GEOPHYSICAL RESEARCH LETTERS, VOL. 36, L12701, doi:10.1029/2009GL038868, 2009

Drought return times in the Sahel: A question of attribution

Arthur M. Greene,[1] Alessandra Giannini,[1] and Stephen E. Zebiak[1]

Received 27 April 2009; accepted 15 May 2009; published 20 June 2009.

[1] Recurrence times for extreme drought events in the African Sahel are estimated using a classical peaks-over-threshold model. Results, which are computed for both mean seasonal rainfall and fractional area in drought, suggest that the distribution of dry extremes after about 1970 is statistically distinct from that of preceding years. This finding throws into relief the critical role played by attribution of causes, and the necessity of improving our understanding of the physical processes driving precipitation variability in the Sahel. **Citation:** Greene, A. M., A. Giannini, and S. E. Zebiak (2009), Drought return times in the Sahel: A question of attribution, *Geophys. Res. Lett.*, *36*, L12701, doi:10.1029/2009GL038868.

1. Introduction

[2] The Sahel — the semiarid region of sub-Saharan Africa lying between about 10° and 20° N and extending across the width of the continent — experienced a dramatic and rapid drying during the decade of the 1960s [see, e.g., *Hulme*, 2001; *Held et al.*, 2005]. This climatic shift had profound human consequences [*Mortimore and Adams*, 2001] and engendered a great deal of interest in both the climate and social science communities (see, e.g., *Batterbury and Warren* [2001] and other articles in that issue). In several individual years, among them 1972 and 1984, drought was both particularly intense and widespread, motivating the question that is addressed herein: Given the known precipitation history of the Sahel, what can be said about the expected recurrence time for events of such severity? This issue is clearly of more than academic interest, bearing as it does on matters of adaptation, climate risk management and more specifically, food security, in coming decades.

[3] In this analysis, extreme-value distributions for threshold exceedances are fit to both area-averaged seasonal rainfall and fractional Sahel area in drought; results indicate that the severe dry extremes of the post-1970 period are very unlikely to have occurred given the characteristics of the preceding climate regime, i.e., that pre- and post-1970 precipitation climates of the Sahel are statistically distinct with respect to drought extremes. This finding leads ineluctably to the question of attribution, since the imputation of anthropogenic influence, or conversely, of its absence, would lead to differing expectations for future Sahelian climate, with corresponding implications for adaptation. The analysis described does not address the attribution problem itself, however (see *Giannini et al.* [2003] and *Biasutti and Giannini* [2006] for examples of recent attribution studies).

[4] In Section 2 the two drought metrics utilized and the data from which they are derived are presented, and the shift toward a more arid climate is discussed. The analysis is taken up in Section 3, which presents, in addition to return time estimates, some theoretical and practical considerations. Section 4 presents conclusions, and considers the return time problem in the larger context of attribution.

2. Characterization of Drought

2.1. Severity Versus Extent

[5] Water stress can be gauged using a variety of metrics, two of which are shown in Figure 1: Jul–Sep (JAS) rainfall averaged over the Sahel region, here taken as 20°W–40°E, 12°–20°N, and the area fraction of this region having JAS rainfall in the lowest (locally defined) quintile. JAS is the Sahelian rainy season, that of the west African monsoon, while the fractional-area measure indicates the spatial extent of drought impacts. While these complementary metrics are fairly well-correlated ($r = -0.83$) and can thus be expected to convey similar information, it is nonetheless of interest to see how analytical results for the two might differ. Correlation tests indicate that results are unlikely to be sensitive to the precise seasonal demarcation.

[6] Both data series are derived from the quality-controlled TS2p1 monthly data product of the Climatic Research Unit, University of East Anglia [*Mitchell and Jones*, 2005]. Missing values here have been filled, facilitating computation of the fractional-area metric. A related but unfilled dataset [*Hulme et al.*, 1998] exhibits similar features, to the extent that the relevant computations can be replicated. The area-averaged TS2p1 precipitation is also well-correlated with the Sahel rainfall indices of *Lamb and Peppler* [1992] (common years 1941–2002, $r = 0.86$), *L'Hôte et al.* [2002] (common years 1901–2000, $r = 0.81$) and *Nicholson* [1993]; *Nicholson et al.* [1996] (common years 1901–1994, $r = 0.83$). None of these series is sensitive to changing station populations. The conclusions reached herein are therefore likely to be robust with respect to data set employed.

2.2. Transition to Drier Conditions

[7] Figure 1 shows clearly the mid-century aridification of the Sahel, with both variables reaching extreme levels in the later part of the record that are not seen before about 1970. (Dry extremes are represented by minima in Figure 1a.) For the present analysis the precise timing of this drying or the rapidity with which it occurred are not critical: It is sufficient that 1970 can reasonably be taken as a point that separates, if approximately, the wetter climate of preceding years from the drier conditions prevailing afterward. Such a partitioning has been recognized in prior

[1]International Research Institute for Climate and Society, Earth Institute at Columbia University, Palisades, New York, USA.

0094-8276/09/2009GL038868$05.00

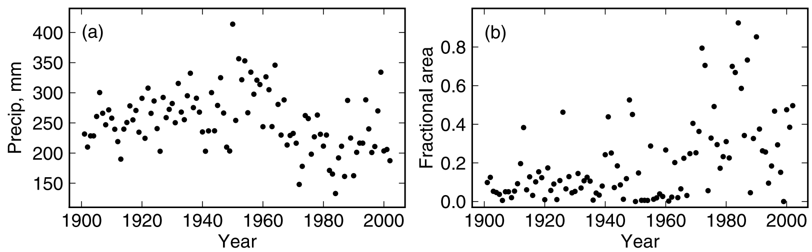

Figure 1. (a) Sahel-averaged JAS rainfall for 1901–2002. (b) Fractional Sahel area with JAS rainfall in the lowest quintile.

studies [*Nicholson*, 1981; *Lamb*, 1982; *Katz and Glantz*, 1986; *Hubert*, 2000], and a sensitivity analysis showed little dependence of analytical results on the precise year used as a dividing point.

[8] Careful inspection of Figure 1 suggests that it may not be strictly correct to treat the climate record as consisting simply of two stationary segments, separated by a more or less abrupt transition. For example, there is a modest upward trend in the early part of the rainfall series, culminating in the relatively high values of the 1950s. Change-points in Sahel precipitation prior to 1970 have also been identified [*Hubert et al.*, 2007]. Such higher-order data characteristics are discussed in Section 3.4, but are not taken into account in the basic treatment described herein. The degree to which rainfall minima follow the mean trend is also open to question.

3. Analysis

3.1. Theoretical Basis

[9] Return times are estimated using a peaks-over-threshold model, in which the generalized Pareto distribution (GPD) is fit to those data whose values exceed a specified level. Given certain distributional assumptions, the GPD is asymptotically optimal for the parametric modeling of threshold exceedances [*Coles*, 2001].

[10] The choice of threshold to be used in estimating model parameters involves a classical balance between bias and variance: The greater the number of exceedances, the more precise will be estimation (lower variance). However, the assumptions on which such models rest are asymptotic, and apply better at higher thresholds (lower bias). Threshold choice was accomplished with the aid of diagnostics that probe this balance, including plots of mean residual life (the average by which extremes exceed a given threshold) and scale and shape parameters of the GPD against threshold, and confirmed with goodness-of-fit metrics to be discussed. Optimally selected thresholds should provide the best return-time estimates for extremes at all levels; it is not necessary that they correspond to those deemed most relevant for impacts or adaptation.

3.2. Application

[11] The GPD was initially fit to the pre-1970 data, using thresholds of 240 mm and 0.1 for precipitation and fractional area, respectively (values apply to the JAS season as a whole). Inference is carried out via maximum likelihood, following the exposition given by *Coles* [2001].

[12] Diagnostics are shown in Figure 2, with precipitation and fractional area models in the top and bottom rows, respectively. These are tests of model fit, in which aspects of the inferred distribution functions are compared with their empirical counterparts. In Figures 2a and 2d the modeled cumulative probability densities are plotted against their empirical estimates, as determined from observed frequencies. In Figures 2b and 2e empirical quantiles are plotted against

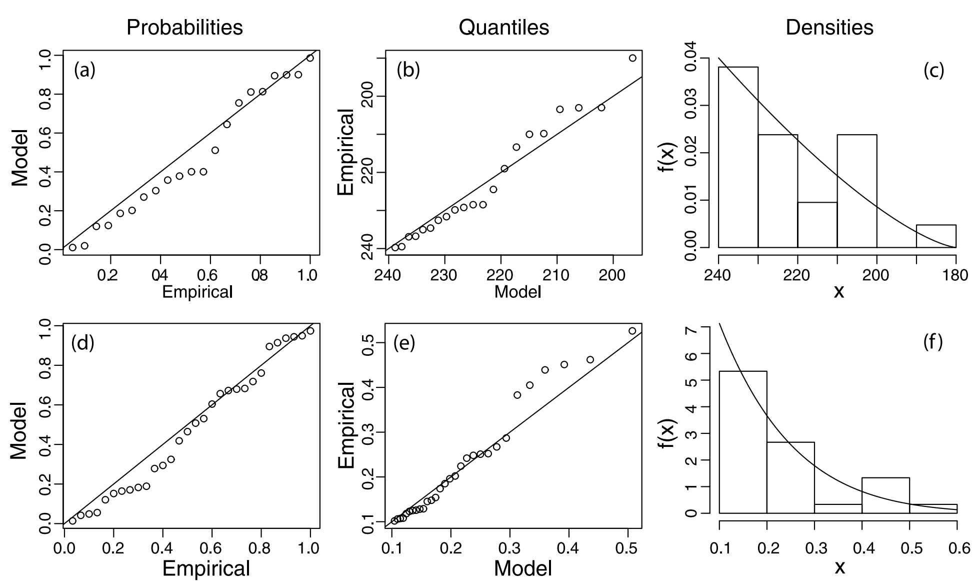

Figure 2. Diagnostic plots for the GPD fit to the pre-1970 precipitation (a–c) and fractional area data (d–f): (a) and (d) probability plots, (b) and (e) quantile-quantile plots, and (c) and (f) data histograms and fitted density functions.

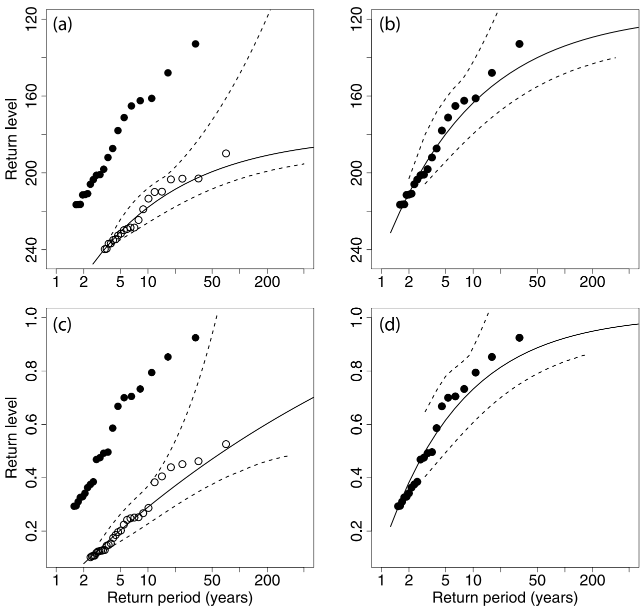

Figure 3. Return-level plots for pre- and post-1970 rainfall (a–b) and fractional drought area (c–d). Circles represent pre-1970 empirical values, solid circles the post-1970 values. Dotted lines show 95% confidence limits.

those of the fitted distributions. In Figures 2c and 2f fitted density functions are plotted on corresponding histograms. These plots indicate that the models fit the data reasonably well, and confirm the reasonableness of chosen thresholds.

[13] Return-level curves and associated 95% confidence intervals (CI) are shown by the solid and dashed lines, respectively, in Figures 3a (rainfall) and 3c (area fraction). CI are computed using the "profile likelihood" method, in which model log-likelihood is maximized conditional on return period. The open circles represent empirical return levels for the data used to fit the models and are seen to fall within the CI, indicating that the inferred models provide reasonably good fits to data. The return-level curves are convex up, implying finite upper bounds for the modeled distributions. This inference is clearly correct, since precipitation cannot fall below zero, while area fraction is bounded above by unity. (Note that ordinate values decrease upward on the rainfall plots.) Finally, asymmetry of the CI reflects the greater uncertainty attaching to higher extremes, for which fewer data are available for estimation.

[14] Empirical return levels for the post-1970 period are also plotted in Figures 3a and 3c, as solid circles, and are seen to fall completely outside the modeled CI. For rainfall, the return-level curve approaches an asymptotic limit of 180 mm. Seven of the post-1970 points lie above this level, indicating return times, based on the pre-1970 model, that are nominally infinite (i.e., the model predicts that such events will never occur). For the fractional area data the disagreement is not quite so stark: the inferred asymptotic limit is actually somewhat greater than unity, which, by construction, none of the empirical values can exceed. The post-1970 distribution nonetheless comports poorly with the pre-1970 fit. These discordances strongly suggest that pre- and post-1970 Sahelian hydroclimates are statistically distinct, and that it would be fundamentally inconsistent to estimate post-1970 return times using models fitted to the earlier data.

3.3. Return Times

[15] Return-level curves estimated from the post-1970 data are shown in Figures 3b and 3d, where empirical values are now seen to agree well with the fitted models. The thresholds utilized here were 220 mm and 0.29, producing 20 and 21 exceedances for the rainfall and fractional area data, respectively. Since there are only 32 a of data, implied exceedance rates are rather high. However, diagnostics (not shown) are quite similar in character to those illustrated in Figure 2, suggesting that these thresholds are reasonable.

[16] Area-averaged JAS precipitation for the three most extreme years, 1984, 1972 and 1987, was 133, 148 and 161 mm, respectively. Taking the median of these values, 148 mm, as representative of the more severe events of the post-1970 epoch, we find a return time of 25 a, with a 95% CI of (9, 103) a. Corresponding values for the fractional-area metric, with a median extreme of 0.85, are similar: recurrence time 31 a, CI (9, 141) a.

3.4. Trailing Analytical Considerations

[17] The analysis presented assumes independence of the exceedance data, e.g., that the occurrence of drought in a particular year does not increase the probability that the following year will also be dry. Inspection of Figure 1 suggests that this assumption could be imperfectly realized, in that there may be some tendency, particularly in the precipitation data, for exceedances to cluster. In part, this can be viewed as a result of ignoring the upward trend in the early part of the record: There are no exceedences between 1949 and 1966, in effect "crowding" those events that do cross the threshold into a 54-year interval (for the pre-1970 data). Statistical tests for proportions, as well as experiments with "partial declustering" (deletion of a single extreme that was part of a cluster), were therefore performed. These suggested that while there may be a slight tendency for an exceedance to be followed by another in the pre-1970 rainfall data (this is not true of the post-1970 rainfall, or of the fractional-area data in general), the conclusions presented have not been compromised thereby.

[18] Regarding the fractional-area metric, the lowest quintile was chosen after some experimentation: Use of the lowest decile was found to produce a record that tends to "saturate" at low values, while the driest years are represented by fractional drought areas <0.8, leaving considerable "headroom" at the upper end of the scale. Use of the quintile metric reduces the saturation, better centers the data and improves goodness-of-fit of the model. The degree to which such a metric might be utilized directly in adaptation work remains to be determined, but its use here is not limiting in this regard.

[19] This analysis does not account for spatial structure in the precipitation record. A treatment that resolved such structure would likely be informative, and could prove helpful in the anticipation, and ultimately management, of the climate risks faced by this vulnerable region.

4. Conclusions

[20] The foregoing analysis indicates that pre- and post-1970 Sahelian hydroclimates differ; as a consequence,

return times have been estimated utilizing the latter period, on the assumption that the more recent past represents a better analog for the Sahelian climate of the future. Of course it is not known with certainty that this is the case, and indeed, there is speculation that a shift toward a less arid state may already be underway [cf. *L'Hôte et al.*, 2002; *Ozer et al.*, 2003].

[21] This brings us face to face with the question of attribution: If we could say, for example, that anthropogenically-imposed changes in the climate system were likely to bring about a drying such as that observed, we might expect the aridity of the post-1970 period to continue or even intensify — just as we expect global temperatures to continue rising as concentrations of atmospheric greenhouse gases increase. Return times would then be relatively short. On the other hand, were such fluctuations known to represent natural (i.e., unforced) variability intrinsic to the climate system, the post-1970 regime might be viewed as one of many similar excursions, not particularly likely to persist. Return-time estimation in the latter case would be more complex, depending as it would on the statistical properties of the fluctuating system. It is also possible that variations in Sahelian climate are attributable to some combination of anthropogenic and natural causes [see, e.g., *Ting et al.*, 2009]. Without being able to disentangle these influences, however, it becomes difficult to anticipate changes in the statistics on which return-time estimates are based. Ultimately then, proper attribution of observed climate variability may be viewed as a necessary (if perhaps not sufficient) condition for the comprehensive modeling and estimation of drought recurrence times in the Sahel.

[22] **Acknowledgments.** We are grateful for helpful comments and suggestions provided by Upmanu Lall, Brad Lyon, Ousmane Ndiaye, Michael Tippett, and Sylwia Trzaska of the IRI, Philippe Naveau of the Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France, and two anonymous reviewers. This paper is funded by a cooperative agreement with the National Oceanic and Atmospheric Administration, NA05OAR431004. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its sub-agencies.

References

Batterbury, S., and A. Warren (2001), The African Sahel 25 years after the great drought: Assessing progress and moving towards new agendas and approaches, *Global Environ. Change*, *11*, 1–8.

Biasutti, M., and A. Giannini (2006), Robust sahel drying in response to late 20th century forcings, *Geophys. Res. Lett.*, *33*, L11706, doi:10.1029/2006GL026067.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, 208 pp., Springer, London.

Giannini, A., R. Saravanan, and P. Chang (2003), Oceanic forcing of sahel rainfall on interannual to interdecadal time scales, *Science*, *302*, 1027–1030, doi:10.1126/science.1089357.

Held, I. M., T. L. Delworth, J. Lu, K. L. Findell, and T. R. Knutson (2005), Simulation of Sahel drought in the 20th and 21st centuries, *Proc. Natl. Acad. Sci. U.S.A.*, *102*, 17,891–17,896.

Hubert, P. (2000), Segmentation, in *Detecting Trend and Other Changes in Hydrological Data*, edited by Z. W. Kundzewicz and A. Robson, *WCDMP-45, WMO/TD 1013*, pp. 110–116, World Meteorol. Org., Geneva.

Hubert, P., J.-C. Bader, and H. Bendjoudi (2007), Un siècle de débits annuels du fleuve Sénégal, *Hydrol. Sci. J.*, *52*, 68–73.

Hulme, M. (2001), Climatic perspectives on Sahelian desiccation: 1973–1998, *Global Environ. Change*, *11*, 19–29, doi:10.1016/S0959-3780(00)00042-X.

Hulme, M., T. Osborn, and T. Johns (1998), Precipitation sensitivity to global warming: Comparison of observations with HadCM2 simulations, *Geophys. Res. Lett.*, *25*, 3379–3382.

Katz, R. W., and M. H. Glantz (1986), Anatomy of a rainfall index, *Mon. Weather Rev.*, *114*, 764–771.

Lamb, P. J. (1982), Persistence of Subsaharan drought, *Nature*, *299*, 46–48.

Lamb, P., and R. Peppler (1992), Further case studies of tropical Atlantic surface atmospheric and oceanic patterns associated with sub-Saharan drought, *J. Clim.*, *5*, 476–488.

L'Hôte, Y., G. Mahé, B. Somé, and J. P. Triboulet (2002), Analysis of a Sahelian annual rainfall index from 1896 to 2000: The drought continues, *Hydrol. Sci. J.*, *47*, 563–572.

Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, *25*, 693–712.

Mortimore, M. J., and W. M. Adams (2001), Farmer adaptation, change and 'crisis' in the Sahel, *Global Environ. Change*, *11*, 49–57.

Nicholson, S. E. (1981), Rainfall and atmospheric circulation during drought periods and wetter years in West Africa, *Mon. Weather Rev.*, *109*, 2191–2208.

Nicholson, S. (1993), An overview of African rainfall fluctuations of the last decade, *J. Clim.*, *6*, 1463–1466.

Nicholson, S., M. Ba, and J. Kim (1996), Rainfall in the Sahel during 1994, *J. Clim.*, *9*, 1673–1676.

Ozer, P., M. Erpicum, G. Demarée, and M. Vandiepenbeeck (2003), The Sahelian drought may have ended during the 1990s, *Hydro. Sci. J.*, *48*, 489–492.

Ting, M., Y. Kushnir, R. Seager, and C. Li (2009), Forced and internal 20th century SST trends in the North Atlantic, *J. Clim.*, *22*, 1469–1481, doi:10.1175/2008JCLI2561.1.

A. Giannini, A. M. Greene, and S. E. Zebiak, International Research Institute for Climate and Society, Lamont Campus, Columbia University, Palisades, NY 10964, USA. (alesall@iri.columbia.edu; amg@iri.columbia.edu; steve@iri.columbia.edu)