

Stochastic Space-Time Regional Rainfall Modeling Adapted to Historical Rain Gauge Data

Zepu Zhang^a and Paul Switzer^b

Abstract

Stochastic rainfall models are important tools both for practical issues and in studies of weather- and climate-sensitive systems. We propose an event-based model, continuous in space (two-dimensional) and time, that describes regional-scale, ground-observed storms by a Boolean random field of rain patches. The model creates complex space-time structures with a mathematically tractable framework. The estimation method relates temporal observations at fixed sites to the movement of the model storm rain field, thereby making historical rain gauge data suitable for model fitting. The model is estimated using hourly historical data at eight rain gauges in Alabama and tested for its capabilities in capturing statistical characteristics of the historical data, including rainfall intensity, rainfall intensity extremes, temporal correlation, effects of temporal aggregation, spatial coverage, and spatial correlation.

Keywords: precipitation, regional modeling, mesoscale meteorology, stochastic hydrology, space-time structures, Boolean model, stereology, rain gauge data.

1 Introduction

Weather sequences generated by stochastic models are often used in process simulations because historical weather data may be inadequate in terms of length, spatial coverage, and completeness. Moreover, weather sequences generated from stochastic models provide a mechanism for investigating the implications of weather uncertainty in process models.

Because of its important role in a broad range of land surface processes, rainfall has been one of the most actively investigated elements in weather generator models. Over the past four decades, stochastic rainfall models have evolved through several generations; see review articles by, for example, Wilks and Wilby [1999], Onof et al. [2000], and Wheeler et al. [2005]. Among the more popular types are alternating renewal models [Green, 1964; Roldan and Woolhiser, 1982], Markov chain models [Chin, 1977; Katz, 1977; Richardson, 1981; Chandler and Wheeler, 2002], clustered point process models [Kavvas and Delleur, 1981; Smith and Karr, 1983; Waymire et al., 1984; Rodriguez-Iturbe et al., 1987; Cox and Isham,

^a Center for Integrating Statistical and Environmental Science, University of Chicago
(zpzhang@uchicago.edu)

^b Department of Statistics, Stanford University (switzer@stanford.edu)

Water Resources Research, submitted 10 October 2005; revised 24 July 2006; revised 15 October 2006; accepted 17 October 2006. <http://dx.doi.org/10.1029/2005WR004654> Copyright (2007) American Geophysical Union.

1988; Cowpertwait, 1995; Northrop, 1998], and downscaling models [Wilby et al., 1998; Ferraris et al., 2003]. Many of the early models, such as those in the Markovian process framework, are “observation-based” in the sense that they make statistical assumptions on certain properties of rainfall, then construct the model and estimate parameters based directly on statistical analysis of observed data. Some more recent approaches, exemplified by clustered point process models, may be called “event-based” as they describe and simulate rainfall starting from a simplified prescription of how the storm event actually occurs and develops.

These two types of modeling approaches differ in two basic respects: discreteness and spatial coverage. Observation-based models naturally arose from analysis of daily or hourly rainfall records at a single station. When extended to spatial models, they concern rainfall at multiple discrete locations as opposed to operating in a continuous spatial domain. In order to express the inter-correlation of rainfall between locations, a complex covariance structure and a large number of parameters are often needed [Smith, 1994; Wilks, 1998]. In contrast, event-based models can be intuitively played out in a continuous spatial domain, evolving seamlessly in time without artificial aggregation with regard to clock-time intervals. This major advantage is possible because at the heart of these models is a quasi-physical picture of the rainfall process. Prescription of this simplified rainfall mechanism is where much effort is devoted in developing these models.

In this paper we propose an event-based regional model along the line of point process models. At the center of the idea is a Boolean model [Matheron, 1975; Serra, 1982; Stoyan et al., 1995; Molchanov, 1997], which consists of a spatial Poisson point process and additional properties attached to the points. The points are the center of rain patches, within which rainfall intensity varies according to a prescribed profile. This model has a clear spatiotemporal structure, works in continuous spatial and temporal scales, and can be estimated with widely available long term historical data. To estimate the model parameters, stereological relations of the Boolean field are used. The proposed estimation strategy is well suited to the way rainfall is observed and recorded at rain gauges.

We first describe the formulation of the model, then give a full account of the model fitting procedure. To validate the model against historical data, we use simulations and analytic derivations to examine several statistical properties of rainfall that such stochastic models are expected to capture. The presentation is illustrated throughout using a historical hourly rainfall data set, which we introduce now, before turning to the model itself.

2 The Illustrative Data

Ground-based observational climate data are maintained by the National Climatic Data Center (NCDC, www.ncdc.noaa.gov) in Asheville, NC. Precipitation data come in several time resolutions such as daily, hourly, and quarter-hourly, with different time coverage at different stations throughout the nation. We chose data from the hourly data set TD3240 at eight stations in Alabama (Figure 1) in the years from 1949–1961 for illustrations. These data were chosen because the region has a relatively simple topology and multiple stations distributed in a balanced configuration. The record period, from 1949 to 1961, corresponds to the period when all stations provided high-resolution data.

Fig 1

Records in this data set are hourly precipitation amounts, accurate to 0.254 mm (0.01 inch). These hourly aggregated data derive from the standard tipping-bucket rain gauge, which consists of a bucket that collects rainfall and tips every time it fills. The number of tips during each hour enters the historical record and reflects the rain amount discretized as multiples of the bucket capacity, 0.254 mm.

Common dry periods across all eight stations exceeding 12 hours are taken to be intervals between storms. By this empirical criterion, 918 storms are identified in the illustrative data. The particular choice of 12-hour dry period as a storm separator precludes individual storms with extended dry periods. The choice to model storms without extended dry periods contrasts with rainfall models having clustering mechanisms.

For each year, data for the period between April 1 and October 31 were extracted for model estimation and validation. Model parameters during this summer season were treated as temporally homogeneous.

3 Formulation of the Regional Boolean Rainfall Model

We model the regional time-varying two-dimensional surface rainfall intensity field for a single storm by “rain patches” moving across the region of interest (Figure 2). These model rain patches should be interpreted as elementary, possibly overlapping, rain areas as observed on the ground rather than as cloud structures. The geographic scale of the model rain field, consisting of rain patches, is taken to be large relative to the monitoring region of interest. Storm size is the length along the direction of storm movement, i.e., storm size is the product of storm speed and duration. A modeled storm has a fixed velocity (including speed and direction).

Fig 2

The internal structure of the modeled rain field is described by an isotropic Boolean model, which is commonly used in stereology and stochastic geometry. The Boolean model is relatively simple, yet its general and flexible structure can accommodate specific variants (see [Matheron \[1975\]](#), [Serra \[1982\]](#), [Hall \[1988\]](#), [Stoyan et al. \[1995\]](#), and [Molchanov \[1997\]](#)). Realizations of the field may appear complex and irregular, but the model is mathematically tractable in many respects and Boolean fields are straightforward to simulate (see Section 5).

In our context, circular rain patches are randomly located with centers forming a homogeneous spatial Poisson point process on the plane. Each rain patch has a random radius and a random mean rainfall intensity, both assumed independent of the location of the patch. Different rain patches are independent of each other. In particular, the patches are free to overlap, thus allowing for complex shapes of connected rain areas on the ground. The rainfall intensity at a location covered by overlapping patches is the sum of the intensities at the corresponding location within each participating patch.

The rainfall intensity within a rain patch is assigned a time-invariant profile which peaks at the patch center and decreases linearly towards zero on the edge of the patch. This is more realistic [[Konrad, 1978](#); [Goldhirsh, 1983](#)] than assuming a constant intensity over the entire patch [[Goldhirsh, 1986](#); [Cox and Isham, 1988](#)]. We make no assumption of independence between the size and rainfall intensity of a patch. The model uses an empirically derived joint distribution of patch size and intensity.

This simple model of a storm, that is, a spatial Boolean field of rain patches moving at a constant velocity, introduces complex space-time structures. A fixed location on the ground, such as a rain gauge, experiences an alternation of dry and wet periods corresponding to a linear transect of the planar field. We define the following geometric attributes on the transect:

- Chord: intersection of the transect with a rain patch.
- Clump: a contiguous wet interval on the transect.
- Gap: a dry interval on the transect.

Figure 2 contains examples of these attributes along two transects **AB** and **CD** through the rain field. A clump consists of a single rain patch chord or overlapping chords. Gaps are dry segments between consecutive clumps. The relations between patches, chords, gaps, and clumps form the basis of inferring properties of rain patches from dry and wet intervals observed at rain gauges.

4 Estimation of the Model

The regional event-based model is characterized by

1. The joint distribution of storm velocities and storm sizes.
2. The Poisson density of patch centers in the model rain field for a storm.
3. The joint distribution of patch radii and average patch rainfall intensities.

Each storm is a random realization of a moving Boolean field. The available historical data used for model estimation consist of the start and stop times of continuous rain periods (clumps) at each of the monitoring stations, together with the rainfall total in each rain period.

The size and velocity will be estimated for each storm in the historical data, therefore we will have empirical distributions of storm size and storm velocity. Estimating storm velocity entails comparing time-series observations at multiple rain gauges that are in a fixed geographic configuration. The estimated velocity is used to convert dry and wet durations in the historical time series data to lengths of gaps and clumps in the modeled Boolean field. Storm size is derived from storm velocity and duration.

Since the model is regionally homogeneous, we pool the gaps and clumps from all the monitoring stations in the region, and over all storms, to form combined samples. Statistics of gaps and clumps of the Boolean field are related stereologically to the density and size of the patches, and these relations are used to estimate the Poisson patch density and patch size distribution. Estimation of the patch rainfall intensity distribution is based on simulations using the estimated patch density and patch size distribution, together with the observed rainfall amounts in wet periods.

Below we discuss issues in the order they occur in the estimation process. Notation used in the model is listed in Table 1.

Table 1: Notation used in the model, especially in Section 4. The symbols are grouped by the geometric object they are used for.

symbol	source	meaning
D	observed	(dry) gap size, km
μ_D	estimated	mean D , km
W	observed	(wet) clump size, km
F_W	empirical	distribution function of W
u	estimated	storm speed, km/hr
θ	estimated	storm direction
R	estimated	patch radius, km
λ_R	estimated	Poisson density of patch centers, km^{-2}
F_R	estimated	distribution function of R
μ_R	derived	mean R , km
r_0	specified	minimum R , km
q	(multiple)	rainfall intensity
C	estimated	chord length, km
λ_C	estimated	Poisson density of chord centers on linear transects, km^{-1}
F_C	derived	distribution function of C
f_C	derived	probability density function of C

4.1 Estimating Storm Velocity

The storm velocity here refers to the velocity of the modeled spatial rain field. The storm velocity is estimated from the rainfall data by analyzing relations between observations of the same storm at multiple stations. Several commonly used methods dealing with this problem are reviewed in Niemczynowicz [1987]. The reviewed methods all rely on comparing detailed rainfall intensity processes at the rain gauges. This imposes some requirements that the data used in this study do not satisfy, including “the time resolution must be of the order of one or two minutes,” and “distance between gauges must be of the order of one km” [Niemczynowicz, 1987, p. 137]. Another requirement is on the model. According to Niemczynowicz [1987, p. 138], “most of the known methods ... often fail when more than one rain cell is present over the gauge network at the same time.” Since our model does not require that different locations experience the same set of rain patches, these methods for estimation of storm velocity, based on detailed comparisons of rainfall intensities, are not applicable even with data of higher spatial or temporal resolution.

The key to the method we propose below is that observations of the storm at a pair of stations have a time difference that, under the model, is determined by the location relation of the two stations and the velocity of the storm. This method uses the geometry of the rain gauge network together with starting and stopping times of the storm at each rain gauge.

Suppose a storm is recorded by k stations $\mathbf{P}_i : (x_i, y_i)$, $i = 1, \dots, k$. For each station there is an observed reference time t_i , taken as the average of the beginning and end of the storm at that station. For each of the $k(k-1)/2$ station pairs, the difference of the two observed reference times is related to the storm speed u and direction θ via

$$\begin{aligned} t_i - t_j &= \frac{d_{j,i} \cos(\theta - \alpha_{j,i})}{u} + \delta_{j,i} \\ &= (x_i - x_j) \frac{\cos \theta}{u} + (y_i - y_j) \frac{\sin \theta}{u} + \delta_{j,i}, \end{aligned} \quad (1)$$

where $1 \leq i < j \leq k$, $\alpha_{j,i}$ denotes the direction of the line from \mathbf{P}_j to \mathbf{P}_i , and $d_{j,i}$ is the distance between the two stations (see Figure 3). The error term $\delta_{j,i}$ exists because the reference times observed at different stations are, although conceptually equivalent, subject to random disagreement.

Fig 3

From system (1) we derive the least squares estimates of $(\cos \theta)/u$ and $(\sin \theta)/u$:

$$\begin{aligned} \begin{bmatrix} (\cos \theta)/u \\ (\sin \theta)/u \end{bmatrix} &= \begin{bmatrix} \sum (x_i - x_j)^2 & \sum (x_i - x_j)(y_i - y_j) \\ \sum (x_i - x_j)(y_i - y_j) & \sum (y_i - y_j)^2 \end{bmatrix}^{-1} \\ &\cdot \begin{bmatrix} \sum (t_i - t_j)(x_i - x_j) \\ \sum (t_i - t_j)(y_i - y_j) \end{bmatrix}, \end{aligned} \quad (2)$$

where all the sums are for all the station pairs. Estimates of the storm speed u and direction θ follow. The storm size is then estimated by the product of the estimated storm speed and the longest storm duration across all stations.

The nonlinear relation $u = \left(\sqrt{(\cos \theta/u)^2 + (\sin \theta/u)^2} \right)^{-1}$ causes an under-estimation of u . We have obtained a correction factor of 1.3 through simulations and applied it to the speed estimates.

System (1) requires observations at three or more stations in order for the velocity of a storm to be estimated. Of the 918 storms identified in the illustrative Alabama data, the velocity and size of 535 were estimated, shown in Figures 4–5.

Fig 4–5

4.2 Estimating the Mean Gap Size

Using the estimated storm speed, we convert within-storm dry periods in the time series data to gap lengths in the Boolean field. Each station with no record for a particular storm contributes a gap equivalent to the size of the storm. We similarly account for gaps from the beginning of a storm to the first rainfall clump and from the last rainfall clump to the end of the storm. Gaps in the Boolean field have an exponential distribution [Stoyan et al., 1995, p. 82]. The distribution of gaps derived from the data is compared with the exponential distribution in Figure 6.

Fig 6

Because large gaps are under-represented in the empirical gap size data, we used a trimmed mean of these data to estimate the model mean gap size, μ_D , trimming the data at the 5th and 95th percentiles, which are denoted by d_5 and d_{95} , respectively. The expected value of the trimmed mean is related to μ_D by

$$\mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n d_i \right] = \frac{d_5 \cdot \exp\{-d_5/\mu_D\} - d_{95} \cdot \exp\{-d_{95}/\mu_D\}}{\exp\{-d_5/\mu_D\} - \exp\{-d_{95}/\mu_D\}}. \quad (3)$$

Solving the nonlinear equation above provides an estimate of μ_D . This estimate from the Alabama data is 153 km.

4.3 Estimating the Density and Size Distribution of Rain Patches

The spatial Boolean model undergoes two reductions to yield data at the monitoring network. First, the 2-D Boolean model of patches reduces to a 1-D model of patch chords on a linear transect corresponding to a monitoring location. Then contiguous wet segments on the transect are observed as clumps, whereas dry segments are gaps. (See Section 3.) Each clump is either a single isolated wet chord or the union of overlapping wet chords.

We infer the 1-D transect model of patch chords from the observed clumps and gaps in the historical data, then infer the 2-D Boolean model of patches from the 1-D transect model of patch chords. The relation between the 1-D model and clumps and gaps is somewhat involved, but practical approximations are available.

4.3.1 Estimating the 1-D patch chord model from clumps and gaps

On a linear transect through the 2-D Boolean field, wet patch chords (locations represented by their mid-points) form a 1-D Boolean field [Matheron, 1975, p. 140; Stoyan et al., 1995, p. 81] with density λ_C (number of chords per unit length of transect) and chord length distribution F_C . These are the two parameters for the 1-D patch chord model that we need to estimate.

The chord density in the 1-D model is determined by the 2-D model as [Mecke and Stoyan, 1980; Stoyan et al., 1995, p. 81, 354]

$$\lambda_C = 2\lambda_R\mu_R. \quad (4)$$

(See Table 1 for notation.) This is connected to statistics of gaps, which have an exponential distribution with expected value [Lu and Torquato, 1993; Stoyan et al., 1995, p. 82]

$$\mu_D = \frac{1}{2\lambda_R\mu_R}. \quad (5)$$

Therefore

$$\lambda_C = \frac{1}{\mu_D}, \quad (6)$$

where mean gap size μ_D has been estimated in Section 4.2.

The distribution of clumps, F_W , is a known function of the chord properties λ_C and F_C [Hall, 1988, p. 91; Quintanilla and Torquato, 1996]. Handley [1999] describes a convenient discrete recursive approximation to this relation. Let

$$P(x) = F_W(x + \delta/2) - F_W(x - \delta/2), \quad (7)$$

where δ is a properly chosen small value, and

$$G(x) = \frac{P(x) + \sum_{i=1}^{\lfloor x/\delta \rfloor} G((i-1)\delta)Q(x)}{\sum_{i=1}^{\lfloor x/\delta \rfloor} Q(x)}, \quad (8)$$

where $Q(x) = P(x - i\delta) \prod_{j=1}^{i-1} G((j-1)\delta)$, and $G(0) = 1 - \delta\lambda_C$. The distribution of chord lengths is calculated from G as

$$F_C(x) = 1 - \frac{1 - G(x)}{\delta\lambda_C}. \quad (9)$$

We use the empirical distribution of data-derived clumps as an estimate of F_W .

4.3.2 Estimating the 2-D patch model from the 1-D patch chord model

The distributions of patch radii and chord lengths are related stereologically [Mecke and Stoyan, 1980; Stoyan et al., 1995, p. 354] by

$$F_R(x) = 1 - \frac{4\mu_R}{\pi} \int_{2x}^{\infty} \frac{f_C(t)}{\sqrt{t^2 - 4x^2}} dt, \quad x > 0, \quad (10)$$

where f_C is the probability density function of chord lengths.

The mean patch radius μ_R can be derived from the expression above. If r_0 is a specified minimum patch radius, then $F_R(r_0) = 0$ by definition. Substituting $F_R(r_0)$ and r_0 into relation (10) yields the mean patch radius expressed in terms of chord lengths:

$$\mu_R = \frac{\pi}{4} \left(\int_{2r_0}^{\infty} \frac{f_C(t)}{\sqrt{t^2 - 4r_0^2}} dt \right)^{-1}. \quad (11)$$

With μ_R obtained by (11), the patch radius distribution $F_R(x)$ can be obtained via (10). (The distribution of patch radii estimated from the Alabama data is shown as part of Figure 13, which will be introduced and discussed in Section 5.1.) According to the estimated distribution, the median patch radius is 3.3 km. The patch density λ_R is estimated via relation (5). Its estimate from the Alabama data is 0.00043 rain patches per square kilometer.

Setting r_0 to a positive value prevents instability in the estimates of F_R and λ_R . The value of 0.25 km was used in this study.

4.4 Estimating the Distribution of Patch Rainfall Intensities

Historical data contain wet clumps, each with an observed average rainfall intensity and an estimated length. Each clump is implicitly composed of one or more chords, each corresponding to a rain patch. We use a simulation method to estimate the spatially-averaged patch rainfall intensity distribution from these rain clump statistics: the estimated Boolean model is used to generate rain patch fields and their derived clumps; simulated clumps are assigned average intensities, based on their length, from the observed joint distribution of clump lengths and intensities.

Suppose that a simulated clump is assigned an average intensity q according to its length w using the empirical joint distribution. The simulated clump consists of n chords, say, of length c_i and average rainfall intensity q_i , $i = 1, \dots, n$. Then

$$\sum_{i=1}^n c_i q_i = wq. \quad (12)$$

We assign a common average intensity x to all the n patches corresponding to the chords contributing to this clump. Let r_i , $i = 1, \dots, n$, be the radii of the patches. Because a modeled rain patch has a conic intensity profile, the average intensity on the i th chord is

$$q_i = \frac{3x}{2} \left(1 - \frac{4r_i^2 - c_i^2}{4r_i c_i} \log \frac{2r_i + c_i}{2r_i - c_i} \right). \quad (13)$$

Combining relations (13) and (12), we get the common average rainfall intensity for the n patches:

$$x = \frac{2wq}{3} \left(\sum_{i=1}^n \left(c_i - \frac{4r_i^2 - c_i^2}{4r_i} \log \frac{2r_i + c_i}{2r_i - c_i} \right) \right)^{-1}. \quad (14)$$

This procedure generates a sample of n (identical) patch rainfall intensities with their corresponding radii. By repeating this procedure for multiple clumps in a simulated rain patch field, we generate a joint distribution for rain patch radius and intensity.

This method does not assume an *a priori* relation between the size and rainfall intensity of a patch. In the simulated sample of patch (**radius**, **intensity**) pairs, the two properties of the patch may be correlated or uncorrelated. The derived patch (**radius**, **intensity**) distribution for the Alabama data is shown in Figure 7, which exhibits the quartiles of patch rainfall intensities as the size of the patch varies. This joint distribution corresponds to 10 000 simulated patches based on parameter estimates derived from the Alabama data.

Fig 7

4.5 Summary of Model Estimation

In summary, we estimate the regional rainfall model in three steps.

First, storm velocity and size are estimated using rainfall time series at multiple stations. The estimated storm speed converts observed wet and dry durations to clump and gap lengths in the Boolean rain field. Mean gap size is estimated with the derived gap lengths.

Second, the Poisson density of patch centers and the distribution of patch radii are estimated using stereological relations between statistics of rainfall clumps and gaps and those of the underlying rain patches.

Third, the distribution of patch average rainfall intensity across patches is estimated using clump rainfall intensities in the data together with Boolean model simulations using parameters estimated in the previous step.

This procedure non-parametrically estimates the distributions of patch sizes and patch rainfall intensities. Relations in the first two steps are summarized in Figure 8.

Fig 8

The components of this abstracted model interact to produce complex spatiotemporal rainfall patterns on the ground. For example, the Poisson density and the size distribution of rain patches together determine the degree of patch overlapping. By allowing elementary rain patches to overlap, the model presents surface rain areas of complex shapes. The movement of this rain field then translates into dry and wet durations in the time series observations at the monitoring network.

In time-aggregated data such as the hourly records used in the Alabama illustration, the time-averaged rainfall intensities in wet periods will under-represent extreme rainfall intensities. Therefore, patch rainfall intensity parameters estimated directly from wet periods (see

Section 4.4) may not appropriately capture peak intensities. As a solution, we reconstructed continuous-time rainfall series from the aggregated historical data using the procedure described in Zhang [2004], and used the reconstructed data in estimating the mean gap size, Poisson patch density, patch size distribution, and patch rainfall intensity distribution. With data of higher time resolution, this reconstruction will be less consequential. We used the original data to estimate storm speeds, because this continuous reconstruction does not affect the storm speed estimation.

5 Assessment of the Model

Simulations provide a platform for investigation of rainfall characteristics and for model validation. We conducted 100 simulations, each consisting of 535 storms with velocity and size taken from each of the 535 storms in the Alabama data that had their velocities and sizes estimated. The procedure for simulating a storm with the estimated model is outlined in Figure 9. The patch centers of each simulated storm are distributed in a rectangular field of desired length (i.e., storm size) and sufficient width. Since storm frequency is not a subject of this study, we inserted long dry periods between simulated storms. Figure 10 shows the spatial rain field of one storm thus simulated.

Fig 9

Fig 10

Using the Boolean model with parameters estimated from the Alabama data, the simulated storms result in continuous rainfall time series at the eight Alabama rain stations. By imitating the aggregating mechanism of the tipping-bucket rain gauge, we converted the simulated continuous data to aggregated hourly data comparable to real historical data. Simulated time-aggregated hourly series at the eight monitoring sites are referred to as the “historical-like” data.

Historical-like data from repeated simulations are used to assess the statistical variability of the model estimation procedure described above. We also used historical-like data to compare model output rainfall with the historical data for selected statistics that were not directly used in model fitting.

5.1 Statistical Properties of Parameter Estimators

The 100 runs of simulated historical-like rainfall data were used to repeatedly estimate storm speed and size as discussed in Section 4. The simulated continuous data records at the eight monitoring sites were used to estimate the remaining components of the model, in lieu of the continuous reconstruction device used for the original parameter estimation. Thus 100 sets of models were re-estimated from simulated data. Examinations of these re-estimated models reveal how much the estimates would fluctuate based on data of 535 storms at these 8 monitoring stations, assuming the model to be correct.

Storms were identified in the simulated data, using the 12-hour storm separator and ignoring knowledge of the storm simulations. Most of the simulated storms were identified as individual storms by this criterion. Each of these storms has an assigned velocity for the simulation and a re-estimated velocity from the simulated data. The assigned and re-estimated speeds and directions of each storm are compared in Figures 11 and 12. One can see that the bias of the estimated storm direction can be as big as 180 degrees. This may happen when, for example, the storm is recorded by few stations and the temporal sequence

Fig 11, 12

of the records for this storm at the stations happen to contradict with the actual movement of the storm. This scenario is more likely if the storm's movement is perpendicular to the configuration of its covered stations. Furthermore, the tipping bucket aggregation also plays a role in distorting the data. It should be noted that once the direction estimate is seriously biased, the velocity estimate can not be trusted. Estimates of storm-specific velocities exhibit substantial variability at the individual-storm level because of the limited data that are available for estimation from single storms.

Compared with their original values in the model, the patch density λ is over-estimated by about 59% whereas the mean gap size is under-estimated by about 14%. These two quantities are connected through relation (5). Their estimation biases are caused mainly by the fact that low-coverage storms that deposit rain at fewer than 3 of the 8 monitoring sites are not used in model estimation. We are investigating ways to correct these estimation biases.

Frequency distributions of patch radius estimates, based on the repeated simulations, are shown in Figure 13 and compared to the patch size distribution calculated from the historical data. The box plot for each patch size interval indicates the range of statistical variation across the 100 simulations. Similarly, frequency distributions of patch rainfall intensities, based on simulations, are shown in Figure 14 and compared to their counterparts based on the Alabama data. Both patch size and patch rainfall intensity in the re-estimated models have some negative bias compared to the original model. The statistical variation of the re-estimates is small.

Fig 13

Fig 14

5.2 Distribution and Extremes of Rainfall Intensities

The fitted Boolean patch model, which simulates spatial-temporal storm fields, thus generates simulated rainfall time series at the monitoring sites and thus indirectly generates rainfall intensity data. The degree of statistical agreement between measured rainfall intensities in the historical data and those indirectly obtained from Boolean patch model simulations is an important indicator of model performance.

For model-simulated historical-like data, we obtained the distribution of rainfall intensities in individual wet hours and compared it with its counterpart in the historical data, as shown in the quantile-quantile plot of Figure 15. This distribution in its entire range is reasonably reproduced. In particular, the reproduction of the long tail of extreme hourly intensities is important for examining rarity of extreme rainfall events using simulations. Similar agreement was seen in the comparison of distributions of wet spell rainfall intensities, i.e., intensities in consecutive wet hours.

Fig 15

In the historical and simulated aggregated data, we found the maximum single-hour rainfall intensity among all eight stations in n storms, where n is 10, 50, 100, or 200. The n storms were sampled at random; this sampling was repeated 100 times. Figure 16 is a summary of these extreme values. Each box on this plot summarizes the 100 rainfall intensity maxima corresponding to 100 random sampling of the group of n storms.

Fig 16

5.3 Temporal Persistence of Rain at a Fixed Location

The rainfall process observed at any location demonstrates strong continuity along time. We consider two measures of temporal persistence to check the model’s performance in this regard—a lagged conditional probability of rain and a lagged covariance between rainfall intensities.

The time-lagged conditional probability of rain, given rain at the same site at the earlier time separated by time lag x , is defined as

$$f(x) = \Pr[q(t+x) > 0 \mid q(t) > 0], \quad (15)$$

where $q(t)$ is rainfall intensity at time t . Comparison of this persistence function between historical and historical-like model-simulated data is shown in Figure 17. The comparison demonstrates reasonable agreement. Fig 17

We could have worked on the correlation of rain status at two moments separated by time lag x instead of the “persistence of rain” defined above. However, to determine the correlation we need the marginal probability of rain. Because we are only interested in times within a raining period (the “rain/dry” status in inter-storm periods would have a strong auto-correlation), this raises the question of how to determine the time boundaries of a rain event.

The type of lagged covariance used here is the expected product of individual wet-hour rainfall intensities at a monitoring site, lagged by x hours, conditional on rain at the earlier time:

$$g(x) = E[q(t)q(t+x) \mid q(t) > 0]. \quad (16)$$

The plot of this lagged covariance for the historical Alabama data and model-simulated historical-like data is shown in Figure 18. It is seen that for short time lags the temporal persistence of rainfall intensity in the model-simulated data is stronger than in the historical data. A main reason for this disagreement is likely to be the model assumption that rain is deposited by a time-invariant structure of rain patches moving over the monitoring sites. Possible extensions of the patch model with limited rain patch life-time might bring the model temporal persistence more into line with what was observed. (See [Cox and Isham \[1988\]](#) for discussions about the relative influences of the speed of storm movement and the speed of rain patch death on spatial and temporal covariances of rainfall intensity.) In both Figures 17 and 18, the comparison at large time lags should be taken with caution, because the between-storm dry durations were not generated based on any model. Fig 18

5.4 Spatial Coverage and Correlation of Rain at a Fixed Time

In this section, we discuss two observable spatial properties of the rain field that can be calculated analytically from the Boolean model. The model-calculated spatial properties are compared with their empirical values estimated from the historical Alabama data.

The first property concerns the spatial coverage of rain patches. Let p denote the expected fraction of the ground covered by (possibly overlapping) rain patches at a moment in time during a storm. Then [[Stoyan et al., 1995](#), p. 67]

$$p = 1 - \exp(-\lambda_R \pi E[R^2]), \quad (17)$$

where $\pi E[R^2]$ is the expected area of a circular rain patch. This spatial coverage calculated according to the Boolean model estimated from the Alabama data is 22%.

Since this spatial coverage is equal to the linear coverage on random transects [Badddeley and Vedel Jensen, 2005, p. 11], we can estimate p from historical time series data at fixed monitoring sites, which are the equivalent of linear transects of the spatial rain fields. In estimating p from the historical data, we used the durations of dry and wet periods in each storm at each station as a substitute for clump and gaps sizes, therefore partly avoided relying on the estimated storm speeds. The storm- and station-wise empirical values demonstrate considerable variation, with a median value of 24%. Part of the observed variability derives from the small number of recorded dry and wet periods during each storm.

The second spatial property concerns contemporaneous rain-status correlations between two locations in the rain field. Let $C_s(r)$ be the probability that two locations separated by distance r are simultaneously raining during a storm. This probability can be calculated from the Boolean patch model as a function of the spatial separation of the two locations [Stoyan et al., 1995, p. 83]:

$$C_s(r) = 2p - 1 + (1 - p)^2 \cdot \exp\left(\lambda_R \int_{r/2}^{\infty} (2x^2 \arccos \frac{r}{2x} - \frac{r}{2} \sqrt{4x^2 - r^2}) dF_R(x)\right),$$

$$r \geq 0,$$

where $F_R(x)$ is the distribution function of patch radii, λ_R is the Poisson patch density, and p is the spatial coverage of rain patches. At the two extremes, $C_s(0) = p$ and $C_s(\infty) = p^2$. Using the fact that the mean and variance of the point rain status (0 or 1) are p and $p(1 - p)$, respectively, we easily derive the rain-status correlation from $C_s(r)$:

$$C(r) = \frac{C_s(r) - p^2}{p(1 - p)}. \quad (18)$$

We obtained empirical values of $C(r)$ for 28 spatial lags between the 8 rain stations in the Alabama data. The fraction of common rainy time for each pair of stations during each storm was calculated. Part of the results of storm velocity estimation was used in this calculation to determine the period in which both stations of the pair were within the storm field. The empirical, storm-wise value of p was used in estimating empirical values of $C(r)$. The empirical values, averaged across storms for each spatial lag, and model-theoretic values of $C(r)$ are compared in Figure 19. The comparison is limited by the small number of empirical values; a few pairs of stations that are closer to each other would provide more revealing comparisons in the steeper section of the curve. The figure shows generally good agreement between the model and the data. However, the empirical values are affected by the storm velocity estimation, and Section 5.1 has only established that storm velocity estimation works if the storms satisfy the model assumptions.

Fig 19

5.5 Effect of Temporal Data Aggregation

It is useful for the storm rainfall model to be able to capture statistical properties of rainfall at different time aggregation levels such as multi-minute or multi-hour levels. Since storm

frequency is not part of this study, we do not examine aggregations for longer periods such as days or months. As an exercise, we aggregated the hourly historical data and simulated historical-like data to the 4-hour aggregation level in order to compare selected properties of model simulations against the data.

Figure 20 is a quantile-quantile plot comparing the distributions of wet spell (i.e., consecutive wet aggregation units) lengths in historical and simulated historical-like data. Due to the limited number of quantile values, which are all integer multiples of four hours, multiple data symbols on this plot could overlap and appear as only one. Subsequently, a genuine single symbol and an apparent single symbol would carry the same weight to the viewer and convey misleading messages. To mitigate this problem, we added small random noises to the spell lengths before retrieving their quantiles. One can see that the simulated wet spells are slightly more concentrated on intermediate values. Figure 21 shows the comparison of rain status time-lagged persistence defined in (15). There is good agreement between the results from the model-simulated historical-like data and the historical Alabama data up to time lags comparable to storm durations.

Fig 20

Fig 21

6 Concluding Remarks

We have introduced a regional stochastic spatiotemporal model for surface rain patterns during a storm. The storm model is built from a moving Boolean field of rain-generating patches, each of a random size, location, and rainfall intensity. A storm moves at a constant speed in a fixed direction. Storms differ because of differences in the realized randomness of their size, motion, component patches, and patch rainfall intensities.

We emphasize simplicity of the model structure and intuitive connections between the modeled spatial process and temporal observations at individual rain gauges. Compared with previously published models, the proposed model does not explicitly incorporate clustering mechanisms or time-dependent rain patch characteristics. A two-layer clustered point process has been used elsewhere to reproduce rainfall characteristics at a range of time scales and aggregation levels [Rodriguez-Iturbe et al., 1987]. Our limited tests of the model described here support scalability of the model (Section 5.5), at least in the context in which it was applied.

Model assessment has indicated that a potential generalization of the model would allow for time-dependent evolution of rain patches. The first step in this direction could allow a rain patch to have time-invariant properties during a limited, random patch life. This is similar to Cox and Isham [1988]. Another extension involves a richer parameterization of the intensity profile for rain patches. As a generalization of the conic shape adopted here, we may take the intensity to be linearly decreasing from the patch center, reaching at the patch border a fixed fraction of the peak value. Such a parameterization will include conic patches and cylindrical patches as special cases, and will require only small changes to relations (13) and (14). Since the intensity profile enters the model estimation only through the numerical construction of patch rainfall intensities (Section 4.4), it is possible to incorporate more complex profiles [Capsoni et al., 1987; Kawamura et al., 1997; Luyckx et al., 1998; Willems, 2001; von Hardenberg et al., 2003].

A main part of this work is model estimation, which has been integral to the model design. Our bottom line is to use conventional historical rain gauge data to estimate model parameters. The estimation procedure we proposed consists of three components: storm velocity and size, geometric aspects of the storm rain field (patch size distribution and patch density), and rainfall intensities (distribution of patch rainfall intensity). The latter two are encompassed in the Boolean model framework. The procedure for estimating geometric aspects of the model is tightly integrated mathematically. Variations or extensions of the basic Boolean framework will affect strategies for estimation of model parameters.

We use calculated velocities for individual storms based on the available rain gauge network data. The method involves certain approximations that may be reasonable only where the spatial extent of the gauge network does not exceed the typical storm size. The bias in the storm speed estimation, caused mainly by a nonlinear expression used in the least square estimator, needs further investigations. However, estimation of patch characteristics in the Boolean model is separated from the estimation of storm size and velocity.

We made minimal assumptions regarding distributional shapes for random components of the model that are not directly observable, allowing the historical data to nonparametrically contain these distributions, such as the joint patch size and patch intensity distribution. On the other hand, the basic Boolean model does imply some distributional properties that are calculable, such as those considered by [Cox and Isham \[1988\]](#), for example, the exponential distribution of gaps.

The model estimation strategy contains some necessary or convenient heuristics such as the minimum patch radius, r_0 , which is introduced for numerical stability. However, we did see encouraging results in the model validation probes described in Section 5. Simulations further indicate that the model fitting strategy appears reliable in the context of the underlying Boolean patch model for regional historical data sets of the size and scope of the illustrative Alabama regional monitoring network. There is reasonable agreement between simulated and historical data on rainfall intensity distributions and extremes. The degree of temporal persistence in single-site rainfall observations is reproduced and, to a lesser extent, so is the spatial persistence at a fixed point in time.

Our storm event model needs to be combined with a storm frequency model in order to simulate long-term rainfall series. In such simulations seasonal variations of rain storm characteristics need to be considered. We may incorporate seasonality by fitting the model separately for different seasons and smoothly varying the model parameters in an annual cycle [[Stern and Coe, 1984](#)]. In a related treatment, the model parameters can be made specific to storm types.

This model is relevant for applications because of its mathematically tractable structure and the ease with which continuous spatiotemporal rainfall scenarios can be simulated. Simulated scenarios may be used to drive rainfall-sensitive process simulations in studies of environmental problems, ecological dynamics and hydrology. In other applications, simulations allow quantitative inquiries into space-time statistical properties of rainfall that are not readily obtainable from relatively short records, such as temporal and spatial structures, and frequency of extreme events. Of fundamental relevance to both types of applications is the ability of the stochastic model of the kind considered here to generate multiple realistic rainfall scenarios that imitate the statistics of historical network monitoring data. In its

current, quite basic form our regional rain storm modeling approach is likely to be most useful in situations comparable to the illustration that we used, i.e., where

1. The region of interest is small relative to the spatial extent of a typical storm, and statistical rainfall characteristics are approximately homogeneous across the region.
2. Storms are relatively stable structures seen on the regional scale of the monitoring network, moving across the region without large change in direction or speed.
3. The time resolution of the historical data used for model estimation is not much coarser than the typical duration of continuous rain periods.

Acknowledgments. The authors thank David Freyberg, Mark Jacobsin, and Keith Loague for helpful comments. We also thank the Associate Editor and the referees for their many comments that helped us improve the manuscript. This paper was finished when Zhang was on the staff of the Center for Integrating Statistical and Environmental Science at the University of Chicago. Although the research described in this article has been funded wholly or in part by the United States Environmental Protection Agency through STAR Cooperative Agreement #R-82940201-0 to the University of Chicago, it has not been subjected to the Agency’s required peer and policy review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

References

- A. Baddeley and E. Vedel Jensen. *Stereology for Statisticians*. Monographs on Statistics and Applied Probability 103. Chapman & Hall/CRC, 2005.
- C. Capsoni, F. Fedi, C. Magistroni, and Others. Data and theory for a new model of the horizontal structure of rain cells for propagation applications. *Radio Science*, 22(3):395–404, May-June 1987.
- R. E. Chandler and H. S. Wheeler. Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland. *Water Resour. Res.*, 38(10):1192, 2002. doi: 10.1029/2001WR000906.
- E. H. Chin. Modeling daily precipitation occurrence process with Markov chain. *Water Resour. Res.*, 13(6):949–956, 1977.
- P. S. P. Cowpertwait. A generalized spatial-temporal model of rainfall based on a clustered point process. *Proc. R. Soc. Lond. A*, 450(1929):163–175, Oct. 1995.
- D. R. Cox and V. Isham. A simple spatial-temporal model of rainfall. *Proc. R. Soc. Lond. A*, 415(1849):317–328, Feb. 1988.
- L. Ferraris, S. Gabellani, N. Rebora, and A. Provenzale. A comparison of stochastic models for spatial rainfall downscaling. *Water Resour. Res.*, 39(12):1368–, Dec. 2003.
- J. Goldhirsh. Rain cell size statistics as a function of rain rate for attenuation modeling. *IEEE T. Antenn. Propag.*, AP-31(5):799–801, Sept. 1983.
- J. Goldhirsh. Rain cell size statistics derived from radar observations at Wallops Island, Virginia. *IEEE T. Geosci. Remote.*, 24(6):947–954, 1986.

- J. R. Green. A model for rainfall occurrence. *J. R. Stat. Soc., B*, 26:345–353, 1964.
- P. Hall. *Introduction to The Theory of Coverage Processes*. Wiley, New York, 1988.
- J. C. Handley. Discrete approximation of the linear boolean model of heterogeneous materials. *Physical Review E.*, 60(5):6150–6152, Nov. 1999.
- R. W. Katz. Precipitation as a chain-dependent process. *J. Appl. Meteorol.*, 16(7):671–676, July 1977.
- M. L. Kavvas and J. W. Delleur. A stochastic cluster model of daily rainfall sequences. *Water Resour. Res.*, 17(4):1151–1160, 1981.
- A. Kawamura, K. Jinno, R. Berndtsson, and T. Furukawa. Real-time tracking of convective rainfall properties using a two-dimensional advection-diffusion model. *J. Hydrol.*, 203(1-4): 109–118, December 1997.
- T. G. Konrad. Statistical models of summer rainshowers derived from fine-scale radar observations. *J. Appl. Meteorol.*, 17:171–188, Feb. 1978.
- B. L. Lu and S. Torquato. Chord-length and free-path distribution-functions for many-body systems. *Journal of Chemical Physics*, 98(8):6472–6482, Apr. 1993.
- G. Luyckx, P. Willems, and J. Berlamont. Influence of the spatial variability of rainfall on sewer system design. In H. Wheater and C. Kirby, editors, *Hydrology in A Changing Environment*, volume 3, pages 339–349. Wiley, Chichester, 1998.
- G. Matheron. *Random Sets and Integral Geometry*. John Wiley & Sons, New York, 1975.
- J. Mecke and D. Stoyan. Stereological problems for spherical particles. *Math. Nachr.*, 96: 311–317, 1980.
- I. Molchanov. *Statistics of The Boolean Model for Practitioners and Mathematicians*. John Wiley & Sons, New York, 1997.
- J. Niemczynowicz. Storm tracking using rain gauge data. *J. Hydrol.*, 93:135–152, 1987.
- P. Northrop. A clustered spatial-temporal model of rainfall. *Proc. Roy. Soc. Lond. A.*, 454: 1875–1888, 1998.
- C. Onof, R. E. Chandler, A. Kakou, P. Northrop, H. S. Wheater, and V. Isham. Rainfall modelling using Poisson-cluster processes: A review of developments. *Stoch. Environ. Res. Risk Assess.*, 14:384–411, 2000.
- J. Quintanilla and S. Torquato. Lineal measures of clustering in overlapping particle systems. *Physical Review E.*, 54(4):4027–4036, Oct. 1996.
- C. W. Richardson. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour. Res.*, 17(1):182–190, Feb. 1981.
- I. Rodriguez-Iturbe, D. R. Cox, F. R. S., and V. Isham. Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond. A*, 410:269–288, 1987.
- J. Roldan and D. A. Woolhiser. Stochastic daily precipitation models 1. a comparison of occurrence processes. *Water Resour. Res.*, 18(5):1451–1459, Oct. 1982.
- J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.
- J. A. Smith and A. F. Karr. A point process model of summer season rainfall occurrences. *Water Resour. Res.*, 19(1):95–103, 1983.

- R. L. Smith. Spatial modelling of rainfall data. In V. Barnett and K. F. Turkman, editors, *Statistics for The Environment 2: Water Related Issues*, pages 19–40. John Wiley & Son Ltd., 1994.
- R. D. Stern and R. Coe. A model fitting analysis of daily rainfall data. *J. R. Stat. Soc., A*, 147(1):1–34, 1984.
- D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications*. John Wiley & Sons, Chichester, England, 2nd edition, 1995.
- J. von Hardenberg, L. Ferraris, and A. Provenzale. The shape of convective rain cells. *Geophys. Res. Lett.*, 30(24):2280–2283, Dec. 2003.
- E. D. Waymire, V. K. Gupta, and I. Rodriguez-Iturbe. A spectral theory of rainfall intensity at the meso- β scale. *Water Resour. Res.*, 20(10):1453–1465, Oct. 1984.
- H. S. Wheeler, R. E. Chandler, C. J. Onof, E. B. V. S. Isham, C. Yang, D. Lekkas, G. Lourmas, and M.-L. Segond. Spatial-temporal rainfall modelling for flood risk estimation. *Stoch. Environ. Res. Risk Assess.*, 19:403–416, 2005. doi: 10.1007/s00477-005-0011-8.
- R. L. Wilby, T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks. Statistical downscaling of general circulation model output: A comparison of models. *Water Resour. Res.*, 34(11):2995–3008, Nov. 1998.
- D. S. Wilks. Multisite generalization of a daily stochastic precipitation generation model. *J. Hydrol.*, 210:178–191, 1998.
- D. S. Wilks and R. L. Wilby. The weather generation game: A review of stochastic weather models. *Prog. Phys. Geog.*, 23(3):329–357, Sept. 1999.
- P. Willems. A spatial rainfall generator for small spatial scales. *J. Hydrol.*, 252(1-4):126–144, Oct. 2001.
- Z. Zhang. *Continuous Spatial-temporal Stochastic Rainfall Models Based on Aggregated Historical Data*. PhD thesis, Stanford University, 2004.

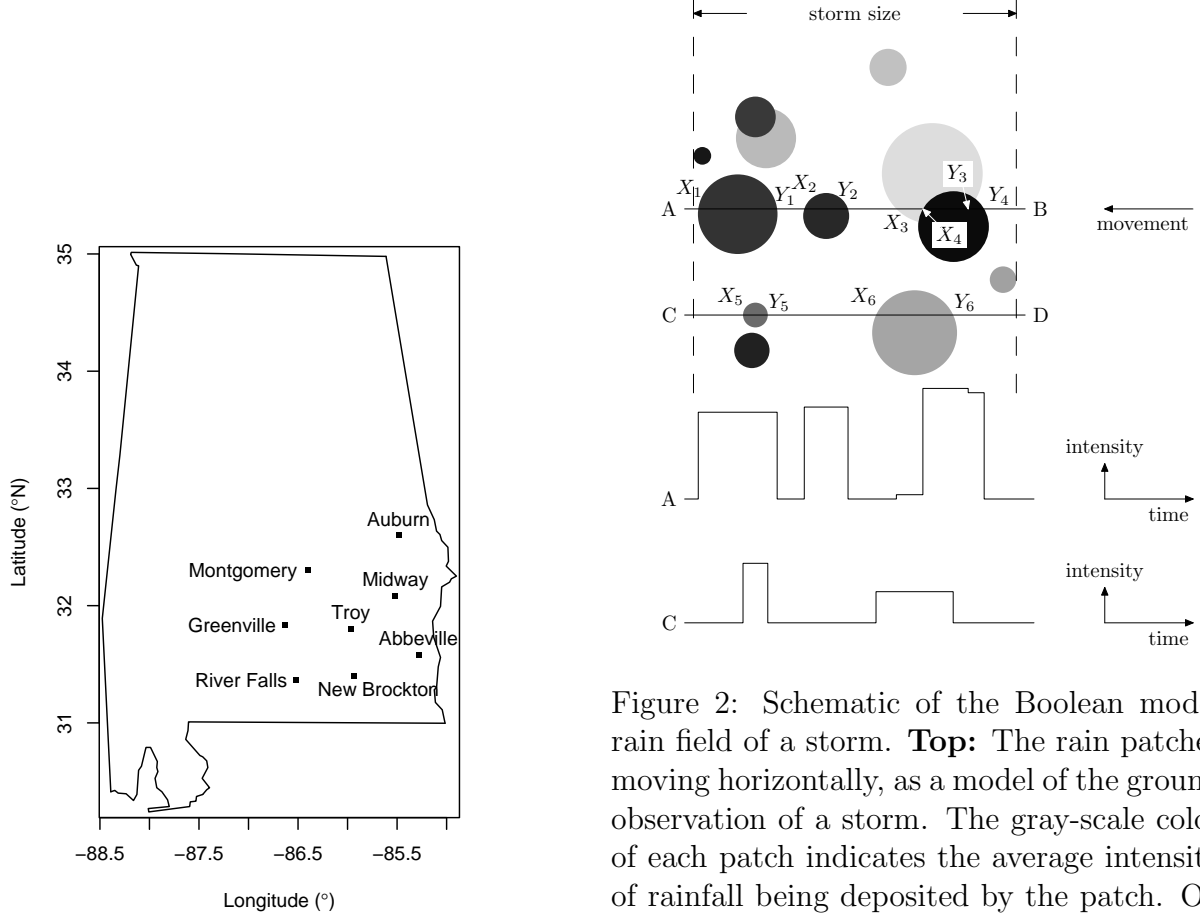


Figure 1: The state of Alabama and locations of the 8 rain gauges for the illustrative data set. The smallest inter-station distance is 44.6 km between New Brockton and Troy. The largest is 168 km between Auburn and River Falls.

Figure 2: Schematic of the Boolean model rain field of a storm. **Top:** The rain patches moving horizontally, as a model of the ground observation of a storm. The gray-scale color of each patch indicates the average intensity of rainfall being deposited by the patch. On the transects AB and CD along the storm's movement, X_1Y_1 , X_2Y_2 , X_3Y_3 , X_4Y_4 , X_5Y_5 , and X_6Y_6 are rain patch chords; X_1Y_1 , X_2Y_2 , X_3Y_4 , X_5Y_5 , and X_6Y_6 are clumps; Y_1X_2 , Y_2X_3 , and Y_5X_6 are gaps. **Middle and bottom:** The rainfall time series observed at locations A and C , assuming constant rainfall intensity within a rain patch. Wet and dry durations observed are proportional to the lengths of clumps and gaps on the linear transects AB and CD .

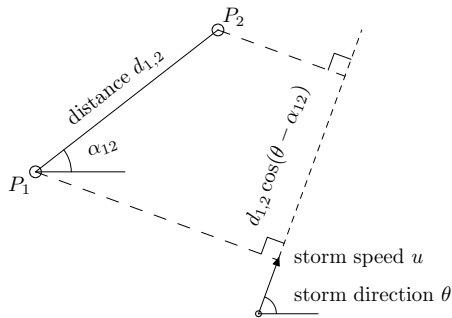


Figure 3: Relations between the locations of two sites, P_1 and P_2 , the movement of the storm, and the travel distance of the storm between the two sites. See equation (1).

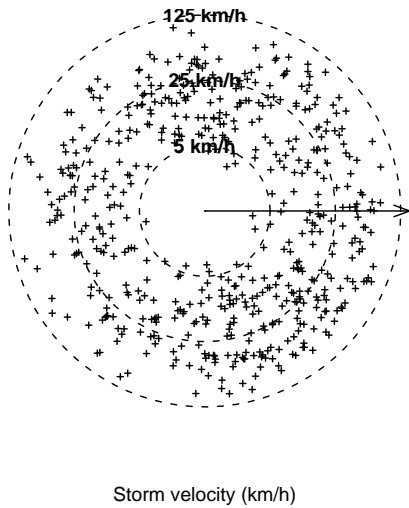


Figure 4: Storm velocities estimated from the Alabama data. Each cross represents a storm moving in the direction of the cross away from the origin, at a speed indicated by its distance to the origin.

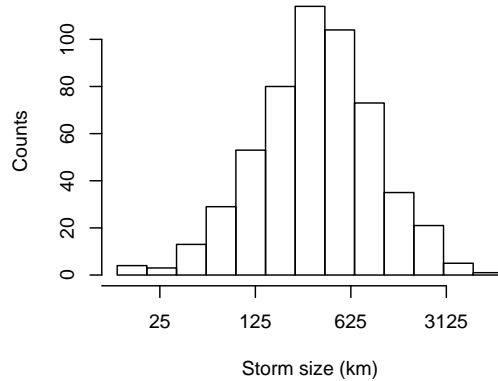


Figure 5: Frequency distribution of storm sizes estimated from the Alabama data. Because the distribution is right-skewed with a very long tail, we use a log scale on the abscissa to help show the full range of the distribution as well as details around where the mass of the distribution concentrates. Similar comments apply to Figures 6, 13, 14, 15, and 20.

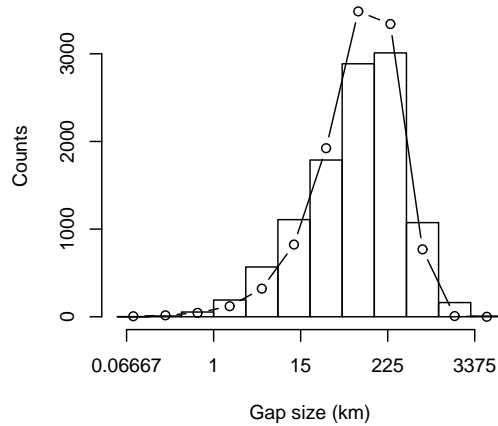


Figure 6: Frequency distribution of gap sizes derived from the Alabama data, compared to the histogram polygon of the exponential distribution whose mean equals the estimated mean gap size, μ_D . (Log scale on abscissa.)

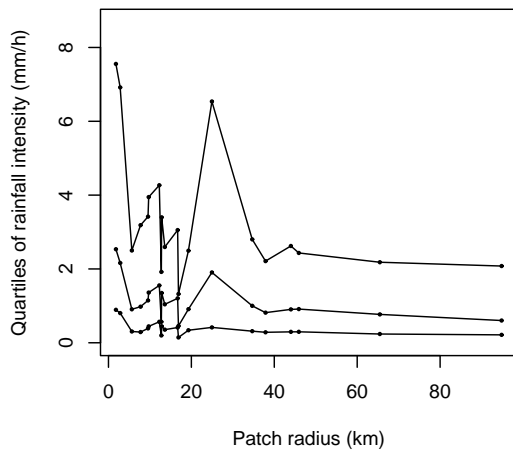


Figure 7: Quartiles of patch rainfall intensity as patch size increases, estimated from the Alabama data. Each patch size interval contains 500 simulated patches, as described in Section 4.4, from which the quartiles of their rainfall intensities are obtained.

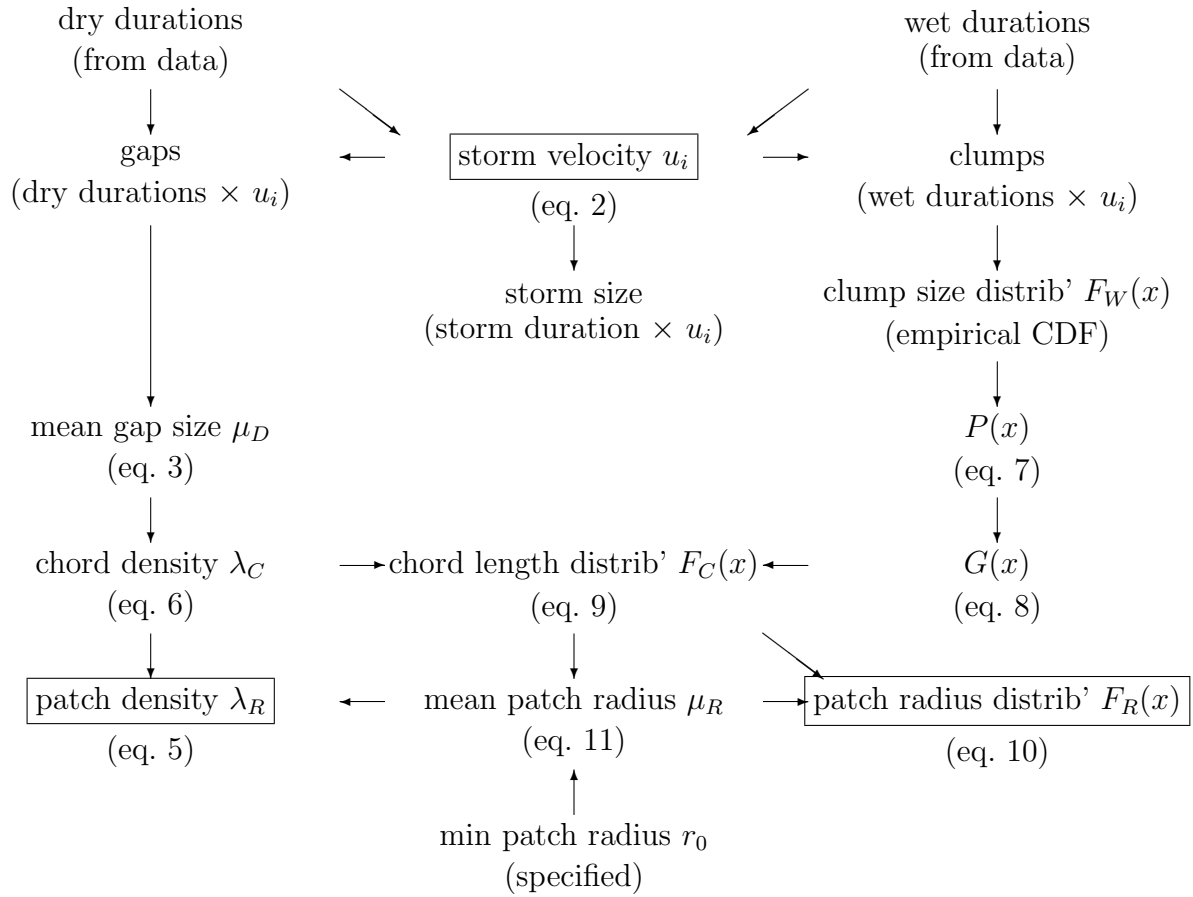


Figure 8: Relation map for the model estimation procedure, prior to the estimation of patch rainfall intensities. Primary model parameters are framed.

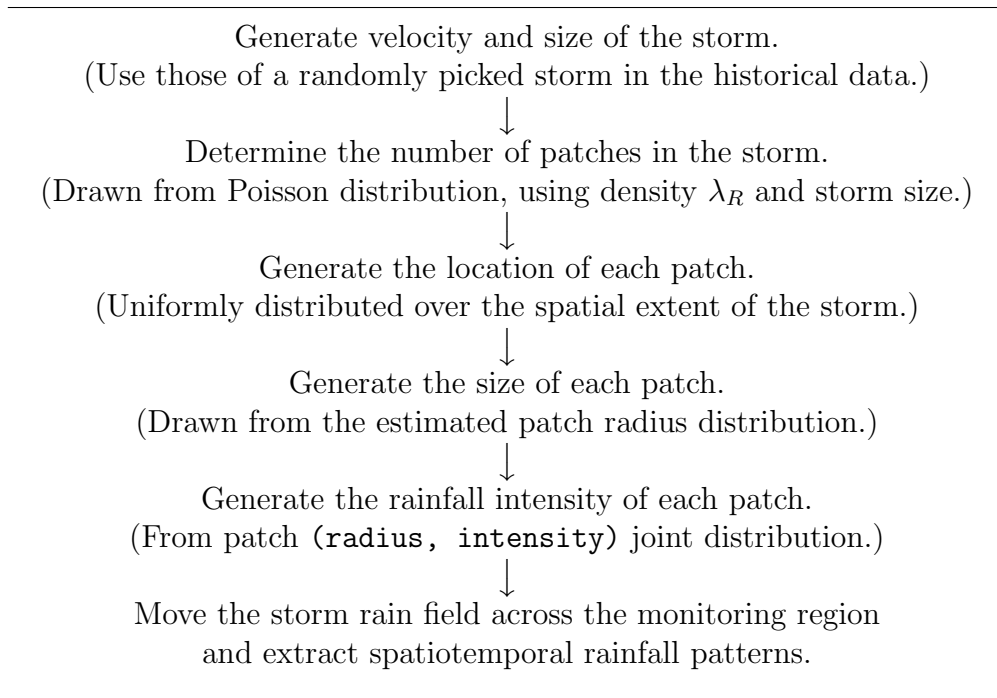


Figure 9: The procedure for simulating a storm with the Boolean storm model.

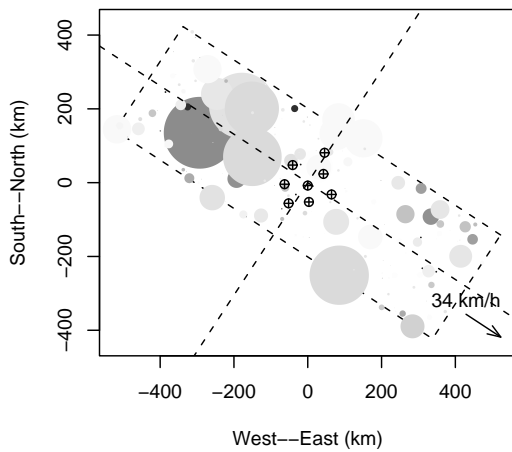


Figure 10: Example of a simulated Boolean storm rain field. The storm is moving as indicated by the arrow in the lower right corner of the plot. The gray-scale color of each patch indicates the average rainfall intensity in the patch. The crossed circles indicate the locations of the 8 monitoring sites for the Alabama data set, from which the model was estimated.

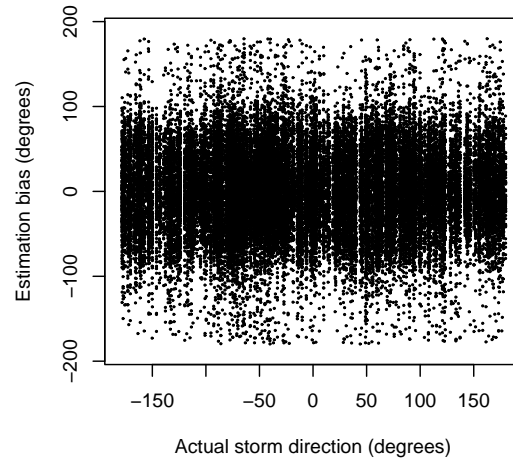


Figure 11: The actual (for simulations) moving directions of 5000 simulated storms and the biases of their estimates from the simulated data (estimated minus actual). Each dot represents one storm.

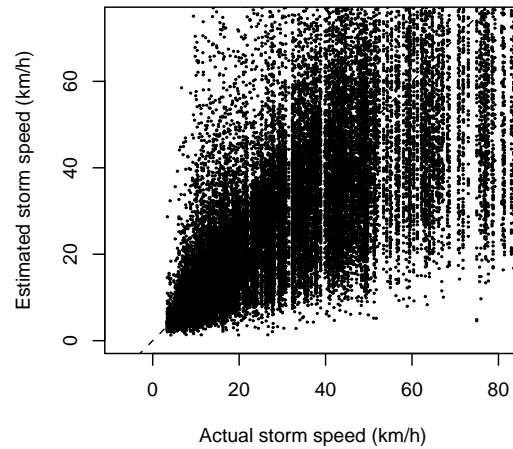


Figure 12: Comparison of the actual (for simulations) and the estimated (from simulated data) speeds of 5000 simulated storms. Each dot represents one storm.

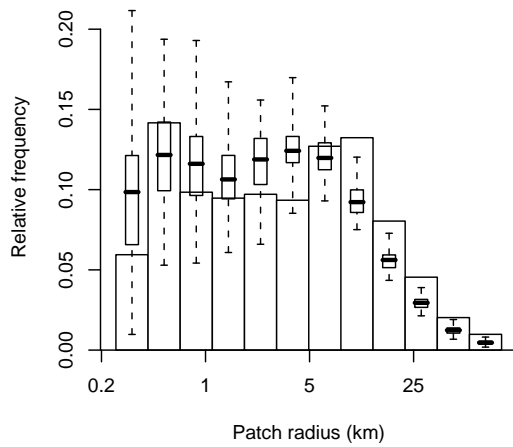


Figure 13: Frequency distribution of patch radii in the model estimated from the Alabama data, overlaid with box plots of the 100 simulation estimates of the same quantities. Each box plot indicates the median, quartiles, and extremes of the distribution (same in subsequent box plots). (Log scale on abscissa.)

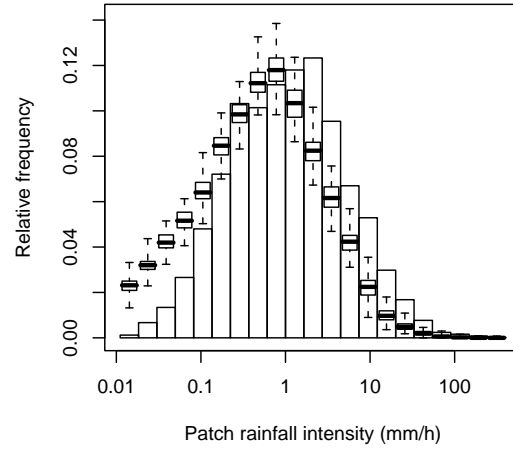


Figure 14: Frequency distribution of patch rainfall intensities in the model estimated from the Alabama data, overlaid with box plots of the 100 simulation estimates of the same quantities. (Log scale on abscissa.)

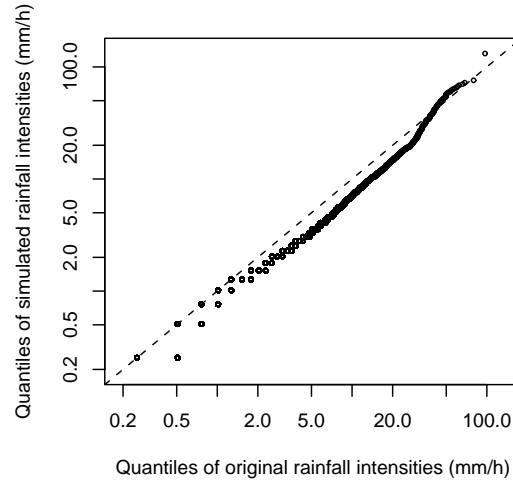


Figure 15: Quantile-quantile plot of individual wet-hour rainfall intensities in the simulated historical-like data versus historical data. (Log scales on both axis.)

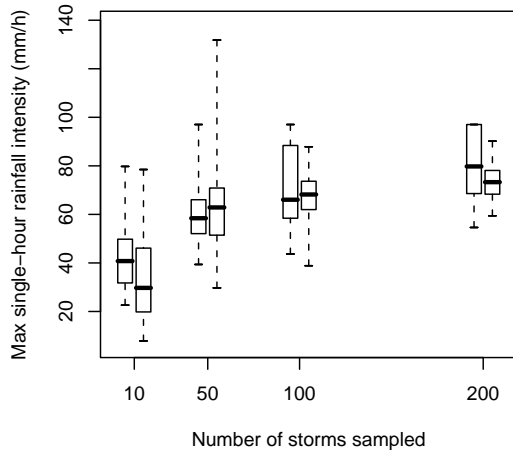


Figure 16: Comparisons of the maximum individual wet-hour rainfall intensities across all the 8 monitoring sites in n ($n = 10, 50, 100, 200$) storms in simulated historical-like data versus historical data. Each box plot summarizes 100 such intensity maxima corresponding to 100 random re-sampling of the n storms. For each n , the box on the left is for historical data, the box on the right is for simulated data.

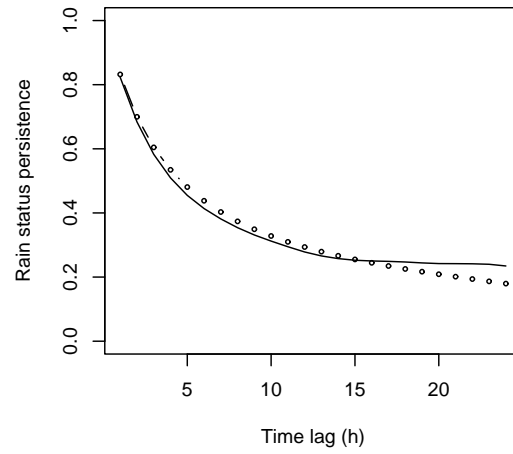


Figure 17: Time-lagged persistence of rain status at a monitoring site for the historical Alabama data (solid curve) and simulated historical-like data (circled curve). See definition in equation (15).

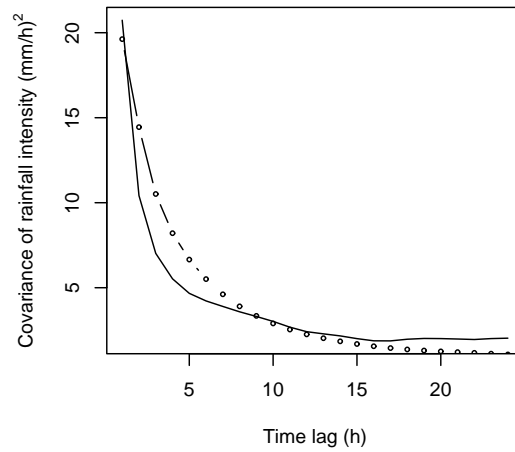


Figure 18: Time-lagged covariance of individual wet-hour rainfall intensities at a monitoring site for the historical Alabama data (solid curve) and simulated historical-like data (circled curve). See definition in equation (16).

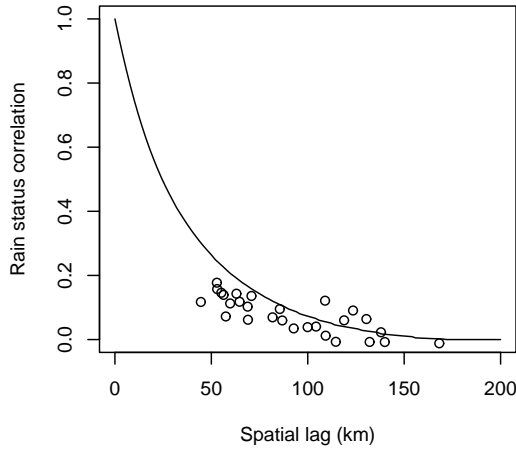


Figure 19: Correlation of contemporaneous rain status as a function of spatial distance. The solid curve is the theoretical value, by equation (18), for the model estimated from the Alabama data. The circles are empirical values estimated from the Alabama data.

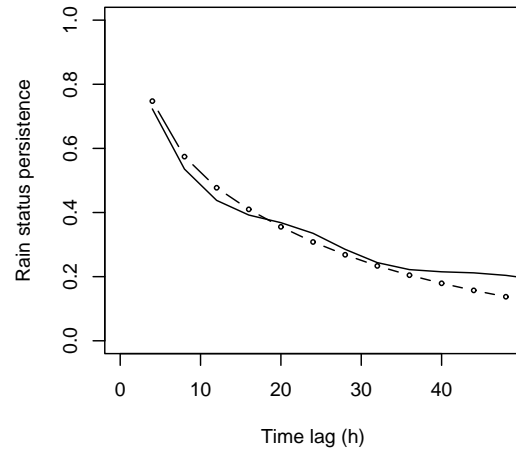


Figure 21: Time-lagged persistence of rain status at a monitoring site for the historical Alabama data (solid curve) and simulated historical-like data (circled curve). Both data sets have been converted to 4-hour aggregated level. See definition in equation (15).

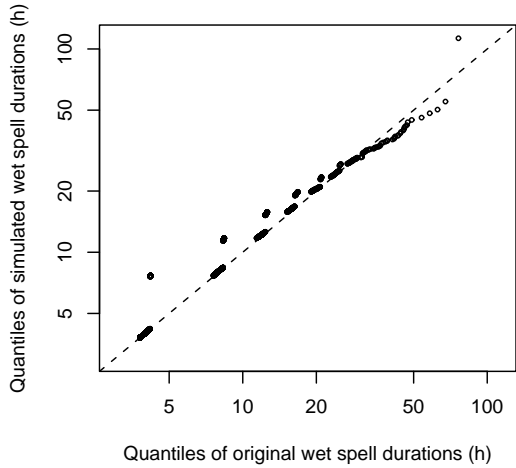


Figure 20: Quantile-quantile plot of the durations of wet periods in historical and simulated historical-like data, both aggregated to the 4-hour aggregation level. (Log scales on both axis.)