New Gridded Daily Rainfall Data For The Sahel

Andrew Matthews
Tom Holt
Jean Palutikof

Introduction

Rainfall in the Sahel, West Africa, is characterised by high interseasonal variability and has suffered frequent severe droughts since the late 1960s, suggesting a change in the rainfall regime (Hulme, 2001). In order to create a spatiially homogenous daily rainfall dataset suitable for unbiased statistical analysis, a gridded dataset has been created using stations in Burkina Faso, Mali, Niger and Senegal. A smoothing thin plate spline has been used to create a daily gridded rainfall series at a 1° resolution.

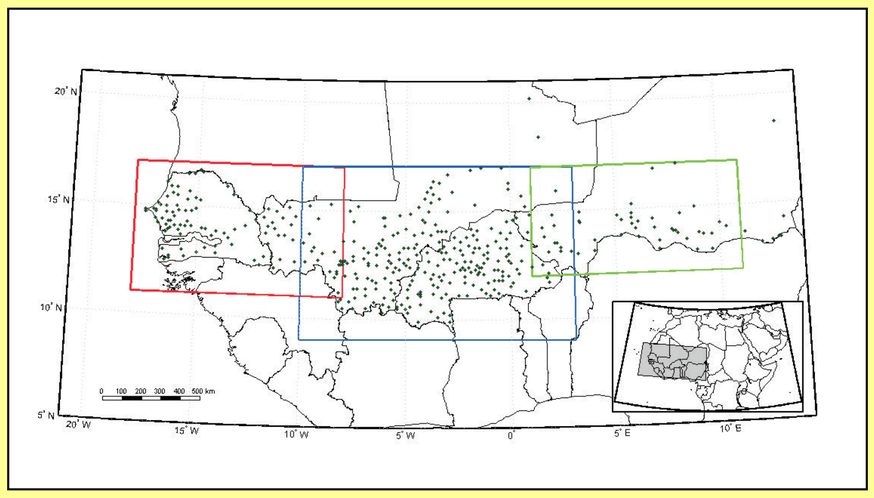


Figure 1. The 524 stations used in the creation of the gridded data sets. The three regions are illustrated by the coloured boxes.

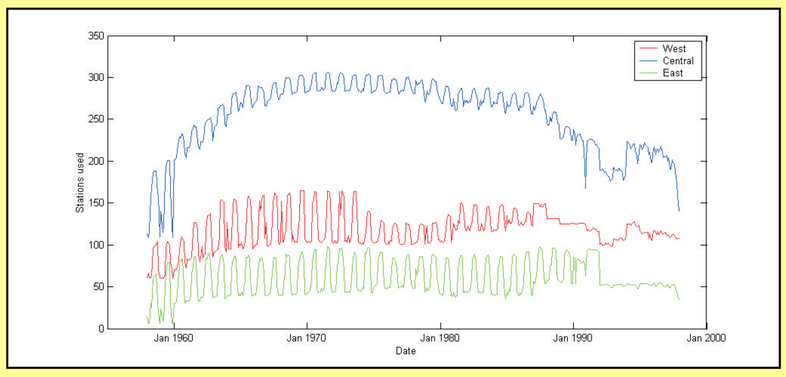
The dataset

Daily rainfall data was available for 524 locations displayed in figure 1 for the period 1958-1997. The station data set was obtained primarily from Dr P. Lamb of the University of Oklahoma. All have records containing a minimum of 50% of values present for rainfall in the wet season (June-September) over the given period. The spatial distribution of the stations allowed for three separate regions to be gridded:

- A western region (11-17°N, 18-8°W) focused upon Senegal
- A central region (9-17°N, 10°W-3°E) focused upon Burkina Faso and Mali
- An eastern region (12-17°N, 1-11°E) focused upon Niger

Each region was gridded at a 1° resolution. The selection of the size of the databox has no effect on the fitted surface, and so were designed to be large enough to study edge effects and to allow validation between regions. The time series of stations contributing to the gridding for each region is show in figure 2.

Figure 2. The time series of the number of stations used to create each gridded series over the period 1958-1997.



Characteristics of Sahelian Rainfall

Figures 3 and 4 represent some of the characteristics of the daily gridded rainfall data, showing the median yearly rainfall and the typical evolution of the wet season throughout the year.

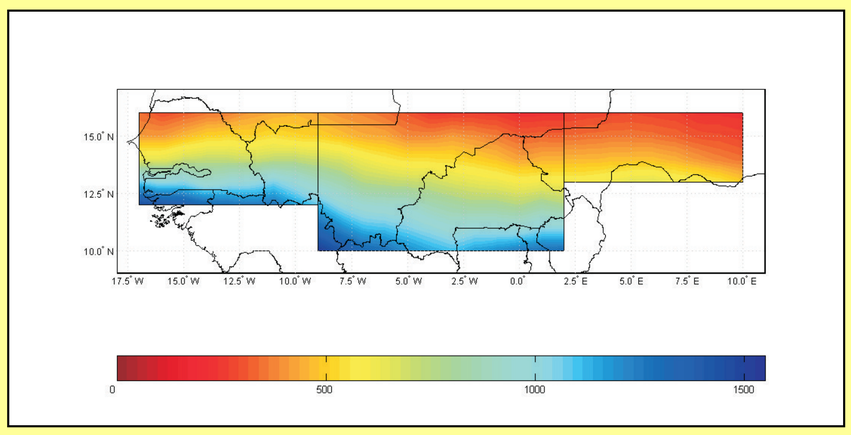
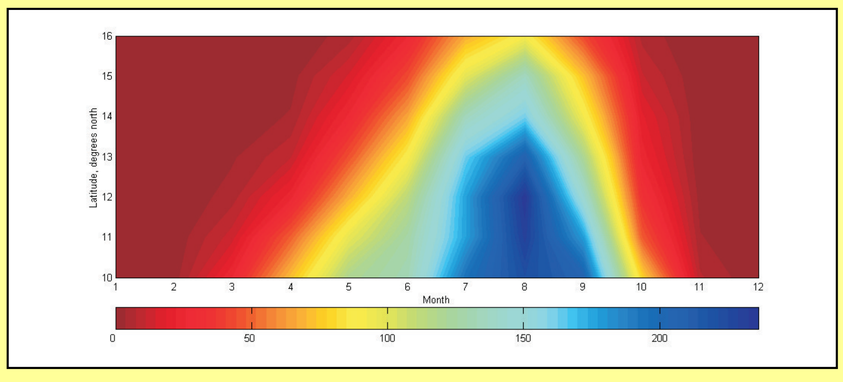


Figure 3.

Median yearly rainfall (in mm) of the gridded series for 1958-1997.

Figure 4.

Plot of median monthly total rainfall (in mm) for each month for a section across 3° West.



The interpolation method

The smoothing thin-plate spline (Wahba, 1990) has often been used in the gridding of climatological datasets including daily rainfall (Hutchinson, 1998, New et al., 1999) and operates by fitting a surface which minimises the following function:

$$\min_{s} \; v \int_{\mathbf{R}^2} \left(\frac{\partial^2 s}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 s}{\partial x \, \partial y} \right)^2 + \left(\frac{\partial^2 s}{\partial y^2} \right)^2 dx dy + \sum_{n=1}^{N} \left(f_n - s(x_n) \right)^2$$

Where s represents the surface to be fitted to the set of points $\{x_n : n = 1, \ldots N\}$, and where f_n is the actual value of the data at point n.

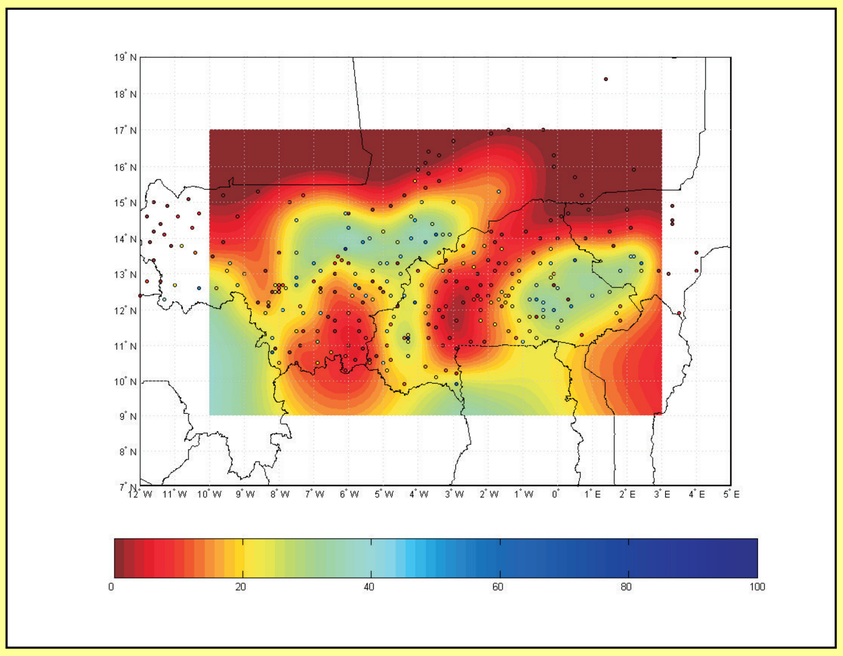
Thus the process aims to minimise two properties for the surface:

- Surface roughness (the first term of the equation)
- Least-squares error at data points (the second term of the equation)

The parameter v regulates the trade-off between smoothness and goodness of fit, and an optimal value can be calculated (Billings et al., 2002). This is computationally expensive, but feasible for datasets of this size.

The resulting dataset represents the areal average of rainfall in a given gridbox. Surfaces were fitted daily for each region using all stations within 2° of the region boundaries. Figure 5 shows a typical fitted surface.

Figure 5. An example of the gridded procedure representing rainfall in the central region on the 17th August 1970. The circles represent the rainfall at the stations used in the procedure, and the contours represent the fitted surface.



Validation of the gridded data

A successful gridding process will represent the variance present in the original data. Therefore validation attempts focused on ensuring the gridded data exhibited a similar distribution to the original data, and examining the residual error of the fitted surface.

Figure 6 shows two quantile-quantile plots of the gridded data for the gridbox 15°N, 17°W against the station data withheld from the gridding process. The left plot represents the distribution of the gridded dataset compared to one station, the right plot the gridded dataset compared to the mean of 10 stations.

The proximity of the plotted values to the line of equality in the right-hand plot suggests the gridded data has a similar distribution to the mean of the ten stations. The left-hand plot indicates the gridded process has a reduced variability compared to station data. A similar distribution is suggested, as the plotted points are roughly in a straight line, at least for non-extreme values.

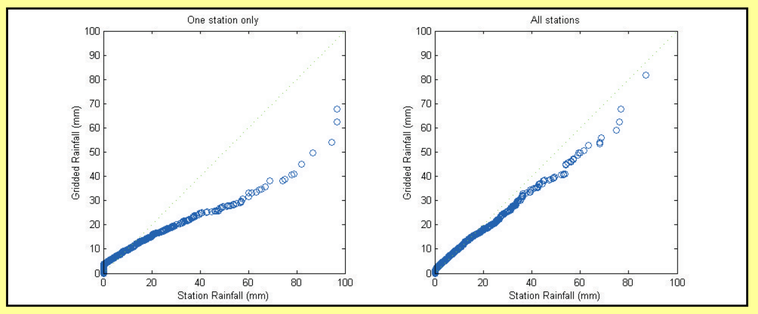


Figure 6. Quantile-quantile plots of the gridded series against original rainfall series for the gridbox 15°N, 17°W. The distribution of the gridded data is plotted against the distribution of rainfall data withheld from the gridding process.

The goodness of fit of the gridded data can be investigated through the errors of the fitted surface at the station locations used in the gridding process. Table 1 gives details of these statistics for one station (Ouagadougou) and for all fifteen stations in the box 12°N, 2°W in the central region.

	One Station (Ouagadougou)	All Stations (15)
Daily Mean Absolute Error, mm	5.97	1.26
Yearly Mean Absolute Error, mm	82.21	25.44
Daily Mean Relative Error	2.44	1.94
Yearly Mean Relative Error	0.10	0.03

Table 1. Daily and Yearly error statistics for the gridbox 12°N, 2°W (rain days only)

The gridded data will be used to investigate the relationship between rainfall and the underpinning atmospheric processes.

References

Billings, S. D., Newsam, G. N. and Beatson, R. K. (2002), Smooth fitting of geophysical data using continuous global surfaces., *Geophysics*, 67, 1823-1834.
Hulme, M. (2001), Climatic perspectives on Sahelian desiccation: 1973-1998., *Global Environmental Change*, 11, 19-29.
Hutchinson, M. F. (1998), Interpolation of Rainfall Data with Thin Plate Smoothing Splines: I Two Dimensional Smoothing of Data with Short Range Correlation., *Journal of Geographic Information and Decision Analysis*, 2, 139-151.
New, M., Hulme, M. and Jones, P. (1999), Representing Twentieth-Century Space-Time Climate Variability. Part I: Development of a 1961-90 Mean Monthly Terrestrial Climatology., *Journal of Climate*, 12, 829-856, 1999.
Wahba (1990) *Spline models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics.

Acknowledgements

Many thanks to Peter Lamb, Issa Lele Mouhamadou and Daouda Diarra for their assistance in obtaining the station data

Andrew Matthews
andrew.matthews@uea.ac.uk
+44 (0)1603 593161

www.cru.uea.ac.uk

Climatic Research Unit
University of East Anglia
Norwich, United Kingdom
NR4 7TJ