

# Increase of extreme events in a warming world

Stefan Rahmstorf<sup>1</sup> and Dim Coumou

Potsdam Institute for Climate Impact Research, PO Box 601203, 14412 Potsdam, Germany

Edited by William C. Clark, Harvard University, Cambridge, MA, and approved September 27, 2011 (received for review February 2, 2011)

**We develop a theoretical approach to quantify the effect of long-term trends on the expected number of extremes in generic time series, using analytical solutions and Monte Carlo simulations. We apply our method to study the effect of warming trends on heat records. We find that the number of record-breaking events increases approximately in proportion to the ratio of warming trend to short-term standard deviation. Short-term variability thus decreases the number of heat extremes, whereas a climatic warming increases it. For extremes exceeding a predefined threshold, the dependence on the warming trend is highly nonlinear. We further find that the sum of warm plus cold extremes increases with any climate change, whether warming or cooling. We estimate that climatic warming has increased the number of new global-mean temperature records expected in the last decade from 0.1 to 2.8. For July temperature in Moscow, we estimate that the local warming trend has increased the number of records expected in the past decade fivefold, which implies an approximate 80% probability that the 2010 July heat record would not have occurred without climate warming.**

global warming | heat wave

In summer 2010, Moscow suffered a heat wave with an average July temperature unprecedented at least since measurements began in the 19th Century (1). Many other unprecedented meteorological extremes have occurred over the past decade (2). At the same time, global-mean temperature in 2010 reached record values (3) and the past decade overall was the warmest decade since measurements began in the 19th century (4). The large number of weather extremes in this exceptionally warm decade raises the question: **How is the number of extremes linked to climatic warming?**

Early theoretical studies on the statistics of record events focused on independent, identically distributed (IID) and thus stationary variables (5, 6). One of the key findings of record theory for IID variables is that the probability of a record is independent of the underlying distribution: It is given by  $1/n$ , where  $n$  is the number of previous data points in the series.

Recent observational studies have shown that, for heat records, the stationarity assumption does not hold. For example, **the observed number of heat records of mean monthly temperatures at stations around the world is significantly higher than that expected in a stationary climate (7).** Also, the ratio of daily record highs to record lows is currently twice as large as what would be expected in a stationary climate in both the United States (8) and Australia (9).

In climate data, nonstationarity can have two principal causes: a shifting mean value or a changing shape of the probability distribution with time (5, 10). For the former, Ballerini and Resnick (11) showed that, for time series with a linear trend, the record rate becomes constant in the asymptotic limit. For Gaussian distributions, the overshoot of a new record over the previous one should become smaller over time (12), both for stationary time series and for time series with a linear trend (12). Further, Krug (5) derived record statistics for time series with a systematic broadening of the probability distribution. In this case, the number of records only increases from the stationary  $1/n$  to  $\ln(n)/n$  in the asymptotic limit (5, 13).

Recently, record statistics of time series with a linear trend have been applied to observed daily temperature maxima (12, 13). For a single weather station (Philadelphia), Redner and Peterson (12) conclude that the current warming rate is insufficient to measurably influence the frequency of daily record temperatures. Using many weather stations in both Europe and the United States, Wergen and Krug (13) find a moderate but significant increase in the number of daily temperature records, in agreement with previous observations (8). They show that this increase can be explained by a systematic change in the mean temperature.

Following on from these studies, **we derive distribution functions for the number of extremes expected in nonstationary climates, both for extremes exceeding a fixed threshold and for record-breaking events** (or “records,” i.e., a value higher—or lower for cold records—than any previous value in the dataset). To analyze the effect of nonlinear climatic trends, we employ Monte Carlo simulations. As two prominent applications, we analyze the global-mean annual-mean temperature series (as the prime metric for global warming) as well as the monthly mean July temperature for Moscow (as an example for a recent record event with major societal impact).

## Statistical Model.

As a first step, we performed Monte Carlo simulations to generate synthetic temperature time series of 100 values each (representing 100 y) as random uncorrelated “noise” with various trends added. Fig. 1A shows one realization of such a series without trend, whereas Fig. 1B shows another realization with an added linear trend of 0.078/y. Temperatures are nondimensionalized by using the standard deviation of their interannual variability as unit (see *Methods*), so this trend implies that the mean temperature (averaged over many realizations) increases by 7.8 standard deviations over the 100 y simulated.

In a stationary climate (Fig. 1A), the number of records of a given type (e.g., heat extremes) declines as  $1/n$ , so the expected number in the whole time series is  $\sum_{n=1}^{100} 1/n = 5.2$ , most of which are found near the beginning of the time series (2.9 in the first decade, 4.5 in the first half). The expected number of records in the last decade is 0.105 because the probability that the highest value occurs is  $1/10$  and that the second-highest value *also* occurs is  $1/100$ . The latter gives a second record in 1 out of 200 realizations, namely, only when the second-highest value precedes the highest value.

The case with warming trend (Fig. 1B) has more unprecedented heat extremes overall, in particular in the last decades of the series. The analytical solution for the expected number of heat records ( $P_{\text{rec}}$ ) in a Gaussian time series with linear long-term trend  $\mu_1$  (derived in *Methods*) is

Author contributions: S.R. designed research; S.R. and D.C. performed research; S.R. analyzed data; and S.R. and D.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. E-mail: stefan@pik-potsdam.de.





thus explains why records occurred only in the 1930s and in the last decade. Fig. 4 clearly shows that the warming trend after 1980 has multiplied the likelihood of a new heat record in Moscow and would have provided a strong reason to expect it before it occurred. Our results thus explicitly contradict those of Dole et al. (16), who did not find any basis for anticipating the Russian heat record of July 2010.

An interesting point to make in the context of Fig. 4 is that, for zero trend, the number of new records declines as  $1/n$  and eventually approaches zero, whereas for a constant warming trend, the number of records approaches a nonzero steady-state level (11) on a timescale roughly equal to the ratio of variability to trend (17). In the example of Fig. 1B, this steady state is reached on a timescale of approximately 15 y. In this steady state, the expected absolute number of records is constant, whereas the fraction attributable to the trend gradually approaches 100%, because the fraction that would have occurred also in a steady climate declines to zero. The strong recent increase in the expected absolute number of records in Moscow shown in Fig. 4 is due to the fact that the temperature evolution there is *not* properly described by a linear 100-y trend but is characterized by a recent warming since the 1980s.

Our analysis of how the expected number of extremes is linked to climate trends does not say anything about the physical causes of the trend. However, the post-1980 warming in Moscow coincides with the bulk of the global-mean warming of the past 100 y, of which approximately  $0.5^{\circ}\text{C}$  occurred over the past three decades (Fig. 1D), most of which the Intergovernmental Panel on Climate Change has attributed to anthropogenic greenhouse gas emissions (4). However, the July warming trend in Moscow over the past three decades was much larger at approximately  $2^{\circ}\text{C}$  (even excluding the extreme 2010 value). A greatly above-average warming trend in the interior of continents is a general feature of both observed warming and of the response to greenhouse gas forcing found in climate models (4). On the other hand, weather stations located in large cities like Moscow are known to be affected by the urban heat island effect (18). We therefore analyzed the warming trend in the surrounding region in microwave sounding unit satellite data (15). These satellite data (running since 1979) show that western Russia has seen a strong July warming of about  $1.4^{\circ}\text{C}$  over this 30 y period (Fig. 5), suggesting that only about one third of the relevant Moscow warming is related to a local urban effect. The fact that observed warming in western Russia is over twice the global-mean warming is consistent with observations from other continental interior areas as well as with model predictions for western Russia under greenhouse gas scenarios (4). Hence, we conclude that the warming trend that has multiplied the likelihood of a new heat record in Moscow is probably largely anthropogenic: a smaller part due to

the Moscow urban heat island, a larger part due to greenhouse warming. That the urban heat island played a smaller role is consistent with the recent finding that the 2010 summer heat wave set a new record not just for Moscow but for Europe (19).

## Discussion

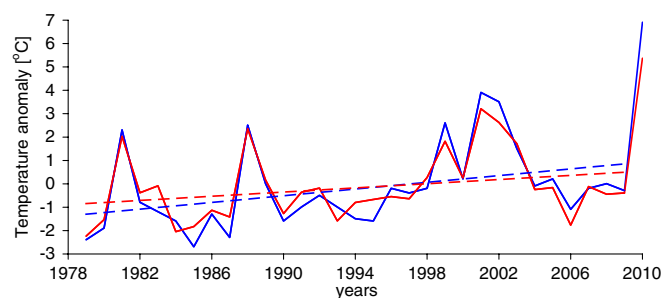
The stochastic model discussed above assumes that the statistical distribution of temperature anomalies does not change but is merely shifted toward warmer temperatures, which holds for the two datasets we analyzed here (see *Methods*). In addition, the distribution can also change itself—possibly it could widen in a warming climate, leading to even more extremes (5, 20). Specific physical processes can alter the distribution of extremes. For example, for the European heat wave of 2003, a feedback with soil moisture has been invoked: Once the soil has dried out, surface temperatures go up disproportionately as less heat is converted to latent heat near the ground; in other words, evaporative cooling becomes ineffective (20).

Although such mechanisms may play an important role and possibly aggravate extremes, it is nevertheless instructive to consider the first-order baseline discussed in this paper, namely, the effect of a simple shift of a random distribution toward warmer temperatures, “all else remaining equal.” Even this simple case demonstrates that large changes in the number of records are expected to arise due to climatic warming. A central result is that the increase in extremes depends on the ratio of climate trend to short-term variability in the data. For the case of records, this dependence is near-linear, whereas for threshold-exceeding extremes it is highly nonlinear. In either case that means that in data with large variability compared to the trend, such as daily station data, the climate-related increase will be relatively small. This theoretical result explains the previous finding that daily data from a single weather station may not yet show a major change in temperature extremes due to global warming (12). In data where the standard deviation has a similar magnitude as the trend, such as monthly mean station data, the expected number of extremes is now several times larger than that in a stationary climate, so the majority of monthly records like the Moscow heat wave must be considered due to the warming trend. In highly aggregated data with small variability compared to the trend, like the global-mean temperature, almost all recent records are due to climate warming. Our statistical method does not consider the causes of climatic trends, but given the strong evidence that most of the warming of the past fifty years is anthropogenic (4), most of the recent extremes in monthly or annual temperature data would probably not have occurred without human influence on climate.

## Methods

**Data Analysis.** Global-average annual data and Moscow station data were obtained from the NASA GISS website (21). Mean July data for Moscow are continuous except for one missing year (1993) which was filled using a neighboring station (Vologda) and its comparison to Moscow in the previous and next year; results are not sensitive to this single year.

A smooth nonlinear climate trend was calculated for these data by using a low-pass filter (22) with half-width of 15 y; we use this to separate the series into a “trend process” and a “noise process” (14). The residual after subtracting this trend is what is defined as “interannual variability” (or noise) for the purpose of this work. This noise is not exactly Gaussian but reasonably close to it (Fig. 3). We have tested that this noise is stationary—i.e., we found no significant changes in its variance over time. If we subdivide the data into four equal intervals, all of these have very similar standard deviations which vary randomly between 1.62 and 1.82 K for the Moscow July data and between 0.079 and 0.099 K for the global annual data. We also tested the interannual variability for serial dependence by computing lagged autocorrelations. For both datasets, there is no statistically significant autocorrelation even for lag 1; the lag-1 autocorrelation coefficients are  $r = -0.04$  for the Moscow data and  $r = 0.17$  for the global data, which both do not differ from the null hypothesis of uncorrelated data at the 95% level. Overall, we conclude that any deviations from stationary, Gaussian white noise in the variability of the datasets we analyze are minor and would not affect our analysis.



**Fig. 5.** Comparison of temperature anomalies from remote sensing systems surface data (red; ref. 15) over the Moscow region ( $35^{\circ}\text{E}$ – $40^{\circ}\text{E}$ ,  $54^{\circ}\text{N}$ – $58^{\circ}\text{N}$ ) versus Moscow station data (blue; ref. 21). The solid lines show the average July value for each year, whereas the dashed lines show the linear trend of these data for 1979–2009 (i.e., excluding the record 2010 value). The satellite data have a trend of  $0.45^{\circ}\text{C}$  per decade for 1979–2009, as compared to  $0.72^{\circ}\text{C}$  per decade for the Moscow station data.



## Corrections

### BIOCHEMISTRY

Correction for “Differential regulation of the activity of deleted in liver cancer 1 (DLC1) by tensin controls cell migration and transformation,” by Xuan Cao, Courtney Voss, Bing Zhao, Tomonori Kaneko, and Shawn Shun-Cheng Li, which appeared in issue 5, January 31, 2012 of *Proc Natl Acad Sci USA* (109:1455–1460; first published January 17, 2012; 10.1073/pnas.1114368109).

The authors note that, due to a printer's error, on page 1455, left column, line 4 of the abstract, “rhodopsin-GTPase-activating (Rho-GAP) activity” should instead appear as “Ras-homology-GTPase-activating protein (Rho-GAP) activity.”

Also, on page 1455, left column, first full paragraph, lines 3–4, “The rhodopsin (Rho) family small GTPases” should instead appear as “The Ras-homology (Rho) family small GTPases.”

[www.pnas.org/cgi/doi/10.1073/pnas.1201491109](http://www.pnas.org/cgi/doi/10.1073/pnas.1201491109)

### NEUROSCIENCE

Correction for “3 $\beta$ -Methoxy-pregnenolone (MAP4343) as an innovative therapeutic approach for depressive disorders,” by Massimiliano Bianchi and Etienne-Emile Baulieu, which appeared in issue 5, January 31, 2012, of *Proc Natl Acad Sci USA* (109:1713–1718; first published January 19, 2012; 10.1073/pnas.1121485109).

The authors note that their conflict of interest statement was omitted during publication. The authors declare that E.-E.B. is president and one of the founders of MAPREG and M.B. is an employee of MAPREG. MAPREG holds a patent on MAP4343 and is currently investigating its use for the treatment of spinal cord injury.

[www.pnas.org/cgi/doi/10.1073/pnas.1202752109](http://www.pnas.org/cgi/doi/10.1073/pnas.1202752109)

### NEUROSCIENCE

Correction for “A complex between contactin-1 and the protein tyrosine phosphatase PTPRZ controls the development of oligodendrocyte precursor cells,” by Smaragda Lamprianou, Elli Chatzopoulou, Jean-Léon Thomas, Samuel Bouyain, and Sheila Harroch, which appeared in issue 42, October 18, 2011, of *Proc Natl Acad Sci USA* (108:17498–17503; first published October 3, 2011; 10.1073/pnas.1108774108).

The authors note that the following statement should be added to the Acknowledgments: “We thank Dr. Barbara Maisson for her help with cell immunostaining.”

[www.pnas.org/cgi/doi/10.1073/pnas.1202480109](http://www.pnas.org/cgi/doi/10.1073/pnas.1202480109)

### GEOPHYSICS

Correction for “Increase of extreme events in a warming world,” by Stefan Rahmstorf and Dim Coumou, which appeared in issue 44, November 1, 2011, of *Proc Natl Acad Sci USA* (108:17905–17909; first published October 24, 2011; 10.1073/pnas.1101766108).

The authors note that on page 17906, left column, the equation

$$P_{\text{rec}}(t_n) = \int_{-\infty}^{\infty} \left[ \frac{1}{2} - \frac{1}{2} \operatorname{erf} \left( \frac{x - \mu_0 - \mu_1 t_n}{\sigma \sqrt{2}} \right) \right] \prod_{i=1}^{n-1} \left[ \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{x - \mu_0 - \mu_1 t_i}{\sigma \sqrt{2}} \right) \right] dx$$

should instead appear as

$$P_{\text{rec}}(t_n) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x - \mu_0 - \mu_1 t_n)^2}{2\sigma^2}} \prod_{i=1}^{n-1} \left( \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{x - \mu_0 - \mu_1 t_i}{\sigma \sqrt{2}} \right) \right) dx.$$

On page 17909, right column, first full paragraph, lines 4–6: “Second of all, the event needs to actually occur at timestep  $t_n$ , the probability of which is given by  $f(x, t_n)$ ,” should instead appear as: “Second of all, the event needs to actually occur at timestep  $t_n$ , the probability of which is given by  $p(x, t_n)$ .” Additionally, on page 17909, right column, the equation

$$P_{\text{rec}}(t_n) = \int_{-\infty}^{\infty} f(x, t_n) \prod_{i=1}^{n-1} (1 - f(x, t_i)) dx$$

should instead appear as

$$P_{\text{rec}}(t_n) = \int_{-\infty}^{\infty} p(x, t_n) \prod_{i=1}^{n-1} \left( \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{x - \mu_0 - \mu_1 t_i}{\sigma \sqrt{2}} \right) \right) dx.$$

These errors do not affect the conclusions of the article.

[www.pnas.org/cgi/doi/10.1073/pnas.1201163109](http://www.pnas.org/cgi/doi/10.1073/pnas.1201163109)