

RESEARCH LETTER

10.1002/2015GL066307

Key Points:

- Conventional normalization of spatiotemporal data sets with respect to a reference period induces artifacts
- Normalization-induced artifacts are most severe if variability or extremes are under scrutiny
- The study provides an analytical correction and accurate estimate of variability and extremes

Supporting Information:

- Supporting Information S1

Correspondence to:

S. Sippel,
ssippel@bgc-jena.mpg.de

Citation:

Sippel, S., J. Zscheischler, M. Heimann, F. E. L. Otto, J. Peters, and M. D. Mahecha (2015), Quantifying changes in climate variability and extremes: Pitfalls and their overcoming, *Geophys. Res. Lett.*, 42, 9990–9998, doi:10.1002/2015GL066307.

Received 23 SEP 2015

Accepted 3 NOV 2015

Accepted article online 6 NOV 2015

Published online 25 NOV 2015

Quantifying changes in climate variability and extremes: Pitfalls and their overcoming

Sebastian Sippel^{1,2}, Jakob Zscheischler^{1,2}, Martin Heimann¹, Friederike E. L. Otto³, Jonas Peters⁴, and Miguel D. Mahecha^{1,5}
¹Max Planck Institute for Biogeochemistry, Jena, Germany, ²Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland, ³Environmental Change Institute, University of Oxford, Oxford, UK, ⁴Max Planck Institute for Intelligent Systems, Tübingen, Germany, ⁵German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Leipzig, Germany

Abstract Hot temperature extremes have increased substantially in frequency and magnitude over past decades. A widely used approach to quantify this phenomenon is standardizing temperature data relative to the local mean and variability of a reference period. Here we demonstrate that this conventional procedure leads to exaggerated estimates of increasing temperature variability and extremes. For example, the occurrence of “two-sigma extremes” would be overestimated by 48.2% compared to a given reference period of 30 years with time-invariant simulated Gaussian data. This corresponds to an increase from a 2.0% to 2.9% probability of such events. We derive an analytical correction revealing that these artifacts prevail in recent studies. Our analyses lead to a revision of earlier reports: For instance, we show that there is no evidence for a recent increase in normalized temperature variability. In conclusion, we provide an analytical pathway to describe changes in variability and extremes in climate observations and model simulations.

1. Introduction

Quantifying to what extent the magnitude and frequency of extreme events are changing is a priority in climate change research [International Panel on Climate Change (IPCC), 2012; Seneviratne et al., 2014]. In recent years, unusually hot temperature extremes have occurred and these events are increasingly exceeding the range of historical variability [Rahmstorf and Coumou, 2011; Mora et al., 2013]. Considerable scientific debate has sparked around whether present-day changes in extreme events are due to the shifting mean climatology or whether we are also confronted with changing variability [Hansen et al., 2012; Huntingford et al., 2013; Alexander and Perkins, 2013; Mora et al., 2013; Seneviratne et al., 2014]. Of particular focus in this context are changes in temperature extremes, which have direct impacts upon human wellbeing and likewise affect ecosystem services and global biogeochemical cycles [IPCC, 2012; Reichstein et al., 2013].

A widely used approach to address this question relies on normalizing climate data relative to a reference period [Hansen et al., 2012; Coumou and Robinson, 2013; Huntingford et al., 2013; Kamae et al., 2014; Curry et al., 2014] aiming to objectively compare temperature variability and extremes across space and time. This approach conventionally derives standardized anomalies by locally subtracting the mean (μ_{ref}) from and dividing the observations by the standard deviation (σ_{ref}) estimated from some reference period:

$$z = \frac{X - \mu_{\text{ref}}}{\sigma_{\text{ref}}} \quad (1)$$

The idea is to rank or count events based on departures from the local climatology (as defined by the reference period) in units of standard deviation (σ). Transformations of this kind underpin studies of changes in the occurrence of monthly or seasonal temperature extremes [Hansen et al., 2012; Coumou and Robinson, 2013; Kamae et al., 2014; Curry et al., 2014] and variability [Huntingford et al., 2013]. Further, so-derived standardized anomalies have been used to determine continental-scale rankings of the most significant meteorological or geophysical extreme events [Grumm and Hart, 2001; Hart and Grumm, 2001; Root et al., 2007; Graham and Grumm, 2010], and Kodra and Ganguly [2014] study asymmetry in the distributions of temperature extremes using a variant of this methodology.

In this paper, we demonstrate that this conventional normalization procedure inevitably leads to erroneous and exaggerated estimates of temperature extremes and variability outside a specified “reference period.”

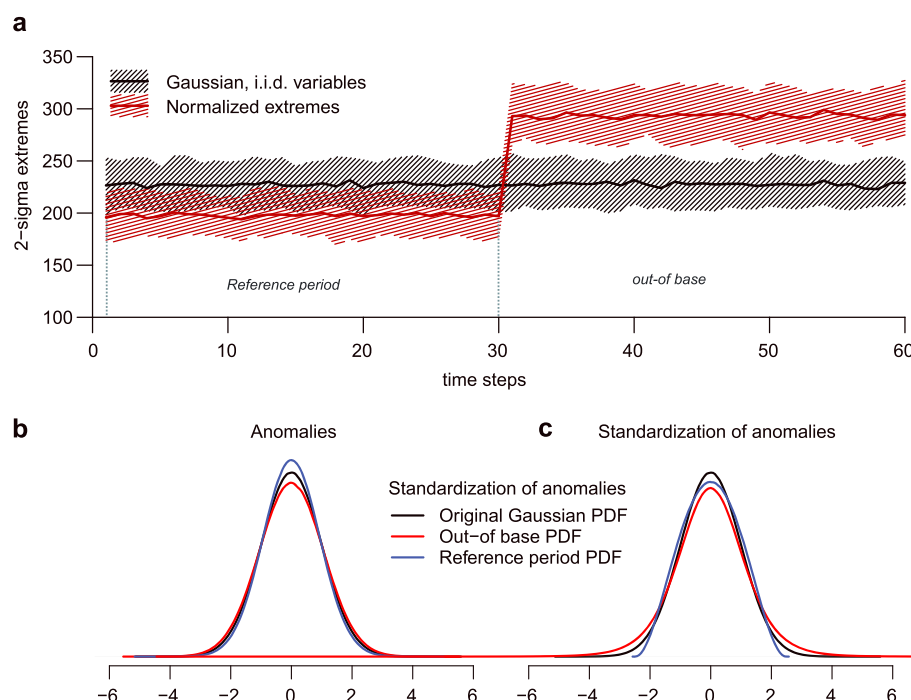


Figure 1. Biases in the detection of extreme events in stationary and independent Gaussian data induced by normalization. (a) Occurrences of positive two-sigma extremes in artificial Gaussian time series based on 10,000 replicates over 60 time points before normalizing the data (black line) and after normalizing each replicate using the first 30 samples as reference period. (b) Illustration of variance inflation and reduction through the generation of anomalies in the out-of-base (blue) versus reference period (red) probability density function (PDF) ($n_{\text{ref}} = 8$ for illustration). (c) Changing tails in normalized (i.e., divided by the SD estimate) Gaussian variables ($n_{\text{ref}} = 8$ for illustration). Colored shading in Figure 1a indicates the 5th to 95th percentile in repeated simulations.

Furthermore, we derive an analytical correction that accounts for these statistical artifacts and allows for an accurate quantification of large-scale climate variability and extremes.

2. Methodology and Results

2.1. Normalization-Induced Artifacts and an Analytical Correction for Quantifying Extremes

To test the suitability of the reference period normalization, we conduct Monte Carlo simulations with independent and identically distributed random variables drawn from a standard Gaussian distribution ($\mathcal{N}(\mu = 0, \sigma^2 = 1)$). This numerical experiment is set up in analogy to investigations of monthly or seasonally standardized extremes (see Hansen *et al.* [2012], for an example) in gridded temperature data with $k = 10^4$ time series ("grid cells") and $n = 60$ data points per time series ("years of data") but consisting of purely random Gaussian variables (independent and identically distributed (i.i.d.)). For each time series we generate anomalies and subsequently standardize these based on the conventional procedure (equation (1)). Both mean ($\hat{\mu}_{\text{ref}}$) and standard deviation ($\hat{\sigma}_{\text{ref}}$) are estimated from each time series' first 30 values (i.e., $n_{\text{ref}} = 30$). The number of values exceeding a given σ threshold (" σ extremes") are counted at each time step in the original and normalized data set (Figure 1, grey and red lines, respectively).

Given that the statistical properties of the artificial data are time invariant, there should be no change in the number of extremes across the data set. However, in fact, we find substantial increases in the number of extreme events outside the reference period along with a reduction in extremes within the reference period (Figure 1a, R code to reproduce these results in Text S1 in the supporting information). A quantification of 2σ extremes across all grid cells in the artificial data set leads to a considerable increase (red line in Figure 1a) in the out-of-base period relative to the reference period of about 48.2%. Considering only the out-of-base period, the number of 2σ (3σ) events would be overestimated by 29.1% (131.0%) relative to the original Gaussian data (black line in Figure 1a), which corresponds to an increase from a 2.3% (1.3‰) chance to 2.9% (3.1‰). For illustration purposes, the distributions at a random time step inside and outside the reference period across all time series is shown in Figures 1b and 1c for anomalies and standardized

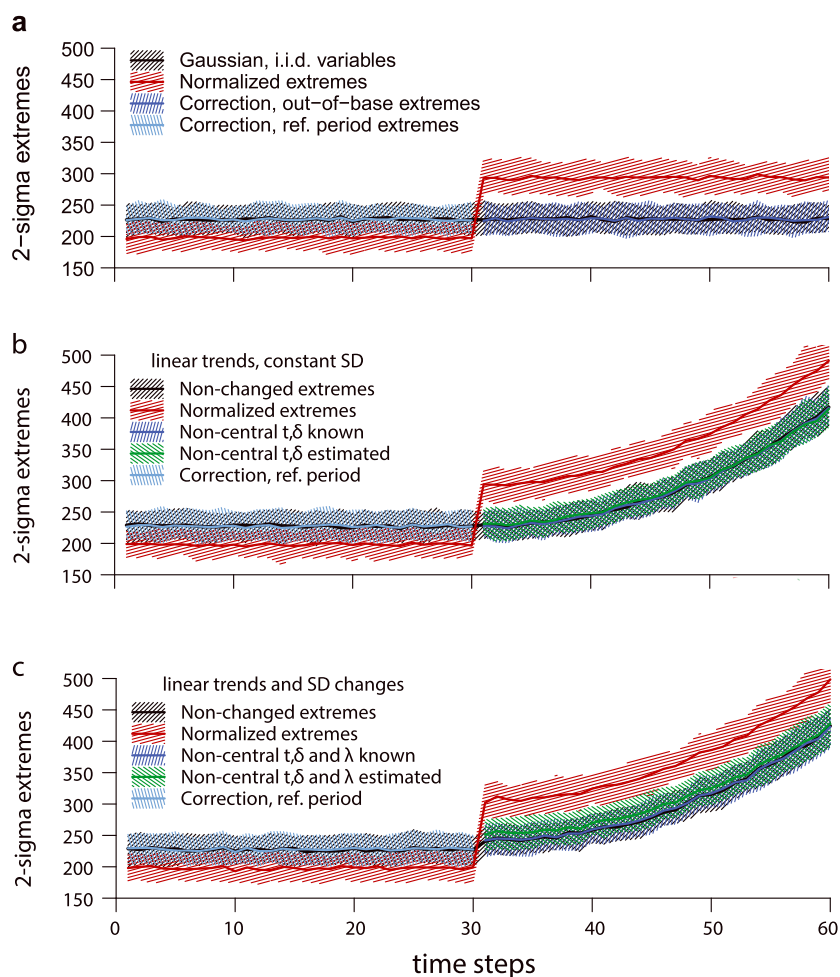


Figure 2. Correction of normalization-induced biases in stationary and nonstationary time series consisting of independent random variables. Detecting two-sigma extreme events in (a) stationary Gaussian time series, (b) Gaussian time series with random linear trends added in the out-of-base period ($-1 < \delta_{t=60} < 1$, in units of σ), and (c) Gaussian time series with random linear trends ($-1 < \delta_{t=60} < 1$, in units of σ) and changing variance ($0.8\sigma_{\text{ref}} < \lambda\sigma_{\text{ref}} < 1.2\sigma_{\text{ref}}$) in the out-of-base period. In each panel, colored shading indicates the 5th to 95th percentile in repeated simulations ($k = 10^4$ simulated time series in all panels).

variables, respectively. Overall, the artificial experiment reveals potentially severe artifacts in the widely applied reference period normalization. In the following paragraphs, we reveal the consequences of this conventional normalization and derive an analytical solution for the induced artifacts.

To understand the origin of the apparent increase in extremes, we have to consider that the “true” values for mean and variability are inherently unknown, which changes equation (1) to

$$z = \frac{X - \hat{\mu}_{\text{ref}}}{\hat{\sigma}_{\text{ref}}}. \quad (2)$$

The estimates of the mean ($\hat{\mu}_{\text{ref}}$) and standard deviation ($\hat{\sigma}_{\text{ref}}$) are random variables with well-known statistical properties [Von Storch and Zwiers, 2001], drawn from an independent sample in case of analyzing the out-of-base period [Zhang et al., 2005] (see Text S2 for a detailed statistical description) and subsequently pooled in space. Consequently, the biases between both periods are induced by a combination of two effects; first, the generation of anomalies ($X_{\text{anom}} = X - \hat{\mu}_{\text{ref}}$) and second, the standardization ($z = \frac{X_{\text{anom}}}{\hat{\sigma}_{\text{ref}}}$) (Figures 1b and 1c): The generation of anomalies systematically increases (decreases) the variance across grid cells in the out-of-base (reference) period [Tingley, 2012] but does not affect the underlying distribution (Text S2). However, the local standardization of each time series induces qualitative changes to the (spatial) distribution (for an analytical derivation see Text S2) such that heavier tails outside the reference period are induced

(Figure 1c). This qualitative difference stems from the fact that any time point in the out-of-base period follows a t -distribution with $n_{\text{ref}} - 1$ degrees of freedom (Text S2). Hence, the heavier tails generated by the conventional standardization lead to a consistent and potentially severe overestimation of extreme events in the out-of-base period (Figure 1a) for relatively short, but in practice often used, sometimes unavoidable, reference periods. However, the distribution after normalization can be derived analytically (Text S2), and hence, the biases can be rectified separately both for the reference and the out-of-base periods. Specifically, instead of counting 2σ (3σ) extremes in the out-of-base period, a search for the corresponding percentile threshold in the variance-adjusted t -distribution (2.12σ (3.32σ), respectively, if $n = 30$) would allow for the detection of the correct number of events (Figures 2a and S1 for an illustration of the correction procedure). Further, it is worth noting that even with an increasing number of samples in the reference period, the convergence to small biases is slow. For autocorrelated data the artifacts are even more pronounced owing to a smaller effective sample size (Figures S2a and S2b, respectively).

Before applying the proposed analytical correction we have to consider that temperatures at monthly or seasonal time scales are typically nonstationary [Ji *et al.*, 2014], i.e., simulated or observed time series might contain spatially and temporarily diverse trends. Using Monte Carlo type simulations of normalized Gaussian time series with changing trends and variability, we find that both exert strong influence on the magnitude of the biases (Text S3). Increasing (decreasing) trends or variability in the out-of-base period severely deflates (inflates) the biases for the upper tail (Figures S2a and S2b). These insights are equally applicable to the lower tail of the distribution if the sign of the trend is reversed. To assess the issue of nonstationarity in more detail, we consider trends and changes in variability in the artificial data set introduced in Figure 1. First, random linear trends are added in the out-of-base period to each random Gaussian time series, where the magnitudes of the trends at the last time step are drawn randomly for each grid cell from a uniform distribution in the interval $[-1 \leq \delta \leq 1]$ in units of σ (Figure 2b). Second, we investigate a trend in the out-of-base period coinciding with randomly assigned changes in variability ($0.8 \leq \sigma \leq 1.2$, Figure 2c).

Following the solution for stationary time series outlined above, we offer an analytical correction that allows handling of the additional artifacts induced by nonstationarities (Text S4). In essence, normalizing nonstationary data induces a noncentral version of Student's t -distribution. This analytical distribution can be used to avoid normalization-induced biases entirely if changes in the trend or variability are known (Figures 2b and 2c). Likewise, estimating the trend and/or changes in variability largely allows for removing the biases (Figures 2b and 2c). As above, σ extremes are counted based on the biased estimate of the conventional procedure (red line) and based on the application of the suggested correction procedure using known (blue) and estimated (green) trends and changes in variability. Throughout this paper, singular spectrum analysis (SSA), a nonlinear spectral decomposition methodology [Golyandina and Zhigljavsky, 2013; von Buttler *et al.*, 2014] is used to estimate trend components, before the analytical correction procedure based on the noncentral t -distribution is applied. Trends are extracted as 31 years and larger components using a 45 year SSA window length ($L = 45$).

2.2. Quantifying Extremes in Earth Observation Data

In this subsection, we assess how monthly temperature extremes on land have changed over the second half of the twentieth century in the Northern Hemisphere up to present by applying the statistical approach outlined above. In order to avoid potential inhomogeneities related to gridded observations, we analyze the state-of-the-art Twentieth Century Reanalysis data set [Compo *et al.*, 2011] (version 2). The reanalysis data set assimilates only surface pressure measurements and monthly sea surface temperatures into an atmosphere and land general circulation model [Compo *et al.*, 2011] and is hence independent from station temperature measurements. The data set has been specifically designed to assess climate variability and extremes statistics "spanning the instrumental record" and has been demonstrated to reproduce the observed temperature trends and variability to a very large extent [Compo *et al.*, 2011].

In our analysis, we first interpolate the data set to a $2^\circ \times 2^\circ$ regular latitude-longitude grid and mask ocean pixels. Second, we estimate separately for each month and grid cell the trend component, local mean, and (nondetrended and detrended) standard deviation in two different reference periods (1921–1950 and 1951–1980). Third, each pixel time series is normalized using both reference periods and the detrended and nondetrended σ_{ref} estimates. For each month we calculate the area affected by 2σ and 3σ extremes, using the conventional normalization approach and our correction. We use the trend estimates for our correction but assume an approximately unchanged variance over the past decades [Huntingford *et al.*, 2013]. Last, we derive

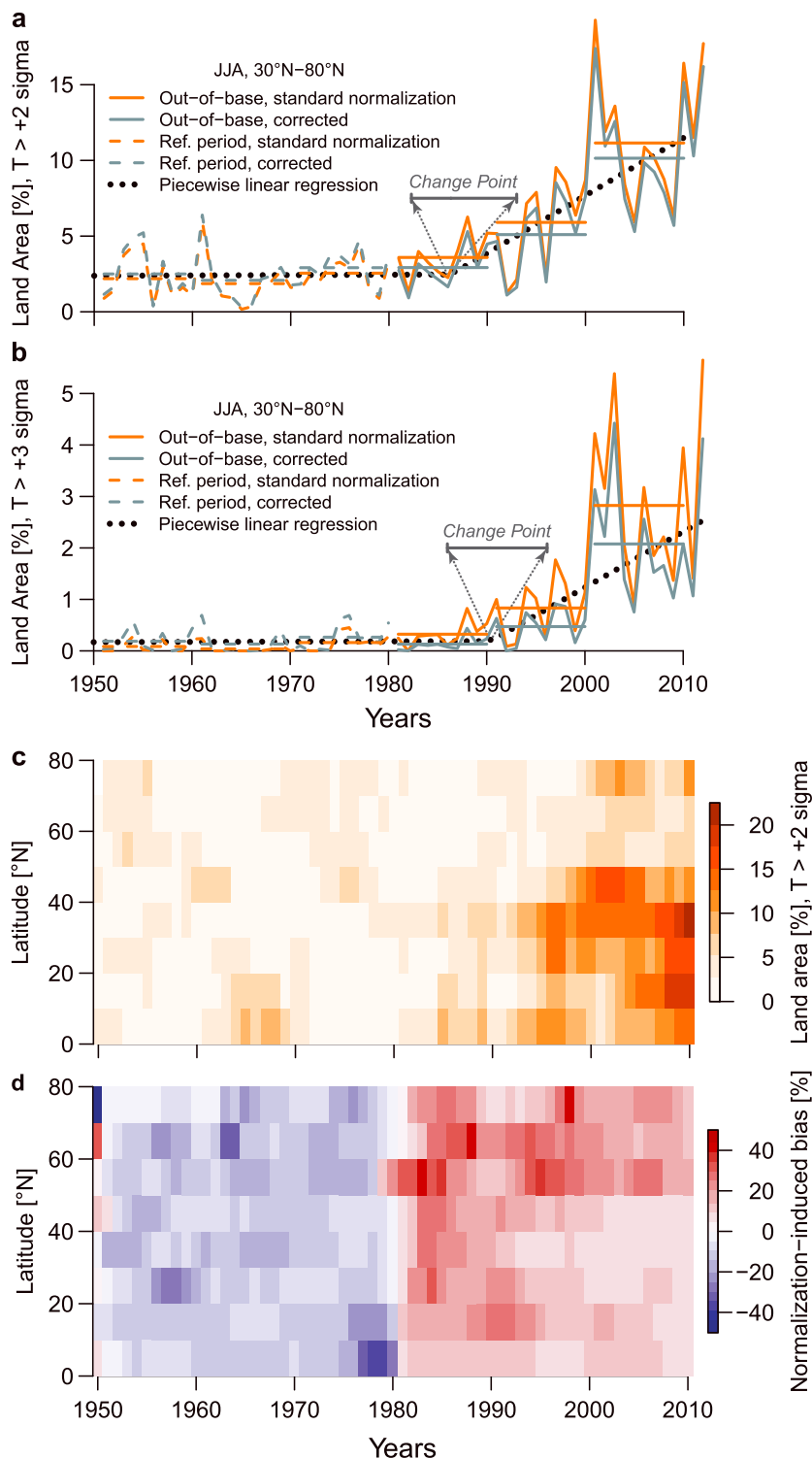


Figure 3. Increase in normalized hot temperature extremes in a spatiotemporal data set (Twentieth Century Reanalysis). Time series of fraction of extratropical Northern Hemisphere land area covered by positive monthly (a) 2σ and (b) 3σ events in summer (reference period: 1951–1980). Horizontal lines indicate decadal averages for the conventional normalization procedure (light blue) and our proposed correction (orange). (c) Zonal evolution of fraction of land area covered by monthly positive 2σ extremes in Northern Hemisphere summer. (d) Zonal evolution of relative biases induced by the conventional normalization approach.

seasonal averages of the “area affected by extremes” for Northern Hemisphere summer (June, July, and August, Figure 3).

Our analysis reveals that the exceedance of monthly 2σ and 3σ temperature extremes in summer has indeed increased substantially over the Northern Hemisphere (Figures 3a and 3b for land areas in the Northern Hemisphere outer tropics). However, the bias-adjusted time series show a consistently slower and smoother increase as compared to the conventionally applied uncorrected normalization procedure. A breakpoint analysis using piecewise linear regression [Toms and Lesperance, 2003] based on our revised figures indicates that the recent rapid increase in hot summer months in the Northern Hemisphere (2σ and 3σ events) started to emerge around the late 1980s or early 1990s (Figure 3b).

The magnitude of the biases and the discontinuities at the reference and out-of-base period are robust across different reference periods and also hold if trends are subtracted before estimating local variability [Coumou and Robinson, 2013] (Figures S3 and S4). Increases in extremes relative to local variability show a clear zonal pattern (Figure 3c) with the largest increases in the tropics and subtropics. Therefore, biases induced by the normalization are largest in areas where the trend is relatively small compared to local variability (Figure 3d). However, it is worth noting that peculiarities of the station-based observational record such as urban heat islands or local land use changes are not accounted for in the Twentieth Century Reanalysis [Parker, 2011]. In addition, the availability of pressure observations varies through time [Compo et al., 2011]. As such, the main purpose of the present analysis is to illustrate the potential biases induced by reference period standardization in spatiotemporal data sets.

2.3. Implications for Large-Scale Assessments of Variability and Asymmetry

Normalization-induced biases are not only relevant for assessments of extremes, but a careful consideration of such statistical preprocessing techniques is equally important for analysis of variability and asymmetry in spatiotemporal data sets. An example is provided by a recent study that investigated whether temperature variability has changed over the second half of the twentieth century on global and continental scales [Huntingford et al., 2013]. The authors argue that annual temperatures in low-variance regions have become more variable over the past decades, while global temperature variability has remained near constant. This explanation stems from the authors' observation that normalized variability has increased more than absolute (spatial) variability (16% versus 2% increases between 1963–1980 and 1981–1996). Using the Twentieth Century Reanalysis data set, we reproduce the increases in the annual, global, area-weighted standard deviation (12.9% versus 1.8% increases, when using the conventional data processing scheme [Huntingford et al., 2013], Figure 4).

However, an artificial experiment in analogy to the previous subsection shows that the conventional normalization procedure changes the standard deviation of the data (Figure 4a) and in particular yields an increase in standard deviation between the reference and the out-of-base period. Therefore, we correct the conventionally normalized standard deviation of annual temperatures in the Twentieth Century Reanalysis data set empirically and analytically. The former is achieved by simulating the reduction in standard deviation in artificial Gaussian data (Figure 4a), whereas the latter is achieved by using an earlier reference period (1921–1950) and the application of our analytical correction. The empirical and analytical corrections reduce the increase in normalized variability from 12.9% to 5.6% and 6.0%, respectively (see Figure 4b). A permutation-based significance test [Fay and Shaw, 2010] shows that the increases in mean corrected normalized standard deviation between both periods are not significant ($p_{\text{empirical}} = 0.147$ and $p_{\text{analytical}} = 0.110$), whereas conventional normalization yields a highly significant increase ($p_{\text{conventional}} = 0.004$). Hence, the relatively small and non-significant difference between the increases in standardized and absolute variability might indeed be due to the explanation offered previously [Huntingford et al., 2013] and potentially related to major El Niño events in the latter period [Fedorov and Philander, 2000]. If the periods before and after 1980 are extended to derive a larger sample, this reduces the increase in normalized variability to only 2% (1981–2006 versus 1955–1980). Thus, based on our proposed normalization, we cannot confirm that changes across low-variance regions have occurred over the past decades. Nonetheless, our results underpin that global temperature variability has not changed [Huntingford et al., 2013] and additionally show that this finding holds both in absolute and normalized terms.

Finally, another recent study [Kodra and Ganguly, 2014] reports that asymmetry in temperature distributions of seasonal extreme values at daily time scale (both minima and maxima, i.e., the hottest and coldest day per season) is strongly increasing toward both the cold and hot tails in model projections of future climate

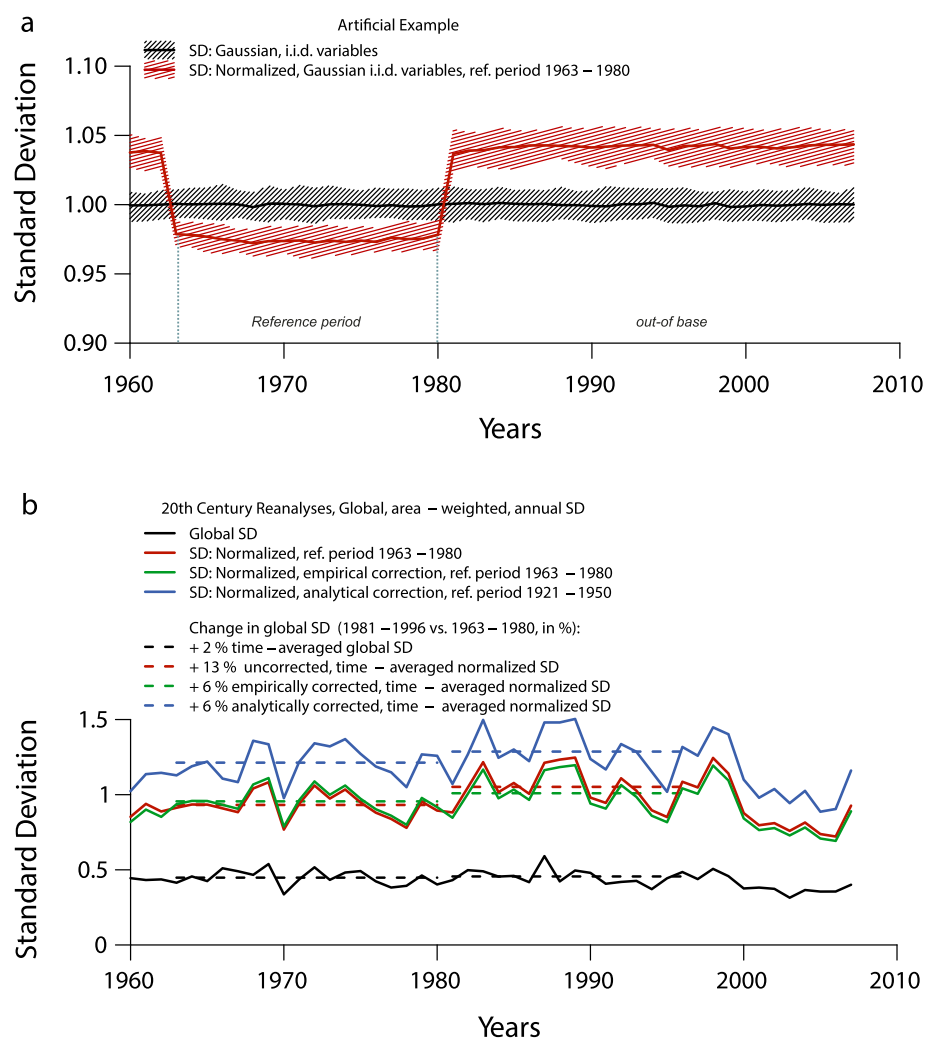


Figure 4. Normalization-induced changes in variability. Time series of normalized variability following the data processing scheme of *Huntingford et al.* [2013] in an artificial example ($k = 10^4$ time series) with (a) i.i.d. Gaussian variables and (b) in the Twentieth Century Reanalysis data set.

conditions relative to a recent period. As a preprocessing step, the authors derive “anomalies” of seasonal extremes by subtracting the mean of the recent (historical) climatology of seasonal extremes from both periods. This procedure leads to narrower distributions in the reference period and a broader distribution in the future (independent) period (see Text S2). This variance inflation in skewed extreme value distributions leads to the observed effect even in stationary time series and should hence be interpreted with caution (Figure S5 and Text S6).

3. Outlook and Conclusion

The observation that a commonly used normalization of temperature data is inappropriate for assessing changes in variability, extremes, and asymmetry is of general validity and should also be considered in investigations of other climatological and Earth observations. The steadily growing archives of Earth observations derived from both ground-based and satellite remote sensing data require reconsidering conventional data analytic approaches such as standardization. For instance, extremes in gridded standardized anomalies of rainfall and storms [Grumm and Hart, 2001; Hart and Grumm, 2001; Root et al., 2007; Graham and Grumm, 2010; Curry et al., 2014] have been studied using varieties of the conventional standardization procedure and are potentially distorted by the artifacts discussed in this paper. Further, our results might facilitate the interpretation of single climatic extreme events or trends that are frequently characterized in terms of standardized departure from climatology, both inside and/or outside the climatological reference period [Schär et al., 2004;

Barriopedro *et al.*, 2011; Xu *et al.*, 2012; Ramos *et al.*, 2014; Cook *et al.*, 2015]. Although our analytical treatment using the *t*-distribution is confined to distributions that can be approximated as Gaussian, we emphasize that the induction of biases in the tails due to dependent/independent estimators of location and scale is fundamental and holds indeed across a wide range of distributions. Furthermore, because temperature extremes are bounded [Nogaj *et al.*, 2006], approximations of temperature values by distributions with infinite tails (such as Gaussian and the *t*-distribution) might poorly estimate the most extreme temperatures. Here we offer a correction which adjusts biases in variability and extremes induced by a widely used data preprocessing approach. Alternatively, statistically more advanced but readily available tools, such as the theory of extreme values [Katz *et al.*, 2013; Nogaj *et al.*, 2006], offer complementary approaches to quantify extreme events under nonstationary conditions that are not affected by the statistical issues reported in this paper.

In conclusion, data normalization for the detection of changes in extremes or variability has to be applied with caution: otherwise, there is a risk to arbitrarily inflate both extremes and variability in the time periods under scrutiny. Our study demonstrates how to avoid biases of this kind. However, our analyses do not call into question the major qualitative results that were outlined in previous studies [Hansen *et al.*, 2012; Seneviratne *et al.*, 2014]: hot temperature extremes have increased considerably on the global scale, a trend which is most likely to continue throughout the 21st century [Coumou and Robinson, 2013; Sillmann *et al.*, 2013].

Acknowledgments

We thank Fabian Gans, Jannis von Buttler, Jens Schumacher, Markus Reichstein, Sonia Seneviratne, Myles Allen, and two anonymous referees for their helpful comments that strongly helped to improve the manuscript, and we are grateful to Silvana Schott for the graphical support. This study was supported by the European Space Agency ESA support to science element CAB-LAB: Coupled Atmosphere Biosphere Virtual Laboratory and the European Commission project BACI "Detecting changes in essential ecosystem and biodiversity properties towards a Biosphere Atmosphere Change Index" (grant ID 640176). S.S. is grateful to the 'German National Academic Foundation' (Studienstiftung des Deutschen Volkes) for the support and the International Max Planck Research School for Global Biogeochemical Cycles (IMPRS-gBGC) for the training. Support for the Twentieth Century Reanalysis Project data set (http://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2.html) is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office.

References

- Alexander, L., and S. Perkins (2013), Debate heating up over changes in climate variability, *Environ. Res. Lett.*, **8**(4), 041001.
- Barriopedro, D., E. Fischer, J. Luterbacher, R. Trigo, and R. García-Herrera (2011), The hot summer of 2010: Redrawing the temperature record map of Europe, *Science*, **332**(6026), 220–224.
- Compo, G. P., et al. (2011), The twentieth century reanalysis project, *Q. J. R. Meteorol. Soc.*, **137**(654), 1–28.
- Cook, B., T. Ault, and J. Smerdon (2015), Unprecedented 21st century drought risk in the American Southwest and Central Plains, *Sci. Adv.*, **1**(1), e1400082.
- Coumou, D., and A. Robinson (2013), Historic and future increase in the global land area affected by monthly heat extremes, *Environ. Res. Lett.*, **8**(3), 034018.
- Curry, C., et al. (2014), A multimodel examination of climate extremes in an idealized geoengineering experiment, *J. Geophys. Res. Atmos.*, **119**, 3900–3923, doi:10.1002/2013JD020648a.
- Fay, M. P., and P. A. Shaw (2010), Exact and asymptotic weighted logrank tests for interval censored data: The interval R package, *J. Stat. Software*, **36**(2), 1–34.
- Fedorov, A. V., and S. G. Philander (2000), Is El Niño changing?, *Science*, **288**(5473), 1997–2002.
- Golyandina, N., and A. Zhigljavsky (2013), *Singular Spectrum Analysis for Time Series*, Springer, Berlin.
- Graham, R. A., and R. H. Grumm (2010), Utilizing normalized anomalies to assess synoptic-scale weather events in the Western United States, *Weather Forecasting*, **25**(2), 428–445.
- Grumm, R. H., and R. Hart (2001), Standardized anomalies applied to significant cold season weather events: Preliminary findings, *Weather Forecasting*, **16**(6), 736–754.
- Hansen, J., M. Sato, and R. Ruedy (2012), Perception of climate change, *Proc. Natl. Acad. Sci. U.S.A.*, **109**(37), E2415–E2423.
- Hart, R. E., and R. H. Grumm (2001), Using normalized climatological anomalies to rank synoptic-scale events objectively, *Mon. Weather Rev.*, **129**(9), 2426–2442.
- Huntingford, C., P. D. Jones, V. N. Livina, T. M. Lenton, and P. M. Cox (2013), No increase in global temperature variability despite changing regional patterns, *Nature*, **500**(7462), 327–331.
- International Panel on Climate Change (IPCC) (2012), Summary for policymakers, in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*, edited by C. B. Field et al., pp. 1–19, Cambridge Univ. Press, Cambridge, U. K., and New York.
- Ji, F., Z. Wu, J. Huang, and E. P. Chassignet (2014), Evolution of land surface air temperature trend, *Nat. Clim. Change*, **4**, 462–466.
- Kamae, Y., H. Shiogama, M. Watanabe, and M. Kimoto (2014), Attributing the increase in Northern Hemisphere hot summers since the late 20th century, *Geophys. Res. Lett.*, **41**, 5192–5199, doi:10.1002/2014GL061062.
- Katz, R. W., P. F. Craigmile, P. Guttorp, M. Haran, B. Sanso, and M. L. Stein (2013), Uncertainty analysis in climate change assessments, *Nat. Clim. Change*, **3**(9), 769–771.
- Kodra, E., and A. R. Ganguly (2014), Asymmetry of projected increases in extreme temperature distributions, *Sci. Rep.*, **4**, 5884.
- Mora, C., et al. (2013), The projected timing of climate departure from recent variability, *Nature*, **502**(7470), 183–187.
- Nogaj, M., P. Yiou, S. Parey, F. Malek, and P. Naveau (2006), Amplitude and frequency of temperature extremes over the North Atlantic region, *Geophys. Res. Lett.*, **33**, L10801, doi:10.1029/2005GL024251.
- Parker, D. E. (2011), Recent land surface air temperature trends assessed using the 20th Century Reanalysis, *J. Geophys. Res.*, **116**, D20125, doi:10.1029/2011JD016438.
- Rahmstorf, S., and D. Coumou (2011), Increase of extreme events in a warming world, *Proc. Natl. Acad. Sci. U.S.A.*, **108**(44), 17,905–17,909.
- Ramos, A. M., R. M. Trigo, and M. L. R. Liberato (2014), A ranking of high-resolution daily precipitation extreme events for the Iberian Peninsula, *Atmos. Sci. Lett.*, **15**(4), 328–334.
- Reichstein, M., et al. (2013), Climate extremes and the carbon cycle, *Nature*, **500**(7462), 287–295.
- Root, B., P. Knight, G. Young, S. Greybush, R. Grumm, R. Holmes, and J. Ross (2007), A fingerprinting technique for major weather events, *J. Appl. Meteorol.*, **46**(7), 1053–1066.
- Schär, C., P. Vidale, D. Lüthi, C. Frei, C. Häberli, M. Liniger, and C. Appenzeller (2004), The role of increasing temperature variability in European summer heatwaves, *Nature*, **427**(6972), 332–336.
- Seneviratne, S. I., M. G. Donat, B. Mueller, and L. V. Alexander (2014), No pause in the increase of hot temperature extremes, *Nat. Clim. Change*, **4**(3), 161–163.

- Sillmann, J., V. Kharin, F. Zwiers, X. Zhang, and D. Bronaugh (2013), Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, *J. Geophys. Res. Atmos.*, *118*, 2473–2493, doi:10.1002/jgrd.50188.
- Tingley, M. P. (2012), A Bayesian ANOVA scheme for calculating climate anomalies, with applications to the instrumental temperature record, *J. Clim.*, *25*(2), 777–791.
- Toms, J., and M. Lesperance (2003), Piecewise regression: A tool for identifying ecological thresholds, *Ecology*, *84*(8), 2034–2041.
- von Buttlar, J., J. Zscheischler, and M. Mahecha (2014), An extended approach for spatiotemporal gapfilling: Dealing with large and systematic gaps in geoscientific datasets, *Nonlinear Processes Geophys.*, *21*, 203–215.
- Von Storch, H., and F. W. Zwiers (2001), *Statistical Analysis in Climate Research*, Cambridge Univ. Press, Cambridge, U. K.
- Xu, X. T., S. L. Piao, X. H. Wang, A. P. Chen, P. Ciais, and R. B. Myneni (2012), Spatio-temporal patterns of the area experiencing negative vegetation growth anomalies in China over the last three decades, *Environ. Res. Lett.*, *7*(3), 035701.
- Zhang, X. B., G. Hegerl, F. W. Zwiers, and J. Kenyon (2005), Avoiding inhomogeneity in percentile-based indices of temperature extremes, *J. Clim.*, *18*(11), 1641–1651.