

# Bedeutung und Nutzen von Mittelwerten bei der Anwendung auf Klima

geschrieben von Chris Frey | 3. Mai 2017

[Bemerkung: Im folgenden Beitrag hat der Autor bei allen Temperaturangaben (in °F) die Zehntel-Ziffer unterstrichen. Ich habe sämtliche Angaben in Grad Celsius umgerechnet, aber auf die Unterstreichung der Zehntel-Ziffern verzichtet. Aus dem Beitrag geht eindeutig hervor, was der Autor damit hervorheben will. Anm. d. Übers.]

## Mittelwerte

Mittelwerte sind vielfach nützlich. Ein allgemeiner Nutzen ist es, die Berechnung einer **fest vorgegebenen** Beschaffenheit genauer und präziser zu machen, wie etwa eine physikalische Dimension. Dies wird erreicht durch die Eingrenzung aller zufälligen Fehler beim Prozess der Messung. Unter geeigneten Umständen wie etwa der Berechnung des Durchmessers eines Balles mittels einer Schublehre können multiple Messungen einen präziseren mittleren Durchmesser ergeben. Grund hierfür ist, dass sich die Zufallsfehler bei der Ablesung auf der Schublehre gegenseitig herausmitteln und die Präzision durch die Standardabweichung gegeben wird, welche invers abhängig ist von der Quadratwurzel der Anzahl der Messungen.

Ein anderer allgemeiner Zweck ist es, die Eigenschaften einer **Variablen** zu charakterisieren, indem man multiple repräsentative Messungen durchführt und die Häufigkeitsverteilung der Messungen beschreibt. Dies kann graphisch erfolgen oder mittels statistischer Parameter wie **Mittelwert, Standardabweichung und Asymmetrie/Wölbung [skewness/kurtosis]**. Da jedoch die gemessenen Eigenschaften variieren, wird es problematisch, die Fehler bei den Messungen von der Variabilität der Eigenschaften zu trennen. Folglich erfahren wir mehr darüber, wie die Eigenschaft variiert, als über die Verteilung des zentralen Wertes. Und doch konzentrieren sich Klimatologen auf die arithmetischen Mittel sowie die daraus errechneten Anomalien. Mittelwerte können Informationen verschleiern, sowohl absichtlich als auch unabsichtlich.

Damit im Hinterkopf müssen wir untersuchen, ob zahlreiche Messungen der Temperaturen auf dem Festland, den Ozeanen und der Luft uns wirklich einen präzisen Wert für die „Temperatur“ der Erde vermitteln können.

## Die „Temperatur“ der Erde

Der Konvention zufolge ist das Klima üblicherweise definiert als das Mittel meteorologischer Parameter über einen Zeitraum von 30 Jahren. Wie können wir die verfügbaren Temperaturdaten verwenden, welche zum Zwecke der Wetterüberwachung und -vorhersage ermittelt werden, um das Klima zu

charakterisieren? Das gegenwärtig angewendete Verfahren ist die Berechnung eines arithmetischen Mittels für eine willkürliche Referenzperiode und dann die Subtraktion tatsächlicher Temperaturmessungen (entweder individuelle Werte oder Mittelwerte) von dieser Referenzperiode, um etwas zu erhalten, was man Anomalie nennt. Was jedoch bedeutet es, alle Temperaturdaten zu sammeln und das Mittel zu berechnen?

Befände sich die Erde im thermodynamischen Gleichgewicht, würde es nur eine Temperatur geben, die zu messen recht einfach wäre. Die Erde hat aber **nicht** nur eine Temperatur, sondern eine unendliche Vielfalt von Temperaturen. Tatsächlich variiert die Temperatur ununterbrochen horizontal, vertikal und mit der Zeit, was eine unendliche Anzahl von Temperaturen ergibt. Die offensichtliche Rekord-Tiefsttemperatur beträgt  $-135,8^{\circ}\text{F}$  [ca.  $-93,2^{\circ}\text{C}$ ] und die aufgezeichnete Höchsttemperatur  $159,3^{\circ}\text{F}$  [ca.  $70,7^{\circ}\text{C}$ ]. Die maximale Bandbreite beträgt also  $295,1^{\circ}\text{F}$ , die geschätzte Standardabweichung etwa  $74^{\circ}\text{F}$ , dem Empirischen Gesetz zufolge. Änderungen von weniger als einem Jahr sind sowohl zufällig als auch saisonal, längere Zeitreihen enthalten periodische Änderungen. Die Frage lautet, ob das Sammeln [von Daten an] einigen tausend Stellen über eine Periode von Jahren uns einen Mittelwert geben kann, welcher einen vertretbaren Wert beim Aufzeigen einer geringen Änderungsrate liefert?

Eines der Probleme ist, dass Wassertemperaturen dazu neigen, geschichtet aufzutreten. Die Wassertemperatur an der Oberfläche neigt dazu, die wärmste Schicht zu sein, darunter nimmt die Temperatur mit der Tiefe ab. Oftmals gibt es eine abrupte Änderung der Temperatur, Thermokline genannt. Außerdem kann aufwallendes Tiefenwasser kaltes Wasser an die Oberfläche bringen, vor allem entlang von Küsten. Daher ist die Lokalisierung und Tiefe der Datenproben entscheidend bei der Berechnung so genannter Meeresoberflächen-Temperaturen (SST). Des Weiteren muss berücksichtigt werden, dass – weil Wasser eine um 2 bis 5 mal höhere spezifische Wärme aufweist als normale Flüssigkeiten und eine vier mal höhere als Luft – es sich viel langsamer erwärmt als das Festland! Es ist unangebracht, SSTs mit Festlandstemperaturen zu mitteln. Das ist ein klassischer Fall eines Vergleiches von Äpfeln und Birnen. Falls jemand Trends der sich ändernden Temperatur erkennen will, können sie auf dem Festland offensichtlicher hervortreten als über den Ozeanen, obwohl Wassertemperaturen dazu neigen, Zufalls-Fluktuationen zu unterdrücken. Es ist vermutlich am besten, SSTs mit einer Skala zu plotten, die vier mal so groß ist wie bei den Festlandstemperaturen und beide in die gleiche Graphik einzutragen zum Vergleich.

Bei Temperaturen auf dem Festland ergibt sich oftmals das ähnlich gelagerte Problem von Temperatur-Inversionen, das heißt, es ist in den bodennahen Luftschichten kälter als in größerer Höhe darüber. Dies ist das Gegenteil dessen, was die Adiabate vorgibt, dass nämlich die Temperatur in der Troposphäre mit der Höhe abzunehmen hat. Aber dies stellt uns vor ein weiteres Problem. Temperaturen werden in Höhen aufgezeichnet, die in einem Bereich von unter dem Meeresspiegel (Death

Valley) bis über 3000 Metern Seehöhe liegen. Anders als es die Allgemeine Gasgleichung vorgibt, welche die Eigenschaften eines Gases bei Standard-Temperatur und -Druck festlegt, werden alle Wetter-Temperaturmessungen bei der Mittelung in einen Topf geworfen, um ein arithmetisches Mittel der globalen Temperatur zu erhalten ohne Berücksichtigung der Standard-Drücke. Dies ist wichtig, weil die Allgemeine Gasgleichung vorgibt, dass die Temperatur eines Luftpaketes mit abnehmenden Druck abnimmt, was das Temperaturgefälle steigen lässt.

Historische Aufzeichnungen (vor dem 20. Jahrhundert) sind besonders problematisch, weil Temperaturen nur auf das nächste ganze Grad Fahrenheit abgelesen wurden – von Freiwilligen, die keine professionellen Meteorologen waren. Außerdem war die Technologie von Temperaturmessungen nicht ausgereift, besonders hinsichtlich standardisierter Thermometer.

Klimatologen haben versucht, die oben beschriebenen zusammengewürfelten Faktoren zu umschreiben mit der Überlegung, dass Genauigkeit und damit Präzision verbessert werden kann durch Mittelbildung. Als Grundlage ziehen sie 30-Jahres-Mittel jährlicher Mittel monatlicher Mittel heran, womit sie die Daten glätten und Informationen verlieren! Tatsächlich besagt das ‚Gesetz über Große Zahlen‘, dass die Genauigkeit gesammelter Stichproben verbessert werden kann (falls keine systematischen Verzerrungen präsent sind!). Dies gilt besonders für probabilistische Ereignisse wie etwa das Werfen einer Münze. Falls jedoch jährliche Mittelwerte aus monatlichen Mittelwerten abgeleitet werden anstatt aus täglichen Mittelwerten, dann sollten die Monate gewichtet werden der Anzahl der Tage in jenem Monat gemäß. Es ist unklar, ob dies gemacht wird. Allerdings werden selbst tägliche Mittelwerte extrem hohe und tiefe Temperaturen glätten und somit die vorhandene Standardabweichung reduzieren.

Jedoch selbst wenn man die oben beschriebenen Probleme nur vorübergehend ignoriert, gibt es das fundamentale Problem zu versuchen, Lufttemperaturen an der Erdoberfläche präziser und genauer zu erfassen. Anders als der Ball mit im Wesentlichen einem Durchmesser (mit minimaler Exzentrizität) ändert sich die Temperatur an jedem Punkt der Erdoberfläche ständig. Es gibt keine einheitliche Temperatur für irgendeine Stelle zu irgendeinem Zeitpunkt. Und man hat nur eine Gelegenheit, jene vergängliche Temperatur zu messen. Man kann nicht multiple Messungen durchführen, um die Präzision einer bestimmten Messung der Lufttemperatur präziser machen!

## Temperatur-Messungen

Höhlen sind dafür bekannt, stabile Temperaturen aufzuweisen. Viele variieren um weniger als  $\pm 0,5^{\circ}\text{F}$  jährlich. Allgemein wird angenommen, dass die Temperatur der Höhle eine mittlere jährliche Lufttemperatur an der Oberfläche reflektieren, jedenfalls an der Stelle, an der sich die Höhle befindet. Während die Lage ein wenig komplexer ist, ist es doch

eine gute Approximation erster Ordnung. (Zufälligerweise gibt es einen interessanten Beitrag von Perrier *et al.* (2005) über einige sehr frühe Arbeiten in Frankreich hinsichtlich Temperaturen im Untergrund. Zur Illustration wollen wir einmal annehmen, dass ein Forscher die Notwendigkeit sieht, die Temperatur einer Höhle während einer bestimmten Jahreszeit zu bestimmen. Der Forscher möchte diese Temperatur mit größerer Präzision ermitteln als es ein durch die Gänge getragenes Thermometer vermag. Der Wert sollte nahe dem nächsten Zehntelgrad Fahrenheit liegen. Diese Situation ist ein ziemlich guter Kandidat für multiple Messungen zur Verbesserung der Präzision, weil es über einen Zeitraum von zwei oder drei Monaten kaum Temperaturänderungen geben dürfte und die Wahrscheinlichkeit hoch ist, dass die gemessenen Werte normalverteilt sind. Die bekannte jährliche Bandbreite gibt an, dass die Standardabweichung unter  $(50,5 - 49,5)/4$  oder etwa  $0,3^{\circ}\text{F}$  liegen sollte. Daher ist die Standardabweichung der jährlichen Temperaturänderung von gleicher Größenordnung wie die Auflösung des Thermometers. Weiter wollen wir annehmen, dass an jedem Tag, wenn die Stelle aufgesucht wird, der Forscher als Erstes und als Letztes die Temperatur abliest. Nach 100 Temperaturmessungen werden die Standardabweichung und der Standardfehler des Mittelwertes berechnet. Unter der Voraussetzung, dass es keine Ausreißer gibt und das alle Messungen um wenige Zehntelgrad um den Mittelwert verteilt sind, ist der Forscher zuversichtlich, dass es gerechtfertigt ist, den Mittelwert bekanntzugeben mit einer signifikanten Zahl mehr als das Thermometer direkt zu messen in der Lage war.

Jetzt wollen wir dies mit der allgemeinen Praxis im Bereich Klimatologie kontrastieren. Klimatologen verwenden meteorologische Temperaturen, die vielleicht von Individuen abgelesen werden mit geringerem Interesse an gewissenhaften Beobachtungen als die einschlägige Forscher haben. Oder Temperaturen wie etwa von den automatisierten ASOS können zum nächsten ganzen Grad Fahrenheit gerundet und zusammen gebracht werden mit Temperaturwerten, die zum nächsten Zehntelgrad Fahrenheit gerundet sind. (Zu allermindest sollten die Einzelwerte invers zu ihrer Präzision gewichtet werden). Weil die Daten eine Mittelung (Glättung) durchlaufen, bevor das 30-Jahre-Referenzmittel berechnet wird, scheint die Datenverteilung zusätzlich weniger verzerrt und mehr normalverteilt, und die berechnete Standardabweichung ist kleiner als sie es bei Verwendung der Rohdaten wäre. Es ist nicht nur die Mitteltemperatur, die sich jährlich ändert. Die Standardabweichung und Verzerrung ändert sich mit Sicherheit ebenfalls, aber dazu werden keine Angaben gemacht. Sind die Änderungen der Standardabweichung und der Verzerrung zufälliger Natur oder gibt es einen Trend? Falls es einen Trend gibt, was ist dafür die Ursache? Was bedeutet das, falls es überhaupt etwas bedeutet? Da sind Informationen, die nicht untersucht und bekannt gegeben werden, obwohl sie Einsichten in die Dynamik des Systems vermitteln können.

Man gewinnt sofort den Eindruck, dass die bekannten höchsten und tiefsten Temperaturen (siehe oben) zeigen, dass die Datensammlung eine Bandbreite bis zu  $300^{\circ}\text{F}$  haben kann, wenngleich auch näher bei  $250^{\circ}\text{F}$ .

Wendet man zur Schätzung der Standardabweichung die Empirische Regel an, würde man einen Wert über 70°F dafür vorhersagen. Geht man konservativer vor und zieht man das Tschbycheff'sche Theorem heran und teilt durch 8 anstatt durch 4, ergibt sich immer noch eine Schätzung über 31°F. Außerdem gibt es gute Gründe für die Annahme, dass die Häufigkeitsverteilung der Temperaturen verzerrt ist mit einem langen Schwanz auf der kalten Seite. Im Kern dieses Arguments steht die Offensichtlichkeit, dass Temperaturen unter 50°F unter Null normaler sind als Temperaturen über 150°F, während das bekannt gemachte Mittel der globalen Festlands-Temperaturen nahe 50°F liegt.

Im Folgenden wird gezeigt, wie meiner Ansicht nach die typischen jährlichen Rohdaten aussehen sollten, wenn sie als Häufigkeitsverteilung geplottet werden, wobei man die bekannte Bandbreite, die geschätzte Standardabweichung und das veröffentlichte Mittel berücksichtigt:



Die dicke, rote Linie repräsentiert die typischen Temperaturen eines Jahres, die kurze grüne Säule (in der Skala eingeordnet) repräsentiert die Temperatur in einer Höhle bei obigem Temperatur-Szenario. Ich bin zuversichtlich, dass das Mittel der Höhlentemperatur bis zu einem Hundertstel Grad Fahrenheit präzise ist, aber trotz der gewaltigen Anzahl von Messungen der Temperatur auf der Erde bringe ich Gestalt und Verteilung der globalen Daten längst nicht die gleiche Zuversicht für die globalen Temperaturen auf! Es ist offensichtlich, dass die Verteilung eine **erheblich** größere Standardabweichung aufweist als die Höhlentemperatur, und die Rationalisierung mittels Teilung der Quadratwurzel der Anzahl kann nicht gerechtfertigt sein, um Zufallsfehler zu eliminieren, wenn der gemessene Parameter niemals zweimal den gleichen Wert aufweist. Die multiplen Schritte der Mittelung der Daten reduziert Extremwerte und die Standardabweichung. Die Frage lautet: „Ist die behauptete Präzision ein Artefakt der Glättung, oder ergibt das Glättungsverfahren einen präziseren Wert?“ Darauf weiß ich keine Antwort. Es ist jedoch mit Sicherheit etwas, das diejenigen beantworten und rechtfertigen sollten, die die Temperatur-Datenbasis erstellen!

## Zusammenfassung

Die Theorie der anthropogenen globalen Erwärmung prophezeit, dass man die stärksten Auswirkungen nachts und im Winter sehen wird. Das heißt, der kalte Schwanz in der Kurve der Häufigkeitsverteilung sollte sich verkürzen und die Verteilung symmetrischer werden. Dies würde die berechnete globale Mitteltemperatur zunehmen lassen, selbst wenn sich die Temperaturen im hohen und mittleren Bereich gar nicht ändern. Die Prophezeiungen zukünftiger katastrophaler Hitzewellen basieren auf der unausgesprochenen Hypothese, dass mit der Zunahme des globalen Mittels sich die gesamte Kurve der Häufigkeitsverteilung hin zu höheren Temperaturwerten verschieben würde. Das ist keine gewünschte Hypothese,

weil die Differenz zwischen täglichen Höchst- und Tiefsttemperaturen während des 20. Jahrhunderts nicht konstant war. Sie bewegen sich nicht in Stufen, möglicherweise weil es unterschiedliche Faktoren gibt, welche die Höchst- und Tiefstwerte beeinflussen. Tatsächlich waren einige der tiefsten Tiefsttemperaturen in der Gegenwart gemessen worden! In jedem Falle ist eine globale Mitteltemperatur keine gute Maßzahl für das, was mit den globalen Temperaturen passiert. Wir sollten die Trends der täglichen Höchst- und Tiefsttemperaturen in allen Klimazonen betrachten, wie sie durch physikalische Geographen definiert worden sind. Wir sollten auch die Form der Häufigkeits-Verteilungskurven für verschiedene Zeiträume analysieren. Der Versuch, das Verhalten des ‚Klimas‘ der Erde mit einer einzelnen Zahl zu charakterisieren, ist keine gute Wissenschaft, ob man nun an Wissenschaft glaubt oder nicht.

Link:

<https://wattsupwiththat.com/2017/04/23/the-meaning-and-utility-of-averages-as-it-applies-to-climate/>

Übersetzt von Chris Frey EIKE