A Bayesian Hierarchical Nonoverlapping Random Disc Growth Model

Athanasios C. Micheas and Christopher K. Wikle

A methodology is proposed to efficiently model a random set via a multistage hierarchical Bayesian model. We define a NonOverlapping Random Disk Model (NORDM), which is similar in spirit to the well-known Poisson–Boolean model. This model is formulated in a conditional setting that facilitates Bayesian sampling of important parameters in the model. This framework can accommodate any object, not just those with disk shapes, although the model can be easily extended to include any known compact convex set instead of the disc (e.g., polygons or ellipses). We further propose a growth model that is conceptually simple and allows straightforward estimation of parameters, without the need for tedious calculations of hitting or inclusion probabilities. The model is applied to severe storm cell development as obtained from weather radar.

KEY WORDS: Growth model; Hierarchical Bayesian estimation; Nonoverlapping random disc model; Poisson–Boolean model.

1. INTRODUCTION

Growth, or change (evolution) in objects over time, is ubiquitous in nature. From observing tumor growth to assessing the effectiveness of a medical treatment, studying the propagation and development of storm systems from radar images, to the investigation of an epidemic as it spreads through a populated area, defining and studying growth/evolution models has become an integral component of scientific research over the last century.

Growth can be modeled in a number of different ways. Many approaches have been deterministic, for example, involving differential equations. Such models require a complete understanding of the phenomenon of interest, which is often unattainable. Only in the past few decades has there been a serious attempt to model growth of spatial objects stochastically and, more importantly, take the spatial component into consideration in the model. The use of set models, such as the Boolean model (e.g., see Cressie, 1993), was a major step toward understanding the evolution of an object, but introduced major mathematical and computational difficulties. Moreover, most important applications are not one-dimensional and in two and higher dimensions there is no clear way of specifying a stochastic mechanism to govern a random set. Defining a random variable, vector, or matrix is a well-understood concept that leaves little room for questioning. This is not the case with random sets.

There are two primary approaches in the literature to define a random set $X$. The “hitting function” approach involves calculation of the function $T_K(X) = P(X \cap K \neq \emptyset)$, for compact convex sets $K$, referred to as “test sets.” For different selection of the sets $K$ we can obtain several such probabilities $T_K(X)$, namely the probability that a random set $X$ “hits” the test set $K$. Popular choices for test sets include the disc and the parallelogram in $\mathbb{R}^2$. A second way to describe a probabilistic law for $X$ is through the inclusion probabilities $P(X \subseteq K)$, where again $K$ is a member of an appropriate class of compact convex sets. One of the major difficulties with these approaches is that likelihood cannot be constructed in a simple way to describe the set and thus provide an easy way to perform inference on the parameters involved in the model. More discussion of the two approaches can be found in Cressie (1993), Schneider (1993), Stoyan, Kendall, and Mecke (1995), and Barndorff-Nielsen, Kendell, and van Lieshout (1999).

We will consider random sets that are created as a collection of random disks. We suggest an approach that models a point in the random set and not the entire set directly, where the points observed can be thought of as parts of the random discs. That is, clusters of points are viewed as parts of the realization of a random disc in the model. The centers of these discs are known as “foci” and the disks themselves as “grains” in Boolean model terminology. For more information on the Boolean model and its many applications we refer to Cressie (1993), van den Berg, Meester, and White (1997), Molchanov (1995, 1997), and Khazaee (2003).

Random set models, such as the Poisson–Boolean model, are based on (marked) Markov processes on the foci. Markov models for inhomogeneous point processes have been studied extensively in recent years; see Stoyan and Stoyan (1998), Baddeley, Möller, and Waagepetersen (2000), Jensen and Nielsen (2000), Hahn, Jensen, van Lieshout, and Nielsen (2003), and the references therein.

In our approach, we consider modeling the foci in a multistage hierarchical Bayesian framework that can include, as a special case, the commonly used homogeneous or inhomogeneous Poisson process to model the foci (i.e., the Poisson–Boolean model). Our method is somewhat similar to the processes considered in Baddeley and Möller (1989). However, our approach is different in several aspects, the most important of which is that we consider a hierarchical Bayesian approach, which facilitates inference without the need for tedious calculation of the hitting or inclusion probabilities. Furthermore, modeling a point from the random set instead of the set itself allows straightforward inference of the parameters of the models, via Markov chain Monte Carlo (MCMC), which allows us to make inference about individual random discs in the model, whereas methods utilizing hitting or inclusion probabilities can only obtain overall results, such as the average area of all the grains. The modeling approach we undertake...
treats different subsets of the observed points as the points that would realize the corresponding nonoverlapping discs. Thus, the models created are nonoverlapping random disc models (NORDM), and are obtained by modeling the distribution of the foci and corresponding radii.

A very important application of the proposed methodology is its use for modeling the growth of objects, not just modeling stochastically an observed state of an object. This situation arises in hydrology and meteorology, where we observe radar images of storm systems (cells) and we are interested in obtaining a short-time forecast (nowcast) of cell location, size, and intensity within the next hour or so. Once a forecast is obtained and the true image is observed, verification methods are required to assess the effectiveness of the nowcast. In this context, higher intensities in the image indicate more severe storm systems. Setting an intensity threshold, and removing intensities below this level, allows us to identify several cells in the radar image, which are almost always noncircular. We can still model these cells with our approach once we estimate the foci and radii of the NORDM based on these clusters of points. Modeling evolution in this context equates to modeling the propagation of the storm systems or a situation where two cells are joined in the next time period as well as when one cell splits into multiple subcells.

Some of the recent literature on nowcasting and verification includes Johnson et al. (1993), Ebert and McBride (2000), Lakshmanan, Rabin, and DeBrunner (2003), Seed (2003), Ebert et al. (2004), Fox and Wikle (2005), Pierce et al. (2004, 2005), Xu, Wikle, and Fox (2005), and Micheas, Fox, Lack, and Wikle (2007). In addition, there has been a vast literature over the past 30 years, in the areas of probability, statistics, and mathematical biology that consider growth models. This literature ranges from extremely theoretical to very applied, with the use of random convex sets, the Boolean model, hitting probabilities, and Markov point processes, as described, for example, in Cressie and Hulting (1992), Cressie (1993), Chaplain, Singh, and McClachlan (1999), Liggett (1999), Capasso et al. (2002), Deijfen (2003), Jónsdóttir and Jensen (2005), Newton (2006), Grenander, Srivastava, and Saini (2007), and the references therein. For example, in the latter article, the authors investigate mathematical and statistical models of biological growth to analyze changes in anatomical parts using images collected over time. They obtain a structured model that treats a cumulative growth deformation as a composition of several elementary deformations and demonstrate their mathematical framework using magnetic resonance imaging (MRI) image data of a rat’s brain growth.

The remainder of the article is organized as follows. Section 2 is devoted to the definition of a NORDM in a Bayesian framework. The proposed methodology for modeling growth or evolution is then described and studied in Section 3. We illustrate the model in Section 4, with an application to nowcasting radar precipitation. Section 5 contains some concluding remarks and discussion of future directions.

2. NONOVERLAPPING RANDOM DISCS MODEL

We define a NonOverlapping Random Discs Model (NORDM) as a collection of random discs on the plane as follows. Assume that we observe \( n \) points \( s_1, \ldots, s_n \in \mathcal{O} \subseteq \mathbb{R}^2 \) on the domain of observation \( \mathcal{O} \), treated as the centers of discs with radii \( R_1, \ldots, R_n > 0 \), respectively, according to some distribution \( f(s, R) \), where \( s = (s_1, \ldots, s_n) \), and \( R = (R_1, \ldots, R_n) \). The NORDM is then defined as the random set

\[
X = \bigcup_{i=1}^{n} B(s_i, R_i),
\]

where \( B(s, R) = \{ x \in \mathcal{O} \mid \|x - s\| \leq R \} \) is the circle centered at \( s \) of radius \( R \), with \( B(s_i, R_i) \cap B(s_j, R_j) = \emptyset \), \( i \neq j \). We discuss later the details needed in model (1) and the associated probability distribution in order for the discs to be nonoverlapping. The points \( s_i \) can be considered as the foci (or germs) where growth of the object (or grain) \( B(s_i, R_i) \) occurs. Furthermore, although the definition (and everything that follows later) can be easily extended to more than two dimensions, we focus on the two-dimensional case for clarity of exposition.

Note that if we assume the \( n \) points arise as a realization of a homogeneous or inhomogeneous Poisson process on \( \mathcal{O} \) (with discs allowed to overlap), then \( X \) is known as the Boolean model in the literature. An alternative but equivalent definition to the Boolean model can be based on marked point processes (e.g., Baddeley and Møller, 1989).

Our first concern is to describe stochastically the NORDM, via modeling of random foci and radii, in a Bayesian framework. The resulting model should provide a clear method to help us perform inference on model parameters. For this reason, we consider an approach that models a point in the NORDM and not the entire set directly.

One way to describe the probability distribution of the disc \( B(s, R) \), is with the density

\[
f(x|s, R) = \frac{I_{B(s,R)}(x)}{\pi R^2}, \quad x \in \mathcal{O},
\]

where \( I_{B(s,R)}(x) = 1 \), if \( x \in B(s, R) \), and 0 otherwise. That is, the point \( x \) is sampled uniformly from the disc. Furthermore, for any point \( x \in \mathcal{O} \), in order for \( x \in X \) (i.e., \( x \) is part of a realization of the NORDM) the vector \( x \) must belong to one of the discs \( B(s_i, R_i) \), and thus, \( x \in X \) if and only if \( x \in B(s_i, R_i) \), for one \( i = 1, 2, \ldots, n \). In Figure 1 (left), we display a realization of a NORDM with \( n = 3 \). Herein, we illustrate the methodology for the model (2). Alternatively, we may assume that \( f(x|s, R) \propto g(s|x, R) I_{B(s,R)}(x) \), where \( g \) is an unnormalized bivariate density (e.g., bivariate normal).

2.1 Defining Nonoverlapping Discs

Consider the assumption of nonoverlapping discs in the model. In Figure 1 (right) we display the cases of overlapping and nonoverlapping discs \( B(s_1, R_1) \) and \( B(s_2, R_2) \). Note that to have nonoverlapping discs, the distance \( l = \|s_1 - s_2\| \) between the centers \( s_1 \) and \( s_2 \) must be larger than the sum of the two radii, namely \( \|s_1 - s_2\| > R_1 + R_2 \). Then, to model stochastically the probability distribution of a point \( x \) from the NORDM, we consider the mixture distribution

\[
f_x(x|s, R, p, n) = \sum_{i=1}^{n} \frac{p_i}{\pi R_i^2} I_{B(s_i,R_i)}(x),
\]
where \( \sum_{i=1}^n p_i = 1 \) and \( p_i = P(\text{the } i\text{th disc is selected}) \), \( s = (s_1, \ldots, s_n) \), \( R = (R_1, \ldots, R_n) \), \( p = (p_1, \ldots, p_n) \), and under the conditions
\[
\|s_i - s_j\| > R_i + R_j, \quad i < j, \quad i, j = 1, \ldots, n,
\]
which guarantee nonoverlapping discs. Moreover, define \( I(\|s_i - s_j\| > R_i + R_j) = 1 \) if \( \|s_i - s_j\| > R_i + R_j \), and 0 otherwise.

2.2 Bayesian Formulation

Suppose we observe \( M \) points from the NORDM, \( P = \{x_1, \ldots, x_M\} \), on the plane, where different clusters of points realize the corresponding nonoverlapping discs. In most real life experiments, we do not have information about which subset of points can be used to realize a disc from the model, and a data-augmentation approach (e.g., see Tanner and Wong, 1987; Diebolt and Robert, 1994) can be used with the mixture model (3). However, for the applications we consider here we do not encounter this difficulty, as we illustrate in the following application.

Assume that we observe \( x_{i,j}, j = 1, \ldots, q(i), i = 1, 2, \ldots, n \), points from the \( j\)th grain. Note that \( \sum_{j=1}^{q(i)} = M \). Then, the likelihood based on the sampling distribution (3) can be written as
\[
f(x|s, R, p, n) = \prod_{i=1}^n \frac{p_i^{q(i)}}{\pi(\|s_{|i}\|, R_{|i})} \prod_{j=1}^{q(i)} I_{|s_{|i}, R_{|i}}(x_{i,j}),
\]

since we always know from which disc we observed the \( j\)th point.

The Bayesian paradigm suggests that we model our prior beliefs on \( s = (s_1, \ldots, s_n) \), \( R = (R_1, \ldots, R_n) \), and \( p = (p_1, \ldots, p_n) \), where the conditions (4) must clearly govern the stochastic mechanism of the foci \( s \). If we have observed several states of the NORDM, we know the value of \( n \) in each one, and hence \( n \) is treated as being part of the observed data. However, a time-varying model for \( n \) may be required in a forecasting scenario.

We begin by modeling the conditional distribution of the foci given \( n \) and the radii, namely \( \pi(s|n, R) \). We will consider pairwise interaction processes similar to those defined in Baddeley and Moller (1989) that will account for conditions (4). The distribution is of the form
\[
\pi(s|n, R) \propto \prod_{i<j} g(\|s_i - s_j\|, R_i + R_j),
\]
where the function \( g(u, R) \) governs pairwise interaction between foci. For example, the Strauss-like interaction defined as \( g(u, R) = cl(u < R) + l(u > R) \), where \( 0 \leq c \leq 1 \), yields nonoverlapping discs for \( c = 0 \) and the Poisson–Boolean model for \( c = 1 \), with \( l(u > R) = 1 \) if \( u > R \), 0 otherwise. An interaction of the form \( g(u, R) = cl(uR) + e^{-\beta/2u}I(u>R) \), for \( c = 0 \) yields again a NORDM, but simulations show that this model tends to cluster the discs toward the average of the foci. For this reason, we will adapt the Strauss-like interaction (Baddeley and Moller, 1989) to allow us to sample from the whole domain of observation.

Now, set \( I_{ij} = I(\|s_i - s_j\| > R_i + R_j) \), for all \( i \neq j \). The model for the foci conditional on the radii can be written as
\[
\pi(s|n, R) \propto \prod_{i=1}^n \prod_{j=1}^n I_{ij}. \tag{6}
\]

We consider modeling the distribution of each radius independently as \( R_i \sim f_r(q(a_i, b_i)), \) with (the possibly improper) joint hyperpriors for \( a_i \) and \( b_i \), modeled as
\[
\pi(a_i, b_i|a_{i1}, b_{i1}, b_{i2}) \propto \pi(a_i|a_{i1}, a_{i2}) \pi(b_i|b_{i1}, b_{i2}) \propto \Gamma(a_i + 2M) e^{-a_i b_i} \frac{b_i}{b_{i1} + 1},
\]
for \( i = 1, \ldots, n \), respectively. To allow the data to provide information about the weights \( p = (p_1, \ldots, p_n) \), we consider a Dirichlet prior \( p|d \sim Dirichlet(d) \), where \( d = (d_1, d_2, \ldots, d_n) \).

Now, we can write the posterior distribution of \( s = (s_1, \ldots, s_n) \), \( R = (R_1, \ldots, R_n) \), \( p = (p_1, \ldots, p_n) \), \( a = (a_1, \ldots, a_n) \), and \( b = (b_1, \ldots, b_n) \) given the observed points and number of foci as
\[
\pi(s, R, p, a, b|x_1, \ldots, x_M, n) \propto \prod_{j=1}^M f_s(x_j|s, R, p, n) \pi(s|n, R) \pi(R|n, a, b)
\times \pi(p|n, d) \prod_{i=1}^n \pi(a_i, b_i|a_{i1}, b_{i1}, b_{i2}), \tag{7}
\]
and hence,
\[
\pi(s, R, p, a, b|X, n) \propto \prod_{i=1}^n \frac{p_i^{q(i)}}{\pi(\|s_{|i}\|, R_{|i})} \prod_{j=1}^{q(i)} I_{|s_{|i}, R_{|i}}(x_{i,j})
\times \prod_{i=1}^n \prod_{j=1}^n I_{ij}
\times \prod_{i=1}^n R_i^{a_i+2M-1} e^{-R_i b_i} \Gamma(a_i + 2M) e^{-a_i b_i} \frac{b_i}{b_{i1} + 1}, \tag{8}
\]
which leads to the full conditional distributions.
\[ \pi(s_k^j) \propto \frac{q(k)}{\prod_{j=1}^n I_{B(s_k^j)R_k} (x_k) \prod_{j=1}^n I_{s_k^j}}, \]

\[ \pi(R_k^j) \propto R_k^{2M-2q(i)-1} e^{\frac{R_k}{2} I(R_k > \max_{i=1}^{M} ||s_k - s_k||)} \times I(R_k < \min_{i=1}^{M} (||s_k - s_k|| - R_k)), \]

\[ \pi(a_k^j) \propto R_k^j e^{-a_k^j} b_k^j \sim IG(b_k + a_k + 2M, b_k + R_k), \]

\[ \text{p(d.q)~Dirichlet(d + q),} \]

where \( q = (q_1, \ldots, q(n)), k = 1, 2, \ldots, n. \)

It is critical to note that \( L_M = \prod_{t=1}^{M} f_X(s, x, R, p, n) \) is not the likelihood of the NORDM, but only part of the likelihood. To build the likelihood function of a NORDM, we must visualize the random set as the collection of random functions \( C = \{L(x): x \in \mathcal{O}\} \). That is, we think of a realization of the NORDM as a collection of \( 0-1 \) values throughout \( \mathcal{O} \). The quantity \( L_M \) provides the likelihood for only \( M \) randomly selected points out of \( \text{card} \mathcal{O} \) points. See Cressie (1993, sec. 9.6) for a discussion on the difficulties one encounters in constructing this likelihood directly, which makes the implementation of classical statistical inference approaches very difficult. The hierarchical Bayesian approach does not suffer from these difficulties, because the data model is used to update the initial (prior) beliefs about the distribution of the foci and corresponding radii, resulting in the posterior distribution that can be used to conduct inference. In summary, the hierarchy of the model can be described as [Data] NORDM[NORDM] Parameters, which emphasizes that the data are just a finite set of points that informs about the NORDM.

3. MODELING GROWTH OR EVOLUTION USING A NORDM

Now, suppose that we observe several states of some object in \( \mathcal{O} \) denoted by \( X_1, X_2, \ldots, X_t \). Our goal is to describe mathematically (using a NORDM) the nature of the clusters of points observed in each state, as well as the evolution of these points across time. To model growth using a NORDM, we can define conditional upon \( n_{t+1} \),

\[ X_{t+1} = \bigcup_{i=1}^{n_{t+1}} \{B(s^{(t+1)}_i, R^{(t+1)}_i): s^{(t+1)}_i \in X_t \}. \]

That is, growth (a new grain) occurs only at foci that belong to a grain in the previous state. Hence, suppose that we model the NORDM \( X_t \) as described in Section 2, and obtain (samples from) the full posterior distribution \( \pi(s^{(t)}_i, R^{(t)}_i, p^{(t)}_i, a^{(t)}_i, b^{(t)}_i | x^{(t)}, n^{(t)}) \), where \( x^{(t)}_i \) are \( M_t \) realizations from \( X_t \). We can perform inference on the parameters \( s^{(t)}_i, R^{(t)}_i, p^{(t)}_i, a^{(t)}_i, b^{(t)}_i \) using standard Bayesian methods, for example posterior means, variances, and credible sets. To completely characterize the observed next state \( X_{t+1} \) in the growth model, we need to estimate \( s^{(t+1)}_i \) and \( R^{(t+1)}_i, i = 1, \ldots, n_{t+1} \). The mixture distribution of (3) can be written as

\[ f_X(x | s^{(t+1)}_i, R^{(t+1)}_i, p^{(t+1)}_i, n_{t+1}) \]

\[ = \sum_{i=1}^{n_{t+1}} \pi(R^{(t+1)}_i)^{2J} B(s^{(t+1)}_i, R^{(t+1)}_i)(x), \]

with \( s^{(t+1)}_i \in X_t, i = 1, 2, \ldots, n_{t+1} \).

Our goal is to perform statistical inference on the parameters of the growth model (9) based on sampled points from the NORDM, which is modeled using (10). We obtain \( \pi(s^{(t+1)}_i, R^{(t+1)}_i, p^{(t+1)}_i, a^{(t+1)}_i, b^{(t+1)}_i | x^{(t+1)}_i, n_{t+1}, s^{(t)}_i, R^{(t)}_i, n^{(t)}) \) using similar priors on the parameters as previously described, with a major difference in the priors for \( s^{(t+1)}_i \) and \( R^{(t+1)}_i \), to account for the fact that the foci now arise from the previous state. Hence, we will consider

\[ \pi(s^{(t+1)}_i | p_{i,t}, n_{t+1}, s^{(t)}_i, R^{(t)}_i, R^{(t+1)}_i) \]

\[ \propto \prod_{i=1}^{n_{t+1}} \left( \sum_{j=1}^{n_t} I \left( \|s^{(t+1)}_i - s^{(t)}_j\| \leq R^{(t)}_j \right) \right) \]

\[ \times \prod_{i=1}^{n_{t+1}} I \left( \|s^{(t+1)}_i - s^{(t)}_j\| \leq R^{(t+1)}_i \right), \]

where \( I(\cdot) \) is an indicator function that equals 1 if \( \cdot \), the \( \text{th} \) focus of state \( t+1 \) appeared in the \( \text{th} \) disc centered at \( s^{(t)}_j \), with radius \( R^{(t)}_j \), from the previous state [i.e., \( s^{(t)}_j \in X_t \), and \( I^{(t+1)}_i = I(\|s^{(t+1)}_i - s^{(t)}_j\| > R^{(t)}_j + R^{(t+1)}_i) \) for all \( i \neq k \), the nonoverlapping discs conditions].

To complete the hierarchical formulation, we require a model for the prior \( \pi(R^{(t+1)} | n_{t+1}, s^{(t)}_i, R^{(t)}_i, s^{(t)}_j) \). This can be done in several ways. If we consider location to be of primary interest instead of the size of the grains (for example, the center of a storm system is of great interest because heavier precipitation often occurs at that location), then it is more appropriate to assume gamma prior distributions for the radii not affected by the previous radii and foci. In this way, we let locations influence more the NORDM distribution for the next time period. We will adapt this approach in our illustrations herein.

3.1 Estimating the Next State of a NORDM

Consider two observed states, \( X_t \) and \( X_{t+1} \), conditional on observed \( n_t \) and \( n_{t+1} \). Suppose that we observed \( M_t \) points from \( X_t \), say \( x_{t,k} \), \( j = 1, \ldots, q(t) \), \( i = 1, 2, \ldots, n_t \), and \( M_{t+1} \) points from \( X_{t+1} \), say \( x_{t+1,k} \), \( j = 1, \ldots, q(t+1) \), \( i = 1, 2, \ldots, n_{t+1} \), where \( \sum_{i=1}^{n_t} q(i) = M_t \) and \( \sum_{i=1}^{n_{t+1}} q(i) = M_{t+1} \). Let \( X^{(t)} \) denote all observed points at time \( t \). Then the likelihood based on the sampling distribution (3) for \( X_t \) can be written as

\[ f_X(x^{(t)} | s^{(t)}_i, R^{(t)}_i, p^{(t)}_i, n^{(t)}) \]

\[ = \prod_{i=1}^{n_t} \pi(p^{(t)}_i)^{q(i)} \]

\[ \times \prod_{j=1}^{q(i)} I(B(s^{(t)}_j, R^{(t)}_j)(x_{t,k})), \]

while the (growth) likelihood of (10) for \( X_{t+1} \) is given by

\[ f_X(x^{(t+1)} | s^{(t+1)}_i, R^{(t)}_i, R^{(t+1)}_i, p^{(t+1)}_i, n_{t+1}) \]

\[ = \prod_{i=1}^{n_{t+1}} \pi(p^{(t+1)}_i)^{q(i)} \]

\[ \times \prod_{j=1}^{q(i)} I(B(s^{(t+1)}_j, R^{(t+1)}_j)(x_{t+1,k})), \]

Note that if the conditions \( s^{(t+1)}_i \in X_t, i = 1, 2, \ldots, n_{t+1} \), are included in the sampling scheme, then there is no need to include them in the prior (11). We presented them in (11) to illustrate how one can go about generating the next state foci.
for forecasting purposes (i.e., when no observed points are available from the next state \( t + 1 \)).

A comment is in order about our approach related to a sequential Monte Carlo algorithm. In that context, Equation (12) is the data model, and the state evolution model is written (hierarchically) in the priors for the future foci and radii given in Equation (11) and the associated discussion. In traditional sequential algorithms, the predictive distribution of the future states given the data through the present can be derived from Bayes’ rule (since it is just the posterior predictive distribution).

In our case, as in general sequential sampling, this cannot be done analytically, but could be sampled using one’s favorite importance sampling algorithm (e.g., particle filters, and so on). However, we note that our current algorithm is quite efficient computationally, and we would not see appreciable computational benefits from a sequential algorithm.

Consequently, Bayesian estimation of the foci and grains of the next state proceeds now as usual via MCMC. The full posterior distribution for state \( t + 1 \) is given by

\[
\pi(s^{(t+1)}, R^{(t+1)}, p^{(t+1)}, a, b | X^{(t+1)}, s^{(t)}, R^{(t)}, n_t, n_{t+1}) \propto f_X(s^{(t+1)}, R^{(t+1)}, p^{(t+1)}, n_t, n_{t+1})
\]

\[
\times \pi(s^{(t+1)} | n_t, n_{t+1}, s^{(t)}, R^{(t)}, p^{(t)}) \pi(R^{(t+1)} | n_{t+1}, a, b) 
\]

\[
\times \pi(p^{(t+1)} | n_{t+1}, d) \prod_{i=1}^n \pi(a_i, b_i | a_{i1}, b_{i1}, b_{i2}),
\]

where \( \hat{s}^{(t)} \), \( \hat{R}^{(t)} \) are the estimated foci and radii that describe \( X_t \), and are computed using the modeling framework of Section 2.2. The full conditional distributions are easily obtained based on (14) as follows:

\[
\pi(s_k^{(t+1)} | . ) \propto \left( \sum_{i=1}^n I( \| \hat{s}_k^{(t+1)} - \hat{s}_i^{(t)} \| \leq \hat{R}_i^{(t)} ) \right)
\]

\[
\times \prod_{j=1}^{n_{t+1}} I_{B_k^{(t+1)}, R_k^{(t+1)}}(x_{t+1, k}) 
\]

\[
\times \prod_{i=1, i \neq k} I( \| s_k^{(t+1)} - s_{i}^{(t+1)} \| > R_i^{(t+1)} + R_k^{(t+1)},
\]

\[
\pi(R_k^{(t+1)} | . ) \propto R_k^{(t+1)^{a_k + 2M - 2q_{1+k}} - 1} e^{-\frac{R_k^{(t+1)}}{b_k}}
\]

\[
I(R_k^{(t+1)} > \max_{j=1, q_{1+k}} \| s_k^{(t+1)} - x_{t+1, k} \| 
\]

\[
\times I(R_k^{(t+1)} < \min_{i \neq k} \| s_i^{(t+1)} - s_k^{(t+1)} \| - R_i^{(t+1)} ),
\]

\[
\pi(a_k | . ) \propto R_k^{(t+1)^{a_k}} e^{-q_{a_k} a_k},
\]

\[
b_k | . \sim IG(a_k + 2M + b_{k1}, b_{k2} + R_k^{(t+1)}),
\]

\[
p^{(t+1)} | . \sim Dirichlet(d + q),
\]

where \( q = (q_{t+1} (1), \ldots, q_{t+1} (n_{t+1})), k = 1, 2, \ldots, n_{t+1} \).

3.2 Capturing Growth Between Two Observed Realizations

In the context of modeling storm system growth, we are interested in the area of the cell, its growth ratio with respect to the previous time period and, of course, the expected area and value of the cell in the following time periods. If we have two observed states, then we can estimate parameters in both and compute the ratio of their areas, \( |X_t | / |X_{t+1} | \). This ratio provides a measure of relative growth that can be used, for example, to help forecasters assert if the storm seems to be reducing in strength.

The area of a NORDM can be easily defined as follows. Using (1) we see that the area of a NORDM is

![Figure 2. Observed storm systems at seven equally spaced times (states) between 2:30 UTC to 3:30 UTC. (left images) Radar reflectivities and (right images) lower than threshold (35) intensities removed.](image)
\[ |X_j| = \sum_{i=1}^{n} \left| B(s^{(i)}, R^{(j)}) \right| = \pi \sum_{i=1}^{n} (R^{(j)})^2, \]

because the discs do not overlap. Hence, the expected area ratio of the two consecutive observed NORDM’s is \( E(|X_j|/|X_{j+1}|) \), which is easy to compute using posterior samples from the radii of each state of the NORDM.

3.3 Model Adequacy Tests for a NORDM

We consider model adequacy tests to assess whether the NORDM fits the data well, with respect to choice of grain (disc) and the distribution used to describe points sampled over the discs. The method used to assess these important assumptions of the model are Monte Carlo tests, similar in spirit to those in Gelman, Carlin, Stern, and Rubin (2004, p. 477). We consider four different test statistics that are computed based on the observed points, \( x_{kj}, j = 1, \ldots, q(i); i = 1, 2, \ldots, n \) (e.g., the clusters of points in the processed images of Section 4) and posterior predictive samples \( x_{ukj}, j = 1, \ldots, q(i); i = 1, 2, \ldots, n; \ u = 1, 2, \ldots, U \). The latter are obtained as follows: we fit the NORDM on the original data \( x_{kj} \) and obtain \( U \) iterations (i.e., \( U \) realizations of the \( n \) foci and corresponding radii), and then generate the predictive samples. Using the original and the sampled data we compute the following statistics and the Monte Carlo \( p \)-values: \( T_{\text{min}} \) = the average (for all foci) of the smallest distance of the points from the focus in a cluster, \( T_{\text{max}} \) = the average (for all foci) of the largest distance of the points from the focus in a cluster, \( T_{\text{avg}} \) = the average (for all foci) of the average distance of the points from the focus in a cluster, and \( T_{\text{var}} \) = the average (for all foci) of the sample variance of the distances of the points from the focus in a cluster. The test can assess, by construction, all the major assumptions of the model simultaneously, such as the choice of the grain being a disc with distribution being uniform over the disc.

Figure 3. Observed storm systems with lower than threshold (35 dBZ) intensities removed. The fitted NORDM’s are also displayed with solid discs.
4. APPLICATION: MODELING EVOLUTION OF SEVERE STORM SYSTEMS

Monitoring the development of a severe storm is of great importance to the public in terms of potential loss of life and property. The proposed methodology does not only help to better understand the growth of such systems, but also provides a measure of uncertainty in the forecasted storm (random set) for the next time period.

To illustrate the methods, we study data from the San Antonio, TX area on 5 July 2002. This day saw a series of intense convective storm cells track over the area resulting in some flash flooding. A mosaic of three radars was used to capture the event, including Corpus Christi, TX (KCRP), Brownsville, TX (KBRO), and San Antonio, TX (KEWX). The data were reduced to a $32 \times 32$ grid of intensities, where the resolution of each pixel is roughly four square kilometers. The reflectivities (intensities) are measured in units of decibels (dBZ). We observe the area from 2:30 UTC to 3:30 UTC, where UTC means Coordinated Universal Time, with a 10 minute interval between observed radar images. Hence, we will consider seven observed progressive realizations of the storm system over the time period of interest.

The nature of the radar data in the processed radar image (with low intensities removed) displays the characteristics of a NORDM (i.e., different clusters of points, with each corresponding to a storm system). Once we fit the NORDM, we can use the estimated foci and corresponding radii to forecast the next state of the storm system. Hence, the purpose of the modeling approach is to describe mathematically (using a NORDM) the nature of the clusters of points observed as well as the evolution of these points across different time periods, which is of great interest to atmospheric scientists. For example, the next state of the NORDM allows for better understanding of important characteristics of the storm system, such as the locations of intense rainfall (areas near the foci) in the next time period and their areas, which can be used to provide adequate warnings to the population that will be affected and might be in imminent danger.

High intensities indicate more severe storm systems. Typically, atmospheric scientists know what threshold value is of interest for a particular type of precipitation system, and the decision with regard to the choice of threshold is determined by the forecaster. In our case, intensities higher than 35 dBZ are of great interest, because they have the potential to cause intense rainfall.

Figure 2 illustrates the situation at successive times (or states) from 2:30 UTC to 3:30 UTC (states 1, 2, . . . , 7). The left side shows the full reflectivity fields for the observed storm systems and the right side the corresponding images after intensities lower than 35 dBZ are removed. For example, notice that there are four identifiable grains (clusters of pixels) in state 1. Each cluster of pixels can be thought of as part of a random grain (a disc), and the collection of pixel centroids is modeled as a set of points from a realization of a NORDM (the union of the discs). Grains with less than five pixels are not considered to be valid. This is reasonable for the current application, as the scientific concern is the relatively large-scale precipitation and associated flash flooding potential.

The model results were obtained by running 50,000 iterations of the MCMC (with a 10,000 iteration burn-in period) on a Pentium 1.862 Mhz, with 1 GB Ram, with the sampler completed after about 35 min. Updating was straightforward for all parameters, except for the hyperparameters $\alpha_k$, which required (random walk) Metropolis–Hastings steps. From our simulations, we found that for each additional disc we introduced, the processing time for fitting a NORDM on a single image was approximately the same.

We first present the fitted NORDM in Figure 3 using solid discs. For example, in Figure 2 (top image), we show the radar image of the area at time 2:30 UTC along with the processed image to identify the grains. The corresponding Figure 3 (state 1) displays the points of the grain(s) and the discs (solid line) we fit based on model (8). The remaining images illustrate the NORDM fits using the growth model (14). We display the posterior means for the foci and radii in Table 1, without their credible sets for the sake of brevity.

In Figure 4 we present the fits from current and previous states for comparison and to illustrate how our model can capture this type of growth; namely, the next grain always arises at a focus that belongs to a grain from the previous state. The solid discs are the current (fitted state) based on the previous state (dotted discs). We notice that each focus of the next state belongs to a grain from the previous state, as expected.

Next, we compute relative growth measures for the threshold value of 35 dBZ. The measures based on the ratio of expected areas, defined in Section 3.2, between states 1 and 2, 2 and 3, and so on, are 1.2328, with 95% Credible Set (C.S.) (1.2166, 1.2551), 1.3397, with C.S. (1.3092, 1.3790), 1.6054, with C.S. (1.5606, 1.6438), 1.2735, with C.S. (1.2408, 1.4038), 0.6402, with C.S. (0.5890, 0.6664), and 0.9803, with C.S. (0.9601, 0.9901).

### Table 1. Bayesian point estimates (posterior means) of foci coordinates and corresponding radii from the fitted growth models between all states 1–2, 2–3, and so on

<table>
<thead>
<tr>
<th>State</th>
<th>Point estimators for (x, y) foci cords for each grain</th>
<th>Point estimator of the corresponding radius for each grain</th>
<th>Grains in state</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>(3.93, 27.02)</td>
<td>5.20</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>(7.93, 10.85)</td>
<td>4.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(15.83, 2.35)</td>
<td>5.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(13.64, 28.63)</td>
<td>2.03</td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>(2.41, 23.91)</td>
<td>1.11</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>(6.95, 29.06)</td>
<td>4.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.06, 10.50)</td>
<td>2.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(16.13, 2.67)</td>
<td>5.16</td>
<td></td>
</tr>
<tr>
<td>3rd</td>
<td>(7.06, 28.40)</td>
<td>5.30</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(16.61, 3.68)</td>
<td>3.75</td>
<td></td>
</tr>
<tr>
<td>4th</td>
<td>(5.77, 24.17)</td>
<td>1.17</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>(10.08, 28.89)</td>
<td>4.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(16.33, 3.29)</td>
<td>3.26</td>
<td></td>
</tr>
<tr>
<td>5th</td>
<td>(8.13, 27.61)</td>
<td>2.24</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(16.33, 3.28)</td>
<td>3.95</td>
<td></td>
</tr>
<tr>
<td>6th</td>
<td>(8.06, 27.64)</td>
<td>4.60</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(16.03, 2.42)</td>
<td>3.27</td>
<td></td>
</tr>
<tr>
<td>7th</td>
<td>(6.68, 27.19)</td>
<td>4.98</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(16.51, 2.09)</td>
<td>2.74</td>
<td></td>
</tr>
</tbody>
</table>
0.9866). Hence, the measure indicates overall reduction in area for the first five states, and growth for the remaining cases, meaning that the storm reduced in area at times 2:40–3:00 UTC and increased in area after time 3:10 UTC. Alternatively, we could perform a cluster-to-cluster comparison deterministically and obtain a subjective growth measure as is done by operational nowcasters. However, our approach allows us to capture the variability of such measures as shown in the credible sets. Such a well-defined measure of growth and associated uncertainty measures is not currently available to the operational forecaster. Thus, the methodology provides a new tool for use in nowcasting.

Clearly, selecting a different threshold level will affect the results, and studying parameter estimate values obtained for different threshold levels reveals that estimation procedures are not robust to the choice of threshold. This is not unusual, and meteorologists are well aware of the effect of selecting different threshold values. Typically, atmospheric scientists pro-

Figure 4. Fitted NORDM at times 2:30 UTC to 3:30 UTC. Displaying the fitted (solid) discs from the NORDM growth model. The previous state is also displayed as dashed discs.
pose the threshold value they deem to be appropriate prior to any statistical analysis. Some discussion about the effect of threshold selection can be found in Micheas et al. (2007), and this is an ongoing research problem that is under investigation.

4.1 NORDM Adequacy Tests for Radar Data

We consider the model adequacy test described in Section 3.3, to assess whether the NORDM fits the radar data well, with respect to choice of grain (disc) and the distribution used to describe points sampled over the discs. We apply the procedure for each state and report our findings in Table 2. Notice that when considering statistics $T_{\text{max}}$ and $T_{\text{avg}}$, we have strong support that a NORDM (with all its assumptions) would fit the data well. That is not the case with statistics $T_{\text{min}}$ and $T_{\text{var}}$, where for the first state there is support that a NORDM would fit the data, but not for the remaining states. This is not unusual in this context, and the results highly depend on the nature of the data under consideration. To further assess the validity of our choice of test statistics for this data, we generated simulated data from known NORDM’s and tested model adequacy. Even in this case, where we know the NORDM to be true, the values of the test statistics $T_{\text{min}}$ and $T_{\text{var}}$ were very small, with $p$-values below 0.2. The statistics $T_{\text{max}}$ and $T_{\text{avg}}$ performed well again, with $p$-values over 0.8. This leads us to believe that $T_{\text{max}}$ and $T_{\text{avg}}$ are the appropriate statistics to use when testing model adequacy for a NORDM, and hence support the validity of Table 2. Using the results based on $T_{\text{max}}$ and $T_{\text{avg}}$ we see that the model fits well in all cases, showing that a NORDM is a reasonable model to use when we study propagation of severe storm systems.

5. SUMMARY

The proposed methodology allows us to better understand the growth or evolution of objects such as tumors, severe storm cells, or an epidemic as it spreads through a populated area. The methods are very amenable to changes in the grains (discs, ellipses, polygons), the sampling scheme for a point from the random set, and the selection of prior information on the foci and radii, or other random variables that may describe a grain, such as the vertices and their distances from the centroid for polygons. The models for the priors also allow inclusion of relevant covariates. For example, in a nowcasting scenario, variables such as wind (i.e., advection) or atmospheric stability might be important covariates.

Herein, we explored the case of nonoverlapping grains, which is attractive both mathematically and, more importantly, from an application point of view. Under this assumption, defining a probability distribution for a sampled point from the random set is a straightforward task, but without it, one would have to carefully follow the tedious route of using hitting probabilities. Furthermore, it can be argued that allowing cells to overlap is inappropriate in many cases. Consider, for instance, observation of the growth of white blood cells (non-overlapping grains scenario). Our approach allows for cells to either merge or split between observed time periods, a situation that cannot be captured adequately using overlapping discs.

Indeed, such dynamics in the case of severe weather nowcasting motivated the development of this methodology.

There are several extensions to this work that could be considered. We could easily accommodate alternative hierarchical Bayesian multistage NORDM’s than the one presented with regards to choice of data model, priors, and hyperpriors. This would allow us to study more complicated data structures. In addition, one could investigate the modeling of grains other than the disc, for example, random ellipses, parallelograms, and polygons. If we can describe any compact convex set using a finite number of parameters, then we can easily treat it as a grain. The major assumption here was that of a disc, where we model the centers (foci) and their radii. A more realistic convex and compact set would be a polygon with a random number of vertices. Clearly, such a set could be modeled with our approach once the centroid, the number of vertices and their distance from the center, was known. For example, a random polygon with $n$ vertices could be described by its center and the $n$ vertices, $2 + 2n$ parameters in all. Note that canonical polygons (those having the same distance between any two vertices) have only $2 + 2 + 1$ parameters to consider for the centroid, the angle between successive vertices and the starting vertex.

The extension to ellipses or polygons is not hard to model, provided that we have a method to model the distribution of a point from these sets. For example, for ellipses a uniform distribution can be analogous to $1/(\pi ab)$, where $a$ and $b$ are the semiaxes lengths of the ellipse. (Similarly for polygons.) This introduces one more parameter (instead of just a radius), which

<table>
<thead>
<tr>
<th>State</th>
<th>$T_{\text{min}}$</th>
<th>$T_{\text{max}}$</th>
<th>$T_{\text{avg}}$</th>
<th>$T_{\text{var}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>0.4865 (0.0604)</td>
<td>4.2230 (0.8810)</td>
<td>2.2457 (0.9945)</td>
<td>1.1446 (0.0716)</td>
</tr>
<tr>
<td>2nd</td>
<td>0.6299 (0.0014)</td>
<td>3.4449 (0.9490)</td>
<td>2.0557 (0.7139)</td>
<td>0.8811 (0.0271)</td>
</tr>
<tr>
<td>3rd</td>
<td>0.5348 (0.0067)</td>
<td>4.5300 (0.9812)</td>
<td>2.2795 (1.0000)</td>
<td>1.1572 (0.4111)</td>
</tr>
<tr>
<td>4th</td>
<td>0.4581 (0)</td>
<td>2.8203 (0.9914)</td>
<td>1.6194 (0.9959)</td>
<td>0.6313 (0.0004)</td>
</tr>
<tr>
<td>5th</td>
<td>0.4854 (0)</td>
<td>3.1013 (0.9869)</td>
<td>1.6980 (0.9688)</td>
<td>0.6028 (0.1075)</td>
</tr>
<tr>
<td>6th</td>
<td>0.5377 (0)</td>
<td>3.9398 (0.9227)</td>
<td>2.2583 (0.7888)</td>
<td>1.0717 (0.0002)</td>
</tr>
<tr>
<td>7th</td>
<td>0.4393 (0.0061)</td>
<td>3.8648 (0.9216)</td>
<td>2.1488 (0.9219)</td>
<td>1.1103 (0)</td>
</tr>
</tbody>
</table>

NOTE: Presenting the values of the test statistics with their $p$-values. Notice that the model fits well each state if we consider statistics $T_{\text{max}}$ and $T_{\text{avg}}$. The statistics are defined as: $T_{\text{min}} = \frac{\text{the average (for all foci) of the smallest distance of the points from the focus in a cluster}}{\text{the average (for all foci) of the largest distance of the points from the focus in a cluster}}$, $T_{\text{max}} = \frac{\text{the average (for all foci) of the largest distance of the points from the focus in a cluster}}{\text{the average (for all foci) of the sample variance of the distances of the points from the focus in a cluster}}$. $T_{\text{avg}}$ performed well again.
can be easily modeled with our approach. The choice of truncated normal instead of uniform can be incorporated by our modeling approach, with appropriate changes across Equations (2), (3), (8), and (10)–(14), and the data (cluster of points) would typically govern the choice of such distribution. Now, if we had observed realizations of the NORDM, we could easily obtain the number of foci $n_j$. For forecasting purposes (no observed data from the next state), we could consider modeling $n_j$ as $n_j \sim \text{Poisson}(\lambda_j)$, with an evolution model on $\lambda_j$. For example, as is common with time series, we could allow $\log(\lambda_j)$ to follow an AR(1) process: $\log(\lambda_{j+1}) = \phi \log(\lambda_j) + \epsilon_j$, where $\epsilon_j \sim N(0, \sigma^2)$, and $\phi$ the autoregression coefficient, with possible hyperpriors $\phi \sim U(-1, 1)$ and $\sigma^2 \sim IG(\alpha_p, \beta_p)$. The foci and radii of the one-step forecast for time $(j + 1)$, could easily be modeled using (11) (see discussion before Section 3.1) and conditional on $n_{j+1}$. Furthermore, various types of Markov models could be considered for modeling $n_j$ directly.

Finally, the choice of threshold along with $m$, the minimum number of pixels used to retain a grain, merit further investigation. If we include $m$ as a parameter in the model, then the usual MCMC is inappropriate, because the number of grains $n_j$ would change, possibly with each iteration. In this case, methods such as a reversible-jump (Robert and Casella, 2004, p. 540) MCMC are required. The latter algorithm can also be helpful in terms of the choice of threshold, a case in which we also observe a different number of $n_j$.

These are subjects for future research and will be described elsewhere.

[Received November 2007. Revised November 2008.]

REFERENCES


