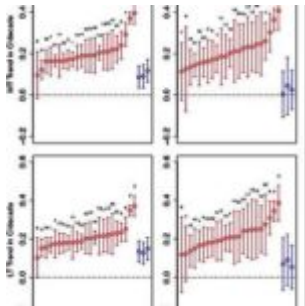


# Neue Studie – Vergleich: Modell mit Beobachtungen in der tropischen Troposphäre



Unsere Studie trägt den Titel [HAC-Robust Trend Comparisons Among Climate Series With Possible Level Shifts.](#) Sie wurde veröffentlicht in *Environmetrics*, und ist frei verfügbar durch eine finanzielle Unterstützung von [CIGI/INET](#). Daten und Code sind [hier](#).

## Tropische Troposphäre wieder aufgegriffen

Das Thema Modell ↔ Beobachtungen in der Troposphäre über den Tropen wurde schon vielfach behandelt, auch hier bei ClimateAudit. Kurz zur Erinnerung:

- Alle Klimamodelle (GCMs) sagen voraus, dass die Erwärmung als Folge des steigenden CO<sub>2</sub>-Niveaus in der Troposphäre über den Tropen rapide und zunehmend verstärkt auftreten wird. Siehe [Abbildung 9.1](#) im AR 4 und die [begleitende Diskussion](#) sowie den Text im AR 4 zur [Abbildung 10.7](#).
- Das richtige Abbilden der tropischen Troposphäre in einem Modell ist wichtig, weil dort der größte Anteil der Solarstrahlung in das Klimasystem eintritt. Außerdem gibt es dort eine hohe Konzentration von Wasserdampf, und die stärksten Rückkopplungen sind aktiv. In vereinfachten Modellen finden als Folge einer gleichmäßigen Erwärmung mit konstanter relativer Feuchtigkeit etwa 55% der Verstärkung der gesamten Erwärmung in der tropischen Troposphäre statt im Vergleich zu 10% am Boden und 35% in der Troposphäre außerhalb der Tropen. Und innerhalb der Tropen gibt es etwa zwei Drittel der zusätzlichen Erwärmung in der oberen Schicht und ein Drittel in der unteren Schicht. ([Soden & Held](#) p. 464).
- Weder Wettersatelliten noch Radiosonden haben viel Erwärmung in der tropischen Troposphäre gefunden, wenn überhaupt irgendeine, besonders im Vergleich zu dem, was die GCMs vorhergesagt haben. Im Bericht des US Climate Change Science Program ([Karl et al 2006](#)) wird dies als eine „potentiell ernste Inkonsistenz“ gelistet (S. 11). Ich denke, es ist an der Zeit, das Wort „potentiell“ fallen zu lassen.
- Der fehlende Hot Spot hat zu vielen Diskussionen in der Blogosphäre (z. B. [hier bei Jo Nova](#)) und unter Experten (z. B. [hier](#)) geführt. Es gibt zwei mit dem Hot Spot in Verbindung stehende Dinge: Verstärkung und Sensitivität.

Ersteres steht dafür, ob das Verhältnis der troposphärischen zur bodennahen Erwärmung größer ist als eins; letzteres bezieht sich auf die Frage, ob es eine starke troposphärische Erwärmungsrate gibt. Unsere Analyse konzentrierte sich auf den Aspekt Sensitivität, nicht auf die Verstärkung. Um die Verstärkung zu belegen, hätte es in der Höhe eine Menge Erwärmung geben müssen, was jedoch, wie sich herausgestellt hat, nicht der Fall war. Die Sensitivität kann direkt getestet werden, und genau das haben wir getan. Außerdem ist es in jedem Falle die relevantere Frage zur Messung der globalen Erwärmungsrate.

- Im Jahre 2007 haben [Douglass et al.](#) eine Studie im IJOC veröffentlicht, die gezeigt hat, dass die Modelle die Erwärmungstrends in jeder Schicht der tropischen Troposphäre überschätzt haben. [Santer et al.](#) (2008) erwiderten, falls man Autokorrelationen in den Daten kontrolliert, sind die Trend-Differenzen statistisch nicht signifikant. Dieses Ergebnis war sehr einflussreich. Die EPA hat sich darauf [verlassen](#) in ihrer Antwort auf Kritiken ihrer Projektionen der Klimaschäden im Technical Support Document hinter der „Gefahren-Findung“. Dies war die Grundlage ihrer Verkündung ständig neuer Vorschriften bzgl. Treibhausgasen. Es war auch die Grundlage für das Ergebnis von [Thorne et al.](#) (2011), deren Begutachtung „keine vernünftigen Beweise einer fundamentalen Nichtübereinstimmung zwischen Modellen und Beobachtungen“ in der tropischen Troposphäre ergeben hatte.

- Aber aus irgendwelchen Gründen haben Santer et al. ihre Daten im Jahre 1999 abgeschnitten, gerade am Ende eines starken El Nino. Steve und ich sandten einen [Kommentar](#) an IJOC, in dem wir darauf hinwiesen, dass falls sie ihr Verfahren auf die gesamte Länge der damals vorhandenen Daten ausgedehnt hätten, sie ein sehr unterschiedliches Ergebnis bekommen hätten, nämlich eine signifikante Über-Vorhersage der Modelle. Das IJOC wollte unseren Kommentar nicht veröffentlichen.

- Später habe ich diese Analyse noch einmal durchgeführt, und zwar mit der vollen Länge der verfügbaren Daten und mittels einer konventionellen Regressionsmethode und einer neueren robusteren Vergleichsmethodik, nämlich der non-parametric HAC (heteroskedasticity and autocorrelation)-robust estimator, entwickelt von Tim Vogelsang und Philip Hans Franses ([VF2005](#)). Ich zeigte, dass die Klimamodelle im Intervall 1979 bis 2009 im Mittel eine zwei- bis viermal zu starke Erwärmung in der unteren und mittleren tropischen Troposphäre vorhergesagt haben, und die Diskrepanzen waren sehr wohl statistisch signifikant. Diese Studie wurde als [MMH2010](#) in Atmospheric Science Letters veröffentlicht.

- Im AR 5 ist das IPCC erfreulich offen zu dem Thema ([Seiten 772-73](#)). Sie bestätigen die Ergebnisse in [MMH2010](#) (und auch andere Studien seitdem haben den Punkt bestätigt) und schließen daraus, dass die Modelle während der Satellitenära (nach 1979) die troposphärische Erwärmung überschätzt haben. Allerdings behauptet es auch, dass der größte Teil des Bias' der Modell-Überschätzung der Erwärmung der Wassertemperatur in den Tropen geschuldet ist. Aus dem Text geht nicht hervor, wie es zu dieser Erkenntnis gekommen ist. Da der Bias deutlich unter den Modellen schwankt, scheint es mir wahrscheinlich, dass es etwas mit der falschen Parametrisierung der Rückkopplungen zu tun hat. Auch besteht das Problem selbst in den Studien,

die die Modelle auf das beobachtete Niveau der Wassertemperaturen beschränken.

- Trotz des Scheiterns der Modelle, die tropische Troposphäre richtig abzubilden, erklärt die [SPM of the AR5](#) *sehr hohes Vertrauen* in die Klimamodelle, wenn es zur Diskussion über die Genauigkeit der Temperaturtrends kommt (S. 15). Sie erklären aber auch ihr *geringes Vertrauen* in ihrer Behandlung der Wolken (S. 16), was nur schwer in Einklang zu bringen ist mit ihrer Behauptung des *sehr hohen Vertrauens* in die Modelle allgemein. Trend-Diskrepanzen in Zeiträumen von 10 bis 15 Jahren scheinen ihnen nichts auszumachen (S. 15). Schauen wir, was sie zu Diskrepanzen über 55 Jahre sagen.

## **Die Ballon-Aufzeichnungen**

Nach der Veröffentlichung von MMH2010 beschloss ich, die Analyse auszuweiten bis zum Beginn der Aufzeichnungen von Wetterballonen im Jahre 1958. Ich wusste, dass ich es mit der Pazifischen Klima-Verschiebung [Pacific Climate Shift] Ende der siebziger Jahre zu tun haben würde. Dies ist ein gut dokumentiertes Phänomen, bei dem eine grundlegende Umordnung der Meeresströme eine sprunghafte Änderung in vielen Temperaturreihen an den Rändern des Pazifiks zur Folge hatte, einschließlich der Aufzeichnungen von Wetterballonen. Legt man einen linearen Trend durch eine Reihe mit einer positiven ruckartigen Änderung in der Mitte, werden die Neigungs-Koeffizienten nach oben verzerrt. Als ich Tim fragte, ob die VF-Methode in einer Anwendung einer vermuteten mittleren Verschiebung angewendet werden könne, sagte er nein. Dies würde die Ableitung einer neuen asymptotischen Verteilung und kritischer Werte erfordern, wobei man die Möglichkeit bekannter oder unbekannter Unstetigkeiten in Kauf nehmen müsste. Er stimmte zu, mit der theoretischen Arbeit weiterzumachen, und wir begannen die Studie zu untersuchen.

Breiten Raum in der Studie nimmt die Ableitung des Verfahrens und die Etablierung von dessen Validität ein. Für Leser, die diesen Teil überspringen und sich fragen, warum dies notwendig ist, lautet die Antwort, dass es hier um seriöse empirische Disziplinen geht und dass man dabei solche Verfahren braucht, um statistische Werkzeuge zu validieren, die man später anwendet.

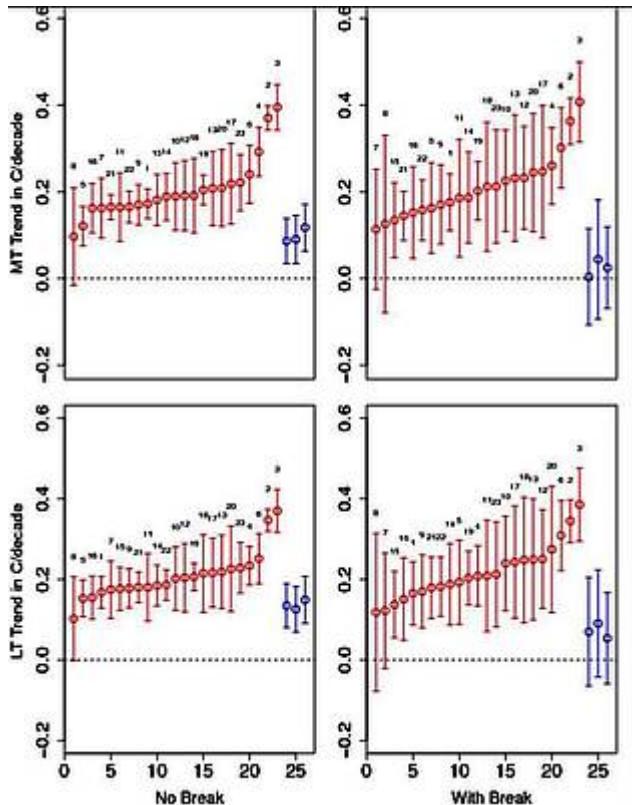
Unsere Studie zeigt einen Trend-Kalkulator und testet Statistiken auf der Grundlage von Standardfehlern, die in Gegenwart von Reihen-Korrelationen jeder Art gültig sind, aber nicht Grundelemente einschließt, die vom User nicht verlangen Tuning-Parameter zu wählen wie Bandbreite und Länge der Verzögerung; und die robust sind gegenüber der möglichen Gegenwart eines Verschiebungs-Terms zu einem bekannten oder unbekanntem Bruchpunkt. In der Studie präsentieren wir verschiedene Gruppen von Ergebnissen auf der Grundlage von drei möglichen Spezifikationen: i) es gibt keinen Verschiebungs-Term in den Daten, ii) es gibt eine Verschiebung zu einem bekannten Zeitpunkt (wir wählten Dezember 1977) und iii) es gibt einen möglichen Verschiebungs-Term, aber wir wissen nicht, wann dieser auftritt.

## **Ergebnisse**

- Alle Klimamodelle außer einem charakterisieren das Intervall von 1958 bis 2012 als einen signifikanten aufwärts gerichteten Temperaturtrend ausweisend. Die ruckartige Änderung Ende der siebziger Jahre hat grundsätzlich keine Auswirkungen in den vom Modell erzeugten Reihen. Die Hälfte der Modelle deutet einen kleinen positiven Ruck an und die andere Hälfte einen kleinen negativen Ruck, aber alle außer zwei zeigen einen erheblichen positiven und signifikanten Trend darum herum. Tatsächlich wird der Trend in der Hälfte aller Fälle sogar noch größer, wenn wir eine ruckartige Änderung zulassen. Im Ensemble-Mittel der GCMs gibt es Ende der siebziger Jahre keine ruckartige Änderung, sondern lediglich einen starken, ununterbrochenen und signifikanten Aufwärtstrend.
- Während des gleichen Zeitraumes finden wir signifikante Aufwärtstrends der unteren und mittleren tropischen Troposphäre, wenn wir die Beobachtungen nicht auf eine ruckartige Änderung abklopfen, obwohl der mittlere beobachtete Trend signifikant kleiner ist als der mittlere modellierte Trend.
- Wenn wir eine ruckartige Änderung Ende der siebziger Jahre in allen Radiosonden-Reihen zulassen, konzentriert sich der größte Teil der Temperaturzunahme nach 1958 sowohl in der unteren als auch in der mittleren Troposphäre auf die ruckartige Änderung, und die Trendneigung wird essentiell Null.
- Klimamodelle projizieren eine viel stärkere Erwärmung im Zeitraum 1958 bis 2012, als weder in der unteren noch in der mittleren Troposphäre beobachtet, und die Inkonsistenz ist statistisch signifikant, egal ob wir eine ruckartige Änderung zulassen oder nicht, aber wenn wir einen Verschiebungs-Term einführen, werden die Modelle auf niedrigere Signifikanz-Niveaus zurückgesetzt.
- Wenn wir den Bruchpunkt als unbekannt ansehen und einen Prozess mit den Daten zulassen, diesen zu identifizieren, ist der Verschiebungs-Term in der unteren Troposphäre kaum signifikant und in der mittleren Troposphäre signifikant, wobei der Bruchpunkt Mitte 1979 geschätzt wird.

Als wir vor ein paar Jahren mit der Arbeit an dieser Studie begannen, stammten die Daten aus der CMIP3-Modellbibliothek, welche wir in der Studie verwenden. Im AR 5 wurde die CMIP5-Bibliothek verwendet, so dass ich die Ergebnisse jener Läufe später erzeugen möchte. Für jetzt jedoch befasse ich mich mit den CMIP3-Ergebnissen.

Wir haben 23 CMIP3-Modelle und 3 Beobachtungsreihen. Hier folgt Abbildung 4 aus unserer Studie:

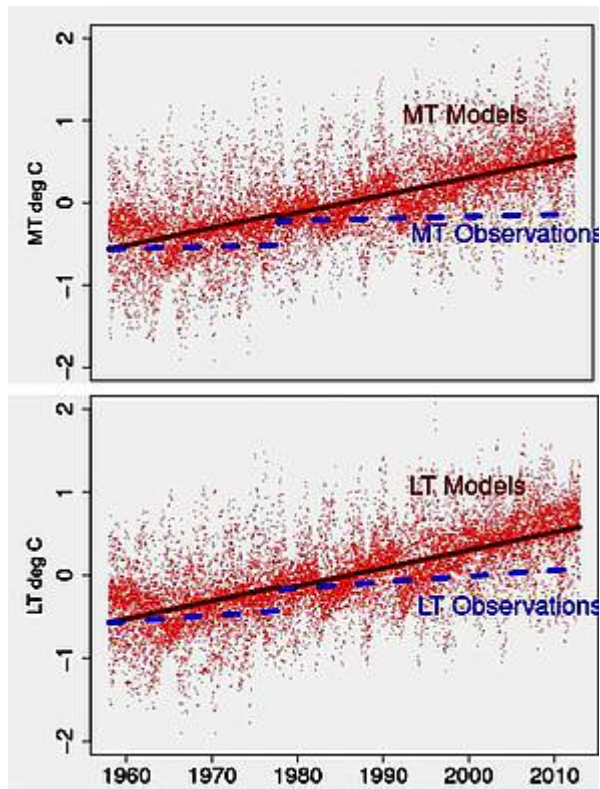


Jedes Paneel zeigt die Trend-Terme ( $^{\circ}\text{C}$  pro Jahrzehnt) und HAC-robuste Vertrauensintervalle für die CMIP-Modelle 1 bis 23 (rot) und die drei Wetterballon-Reihen (blau). Die linke Spalte zeigt den Fall ohne Abklopfen der Daten auf eine ruckartige Änderung. Die rechte Spalte zeigt den Fall, wenn wir das tun und den Bruchpunkt im Dezember 1977 datieren. Die obere Reihe zeigt die mittlere, die untere Reihe die untere Troposphäre.

Man erkennt, dass die Modelltrends etwa die Gleichen sind mit oder ohne den Verschiebungs-Term, obwohl sich die Vertrauensintervalle erweitern, wenn man eine Verschiebung zulässt. Lassen wir das nicht zu (linke Spalte) zeigen alle 6 Ballon-Reihen kleine, aber signifikante Trends. Wenn wir einen Sprung zulassen (rechte Spalte), datiert im Dezember 1977 werden alle beobachteten Trends sehr klein und statistisch unbedeutend. Alle außer zwei Modelle (GFDL 2.0 {#7} und GFDL 2.1 {#8}) zeigen in beiden Fällen positive und signifikante Trends.

### Mainstream gegen Realität

Abbildung 3 aus unserer Studie (unten) zeigt die vom Modell generierten Temperaturdaten, den mittleren GCM-Trend (rote Linie) und den angepassten [fitted] mittleren Ballon-Trend (blaue gestrichelte Linie) während des gesamten Zeitraumes. In allen Reihen (einschließlich derjenigen aller Klimamodelle) lassen wir eine Verschiebung im Dezember 1977 zu. Oben: Mittlere, unten: untere Troposphäre.



Die dunkelrote Linie zeigt den Trend des Ensemble-Mittels der Modelle. Da dies die zentrale Tendenz der Klimamodelle zeigt, können wir es als zentrale Tendenz des Mainstream-Denkens über die Klimadynamik ansehen und im Besonderen, wie das Klima auf den Antrieb durch steigende Treibhausgase reagiert. Die gestrichelte blaue Linie ist der angepasste Trend in den Beobachtungen, d. h. in der Realität. Angesichts der Größe und Dauer der Diskrepanz und der Tatsache, dass die Trends in der unteren und mittleren Troposphäre kaum von Null unterscheidbar sind, kann ich mir persönlich nicht vorstellen, wie das „Mainstream“-Denken korrekt sein kann, wenn man die Prozesse berücksichtigt, die die Gesamt-Reaktion der Atmosphäre auf steigende CO<sub>2</sub>-Niveaus betrachtet. Wie in der Begutachtung von Thorne et al. erwähnt, würde ein Fehlen einer troposphärischen Erwärmung „fundamentale und weit reichende Implikationen für das Verständnis des Klimasystems“ haben.

Die Abbildungen zeigen nicht wirklich die Klarheit unserer Ergebnisse; man muss die Zahlen sehen. Tabelle 7 fasst die wesentlichen Testzahlen zusammen, auf denen unsere Schlussfolgerungen beruhen:

Table 7. Results of hypothesis tests using VF statistic

Test	Null hypothesis	Test score	Bootstrap p-value
<b>Trend in</b>			
No level shift			
Hadley LT (0.135)	Trend = 0	260.4***	<0.001
RICH LT (0.126)	Trend = 0	212.0***	<0.001
RAOBCORE LT (0.149)	Trend = 0	281.4***	<0.001
Hadley MT (0.086)	Trend = 0	117.0***	<0.004
RICH MT (0.090)	Trend = 0	111.4***	<0.005
RAOBCORE MT(0.117)	Trend = 0	197.4***	<0.001
LT average	Models = observed	97.2***	<0.007
MT average	Models = observed	167.0***	<0.002
With level shift at date (assumed known): December 1977			
Hadley LT (0.064)	Trend = 0	13.2	0.273
RICH LT (0.093)	Trend = 0	22.6	0.160
RAOBCORE LT (0.065)	Trend = 0	11.2	0.311
Hadley MT (-0.001)	Trend = 0	0.1	0.925
RICH MT (0.048)	Trend = 0	5.1	0.485
RAOBCORE MT (0.042)	Trend = 0	3.4	0.563
LT average	Models = observed	354.5***	<0.001
MT average	Models = observed	685.9***	<0.001
<b>Shift term in</b>			
LT average	Avg obs shift term = 0	19.4	0.189
MT average	Avg obs shift term = 0	30.8	0.107
<b>Trend in</b>			
With level shift at date assumed unknown			
LT average	Models = observed	495.6***	<0.001
MT average	Models = observed	937.7***	<0.001
<b>Shift term in</b>			
LT average	Avg obs shift term = 0	180.4*	0.061*
MT average	Avg obs shift term = 0	259.8**	<0.025

Sample period (monthly): January 1958 to December 2012. The bootstrap p-value is computed using the method described in Section 3 using 10,000 bootstrap replications. VF critical values: Without level shift, 27.14 (10%, denoted \*), 41.53 (5%, denoted \*\*), 83.96 (1%, denoted \*\*\*). With level shift at known date, 32.35 (10%, denoted \*), 49.40 (5%, denoted \*\*), 99.98 (1%, denoted \*\*\*). With level shift at unknown date, 150.99 (10%, denoted \*), 188.68 (5%, denoted \*\*), 279.85 (1%, denoted \*\*\*).  
 LT, lower troposphere; MT, mid-troposphere; RAOBCORE, Radiosonde Observation Bias Correction using Reanalysis; RICH, Radiosonde Innovation Composite Homogenization.  
 \*Interpolated using critical values in Table 2.

Die erste Spalte zeigt die getesteten Datenreihen. Die zweite Spalte listet die Null-Hypothese. Die dritte Spalte gibt die VF-Werte, aber man beachte, dass diese Statistik einer Nicht-Standard-Verteilung folgt, und kritische Werte müssen entweder simuliert oder urgeladen [bootstrapped?] werden. Die letzte Spalte zeigt den p-Wert.

Der erste Block zeigt Ergebnisse ohne Verschiebungs-Term, der in den geschätzten Modellen enthalten ist. Die ersten 6 Zeilen zeigen die 3 Trends der unteren Troposphäre (mit dem Trendkoeffizienten in °C pro Jahrzehnt in Klammern), gefolgt von den 3 Trends der mittleren Troposphäre. Der Test eines Null-Trends in jedem einzelnen Fall widerspricht sich (in diesem Falle ist der kritische 5%-Wert 41,53 und 1% 83,96). Die nächsten beiden Zeilen testen den mittleren Modelltrend gegen den mittleren beobachteten Trend. Auch diese widersprechen sich und ignorieren sogar den Verschiebungs-Term.

Der zweite Block wiederholt diese Ergebnisse mit einer Verschiebung im Dezember 1977. Hier erkennt man die dramatische Auswirkung der Kontrolle auf die Pazifische Klima-Verschiebung. Die VF-Werte für den Null-Trend kollabieren und die p-Werte steigen rasant. Mit anderen Worten: die Trends verschwinden und werden praktisch und statistisch insignifikant. Die Äquivalenz zwischen Modell- und Beobachtungs-Trend widersprechen sich [reject] auch wieder stark.

Die nächsten beiden Zeilen zeigen, dass die Verschiebungs-Terme in diesem Falle nicht signifikant sind. Teilweise liegt dies daran, dass die Verschiebungs-Terme schwerer zu identifizieren sind als die Trends in den Daten der Zeitreihen.

Der letzte Abschnitt der Studie beschreibt die Ergebnisse, wenn wir einen die Daten untersuchenden Algorithmus verwenden, um das Datum der Verschiebung zu

identifizieren. Damit adjustieren wir die kritischen Werte, die im Suchvorgang in Betracht kommen. Erneut widersprechen sich die Trend-Äquivalents-Tests zwischen den Modellen und den Beobachtungen erheblich, und dieses Mal wird der Verschiebungs-Term signifikant oder schwach signifikant.

Wir listen auch die Ergebnisse Modell für Modell in der Studie auf. Einige GCMs widersprechen sich nicht individuell, einige tun das immer, und bei einigen kommt es auf die Spezifizierung an. Fügt man einen Verschiebungs-Term hinzu, nehmen die VF-Testzahlen zu, aber auch die kritischen Werte, so dass dies nicht immer zu kleineren p-Werten führt.

### **Warum wird das Ensemble-Mittel und dessen Verteilung getestet?**

Das IPCC sagt auf Seite 772, dass die Beobachtungen eher getestet werden sollten gegen die Spanne des gesamten Ensembles der Modellläufe als gegen das Mittel. In einer Hinsicht tun wir das: Die Ergebnisse Modell für Modell sind in der Studie gelistet. Aber wir stellen dieses Verfahren auch in Frage, weil die Bandbreite des Ensembles beliebig erweitert werden kann einfach durch das Hinzufügen weiterer Läufe mit alternativen Parametrisierungen. Die Durchführung eines Tests, dass Daten erfordert, die außerhalb dieser Bandbreite liegen, die man so weit machen kann wie man will, macht die Theorie im Endeffekt unfalsifizierbar. Auch redet das IPCC (und alle anderen) über Klimamodelle als eine *Gruppe* oder ein methodisches Genre. Aber das Genre wird in keiner Weise durch die Beobachtung unterstützt, dass sich ein einzelnes außerhalb liegendes GCM mit den Beobachtungen überlappt, während alle anderen dazu tendieren, weit entfernt von den Beobachtungen zu liegen. Klimamodelle sind wie alle Modelle (einschließlich von Wirtschaftsmodellen) ultimativ groß und konstruieren numerische Hypothesen: *Falls* die Welt derartig funktioniert und *falls* die Input-Variablen sich so und so ändern, *dann* werden sich die Output-Variablen so und so ändern. Um die „Modelle“ kollektiv zu verteidigen, d. h. sie als einen untereinander in Beziehung stehenden Satz physischer Hypothesen zu betrachten, wie die Welt funktioniert, erfordert einen Test der Maßzahl ihrer zentralen Tendenz, wofür wir hier das Ensemble-Mittel halten.

In gleicher Weise hat [James Annan](#) die MMH20-Ergebnisse verworfen und gesagt, dass es bedeutungslos sei, das Modell-Mittel mit den Daten zu vergleichen. Sein Argument lautete, dass einige *Modelle* sich auch widersprechen, wenn man sie mit dem mittleren Modell vergleicht, und es macht keinen Sinn zu sagen, dass die Modelle inkonsistent mit den Modellen sind. Daher ist der ganze Test falsch. Aber dies ist ein Gedankensprung. Sogar falls ein oder mehrere individuelle Modelle so weit außerhalb liegen, dass sie dem Modell-Mittel widersprechen, negiert dies nicht das Ergebnis, dass sich das mittlere Modell mit den beobachteten Daten widerspricht. Falls die zentrale Tendenz der Modelle signifikant weit entfernt von der Realität liegt, ist die zentrale Tendenz der Modelle falsch. Punkt! Dass das einzige Modell, das zuverlässig nicht den Daten widerspricht (in diesem Falle GFDL 2.1), ein Ausreißer unter den GCMs ist, verstärkt nur die Beweislage, dass die Modelle systematisch verzerrt sind.

Es gibt ein subtileres Problem in Annans Rhetorik, wenn er sagt: „Argumentiert wirklich einer ernsthaft auf der Grundlage, dass die Modelle



nicht einmal ihr eigenes Verhalten vorhersagen?“ Indem er dies sagt, glossiert er die Unterscheidung zwischen einem einzelnen Ausreißer-Modell und „den Modellen“ als *Gruppe*, nämlich als ein methodisches Genre. Sich auf „die Modelle“ als eine Einheit zu beziehen heißt, die Vermutung eines gemeinsamen Satzes von Hypothesen herbeizurufen darüber, wie das Klima funktioniert. Modellierer weisen oft darauf hin, dass die GCMs auf bekannter Physik basieren. Vermutlich sind die Gesetze der Physik für alle gleich, auch für die Modellierer. Einige klimatische Prozesse sind mit Grundprinzipien nicht lösbar und müssen als empirische Annäherungen und Parametrisierungen präsentiert werden. Folglich gibt es Differenzen unter den spezifischen Modellen und spezifischen Modellläufen. Das Ensemble-Mittel des Modells (und deren Standardabweichung) scheint mir der beste Weg, die gemeinsame zentrale Tendenz der Modelle zu charakterisieren. Ist ein Modell ein Ausreißer vom Mittel, ist es immer weniger repräsentativ für die Modelle allgemein. Also scheint das Mittel unter den Modellen genauso gut ein Weg zu sein wie alle anderen, um ihre zentrale Tendenz zu repräsentieren, und tatsächlich wäre es schwierig, irgendeine Alternative zu finden.

### **Unter dem Strich**

Während der 55 Jahre von 1958 bis 2012 haben die Klimamodelle nicht nur die beobachtete Erwärmung in der tropischen Troposphäre signifikant überbewertet, sondern diese auch auf fundamental andere Weise repräsentiert als beobachtet. Die Modelle repräsentieren das Intervall als einen glatten Aufwärtstrend ohne ruckartigen Änderung. Den Beobachtungen zufolge konzentriert sich die gesamte Erwärmung jedoch auf eine einzige ruckartige Veränderung Ende der siebziger Jahre, was mit einem bekannten Ereignis zusammenfällt (der Pazifischen Klima-Verschiebung). Davor und danach zeigt sich in den Beobachtungen kein signifikanter Trend. Meiner Ansicht nach ist die einfachste und wahrscheinlichste Interpretation dieser Ergebnisse, dass die Klimamodelle im Mittel versagen zu zeigen, welche Gründe diese ruckartige Veränderung Ende der siebziger Jahre hatte, und sie haben die Gesamt-Reaktion der Atmosphäre auf steigende CO<sub>2</sub>-Niveaus signifikant überschätzt.

Link:

<http://climateaudit.org/2014/07/24/new-paper-by-mckitrick-and-vogelsang-comparing-models-and-observations-in-the-tropical-troposphere/>

Übersetzt von [Chris Frey](#) EIKE