

Die schwierige, nimmer endende Fehlerdiskussion! Oder wie gut sind historische meteorologische Temperaturdaten? Wie wurden sie ermittelt, wie zusammengestellt und wie verarbeitet? (Teil 1)

Eine Übersicht in 3 Teilen von Michael Limburg

: "...Using data from all these poor stations, the U.N.'s Intergovernmental Panel on Climate Change estimates an average global 0.64°C temperature rise in the past 50 years, "most" of which the IPCC says is due to humans. Yet the margin of error for the stations is at least three times larger than the estimated warming" BEST Studienleiter Richard Muller 21.10.2011 Wallstreet Journal ¹

Mein Beitrag über zwei kleine Anfragen der AfD ([hier](#)) zur ungelösten Problematik der absoluten globalen Mitteltemperatur in jüngerer historischer Zeit – die Pariser Klimaübereinkunft bezieht sich sogar auf „Werte vorindustrieller Zeit“ ohne diese auch nur ansatzweise zu beziffern- hat eine große, manchmal hitzige, Diskussion ausgelöst. Mit aktuell 127 Kommentaren und über 3000 Aufrufen (Stand 27.4.20 12:00 Uhr) gehört er sicher zur Klasse der Meistgelesenen. Soweit so gut.

Die z.T. hitzigen Kommentare zeigen m.E.n. aber auch, dass viele Leser mit der zugrunde liegenden, manchmal komplizierten, meteorologischen Messtechnik nicht so vertraut sind, wie sie der Artikel vielleicht voraussetzt. Das hat zur Folge, dass sie nicht auf Anhieb verstehen, warum der Artikel Versäumnisse – gar Falschinformationen- der Bundesregierung – beanstandet.

Sie vertrauen dabei auf ihre eigenen, oft in wissenschaftlicher Arbeit gemachten Erfahrungen aus eigener Anwendung der Messtechnik und ihrer Behandlung von Messdaten. Diese stimmen aber – so meine eigene Erfahrung – wenig bis nicht mit den hier angewendeten Methoden und Verfahren, insbesondere bei der Erstellung und Bewertung von Zeitreihen, überein. Und sind deshalb in ihrer Schlichtheit selten auf diese Messdaten anzuwenden.

Dies voraus geschickt will ich die Leser jetzt ein wenig in die meteorologische Messtechnik von Temperaturdaten und deren Auswertung einführen. Dabei sollen die alten wissenschaftlichen Grundsätze gelten, klassifiziere erst und vergleiche dann, aber vergleiche immer nur Vergleichbares.

Zunächst schauen wir uns an, wie, wo, wann und wie oft die historischen Temperaturdaten gemessen und vorverarbeitet wurden.

Dann werden wir uns die Bildung von Anomalien aus diesen bereits vorverarbeiteten Daten und die sich daraus ergebenden Besonderheiten anschauen und zum Schluss die Bildung von Zeitreihen, und wie diese auf Fehlerkompensation bzw. Nichtkompensation Einfluss nehmen.

Was nicht behandelt wird!

¹ <https://www.wsj.com/articles/SB10001424052970204422404576594872796327348#printMode>

Gänzlich außen vor lassen wir hier den magischen Trick der Klimaforscher, aus bekannten absoluten Unsicherheiten der Messwerte durch Bildung von Anomalien aus vielen Temperaturdaten statistische Wahrscheinlichkeiten werden zu lassen, die sich mit der bekannten Wahrscheinlichkeitsverteilung der Normalverteilung über ein oder zwei Sigma berechnen ließen. Darüber hat Kip Hansen eine sehr eindrückliche Kritik geschrieben. Sie finden Sie hier unter dem Titel „[Der „Trick“ anomaler Temperatur-Anomalien](#)“ Weil es das Thema viel zu weit ausdehnen würde, lassen wir auch die merkwürdige Behauptung des Stammvaters aller Klimaängste, James Hansen außen vor, der in einem paper zusammen mit seinem Kollegen Lebedeff 1987² behauptet, es wäre möglich auch mit einer sehr geringen Zahl von ca. 80 Gitterzellen gleicher Fläche von 2500 x 2500 km die Temperaturentwicklung der gesamten Erde zu bestimmen. Und dies obwohl in den meisten Zellen nur sehr wenige, in anderen überhaupt keine Messstationen vorhanden waren. Dieser Artikel war die Grundlage und ist es noch heute, für die Behauptung, dass die Anomalienbildung und deren Korrelation über große Flächen hinweg aussagefähig für die Temperaturentwicklung sehr große Flächen seien. Diese Idee, die jeder Korrelations- und Fehlerrechnung widerspricht, wird hier unter dem Titel [Von Korrelationen, Trends bei Anomalien!](#) besprochen Und ebenfalls lassen wir die Erkenntnis von Dresdner Statistikern außen vor, die nachwiesen, dass selbst echt zufällige Fehler, sofern sie sich in Daten befinden, die für autokorreliert sind und in entsprechenden Zeitreihen verwendet werden, systematische Komponenten in der Größe von einigen Zehntel Grad entwickelten. Details dazu finden Sie im Artikel mit dem Titel „[Konfidenz-Intervalle* für zeitliche Mittelwerte bei langfristigen Korrelationen: eine Fallstudie zu Anomalien der Temperatur auf der Erde](#)“

Am Schluss der Vorbemerkung erlaube ich mir noch den Hinweis auf die Tatsache, dass mit der Bildung von raum- und zeitüberspannenden Durchschnittswerten, die darin enthaltenen Temperaturwerte ihre physikalische Bedeutung als Temperatur verlieren. Es entsteht stattdessen irgendeine Art von Index. Er ist für statistische Zwecke begrenzt brauchbar, wenn man die Randbedingungen beachtet, aber keine physikalische energetisch relevante Größe mehr. Wer dazu genaueres wissen möchte, der sei auf den schöne Beitrag von C. Essex und R. McKittrick von 2006 verwiesen mit dem schönen Titel „[Does A Global Mean Temperature exist](#)“³

[Wie, wo und wann wurden die historischen meteorologischen Messdaten gewonnen?](#)

1. Thermometer

Am Anfang stand und steht das Thermometer. Dabei ist es zunächst egal, ob es ein mechanisches oder ein elektronisches ist, obwohl jede Art einen erheblichen Einfluss auf die (systematischen) Fehler hat, natürlich auch auf die erzielbare Genauigkeit.

2. Messumgebung

² HANSEN, JAMES LEBEDEFF, SERGEI (1987) Global Trends of Measured Surface Air Temperature. JOURNAL OF GEOPHYSICAL RESEARCH 92:13345 13372.

³ Essex, C, Mc Kittrick, R. Andresen, B. (2006) Does A Global Temperature Exist? Non-Equilibrium Thermodynamics

Doch mindestens ebenso wichtig ist die Messumgebung. Diese wird bestimmt durch das Gehäuse (Messhütte), welches das Thermometer vor verfälschenden Umgebungseinflüssen, wie Strahlung, Wind oder Niederschlag aber auch Beschädigung schützen soll, aber auch in starkem Maße von der unmittelbaren Umgebung. Um vergleichbare Ergebnisse wenigstens annähernd zu erreichen, muss dies bestimmten überall vergleichbaren Standards genügen und sorgfältig gebaut und gepflegt werden.

3. Messregime

Und auch das Messregime ist von entscheidender Bedeutung, nämlich wann und wie oft, von wem und wie das Thermometer abgelesen wird.

4. Die geografische und zeitliche Abdeckung

Und da das Wetter- wie das Klimageschehen nicht an nationalen Grenzen haltmacht, muss die geografische Abdeckung mit Messstationen gesichert sein, und das sowohl nach Zahl als auch nach Ort und Zeit. Alle drei Anforderungen werden erst seit dem Aufkommen der Satelliten hinreichend erfüllt. Dabei tauscht man die Vorzüge der Satellitenmessung gegen einen enormen Nachteil ein, denn man musste einen unangenehmen Kompromiss akzeptieren. Es ist die sehr schlechte vertikale Auflösung. Eine Satellitenmessung ist daher immer der Mittelwert von Temperaturänderung über eine Distanz von ca. 5000 m. (z.B. der lower troposphere)

5. Was geschieht mit den meteorologischen Messdaten? Die Weiterverarbeitung!

Bis zur „Erfindung“ der Idee von der Notwendigkeit einer weltweiten Mitteltemperatur, gab es nur wenige Meteorologen, die die vorhandenen Messdaten nutzen wollten, um evtl. aus ihnen eine weltweite Mitteltemperatur zu extrahieren. Man begnügte sich damit – nachdem solche großartigen Wissenschaftler der Meteorologie wie Köppen die Definition von Klimazonen etabliert hatten- für diese Klimazonen oder Regionen Mitteltemperaturen und deren Jahresgänge zu berechnen. Normalerweise dienten die meteorologischen Messdaten aber vor allem der örtlichen Wettervorhersage und bez. ihrer Langzeit-Analyse auch dazu ggf. lokale Klimaänderungen im Nachhinein zu erkennen.

Anmerkung: Achten Sie mal auf die örtlichen Wetterberichte, dort wird man ihnen so gut wie nie die erwartete Durchschnittstemperatur der nächsten Tage nennen, sondern ausschließlich deren Maxima und Minimawerte.

Dazu erstellte man Zeitreihen (hier Temperatur-Ganglinien) entweder aus absoluten Werten oder von Anomalien je nach Bedarf, deren erzielbare Genauigkeit für die gewünschten Zwecke völlig ausreichte.

Zu 1. Thermometer



Abbildung 1 zeigt eine klassische, deutsche Wetterhütte (erweiterter Stevenson Screen) und (rechts) ein darin befindliches modernes Thermometerensemble gezeigt, wie es immer noch in vielfältiger Verwendung ist. Die senkrechten Thermometer sind (links) das Assmansche Aspirationsthermometer zur Temperaturmessung, rechts daneben ein befeuchtetes und belüftetes Psychrometer zur Messung der Luftfeuchte. Quer darunter befinden sich zwei liegende Thermometer zur Messung von Max. und Min. Temperaturen. Die runde Plastiktrommel enthält den Federmotor mit Ventilator zur Erzeugung des Luftstromes mit \varnothing 2,5 m/s vor und während der Messung. Im März 2009 vom Autor im Meteorologischen Institut der Freien Universität Berlin aufgenommen

Obwohl schon Galileo erste Versuche unternahm mit einem Thermometer meteorologische Messungen zu machen, und sich danach ein ziemlicher Wildwuchs an solchen Thermometer ausbreitete, weil viele Landesfürsten die aufkommende Meteorologie aus naheliegenden Gründen für äußerst wichtig hielten, kann man erst ab 1892 von einem Standardinstrument für meteorologische Zwecke sprechen. Es handelt sich dabei um das zwangsbelüftete Assmansche (Quecksilber gefüllte) Aspirationsthermometer (linkes, senkrechtes Thermometer). Die zuvor verwendeten tlw. mit Alkohol gefüllten Thermometer mit mehr als ± 1 Kelvin Gesamt-Unsicherheit wurden nach und nach weltweit ersetzt. In den angelsächsisch orientierten tlw. auch regierten Ländern setzte sich hingegen oft zusätzlich das Max-Min-Thermometer durch.

Doch allein die Wahl des Thermometers, sein jeweiliger Zustand, seine Montage, sein Betrieb und seine Ablesung sorgen für eine ganze Reihe von zufälligen wie systematischen Fehlern. Diese wurden vielfach sorgfältig untersucht, aber nur in geringer Anzahl auch quantitativ bestimmt und damit überhaupt erst theoretisch korrigierbar.

Zu 2. Messumgebung

Für die Messumgebung war von Anfang an klar, dass dieses zuallerst eine Art Schutzhütte sein sollte, von denen besonders im 19 Jahrhundert eine ganze Reihe von Konstruktionen erprobt wurden, wobei sich letztendlich die erweiterte englische Hütte als Standard durchsetzte. Doch bis dahin waren die (wenigen) Messergebnisse praktisch unvergleichbar, weil die Hütte auf ihre eigene Art die Messergebnisse beeinflusste.

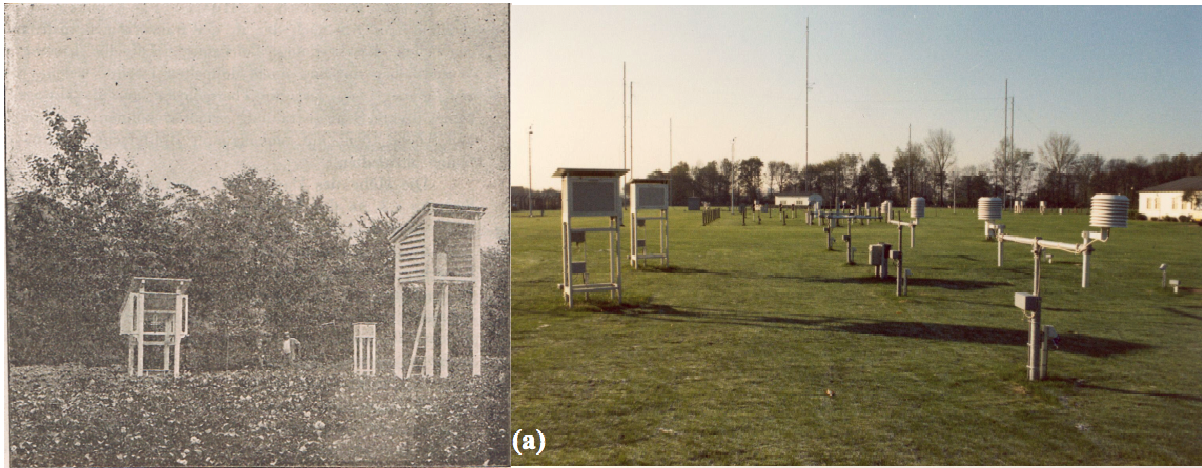


Abbildung 2 Darstellung der zum Ende des 19. Jahrhunderts und bis heute verwendeten Hüttentypen bzw. Screens von links: französische Hütte, daneben (kaum erkennbar) Mitarbeiter zum Größenvergleich, daneben (Original) Stevenson Hütte, dann Wildsche Hütte, dann mehrere erweiterte Stevenson Hütten und moderne runde kleine Screens. (Quellen. Linkes Bild [Sprung, 1890] rechtes Bild: KNMI Messfeld von van der Meulen für die Zeit von 1989-1995)

*Praktisch unkorrigiert, obwohl seit langem bekannt, ist der systematische Fehler, der allein durch die Wärmeträgheit der Hütte entsteht, und der ca + 1 bis 2 Kelvin zu jedem Ergebnis der eigentlich gesuchten Außentemperatur hinzuaddiert. Praktischerweise definieren jedoch die Meteorologen **die Temperatur in der Hütte**⁴, als die gesuchte Außentemperatur. Woraus geschlossen werden muss, dass jede durch diese Messungen gewonnene absolute Temperatur im Mittel im 1 bis 1,5 Kelvin gegenüber der wahren Außentemperatur zu hoch sein muss.*

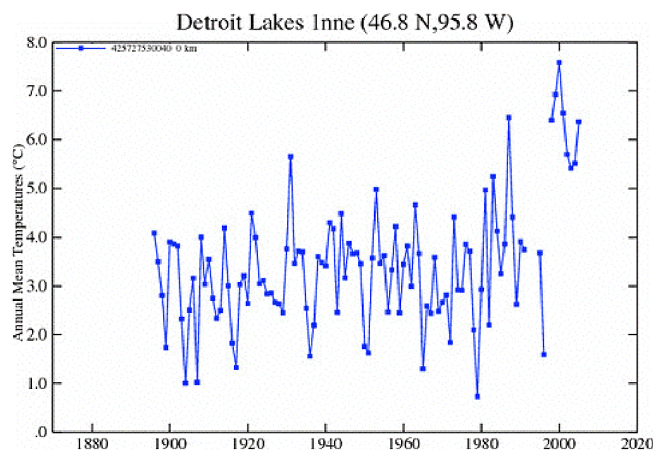


Abbildung 3(links): Englische Hütte und moderne Messstation am Messort Detroit Lakes USA, kurz vorher aufgenommen und ins Netz gestellt am 25.07.07 von A. Watts; (rechts) Temperaturverlauf aus der GISS Datenbank zur Messstelle Detroit Lakes USA

Beispiel für den Zustand und Umgebung einer meteorologische US Messstation sowie deren Zeitreihe ihrer Jahresmittelwerte, die vom Projektteam untersucht wurde. Man beachte den Sprung ab etwa dem Jahr 2000 und die Entlüftungen der Klimaanlage in unmittelbarer Nähe.

Nicht weniger Einfluss hat die unmittelbare Umgebung in der die Messhütte.

Zur Klärung dieses Sachverhaltes hat sich der Meteorologe und Betreiber des inzwischen weltweit erfolgreichsten Klimarealisten- Blogs Wattsupwiththat Antony Watts große Meriten erworben. Er rief 2007 das SurfaceStationorg. Netzwerk ins Leben, das 2009 erfolgreich mit

⁴ Weischet [Weischet, 1995] S115 - S 109 ff

einer Veröffentlichung⁵ abgeschlossen wurde. Freiwillige im ganzen Land untersuchten über 850 der 1200 meteorologischen Messstationen, die von der US-Wetter- und Ozeanbehörde NOAA betrieben wurden.

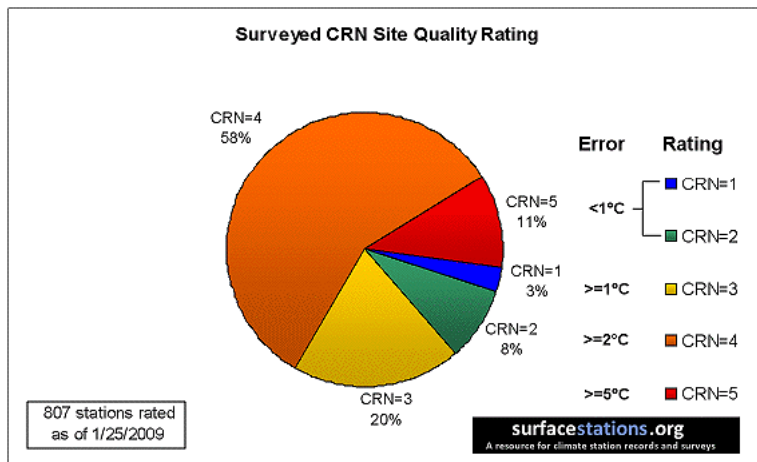


Abbildung 4 Untersuchung von bisher (Ende Februar 2009) 854 von 1221 offiziell verwendeten Messstationen der unterschiedlichsten Art. Man beachte, dass nur 3 % aller Stationen dem Standard voll entsprechen, weitere 8 % können evtl. noch als dem Standard genügend zugeordnet werden. Die restlichen 89 % zeigen jedoch potentielle Abweichungen nach den Vorgaben der CRN von mindestens 1 °C (20 %) über 2 °C (58%) bis 5 °C (11%).

Auswertung der Überprüfung von 807 US Messstationen in Bezug auf die offizielle Einstufung ihrer Qualität und Umgebung (CRN).

Das Ergebnis war erschreckend. Der Zustand der Messstationen und damit die Messergebnisse waren miserabel. 89 % von ihnen lagen oberhalb der Fehlerklasse 2, was bedeutet, dass sie > 2 °C bis > 5 °C fehlerhafte Ergebnisse zeigten. Nur 3 % lagen gleich oder unterhalb von 1 °C. Und noch schlimmer war, dass diese Stationen für die gesamte Welt sozusagen als Goldstandard in Punkto Qualität galten. D.h. die vom IPCC und ihren Zuarbeitern immer wieder herangezogenen Daten aus den anderen Kontinenten waren mit hoher Wahrscheinlichkeit noch schlechter.

Es würde hier zu weit gehen, die vielfältigen Fehlerquellen allein der landbasierten Messstationen aufzulisten, daher beschränke ich mich auf zwei kurze Ergänzungen.

1. Selbst bei identischer Konstruktion von Hütte und Messinstrument, ergeben sich große Unterschiede allein durch das Material und seinen Anstrich, bzw. dessen Veränderung über der Zeit, z.B. durch Alterung
2. Sind die landbasierten Messungen schon oft mit mehr (systematischen) Fehlern behaftet als ein Hund Flöhe hat, so ist das bei den seegestützten Messungen noch weit schlimmer. Und weil das Land nur 29 % der Erdoberfläche bedeckt, aber 71 % der Erde Ozeane sind, ist dieser Umstand eigentlich noch schlimmer. Denn dort sind die Messungen wirklich nur punktuell, d.h. äußerst selten sowohl in räumlicher als auch in zeitlicher Hinsicht. Ein wenig wird dieser Mangel gelindert durch die Tatsache, dass statt der eigentlich gesuchten Lufttemperatur, stattdessen, als Proxy,

⁵ Is the U.S. Surface Temperature Record Reliable? By Antony Watts SurfaceStations.org ISBN 13: 978-1-934791-29-5

die Temperatur der Seewasser Oberfläche (SST Sea Water Surface Temperature) verwendet wird. Wegen der großen Wärmeträgheit und guten -Leitfähigkeit des Wassers deckt ein Messpunkt im Vergleich zum Land viel größere Gebiete ab.

Wer Interesse daran hat, diese Fehler etwas genauer behandelt zu sehen, der sei auf das Video „Wie glaubwürdig sind die Zeitreihen historischer Klimadaten“ ([hier](#)) verwiesen, dass einige der wichtigsten Aspekte zu diesen Fehlern aufzeigt.

Ende Teil 1

Teil 2

„...sie haben vollkommen die systematischen Messfehler ignoriert, welche die Temperaturaufzeichnung durchziehen und diese ungeeignet macht, um irgendetwas über das historische Klima daraus abzulesen, siehe [hier](#), [hier](#) und [hier](#).“ (Pat Frank 9.12.2019⁶)

Zu 3 Messregime

Da es praktisch undurchführbar war die gemessene Temperatur in kurzen Zeitabständen abzulesen, wünschenswert wären bspw. alle 5 Minuten gewesen, definierte man entweder wenige feste Tageszeiten, an denen abgelesen (und notiert) wurde, dies wurde insbesondere in Zentraleuropa – aber zu verschiedenen Zeiten und auch unterschiedlich oft- so gemacht, oder man las nur die Max. und die Min-Werte von speziellen Thermometern ab, und war dadurch zeitlich nicht so gebunden. Man handelte sich aber damit die Unkenntnis ein, nicht genau zu wissen, wann am Tag der Max. bzw. Min. Wert aufgetreten war. Das führte zu einem speziellen, aber kräftigen Fehler bei der Auswertung, der sogar einen eigenen Namen bekam, der Time of Observation Error (TOBs). Vor allem die praktischen Briten und mit ihnen dann alle ihre Kolonien und Einflussgebiete benutzten bis zum Ende des vergangenen Jahrhunderts diese Methode.

⁶ Zitate in deutscher Übersetzung von Pat Frank Quelle: <https://www.eike-klima-energie.eu/2019/09/09/fehler-fortpflanzung-und-die-zuverlaessigkeit-von-projektionen-der-globalen-lufttemperatur/>

<u>Mittelwertbildung nach Mannheimer Stunden</u>	
1.	$\bar{T}_a = \frac{T_7 + T_{14} + 2T_{21}}{4}$
mit T_7, T_{14}, T_{21} als Ablesungen um 7,14, und 21 Uhr MOZ	
<u>Mittelwertbildung aus den Extremwerten</u>	
2.	$\bar{T}_a = \frac{T_{\min} + T_{\max}}{2}$
<u>Mittelwertbildung aus den Sechsstundenwerten</u>	
3.	$\bar{T}_a = \frac{T_0 + T_6 + T_{12} + T_{18}}{4}$

Abbildung 5 Kleine Auswahl aus historisch vielfach verwendeten Mittelwertbildungsalgorithmen

Aus diesen punktuellen Messwerten wurden dann arithmetische Mittelwerte gebildet, das sog. Tagesmittel. Diese (verschiedenen) Tagesmittel erlaubten nun Monatsgänge der Tagestemperaturen zu bilden. Sie wurden und werden dann weiter zu arithmetischen Monatsmitteln zusammengefasst, die es wiederum erlaubten Jahresgänge zu bilden, welche den Verlauf der Jahreszeiten anhand der Mittel-Temperatur zeigten. Diese können dann zu Zeitreihen für langjährige klimatologisch interessante Zeiten zusammengefasst werden, was auch geschieht.

Es geschieht aber noch etwas anderes mit diesen Daten, sie werden in „Anomalien“ umgerechnet, d.h. es werden Differenzen aus dem aktuellen Wert abzüglich eines Referenzwertes gebildet. Dies geschieht wegen der berechtigten Annahme, dass Anomalien die eigentlich gesuchten Veränderungen präziser und weiträumig vergleichbarer abbilden, als die der absoluten Temperaturen, die zusätzlich noch weiteren Einflüssen wie Höhenlagen, windhöflicher oder stiller Messort etc. ausgesetzt sind. Doch damit handelt man sich neue Fehlerquellen ein, von denen die wenigsten Klimaforscher etwas wissen, ja schlimmer noch, die wegen falscher, wenn auch plausibler Analogien, zu völlig falschen Ergebnissen führen müssen. Welche das im Einzelnen sind und was das bedeutet wird im Teil „Weiterverarbeitung“ ausführlich beleuchtet.

Zu 4. Die geografische und zeitliche Abdeckung

Sobald die so gewonnenen Daten etwas mehr zur Beobachtung und Deutung von großräumigem Wettergeschehen herangezogen werden sollen, was in der Natur der Sache liegt, denn Wettergeschehen ist großräumig, kommt es unmittelbar auf eine genügende räumliche wie zeitliche Abdeckung der in Frage kommende Gebiete an. Doch da setzen ebenso technische wie finanzielle Machbarkeit enge Grenzen. Errichtung und Betrieb von Wetterstationen ist nicht nur teuer, sondern verlangt neben der genügenden Verfügbarkeit der richtigen Messmittel, auch genügend personelle Kapazität, großes Know How und

Disziplin für ihren Betrieb. Und das permanent über lange, im politischen Maßstab sehr lange Zeiträume. All das war von Anfang an nicht selbstverständlich und musste erst langsam nach dem Entstehen reifen. So wundert es nicht, dass die Stationsdichte sich ausnahmslos in früh entwickelten Ländern konzentriert, während weite Teile der Welt bis Anfang der 50er Jahre praktisch ohne Messstationen blieben. Das gilt für Südamerika, Afrika und weite Teile Asiens, ebenso wie die mehr als doppelt so große Ozeanfläche dieser Welt.

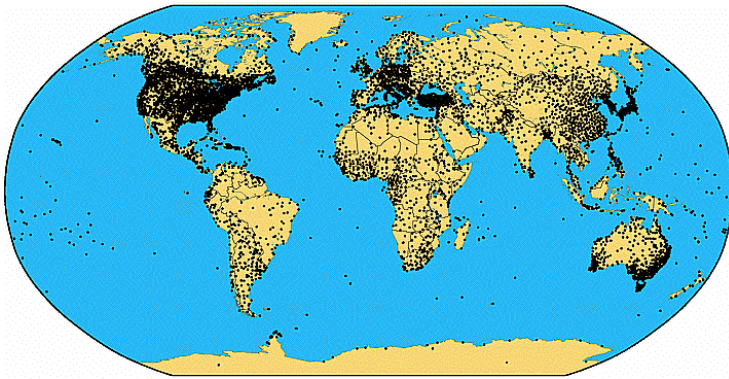


Abbildung 6 Aktueller Stand der im Global Historical Climate Networks GHCN der NOAA mit Monatsmittelwerten der Temperatur insgesamt zuliefernden 7.280 Stationen mit mindesten 10 Jahren kontinuierlicher Datenreihen. Für Stationen mit Max/Min Erfassung gibt es ähnliche Grafiken. Quelle (<http://www.ncdc.noaa.gov/oa/climate/ghcn-monthly/index.php> fig. 1). Man beachte jedoch, dass von GISS und CRU davon nur ungefähr 1/7 ausgewertet werden.

Quelle <https://climateaudit.org/2008/02/10/historical-station-distribution/>

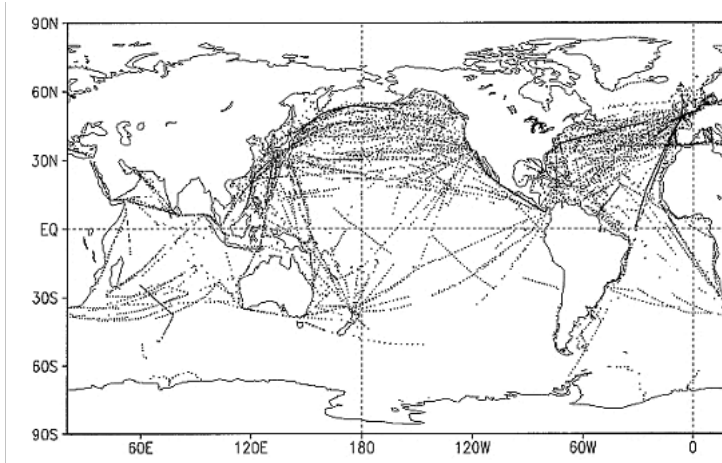


Abbildung 7 Darstellung der gemeldeten SST's für die Woche vom 2. bis 8. Januar 2000. Sie zeigt die dünne messtechnische Abdeckung der Wasserfläche der Erde. Frühere Abdeckungen der Wasserfläche sind mit großer Wahrscheinlichkeit noch deutlich schlechter gewesen. Aus [REYNOLDS, R. W. & RAYNER., 2002] Seite 1611

Zu 5. Die Weiterverarbeitung! Was geschieht mit den meteorologischen Messdaten?

Wie zuvor schon kurz angesprochen werden die Einzeldatensätze erst einer umfassenden Prüfung auf Konsistenz, Dauer, und Plausibilität, mit anschließender Weiterverarbeitung unterzogen. Zunächst bleibt es bei den einzelnen Stationen und ihren Daten. Je nach Verwendungszweck wird aus diesen erst Tagesmittel und daraus das Monatsmittel gebildet und von diesem eines als „Station Normal“ bezeichneten langjährigem (bspw. über 360 Monate also 30 Jahre) Mittel abgezogen. Die so entstehende Differenz wird als so gut wie

fehlerfrei betrachtet, weil – so die Annahme – die lokalen systematischen Fehler ja auch im „Station Normal“ vorhanden seien, und, wenn man beide Werte voneinander abzöge, sich der Fehler kompensierte (stellvertretend dafür [Brohan et al. 2006])⁷,⁸. Manchmal werden aber auch zunächst die Jahresmittel gebildet und diese vom „Station Normal“ abgezogen. Die Grundannahme aber, von der Fehler selbstkorrektur bei Bildung von Anomalien, bleibt aber dieselbe.

Und daher behaupten so gut wie alle Klimaforscher, auch Stefan Rahmstorf vom PIK unwidersprochen⁹, dass Anomalien viel genauer seien als die ihnen zugrunde liegenden Absolutwerte. Liest man seinen Text genauer, dann muss man feststellen, dass er sich mit der Materie nicht beschäftigt hat, sondern nur auf die Korrelation zwischen den Trends von absoluten Temperaturdaten verweist, wie sie schon von James Hansen für seine Schwarze Kunst verweist. Es bleibt ein bitterer Treppenwitz der Wissenschaftsgeschichte, dass Annahmen über Korrelationskoeffizienten herunter bis 0,33 (Hansens Grenzwerte) die Politik in ein nie dagewesenes Desaster treiben konnte. Wer Näheres dazu wissen möchte sei auf diesen Beitrag verwiesen: **Von Korrelationen, Trends bei Anomalien!**¹⁰

Doch diese Annahme ist – bei Zeitreihen von Anomalien - und das werde ich im Folgenden belegen, bis auf eine Ausnahme, immer falsch!

Anmerkung: Der folgende Text ist einem Artikel entnommen, den ich am 7. Februar 2011 bei EIKE einstellte, und der am 23.11.2018 aktualisiert wurde.

In der Klimatologie werden zur Bestimmung des Referenzwertes, die errechneten Monatsmittel oder Jahresmittel der gemessenen Temperatur einer Station, über den Zeitraum von 1961 bis 1990 aufaddiert und durch die Anzahl aller Monats- oder Jahresmittel dieses Zeitraumes geteilt. Der so enthaltene Wert entspricht einer Geraden gleicher „Temperatur“ über den Referenz-Zeitraum von 1961 bis 1990. (Es kann im Prinzip aber auch jeder andere Zeitraum sein.) Der zugehörige Referenz-Wert wird oft das **Stationsnormal** genannt. Folgerichtig heißen die Abweichungen vom (Stations)-Normal **Anomalien**.

Nun bildet man die Differenz aus allen absoluten Jahresmitteln T_x der Zeitreihe und dem **Stationsnormal** T_{ref} und ermittelt auf diese Weise eine neue Zeitreihe der **Anomalien** eben dieser Station. Dabei geht man von der Annahme aus, dass diese, wenn sie einen Fehler C enthalten, dieser auch im Stationsnormal vorhanden ist. Ferner

⁷ Brohan, P.K., J. J. Harris, I., Tett S. F. B.; & Jones, P. D. (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. HadCRUT 3 HadCRUT 3:1 to 35.

⁸ Brohan et. al geben auf Seite 6 den allgemeinen Glauben an die Fehlerfreiheit der Anomalien wider, in dem sie schreiben: „...*There will be a difference between the true mean monthly temperature (i.e. from 1 minute averages) and the average calculated by each station from measurements made less often; but this difference will also be present in the station normal and will cancel in the anomaly. So this doesn't contribute to the measurement error.*“ Im Folgenden beschreiben die Autoren wie die Anforderungen an die Bildung der „station normals“ also der Referenzwerte (lt. WMO der Mittelwert jeder Station über die Jahre 1961-1990) aufgeweicht (relaxed) wurden. Die durchgeführte Reduktion der Anforderungen zur Bildung des Referenzwertes (Station Normal) auf nur noch 15 Jahre und daraus 3 x 4 Jahre in jeder Dekade, erhöht den Fehler, der schon in der Verwendung des Station Normals steckt, weiter.

⁹ Verwirrspiel um die absolute globale Mitteltemperatur

<https://scilogs.spektrum.de/klimalounge/verwirrspiel-um-die-absolute-globale-mitteltemperatur/>

¹⁰ <https://www.eike-klima-energie.eu/2018/01/09/von-korrelationen-trends-bei-anomalien/>

unterstellt man, dass C eine konstante Größe hat. Bspw. Brohan und Jones¹¹ formulieren diese Annahme (in Bezug auf die Fehler durch unterschiedliche Mittelwertbildung) so:

„...There will be a difference between the true mean monthly temperature (i.e. from 1 minute averages) and the average calculated by each station from measurements made less often; but this difference will also be present in the station normal and will cancel in the anomaly.

Berechnet man jetzt die Anomalie $A_x = (T_x + C) - (T_{ref} + C)$, dann hebt sich C auf. Das ist schlichte Algebra. So weit so gut! Die große Frage ist nun, funktioniert das bei Zeitreihen auch so? Warum denn nicht, wird mancher vorschnell antworten. Die Algebra ist schließlich immer gleich! Oder?

Es zeigt sich sehr schnell, dass die Idee der Fehlerkompensation, die hinter dieser einfachen Methode steckt, ebenso plausibel wie falsch ist, bzw. sie stimmt nur innerhalb ganz enger Randbedingungen, die in der hier interessierenden Praxis nicht eintreten.

Ende Teil 2

Teil 3

Allerdings hält sich diese Annahme hartnäckig, wie ein ebenso frischer wie falscher Aufsatz von Stefan Rahmstorf in seinem Blog Scilog zeigt.

Zitat: Um in Paris zu bleiben: gehe ich auf dem Eiffelturm eine Treppenstufe hinauf, kann ich leicht mit einem Lineal messen, wieviel höher ich dann stehe – aber wie viele Meter über dem Erdboden ich insgesamt bin, lässt sich schon schwerer feststellen.

Stefan Rahmstorf am 12.2.18¹² zur Behauptung warum Anomalien genauer sind als die ihnen zugrunde liegenden Absolutwerte. Er hat bis heute nicht begriffen, dass er in diesem Beispiel Absolutwerte misst und diese als Anomalien ausgibt. Klima-anomalien werden jedoch nicht gemessen, sondern aus den Absolutwerten mit all ihren Fehlern berechnet. Und da gilt nun mal das Gesetz von der Fehlerfortpflanzung. Er betreibt also – bewusst oder unbewusst- Etikettenschwindel.

Anomalien und sprunghafter Fehler

Machen wir dazu die Probe aufs Exempel, indem wir einen Anomalienverlauf aus einem (künstlichen) sinusförmigen Temperaturverlauf von 1850 bis 2009 erzeugen. Blau sei der Anomalienverlauf ohne Fehler. Die Anomalie verläuft fast gradlinig, wie die blaue Trendlinie

¹¹ Brohan, P.K., J. J. Harris, I., Tett S. F. B.; & Jones, P. D. (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. HadCRUT 3 HadCRUT 3:1 to 35.

¹² Quelle: „Verwirrspiel um die absolute globale Mitteltemperatur“

<https://scilogs.spektrum.de/klimalounge/verwirrspiel-um-die-absolute-globale-mitteltemperatur/>

anzeigt. Ab 1940 wird nun ein Fehler von + 1 K eingefügt. Gemäß Rechenanweisung w.o. wird dann aus beiden Werten für alle Jahre (x) die Differenz gebildet. Ab 1940 hebt sich der Fehler auf. Beide Verläufe gehen ineinander über. Vor 1940 jedoch wird allein durch den Rechenprozess die Temperaturanomalie um 1 k abgesenkt. Wir sehen, die Annahme von vorhin gilt nur für die Zeit nach 1940. **Vorher nicht.** Lässt man Excel eine Trendgerade durch diese Werte legen, dann sehen wir, dass sich beide erheblich unterscheiden.

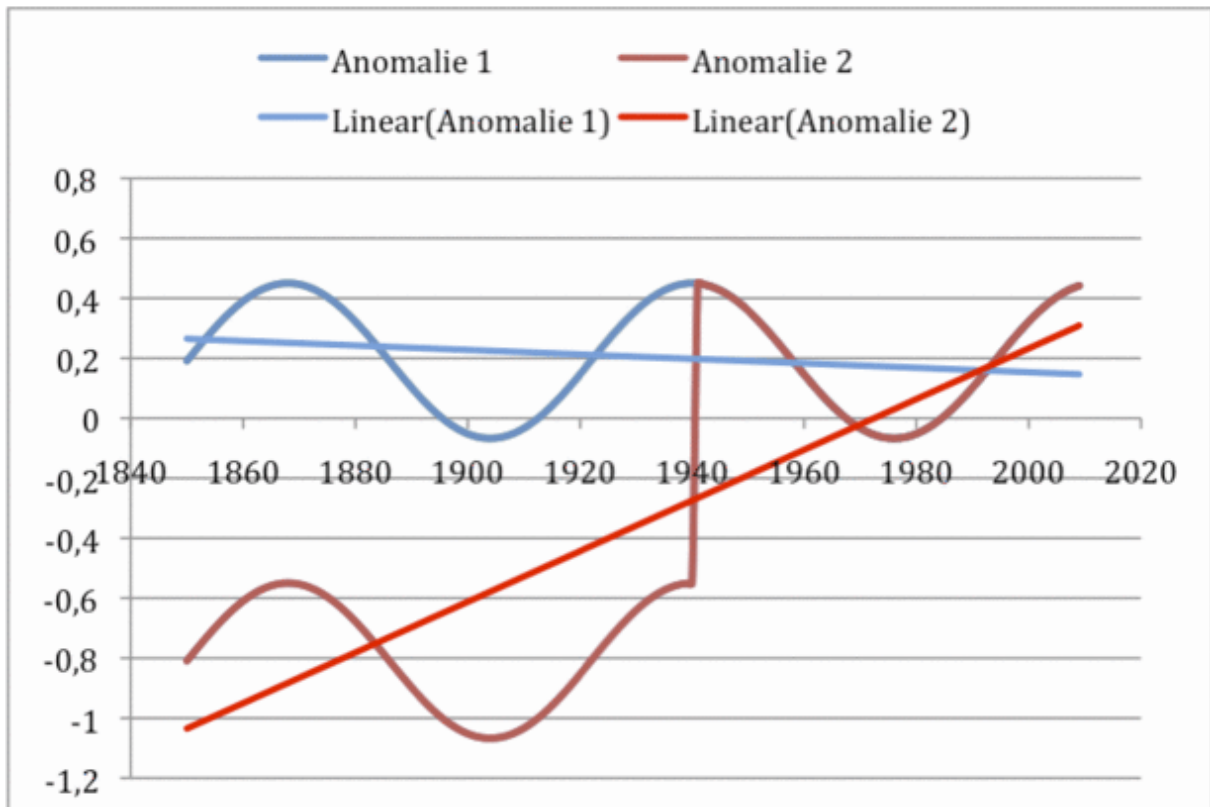


Abbildung 8 Ein systematischer Fehler von + 1 K tritt ab 1940 sprunghaft auf, weil z.B. die Station neu gestrichen wurde. Von dieser Veränderung sind das „Station Normal“ wie alle aktuellen Werte ab 1940 betroffen. Ab 1940 gleicht sich der Fehler aus, vorher aber nicht. Der Trend verschiebt sich.

Daran ändert sich auch nichts, wenn wir den Eintritt des Fehler auf später verlegen. Wie das nächste Bild zeigt. Dort werden die Kurven Anomalie 1 und 3 genannt.

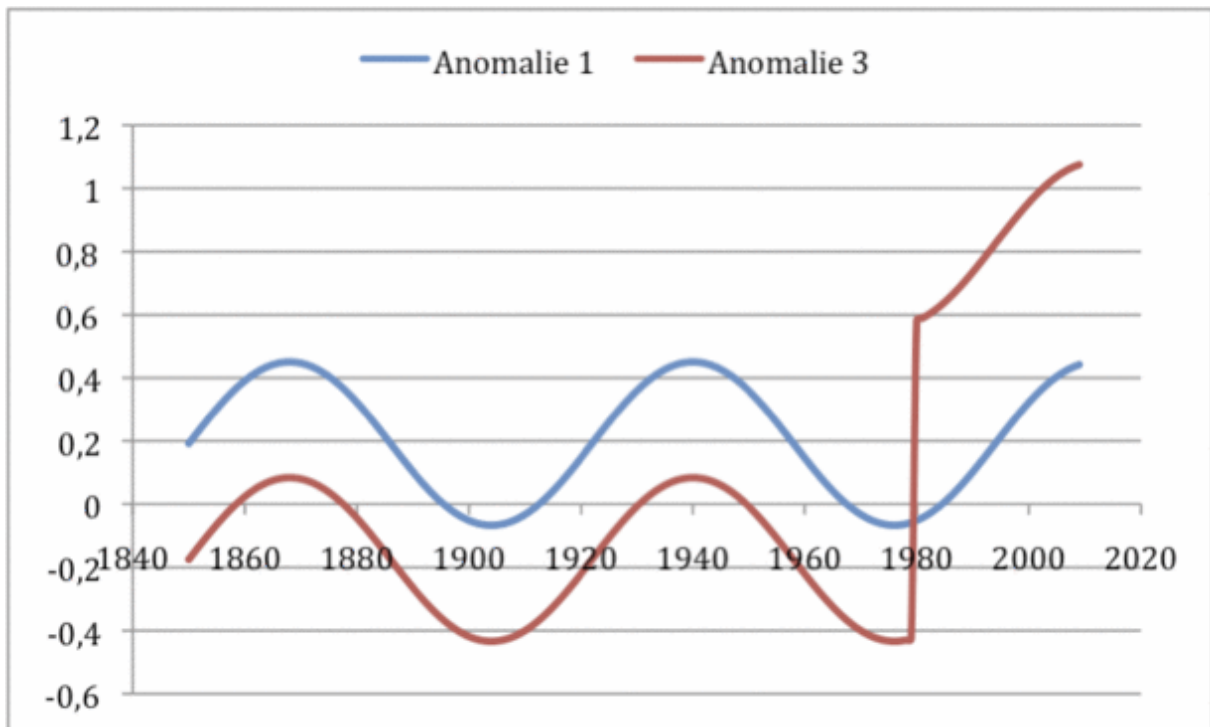


Abbildung 9 Ein systematischer Fehler von + 1 K tritt jetzt 1980 sprunghaft auf, weil z.B. die Station neu gestrichen wurde. Von dieser Veränderung sind das „Station Normal“ zur Hälfte und alle aktuellen Werte ab 1980 betroffen. Der Fehler gleicht sich nicht mehr aus, weder vorher noch nachher. Der Trend verschiebt sich.

Auch hier hebt sich der Fehler nicht auf, sondern ist weiter voll vorhanden und wirksam, natürlich erst, nachdem er auftritt. Der Unterschied zum ersten Beispiel ist, dass nun überhaupt keine „wahre“ Temperaturanomalie mehr gezeigt wird, sondern nur noch die fehlerbehaftete Summe.

Schleichender Fehler

Genauso entwickelt sich das Ergebnis, wenn man anstatt (oder zusätzlich) einen schleichenden Fehler in den zuvor fehlerfreien Verlauf einbringt. Ich habe das in der folgenden Abbildung getan. Dort wurde von Anfang an ein schleichender systematischer Fehler von +0,1 K/Dekade eingebracht, wie er z.B. vom städtischen Wärmeinseleffekt hervorgerufen werden

kann.

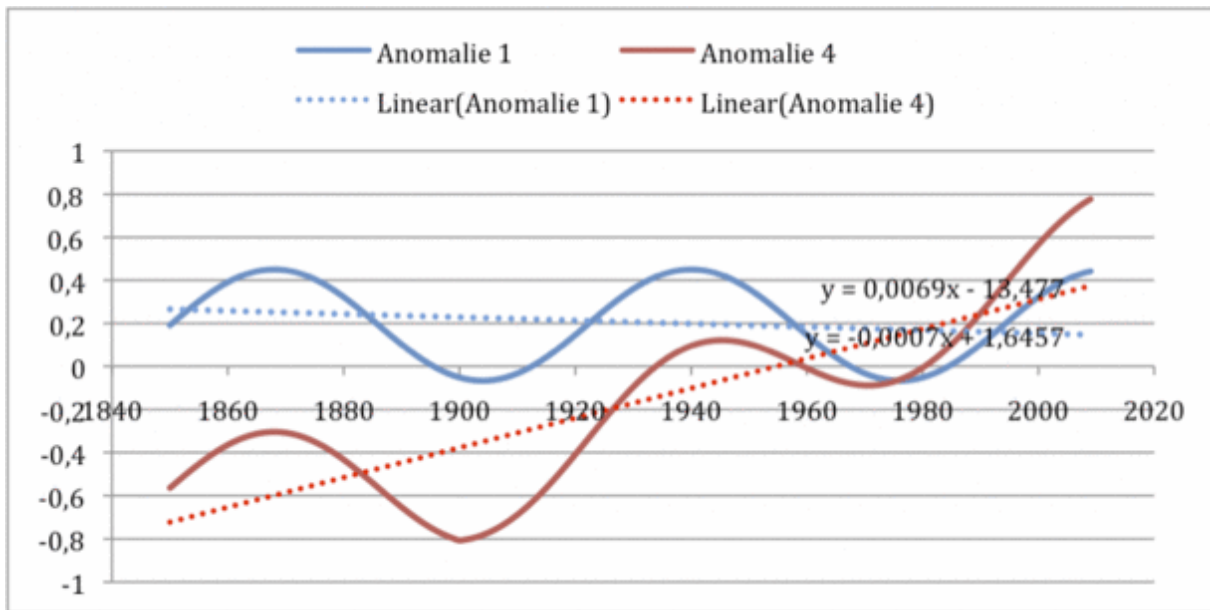


Abbildung 10 Ein systematischer Fehler von + 1 K tritt schleichend ab Beginn auf, weil z.B. die Station altert. Von dieser Veränderung sind das „Station Normal“ entsprechend seinem linearen Anteil ebenso wie alle aktuellen Werte von Anfang an betroffen. Der Fehler gleicht sich nicht mehr aus, weder vorher noch nachher. Der Trend verschiebt sich.

Wir sehen einen schön **ansteigenden** Verlauf (zuvor war er fast gerade) – verursacht allein durch den schleichenden systematischen Fehler- z.B den UHI. Nur kann jetzt überhaupt nicht unterschieden werden, ob ein systematischer Fehler vorliegt, oder ob sich die Umgebungstemperatur z.B. durch den Treibhauseffekt erhöht hat.

Man könnte nun beanstanden, dass die Fehlergröße in diesen Beispielen etwas hoch gewählt wurde. Dem ist aber nicht so, wie die zahlreichen Untersuchungen z.B. von Watts (<http://www.surfacestations.org/>) zeigen. Denn Fehler dieser Art gab und gibt es zahlreich. Sie werden durch Stationsverlegungen, Thermometertausch, Änderung der Farbbeschichtung der Station, Änderung des Algorithmus für die Berechnung des Mittelwertes u.v.a. mehr eingebracht. Es wäre nun vielleicht möglich den Anstieg im obigen Beispiel als Ausreißer zu erkennen, weil er einmalig und sprunghaft -wenn auch konstant- auftritt, und ihn durch entsprechende Rechnungen zu kompensieren. Das geschieht aber nur sehr, sehr selten, weil sich bei den abertausenden von Datensätzen der Vergangenheit kaum jemand diese Mühe macht, bzw. machen kann

Kein Fehlerausgleich möglich

Eine Korrektur unterbleibt hauptsächlich deswegen, weil man die Binsenweisheit (siehe Brohan et al 2006) von zuvor glaubt, dass sich der Fehler bei Anomalienbildung von selbst ausgleicht. Das ist aber, wie wir gesehen haben, grottenfalsch!

Eine Korrektur unterbleibt aber auch in den allermeisten Fällen deshalb, weil die dazu erforderlichen sog. „Metadaten“ fehlen und auch nicht mehr herbeigeschafft werden können. Diese beschreiben die Umgebungsbedingungen, Maßnahmen, Algorithmen und vieles anderes, was in und um die Station über den Zeitraum passiert ist. (Siehe dazu bspw. Harrys Read Me Files des Meteorologen und Programmierers bei der CRU Harry: „HARRY_READ_Me.txt.“ z.B. hier http://www.anenglishmanscastle.com/HARRY_READ_ME.txt. Diese ist 274 Seiten lang. Die dazugehörige Datenbasis enthält über 11.000 Dateien aus den Jahren 2006 bis 2009¹³)

Allgemein gilt daher, die Annahme, dass sich bei Anomalienbildung die Fehler aufheben, ist nur dann richtig, wenn der gemeinsame, gleich große und richtungsgleiche Fehler **vor dem Beginn** der untersuchten Zeitspanne eintritt und dann so bleibt. In unserem Falle also vor 1850. Das liegt jedoch weder in unserem Ermessen, noch haben wir davon Kenntnis, sondern es wird allein durch die Realität bestimmt. Deshalb kann festgehalten werden, dass diese simple Fehlerkorrekturmethode in aller Regel nicht anwendbar ist. Angewendet wird sie aber von so gut wie allen -auch IPCC- Klimatologen trotzdem.

Zusammenfassung

In der Statistik ist es eine gängige Methode Werte von div. Variablen mit Referenzwerten eben dieser Variablen zu vergleichen, um auf diese Weise bei eventuellen Abweichungen u.U. Ähnlichkeiten im Verlauf oder sogar Hinweise auf mögliche Ursache und Wirkungsbeziehungen zu bekommen. Allerdings muss man sich immer im Klaren darüber sein, wie sehr diese Methode von den Randbedingungen abhängt. Es wird gezeigt, dass eine schlichte Anomalienbildung keineswegs ausreichend ist, um schwer bestimmbare variable oder konstante systematische Fehler herauszurechnen. Im Gegenteil, man müsste in jedem Fall diese Fehler bestimmen, sie quantifizieren und einordnen, um sie dann evtl. mehr oder weniger gut rechnerisch ausgleichen zu können. In der Klimatologie ist diese Einschränkung in Bezug auf die Schwächen der Anomalienbildung offensichtlich nicht nur nicht bekannt, sondern wird auch – auf den ersten Anschein hin- negiert.

Der berühmte Physiker und Nobelpreisträger [Dick Feynman](#) würde sagen: „**Sie halten sich selbst zum Narren**“. Nur so lässt sich erklären, dass auch hochangesehene Forscher diese simplen Zusammenhänge oft nicht beachten. Ihre Ergebnisse sind dadurch entsprechend falsch und damit unbrauchbar, bzw. mit wesentlich größeren Fehlern (Unsicherheiten) behaftet als angegeben.

Anhang.

Zur behaupteten Genauigkeit

Die berechnete Globaltemperatur im letzten Jahrhundert hat sich, gemäß Aussage des Intergovernmental Panel of Climate Change (IPCC) um ca. 0,6 ° bis 0,7° C erhöht, wie uns im Bericht *TAR 2001(Third Assessment Report)* mitgeteilt wurde. Um diese Aussage treffen zu können, müssen Veränderungen über lange Zeiträume mit einer Genauigkeit von < 1/10 °C

¹³ Details entnommen aus ‚Botch after botch after botch‘ Leaked ‚climategate‘ documents show huge flaws in the backbone of climate change science By LORRIE GOLDSTEIN

ermittelt, erfasst, dokumentiert und verdichtet werden. Das englische Klimazentrum CRU (Climate Research Unit) -siehe Abbildung 1- zeigt sogar einen Anstieg von 0,8 °C mit einem *Vertrauensintervall* bei 95% von -0,09 bis +0,11 °C (Jahr 1905) bis $\pm 0,08$ ° (Jahr 2005.).¹⁴

Dem steht aber entgegen, dass selbst bei Verwendung der bestgewarteten Messstationen und von gut trainierten Meteorologen so genau wie möglich abgelesenen Temperaturwerte, diese nur die Bestimmung von Tagesmittelwerten mit einer von Genauigkeit ± 2 bis ± 3 K erlauben. Zitat des Meteorologen und Statistikers Jürgen Pelz

Jürgen Pelz: „Anmerkungen zur Prüfung von Daten und Ergebnissen von Modellrechnungen unter Verwendung der Statistik und der ‚Informationstheorie“ Beilage zur Berliner Wetterkarte vom 7.12.1995; S. 5 „*Will man beispielsweise die Tagesmitteltemperatur auf ± 0.1 K genau ermitteln, darf der Abstand der Messungen nicht grösser als 15 Minuten sein. Genügt eine Genauigkeit von ± 2 bis 3 K, reichen die Klimatermine.*“ [Pelz, 1995b]¹⁵

Trotzdem behaupten die Wissenschaftler der Climate Research Unit (CRU) der University von East Anglia, welche die CRU Hadley Reihe produzieren¹⁶: „*How accurate are the hemispheric and global averages? Annual values are approximately accurate to $\pm 0.05^{\circ}\text{C}$ (two standard errors) for the period since 1951. They are about four times as uncertain during the 1850s, with the accuracy improving gradually between 1860 and 1950 except for temporary deteriorations during data-sparse, wartime intervals. Estimating accuracy is a far from a trivial task as the individual grid-boxes are not independent of each other and the accuracy of each grid-box time series varies through time (although the variance adjustment has reduced this influence to a large extent). The issue is discussed extensively by Folland et al. (2001a, b) and Jones et al. (1997). Both Folland et al. (2001a,b) references extend discussion to the estimate of accuracy of trends in the global and hemispheric series, including the additional uncertainties related to homogeneity corrections.*“

¹⁴ Diese Darstellung der Zeitreihe der mittleren Globaltemperatur kommt auf vielfältige Weise in öffentlichen Berichten des IPCC AR 4 vor. So z.B. Im Summary for Policymakers SPM auf Seite 19, In den FAQ's des Berichtes der Working Group 1 Fig. 3.1 auf S 104. In der Technical Summary des Berichtes der Working Group 1 (WG1-TS) als Figure TS.6 auf Seite 37 usw.

¹⁵ Pelz, J b (1995) Anmerkungen zur Prüfung von Daten und Ergebnissen von Modellrechnungen unter Verwendung der Statistik und der Informationstheorie. Beilage zur Berliner Wetterkarte 7.12.1995

¹⁶ Quelle <http://www.cru.uea.ac.uk/cru/data/temperature/#datter>