Regional Hydrological Impacts of Climatic Change—Hydroclimatic Variability
(Proceedings of symposium S6 held during the Seventh IAHS Scientific Assembly at Foz do Iguaçu, Brazil, April 2005). IAHS Publ. 296, 2005.

Spatial disaggregation of Sahelian rain fields: a new application of the Gibbs sampler

HUBERT ONIBON[1,2], THIERRY LEBEL[2], ABEL AFOUDA[1], GILLES GUILLOT[3]

1 *CIPMA-UAC 07 BP 72 Cotonou, Bénin*
hubert.onibon@afriturible.com

2 *LTHE - 1025 BP 53 Domaine Universitaire, F-38041 Grenoble Cedex 9 France*

3 *INRA, Unité de Biométrie Domaine Saint Paul, F-84914 Avignon Cedex 9 France*

Abstract The present study is in the framework of the pending research actions on the hydrological impacts of the drought that struck West Africa for almost 30 years, starting at the end of the 1960s. It centred on solving the spatial-scale problems existing between the rain fields, as they are presently available, and those necessary for the hydrological models. The proposed method is based on the coupling of the Gibbs sampler and the acceptance–rejection algorithm for the simulation of meta-Gaussian rain fields able to characterize an integral value representative for a given area by respecting the statistical properties of point rain depths. It is applied to the conditional simulation of rain fields observed at the Niamey Degree Square raingauge during the EPSAT-Niger campaign. The model performance is evaluated by comparing the simulated statistical rain field properties with those observed for high-resolution Sahelian area. According to the results of this validation, the proposed model provides an elegant and convenient answer to the conditional simulation of Sahelian rain fields.

Key words Gaussian model; Gibbs sampler; model; rainfall; observation; scale; simulation

INTRODUCTION

The drought that struck West Africa for almost 30 years, starting at the end of the 1960s had dramatic consequences on the water resources, the agriculture and the economic development of the sub-region (e.g. Lebel *et al.*, 2000; Le Barbé *et al.*, 2002). Observations of present rainfall regimes and plausible scenarios for future rainfall regimes derived from climatic models have to be considered jointly to characterize the possible impact of climate variability on water resources (Onibon *et al.*, 2002). The disadvantage is that the spatial resolution of the climatic models is too crude to satisfy the needs of hydrology. Stochastic disaggregation models respecting intrinsic statistical properties derived from observations constitute one of the solutions to this problem. Using the Turning Band Method (TBM) a spatial disaggregation model was previously developed, allowing for the non-conditional simulation of rainfall fields for climatic regions where rainfall is mostly associated with the occurrence of mesoscale convective systems (Guillot, 1998). The weakness of this approach is the requirement of a large set of simulation scenarios in order for its use in a conditional context. It is proposed here to revisit this model by coupling the Gibbs sampler (Geman & Geman, 1984) and the acceptance–rejection simulation method

(Von Neuman, 1951). As a preliminary step, the iteration kernel of the proposed approach is briefly described. In order to explore the efficiency of the model, it is tested on a sample of Sahelian rain fields observed during the EPSAT-Niger experiment. The proposed validation of the model permits comparison of observed and simulated rain fields statistics obtained for various spatial scales.

PRESENTATION OF THE MODEL

The Gaussian transformed function

Let X denotes a stationary Gaussian random function and Φ a strictly monotonic and non-decreasing function referred hereafter as anamorphosis. In this study, the original rain field data of interest are assumed to be characterized by a stationary random function $Y \overset{L}{\equiv} \Phi(X)$. Under these conditions, the random function Y is said to be Gaussian transformed and its distribution is characterized by the anamorphosis function Φ, the expectation μ_X and the covariance C_X of the Gaussian function *X*. Assuming that G is the Gaussian probability function of X and F the probability function of *Y*, then $G(X) = F(Y) = F[\Phi(X)]$. A detailed mathematical description of the Gaussian transformed model may be found in Freulon (1992) and Guillot (1999).

Conditional simulation of a Gaussian transformed field

Let *z* denote the spatial average of the event rainfall, $y = (y_i)_{i \in N}$ the random vector representing the values of the process at the nodes of a regular grid and $x_{-i} = (x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n)$ the vector comprising all x except x_i. The purpose here is to implement the Gaussian transformed model, to generate the vector y conditioned by the spatial average value z over a domain A. When estimating z from the set of observations $y = (y_i)_{i \in N}$, the kriging estimate is:

$$\hat{z} = \sum_{i=1}^{n} \lambda_i y_i = \sum_{i=1}^{n} \lambda_i \Phi(x_i) \tag{1}$$

where λ_i are the weighting factors computed by solving the classical kriging matrix (Onibon, 2001). If ε denotes a Gaussian residual of mean zero and standard deviation K, this is equivalent to:

$$z = \hat{z} + \varepsilon = \sum_{i=1}^{n} \lambda_i \Phi(x_i) + \varepsilon \tag{2}$$

In terms of probability density functions (pdf), equation (2) can be expressed as:

$$f(z|x) = f(\varepsilon|x) = g_\varepsilon\left(z - \sum_{i=1}^{n} \lambda_i \Phi(x_i) \right) \tag{3}$$

The conditional density function $f(x|z)$ is unknown and an iterative approach is sought involving the successive computation of all the components x_i of the vector x. By doing so, we should in principle be able to determine the conditional density function $f(x_i|x_{-i},z)$. In general this is not possible. But, based on Bayesian considerations, Onibon (2001) showed that the probability density function of x_i conditional on (x_{-i},z) could be expressed as a function of the product of two known distributions, except for an unknown coefficient, equation (4):

$$f(x_i|x_{-i},z) \quad \propto \quad h_{x_{-i},z}(x_i) = \exp\left(-\frac{1}{2}\left(\frac{1}{\kappa}\left(z-\sum_{i=1}^{n}\lambda_i\Phi(x_i)\right)\right)^2\right)g(x_i|x_{-i}) \tag{4}$$

In order to generate each component of x by respecting equation (4), the random Gaussian noise ε will be generated with the acceptation-rejection algorithm introduced in statistic by Von Neuman (1951) and used to simulate random variables characterized by a density function totally defined up to a multiplicative constant. Finally, the disaggregation algorithm proposed here is based upon the coupling of the Gibbs sampler the general objective of which is to sample an *n*-dimensional random vector $x=(x_1,x_2,.....,x_n)\in IR^n$ from a multivariate *n*-dimensional pdf $f(x)$ when no practical algorithm is available for doing so directly, and the acceptance–rejection algorithm which are well described in Freulon (1992) and Onibon (2001).

The successive steps of the proposed algorithm are as follows:

Build an initial Gaussian vector x^0, at the j^{th} iteration, and for the i^{th} component:

(1) calculate the simple kriging estimate x_i^{SK} and its standard deviation σ_i^{SK},

(2) generate a standard zero-mean normal random variable u,

(3) replace the component x_i^{j-1} by $x_i^j = x_i^{SK} + \sigma_i^{SK}u$,

(4) calculate $\dfrac{h_{x_{-i},z}(x_i)}{g(x_i|x_{-i})} = \exp\left(-\dfrac{1}{2}\left(\dfrac{1}{\kappa}\left(z-\sum_{i=1}^{n}\lambda_i\Phi(x_i)\right)\right)^2\right)$,

(5) generate a uniform random variable w,

(6) if $w \le \dfrac{h_{x_{-i},z}(x_i)}{g(x_i|x_{-i})}$, then accept x_i^j; else, go to (2).

Stop the iterations.

APPLICATION FOR THE SIMULATION OF SAHELIAN RAINFIELDS

Unconditional simulation of the Niamey Degree Square rain fields

Through the following, we are trying to determine whether the statistical properties of Sahelian event rain fields specified in the model were well produced by the algorithm; what would confirm that the algorithmic and the numeric aspects are well mastered as shown in Table 1, this verification made on a grid of 400 points (very often resampled

Table 1 Comparison of the observed and simulated statistical parameters.

	Mean (mm)	Variance (mm^2)	F_0 (%)
Observations	10.9	200.0	25.4
Gibbs	10.6	201.3	23.0
TBM	10.8	190.0	26.0

to 30 points, and this for homogeneity with the EPSAT-Niger network) proved to be satisfactory for the mean and the standard deviation of point rain depths. However, an underestimation of the intermittence (22% instead of 26%) was noticed. It should be noted that at this point our algorithm is less efficient than the Turning Band method (TBM) used by Guillot (1998), whereas theoretically these two methods should equally reproduce the specified properties. In terms of distribution of (1) the points rain depths, (2) the mean and, (3) the standard deviation of the events rainfalls, a good accord between the two observations and simulations could be noticed (Figs 1, 2(a) and 2(b)). However, some differences on the level of the high point rain depth values can be noted. Concerning the spatial structure of the event rain fields, a good resemblance (Fig. 3) can be noted between the mean variogram of the simulated events and the theoretical variogram deducted from the model of covariance imposed in the model following the results obtain by Guillot (1998).

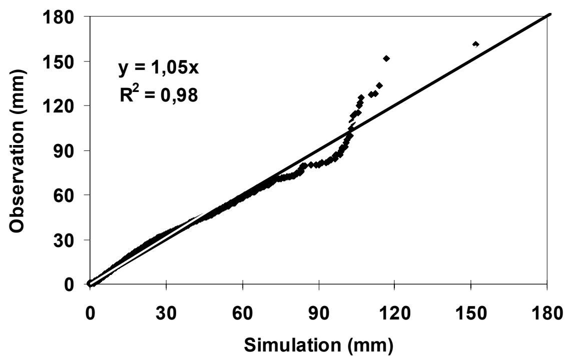

Fig. 1 Quantile–quantile plot of the observed and simulated point rainfalls.

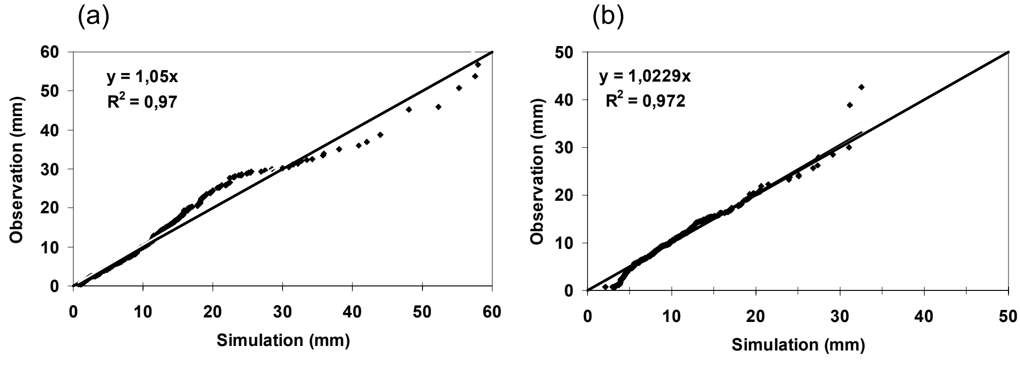

Fig. 2 Quantile–quantile plot of the observed and simulated (a) spatial average rainfall and (b) spatial standard deviation.

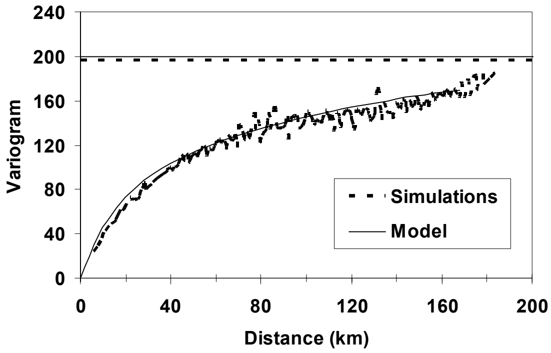

Fig. 3 Comparison of the theoretical variogram with the mean variogram of the simulated events.

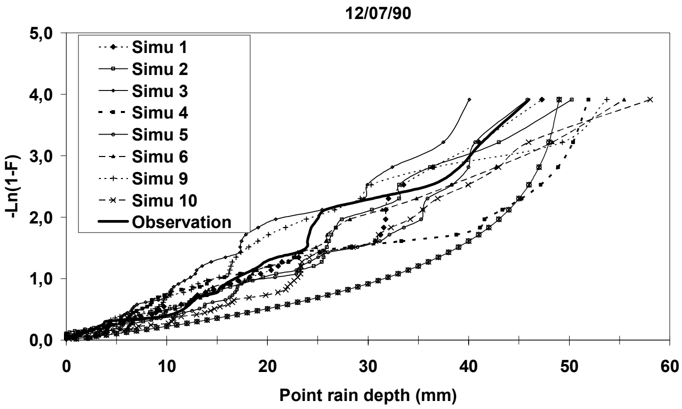

Fig. 4 Cumulative distribution functions: internal variability of the model.

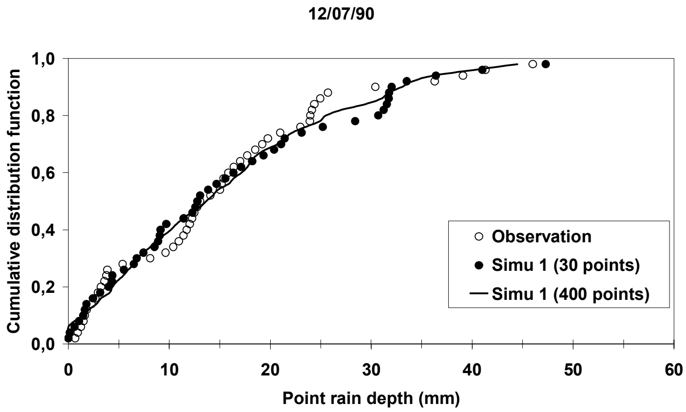

Fig. 5 Cumulative distribution functions: observations compared with simulations.

Conditional simulation of the Niamey Degree Square rain fields

Comparison of the cumulative distribution functions As a preliminary step, we are studying the internal variability of the model in terms of the cumulative distribution of the point rain depths. In conformity with the EPSAT-Niger network geometry, the simulated rain fields are re-sampled to 30 points. In Fig. 4, we compare the cumulative distribution curves obtained from various simulations conditioned by the spatial average of the event of 12 July 1990. It is noticed that the internal variability of the model is strong enough in terms of cumulative frequencies of point rain depths. This, far from being a problem, constitutes a very positive result. It emphasizes an intrinsic property of the numeric method used here, that is, its aptitude to generate, for a given spatial average, different types of distributions of point rain depths. Still using the same logic, we are comparing the empirical frequencies of the point rain depths observed to those of some simulations. We note that despite the internal variability of the model, there is a good agreement between the distributions obtained from some simulations and the observed distributions (Fig. 5).

Spatial structure of the simulated rain fields The basic principle of the spatial disaggregation model proposed in this work is based especially on the respect of the well-known statistical properties of the event rain fields (Onibon, 2001). In the preceding section we studied, in terms of distribution, the realism of rain fields simulated conditionally to a spatial average value. To appreciate the spatial structure of the simulated rain fields, we compared for each event the patterns of the simulated rain fields to those observed. The example presented here is that of the event of 12 July 1990 observed in the study region. As Fig. 6 shows, the isohyets of the point rain depths observed have a tendency to be centred on the network with the high values being localized in the centre. This is equally detected by one of the simulated rain fields (Fig. 7). Since the spatial organization of the event rainfall is characterized by a high degree of randomness, two events characterized by the same spatial average and variance could have very different patterns. When comparing the images shown in Fig. 7, a very strong variability of the spatial distribution of point rain depth values is observed. This result shows the skill of the model in reproducing the large range of patterns associated with events with equal magnitude and dispersion.

CONCLUSIONS

The aim of the study presented in this manuscript, is to develop an approach which allows for spatial disaggregation of low resolution rain fields generated by climatological models, in view of possible use for hydrological models. The new method proposed allows the implementation of the Gaussian anamorphosis model based on the simultaneous use of the Gibbs' sampler and the acceptance–rejection algorithm. The model performance is evaluated by comparing the simulated statistical rain field properties with those observed for the high resolution Sahelian area. According to the result of the application of the model for the unconditional simulations of rain fields, one can conclude that the well-known statistical properties of the observed rain field are well reproduced.

Furthermore, the real value of the proposed approach lies in its ability of simulation conditional on a known spatial value. Even though a complete validation of the model in conditional mode is not possible (it would require several dozen observed realizations with the same spatial average), it is possible to assess its realism by conditionally simulating several rain fields and comparing them to the observed

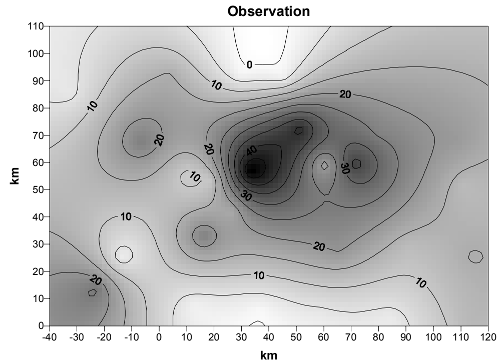

Fig. 6 Spatial distribution of the observed cumulated values (event of 12 July 1990).

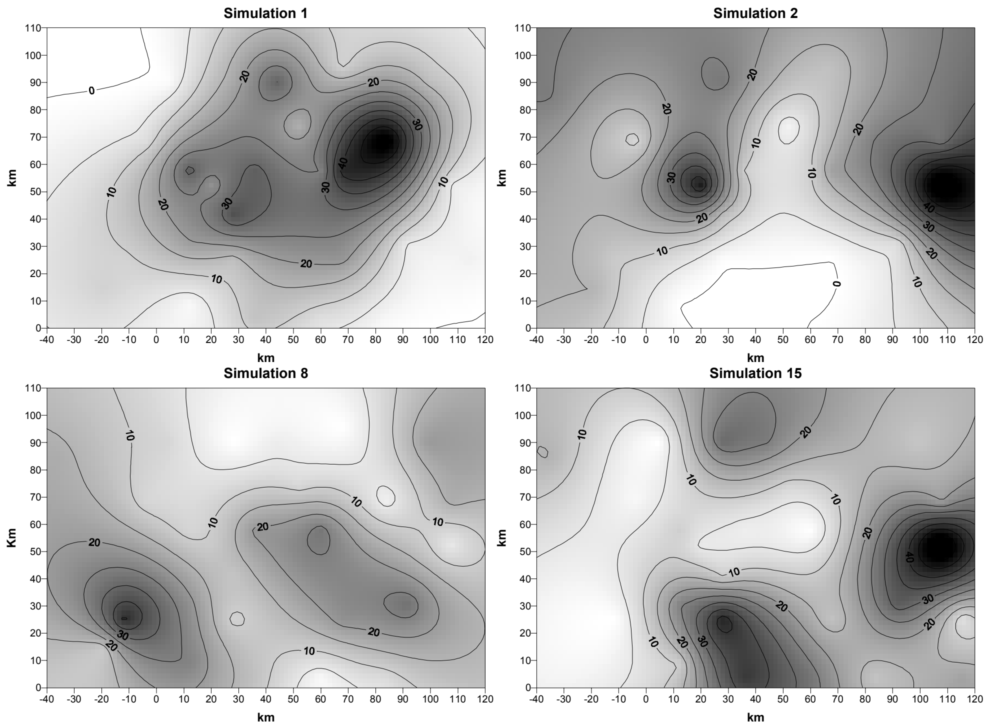

Fig. 7 Examples of the conditional simulations obtained for the event of 12 July 1990.

rain field and to other rain fields of similar magnitude. This comparison was carried out for an event observed with the dense EPSAT-Niger network. For a given class of events, the conditional rain fields have a distribution of point values similar to the distribution of observed point values. At the same time the model is producing a wide range of spatial patterns corresponding to a single area average. This characteristic should allow study of a large spectrum of responses of the hydrological systems to the spatial pattern of rainfall in this region.

REFERENCES

Freulon, X. (1992) Conditionnement du modèle Gaussien par des contraintes ou des randomisées. Thèse de Doctorat, Ecole Nationale Supérieure des Mines, Paris, France.

Galli, A. & Gao, H. (2001) Rate of convergence of the Gibbs sampler in the Gaussian case. *Math. Geol.* **33**(6), 653–677.

Geman, S. & Geman, D. (1984) Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence* **6**, 721–741.

Guillot G. (1998) Modélisation statistique des champs de pluie sahéliens. Application à leur désagrégation spatiale et temporelle. Thèse de Doctorat. Université Joseph Fourier, Grenoble, France.

Guillot, G. (1999) Approximation of Sahelian rainfall fields with meta-Gaussian random functions. Part 1: Model definition and methodology. *Stoch. Environ. Res. Risk Assess*. **13**, 100–112.

Le Barbé, L., Lebel, T. & Tapsoba, D. (2002) Rainfall variability in West Africa during the years 1950–1990. *J. Climate* **15**(2), 187–202.

Lebel, Th., Delclaux, F., Le Barbé, L. & Polcher, J. (2000) From GCM scales to hydrological scales: rainfall variability in West Africa. *Stoch. Environ. Res. Risk Assess*. **14**, 275–295.

Onibon, H. (2001) Simulation conditionnée des champs de pluie au Sahel: application de l'algorithme de Gibbs. Thèse de Doctorat, Institut National Polytechnique de Grenoble, Grenoble, France.

Onibon, H., Lebel, T. & Afouda, A. (2002) Space-time rainfall variability in West Africa derived from observations and GCMs. In: *FRIEND 2002–Regional Hydrology: Bridging the Gap Between Research and Practice* (ed. by H. A. J. van Lanen & S. Demuth), 483–490. IAHS Publ. 274. IAHS Press, Wallingford, UK.

Von Neuman, J. (1951) Various techniques used in connection with random digits. *National Bureau of Standards Appl. Math. Ser.* **2**, 36–38.