


Abbildung 1: alle Schichtdicken-Aufzeichnungen aus TH2013. Die Einheiten wechseln und sind so wie vom Original-Forscher berichtet.

Was also stimmt hier nicht? Nun, eine ganze Menge!

Zu Beginn ist da die infame Korttjarvi-Aufzeichnung. Eine gute Beschreibung hiervon gibt Steve McIntyre:

In Übereinstimmung mit der totalen und vollständigen Sturheit der Paläoklima-Gemeinschaft verwenden sie die berühmteste Reihe von Mann et al. 2008: die kontaminierten Korttjarvi-Sedimente. Die Probleme dieser Reihe sind in Skeptiker-Blogs nur allzu bekannt, und Ross und ich haben darüber in einem Kommentar bei PNAS berichtet. Die Original-Autorin, Mia Tiljander, warnte vor der Verwendung des jüngsten Teils dieser Daten, waren doch die Sedimente durch modernen Brückenbau und die Landwirtschaft kontaminiert worden. Obwohl die Schwächen dieser Reihe als Proxy Lesern „skeptischer“ Blogs gut bekannt sind, haben die Begutachter bei Nature offenbar keine Einwände gegen die Berücksichtigung dieses Proxy für eine Temperaturrekonstruktion gehabt. Hier möchte ich einen Moment innehalten und über Proxies aus Seen sprechen. Auf dem Grund nahezu jeden Sees wird in jedem Jahr eine neue Sedimentschicht abgelegt. Dieses Sediment enthält einen sehr informativen Mix von allem, was jemals während eines bestimmten Jahres in den See gespült worden ist. Man kann Änderungen der lokalen Vegetation beispielsweise durch Pflanzenpollen identifizieren, die als Teil des Sediments abgelagert worden sind. Es gibt eine Menge Informationen, die man aus dem Schlamm am Seegrund gewinnen kann. Eine Informationsquelle, die wir betrachten können, ist die Rate, mit der sich das Sediment akkumuliert. Das nennt man „Schichtdicke“ [varve thickness], wobei eine „Schicht“ [varve] ein Paar dünner Schichten in dem Sediment bedeutet, eine für Sommer und eine für Winter, die zusammen

ein einzelnes Jahres-Sediment ausmachen. Offensichtlich kann diese Dicke ziemlich variieren. Und in einigen Fällen ist es in gewisser Weise mit der Temperatur korreliert.

Allerdings unterscheiden sich See-Proxies in einem entscheidenden Punkt von Proxies aus Eisbohrkernen. Die täglichen Aktivitäten der Menschen ändern nicht die Dicke der Eisschichten, die sich jedes Jahr bilden. Aber alles vom Straßenbau bis zu Änderungen der landwirtschaftlichen Methoden kann die Sediment-Menge in lokalen Gewässern radikal verändern. Und genau das ist das Problem bei Korttajarvi.

Zusätzlich können auch Änderungen der natürlichen Landschaft das Sediment-Niveau verändern. Viele Dinge, vom Verbrennen lokaler Vegetation über Insektenvernichtung bis hin zu Änderungen des lokalen Wasserabflusses können die Sediment-Ablagerungen in einem bestimmten Teil eines bestimmten Sees verändern.

Man betrachte hierzu beispielsweise die Soper-Daten in Abbildung 1. Mehr als offensichtlich werden einige signifikante Änderungen während der ersten Hälfte des 20. Jahrhunderts erkennbar. Nach vier Jahrhunderten eines bestimmten Regimes ist etwas geschehen. Wir wissen nicht was, aber es erscheint zweifelhaft, dass eine graduelle Temperaturänderung eine so plötzliche Änderung der abgelagerten Sedimentmenge in Kombination mit einer Änderung der Variabilität ausgelöst haben kann.

An genau dieser Stelle möchte ich innehalten und betonen, dass allein die Erwähnung dieses Proxys, wenn man die offensichtliche Unsinnigkeit ignoriert, auch Korttajarvi zu betrachten, ausreicht, die ganze Studie total zu disqualifizieren. Es gibt keinerlei Rechtfertigung für die Behauptung, dass diese Änderungen mit der Temperatur zusammenhängen. Ja, ich weiß, später in der Studie wird darauf noch eingegangen, aber bleiben wir realistisch: Dies ist keine Repräsentation der Temperatur. Aber Korttajarvi und Soper sind nicht das einzige Problem. Man betrachte Iceberg, drei separate Aufzeichnungen. Es ist wie eine zweitrangige Quizfrage: „Welcher dieser drei Aufzeichnungen ist anders als die beiden anderen?“ Wie kann man so etwas nur als stichhaltige Proxy betrachten? Wohin führt die Betrachtung all diesen Mülls? Die Autoren erklären es so:

Alle Schichtdicken-Aufzeichnungen, die öffentlich beim Paleolimnology Data Archive der NOAA öffentlich zugänglich sind (Stand Januar 2012) werden betrachtet unter der Voraussetzung, dass sie folgenden Kriterien genügen:

- sie reichen mindestens 200 Jahre zurück,
- sie zeigen eine jährliche Auflösung
- sie werden in Längeneinheiten dargestellt
- die Original-Veröffentlichungen oder andere Referenzen sprechen für eine positive Verbindung mit der Sommertemperatur.

Nun, das klingt alles sehr gut, aber diese Kerle sind so klassisch ... man betrachte den Devon Lake in Abbildung 1, es ist DV09. Sehen Sie, wie weit diese Reihe zurück reicht? Bis 1843, also bis vor 170 Jahren ... so viel zu ihrem 200-Jahres-Kriterium.

Wollen Sie den spaßigen Teil erfahren? Fast hätte ich es übersehen, aber als ich die Kriterien gelesen habe dachte ich „warum gerade 200 Jahre?“

Das schien mir eine Besonderheit zu sein, also betrachtete ich sehr genau die einzige Reihe, auf die es passte, und ... nanu? Das sieht nicht wie 200 Jahre aus. Also untersuchte ich die Daten hier ... 1843, nicht vor 200, sondern vor 170 Jahren.

Mann, je mehr ich schaue, umso mehr finde ich. In dieser Hinsicht haben Sawtooth und Murray wenige kurze separate Abschnitte zum Ende ihrer Hauptdaten. Vielleicht ist es Zufall, aber beide werden den wie auch immer gefälschten Hockeyschläger von Korttajarvi und Soper stützen. Soviel also zum ersten Blick auf die Rohdaten. Jetzt wollen wir betrachten, was sie tatsächlich mit diesen Daten machen. Aus der Studie: *Wie allgemein üblich werden die Schichtdicken vor der Analyse logarithmisch transformiert, was zu Verteilungen führt, die normalverteilter sind und mit der Hypothese in Einklang stehen, die unsere Analyse charakterisiert (siehe folgenden Abschnitt).*

Ich kann mich mit dieser Transformation nicht ganz abfinden. Ich verstehe die zugrunde liegende Rechtfertigung oder Logik nicht, das zu tun. Falls die Schichtdicke in irgendeiner Weise proportional zur Temperatur ist, und das kann gut sein, warum sollte sie proportional zum Logarithmus der Dicke sein?

Wie auch immer, schauen wir, über wie viel „normalverteilter“ wir reden. Hier folgen die Verteilungen der gleichen Aufzeichnungen nach der Log-Transformation [?] und Standardisierung. Ich habe einen „Violin-Plot“ verwendet, um das Aussehen der Verteilung zu untersuchen. Die Entfernung an irgendeinem Punkt steht für die geglättete Zahl der Datenpunkte um diesen Wert. Der weiße Punkt zeigt den Medianwert der Daten. Der schwarze Kasten zeigt die interquartile Bandbreite, die die Hälfte der Daten enthält. Die vertikalen „Schnurrhaare“ [whiskers] erstrecken sich 1,5 mal so weit wie die interquartile Distanz an der Ober- bzw. der Unterseite des Kastens.

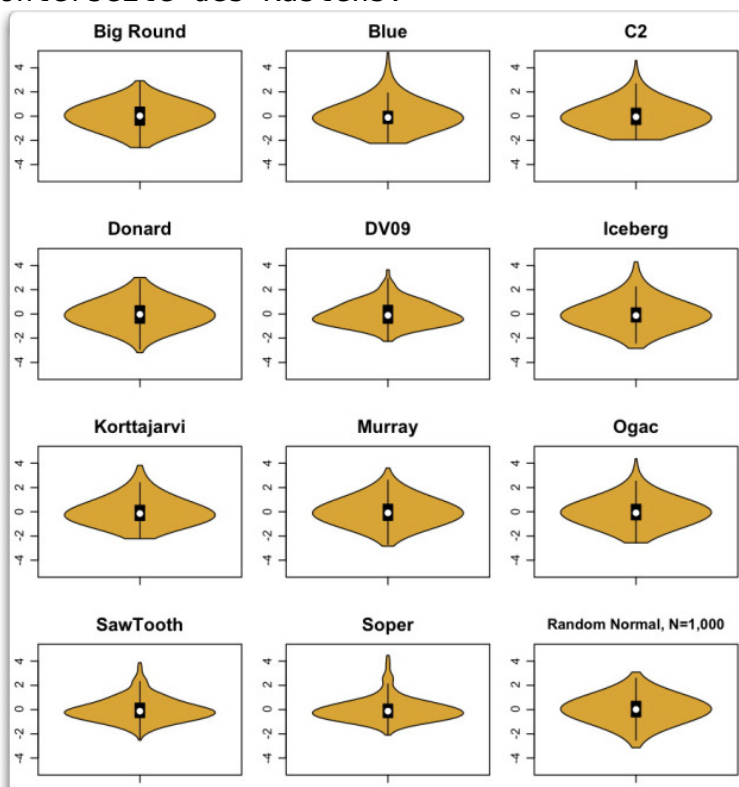


Abbildung 2: Violin-Plots der Daten aus Abbildung 1, aber nach der Log-Transformation und Standardisierung. Die Randwert-Normalverteilung unten rechts dient Vergleichszwecken.

Man beachte die sehr großen Unterschiede zwischen den verschiedenen Datensätzen der Schichtdicke. Man kann die Probleme beim Soper-Datensatz erkennen. Einige Datensätze sind nach der Log-Transformation ziemlich normalverteilt, wie Big Round und Donard. Andere wie DV09 und Soper sind selbst nach der Transformation noch weit von der Normalverteilung entfernt. Viele von ihnen sind stark asymmetrisch mit Ausreißern von vier Standardabweichungen, die in positiver Richtung allgemein sind. Im Gegensatz dazu variieren sie nur halb so stark in negativer Richtung, zwei Standardabweichungen. Und wenn man sie mit einschließt, beeinflussen die Unterschiede, mit denen sie aus dem Normalzustand schwingen (die positiven Ausreißer übertreffen die negativen bei Weitem), sowohl die Ergebnisse als auch deren Unsicherheiten. In jedem Falle sind die Datensätze nach der Log-Transformation und Standardisierung zu einem Mittel von Null und einer Standardabweichung von eins in Abbildung 3 gezeigt:

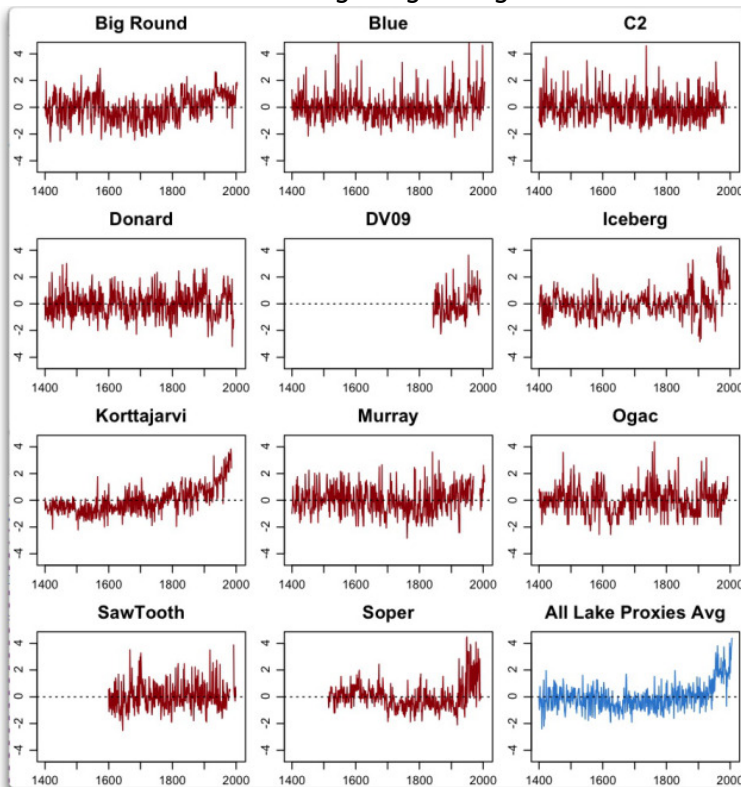


Abbildung 3: Schichtdicken-Aufzeichnungen nach Log-Transformation und Standardisierung.

Wie man sieht, ändert die Log-Transformation nichts an den Problemen von z. B. Soper oder Iceberg. Sie sind immer noch nicht intern konsistent. Als Ergebnis der Mitbetrachtung dieser problematischen Reihen, die allesamt sichtbare Unregelmäßigkeiten in den jüngsten Daten enthalten, zeigt selbst ein einfacher Mittelwert einen vollkommen gefälschten Hockeyschläger.

Tatsächlich zeigt das Mittel die typische Form für diese Art gefälschten Hockeyschläger. Im „Schaft“ des Hockeyschlägers scheinen sich die Rand-Variationen der gewählten Proxies einander wegzumitteln. Im „Blatt“

mitteln sich die Rand-Proxies ebenfalls heraus, und alles was übrig bleibt, sind die wenigen Proxies, die im jüngsten Abschnitt einen Anstieg zeigen.

Meine Schlussfolgerungen lauten (in beliebiger Reihenfolge):

- Man muss den Autoren hinsichtlich der Eindeutigkeit ihrer Datenquellen gratulieren. Sie lassen eine einfache Analyse ihrer Arbeit zu.
- Man muss ihnen auch zu den klaren Definitionen der Kriterien gratulieren, die sie an die Berücksichtigung der Proxies anlegen.
- Traurigerweise folgen sie nicht ihren eigenen Kriterien.
- Die wesentliche Schlussfolgerung jedoch lautet: Klare, eindeutige Kriterien von der Art, wie sie sie verwendet haben, sind ein notwendiger, aber kein ausreichender Teil des Prozesses. Es müssen noch mehr Schritte unternommen werden!

Der zweite Schritt besteht in der Verwendung von Quellen-Dokumentationen und der Literatur um zu sehen, ob es Probleme mit der Verwendung einiger Teile der Daten gibt. Hier Korttajarvi mit einzubeziehen ist ein besonders ungeheuerliches Übersehen. Michael Mann hat es in seiner Analyse aus dem Jahr 2008 umgekehrt verwendet. Er hat anschließend gesagt, dass „das keine Rolle spielt“. Es erscheint hier wieder umgekehrt, und die ursprünglichen Forscher sagten, man verwende es nicht nach 1750 oder so. Es ist absolut erbärmlich, dass nach all den Diskussionen in der Literatur und im Internet einschließlich eines veröffentlichten Briefes an PNAS Korttajarvi wieder einmal in einer Proxy-Rekonstruktion verwendet wird, und wieder einmal wird es umgekehrt verwendet. Das ist unentschuldig.

Der nächste Schritt bei der Auswahl der Proxies muss in der Verwendung des Mark I eyeball [?] bestehen um zu sehen, ob es Lücken, Sprünge der Amplitude, Änderungen der Variabilität oder andere Anzeichen von Problemen in den Daten gibt.

Als nächstes müssen die Auswirkungen der fragwürdigen Daten auf das Endergebnis untersucht werden.

Und schließlich müssen noch die Gründe dargelegt werden für den Ein- oder Ausschluss der fragwürdigen Daten sowie deren Auswirkung auf das Ergebnis der Studie.

Unglücklicherweise haben sie nur dem ersten Teil Genüge getan, indem sie klare Kriterien formuliert haben.

Sehen Sie, man kann nicht einfach ein Bündel Proxies nehmen und diese mitteln, egal ob man Bayesianische Methoden anwendet oder nicht. Die Menge [crowd] der Paläo-Proxy-Forscher hat wieder und immer wieder gezeigt, dass man damit einen Hockeyschläger konstruieren kann, wenn man nur die richtigen Proxies dafür auswählt...

Und jetzt? Alles, was es beweist, ist: ja, wirklich; wenn man Müll hineinsteckt, wird man vermutlich auch Müll herausbekommen. Wenn man sorgfältig den Auswahlprozess der Proxies durchführt, kann man jedes gewünschte Ergebnis bekommen.

Mann, wie sehr ich es satt habe, mich durch diese Art von Müll zu wühlen; künstliche Studien von künstlichen Wissenschaftlern.

Willis Eschenbach

Link: <http://wattsupwiththat.com/2013/04/13/spurious-varvology/>

Übersetzt von Chris Frey EIKE