

Langzeitkorrelationen in der Natur: von Klima, Erbgut und Herzrhythmus

Die Fluktuationsanalyse erlaubt es, Klimamodelle zu testen oder Schlafphasen zu untersuchen

Armin Bunde und Jan W. Kantelhardt

Wichtige Auskünfte über komplexe Systeme in der Natur erhalten wir oft aus Zeitreihen: Aus Temperatur- und Niederschlagsreihen erhoffen wir uns Aufschluss über globale Klimaänderungen, und aus medizinischen Zeitreihen wie der des Herzschlags oder des Blutdrucks kann man oft Rückschlüsse auf den Gesundheitszustand des Patienten ziehen. In den letzten Jahren hat sich herausgestellt, dass in einigen wichtigen Zeitreihen die Schwankungen der einzelnen Elemente um ihren Mittelwert nicht voneinander unabhängig sind. Die Schwankungen sind stattdessen über sehr lange Zeitspannen hinweg untereinander gekoppelt, d. h. die zugrunde liegenden Prozesse besitzen ein oft überraschend langes Gedächtnis. Von diesem Phänomen der Langzeitkorrelationen, die gewisse universelle Züge tragen, handelt dieser Artikel.

Dass die Fluktuationen der Tagestemperaturen um ihre langjährigen, von der Jahreszeit abhängenden Mittelwerte nicht ganz zufällig sind, weiß man schon seit langem. Einem zu warmen Tag folgt eher wieder ein zu warmer Tag als ein zu kalter und umgekehrt. Diese Erhaltungstendenz („Persistenz“) des Wetters wird durch die Existenz von Großwetterlagen bedingt, in denen das Wetter einigermaßen stabil ist [1, 2].

Allgemein spricht man von „Persistenz“, wenn eine Zufallsgröße die Tendenz hat, ihren augenblicklichen Wert beizubehalten. Dies kann über längere Zeiträume zu ausgeprägten positiven oder negativen Abweichungen vom jeweiligen Mittelwert führen. Bei Trends hingegen handelt es sich um systematische Abweichungen vom Mittelwert, die durch äußere Prozesse erzwungen werden. In diesem Sinne sind der Jahresgang und eine mögliche Klimaänderung durch den Treibhauseffekt „Trends“; Großwetterlagen dagegen stellen eine „Persistenz“ auf kürzeren Zeitskalen dar.

Typische Großwetterlagen klingen etwa nach einer Woche ab, und die Frage ist, ob es auf größeren Zeitskalen, nach einem Monat, nach einem Jahr oder nach einer Dekade überhaupt noch eine Persistenz gibt, und wenn ja, wie sie zeitlich abfällt. Wenn dies geklärt ist, wüssten wir auch gern, wie diese Gesetzmäßigkeit vom jeweiligen Ort des Geschehens abhängt.

Zu diesem Zweck betrachten wir Temperaturreihen T_i , die an einem festen Ort – z. B. Prag – gemessen wur-



den (Abb. 1) [3]. Für solche Temperaturreihen verwendet man z. B. mittlere oder maximale Tagestemperaturen. Der Index i läuft über den gesamten Beobachtungszeitraum, für Prag sind das 218 Jahre. Um saisonal bedingte Einflüsse auszuschalten, bestimmen wir zunächst den Jahresgang, indem wir für jeden Jahrestag den über alle Jahre gemittelten Durchschnittswert \bar{T}_i berechnen, und ihn dann von den T_i abziehen. Wir erhalten so eine neue Zeitreihe $\tau_i = T_i - \bar{T}_i$ (Abb. 1a).

Um Persistenz zu quantifizieren, schaut man sich in der Regel die Autokorrelationsfunktion $C(s)$ an, die angibt, wie zwei Zufallszahlen, die s Werte auseinanderliegen, miteinander korreliert sind. Man erhält $C(s)$, indem man das Produkt $\tau_i \tau_{i+s}$ über alle Paare mit festem zeitlichen Abstand s mittelt und durch das mittlere Schwankungsquadrat $\overline{\tau^2}$ dividiert. Wenn $C(s)$ nur für $s=0$ verschieden von 0 ist, sind die τ_i gänzlich unkorreliert. Von *Kurzzeitkorrelationen* sprechen wir, wenn die charakteristische Korrelationszeit $s_x = \int_0^\infty C(s) ds$, mit der $C(s)$ abfällt, verhältnismäßig klein ist. Das einfachste Beispiel hierfür sind exponentiell abklingende Korrelationen, $C(s) = \exp[-s/s_x]$. Von *Langzeitkorrelationen* hingegen sprechen wir, wenn $C(s)$ asymptotisch wie

$$C(s) \sim s^{-\gamma}, \quad 0 < \gamma < 1, \quad (1)$$

abfällt. In diesem Fall divergiert die charakteristische Korrelationszeit und wird damit größer als alle anderen Zeitskalen im System.

Wie hat sich die mittlere Tagestemperatur in Prag zwischen 1775 und 1992 verändert? Eine Auswertung der gemessenen Temperaturdaten mithilfe der trendbeseitigenden Fluktuationsanalyse ermöglicht es, die Erwärmung aufgrund der zunehmenden Verstärkung von Prag von den langzeitkorrelierten Temperaturschwankungen zu trennen.

Prof. Dr. Armin Bunde, Dr. Jan W. Kantelhardt, Universität Gießen, Institut für Theoretische Physik III, Heinrich-Buff-Ring 16, 35392 Gießen

Bei der direkten Bestimmung von $C(s)$ durch Mittelung der Produkte $\tau_i \tau_{i+s}$ tritt folgendes Problem auf: Wenn $C(s)$ schon stark abgefallen ist, schwankt es um 0 herum und die uns interessierenden Gesetzmäßigkeiten für große Zeiten s sind nicht zu erkennen. Das sieht man sehr schön in den Abbildungen 2a und 2b, in denen $C(s)$ doppelt-logarithmisch für eine Reihe kurz- bzw. langreichweitig korrelierter Zahlen aufgetragen ist. Oberhalb der mit roten Pfeilen markierten Skala treten aufgrund der Schwankungen sogar schon negative Werte von $C(s)$ auf, und man kann das Korrelationsverhalten nicht mehr zweifelsfrei bestimmen. Hinzu kommt, dass Trends – zum Beispiel ein Anwachsen der Temperatur durch Verstädterung – bei der Betrachtung von $C(s)$ Langzeitkorrelationen vortäuschen können.

Um das Vorliegen von Langzeitkorrelationen zu untersuchen, empfiehlt es sich also nicht, $C(s)$ direkt zu bestimmen. Geschickter ist es, anstatt der τ_i die kumu-

lierte Zeitreihe $Y_j = \sum_{i=1}^j \tau_i$ zu untersuchen, die man als Position eines diffundierenden Teilchens nach j Schritten auffassen kann: Im i -ten Schritt bewegt sich das Teilchen um die Strecke $|\tau_i|$ nach links (falls $\tau_i < 0$) oder nach rechts (falls $\tau_i > 0$). Wenn wir nun die Größe $F_j^2(s) \equiv (Y_{j+s} - Y_j)^2$ für alle Paare mit festem s bestimmen und darüber mitteln, erhalten wir das mittlere Fluktuationsquadrat $F^2(s)$, das völlig dem mittleren Verschiebungsquadrat des diffundierenden Teilchens nach s Schritten entspricht. Aus dem Verhalten von $F^2(s)$ kann man nun Aufschluss über $C(s)$ gewinnen:

► a) Im Falle von Kurzzeitkorrelationen, wenn die τ_i oberhalb der Korrelationszeit s_x praktisch unkorreliert sind, gilt

$$F^2(s) \sim s \text{ für } s > s_x, \quad (2)$$

siehe Abb. 2a. Das ist das Ficksche Diffusionsgesetz; unterhalb von s_x wächst $F^2(s)$ stärker als linear an.

► b) Falls die τ_i langzeitkorreliert sind wie in Gl. (1), dann gilt

$$F^2(s) \sim s^{2\alpha} \text{ mit } \alpha = 1 - \gamma/2, \quad (3)$$

siehe Abb. 2b.

Das Problem ist allerdings, dass bei dieser einfachen Fluktuationsanalyse (FA) Trends auch $F^2(s)$ verfälschen und damit ein falsches Korrelationsverhalten vorgaukeln können. Das ist in Abb. 2c für künstliche unkorrelierte Zufallszahlen z_i gezeigt, zu denen ein schwacher Trend addiert wurde: $\tau_i = z_i + 5i/10^6$. Durch den Trend tritt ein Übergang (Crossover) bei $s_C \approx 70$ auf. Oberhalb des Crossovers trägt die Steigung nicht mehr $\alpha = 0,5$, was nach (2) und (3) für unkorrelierte Zahlen zu erwarten ist, sondern zunächst $\alpha \approx 0,9$ und später, bei ganz großen s -Werten, $\alpha = 1$. Das eigentlich unkorrelierte Verhalten wird also durch den Trend verschleiert.

Um Langzeitkorrelationen von Trends zu separieren, muss man also das Verfahren noch weiter modifizieren. Eine mögliche Variante stellt die „Detrended Fluctuation Analysis“ (DFA) dar, die in ihrer einfachsten Form 1994 von Peng *et al.* entwickelt wurde, um langreichweitige Korrelationen zwischen den DNA-Bausteinen zu untersuchen (s. u.) [4]. Wie im Infokasten „Die trend-beseitigende Fluktuationsanalyse (DFA)“ erklärt wird, unterscheiden wir verschiedene Ordnungen der DFA. In DFA1 werden lineare Trends in der kumulierten Reihe Y_j eliminiert (welche stückweise konstanten Trends in der Originalreihe τ_i entsprechen), in DFA2 quadratische Trends (entsprechend linearen Trends in den τ_i), usw.

Nach dieser langen Einleitung kehren wir nun zu den mittleren

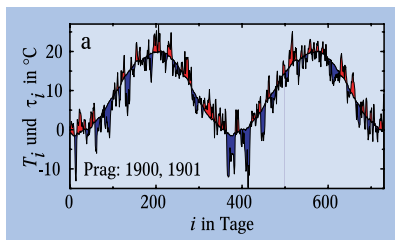


Abb. 1: Analyse der Prager Temperaturdatenreihe, die aus den gemessenen Tagesmittelwerten von 1775 bis 1992 besteht. Teil (a) zeigt für die Jahre 1900 und 1901 die mittlere Tagestemperatur T_p , wobei die Abweichungen vom Jahresgang rot oder blau eingefärbt sind. Die Teile (b) und (c) zeigen die Fluktuationsfunktionen $F(s)$ für FA (\times), DFA1 (\square), DFA2 (\circ), DFA3 (\triangle), DFA4 (∇) und DFA5 (\diamond) für die Originaldaten bzw. für Daten, die nach der Subtraktion des Jahresganges zufällig gemischt wurden, um die langreichweitigen Korrelationen zu zerstören (in Anlehnung an [3]).

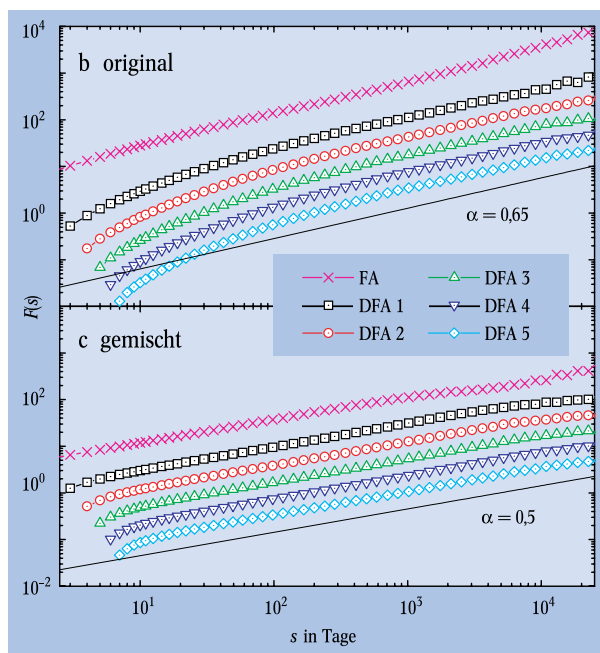
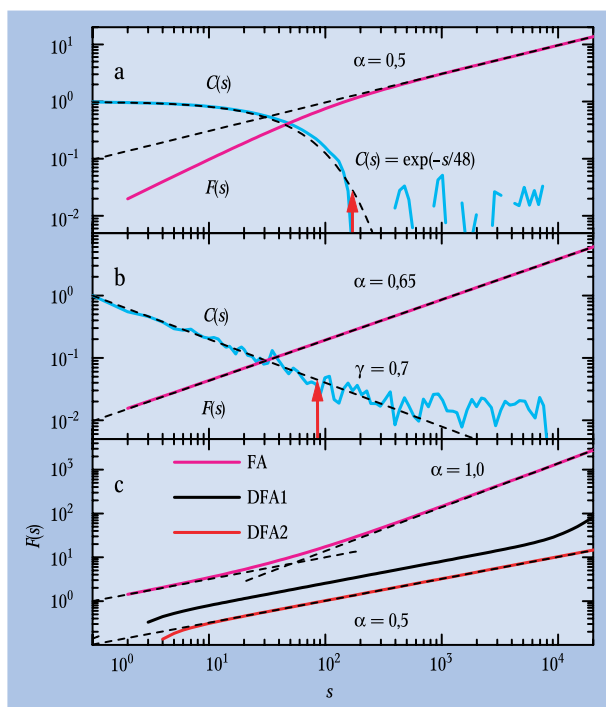


Abb. 2: Beispiele von Autokorrelationsfunktionen $C(s)$ (blau) und den zugehörigen Fluktuationsfunktionen $F(s)$.
 ► a) Analyse von kurzreichweitig korrelierten Zahlen mit exponentiell abfallendem $C(s)$, die man erhält, indem man die τ_i rekursiv mit einer Wahrscheinlichkeit von 0,98 gleich dem vorhergehenden Wert τ_{i-1} setzt und mit einer Wahrscheinlichkeit von 0,02 als neue unabhängige Zufallszahlen wählt.
 ► b) Analyse von langreichweitig korrelierten Zahlen mit potenzgesetzartig ($\gamma = 0,7$) abfallendem $C(s)$, die mit dem in [7] beschriebenen Algorithmus erzeugt wurden.
 ► c) Analyse von unkorrelierten Zahlen, zu denen der Trend $5i/10^6$ addiert ist. $F(s)$ ist über jeweils 500 Reihen aus 10^5 Werten gemittelt, während $C(s)$ jeweils aus nur einer Reihe mit 5×10^4 Werten bestimmt ist. Die roten Pfeile in a) und b) zeigen die kleinsten s an, bei denen das numerische $C(s)$ erstmals negativ wird.



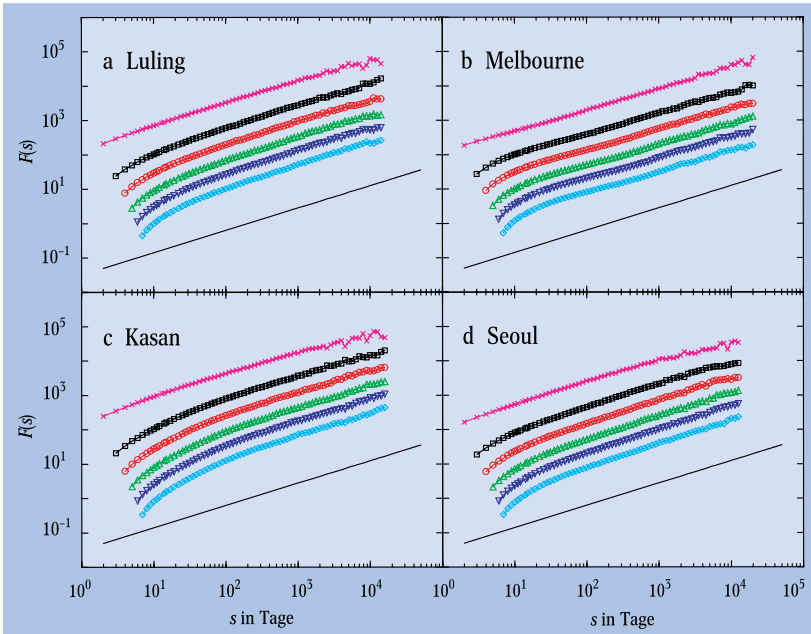


Abb. 3: FA- und DFA-Fluktuationsfunktionen für (a) Luling (Texas, USA, 1902–1991), (b) Melbourne (Australien, 1887–1994), (c) Kasan (Russland, 1895–1990) und (d) Seoul (Korea, 1908–1993). Die Symbole folgen Abb. 1 (nach [3e]).

Tagestemperaturen in der Prager Reihe zurück. Abbildung 1a zeigt die Originaldaten für zwei Jahre und Abb. 1b die Ergebnisse von FA und DFA1–DFA5. Die Flächen in der Abweichung vom Jahresgang in (a) weisen auf Persistenz auf kurzen Zeitskalen hin. Den Einfluss eines Trends, der auf die Stadterwärmung von Prag zurückzuführen ist, sieht man in Abb. 1b noch in der obersten Kurve (FA), die sich oberhalb von ca. 1500 Tagen in dem doppelt-logarithmischen Plot als leicht nach oben gekrümmte Kurve darstellt. Abgesehen von sehr großen Zeitskalen $s > 1500$ d (auf denen der Trend zur erhöhten Steigung von $\alpha \approx 0,8$ führt) und abgesehen von Zeitskalen innerhalb einer Großwetterlage hat die FA-Kurve näherungsweise die Steigung $\alpha = 0,65$. Die gleiche Steigung finden wir für die DFA1-bis DFA5-Kurven. Da die Steigungen systematisch konstant bleiben, schließen wir aus Abb. 1b und unter Verwendung von Gl. (3), dass die trend-bereinigte Temperatur-Korrelationsfunktion auf langen Zeitskalen wie

$$C(s) \sim s^{-\gamma}, \quad \gamma \approx 0,7, \quad (4)$$

abnimmt. Zum Vergleich haben wir, dem Beispiel von [3] folgend, die Daten zufällig durcheinander gemischt. Das Ergebnis ist in Abb. 1c zu sehen: Die Steigung der Fluktuationsfunktionen ist 0,5, wie wir es erwarten können, denn die Korrelationen im zeitlichen Verlauf der τ_i wurden durch das Mischen zerstört. Zwei Dinge sind hier zu beachten: (i) Die *Crossovers* bei ganz kleinen Zeitskalen sind ein Artefakt der DFA-Methode und können weitgehend beseitigt werden [5]. (ii) Die Abweichungen von der Steigung 0,5 bei großen Zeitskalen sind auf unzureichende Statistik zurückzuführen. Sie setzen erst bei etwa 30 Jahren ein und zeigen damit die Grenze unseres Zeitfensters auf, in dem wir verlässliche Aussagen über die Korrelationen mit unseren Methoden machen können.

Abbildung 3 zeigt die entsprechenden Analyseergebnisse für vier weitere Wetterstationen, die in ganz verschiedenen Klimazonen angesiedelt sind. Diesmal wurden die maximalen Tagestemperaturen analysiert. Wie

bei Prag stellen die DFA-Fluktuationsfunktionen in dem doppelt-logarithmischen Plot für große Zeitskalen Geraden dar, deren Steigungen alle mit $\alpha = 0,65$ konsistent sind. Ähnliche Ergebnisse wurden für alle bisher untersuchten langen Temperaturzeitreihen erhalten [3]. Dies deutet darauf hin, dass es sich bei den Langzeitkorrelationen um ein globales Phänomen handelt, das unabhängig von den lokalen Gegebenheiten der jeweiligen Messstationen ist. Ursache könnte die Kopplung der Atmosphäre an die Ozeane sein, die in letzter Zeit – gerade auch bei den zu erwartenden Klimaänderungen – in den Vordergrund des Interesses getreten ist [6].

Das Vorliegen von Langzeitkorrelationen hat wichtige Konsequenzen, sowohl für die Abschätzung

der Eintrittswahrscheinlichkeit von Extremereignissen, die dann untereinander gekoppelt sind (s. u.), als auch für die Trendbestimmung. Konventionell wird die Trendanalyse durch „Moving Averages“ durchgeführt, bei denen die interessierenden Daten über hinreichend

Die trend-beseitigende Fluktuations-Analyse (DFA)

Wir betrachten eine Reihe (x_i) von $i = 1, \dots, N$ äquidistanten Messwerten. Als erstes wird der Mittelwert \bar{x} aller x_i subtrahiert: $\tau_i = x_i - \bar{x}$ und eine kumulierte Reihe konstruiert: $Y_j = \sum_{i=1}^j \tau_i$. Dann wird Y_j in $N_s = [N/s]$ nicht-überlappende Segmente der Skalenlänge s aufgeteilt. Um den meist übrig bleibenden Rest nicht zu vernachlässigen, erfolgt die Aufteilung an zwei identischen Kopien jeweils von dem

$$F_v^2(s) = \frac{1}{s} \sum_{j=(v-1)s+1}^{vs} (Y_j)^2$$

einen und dem anderen Ende her, so dass man insgesamt $2N_s$ Segmente erhält. Als nächstes gilt es, die Fluktuationen in jedem Segment abzuschätzen. Bei der DFA [4, 5] werden Trends dadurch beseitigt, dass die Fluktuationen in jedem Segment aus der mittleren Abweichung von einem an den Verlauf angepassten Polynom bestimmt werden. Dadurch werden Trends auf der Skala s beseitigt. Es wird also in jedem Segment ein Polynom vom Grad n an die kumulierte Reihe angepasst, und aus der Differenz zwischen Y_j und dem Wert des Polynoms die trend-bereinigte kumulierte Reihe Y_j' definiert (s. Abb.). Der Wert von

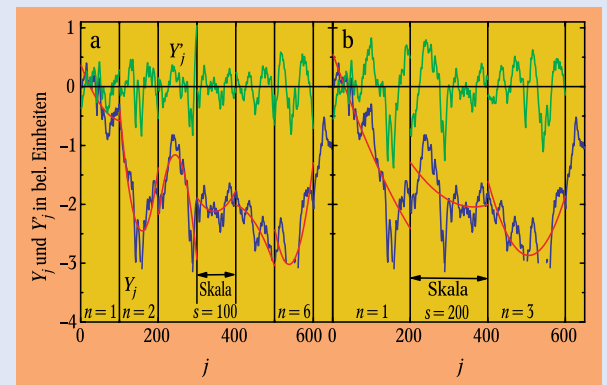


Illustration zur Erklärung der trend-bereinigenden Fluktuationsanalyse 2. Ordnung (DFA2). Für zwei (Zeit-) Fenster $s = 100$ (a) und $s = 200$ (b) sind die kumulierten Reihen Y_j (blau), die daran angepassten, stückweise quadratischen Polynome (rot) und die trend-bereinigten kumulierten Reihen Y_j' (grün) aufgetragen (nach [2]).

In der Abbildung wurden quadratische Polynome benutzt (DFA2). In der DFA der Ordnung n werden Trends der Ordnung n in der kumulierte Reihe und der Ordnung $n - 1$ in den Daten bereinigt, sodass ein Vergleich der Ergebnisse für verschiedene n eine Abschätzung der Ordnung der Trends ermöglicht [5].

große Zeitfenster gemittelt werden. Unter der Annahme, dass sich die zufälligen Fluktuationen dann herausmitteln, kann man Aufschluss über die Trends bekommen, und danach versuchen, durch geeignete Fits eine Vorhersage für die Zukunft zu erhalten. Dass dieses Verfahren im Falle von Langzeitkorrelationen nicht funktioniert, zeigt Abbildung 4, in der ein Ausschnitt aus einer langen Reihe langzeitkorrelierter künstlicher Daten (mit $\alpha=0,8$; $\gamma=0,4$) zu sehen ist (für eine Be-

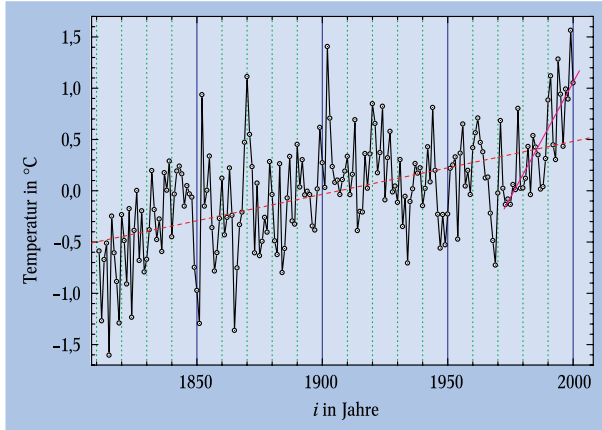


Abb. 4: 190 Punkte („Jahre“) einer künstlichen Reihe von langreichweitig korrelierten Zahlen („Temperaturwerten“; mit $\alpha=0,8$ und $\gamma=0,4$) ohne echten Trend. Der Abschnitt wurde aus einer Reihe mit 10 000 Punkten so ausgewählt, dass eine lineare Regression einen zum Ende hin verstärkten ansteigenden Trend vermuten lässt, wie er in klimatischen Datenreihen beobachtet wird. Die rote gestrichelte Linie ist ein Fit zu allen Punkten mit der Steigung $0,5^\circ\text{C}$ in 100 „Jahren“, während der Fit der letzten 25 Jahre (violette Gerade) sogar eine Steigung von $4,5^\circ\text{C}$ in 100 „Jahren“ aufweist.

schreibung der Erzeugung künstlicher korrelierter Zahlen s. z. B. [7]). Die Daten schwanken um 0 und Trends gibt es nicht. Gleichwohl zeigt die lineare Regression (Ausgleichsgerade) scheinbare Trends an. Nähme man die Analyse des ausgewählten Abschnitts ernst und interpretierte die Daten als Abweichungen einer mittleren Jahrestemperatur vom Langzeitmittel, so würde man bis zum Jahr 2100 eine Erwärmung um $4,5^\circ\text{C}$ vorhersagen und sich damit in diesem Beispiel gewaltig verschätzen.

Die Vorhersagen der Klimatologen basieren denn auch meist nicht auf solch trivialen Abschätzungen. Die heutigen Klimamodelle (General Circulation Models) ziehen bereits vielfältige Beeinflussungen der Atmosphäre z. B. durch die Ozeane und den wachsenden CO_2 -Gehalt der Luft mit in Betracht [6]. Aus Platzgründen können wir auf eine Beschreibung der Modelle und ihrer Errungenschaften nicht näher eingehen.

Da das Persistenzgesetz universell ist und nicht von Details in der Umgebung der jeweiligen Messstationen abhängt, ist es möglich, die Ergebnisse der Modelle kritisch zu prüfen. Eine solche Untersuchung wurde vor kurzem durchgeführt [8], und die Ergebnisse werden hier exemplarisch für das Klimamodell ECHAM4 skizziert, das vom MPI für Meteorologie gemeinsam mit dem Deutschen Klimarechenzentrum in Hamburg entwickelt worden ist. Abbildung 5 zeigt die Analyse der mittleren Monatstemperaturen für Luling (Texas) wie sie durch Interpolation der vier benachbarten Gitterpunkte, für die ECHAM4 Daten lieferte, gewonnen wurden. Die Analyse beschränkte sich auf den Zeitraum bis 1991, für den auch reale Daten vorliegen (Abb. 3a). Die Modelldaten zeigen, im Gegensatz zu

den realen Daten, ein Abknicken der DFA-Funktionen oberhalb von zwei Jahren. Der Exponent ändert sich auf großen Zeitskalen in 0,5 für alle DFA-Kurven, d. h. die Korrelationen verschwinden für große zeitliche Abstände. Das Modell ist also nicht in der Lage, für Texas das universelle Persistenzgesetz zu reproduzieren. Auf großen Zeitskalen liefert es schlicht unkorreliertes Verhalten.

Nichtsdestoweniger zeigt die FA-Fluktuationenfunktion ein schönes Potenzverhalten, mit einem Exponenten $\alpha \approx 0,65$ auf großen Zeitskalen und täuscht so die Existenz von Langzeitkorrelationen $C(s) \sim s^{-0,7}$ vor. Abbildung 5 ist daher auch ein schönes Beispiel dafür, dass Trends in der FA Langzeitkorrelationen vortäuschen können.

Indem wir die Ergebnisse aus FA und DFA miteinander vergleichen, können wir etwas über die Struktur der Trends in den stimulierten Daten lernen. Da alle DFA-Kurven für große s denselben Exponenten $\alpha = 0,5$ liefern, kann der auftretende Trend nur schwach und höchstens linear in der kumulierten Reihe sein, also stückweise konstant in den Originaldaten.

Ganz offensichtlich überschätzt das Modell die Trends auf Kosten der Persistenz. Es ist daher zu erwarten, dass das Modell auch die künftige globale Erwärmung überschätzt. Ähnliche Ergebnisse, wenn auch nicht immer so ausgeprägt wie hier, erhält man für andere Gegenden und von den anderen Modellen. Wenn auch eine weitere Klimaänderung in den nächsten 100 Jahren als wahrscheinlich angesehen werden muss, so besteht doch die Hoffnung, dass die Änderung weniger dramatisch ausfällt als von den heutigen Klimamodellen prognostiziert.

Hydrologie

Schon vor langer Zeit wurden von H. E. Hurst die historischen Tiefststände des Nils untersucht. Er fand die ausgeprägte Tendenz, dass sie sich über sehr lange Zeiträume oberhalb oder unterhalb des langjährigen Mittels bewegen. Mit seiner Analyse, welche der FA-Methode entspricht, erhielt er einen Fluktuationsexponenten $\alpha \approx 0,9$ [9]. Da Hurst unglücklicherweise zur Erzeugung vergleichbarer künstlicher Zahlenreihen ein

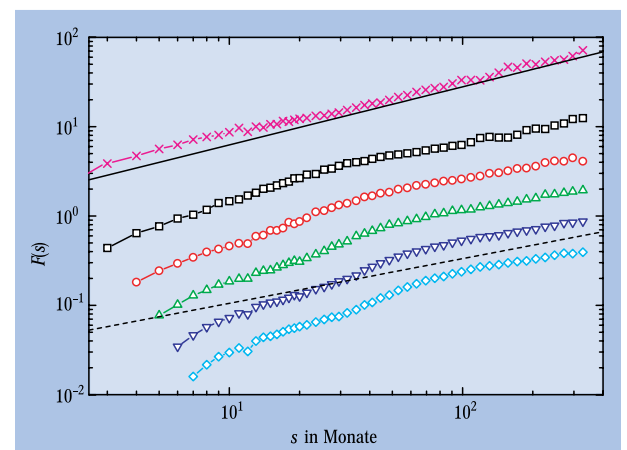


Abb. 5: Ergebnis der Korrelations-Analyse für simulierte Monatsmittel-Temperaturdaten des ECHAM4-Klimamodells zu den Jahren 1860 bis 1991 für Luling (Texas). Aufgetragen sind die FA- und DFA-Fluktuationenfunktionen $F(s)$, wobei dieselben Symbole wie in Abb. 1 benutzt wurden. Die durchgezogene Gerade zeigt zum Vergleich die Steigung $\alpha=0,65$, während die gestrichelte Gerade die für unkorrelierte Daten erwartete Steigung $\alpha=0,5$ aufweist (nach [8]).

stochastisches Modell vorschlug, das nur Trends und kurzreichweitige Korrelationen liefert [10] (was Hurst übersah), wurden seine Analysemethoden und Ergebnisse stark in Zweifel gezogen. Neuere Ergebnisse unterstützen zumindest teilweise die Hurstschen Vorstellungen von Langzeitkorrelationen bei Flusswasserständen und Durchflüssen. Universelles Verhalten ist allerdings (entgegen der Hurstschen Vorstellungen) nicht zu erwarten, denn bei Durchflussreihen spielt der Abstand der Messstation von der Quelle, d. h. die Größe des „Einzugsgebietes“, eine wichtige Rolle, wie auch die Beschaffenheit des Speichermediums Boden im Einzugsgebiet [11].

Die Langzeitkorrelationen haben eine wichtige Konsequenz in der Hochwasserforschung. Bisher werden die Eintrittswahrscheinlichkeiten extremer Durchflüsse aus der Extrapolation der Durchfluss-Verteilungsfunktion zu großen Durchflüssen hin gewonnen. Diese Praxis ist fragwürdig, weil sie außer Stationarität auch die Abwesenheit von Langzeitkorrelationen voraussetzt, also die zeitliche Struktur der Hochwasserhistorie vernachlässigt und damit einen wesentlichen Teil der Informationen über des komplexe System ignoriert. Neue Denkansätze sind nötig, um die Wahrscheinlichkeit von Extremereignissen bei Anwesenheit von Langzeitkorrelationen und in Zeiten des globalen Wandels abzuschätzen. Dasselbe gilt auch für die Klimavorhersagen, denn die Wahrscheinlichkeit extremer Hitze- oder Kältetage lässt sich wegen der Langzeitkorrelationen nicht ohne weiteres auf der Basis einfacher Verteilungsfunktionen sinnvoll abschätzen.

Erbgut (DNA)

Die DNA besteht aus einer Doppelhelix von Phosphat-Zuckern, an denen innen jeweils eine von vier möglichen Basen hängt: Adenin, Guanin (Purine) oder Cytosin, Thymin (Pyrimidine). Die einander gegenüberliegenden Basenpaare sind durch Wasserstoffbrücken derart gebunden, dass sich nur Adenin und Thymin oder Guanin und Cytosin gegenüberstehen können. Die Frage ist, ob an einem Strang der DNA die Purine und die Pyrimidine korreliert oder völlig regellos angeordnet sind. Um das zu sehen, können wir die Abfolge der Basen entlang eines der beiden DNA-Stränge als Datenreihe x_i auffassen, wobei i die Position der Basen angibt. Für die beiden Pyrimidine kann man z. B. $x_i = +1$ setzen und für die beiden Purine $x_i = -1$ (Purin-Pyrimidin-Regel) [12]. Zur Analyse betrachtet man wieder die kumulierte Reihe, $Y_j = \sum_{i=1}^j (x_i - \bar{x})$, die hier als „DNA-Walk“ bezeichnet wird. Obwohl die Häufigkeit der vier Basen insgesamt etwa gleich ist, gibt es in einzelnen Abschnitten der DNA sehr starke, vor allem stückweise lineare Trends. Daher wurde die Methode der DFA1 auch ursprünglich zur Untersuchung der DNA-Walks entwickelt [4].

Die Erbinformation ist in nur etwa fünf Prozent der Basenpaare gespeichert, die die so genannten kodierenden Bereiche der DNA ausmachen. Die restlichen 95 Prozent werden hingegen nicht benutzt. Bei primitiveren Lebewesen ist der Anteil der kodierenden Bereiche wesentlich höher. Die Korrelationsanalyse an DNA-Datenreihen von verschiedenen Lebewesen ergibt, dass die Basenpaare hauptsächlich in den nicht-kodierenden Bereichen Korrelationen wie in Gl. (1) aufweisen ($\alpha \approx 0,6-0,65$; $\gamma \approx 0,7-0,8$). Die kodierenden Bereiche zeichnen sich meist durch eine völlig regellose Anordnung der Basenpaare (bis auf Trends) aus

($\alpha \approx 0,5$). Auf diesen Unterschieden im Korrelationsverhalten beruht der „Coding Sequence Finder“ Algorithmus von Buldyrev *et al.*, mit dem man in guter Näherung die kodierenden von den nicht-kodierenden Bereichen trennen kann [12]. Kürzlich wurde darüber hinaus nachgewiesen, dass die Korrelationen in der DNA völlig verschiedener Zellen (Eukaryoten und Eubakterien) in gleicher, universeller Weise auftreten, während sie nur bei Archaeobakterien fehlen [13]. Dies deutet in Kombination mit weiteren Ergebnissen darauf hin, dass sie von der dreidimensionalen „Verpackung“ der DNA mit den so genannten Nukleosomen, die bei den Archaeobakterien fehlt, herrühren [13].

Herzschlag

Die Zeitdauern zwischen aufeinanderfolgenden Herzschlägen bestimmt man am besten aus dem Elektrokardiogramm (EKG), bei dem die Aktivität des Herzmuskels elektrisch registriert wird. Die scharfen Maxima im EKG (R-Peaks) zeigen die Kontraktion der Hauptkammern an. Aus dem zeitlichen Abstand zwischen zwei Maxima lassen sich die Herzschlagintervalle (mit einer Genauigkeit von etwa 2 %) bestimmen. Auf den ersten Blick erscheinen die Intervalle, die unsere Zeitreihe x_i definieren, bei gesunden Probanden ungefähr konstant zu sein. Aber eine genauere Betrachtung zeigt, dass ständig Fluktuationen um einen

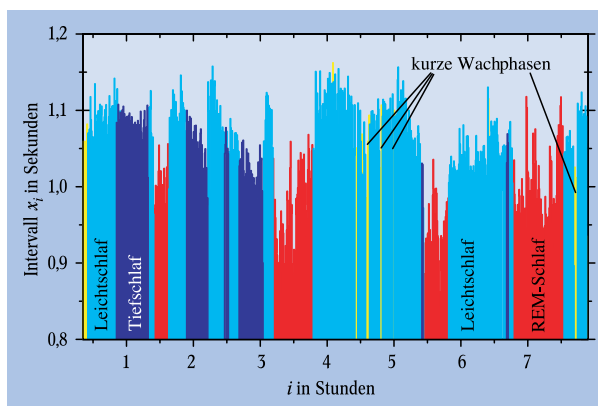


Abb. 6: Darstellung von Herzschlagintervallen und Schlafphasen in einer Nacht für einen typischen gesunden Probanden. Die aufgetragenen Intervalldauern x_i wurden über jeweils 30 Sekunden gemittelt. Die hellblauen Gebiete zeigen den Leichtschlaf an, die dunkelblauen Gebiete den Tiefschlaf und die roten Gebiete den REM-Schlaf. Die kurzen zwischenzeitlichen Wachphasen sind gelb markiert (nach [2]).

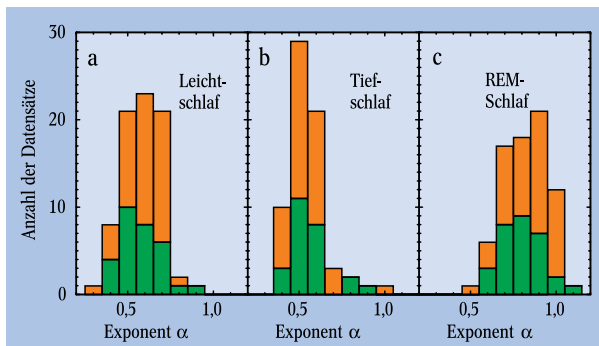
mittleren Abstand von $\bar{x} \approx 0,6$ bis 1,0 Sekunden auftreten. Auch der Mittelwert \bar{x} selbst ist nicht konstant, sondern es sind langsame Trends und andere Nichtstationaritäten wie z. B. Sprünge zu beobachten, die beispielsweise durch unterschiedliche körperliche Belastung oder Aufregung verursacht werden können. Die Frage ist nun, wie sich die Fluktuationen der Intervalle x_i charakterisieren lassen, und was man daraus über die Funktion und den Zustand des Herz-Kreislauf-Systems bzw. eventuelle Erkrankungen lernen kann.

Um die Trends und Nichtstationaritäten im Herzschlag zu vermeiden, die durch willentliche Einflüsse verursacht werden, ist es am geschicktesten, die Fluktuationen während der Schlafphase zu untersuchen. Im Schlaf lassen sich verschiedene Phasen unterscheiden: Leichtschlaf, Tiefschlaf und REM-Schlaf (Abb. 6). Während der Tiefschlaf hauptsächlich der physischen

Erholung dient, benötigt man die REM-Phasen, die ihren Namen von den raschen Bewegungen der (geschlossenen) Augen haben (REM = *rapid eye movement*), hauptsächlich zur psychischen Erholung und Gedächtnisverarbeitung; sie werden häufig mit dem Traumschlaf in Verbindung gebracht. Die Schlafphasen lassen sich in einem medizinischen Schlaflabor bestimmen, indem man die Gehirnströme, die Aktivität bestimmter Muskeln sowie die Augenbewegungen aufzeichnet. Aus den Aufzeichnungen werden dann durch visuelle Untersuchung die Schlafphasen erkannt.

Die Korrelationsanalyse zeigt nun, dass die Herzschlagintervalle in den verschiedenen Schlafphasen einem ganz unterschiedlichen Rhythmus folgen. Während sie im REM-Schlaf stark untereinander korreliert sind, sind sie sowohl im Leicht- als auch im Tiefschlaf oberhalb charakteristischer Zeiten (die näherungsweise dem Atmungszyklus entsprechen) völlig regellos [14].

Abb. 7: Histogramme der DFA2-Fluktuationsexponenten α in den verschiedenen Schlafphasen (a) Leichtschlaf, (b) Tiefschlaf, und (c) REM-Schlaf. Die Datensätze von gesunden Probanden sind in grün eingetragen, und die von Patienten mit schlafbezogenen Atmungsstörungen in orange (nach [14]).



Eine Zusammenfassung der Ergebnisse, die für insgesamt 77 Datensätze von 15 gesunden Probanden und 26 Patienten mit schlafbezogenen Atmungsstörungen mit Hilfe der DFA2 gewonnen wurden, ist in Abb. 7 zu sehen. Die Histogramme der Exponenten α sind im REM-Schlaf bei $\alpha \approx 0,85$, im Tiefschlaf hingegen bei $\alpha = 0,5$ zentriert. Die Tatsache, dass die charakteristischen Zeiten bei Leicht- und Tiefschlaf näherungsweise dem Atmungszyklus entsprechen, ist ein weiteres Indiz für das starke Zusammenspiel zwischen Herzfrequenz und Atmung [15]. Es ist deshalb bemerkenswert, dass bei Patienten mit Atmungsproblemen, die zu einer Fragmentierung des Schlafes führen, die deutlichen Unterschiede zwischen REM- und nicht-REM-Phasen beibehalten werden (Abb. 7).

Die Ergebnisse weisen auch darauf hin, dass das autonome Nervensystem, das u. a. den Herzschlag steuert, nicht so autonom ist, wie man oft denkt. Die Schlafphasen, die im Wesentlichen das Gehirn betreffen, beeinflussen auch den Herzschlag, sodass sich die Regelung im Leicht- und Tiefschlaf deutlich von der im REM-Schlaf unterscheidet. Daher erscheint es durchaus möglich, in Zukunft die Schlafphasen einfach aus dem Herzrhythmus zu bestimmen, ähnlich wie bei der DNA die kodierenden und nicht-kodierenden Teile anhand ihres Korrelationsverhaltens unterschieden werden können. Vielleicht wird es auch in Zukunft möglich sein, aus Störungen des beobachteten Korrelations-Schemas Rückschlüsse auf Herz-Kreislauf-Erkrankungen oder besondere Risiken für zukünftige Erkrankungen zu ziehen.

Zusammenfassung und Ausblick

Wir haben anhand einiger subjektiv ausgewählter Beispiele gezeigt, dass in ganz verschiedenen Systemen in der Natur Langzeitkorrelationen auftreten. Die Liste ist nicht vollständig, so treten Langzeitkorrelationen z. B. auch an der Börse auf, und zwar in den Tages-Handelsvolumina der einzelnen Aktien.

In den Geowissenschaften sind die Konsequenzen der Langzeitkorrelationen besonders gravierend, da sie die relevanten hydrologischen und klimatologischen Fragestellungen unserer Zeit tangieren: Die Eintrittswahrscheinlichkeiten für Extremereignisse können nicht – wie bisher üblich – allein aus dem asymptotischen Verhalten der entsprechenden Verteilungsfunktionen bestimmt werden, und Trends können nicht zuverlässig durch Betrachtung von „Moving Averages“ und lineare Regression abgeschätzt werden. Wir glauben, dass die erforderliche Weiterentwicklung der statistischen Methoden zur Bestimmung von Trends und Wahrscheinlichkeiten für Extremereignisse eine große Herausforderung für die Zukunft darstellt. Positiv ist bereits jetzt zu vermerken, dass die Langzeitkorrelationen in den klimatischen Zeitreihen zum Test von Klimamodellen benutzt werden können und so mit helfen können, die Modelle und damit die Zuverlässigkeit der Klimaprognosen zu verbessern.

Danksagung

Wir möchten uns ganz herzlich bei allen Kollegen und Mitarbeitern bedanken, die durch zahlreiche Diskussionen mit zu diesem Artikel beigetragen haben. Unser besonderer Dank gilt Eva Koscielny-Bunde, Shlomo Havlin, Peter Braun, Plamen Ch. Ivanov, Frank Moss, Thomas Penzel, H. Eduardo Roman, Hans-Joachim Schellnhuber und H. Eugene Stanley.

Literatur

- [1] T. N. Palmer, Rep. Prog. Phys. **63**, 71 (2000).
- [2] The Science of Disaster: Climate Disruptions, Market Crashes, and Heart Attacks, hrsg. von A. Bunde, J. Kropp und H. J. Schellnhuber, Springer, Berlin (2001).
- [3] E. Koscielny-Bunde et al., Physica A **231**, 393 (1996); Phys. Rev. Lett. **81**, 729 (1998); A. Bunde et al., in [2].
- [4] C.-K. Peng et al., Phys. Rev. E **49**, 1685 (1994).
- [5] J. W. Kantelhardt et al., Physica A (im Druck, 2001), Preprint: cond-mat/0102214.
- [6] K. Hasselmann, Science **276**, 914 (1997); H. J. Schellnhuber, Nature **402**, C19 Suppl. (1999); S. Rahmstorf, Nature **399**, 523 (1999); A. Ganopolski und S. Rahmstorf, Nature **409**, 153 (2001); K. Fraedrich et al., Climatic Change **44** (im Druck, 2001).
- [7] Fractals in Science, hrsg. von A. Bunde und S. Havlin, Springer, Berlin (1994).
- [8] R. S. Govindan et al., Preprint (2001).
- [9] H. E. Hurst, R. P. Black und Y. M. Simaika, Long-term storage. An experimental study, Constable, London (1965); B. B. Mandelbrot und J. R. Wallis, Water Resources Research **5**, 321 (1969); s. auch [10].
- [10] J. Feder, Fractals, Plenum Press, New York (1988).
- [11] P. Braun, T. Molnar und H. B. Kleeberg, Hydrol. Processes **11**, 1219 (1997).
- [12] S. V. Buldyrev et al., in [7], und Referenzen darin.
- [13] B. Audit et al., Phys. Rev. Lett. **86**, 2471 (2001).
- [14] A. Bunde et al., Phys. Rev. Lett. **85**, 3736 (2000).
- [15] C. Schäfer et al., Nature **392**, 239 (1998).