Generation of high resolution rainfields in West Africa: evaluation of dynamical interpolation methods
T. Vischel, G. Quantin, T. Lebel, J. Viarre, M. Gosset, F. Cazenave, G. Panthou

To cite this version:

HAL Id: insu-00649355
https://hal-insu.archives-ouvertes.fr/insu-00649355
Submitted on 3 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Generation of High-Resolution Rain Fields in West Africa: Evaluation of Dynamic Interpolation Methods

T. VISCHEL, G. QUANTIN, AND T. LEBEL
LTHE, Université Grenoble 1, IRD, Grenoble, France

J. VIARRE
CNES, LMTG, Toulouse, France

M. GOSSET
GET, IRD/UPS/CNRS, Toulouse, France

F. CAZENAVE AND G. PANTHOU
LTHE, Université Grenoble 1, IRD, Grenoble, France

(Manuscript received 28 October 2010, in final form 6 April 2011)

ABSTRACT

High-resolution rain fields are a prerequisite to many hydrometeorological studies. For some applications, the required resolution may be as fine as 1 km in space and 5 min in time. At these scales, rainfall is strongly intermittent, variable in space, and correlated in time because of the propagation of the rainy systems. This paper compares two interpolation approaches to generate high-resolution rain fields from rain gauge measurements: (i) a classic interpolation technique that consists in interpolating independently the rain intensities at each time step (Eulerian kriging) and (ii) a simple dynamic interpolation technique that incorporates the propagation of the rainy systems (Lagrangian kriging). For this latter approach, three propagation models are tested. The different interpolation techniques are evaluated over three climatically contrasted areas in West Africa where a multiyear 5-min rainfall dataset has been collected during the African Monsoon Multidisciplinary Analyses (AMMA) campaigns. The dynamic interpolation technique is shown to perform better than the classic approach for a majority of the rainy events. The performances of the three propagation models differ from one another, depending on the evaluation criteria used. One of them provides a satisfactory time of arrival of rainfall but slightly smooths the rain intensities. The two others reproduce well the rain intensities, but the time of arrival of the rain is sometimes delayed. The choice of an appropriate propagation algorithm will thus depend on the operational objectives underlying the rain field generation.

1. Introduction

Producing high-resolution rain fields is a key element in several domains: (i) studying the climatology of rainfall at fine space–time scales (e.g. Krajewski et al. 2003; Moszkowicz 2000; Bacchi and Kotegoda 1995); (ii) modeling the hydrological processes on the continental surface because hydrologic, agronomic, or soil–vegetation–atmosphere transfer models require high-resolution forcing rain fields as input (e.g. Michaud and Sorooshian 1994; Gourley and Vieux 2006; Vischel and Lebel 2007; Vischel et al. 2009); and (iii) calibrating/validating satellite rainfall algorithms (e.g. Wolff et al. 2005; Turk et al. 2009). The relevant resolutions for these various purposes may be as fine as 1 km in space and 5 min in time to resolve the convective scale.

Studies dealing with rainfall at small space–time scales are usually based on data obtained from recording rain gauge (RG) networks and/or meteorological radars. The latter provide valuable estimation of rainfall spatial pattern but yield rainfall intensity estimations subject to strong uncertainties (Wilson and Brandes 1979; Krajewski et al. 2010). Rain gages remain so far the most reliable...
sensor producing a direct and accurate measurement of rainfall intensities. However these point measurements need to be spatially interpolated to obtain the areal rain fields required for most of the applications listed above. This interpolation process must take into account two features characterizing small timescale rain fields: (i) their great spatial variability, including an important intermittency, and (ii) a nonnegligible autocorrelation in time. These features are especially marked for regions where rainfall is mainly due to organized propagative convective systems as in tropical regions (Moron et al. 2007). In West Africa, which is the region studied in this paper, Mathon et al. (2002) have calculated that 90% of the annual rainfall is produced by mesoscale convective systems (MCS) for the Sahelian band.

Very few studies dealing with rainfall interpolation at subdaily timescales incorporate information about the rain kinematics (Amani and Lebel 1997; Cheng et al. 2007; Spadavecchia and Williams 2009). Most interpolation methods process each time step independently, neglecting to take into account the valuable information contained in the time autocorrelation of rain fields (e.g. Tsanis et al. 2002; Haberlandt 2007; Tao et al. 2009). Several studies have shown the importance of taking into account the propagation of rainy systems for describing the space–time correlation of rain fields (e.g. Zawadzki 1973; Waymire et al. 1984; Bacchi and Kottegoda 1995).

Acknowledging this, the goal of our study is first to propose and implement a simple dynamic interpolation technique incorporating information on the statistical structure associated with the propagation of the rainy systems. This is done by building on the theoretical framework laid out by Amani and Lebel (1997). The second objective is to assess the performances of this dynamic interpolation approach compared to a classic interpolation technique with no propagation included. More specifically, the paper focuses on the following aspects: (i) three different methods for deriving and accounting for the fields kinematics are tested; (ii) a robust multicriteria evaluation of the interpolation methods is presented, based on a multiyear rainfall dataset from the international African Monsoon Multidisciplinary Analysis (AMMA) project; and (iii) the results are compared over the three dense monitoring networks of the AMMA Couplage de l’Atmosphère Tropicale et du Cycle Hydrologique (AMMA-CATCH; Lebel et al. 2009) observing system to test them under different rainfall regimes.

2. Region and data

a. The AMMA-CATCH observatory

During the 1970s and 1980s, West Africa experienced a generalized drought that is recognized as one of the most significant regional-scale climate anomalies of the twentieth century. This drought associated with a strong anthropogenic pressure has had devastating effects on water resources and agricultural yields, leading to dramatic food and health crises. The AMMA-CATCH observing system, whose data are used here, is a major component of the AMMA field campaigns (Lebel et al. 2010). It has been specifically designed to document the surface processes and their interaction with the atmospheric processes. It is composed of three intensively instrumented mesoscale sites in Mali (Gourma region), Niger (Niamey region), and Benin (Upper Ouémé catchment) that sample the various components of the water cycle across the north–south bioclimatic gradient associated with the West African monsoon (Fig. 1 and Table 1). The Mali and Niger sites are characterized by a Sahelian climate with 150–400 mm (from north to south) and 470–570 mm, respectively, of mean annual rainfall falling within a single monsoon rainy season lasting from June to October and peaking in August. The Benin site is characterized by a Sudanian climate with an annual rainfall ranging from 1200 to 1300 mm and falling from March to October. The rainy season is characterized by two stages: a premonsoon stage (March–mid-June), representing about one-third of the annual total, and the strictly speaking monsoon season (mid-June–October), representing the other two-thirds of the annual total. Over the three studied regions, rainfall is mainly produced by mesoscale convective systems. The associated rain fields display a strong intermittency in time and a strong spatial variability of the intensities.

b. Rainfall data

The AMMA-CATCH rainfall monitoring network is composed of tipping-bucket recording rain gauges available in Mali, Niger, and Benin since 2005, 1990, and 2000, respectively. One tip of the bucket corresponds to 0.5 mm of rain, with its timing being recorded with an accuracy of 1 s. The so-obtained series of tipping times are post treated to produce 5-min rainfall series [see Balme et al. (2006) and Russell et al. (2010) for the detailed algorithm]. In this treatment, a time step is considered as nonrainy when the rain rate is lower than 0.01 mm per 5 min.

The density of the rain gauge networks has varied over the years, but all networks have been reinforced for the extended observed period of AMMA that started in 2005 (for details, see Lebel et al. 2010). The Niger and Benin networks consist of 33 stations in 2005 and 52–56 from 2006. In Mali, 18 rain gauges are available since 2008. The selected periods to evaluate the interpolation methods are 2005–08 for the Niger and the Benin sites and 2008–09 for the Mali site.
Additionally to rain gauges, meteorological radars were installed during the intensive observation period of AMMA in 2006: a C-band radar in Niger [the Massachusetts Institute of Technology (MIT) radar; Russell et al. 2010] and an X-band radar in Benin (X-port radar; Gosset et al. 2010). Radars potentially provide a great potential in characterizing rainfall variability, but specific data treatment is needed to estimate rainfall intensities.

**Table 1. The AMMA-CATCH observatory.**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Niger</th>
<th>Benin</th>
<th>Mali</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Niamay square degree</td>
<td>Ouémé catchment Benin site</td>
<td>Gourma Mali site</td>
</tr>
<tr>
<td>Position</td>
<td>13°–14°N</td>
<td>9°–10°N</td>
<td>14.8°–17.3°N</td>
</tr>
<tr>
<td>Surface</td>
<td>16 000 km²</td>
<td>14 200 km²</td>
<td>27 000 km²</td>
</tr>
<tr>
<td>Environment</td>
<td>Sahelian climate with semiarid vegetation and crops (millet, fallows, and tiger bush)</td>
<td>Sudanian climate (different types of rain systems) and Guinean savanna vegetation</td>
<td>North Sahelian climate (between isohyets 400 and 100 mm); semiarid natural vegetation composed of annual grasses and a sparse tree layer; crops only present in the southern part of the area</td>
</tr>
<tr>
<td>Mean annual rainfall</td>
<td>470–570 mm (from north to south)</td>
<td>1200–1300 mm (from north to south)</td>
<td>150–400 mm (from north to south)</td>
</tr>
</tbody>
</table>
from the raw radar fields. Because the present paper is focused on the evaluation of interpolation methods, radar data will only be used here for a qualitative evaluation of the interpolated rain field patterns.

c. Selection of the rainfall events

The interpolation techniques will be applied to rainy events associated with MCS that passed over the different areas during the periods of study. The identification of these events is based on a two-step process applied on the 5-min rainfall data (adapted from Balme et al. 2006):

(i) The beginning of an event is determined as the first 5-min time step of a period for which at least 30% of the rain gauges record more than 0.5 mm of rain and when at least one station records 1 mm.
(ii) The end of the rain event is defined as the first 5-min time step when no rain is recorded at any of the network stations.

This event definition selects the major rainy systems crossing over the study areas and avoids rainfall due to isolated rain cells. The selected events represent 91% of the total cumulative rainfall in Niger, 85% in Mali, and 88% in Benin. For the period 2005–08, 201 events were identified for the Niger site and 445 events for the Benin site. For the period 2008–09, 88 events were identified for the Mali site.

3. Methodology

a. Classic and dynamic interpolations process

1) PRINCIPLE

The principle of the interpolation methods is illustrated in Fig. 2 with a simplified example. For a comprehensive mathematical formalization, the reader is referred to Amani and Lebel (1997). The system simulated in Fig. 2...
is an MCS characterized by a convective band and a stratiform trail. The system is propagating by translation, with a constant speed and direction. The event hyetograph is stationary in space. The system propagates over four imaginary gages (RG 1–4). The four “measured” hyetographs are plotted in Fig. 2a. They have an identical shape but are delayed in time because of the MCS propagation. In this idealized case, one wants to interpolate the rainfall at the pixel location Pi. With this simplified model, the retrieved hyetograph at Pi should have the same shape as the measured hyetographs RG 1–4 but with a delay in time comprised between RG 2 and 3. The process described for pixel i can be generalized to any pixel or grid in the domain of interest.

Two types of interpolation approaches are compared in the paper (Fig. 2b):

(i) With the classic interpolation approach, each interpolated 5-min rain field is constructed independently of the other time steps.

(ii) The dynamic interpolation approach is a two-step process. First, (Fig. 2b, step 1) the measured hyetographs are shifted in time to maximize the temporal concomitancy or alignment between their shapes. Then the 5-min intensities of the time-shifted hyetographs are interpolated at the pixel of interest. Qualitatively speaking, these shifted hyetographs correspond to the rain intensity signals that an observer would see by following the propagating rain system. This shifting in time comes down to implicitly defining a Lagrangian coordinate system replacing the Eulerian coordinate system used in the classic interpolation technique. The second step (Fig. 2b, step 2) consists in projecting back the interpolated Lagrangian hyetograph into the Eulerian coordinate system. To do so, a time of arrival of the interpolated Lagrangian hyetograph is defined by interpolating the time of arrival of the maximum rain intensity to define the Lagrangian coordinate system. This time of arrival is thus a proxy for modeling the propagation of the rainy system.

In the dynamic interpolation approach, several methods can be used to derive the time shift to be applied to each observed hyetograph to produce the field of Lagrangian hyetograph and thus to retrieve the times of arrival map. Different methods are tested in the paper. They are presented in the following.

2) LAGRANGIAN ORIGIN AND RAIN FIELD INTERPOLATION IN THE DYNAMIC INTERPOLATION APPROACH

Two techniques are commonly used to estimate rainfall propagation from a rain gauge network (Johnson and Bras 1979; Upton 2002; Tsanis et al. 2002). The first technique is based on the detection of the propagation of particular features on the observed hyetographs. The second technique is based on space–time correlation techniques used to assess the velocity and direction propagation of the rainy system. Both techniques are tested here. In our dynamic interpolation approach, we compare three methods for defining the time of arrival of the interpolated Lagrangian hyetographs: two methods (FirstRain and MaxRain) pertain to the first technique, whereas the third method [average synchronized hyetograph (ASH)] pertains to the family of space–time correlation techniques. These methods are illustrated in Fig. 3 for a real case study, a typical MCS that passed over the AMMA-CATCH Niger site on 22 July 2006.

(i) FirstRain method: it is based on the detection at each rain gauge of the first nonzero rain intensity of the event hyetographs. In step 1 of the dynamic interpolation technique (see Fig. 2b), the origin of the Lagrangian coordinate system is defined by adjusting the hyetographs on the first nonzero 5-min rain intensity (Fig. 3a, FirstRain). In step 2, the times of arrival of the first nonzero intensities are then interpolated over the interpolation grid to get a field of propagation of the rainy system.

(ii) MaxRain method: it is based on the detection of the time of arrival of the maximum rain intensity of the hyetographs. The hyetographs are adjusted in the maximum rain intensity to define the Lagrangian coordinate system (Fig. 3a, MaxRain). The field of propagation is obtained by interpolating the times of arrival of the maximum rain intensities (Fig. 3b, MaxRain).

(iii) ASH method: this method was developed by Depraetere et al. (2009) to track the rainfall events over the Benin site. This global method uses an optimization technique to fit a simplified propagating MCS model on the observed hyetographs. The fitted model is similar to the one in Fig. 2, a straight convective band propagating with a constant velocity v and direction d. By exploring the velocity–direction parameter space, the observed hyetographs are shifted in time according to different sets of kinematics parameters $(v_k, d_k)$. The average hyetograph is computed for each parameter set. It is characterized by its peakness defined as the maximum value obtained by a 30-min moving average (Depraetere et al. 2009). The Lagrangian system is determined by the optimal set $(v_{opt}, d_{opt})$ that maximizes the peakness. Then, $(v_{opt}, d_{opt})$ are used to derive the map of time shifts (Fig. 3b, ASH) to be applied to the interpolated Lagrangian rain intensities to come back in the Eulerian coordinate system.
The three models are based on a simplified representation of the rain systems' kinematics. In reality, the MCS motion results from a combination of the mean system propagation with the inner movement and life cycle of the rain cells. By assuming a constant propagation speed of the MCS, the ASH method ignores the rain cell movement. Conversely, the FirstRain and MaxRain methods focus on the identification of the rain cells. The FirstRain method can thus be sensitive to isolated cells preceding the main front of the system; the MaxRain method can be influenced by successive rain cells within the front that might produce multipeak hyetographs. The performances of the dynamic interpolations are expected to be better for well-organized systems without scattered and multiple cell developments and collapses. The dynamic methods will work better for systems with a well-defined space–time organization.

3) CHOICE OF THE INTERPOLATOR

Spatial interpolation is required to produce the fields of rainfall intensities (for both classic and dynamic interpolation techniques) and the fields of time of arrival. From the large variety of existing mathematical interpolators, kriging interpolation has been chosen because...
it gives an unbiased estimation of rainfall and (ii) it is data driven, meaning that it takes into account the spatial structure of the process (here the variogram function is used). Kriging also provides the estimation variance, which is valuable to assess the kriging error; however, this property will not be used here because the transfer of the kriging error from the Lagrangian to the Eulerian coordinate system is theoretically not straightforward and will not be treated in the present paper.

In both Eulerian and Lagrangian coordinate systems, ordinary kriging is used to interpolate the rainfall intensities. Eulerian and Lagrangian empirical climatological variograms of rainfall intensity were computed for the Niger and Benin sites as follows (Lebel and Bastin 1985):

$$\gamma^* = \frac{1}{N_e} \sum_{k=1}^{N_e} \frac{1}{T_k} \sum_{t=1}^{T_k} \frac{\gamma_{kt}^*}{s_{kt}}$$  

where \(\gamma^*\) is the empirical climatological variogram, \(\gamma_{kt}^*\) is the empirical variogram of the rain field at the time step \(t\) of the \(k\)th event, \(T_k\) is the total number of time steps of event \(k\), and \(N_e\) is the total number of events. Exponential models were used to fit the experimental variograms (see example of the climatological variograms computed from the entire event dataset for the Niger site in Fig. 4). The Niger variogram is also used for the Mali site for which an insufficiently long period of measurement hindered a robust estimation of the climatological variogram. The convective and trailing stratiform parts of the rain fields were treated separately. Based on the study of Balme et al. (2006), a mean threshold of 0.5 mm per 5 min was used to separate the convective and the stratiform intensities.

The propagation of the rainy system induces a trend in the fields of the rainfall time of arrival. Therefore, a universal kriging interpolation was preferred. Climatological residual variograms have been estimated by assuming a linear trend adjusted by an ordinary least squares regression. In accordance with the previous works of Guillot and Lebel (1999), an exponential variogram with a 30-km range has been used to represent the structure of the residuals. After kriging, the continuous interpolated values of time of arrival are rounded to the nearest 5-min time step.

The whole data treatment achieved in this study has been coded in the Python programming language (http://www.python.org/) interfaced with the package gstat (Pebesma 2004) implemented in the R environment (R Development Core Team 2009), used here for our kriging-related needs.

### 4) SUMMARY OF THE COMPARED INTERPOLATION METHODS

Four interpolation methods are compared in this study:

- one classic interpolation method based on an Eulerian ordinary kriging of the rain intensities referred to as Eulerian kriging (Eul.Kri) and
- three dynamic interpolation methods based on a Lagrangian ordinary kriging (Lag.Kri) of the rain intensities; these dynamic methods differ from one another by the propagation technique used: namely FirstRain, MaxRain, and ASH. These interpolations are referred to as Lag.Kri FirstRain, Lag.Kri MaxRain, and Lag.Kri ASH.

---

**Fig. 4.** Climatological semivariograms of the (top) convective and (bottom) stratiform 5-min rain intensities into (left) the Eulerian and (right) Lagrangian coordinates for the Niger site.
b. Evaluation of the interpolation methods

The methodologies to evaluate interpolation performances usually make use of (i) independent reference data (e.g., meteorological radar), (ii) a secondary rainfall network, and (iii) a cross-validation process. The first method was inappropriate here for a systematic evaluation because radar data were only available on a limited period and on only two of the three studied regions (see section 2b). The second method would imply to split the rain gauge network into two subnetworks, which would make the evaluation strongly sensitive to the spatial sampling effects (which can be significant as discussed in section 5b). The performances of the four interpolation techniques were thus assessed by a leave-one-out cross-validation procedure. The cross validation is computationally time consuming, but it limits the spatial sampling effect.

The evaluation will first focus on the way the interpolation methods reproduce the main features of the rain fields (propagation, intermittency, and intensity distribution; results presented in section 4b). Then, an event-based evaluation of the best interpolation method will be carried out (results presented in section 4c). The different criteria used for this evaluation and comparison process are presented in the following.

1) Evaluating the Realism of the Propagation Models

The realism of the propagation model is evaluated by analyzing the simulated time of propagation of the events compared to the observed time of propagation. This analysis must help in selecting the events that do not satisfy the preconditions required by the dynamic interpolation concerning the organization of the rainy systems.

2) Evaluating How the Interpolation Methods Reproduce the Intermittency of Rainfall

The ability of the interpolation methods to model the rainfall intermittency is evaluated on the basis of binary rain–no-rain contingency tables. Through the cross-validation procedure, one contingency table per event is built by accounting for all the time steps of all the rain gauges. The contingency tables are analyzed through the computation of the false alarms ratio (FAR), which is defined for a given event as

$$\text{FAR} = \frac{\sum_{i=1}^{T} \sum_{t=1}^{N} \text{FalseAlarms}\left[R_i^a(t), R_i(t)\right]}{\sum_{i=1}^{T} \sum_{t=1}^{N} \text{Hits}\left[R_i^a(t), R_i(t)\right] + \sum_{i=1}^{T} \sum_{t=1}^{N} \text{FalseAlarms}\left[R_i^a(t), R_i(t)\right]},$$

where $T$ is the number of time steps of the event, $N$ is the number of rain gauges, and $R_i(t)$ is the rain intensity at time step $t$ for the station $i$.

Hits correspond to positive rain for both observation and prediction,

$$\text{Hits}\left[R_i^a(t), R_i(t)\right] = \begin{cases} 1 & \text{if } R_i^a(t) > 0 \text{ and } R_i(t) > 0 \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

FalseAlarms corresponds to no rain for observation and rain for prediction,

$$\text{FalseAlarms}\left[R_i^a(t), R_i(t)\right] = \begin{cases} 1 & \text{if } R_i^a(t) > 0 \text{ and } R_i(t) = 0 \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

FAR gives a measure of the model’s tendency to predict rain where none was observed, which is the main drawback of most interpolation methods to produce rain in nonrainy areas. The prediction of zero rain where positive rainfall was observed is negligible. FAR was thus considered consistent to evaluate the ability of the methods to reproduce a realistic intermittency. A 0 value of FAR indicates that the interpolation method satisfactorily estimates the rainfall intermittency, whereas a 1 value indicates a complete inability to predict the zeros. To be coherent with the definition of nonrainy time steps in the observations (see section 2b), all interpolated intensities lower than 0.01 mm per 5 min were considered as 0 values.

3) Evaluating the Distribution of Rain Intensities

The distributions of rain intensities resulting from the interpolation are compared to the observed intensity distribution through quantile–quantile plots.

4) Determining the Best Interpolation Methods

Elected the best interpolation method depends upon the user objectives: what are the rain features that must
be well reproduced for the targeted application? For instance, for the modeling of long-term water resources, one will be interested in predicting the rain intensities that are responsible for the partition between infiltration and runoff; in agricultural applications, the rainfall intermittency will probably be the targeted feature; and, for designing a flood warning system in operational hydrology, one will be interested in well predicting both the timing and the magnitude of the rainy system. There is thus a need to (i) define criteria to compare the observed and modeled hyetographs based on the cross-validation procedure and then (ii) assess which method performs the best according to the user requirements. Three objective functions are used to that end.

(i) The determination coefficient \( r^2 \)

The determination coefficient \( r^2 \) is

\[
r^2 = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} (R_i(t) - \bar{R})(R^*_i(t) - \bar{R}^*_t))^2}{\sum_{i=1}^{N} \sum_{t=1}^{T} (R_i(t) - \bar{R})^2 \sum_{i=1}^{N} (R^*_i(t) - \bar{R}^*_t)^2}, \tag{5}
\]

where \( T \) is the number of time steps of the event, \( N \) is the number of rain gauges, \( R_i(t) \) is the rain intensity at time step \( t \) for the station \( i \) (in mm per 5 min), and \( \bar{R} \) is the mean rain intensity averaged over the whole stations and the whole time steps of the event (in mm per 5 min). The determination coefficient is a good indicator of the time concordance between the observed and modeled hyetographs and is mostly sensitive to the position of the peak of intensity of the hyetograph. It is thus tailored to evaluate the capacity of the interpolation methods to reproduce the timing of the hyetographs.

(ii) The KGE

The Kling–Gupta efficiency (KGE) is

\[
\text{KGE} = 1 - \sqrt{(1 - r)^2 + (1 - \alpha)^2 + (1 - \beta)^2}, \tag{6}
\]

where \( r \) is the linear correlation coefficient [square root of the determination coefficient in Eq. (4)], \( \alpha \) is a measure of relative variability, which is defined as

\[
\alpha = \frac{\sqrt{\sum_{i=1}^{N} \sum_{t=1}^{T} (R^*_i(t) - \bar{R}^*_t)^2}}{\sqrt{\sum_{i=1}^{N} \sum_{t=1}^{T} (R_i(t) - \bar{R})^2}}, \tag{7}
\]

and \( \beta \) is the bias, which is defined as

\[
\beta = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} R_i^*(t)}{\sum_{i=1}^{N} \sum_{t=1}^{T} R_i(t)}. \tag{8}
\]

KGE evaluates the interpolation by equally weighting the correlation, the relative variability, and the bias. It can thus be seen as a compromise between a fair modeling of both timing and shape of the hyetographs. Gupta et al. (2009) proposed it as an alternative objective function to overcome the problems associated with the widely used mean-square error (MSE) and its related normalization, the Nash–Sutcliffe efficiency (NSE; Nash and Sutcliffe 1970). Gupta et al. (2009) indicate that, to maximize the MSE or the NSE, the variability has to be underestimated, which is not the case for the KGE. As a consequence, for a given bias, the MSE and the NSE are much more influenced by the correlation component than by the relative variability. This has been verified in the present study because the MSE and NSE were tested and were shown to behave similarly to \( r^2 \) (not shown here).

(iii) KGE_move

Both \( r^2 \) and KGE remain sensitive to the hyetograph timing. This can be seen as a limitation because, if a hyetograph is modeled with a shape similar to the observed hyetograph but slightly delayed in time, despite an obvious skill of the interpolation to correctly predict the rainfall intensities, the cross-validation will produce low values of \( r^2 \) and KGE because of the time lag between prediction and observation. To strictly evaluate the modeling of the hyetograph shapes independently of the timing, a new objective function called the moving Kling–Gupta efficiency (KGE_move) has been defined. KGE_move is the KGE value obtained by replacing \( R_i^*(t) \) by \( R_i^*(t + \Delta t) \) in Eqs. (8), (7), and (5), with \( \Delta t \) being a time shift adjusted independently at each station \( i \) to maximize the correlation between the observed and the predicted hyetographs.

4. Results

a. Qualitative comparison for the case study of 22 July 2006 in Niger

Figure 5a shows three successive rain fields estimated from the MIT radar data for the event of 22 July 2006 over the Niger. The radar-derived rain fields have their own uncertainties, linked essentially to ambiguous conversion between radar reflectivity and rain rates (Russell et al. 2010). However, radar provides a good rendering of the spatial structure of the fields with a frequent time
FIG. 5. Rain fields at (left) 0810 local time (LT), (middle) 0900 LT, and (right) 0950 LT 22 Jul 2006 estimated by (a) the MIT meteorological radar; (b) classic interpolation: Eul.Kri; (c) dynamic interpolation: Lag.Kri with the ASH kinematic model; (d) dynamic interpolation: Lag.Kri with the MaxRain propagation model; and (e) dynamic interpolation: Lag.Kri with the FirstRain propagation model.
sampling. The fields are characterized by extended non-rainy areas, particularly ahead of the convective front (left of the radar images) and in the rear part of the stratiform trail mainly associated with the external intermittency of the system. Some of the spatial variability and the occurrences of patch of lower rain rates, behind the convective cells, might be exaggerated by the effect of attenuation. Within the system, rainfall intensities are strongly variable in space, but the overall pattern, marked by a line of convective cells and more widespread rain behind, exhibits some similarities among the three images. This illustrates the time correlation between successive rain fields.

Figures 5b–e show the interpolated fields obtained from the different interpolation methods at a 5-min time step. Figure 5b shows that Eul.Kri fails in retrieving the inherent characteristics of the rain fields at small time scales. The intermittency is not correctly reproduced, especially where the rain gauges density is low (see, e.g., the westernmost part of the AMMA-CATCH Niger area). This is linked to the basic property of ordinary kriging, which tends to the mean value of the interpolated process, when the nearest measurement points are too remote to provide information. The lack of resemblance between the spatial patterns of the successive rain fields illustrates the consequence of ignoring the space–time correlation in the interpolation technique. The Lag.Kri methods (Figs. 5c–e) reproduce much better the features of the rain fields: the modeling of the intermittency and the similarity of the successive spatial patterns is improved, and the rainfall intensities are more clearly organized within the convective front and the stratiform trail of the MCS, which are now more clearly distinguishable. The fields are still smooth compared to the higher-resolution radar data but much improved compared to the static interpolation. The three tested propagation models produce different rainfall patterns. The ASH method (Fig. 5e) produces a rectilinear pattern that seems to be a bit artificial as the front of the observed system is obviously not a straight line. The MaxRain and FirstRain methods produce more variable spatial patterns more resembling the radar patterns. A quantitative and systematic characterization and evaluation of the interpolation methods is presented in the following.

b. Retrieval of the features of the rain fields: Propagation, intermittency, and intensity distributions

1) EVALUATION OF THE PROPAGATION MODEL REALISM

A pre-analysis of the propagation fields revealed some difficulties of the FirstRain, MaxRain, and ASH models to retrieve realistic times of propagation for some events. For these events, the models produced unexpectedly long times of propagation (event duration of several days over the study areas instead of the usually observed 3–12 h). Most of these events were associated with scattered rain cells with no clearly organized propagation patterns. As explained in section 3a, for such disorganized systems the use of a dynamic interpolation technique does not make sense and the classic interpolation is preferable by default. It was then considered that a propagation field is failing when a significant part of the total event rain amount is predicted out of the time boundaries of the observed event. The proportion of failing events is reported in Table 2 for different percentages of rainfall amount interpolated outside of the observed temporal boundaries. For each of the three sites, the ASH propagation model has the highest rates of failures. For instance, 12% of the events in Niger have 10% of their rainfall predicted outside of the
observed time boundaries, with 22% in Benin and 46% in Mali. However, the low percentage of the associated cumulative rainfall (3%, 11%, and 26% for Niger, Benin, and Mali, respectively) shows that these events are more likely to be associated with small rainy systems, which are also the most disorganized. According to the statistics of Table 2, the methods MaxRain and FirstRain seem to have less difficulty to keep the time of propagation in a realistic order of magnitude. One will note, however, that this result is not a guarantee against a possible unrealistic modeling of the timing of the rain intensities within the event. The present evaluation just allows the identification of events for which the dynamic approach produces unrealistic propagation fields. The results presented in the following were obtained by selecting for each method only those events for which less than 5% of the total interpolated event rainfall amount is out the observed time boundaries.

2) Evaluation of the Rainfall Intermittency

FAR [Eq. (2)] was computed for events of increasing intensities (Fig. 6) for each of the four interpolation methods. As a common behavior over the three study areas, the FAR values are always lower for the three Lag.Kri methods than for the Eul.Kri method, demonstrating that incorporating some information of the rainy system propagation into the interpolation improves the modeling of the nonrainy areas. In average, the FirstRain method reproduces the best the intermittency, whereas ASH and MaxRain perform variably, depending on the study area (MaxRain always performs better than ASH for the Mali site, whereas ASH performs better than MaxRain for intense events in Niger and Benin). FAR values decrease when the intensity of the event increases, which is coherent with the fact that the most intense events have the greatest spatial extension; for these events, there is thus more chance to correctly predict the intermittency, which is obviously smaller. However, FAR computed on the Benin site displays a singular behavior for the most intense events (quantile 95% corresponding to events with mean cumulative rainfall greater than 50 mm) because the values of all methods surprisingly increase and modify the ranking of the method performances. This trend is associated with a complex space–time organization of the MCS in Sudanian regions, which makes it difficult to identify the rainy structures and the propagation of the system. This aspect will be further discussed in section 5a.

3) Rain Intensity Distribution

To assess how the interpolation methods reproduce the rain intensities, quantile–quantile plots comparing the observed and interpolated rain intensities are presented in Fig. 7. For all plots, the curves are above the first bisector for the lowest rain intensities and below for the highest. This behavior is a well-known shortcoming of the kriging process that tends to smooth the rainfall patterns (see, e.g., Vischel et al. 2009). The underestimation affects the intensities greater than 1 mm per 5 min (12 mm h$^{-1}$ in 5 min, corresponding roughly to the quantile 90%), meaning the highest intensities, which are of first importance to characterize the climatology (Balme et al. 2006) as well as the hydrology in the region (Vischel and Lebel 2007). Comparing the observed values in Niger, the quantile 99% (corresponding to the observed intensity of 3.2 mm per 5 min or 38 mm h$^{-1}$) is 38% lower for Eul.Kri, 27% lower for Lag.Kri ASH and FirstRain, and 19% lower for...
Lag.Kri MaxRain. From Fig. 7, it can be noted that, for the three sites Eul.Kri underestimates the most, the rain intensities. The ASH method is in between the Eul.Kri and the FirstRain method, except for the Mali site, where ASH is under the Eul.Kri curve. For all sites, the underestimation is less pronounced for the MaxRain method. This behavior is coherent with the adjustment of the hyetograph within the Lagrangian coordinate system: by adjusting the maximum intensities all together, it is quite logical that the MaxRain method favors the interpolation of the highest intensities.

c. What is the best interpolation method?

The previous results showed that, although the Lag.Kri approach performs significantly better than the Eul.Kri approach, none of the three dynamic interpolation methods satisfactorily predicts both the intermittency and the intensity distribution of the rainy systems; also, there is not a method that is always performing better than the others. To gain more insight into the behavior of each method, they were ranked for each event and each of the three criteria. The resulting statistics are reported in Table 3.

Whatever the objective function used to compare the methods, most of the time the Lag.Kri techniques perform better than Eul.Kri for the three studied regions. From 82% to 86% of the events (depending on the criterion) are better interpolated by the Lag.Kri techniques over the Niger site. These percentages vary from 53% to 75% in Benin. The performances of the Lag.Kri approach are lower for the Mali site: except for the KGE_move criterion (61%), only 40% (for $r^2$) and 37% (for KGE) of the events are best modeled by the Lag.Kri methods. These lower percentages are mainly explained by the difficulty of the Lag.Kri techniques to detect the rainfall propagation over the Mali network: indeed, a fourth of the events are interpolated with the Eul.Kri by default.

Table 3. Best methods of interpolation according to the $r^2$, KGE, and KGE_move criteria. Eul.Kri default corresponds to the percentage of events for which the propagation modeled by ASH, FirstRain, and MaxRain were considered as nonrealistic in a sense defined in section 4b and Table 2.

<table>
<thead>
<tr>
<th></th>
<th>$r^2$</th>
<th>KGE</th>
<th>KGE_move</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage of events</td>
<td>Percentage of events</td>
<td>Percentage of events</td>
</tr>
<tr>
<td>Niger</td>
<td>Eul.Kri default</td>
<td>17.9</td>
<td>17.5</td>
</tr>
<tr>
<td></td>
<td>Lag.Kri ASH</td>
<td>82.1</td>
<td>82.5</td>
</tr>
<tr>
<td></td>
<td>FirstRain</td>
<td>14.4</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td>MaxRain</td>
<td>8.5</td>
<td>33.3</td>
</tr>
<tr>
<td>Benin</td>
<td>Eul.Kri default</td>
<td>42.7</td>
<td>46.3</td>
</tr>
<tr>
<td></td>
<td>Lag.Kri ASH</td>
<td>57.3</td>
<td>53.7</td>
</tr>
<tr>
<td></td>
<td>FirstRain</td>
<td>10.3</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>MaxRain</td>
<td>6.3</td>
<td>25.8</td>
</tr>
<tr>
<td>Mali</td>
<td>Eul.Kri default</td>
<td>59.1</td>
<td>62.5</td>
</tr>
<tr>
<td></td>
<td>Lag.Kri ASH</td>
<td>40.9</td>
<td>37.5</td>
</tr>
<tr>
<td></td>
<td>FirstRain</td>
<td>11.3</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>MaxRain</td>
<td>11.4</td>
<td>17.0</td>
</tr>
</tbody>
</table>

FIG. 7. Quantile–quantile plots comparing the observed and the interpolated rainfall intensities: (left) Niger, (middle) Benin, and (right) Mali.
Among the three dynamic techniques, the ASH method is the most often selected as the best method according to the $r^2$ criterion. According to KGE, MaxRain is the most frequent best method, but ASH remains a reliable interpolation for a large portion of events, especially in Niger. MaxRain is mostly selected for the KGE_move criterion. These differences result from the varying capacity of the methods to model both the shape and the timing of the hyetographs. In Fig. 8, MaxRain and ASH performances are compared for the case of Niger (similar behaviors were noticed for the two other sites; not shown here). The KGE_move values were lower for ASH than for MaxRain for all events. This means that the hyetograph shapes modeled by ASH are less realistic than those produced by the MaxRain method. Actually, by imposing a constant average propagation of the system, the ASH method tends to spread the rain intensities over the time. This produces less “peaky” hyetographs than the MaxRain method for which the Lagrangian origin adjusted on the maximum rain intensity accentuates the peaks of intensities. However, the better values of $r^2$ obtained by the ASH method show that the timing is more appropriately modeled by ASH than by MaxRain. ASH thus behaves as if seeking a compromise between reproducing the hyetograph shapes and reproducing their time of arrival, whereas MaxRain produces more realistic hyetograph shapes at the price of less accuracy in their timing. The FirstRain method displays an intermediate behavior between ASH and MaxRain.

5. Discussion

Different interpolation performances over the three studied regions have been noticed in section 4. These differences are clearly visible in Fig. 9, where box plots are summarizing the distribution of the values obtained for each criterion on the three sites. Only the values of the best method (regardless of the type) were retained for each event site. Significantly better scores are obtained for the Niger site than for the Benin or Mali sites. Two factors explain this: (i) a larger variety of rain events in Benin and (ii) a pattern and density of the rain gauge network producing larger sampling errors on the Mali site.
a. Effect of rainfall typology on interpolation performances

In the Sahelian region, it has long been known that the fast and well organized MCS are the most active in term of rainfall (see Mathon et al. 2002; Laing et al. 1999). In average, the organized convective systems defined by Mathon et al. (2002) produce 80% of the annual rainfall on the Niger site. The superiority of the Lag.Kri techniques for these rain fields is quite intuitive. Such systems are also responsible for a significant share of the annual rainfall on the Oue´ me´ catchment of Benin, but Depraetere et al. (2009) have shown that large event rainfall may also be produced by situations where several midsize MCS hit different parts of the region of study at close time intervals. The resulting space–time pattern of the rain fields is somewhat complex, with strong changes from one time step to the next and no clear kinematics identifiable from a network of gauges. Two remarkable examples of such events are given in Fig. 10. They are characterized by successive rainy structures spread over several days, but our algorithm of event separation could not distinguish these various rainy structures because it kept raining over at least some part of the area during the whole period. Such events are characterized by a significant intermittency at small time steps, an intermittency that no interpolation method can retrieve properly. However, they are also characterized by a large amount of cumulative rainfall (more than 50 mm). In our sample of 445 Benin site events, more than 10 major events are presenting characteristics similar to those of the events of Fig. 10; they are responsible for the singular shape of the FAR curves in Fig. 6. This relatively small but still significant proportion of “multi MCS” rain is part of the larger diversity of rainy systems characterizing the Sudanian regions, in comparison to the Sahelian region. Rain gauge networks are less suitable for studying these complex rainy structures, and it is difficult to find suitable methods to dissociate these rainy structures from the rain gauge data.

b. Effect of network density and pattern on interpolation performances

Although both the Mali and the Niger sites are characterized by a similar Sahelian rainfall regime where fast moving and well organized MCS dominate, the pattern and density of the rain gauge networks operated on these two sites are very different (Fig. 1). Keeping in mind that the MCS are moving from east-northeast to west-southwest, it is obvious from Fig. 1 that the Mali site network is much less fitted to capture the dynamics of these systems, and it is thus not surprising to obtain significantly less good interpolation scores over this site (Fig. 9), with a large part of the events being by default interpolated by Eul.Kri (25% against 6% for the Niger site; Table 3). Keeping aside the pattern influence, the sole effect of the density was tested on the Niger site by decreasing the number of rain gauges from the initial total of 33 in 2005 and 52–56 in 2006–08 to 30 in a first step and then to 20 (a number close to the 18 rain gauges available on the Mali site). As expected, decreasing the number of rain gauges decreases the interpolation performances. In Fig. 11, the differences of \( r^2 \), KGE, and KGE_move values between the 20 gauges and the full rain gauge networks are on the same order of magnitude of the differences between the Niger and Mali site reported in Fig. 9. This confirms that a large part of the low performances of the interpolation in Mali is due to sampling effect. The spatial undersampling...
also increases the proportion of the event interpolated by the Eul.Kri (Fig. 12). This proportion reaches 17% for the 20 gauges network, which indicates that the still higher proportion (25%) obtained on the Mali site is indeed due to the additional effect of the pattern of the network. It may be concluded that a regular pattern of rain gauges separated by a distance on the order of 20 km allows us to retrieve some useful information on the dynamics of the MCSs, an information allowing dynamic interpolation methods to perform better than classic methods.

6. Conclusions

Rain field estimation at small time steps remains both a key requirement and a challenge in regions of the world where the water cycle is strongly sensitive to small time scale processes, even more so when rainfall is characterized by a strong space–time variability and intermittency. Making use of a large dataset of 734 rain events collected in three climatically contrasted West African regions (Ouémé catchment in Benin, 10.5°N; Niamey region in Niger, 13.5°N; and Gourma region of Mali, 14.5°N), it was shown here that, in most cases, interpolation methods incorporating some information about the time structure of successive 5-min rain fields perform significantly better than a classic interpolation method treating each rain field independently of the previous and next rain fields.

Three algorithms were proposed to account for the time structure of rainfall. Rather than identifying explicitly a time autocorrelation or a 3D covariance function, these three algorithms are based on propagation models allowing us to (i) define a Lagrangian coordinate system in which rainfall intensities are interpolated and (ii) shift the interpolated intensities back into the Eulerian coordinate system. The three propagation algorithms differ by the way the origin of the Lagrangian coordinate system is defined: in the MaxRain and the FirstRain algorithms the hyetographs are adjusted on the maximum and the first rain rates, respectively, and in the ASH algorithm the hyetographs are adjusted according to the determination of a mean and unidirectional speed of the rainy system. Different criteria have been used to intercompare their results, based on a cross-validation procedure. The dynamic interpolation (all algorithms taken together) was shown to perform better than the classic interpolation for 82%–86% of the 201 events in Niger, 57%–75% of the 445 events in Benin, and 40%–61% of the 88 events in Mali, depending on the evaluation criterion used. Among the three algorithms, the MaxRain method is the best in modeling the shape of the hyetograph but the timing is sometimes delayed, and the ASH method provides a good compromise between the shape and the timing but tends to smooth the shape of the hyetographs. The FirstRain method behaves in between the two methods.

Despite their good performances, one should notice that none of the three dynamic algorithms succeeds entirely in accounting properly for the three important properties of the small time step rain fields: that is, rainfall propagation, rain field intermittency, and rain intensity distribution. Also, it was shown that these algorithms fail for some type of rain fields for which the Eulerian ordinary

FIG. 11. Box plot of (left to right) $r^2$, KGE, and KGE_move for a decreasing RG density in Niger.

kriging performed better and that their performances depend on the pattern and density of the measuring network. For these reasons, it would be unwise to blindly use any of these algorithms. One important factor to keep in mind is the user’s requirements: in some applications, a good estimation of the rainfall propagation is a priority, whereas, in others, a good simulation of the intensities and/or of the intermittency is more important.

The significantly better accuracy of the rain fields interpolated through the dynamic method proposed here led to use them as reference rain fields for the ongoing AMMA Land Surface Model Intercomparison Project Phase 2 (ALMIP-2) (Boone et al. 2009). The validation of high-resolution rain products derived from indirect measurements such as weather radar or satellite is another potential use for our dynamically interpolated rain fields. It is indeed unsound to compare the “instant” rain maps produced by remote sensing with the smooth interpolated rain field traditionally derived from gages. The dynamically interpolated rain fields produced 5-min rain fields are much better suited as a reference ground product for these applications, and it is planned to use them in future ground validation exercise, such as for the future Megha-Tropiques mission (Desbois et al. 2007; http://meghatropiques.ipsl.polytechnique.fr/).

In the present study, the ordinary kriging interpolation has been used in both Eulerian and Lagrangian coordinate systems. Obviously, a large variety of other interpolation methods might also provide a suitable framework to model the characteristics of high-resolution rain fields (for a review, see Amani and Lebel 1997). A strong argument for adopting the three proposed dynamic algorithms is that they do not require inferring a complex 3D structure function, making their numerical implementation relatively easy.

Future developments will focus on four aspects: (i) comparing the dynamic kriging with other interpolation methods, (ii) incorporating into the dynamic interpolation methods radar and satellite information, (iii) studying the dependence of the dynamic interpolation performances to the time step of the rainfall data, and (iv) defining a theoretical framework to define an interpolation error model allowing the development of a 5-min rain field stochastic generator conditioned by point values.

Acknowledgments. This research was funded by IRD in the framework of the AMMA-CATCH “ORE” program, as well as by the French ECCO-PNRH program (project “Eau et végétation au Niger”). Based on a French initiative, AMMA was built by an international scientific group and is currently funded by a large number of agencies, especially from France, the United Kingdom, the United States, and Africa. It has been the beneficiary of a major financial contribution from the European Community’s Sixth Framework Research Programme. Detailed information on scientific coordination and funding is available on the AMMA International Web site (http://www.amma-international.org). The authors thank the two anonymous reviewers for their constructive and detailed comments, which helped improve the study and the manuscript.

REFERENCES


