

LA VICENDA DELLA "HOCKEY STICK" SPIEGATA

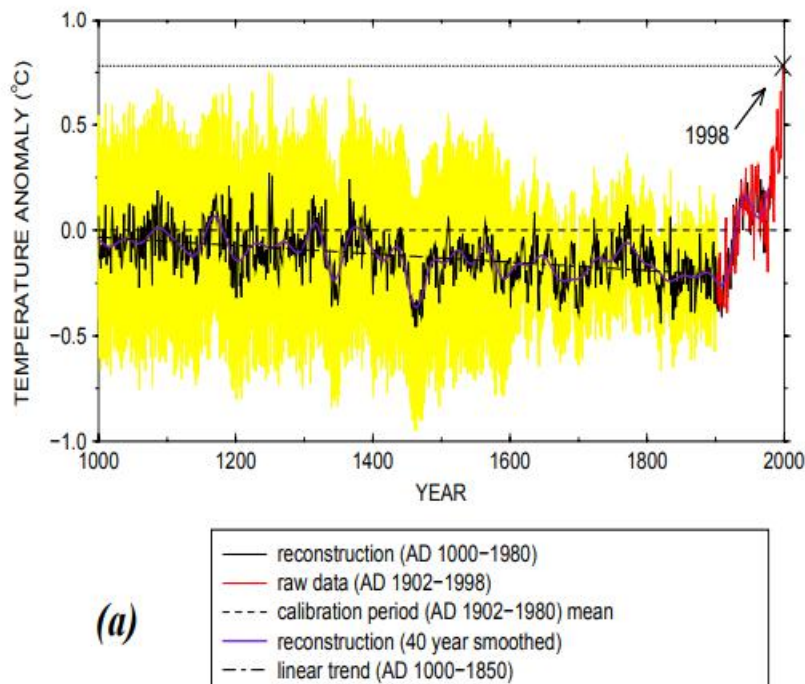
I grafici delle *Hockey Stick* (*mazze da hockey*) sono grafici che dichiarano di presentare il record della temperatura media globale o emisferica degli ultimi 500-2000 anni come mostrato da ricostruzioni climatiche quantitative basate su registrazioni di proxy climatici, prevalentemente carotaggi di ghiaccio e anelli d'albero.

Queste ricostruzioni hanno costantemente mostrato una lenta tendenza al raffreddamento a lungo termine, che si è trasformata in un riscaldamento relativamente rapido nel XX secolo, con il record strumentale della temperatura entro il 2000 che ha superato le temperature precedenti. Il termine "Hockey Stick Graph" è stato reso popolare dal climatologo Jerry Mahlman per descrivere il modello mostrato dalla ricostruzione di Mann, Bradley & Hughes del 1999 (MBH99), che mostra un grafico relativamente piatto con una tendenza al ribasso fino al 1900 - come a formare il "manico" di una mazza da hockey - seguito da un aumento netto e costante corrispondente alla "lama" della mazza stessa. Le ricostruzioni sono state presentate nei rapporti del Gruppo Intergovernativo di esperti sui Cambiamenti Climatici (IPCC) come prova del riscaldamento globale.

In questo documento vogliamo analizzare la questione dell' Hockey Stick usando due approcci cronologici: nella prima parte analizzeremo la vicenda dal lato di Michael E. Mann e dei personaggi principalmente coinvolti nella diatriba, nella seconda parte esamineremo invece la stessa vicenda dal lato di Keith Briffa, involontariamente fonte principale per la ricostruzione presente nel grafico di Mann et al. Le analisi tratteranno sotto forma di elenco puntato i punti salienti delle vicende. Dopo le due analisi cronologiche, offriremo ulteriori informazioni ed espanderemo il campo di analisi per chiarire alcuni punti chiave.

APPROCCIO 1: LA VICENDA DAL LATO DI MICHAEL E. MANN

1- Nel 1998 viene pubblicato un articolo del Dr. Michael E. Mann, all'epoca ricercatore all'Università della Virginia, ora climatologo della Penn State University (Università di stato della Pennsylvania), con co-autori Raymond S. Bradley e Malcolm K. Hughes. Il documento si intitolava "*Global-scale temperature patterns and climate forcing over the past six centuries*"; tale documento diventa subito noto con il nome di MBH98 [0] e verrà seguito l'anno successivo da un altro documento, intitolato "*Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations*" divenuto a sua volta noto come MBH99 [1]. Quest'ultimo presentava un grafico divenuto molto importante e basato sulla ricostruzione climatica in base ad alcuni dataset di anelli d'albero, carotaggi di ghiaccio, e coralli. La conclusione di tale studio sul clima negli ultimi 1000 anni fu che ora siamo nel periodo più caldo di sempre della storia moderna. Il grafico è riportato qui di seguito per praticità.



Stephen McIntyre, un matematico canadese, sospettò che gli anelli d'albero non stessero raffigurando la realtà con quel gigantesco ed improvviso aumento sul lato destro del grafico che descriverebbe il XX secolo come il "*periodo più caldo degli ultimi 1000 anni*" e che gli allarmisti climatici usarono ed usano spesso come prova di un Riscaldamento Globale Antropico. Il grafico viene soprannominato "Hockey Stick" e diventa famoso in tutto il mondo. Al Gore lo usò nel suo film *An Inconvenient Truth*.

2- McIntyre tentò di replicare il lavoro sugli anelli d'albero fatto da Mann nei documenti MBH98/99, ma fu ostacolato dalla mancanza di archiviazione dei dati e dei protocolli di calcolo. Inviò decine di lettere nel corso degli anni cercando di ottenere l'accesso ai dati, ma l'accesso venne sempre negato. McIntyre e Ross McKittrick, dell'Università di Guelph, pubblicarono nel 2003 [2] e nel 2004 [3] due articoli di critica al lavoro di Mann e soci. Per risposta, sempre nel 2004, venne creato un nuovo sito web (chiamato Real Climate), gestito da persone che raccoglievano i dati sugli anelli d'albero, e che denunciavano le critiche di McIntyre e Kitrick definendole 'false accuse' ed attaccando i due studiosi sul lato personale [4].

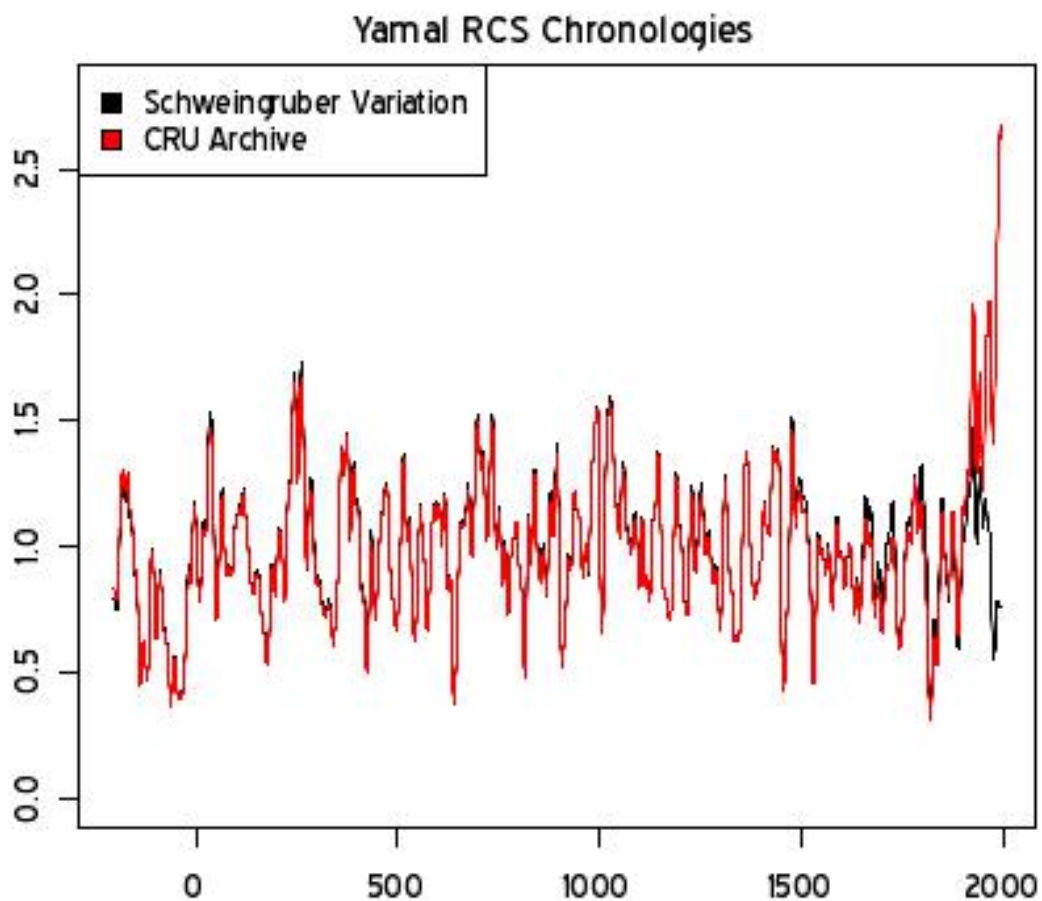
3- Passano 3 anni. McIntyre è ancora ostacolato nel tentativo di ottenere l'accesso ai dati originali in modo da poter replicare la conclusione di Mann del 1998 e del 1999. Nel 2008, Mann pubblica un altro articolo per rafforzare la sua affermazione sugli anelli d'albero a causa di tutte le controversie che lo circondano. Un coautore di Mann e fonte di dati sugli anelli d'albero (il professor Keith Briffa della Hadley UK Climate Research Unit) utilizzò una delle serie di dati sugli anelli d'albero (Yamal in Russia) in un articolo pubblicato su

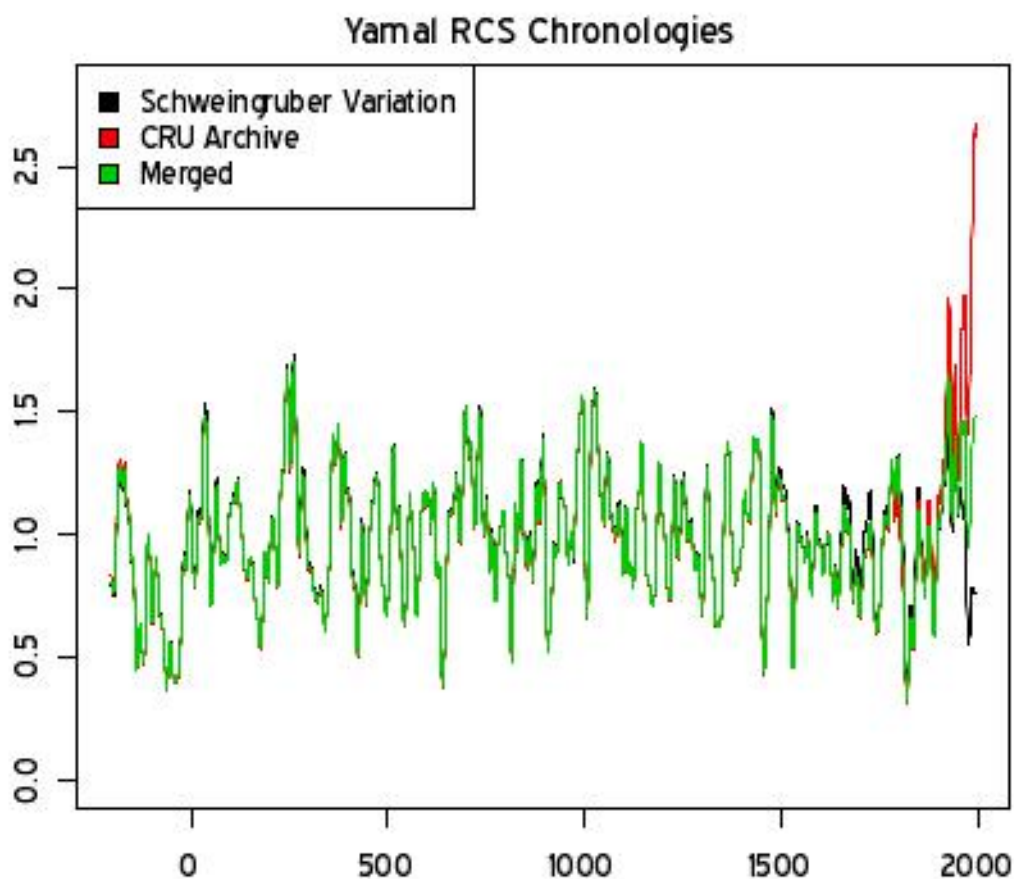
Philosophical Transactions of the Royal Society nel 2008, che ha una rigorosa politica di archiviazione dei dati. Grazie a quella politica, Steve McIntyre riuscì infine ad ottenere l'accesso a quei dati.

4- Avendo i dati di Yamal in forma completa, McIntyre cerca di replicare la procedura descritta da Mann et al - nonostante non avesse a disposizione la procedura di ricostruzione usata da Mann - e scopre che uno dei coautori di Mann, Briffa, aveva selezionato 10 set di dati di alberi da un insieme molto più ampio di alberi campionati a Yamal.

5- Quando vengono tracciati tutti i dati degli anelli d'albero di Yamal, la famosa mazza da hockey scompare. Non solo scompare, ma diventa negativa! La conclusione è secondo McIntyre inevitabile: i dati degli anelli d' albero erano stati scelti accuratamente ottenere il risultato desiderato.

Questi di seguito sono i grafici pertinenti di McIntyre che mostrano ciò che dimostrano i nuovi dati disponibili.





APPROCCIO 2: LA VICENDA DAL LATO DI KEITH BRIFFA

Ora scenderemo molto più in dettaglio, esaminando la vicenda ed estendendo ciò che successe, ma affrontando il discorso osservando l'operato di Keith Briffa durante il corso degli anni.

1. In un articolo pubblicato su Nature nel 1995 ("*Unusual twentieth-century summer warmth in a 1,000-year temperature record from Siberia*", di Briffa, Schweingruber et al.) [5], gli studiosi riportarono che il 1032 fu l'anno più freddo del millennio, proprio nel bel mezzo del periodo caldo medievale (MWP). Ma la ricostruzione dipendeva da 3 brevi nuclei di anelli d'albero degli Urali polari la cui datazione era molto problematica.

2. Negli anni '90, Schweingruber ottenne nuovi dati sugli Urali polari con nuclei datati in modo più sicuro per il MWP. Né Briffa né Schweingruber pubblicarono alcuna nuova cronologia degli Urali polari utilizzando questi dati. Una cronologia aggiornata avrebbe prodotto un quadro molto diverso, vale a dire un'era medievale calda e nessuna anomalia del XX secolo. Invece di utilizzare la serie Polar Urals aggiornata, Briffa calcolò una nuova cronologia dai dati degli alberi della penisola di Yamal, ottenendo una serie che mostrava la forma di mazza da hockey. Dopo la sua pubblicazione, praticamente in ogni studio condotto

dai membri del team legato alla Hockey Stick abbandonò del tutto i dati degli Urali e utilizzò la serie Yamal di Briffa al loro posto. L'eccezione a questo modello fu uno studio di Esper et al su Science (2002), che utilizzò i dati combinati degli Urali polari ma si rifiutò di fornire i suoi dati.

3. Successivamente, sono comparsi innumerevoli studi prodotti dal team che non solo utilizzavano i dati Yamal al posto degli Urali polari, ma in cui Yamal aveva un impatto critico sulla classifica relativa del XX secolo rispetto all'era medievale.

4. Nel frattempo, Briffa rifiutò ripetutamente di rilasciare i dati di misurazione della serie Yamal utilizzati nel suo calcolo nonostante i molteplici usi di questa serie su riviste che affermavano di richiedere l'archiviazione dei dati.

5. Poi, nel 2008, Briffa et al. Parteciparono ad uno studio di Mann utilizzando la serie Yamal, sempre senza archiviazione, ma pubblicarono su una rivista che ha rigide regole di condivisione dei dati. (Vedi cronologia di Mann punto 3)[6]

6. Poco tempo dopo, con l'aiuto degli editori della rivista, i dati poterono essere estratti e apparirono sul sito web della CRU.

7. Si scoprì quindi che la fine del XX secolo - nella serie Yamal - é composta da sole 10 cronologie di anelli d'albero successive al 1990 (e 5 successive al 1995), il che li rendeva un campione troppo esiguo per essere utilizzato secondo le regole convenzionali. Ma il vero problema non era che c'erano solo 5-10 nuclei della fine del XX secolo e che dovevano essercene molti di più. Il problema reale fu che gli studiosi usavano solo un sottoinsieme di 10 core a partire dal 1990 che non furono nemmeno selezionati casualmente.

8. Trovandosi davanti ad un campione nella cronologia Taymir che probabilmente aveva 3-4 volte più serie della cronologia Yamal, Briffa aggiunse dati da campioni di altri ricercatori prelevati nel sito di Avam, campioni prelevati a circa 400 km di distanza. Inoltre, usò i dati del programma di campionamento condotto da Schweingruber intorno al 1990, anch'essi presi a circa 400 km da Taymir. Indipendentemente dai meriti o meno del raggruppamento di campioni da tali luoghi così distanti, ciò stabilisce un precedente in cui Briffa aggiunse i dati di Schweingruber per fornire campioni aggiuntivi. Questo aumentò la solidità dell'Hockey Stick della cronologia (ora chiamata Avam - Taymir).

9. Steve McIntyre cercò dati da altri campioni presso o vicini al sito di Yamal che avrebbero potuto essere utilizzati per aumentare la dimensione del campione nella cronologia Briffa relativa a Yamal, e scoprì rapidamente una vasta serie di 34 campioni raccolti da Schweingruber da alberi viventi. Usando questi al posto dei 12 alberi del gruppo Briffa (e registrati nei dati del CRU) che si estendono fino al presente, si ottiene una completa

divergenza nel XX secolo (curva in nero nella immagine precedente). Quindi, i dati di Schweingruber contraddicono completamente la serie CRU (curva in rosso nell' immagine).

10. Combinando i dati CRU e Schweingruber si ottiene la linea verde nell' immagine. Come è facile notare, la parte incriminata a forma di 'lama' da mazza da hockey ha nella versione combinata un' apparenza molto diversa: non é presente nessun 'picco' di riscaldamento anomalo. Possiamo dire, quindi, che l'ingrediente chiave in molti degli studi che sono stati invocati a sostegno della mazza da hockey, vale a dire la serie Briffa Yamal (linea rossa), dipende dall'influenza di un sottile sottinsieme delle cronologie post-1990 e dall'esclusione della (molto più grande) raccolta di dati di Schweingruber prontamente disponibili per la stessa area.

OTTENERE LA HOCKEY STICK DAL RUMORE BIANCO

Una delle email meno diffuse e più sottovalutate della vicenda del Climategate (lo scandalo che coinvolse alcune università ed alcuni ricercatori - in primis Mann, Briffa, Schmidt, Jones e altri) venne scritta da Rob Wilson, allora un accademico itinerante e molto giovane; tale email (catalogata 4241. 2006-03-07 [8]) riportava simulazioni che supportavano in modo convincente le preoccupazioni di McIntyre sullo screening ex post, un metodo di screening dei dati molto controverso. L'e-mail di Wilson era indirizzata alla maggior parte dei principali dendroclimatologi dell'epoca: Ed Cook, Rosanne D'Arrigo, Gordon Jacoby, Jan Esper, Tim Osborn, Keith Briffa, Ulf Buentgen, David Frank, Brian Luckman ed Emma Watson, nonché Philip Brohan del Met Office. Riportiamo qui di seguito l' email di Wilson completa di traduzione:

Greetings All,

I thought you might be interested in these results. The wonderful thing about being paid properly (i. e. not by the hour) is that I have time to play.

The whole Macintyre issue got me thinking about over-fitting and the potential bias of screening against the target climate parameter. Therefore, I thought I'd play around with some randomly generated time-series and see if I could 'reconstruct' northern hemisphere temperatures.

I first generated 1000 random time-series in Excel – I did not try and approximate the persistence structure in tree-ring data. The autocorrelation therefore of the time-series was close to zero, although it did vary between each time-series. Playing around therefore with the AR persistent structure of these time-series would make a

difference. However, as these series are generally random white noise processes, I thought this would be a conservative test of any potential bias.

I then screened the time-series against NH mean annual temperatures and retained those series that correlated at the 90% C. L. 48 series passed this screening process.

Using three different methods, I developed a NH temperature reconstruction from these data:

- 1. simple mean of all 48 series after they had been normalised to their common period*
- 2. Stepwise multiple regression*
- 3. Principle component regression using a stepwise selection process.*

The results are attached. Interestingly, the averaging method produced the best results, although for each method there is a linear trend in the model residuals – perhaps an end-effect problem of over-fitting.

The reconstructions clearly show a ‘hockey-stick’ trend. I guess this is precisely the phenomenon that Macintyre has been going on about.

It is certainly worrying, but I do not think that it is a problem so long as one screens against LOCAL temperature data and not large scale temperature where trend dominates the correlation. I guess this over-fitting issue will be relevant to studies that rely more on trend coherence rather than inter-annual coherence. It would be interesting to do a similar analysis against the NAO or PDO indices. However, I should work on other things.

Thought you’d might find it interesting though. comments welcome

Rob

Traduzione:

Un saluto a tutti,

Ho pensato che potreste essere interessati a questi risultati. La cosa meravigliosa di essere pagato correttamente (cioè non a ore) è che ho tempo per giocare.

*L'intera questione di Macintyre mi ha fatto pensare all'eccessivo adattamento e al potenziale pregiudizio dello screening rispetto al parametro climatico target. Pertanto, **ho pensato di giocare con alcune serie temporali generate casualmente e vedere se potevo "ricostruire" le temperature dell'emisfero settentrionale.***

Per prima cosa **ho generato 1000 serie temporali casuali** in Excel: non ho cercato di approssimare la struttura di persistenza nei dati degli anelli d'albero. L'autocorrelazione quindi delle serie temporali era vicina allo zero, sebbene variasse tra ciascuna serie temporale. Giocando quindi con la struttura persistente AR di queste serie temporali farebbe la differenza. Tuttavia, **poiché queste serie sono generalmente processi casuali di rumore bianco**, ho pensato che questo sarebbe stato un test conservativo di qualsiasi potenziale distorsione.

Ho quindi **sottoposto a screening le serie temporali** rispetto alle temperature medie annuali di NH e ho **mantenuto quelle serie correlate alla serie CL 48 al 90%** che hanno superato questo processo di screening.

Utilizzando tre diversi metodi, ho sviluppato una ricostruzione della temperatura NH da questi dati:

- media semplice di tutte le 48 serie **dopo che sono state normalizzate** al loro periodo comune
- Regressione multipla graduale
- Regressione dei componenti principali mediante un processo di selezione graduale.

I risultati sono allegati. È interessante notare che **il metodo della media ha prodotto i risultati migliori**, sebbene per ogni metodo vi sia una tendenza lineare nei residui del modello, forse un problema di overfitting.

Le ricostruzioni mostrano chiaramente un andamento 'bastone da hockey'. Immagino che questo sia esattamente il fenomeno di cui parla Macintyre.

È certamente preoccupante, ma non penso che sia un problema **finché si schermano i dati della temperatura LOCALE e non la temperatura su larga scala dove la tendenza domina la correlazione**. Immagino che questo problema di adattamento sarà rilevante per gli studi che si basano più sulla coerenza delle tendenze piuttosto che sulla coerenza interannuale. Sarebbe interessante fare un'analisi simile rispetto agli indici NAO o DOP. Tuttavia, dovrei lavorare su altre cose.

Pensavo che avreste potuto trovarlo interessante però. Commenti benvenuti

Rob

Ho evidenziato in grassetto i passi fondamentali ed importanti di questa email, perchè il lettore possa rendersi conto della sua importanza. Sostanzialmete, Wilson sta informando i

suoi colleghi di un problema relativo allo screening dei dati con il metodo 'ex-post': tramite adattamenti, medie e normalizzazioni, é riuscito a ottenere una Hockey Stick su dati generati casualmente e senz nessun correlazione! Torneremo sul problema dello screening 'ex-post' più avanti, e ne vedremo davvero delle belle.

IL COMITATO DI RIANALISI

Dopo la pubblicazione della critica di McIntyre e McKitrick al lavoro di Mann, ci fu un immediato contrattacco da parte di alcuni climatologi che avevano lavorato a stretto contatto con Mann in passato. L'attacco a McIntyre e alla critica di McKitrick al lavoro di Mann si ridussero a due tattiche: alcuni ricercatori affiliati a Mann cercarono di trovare delle fallacie matematiche e concettuali al lavoro di McIntyre e McKitrick, e parallelamente la comunità affermò che ovviamente l' Hockey Stick scompare se si smette di usare le tecniche di Mann, ma che proprio per questo si devono continuare a usare tali tecniche di Mann!

Alla fine, venne istituita una commissione d'inchiesta del Senato degli Stati Uniti sotto la presidenza di Edward Wegman, un rispettato professore di Matematica e Statistica; il suo rapporto venne pubblicato nel 2006 [9]

Il rapporto esaminò il background dell'articolo di Mann relativo alla Hockey Stick, il documento stesso, la critica di McIntyre e McKitrick, e raccolse prove da tutti i personaggi chiave della vicenda. È interessante notare che il comitato di Wegman commissionò alcune ricerche originali su come funzionava effettivamente il ristretto 'mondo' della Climatologia. Lo studio del social networking del mondo della paleoclimatologia mostrò quanto questo fosse chiuso, e quanto spesso un piccolo gruppo di scienziati sia stato adibito alla peer reviewed dei documenti l' uno dell'altro. Per un lavoro così dipendente dall'uso della Statistica e che faceva affermazioni statistiche sulle tendenze, fu subito sorprendente che nessuno statistico professionista sembrasse mai essere stato coinvolto né nel lavoro di ricerca stesso, né nella sua revisione tra pari.

La scoperta chiave del rapporto Wegman fu questa:

"La nostra commissione ritiene che le valutazioni secondo cui il decennio degli anni '90 sia stato il decennio più caldo in un millennio e che il 1998 sia stato l'anno più caldo in un millennio non possono essere supportate dal MBH98/99 [il nome della carta originale del bastone da hockey di Mann]"

Molto interessanti sono anche le altre conclusioni specifiche del Rapporto Wegman; Le riportiamo per intero qui di seguito:

Conclusione 1: *La politicizzazione del lavoro accademico porta a dibattiti pubblici confusi. Gli articoli accademici pubblicati su riviste peer-reviewed sono considerati il record d'archivio della ricerca. Di solito, non è necessario archiviare materiale supplementare come codice e dati. Di conseguenza, il materiale supplementare per il lavoro accademico è spesso scarsamente documentato e archiviato, e non è sufficientemente robusto per resistere a un intenso dibattito pubblico. Nel presente esempio c'era troppa dipendenza dalla revisione tra pari, che sembrava non essere sufficientemente indipendente.*

Conclusione 2: *La condivisione di materiali di ricerca, dati e risultati è casuale e spesso eseguita a malincuore. Siamo rimasti particolarmente colpiti dall'insistenza del dottor Mann sul fatto che il codice che ha sviluppato fosse sua proprietà intellettuale e che potesse detenerlo legalmente personalmente senza rivelarlo ai colleghi. Quando codice e dati non sono condivisi e la metodologia non è completamente divulgata, i peer non hanno la capacità di replicare il lavoro e quindi la verifica indipendente è impossibile.*

Conclusione 3: *Come statistici, siamo rimasti colpiti dall'isolamento di comunità, come la comunità del Paleoclima che si basa molto su metodi statistici, ma non sembra interagire con la comunità statistica tradizionale. Le implicazioni di politica pubblica di questo dibattito sono sconcertanti dal punto di vista finanziario e tuttavia, a quanto pare, non è stata richiesta o utilizzata alcuna perizia statistica indipendente.*

Conclusione 4: *Sebbene la ricostruzione del Paleoclima abbia raccolto molta pubblicità perché rafforza un'agenda politica, non fornisce informazioni e comprensione dei meccanismi fisici del cambiamento climatico se non nella misura in cui gli anelli d'albero, le carote di ghiaccio e simili forniscono prove fisiche come la prevalenza dei Gas Serra. Ciò che è necessario è una comprensione più profonda dei meccanismi fisici del cambiamento climatico .*

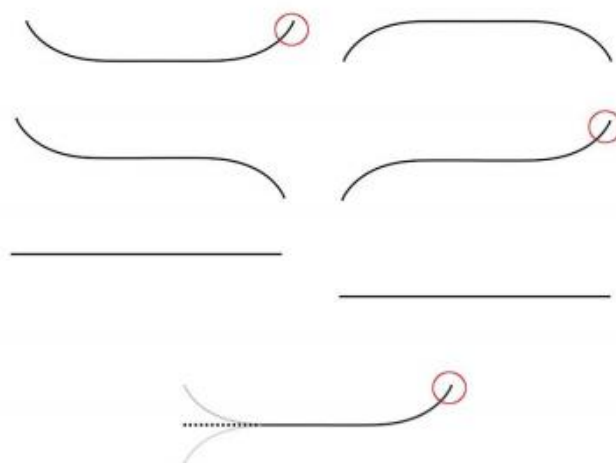
LO SCREENING 'EX-POST' E IL COEFFICIENTE R2

Torniamo per un attimo indietro, e ricollegiamoci all' email scritta da Rob Wilson ed indirizzata a Ed Cook, Rosanne D'Arrigo, Gordon Jacoby, Jan Esper, Tim Osborn, Keith Briffa, Ulf Buentgen, David Frank, Brian Luckman, Emma Watson, e Philip Brohan. Ho evidenziato alcuni passaggi chiave in cui Wilson ammette di aver giocato con dei numeri ed ottenuto un preoccupante andamento a 'mazza da hockey' pur avendo usato numeri casuali (rumore bianco). Fu possibile attuando un approccio 'ex post' al vaglio e alla selezione dei dati. Questa locuzione latina, antitetica ad 'ex ante', indica una operazione di analisi o considerazione 'dopo i fatti', nel caso della analisi dei dati, quindi, indica un controllo o una azione 'dopo

aver ottenuto i risultati'. Quando applicato allo screening, che é il vaglio o la selezione dei dati, la locuzione 'ex post screening' indica dunque una selezione di dati successiva all'ottenimento di un risultato (la composizione di un grafico, la ricerca di una correlazione, etc). Ad un lettore superficiale, questa pratica potrebbe sembrare normale o giustificata, ma non lo é nel caso in cui la selezione ex post porti ad eliminare parte dei dati da un grafico o da una correlazione non soddisfacenti, per eliminare gli ostacoli all'ottenimento di un andamento desiderato. Ed é esattamente ciò che Wilson stava segnalando di aver ottenuto dal rumore bianco: avendo preso dati non correlati, tramite metodi statistici applicati 'ex post', era riuscito a ottenere la Hockey Stick!

Nel campo della ricostruzione paleoclimatica, mentre effettuare screening sulla base della correlazione con la temperatura sembra superficialmente sensato, l'errore è facilmente comprensibile se si ipotizza che uno scienziato farmaceutico utilizzi lo screening ex post: immaginate uno studio farmacologico che riportasse risultati solo per i pazienti le cui condizioni migliorano in seguito ad assunzione di un farmaco. O un gestore di fondi finanziari che rendiconta solo gli investimenti che aumentano di valore. Una tale tecnica sarebbe ridicola ed i risultati fuorvianti. L'errore, nella letteratura statistica più ampia, è comunemente chiamato *screening sulla variabile dipendente*; gli studi che utilizzano questa tecnica fallace non hanno validità statistica. Non c'è niente di sbagliato nell'ipotizzare 'ex ante' che (per esempio) le larghezze degli anelli di abete rosso nei siti al limite del bosco siano un proxy sensibile alla temperatura. Ma una volta fatto, prendi il tuo campione di tutti gli alberi che ritieni adatti e usi tutti i dati risultanti. Se si escludono i siti in cui le larghezze degli anelli diminuiscono nel XX secolo, non è possibile utilizzare il fatto che la serie risultante salga nel XX secolo come prova di nulla, perché una tale tecnica darà la forma di una mazza da hockey anche ai dati che, in media, non ha tale andamento.

La distorsione dello screening ex post è illustrata nel diagramma seguente:



Questo diagramma mostra quattro serie che, se mediate, producono una linea retta. Ma quando vengono selezionate in base al fatto che la serie sia salita nel XX secolo (cerchio rosso), la loro media è una mazza da hockey.

Perchè tutto questo è importante? Sia perchè lo studio di Mann et al commise lo stesso reato di manipolazione statistica 'ex post', sia perchè Wilson stava segnalando che con questa metodologia era possibile ottenere la Hockey Stick praticamente da qualsiasi cosa.

L' email di Wilson fu facilmente e immediatamente liquidata da Ed Cook, che rispose (forse ironicamente) "sei un masochista!" [10]:

You are a masochist. Maybe Tom Melvin has it right: "Controversy about which bull caused mess not relevant. The possibility that the results in all cases were heap of dung has been missed by commentators"

Traduzione:

Sei un masochista. Forse Tom Melvin ha ragione: "La controversia su quale toro ha causato disordine non è rilevante. La possibilità che i risultati in tutti i casi fossero un mucchio di sterco è stata persa dai commentatori".

Lo studio di Mann et al conteneva un secondo tipo di analisi e di operazione statistica: la ricostruzione tramite regressione, validata secondo Mann dal coefficiente di determinazione r^2 . [11] Di cosa si tratta? Non esiste una definizione concordata del coefficiente r^2 : in statistica, esso è una proporzione tra la variabilità dei dati e la correttezza del modello statistico utilizzato. Esso misura la frazione della varianza della variabile dipendente espressa dalla regressione. Il suo valore varia tra 0 ed 1: quando è 0 il modello utilizzato non spiega per nulla i dati; quando è 1 il modello spiega perfettamente i dati.

Quel che però è importante sapere di questo coefficiente, è che questo coefficiente ha grosse limitazioni ben note (si veda ad esempio "*The Coefficient of Determination—Some Limitations*" di J.P. Barrett, 1974 [12]): ad esempio non permette di sapere nè se una variabile sia statisticamente significativa, nè se i regressori sono causa effettiva dei movimenti della variabile dipendente, nè se c'è una distorsione da variabile omessa, nè se è stato scelto il gruppo dei regressori più appropriato.

Ora, vediamo perchè è importante parlare del coefficiente r^2 ai fini della vicenda di Mann. In una famosa (e famigerata) e-mail del 2003 (discussa in dettaglio sul sito Climate Audit entro le prime due settimane di Climategate), Mann inviò alcuni calcoli - prima mai divulgati e provenienti dalla ricostruzione della temperatura di Mann et al 1998 - a Tim Osborn, un "collega fidato", dicendogli che la serie rappresentava i suoi "panni sporchi" e i dati

dovevano essere mantenuti strettamente confidenziali, per non cadere nelle "mani" sbagliate. Di seguito il contenuto dell' email (catalogata: email 1059664704):

From: "Michael E. Mann" <mann@virginia.edu>

To: Tim Osborn <t.osborn@uea.ac.uk>

Subject: Re: reconstruction errors

Date: Thu, 31 Jul 2003 11:18:24 -0400

Tim,

Attached are the calibration residual series for experiments based on available networks

back to:

AD 1000

AD 1400

AD 1600

I can't find the one for the network back to 1820! But basically, you'll see that the residuals are pretty red for the first 2 cases, and then not significantly red for the 3rd case--its even a bit better for the AD 1700 and 1820 cases, but I can't seem to dig them up. In any case, the incremental changes are modest after 1600--its pretty clear that key predictors drop out before AD 1600, hence the redness of the residuals, and the notably larger uncertainties farther back...

You only want to look at the first column (year) and second column (residual) of the files.

I can't even remember what the other columns are!

Let me know if that helps. Thanks,

mike

p.s. I know I probably don't need to mention this, but just to insure absolutely clarify on this, I'm providing these for your own personal use, since you're a trusted colleague. So please don't pass this along to others without checking w/ me first. This is the sort of "dirty laundry" one doesn't want to fall into the hands of those who might potentially try to distort things...

La traduzione:

Da: "Michael E. Mann" <mann@virginia.edu>

A: Tim Osborn <t.osborn@uea.ac.uk>

Oggetto: Re: errori di ricostruzione

Data: gio, 31 luglio 2003 11:18:24 -0400

Tim,

In allegato le serie residue di calibrazione per gli esperimenti basati sulle reti disponibili

torna a:

dC 1000

1400 d.C

1600 d.C

Non riesco a trovare quello per la rete fino al 1820! Ma fondamentalmente, vedrai che i residui sono piuttosto rossi per i primi 2 casi, e quindi non significativamente rossi per il terzo caso - è anche un po' meglio per i casi AD 1700 e 1820, ma non riesco a scavarli. In ogni caso, i cambiamenti incrementali sono modesti dopo il 1600: è abbastanza chiaro che i predittori chiave scompaiono prima del 1600 d.C., da qui il rossore dei residui e le incertezze notevolmente più grandi più indietro...

Si desidera esaminare solo la prima colonna (anno) e la seconda colonna (residuo) dei file.

Non ricordo nemmeno quali sono le altre colonne!

Fammi sapere se questo aiuta. Grazie,

Mike

p.s. So che probabilmente non ho bisogno di menzionarlo, ma solo per essere assolutamente chiaro su questo, sto fornendo questi per il tuo uso personale, dato che sei un collega fidato. Quindi, per favore, non trasmetterlo ad altri senza prima aver controllato con me. Questo è il tipo di "biancheria sporca" che non si vuole far cadere nelle mani di coloro che potrebbero potenzialmente provare a distorcere le cose...

Le serie di dati che Mann chiama "biancheria sporca" sono chiamate *residui*. Sono le differenze tra le stime di ricostruzione tramite proxy delle temperature passate e i record di temperatura osservati durante le fasi di stima del modello (la "calibrazione") e di test (la "verifica"). Poiché tali residui misurano la bontà di adattamento del modello, sono essenziali per calcolare i punteggi dei test di verifica. In questa e-mail, Mann forniva residui per ricostruzioni (che chiama impropriamente "esperimenti") basate sugli intervalli post-1000, post-1400 e post-1600. I primi due sono importantissimi, poiché determinano se è legittimo fare la ricostruzione indietro fino a quel punto. IN sostanza: se le prime due serie si dimostrassero inattendibili, verrebbe a cadere l' appiattimento ottenuto da Mann nel periodo medievale, e verrebbero a mancare circa 400 anni di ricostruzione, rendendo il lavoro di Mann molto meno utile ed attendibile.

Numerose autorità statistiche, comprese quelle citate in Mann et al 1998 (Fritts, Cook e Peters), raccomandano di testare la validità della ricostruzione utilizzando diversi punteggi basati sui residui. Mann dichiarò nel suo articolo del 1998 di aver calcolato due di questi punteggi, la statistica sulla riduzione dell'errore (chiamata RoE o RE) e il punteggio r2, proprio il coefficiente di cui abbiamo parlato poc' anzi. Ma nel suo articolo e nell'archivio allegato Mann elencò solo i valori dell' RE. Non aveva (e non ha mai) pubblicato i punteggi r2. Né potrebbero questi essere facilmente calcolati dalle informazioni divulgate con la pubblicazione originale, perché, contrariamente alla credenza diffusa tra gli scienziati del Clima, l'archivio di Mann non conteneva le ricostruzioni complete per ogni fase temporale. Per la tipica ricostruzione dell'emisfero settentrionale (NH), Mann archiviò solo i 'segmenti di ricostruzione giuntati' in cui, ad ogni passaggio temporale, i risultati di un passaggio

successivo sono stati stampati sui risultati di passaggi precedenti. Senza la serie residua, nessuno sarebbe in grado di ricalcolare i punteggi r^2 non riportati.

Verso la fine del 2003, solo pochi mesi dopo l'e-mail della "biancheria sporca", McIntyre e McKittrick chiesero a Mann di fornire la serie residua per la fase AD1400 della sua ricostruzione. Mann rifiutò. I due ricercatori allora presentarono un reclamo alla rivista Nature, che aveva pubblicato lo studio originale del 1998, appellandosi alle loro politiche di divulgazione sia per le serie residue che per le fasi di ricostruzione. Vergognosamente, anche Nature rifiutò di fornire dati o fare pressione su Mann perchè li fornisse. Il terzo step fu per McIntyre e McKittrick il rivolgersi alla US National Science Foundation, perchè richiedessero a Mann di fornire questi dati. Anche loro si rifiutarono. Nonostante la disinformazione messa in atto dai 'warmisti' colleghi e sostenitori di Mann, i risultati dei singoli passi di Mann rimangono non archiviati fino ad oggi.

Nel 2004, McIntyre e McKittrick finalmente scoprirono il motivo per cui Mann era così irremovibile nel trattenere la sua "biancheria sporca". All'inizio del 2004, nonostante molti ostacoli, i due furono in grado di replicare la metodologia peculiare e scarsamente documentata di Mann abbastanza bene da calcolare le serie residue (e le statistiche di verifica) per il passo AD1400. Scoprirono, con notevole sorpresa, che la statistica di verifica r^2 per il passo AD1400 era disastrosamente bassa (solo 0,018). La verifica tramite coefficiente r^2 è una verifica statistica banale, che dovrebbe essere facilmente superata da qualsiasi ricostruzione che pretende di avere "abilità" statistica nello stabilire un minimo di correlazione. Valori vicini a 1 vengono quasi sempre ottenuti in grafici che mostrano correlazioni reali, anche senza consequenzialità. E' facile, ad esempio, trovare misure correlate non legate da nessi causali che condividono un r^2 anche di 0.6 o 0.8 (si veda ad esempio i grafici ironici del sito "Spurious Correlations"). Come abbiamo già detto, l' r^2 non è una garanzia di validità del modello, ma il suo fallimento (valori molto vicini a 0) è più o meno una garanzia di invalidità del modello. I valori di r^2 ottenuti da McIntyre e McKittrick usando le procedure di Mann, erano praticamente uguali a zero!

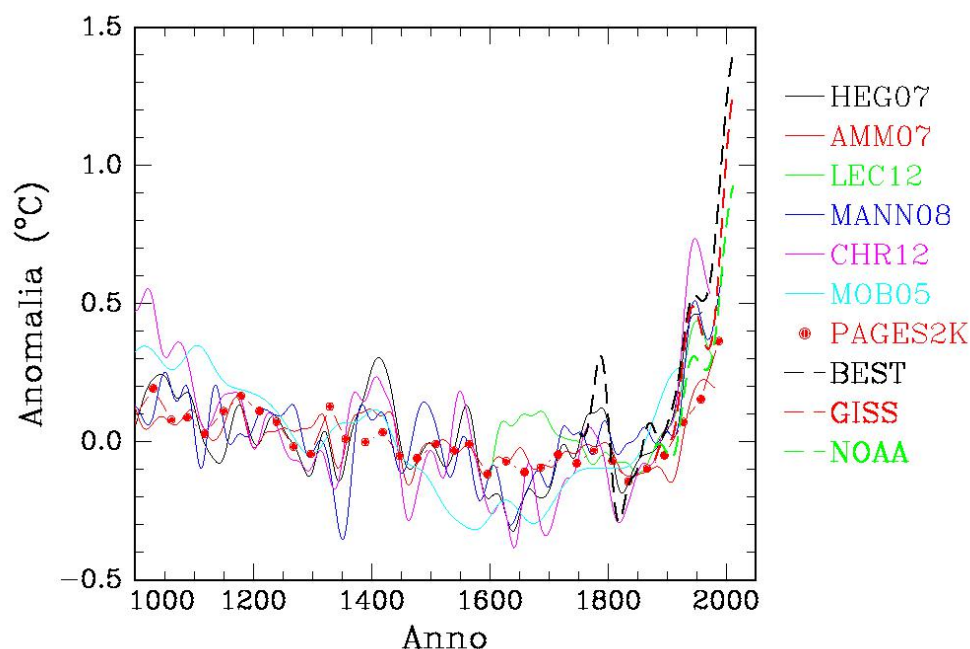
I due riportarono le loro scoperte in due articoli del 2005 ("*Hockey sticks, principal components, and spurious significance*" e "*The M&M critique of the mbh98 northern hemisphere climate index: update and implications*"). Questi due articoli, passati quasi inosservati nella comunità climatica, ma sollevarono diverse reazioni dai media e da enti influenti, compreso un articolo in prima pagina sul Wall St Journal che attirò l'attenzione del Comitato per l'Energia e il Commercio della Camera degli Stati Uniti. Il Comitato inviò una serie di domande a Mann, comprese quelle sul codice sorgente e sulle statistiche di verifica r^2 . Queste domande provocarono vigorose proteste da parte dell'AAAS, dell'AGU e di altre istituzioni scientifiche. Ralph Cicerone, allora presidente della National Academy of Sciences,

scrise al House Energy and Commerce Committee offrendo i propri servizi, incluso, in particolare, l'esame della verifica tramite coefficiente r^2 . Tale richiesta non venne accettata.

In risposta parziale alle domande del Comitato, Mann archiviò online alcune porzioni del codice sorgente per il suo articolo del 1998. Sebbene incompleto, questo codice confermò la supposizione di McInTyre e McKittrick che Mann avesse calcolato le statistiche di verifica r^2 per ogni passaggio della ricostruzione dell' Emisfero Nord (NH) e le avesse trattenute dalla divulgazione. Ad oggi, Mann non ha mai archiviato le ricostruzioni NH per singoli passaggi, le serie residue equivalenti (appunto, la "biancheria sporca") e nemmeno i risultati della verifica tramite r^2 .

LE PRESUNTE VALIDAZIONI DELL' HOCKEY STICK

Nonostante tutte le controversie legate alla vicenda, ancora molti studiosi e molti magazine scientifici e divulgativi continuano a proporre l' Hockey Stick ed a ritenerlo attendibile e conclusivo. Un famoso sito italiano di ambientalisti warmisti, Climalteranti, ha proposto nel 2013 un articolo intitolato "*Quindici anni di mazze da hockey*" [7] nel quale un docente di Fisica (Riccardo Reitano) sostiene che la Hockey Stick di Mann et al sia stata nel corso degli anni validata da numerosi studi. Reitano, ovviamente, fornisce una serie di link a studi verificabili, e fornisce una immagine riassuntiva, che riportiamo qui di seguito:



Sembrerebbe quindi che l' Hockey Stick sia stata verificata da numerosi studi verificabili. Tutto ok? Non proprio, perchè ci sono due punti oscuri in tutta la vicenda.

Il 90% degli studi che pretendono di aver verificato la Hockey Stick si basa sullo stesso set di dendrocronologie utilizzate da Mann e dai suoi collaboratori, utilizzando quindi la stessa fonte inadatta selezionata ad arte. Molti di questi sono diventati famosi per la loro presunta indipendenza poi dimostrata falsa in quanto utilizzavano gli stessi dataset e le stesse metodiche di regressione. Il caso più famoso di analisi indipendente fu il PAGES2K multiproxy networks, che fu prodotto in 3 versioni: nel 2013, nel 2017 e nel 2019.

Questa analisi multiproxy, alla quale partecipò anche Mann, rendendo quindi evidentemente non indipendente lo studio, presenta una serie di problematiche che qui elenchiamo:

- PAGES2K e gli studi simili rimangono principalmente dipendenti da dati problematici e incoerenti sugli anelli d'albero, molti dei quali hanno un trend in discesa nell'ultima metà del XX secolo. Per estrarre una forma di mazza da hockey da dati inconsistenti sugli anelli degli alberi, gli accademici del clima, tra cui appunto gli specialisti del PAGES2K, hanno fatto ricorso a metodi ad hoc (screening ex post, orientamento ex post) che sono condannati dagli statistici tradizionali e dalla letteratura statistica, ma che migliorano la forma a mazza da hockey nella ricostruzione risultante. Lo screening e la manipolazione ex post si estendono anche ai dati utilizzati in rapporti apparentemente tecnici;
- L'uso delle larghezze degli anelli degli alberi come proxy della temperatura è reso ancora più problematico dall'impatto dello straordinario "inverdimento" mondiale negli ultimi 30 anni - attribuito principalmente dagli specialisti alla fertilizzazione con anidride carbonica - sulle larghezze degli anelli. Tale effetto non è considerato in PAGES2K;
- La controversa serie di stripbark bristlecone, su cui si basava lo studio Mann et al 1998, continua ad essere utilizzata nei rapporti PAGES2K, anche se il pannello NAS del 2006 ha raccomandato che tali dati vengano "evitati" nelle ricostruzioni della temperatura;
- Nel loro tentativo di ottenere un grafico a mazza da hockey, gli autori di PAGES2K, come Mann et al 2008, hanno introdotto serie di sedimenti senza preoccuparsi di garantire un collegamento fisico, portando a una serie di imbarazzanti gaffe derivanti da serie contaminate da run-off di costruzione e persino da serie usate al contrario.

Nella loro rete nordamericana (un importante set di dati che conteneva i bristlecones stripbark), c'erano 146 cronologie di anelli d'albero presumibilmente accuratamente selezionate nella rete del 2013. Ma, solo 4 anni, dopo ne vennero eliminate 124 e ne vennero introdotte 129 di nuove. La situazione è riassunta in tabella:

**Table 1. North American Tree Ring Proxy Network,
PAGES 2013 and PAGES 2017**

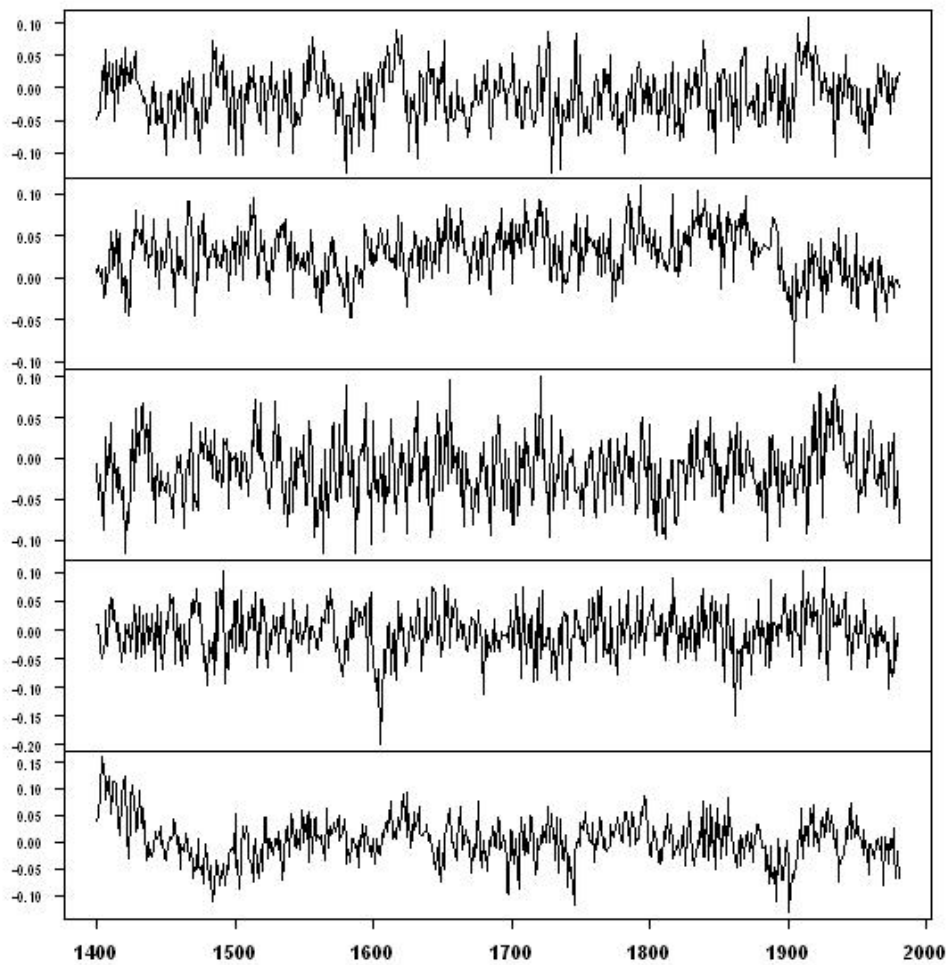
	PAGES (2013)	PAGES (2017)	Combined
PAGES 2013 only	124	.	124
PAGES 2017 only	.	129	129
Both	22	22	22
TOTAL	146	151	275

Questo cambio di serie non derivò da nuove serie rese disponibili tra il 2013 e il 2017, ma da un cambio del sistema di screening e orientamento ex post. Il motivo dell' eliminazione di così tante cronologie PAGES2K-2013 infatti fu l'introduzione del seguente criterio di screening ex post nella rete 2017: :

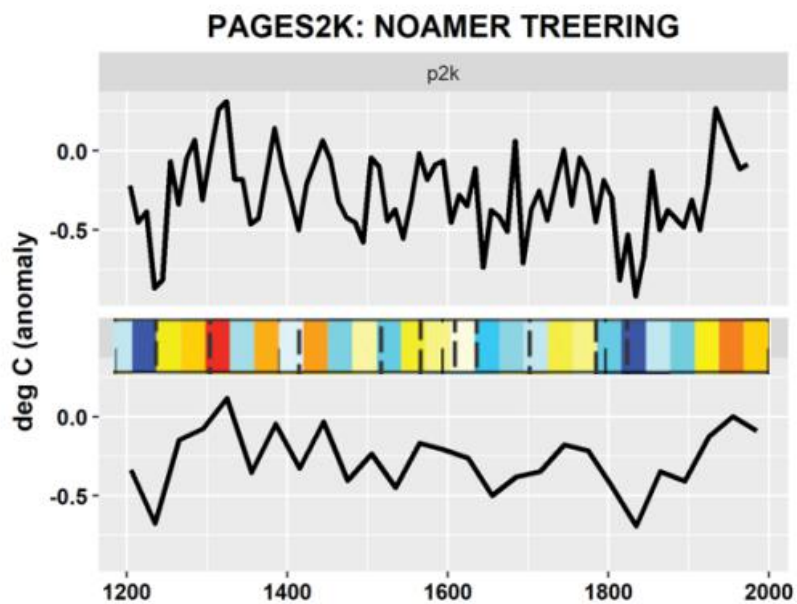
Per essere inclusi nell'attuale database, i dati sugli anelli degli alberi dovevano essere correlati positivamente ($P < 0,05$) con la temperatura locale o regionale (media dell'intero anno o durante la stagione di crescita). [13]

In altre parole, l'85% delle cronologie degli anelli degli alberi nordamericani "scelte con cura" utilizzate in PAGES2K-2013 non era correlato positivamente con la temperatura, anche con quattro possibilità (intero anno o stagione di crescita, locale o regionale). Alcune di queste serie provengono dai siti utilizzati nella ricostruzione originale di Briffa , con il suo inquietante declino nell'ultima metà del XX secolo.

Questo declino era comune a molte serie incluse nel PAGES2K-2013:

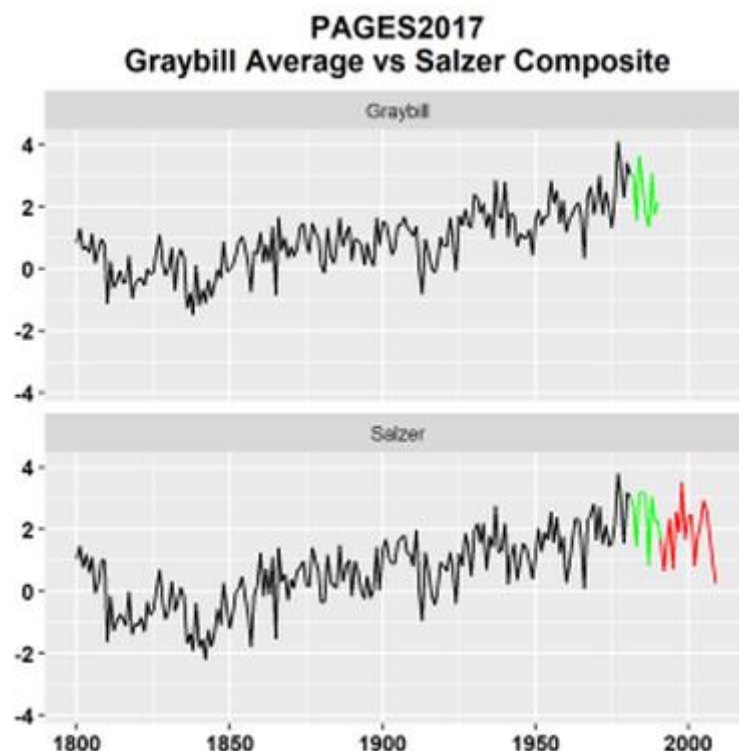


Altre cronologie smentivano invece il fatto che l'andamento in ascesa del XX secolo fosse senza precedenti, mostrando una addirittura maggiore impennata a ridosso del 1300 circa:



Per sostituire le 124 cronologie 'scomode', gli autori di PAGES2K-2017 cercarono nel database ITRDB (International Tree-Ring Data Bank) cronologie sostitutive con una correlazione positiva, individuandone 129. Non riportarono il numero di cronologie esaminate - sebbene sarebbe stato molto rilevante per la determinazione del significato dei dati statistici. Ci sono circa 3100 set di dati di misurazione degli anelli degli alberi nordamericani nel database ITRDB, da cui 151 (4,9%) vennero selezionati come aventi una correlazione $P < 0,05$ con la temperatura. Approssimativamente la stessa percentuale che ci si aspetterebbe da dati casuali. Sorprendentemente, la maggior parte delle cronologie degli anelli degli alberi che sono state utilizzate sia in PAGES2K-2013 che in PAGES2K-2017 erano le controverse *Graybill stripbark chronologies*, che avevano impresso la forma della mazza da hockey sulla ricostruzione di Mann et al 1998 e che la NAS (National Academy of Sciences) nel 2006 aveva consigliato specificatamente di evitare nelle ricostruzioni della temperatura.

Quindi, rispetto a questi proxy più controversi, PAGES2K non costituisce studio indipendente, ma continua a fare affidamento sugli stessi dati che avevano causato precedenti controversie. Non da ultimo, segnaliamo che anche nel PAGES2K-2017 molte cronologie usate per ricostruzioni mostravano un declino di temperatura negli ultimi decenni:



Concludiamo dunque questa sezione riassumendo:

- lo screening ex post basato su tendenze proxy recenti necessariamente distorce i dati risultanti verso una forma a mazza da hockey, una critica mossa più e più volte da molti

autori considerati 'scettici', ma non compresa da Michael Mann e dalla comunità paleoclimatica dell'IPCC;

- la rete di anelli d'albero nordamericana PAGES 2017 è stata severamente selezionata tramite metodiche 'ex post' estraendo a piacere dati da una popolazione di dati molto più ampia: nel corso degli anni, sono state utilizzate circa 983 diverse cronologie di anelli degli alberi nordamericani in MBH98, Mann et al 2008, PAGES2K-2013 e PAGES2K-2017. Vale a dire che fu preso soltanto il 15% circa della intera disponibilità di dati (oltre 3100 set) e questo campione fu selezionato ex post, una procedura che, anche con dati casuali, avrebbe conferito l'Hockey Stick-ness a qualsiasi composito risultante;
- nonostante questo severo screening ex post (sia in PAGES2K-2013 che in PAGES2K-2017), il composito di tutti i dati diversi dai bristlecones stripbark non ha mostrato alcuna Hockey Stick-ness e non rappresenta un proxy della temperatura .
- PAGES2K-2013 e PAGES2K-2017 perpetuano l'uso delle cronologie Graybill stripbark, nonostante la raccomandazione del NAS Panel del 2006 che queste serie problematiche vengano "evitate" nelle ricostruzioni future. PAGES2K-2013 (come Mann et al 2008) usava tutte le 20 cronologie stripbark. PAGES2K-2017 ha continuato l'uso della maggior parte delle cronologie stripbark sia nella versione originale Graybill che in un composito più recente (Salzer et al 2014) ignorando che la ricostruzione negava la presenza di Hockey Stick e la 'mancaza di precedenti' dell' ascesa dei valori di temperatura relativa nel XX secolo.

Possiamo dunque concludere che la Hockey Stick é un grafico fasullo, creato tramite metodi statistici usati fraudolentemente, e tramite la selezione ex-post (a posteriori) di dati appositamente scelti per ottenere un tipo di grafico preconfezionato. I presunti studi indipendenti che sostengono la abbiano validata, non sono nè indipendenti nè autorevoli in quanto basati sulle stesse cronologie scelte a piacere e con manipolazione dei dati. Il metodo statistico utilizzato per produrre la Hockey Stick é inaffidabile poichè capace di produrla anche da 'rumore bianco' (dati casuali non correlati).

ALESSANDRO DEMONTIS

Chimico, tecnico ambientale certificato CEE

Agosto 2021

FONTI CONSULTABILI

- [0] <https://www.nature.com/articles/33859>
- [1] http://www.meteo.psu.edu/holocene/public_html/Mann/research/res_pages/ONLINE-PREPRINTS/Millennium/mbh99.pdf
- [2] <https://climateaudit.files.wordpress.com/2005/09/mcintyre.mckitrick.2003.pdf>
- [3] https://www.researchgate.net/publication/228575617_Hockey_Sticks_Principal_Components_and_Spurious_Significance_Geophys
- [4] <https://www.realclimate.org/index.php/archives/2004/12/false-claims-by-mcintyre-and-mckitrick-regarding-the-mann-et-al-1998reconstruction/>
- [5] https://www.researchgate.net/publication/242878090_Unusual_twentieth-century_summer_warmth_in_a_1000-year_temperature_record_from_Siberia
- [6] <https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2007.2199>
- [7] <https://www.climalteranti.it/2013/07/24/quindici-anni-di-mazze-da-hockey/>
- [8] <http://di2.nu/foia/foia2011/mail/4241.txt>
- [9] https://climateaudit.files.wordpress.com/2007/11/07142006_wegman_report.pdf
- [10] <http://di2.nu/foia/foia2011/mail/4241.txt>
- [11] https://it.wikipedia.org/wiki/Coefficiente_di_determinazione#Interpretazione
- [12] <https://www.tandfonline.com/doi/abs/10.1080/00031305.1974.10479056>
- [13] <https://www.nature.com/articles/sdata201788>