The need for spatial uncertainty in climate data: an example from west African Sahelian rainfall (1930-1990)

A. Chappell[1], C. T. Agnew[2]

[1]School of Environment & Life Sciences, University of Salford, Manchester M5 4WT
Telephone: (+44) 0161 295 3635
Fax: (+44) 0161 295 5015
Email: a.chappell@salford.ac.uk

[2]School of Geography, University of Manchester, Oxford Road, Manchester M13 9PL, UK
Telephone: (+44) 0161 275 3654
Fax: (+44) 0161 275 7878
Email: clive.agnew@man.ac.uk

1. Introduction

For several decades the observed decrease of rainfall in the west African Sahel (WAS) between the late 1960s and 1990 has been the focus of climate research (e.g., Folland et al. 1986, Nicholson 1985, Nicholson and Palao 1993, Zeng et al. 1999). Recently Zeng (2003) identified two main hypotheses: anthropogenic factors such as overgrazing and deforestation that increase surface albedo and reduce moisture supply to the atmosphere and lead to less precipitation; the second invokes large-scale atmospheric circulation changes triggered by multidecadal variation in global sea surface temperature. These are important hypotheses and are typically tested by comparing empirical evidence or model predictions against 'observations' of Sahelian rainfall. The outcomes of testing these hypotheses may have considerable influence on the future direction of scientific research and may provide a foundation for understanding the environmental systems for the region. However, the rainfall observations are assumed to represent the underlying rainfall population. They are aggregated from relatively few stations in space, when considering the area of the WAS and variability of its rainfall in space and time. There have been no published assessments of uncertainty surrounding the spatially aggregated annual rainfall in the WAS.

The aim here is to compare the WAS time series produced using the traditional climatological deterministic approach with that produced using stochastic simulations and the latter approach will be used to quantify uncertainty in the estimates. The implications are discussed of these estimates of uncertainty in the WAS rainfall for previous work that compared the results of empirical and model-based predictions to previous estimates of WAS rainfall.

2. Data and Methods

2.1 Rainfall observation records

We used the Global Historical Climatology Network (GHCN v.2) rainfall data (Petersen et al. 1997) and extracted rainfall stations for the years between 1930 to 1990 inclusively, within the west African Sahel (WAS) following the definition of Nicholson (1993; 10-20°N, -20°W to 20°E). We transformed the station locations held using latitude and longitude to provide an equal area projection of Cartesian co-ordinates. In the WAS most rainfall coincides with the summer months (defined here as June, July, August and September). We extracted from the database only stations that had recorded rainfall for all of the summer months in any one year.

2.2 Rainfall anomalies and inverse-distance weighting

The conventional approach in climatology to removing the influence of location on the estimate of the spatial mean is to calculate the inverse-distance weighted rainfall anomalies. This approach requires that the stations used in the analysis have continuous or nearly continuous (allowing for some missing) data. Thus, the extracted data were filtered using the requirements of the calculation of inverse-distance weighted anomalies. A reference period of 1961 to 1990 was set to achieve the maximum number of stations with the minimum amount of missing data (29%). Rainfall anomalies were calculated using the method described by Jones and Hulme (1996) and the inverse-distance area-weighting scheme following Dai et al. (1997).

2.2 Sequential (median indicator approximation) simulation

An alternative approach to estimating the weights for each station in the calculation of the spatial mean was performed using cell-declustering (Deutsch, 1989). It is commonly used in geostatistical analyses (Isaaks and Srivastava, 1989) and is appropriate to climatological data (Chappell and Ekström, 2005) where spatial data are clustered in either high- or low-valued areas.

Despite strong gradients in rainfall in x- and y-directions, there were insufficient data to reliably model the anisotropic spatial variation (Webster and Oliver, 1992). Consequently, omni-directional (isotropic) indicator variograms of rainfall were used to calculate the average variation at several thresholds in all directions. Unfortunately, sample indicator variograms are not well-defined at the margins or extremes of a distribution because they depend on the spatial distribution of only a few pairs of indicator data. Nevertheless, the magnitude and spatial connectivity of extremes in rainfall data are particularly important to spatial simulations and so too are the number of thresholds used in the simulation. The data are equally distributed at the median and the variogram at thresholds other than the median may be inferred using the mosaic model (Journel, 1984). This approximation was known to be at the expense of less flexibility in comparison with direct indicator coding at several thresholds. However, the need for calculation and model-fitting of variograms for multiple thresholds (e.g., 5 thresholds) for each year between 1930 and 1990 (61 years) would require a considerable amount of modelling work (305 indicator variograms). The mosaic model offered a valuable compromise of reliable variograms at the important margin of the distributions at the expense of restricted spatial structural information at each threshold. Furthermore, the median approximation also reduces the number of order relation deviations caused by the integration of maps produced separately for each threshold (Journel, 1984).

The declustered rainfall cumulative distribution function was calculated for all stations' annual summer rainfall between 1930 and 1990 inclusively (61 cumulative distributions). The rainfall for K=5 quantiles (10, 25, 50, 75 and 90%) were established each year (305 rainfall values) and were also used in the simulations. These percentiles were used to transform the rainfall values into indicator variables.

The median sample indicator isotropic variogram of annual rainfall data was calculated to approximately one third (1500 km) of the maximum separation distance between rainfall stations. These variograms were fitted, using weighted least squares, with several models authorised for kriging (linear, spherical, exponential, power, and circular; Webster and Oliver, 2001). The models that fitted best, in the least-squares sense, (Spherical and Exponential) were selected using the square root of the mean squared difference between the model and the observations (RMSE). Sequential indicator simulation was used here to generate 300 realisations of rainfall and honour the values of the rainfall stations each year, reproduce (approximately) the declustered sample histogram and the covariance models for the five thresholds using the median indicator approximation (Deutsch and Journel, 1998).

3. Results

The means of the 300 rainfall realisations (maps) provided distributions of values for each year. The percentiles of the mean realisations distribution are used to characterise the spatial uncertainty of rainfall over time (Figure 1).

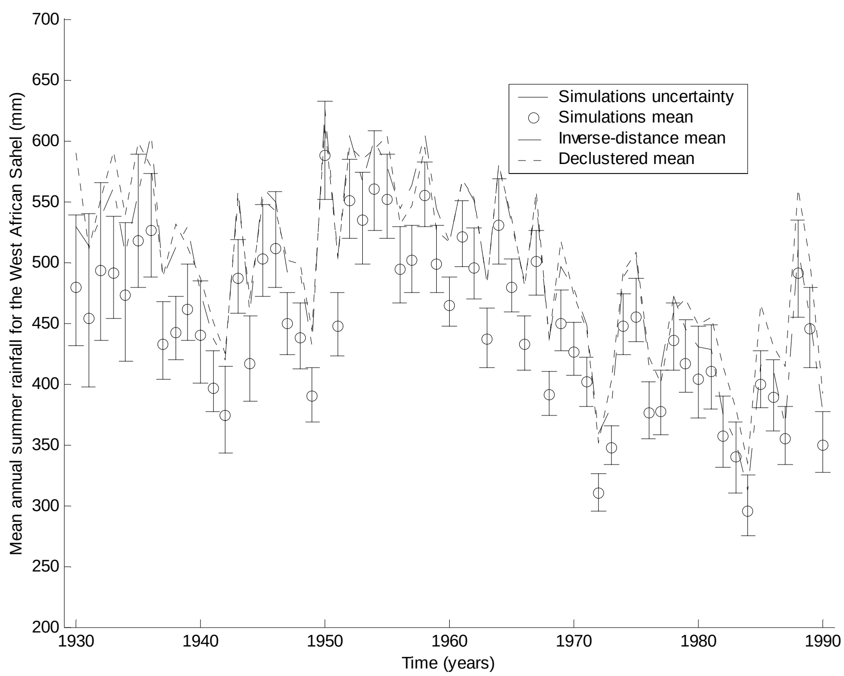

Figure 13. Annual variation in summer rainfall for selected stations using the inverse-distance weighted anomalies, the declustered rainfall and uncertainty error bars represented by the 5th and 95th percentiles of the 300 averages of the simulation realizations.

The median of the mean distribution is shown in the centre of the 5th and 95th percentile error bars. There is a strongly cyclical pattern in annual rainfall. Uncertainty during the period 1930 to 1935 is noticeably larger than other years. There appears to be no evidence of decreasing rainfall. The well known drought years of 1972 and 1984 show much smaller values of rainfall and reduced uncertainty than other rainfall years. The area-weighted time series mostly exceed the 95th percentile of the simulation distribution. Notable exceptions occur between 1978 and 1989 when inverse-distance weighted anomalies are well within the range of uncertainty. This comparison demonstrates that these area-weighting techniques provide time series for the WAS which are largely an extreme realisation of the simulated ensemble distribution.

4. Discussion

The annual rainfall distributions show uncertainty (average inter-quartile range = 26 mm; average standard deviation = 19 mm) in the mean annual summer rainfall for the WAS. That uncertainty is larger between 1930 and 1935 than later in the period and appears to be strongly cyclical. The uncertainty is notably smaller during the notorious drought years of 1972 and 1984 than other years during the period. The uncertainty in the WAS time series is so large that it casts doubt on the existence of desiccation. That pattern of desiccation is also less apparent when those drought years are removed from the time series. The uncertainty surrounding the rainfall estimates complicates its comparison with empirical evidence of related variables e.g., soil dust (Prospero and Lamb, 2003) or model-based predictions of rainfall. Furthermore, the long-term (1930-1990) mean annual rainfall of the area-weighted estimates (490 mm and 501 mm for the inverse-distance and declustered area-weighting techniques, respectively) are approximately 8% and 10%, respectively larger than the long-term mean of the median annual simulation realisations (449 mm).

5. Conclusion

The analysis provided the first estimates of the temporal variation in spatial uncertainty of west African Sahel (WAS) rainfall. The established time series of WAS rainfall mostly exceeded the 95th percentiles of the annual simulation realisations and the area-weighted long-term mean (490 mm) was about 8% larger than that of the simulation realisations (449 mm). The variation in estimates of mean annual rainfall and the 5th and 95th percentiles of the uncertainty distribution was considerable. The results showed that uncertainty in the mean annual summer rainfall was so large that it confused the pattern of desiccation in the region that was previously commonly accepted. The uncertainty in desiccation rendered inconclusive previous comparisons between WAS rainfall and empirical evidence from related variables and model-based predictions. The results have important implications for future climatological work particularly as all published regional, hemispheric, and global time series of temperature and precipitation are based on area-weighted averages without expressions of spatial uncertainty.

6. Acknowledgements

The authors are grateful to NCAR for hosting the provision of the GHCN v.2 data.

7. References

Chappell A, and Ekström M, 2005, The importance of de-clustering and uncertainty in climate data: a case study of west African Sahel rainfall. In: Leuangthong, O. and Deutsch, C. (eds), *Quantitative Geology and Geostatistics*, Vol. 14, XXVIII, 1167 p. Springer (ISBN: 1-4020-3515-2)

Dai AI, Fung Y, and Del Genio AD, 1997, Surface observed global land precipitation variations during 1900-88. *Journal of Climate*, 10, 2943-2962.

Deutsch CV, 1989, DECLUS: A Fortran 77 program for determining optimum spatial declustering weights. *Computers and Geosciences*, 15(3), 325-332.

Deutsch CV, and Journel AG, 1998, *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press, 340p.

Folland CK, Palmer TN, and Parker DE, 1986, Sahel rainfall and worldwide sea temperatures, 1901-85. *Nature*, 320, 602-606.

Isaaks EH, and Srivastava RM, 1989, *An introduction to applied geostatistics*. OUP, Oxford.

Jones PD, and Hulme M, 1996, Calculating regional climatic time series for temperature and precipitation: methods and illustrations. *International Journal of Climatology*, 16, 361-377.

Journel, AG, 1984, The place of non-parametric geostatistics. In: G. Verly, M. David, A.G. Journel and A. Marechal, (eds), *Geostatistics for Natural Resources Characterization*, V1 pages 307-355. Reidel, Dordrecht.

Nicholson SE, 1985, Sub-Saharan rainfall 1981-1984. *J. Climate Appl. Meteor.*, 24, 1388-1391.

Nicholson SE, 1993, An overview of African rainfall fluctuations of the last decade. *Journal of Climate*, 6, 1463-1466.

Nicholson SE, and Palao IM, 1993, A re-evaluation of rainfall variability in the Sahel. *International Journal of Climatology*, 13, 371-389.

Peterson TC, and Vose RS, 1997, An overview of the Global Historical Climatology Network temperature data base. *Bulletin of the American Meteorological Society*, 78, 2837-2849.

Prospero JM, and Lamb PJ, 2003, African droughts and dust transport to the Caribbean: Climate change implications. *Science*, 302, 1024-1027.

Webster R, and Oliver MA, 1992, Sample adequately to estimate variograms of soil properties. *J. Soil Sci*., 43: 177-192.

Webster R, and Oliver MA, 2001, *Geostatistics for environmental scientists*. J. Wiley & Sons, Chichester.

Zeng N, 2003, Drought in the Sahel. *Science*, 302, 999-1000.

Zeng N, Neelin, JD, Lau K-M, and Tucker CJ, 1999, Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. *Science*, 286, 1537-1540