2002

Asymptotic inference for spatial CDFS over time

Jun Zhu
University of Wisconsin

S N. Lahiri
University of Wisconsin

Noel A. Cressie
University of Wollongong, ncressie@uow.edu.au

Asymptotic inference for spatial CDFS over time

Abstract

A spatial cumulative distribution function (SCDF) is a random function that provides a statistical summary of a random process over a spatial domain of interest. In this paper, we consider a spatio-temporal process and establish statistical methodology to analyze changes in the SCDF over time. We develop hypothesis testing to detect a difference in the spatial random processes at two time points, and we construct a prediction interval to quantify such discrepancy in the corresponding SCDFs. Using a spatial subsampling method, we show that our inferences are valid asymptotically. As an illustration, we apply these inference procedures to test and predict changes in the SCDF of an ecological index for foliage condition of red maple trees in the state of Maine in the early 1990s.

Keywords

time, asymptotic, over, cdfs, spatial, inference

Disciplines

Physical Sciences and Mathematics

Publication Details

Zhu, J., Lahiri, S. N. & Cressie, N. A. (2002). Asymptotic inference for spatial CDFS over time. Statistica Sinica, 12 (3), 843-861.

Statistica Sinica **12**(2002), 843-861

ASYMPTOTIC INFERENCE FOR SPATIAL CDFS OVER TIME

Jun Zhu, S. N. Lahiri and Noel Cressie

University of Wisconsin–Madison, Iowa State University and The Ohio State University

Abstract: A spatial cumulative distribution function (SCDF) is a random function that provides a statistical summary of a random process over a spatial domain of interest. In this paper, we consider a spatio-temporal process and establish statistical methodology to analyze changes in the SCDF over time. We develop hypothesis testing to detect a difference in the spatial random processes at two time points, and we construct a prediction interval to quantify such discrepancy in the corresponding SCDFs. Using a spatial subsampling method, we show that our inferences are valid asymptotically. As an illustration, we apply these inference procedures to test and predict changes in the SCDF of an ecological index for foliage condition of red maple trees in the state of Maine in the early 1990s.

Key words and phrases: Environmental resource assessment and monitoring, Functional central limit theorem, Spatial cumulative distribution function, Spatial prediction, Spatial subsampling, Spatio-temporal process.

1. Introduction

Well designed, large-scale, long-term ecological resource monitoring programs allow study of the current status of and changes in the nation's ecological resources on a regional basis, and hence are of great importance. Indicators of ecological resources over regions of interest are often selected and modeled as a spatially distributed random process, known as a random field (r.f.) with continuous spatial index. A spatial cumulative distribution function (SCDF) is a random function that provides a spatial statistical summary of a r.f. In the past, statistical inference for the SCDF at a given time point has been developed to determine the current condition of an ecological resource (e.g., Lahiri, Kaiser, Cressie and Hsu (1999)). This paper extends these earlier results and develops inference for comparing SCDFs over time in order to *detect* and *quantify* changes in the ecological resources at different points in time. The proposed methodology is illustrated with forest-health monitoring data collected from a spatial network of monitoring sites in the state of Maine in the early 1990s.

We consider a r.f. with continuous spatial index, $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in R\}$, where $Z(\boldsymbol{s})$ is a random variable at the spatial location $\boldsymbol{s}$ and $R \subset \mathbb{R}^d$ is a spatial domain of interest. Its SCDF is $F_\infty(z;R) \equiv |R|^{-1}\int_R I(Z(\boldsymbol{s}) \leq z)d\boldsymbol{s},\ z \in I\!\!R$, where $|R| \equiv \int_R d\boldsymbol{s}$ denotes the volume of R and I is the indicator function. Note that F_∞ is a random cumulative distribution function (CDF). Specifically, it is a function of actual *and* potential observations, of which any realization is a CDF that is right continuous and increases from 0 to 1. All spatial moments, areal proportions, and spatial quantiles can be recovered from the SCDF. For example, the regional mean of the r.f. $Z(\cdot)$, $Z(R) \equiv |R|^{-1}\int_R Z(\boldsymbol{s})d\boldsymbol{s}$, is equal to $\int z dF_\infty(z;R)$. The SCDF can be thought of as a basic (random) functional whose properties can be used to indicate resource status.

A commonly used predictor of the SCDF is the empirical CDF (ECDF). Based on a finite sample $\{Z(\boldsymbol{s}_1),\ldots,Z(\boldsymbol{s}_N)\}$, observed at known spatial sampling sites $\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_N\} \subset R$, we define the equal-weight version of the ECDF as $F_N(z;R) \equiv N^{-1}\sum_{i=1}^N I(Z(\boldsymbol{s}_i) \leq z),\ z \in I\!\!R$. Figure 1 gives the ECDFs of foliage condition indices sampled in the state of Maine in 1991, 1992 and 1993. Using an asymptotic framework and a subsampling method, Lahiri, Kaiser, Cressie and Hsu (1999) predict the complete-data version SCDF with the ECDF, and construct large-sample prediction bands for the SCDF that achieve the desired confidence levels asymptotically. Lahiri (1999) gives further theoretical results for the SCDF.

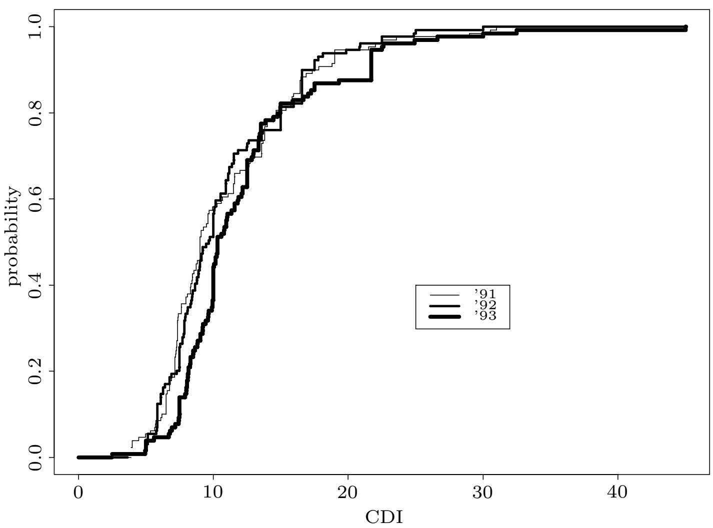

Figure 1. Empirical cumulative distribution functions (ECDF) of the crown defoliation index (CDI) values for 1991, 1992 and 1993.

In this paper we use a similar asymptotic framework, which is a combination of "increasing domain asymptotics" and "infill asymptotics" (Cressie (1993, pp.100-101)). We now describe the asymptotic framework in detail, using the following notation throughout the rest of the paper. For the increasing domain component of the asymptotic structure, let R_0 denote an open connected subset of $(-1/2, 1/2]^d$ that contains the origin $\mathbf{0}$. Then we obtain the sampling region R_n by "inflating" the region R_0 with $R_n \equiv \lambda_n R_0$, where $\{\lambda_n\}$ is a scaling sequence of positive real numbers tending to infinity as $n \to \infty$. For the infill component of the asymptotic structure, we partition the sampling region R_n into equal-volume cubes and denote them as $\Gamma(\boldsymbol{j}) \equiv (\boldsymbol{j} + (0,1]^d)h_n$, where $\boldsymbol{j} \in \mathbb{Z}^d$, $\mathbb{Z}$ denotes the set of integers, $\Gamma(\boldsymbol{j}) \cap R_n \neq \emptyset$, and $\{h_n\}$ is a sequence of positive real numbers decreasing to 0 as $n \to \infty$. Let $\boldsymbol{c}$ denote an arbitrary point in the interior of the unit cube $(0,1]^d$ and assign $\boldsymbol{c}h_n$, in the cube $(0, h_n]^d$, to be the starting sampling site. Then, the sampling sites in R_n form a grid $\{(\boldsymbol{j}+\boldsymbol{c})h_n \in \Gamma(\boldsymbol{j}) \cap R_n : \boldsymbol{j} \in \mathbb{Z}^d\}$; as $n \to \infty$ the number of sampling sites in R_n, denoted by N_n, satisfies the growth condition, $N_n \sim |R_0|\lambda_n^d/h_n^d$ ($u_n \sim v_n$ means $u_n/v_n \to 1$ as $n \to \infty$). An illustration of the asymptotic structure and the sampling design in $\mathbb{R}^2$ is shown in Figure 2. Within the circled sampling regions R_{n-1} and R_n, the sampling sites form a square grid. Note that the sample size increases from N_{n-1} to N_n as a result of both a *larger* sampling region and a *finer* grid of sampling sites.

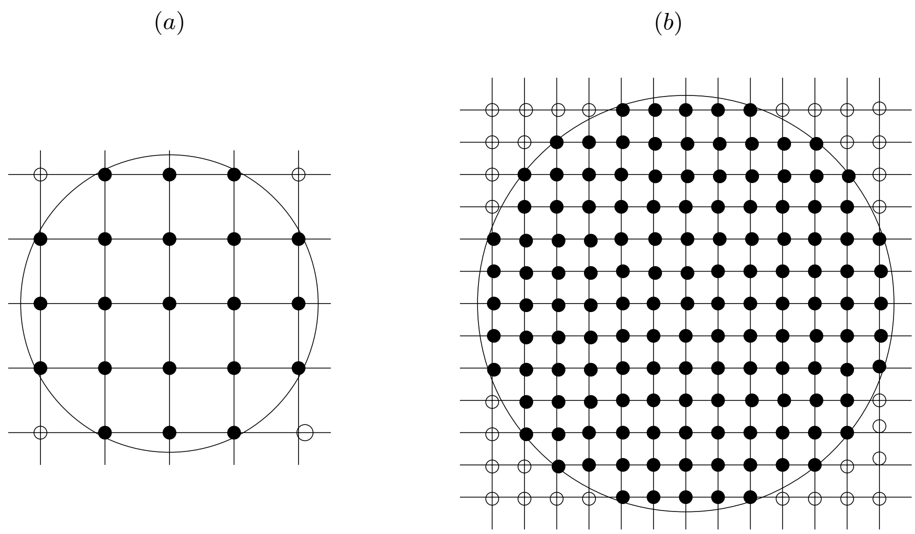

Figure 2. An illustration of the asymptotic structure and the sampling design: (a) sampling region R_{n-1} and a square grid of sampling sites (shown as •) with sample size N_{n-1}; (b) sampling region R_n and a square grid of sampling sites (shown as •) with sample size N_n.

Implicitly, the discussion above about the SCDF concerns a r.f. $Z(\cdot)$ at a given point in time and is restricted to statistical inference for the current status of the process. In order to detect any changes or trends in an ecological resource over time, we consider different r.f.s at various time points, namely $\{Z_t(\boldsymbol{s}) : \boldsymbol{s} \in R_n\}$, where t belongs to an index set of time points and $R_n \subset \mathbb{R}^d$. Associated with the r.f. $Z_t(\cdot)$ we have the SCDF

$$F_{\infty,t}(z) \equiv |R_n|^{-1} \int_{R_n} I(Z_t(\boldsymbol{s}) \leq z) d\boldsymbol{s}, \ z \in \mathbb{R}. \tag{1}$$

Given a finite sample $\{Z_t(\boldsymbol{s}_1), \ldots, Z_t(\boldsymbol{s}_{N_n})\}$ from the r.f. $Z_t(\cdot)$, observed at fixed (over time) sampling sites $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N_n}\}$, the ECDF is

$$F_{n,t}(z) \equiv {N_n}^{-1} \sum_{i=1}^{N_n} I(Z_t(\boldsymbol{s}_i) \leq z), \quad z \in \mathbb{R}, \ t \in \{1, \ldots, T\}, \tag{2}$$

where T is the total number of time points under consideration. In this paper, we extend the statistical methodology for inference on a single SCDF to inference on two SCDFs (with $T = 2$), and we assume that the (joint) spatial random process, $\{\boldsymbol{Z}(\boldsymbol{s}) \equiv (Z_1(\boldsymbol{s}), Z_2(\boldsymbol{s}))' : \boldsymbol{s} \in \mathbb{R}^d\}$, is stationary. Hence the theoretical CDF is location invariant and is denoted

$$G_t(z) \equiv P(Z_t(\mathbf{0}) \leq z), \quad z \in \mathbb{R}, \ t = 1, 2, \tag{3}$$

where $\mathbf{0}$ is the origin of the spatial domain R_n. Henceforth, we call G_t the *invariant CDF* of the r.f. $Z_t(\cdot)$.

The statistical inference for detecting change of the SCDF over time is based on two procedures. The first is a formal test of the null hypothesis $H_o : G_1 = G_2$ versus the alternative hypothesis $H_a : G_1 \neq G_2$, where G_t is the invariant CDF. We use the difference between the two ECDFs, $F_{n,1}(z) - F_{n,2}(z), \ z \in \mathbb{R}$, for our test statistic. We derive the large-sample distribution of the normalized test statistic

$$\xi_n(z) \equiv \lambda_n^{d/2}(F_{n,1}(z) - F_{n,2}(z)), \ z \in \mathbb{R}, \tag{4}$$

and propose a test criterion that asymptotically guarantees a desired significance level.

The second procedure quantifies change by a weighted integrated squared difference (WISD) between the two SCDFs, defined as

$$\mathcal{X}_\infty \equiv \int_{\mathbb{R}} (F_{\infty,1}(z) - F_{\infty,2}(z))^2 \, w(z) dz, \tag{5}$$

where $w(\cdot)$ is a fixed weight function. The predictor for the WISD is the weighted integrated squared distance between the ECDFs, $\mathcal{X}_n \equiv \int_{\mathbb{R}} (F_{n,1}(z) - F_{n,2}(z))^2 \times$

$w(z)dz$. After deriving the large-sample distribution of the normalized and centered predictor, $b_n(\mathcal{X}_n - \mathcal{X}_\infty)$, where the normalizing constant b_n is to be specified, we construct a prediction interval for $\mathcal{X}_\infty$ that asymptotically achieves a desired prediction probability.

Under the mixed asymptotic framework and some fairly general assumptions, we establish weak convergence of the test statistic and the predictor $\mathcal{X}_n$. However, the limiting distributions involve unknown quantities and need to be estimated in order to carry out the test and to compute prediction intervals. For this purpose, we extend the subsampling method developed by Lahiri (1999) to the vector r.f. $\boldsymbol{Z}(\cdot)$. The basic idea is to create several smaller subsampling regions within the sampling region R_n and then to recreate "samples" and "populations" at the level of the subsamples, so that a subsampling estimate for the sampling distributions of the quantities of interest can be obtained.

The smaller subregions in the sampling region R_n are of the form, $\boldsymbol{i}\beta_n + (-1/2, 1/2]^d \lambda_l$, where $\boldsymbol{i} \in \mathbb{Z}^d$ and $l \equiv l_n$ is a sequence of positive integers such that $l \to \infty$ as $n \to \infty$. The constant λ_l determines the size of the subregion, and the constant β_n, $0 < \beta_n \leq \lambda_l$, controls the amount of overlap of the subregions. Note that when $\beta_n = \lambda_l$, the subregions are disjoint. Inside each subregion, we inscribe a subsampling region $\boldsymbol{i}\beta_n + \lambda_l R_0$. As $\boldsymbol{i}$ varies, a collection of subsampling regions is generated that are contained in R_n and are the same shape as R_n. These are denoted $R_{*1}, \ldots, R_{*K_n}$, where K_n is the total number of subregions created. Within each subsampling region, we recreate the effect of "sample" and "population" by imposing a coarser grid $\mathcal{P}_l$ with spacing h_l (i.e., $\mathcal{P}_l \equiv \mathbb{Z}^d h_l$), and a finer grid $\mathcal{P}_n$ with spacing h_n (i.e., $\mathcal{P}_n \equiv \mathbb{Z}^d h_n$) of sampling sites. Recall that h_n is the infill rate of the original sampling grid, which tends to zero. Suppose that $\{h_l\}$ is a sequence of positive real numbers tending to 0 as $n \to \infty$. For simplicity we assume that h_l/h_n is an integer and hence that $\mathcal{P}_l$ is nested in $\mathcal{P}_n$. Figure 3 illustrates the idea of the subsampling design in $\mathbb{R}^2$ on a square grid of sampling sites within the circled sampling region R_n.

Now, a subsampling version of the ECDF $F_{n,t}$ in the subsampling region R_{*i} is

$$F_{n,t}^{*i}(z) \equiv |R_{*i} \cap \mathcal{P}_l|^{-1} \sum_{\boldsymbol{s} \in R_{*i} \cap \mathcal{P}_l} I(Z_t(\boldsymbol{s}) \leq z), \; z \in \mathbb{R}, \tag{6}$$

and a subsampling version of the SCDF $F_{\infty,t}$ in R_{*i} is

$$F_{\infty,t}^{*i}(z) \equiv |R_{*i} \cap \mathcal{P}_n|^{-1} \sum_{\boldsymbol{s} \in R_{*i} \cap \mathcal{P}_n} I(Z_t(\boldsymbol{s}) \leq z), \; z \in \mathbb{R}, \tag{7}$$

where $i = 1, \ldots, K_n$. Let $T_n \equiv T_n(F_{n,1}, F_{\infty,1}, F_{n,2}, F_{\infty,2})$ denote a quantity involving $F_{n,t}$ and $F_{\infty,t}$, $t = 1, 2$. Then the sampling distribution of T_n, namely

$H_n(z) \equiv P(T_n \le z),\ z \in I\!R$, can be estimated by the empirical distribution of the subsampling versions, $T_l^{*i} \equiv T_l(F_{n,1}^{*i}, F_{\infty,1}^{*i}, F_{n,2}^{*i}, F_{\infty,2}^{*i})$, $i = 1, \ldots, K_n$:

$$\hat{H}_n(z) \equiv K_n^{-1} \sum_{i=1}^{K_n} I(T_l^{*i} \le z), \quad z \in I\!R. \tag{8}$$

One of the main contributions of this paper is to develop valid inferences for $T_n(F_{n,1}, F_{\infty,1}, F_{n,2}, F_{\infty,2})$ based on subsampling versions as defined in (6) and (7). This extends the work of Lahiri et al. (1999) to the important problem of multiple time points.

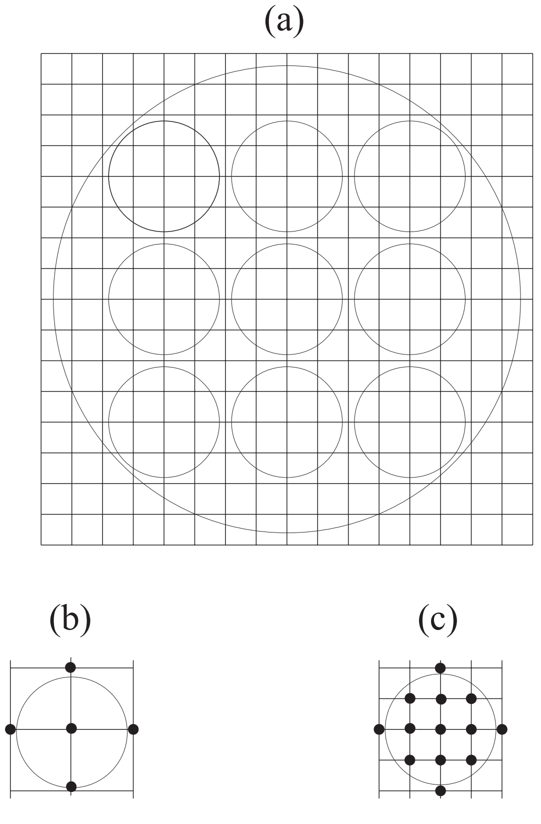

Figure 3. An illustration of the subsampling method: (a) all subsampling regions inscribed in R_n; (b) a coarser grid of subsampling sites (shown as $\bullet$) in one subsampling region; (c) a finer grid of subsampling sites (shown as $\bullet$) in the same subsampling region.

The rest of the paper is organized as follows. In Section 2, we develop an asymptotic test procedure for testing the null hypothesis $H_o : G_1 = G_2$ against the alternative hypothesis $H_o : G_1 \neq G_2$, based on the process defined in

(4). In Section 3, we construct an asymptotic prediction interval for the WISD $\mathcal{X}_\infty$ defined by (5), based on $\mathcal{X}_n$. As an illustration, the inference procedures developed in Sections 2 and 3 are applied to a forest-health-monitoring data set in Section 4. Conclusions are given in Section 5, and the proofs of the theorems are in Section 6.

2. Detection of Changes over Time

The SCDF is a basic (random) functional whose properties can be used to indicate the status of ecological resources. Moreover, comparisons of SCDFs at different points in time over the same region can be used to indicate resource improvement or resource decline over time. Under a spatial stationarity assumption of $Z_t(\cdot)$, we have $E(F_{\infty,t}(z)) = G_t(z)$, where G_t is the invariant CDF of $Z_t(\cdot)$. Then, for two time points, a comparison of two SCDFs could be considered as a test statistic for testing $H_o : G_1 = G_2$ versus $H_a : G_1 \neq G_2$.

Since the spatial sampling sites in ecological monitoring programs are oftentimes fixed, it is reasonable to assume that the finite samples, $\{Z_1(\boldsymbol{s}_1), \ldots, Z_1(\boldsymbol{s}_{N_n})\}$ and $\{Z_2(\boldsymbol{s}_1), \ldots, Z_2(\boldsymbol{s}_{N_n})\}$, are observed at the same set of sampling sites $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N_n}\}$. Recall that the test statistic is based on the normalized ECDF difference, $\xi_n(z) \equiv \lambda_n^{d/2}(F_{n,1}(z) - F_{n,2}(z))$, defined in (4). The subsampling version of ξ_n is $\xi_l^{*i}(z) \equiv \lambda_l^{d/2}\left((F_{n,1}^{*i}(z) - F_{n,2}^{*i}(z)\right)$, $z \in I\!R$, $i = 1, \ldots, K_n$, where recall that $F_{n,t}^{*i}$ is the subsampling version of $F_{n,t}$, λ_l is the growth rate of the subsampling regions, l is a sequence of positive integers such that $\lambda_l \to \infty$, $l \to \infty$ as $n \to \infty$, and K_n is the total number of subsampling regions in R_n.

To establish the limit distribution of $\xi_n(\cdot)$ and the validity of subsampling under the null hypothesis, we use a ρ-mixing condition to specify the dependence structure of the r.f. $\boldsymbol{Z}(\cdot)$. Let $\mathcal{L}_2'(S)$ denote the collection of all random variables with zero mean and finite second moments that are measurable with respect to the σ-field generated by $\{\boldsymbol{Z}(\boldsymbol{s}) : \boldsymbol{s} \in S\}$, $S \subset I\!R^d$. For $S_1, S_2 \subset I\!R^d$, let $\rho_1(S_1, S_2) \equiv \sup\{|E(\zeta\eta)|/\left((E(\zeta^2))^{1/2}(E(\eta^2))^{1/2}\right) : \zeta \in \mathcal{L}_2'(S_1), \eta \in \mathcal{L}_2'(S_2)\}$. Then the ρ-mixing coefficient of the r.f. $\boldsymbol{Z}(\cdot)$ is (Doukhan (1994), Section 1.1) $\rho(k; m) \equiv \sup\{\rho_1(S_1, S_2) : |S_1| \leq m, |S_2| \leq m, \delta(S_1, S_2) \geq k\}$, where $\delta(S_1, S_2) \equiv \inf\{|\boldsymbol{x} - \boldsymbol{s}| : \boldsymbol{x} \in S_1, \boldsymbol{s} \in S_2\}$, and $|\cdot|$ is the ℓ^1-norm. We assume the following.

(A.1) There exist positive real numbers C, τ, θ with $\tau > 3d$, $\theta d < \tau$ such that $\rho(k; m) \leq Ck^{-\tau}m^{\theta}$.

(A.2) The r.f. $\boldsymbol{Z}(\cdot)$ is stationary and the invariant CDF G_1 and G_2 are continuous.

(A.3) For any sequence of positive real numbers $\{a_n\}$ tending to 0 as $n \to \infty$, let R_0 satisfy the condition that the number of cubes of the lattice $a_n\mathbb{Z}^d$ that intersect both R_0 and its complement R_0^c is of the order $(a_n^{-1})^{d-1}$ as $n \to \infty$.

Assumption (A.1) requires the r.f. $\boldsymbol{Z}(\cdot)$ to be short-range dependent. Intuitively, this requires the variables $\boldsymbol{Z}(\boldsymbol{s})$ and $\boldsymbol{Z}(\boldsymbol{x})$ to be approximately independent when the distance between $\boldsymbol{s}$ and $\boldsymbol{x}$ is large. The actual rate of decay is chosen for the convenience of theoretical justification and has been previously used in similar problems (see, e.g., Doukhan (1994); Lahiri (1999)). Assumption (A.2) is a smoothness condition on the invariant CDFs. Note that we do not require G_1 and G_2 to have any specific parametric forms. Finally, assumption (A.3) requires that the number of the sampling sites on the boundary of R_n is negligible compared to the totality of all sampling sites in R_n, ensuring that the edge-effect is negligible in the limit. This condition is satisfied by most sampling regions of practical interest.

The first main result of this section is a functional central limit theorem (FCLT) for the difference process ξ_n in (4), considered as a random element of the space of all real-valued functions on $[-\infty, \infty]$ that are right continuous with left-hand limits, denoted by $D[-\infty, \infty]$ and equipped with the Skorohod metric.

Theorem 2.1. *Under the mixed asymptotic framework described in Section* 1*, suppose that* (*A*.1)−(*A*.3) *and* $H_o : G_1 = G_2$ *hold. Then, as* $n \to \infty$, $\xi_n \stackrel{\mathcal{D}}{\to} \mathcal{W}$, *where* $\mathcal{W}(\cdot)$ *is a zero-mean Gaussian process with covariance function* $\sigma(z_1, z_2) \equiv |R_0|^{-1} \int_{I\!R^d} [G_{11}(z_1, z_2; \boldsymbol{s}) - G_{12}(z_1, z_2; \boldsymbol{s}) - G_{12}(z_2, z_1; -\boldsymbol{s}) + G_{22}(z_1, z_2; \boldsymbol{s})] d\boldsymbol{s}$, $G_{ij}(z_1, z_2; \boldsymbol{s}) \equiv P(Z_i(\boldsymbol{0}) \le z_1, Z_j(\boldsymbol{s}) \le z_2)$ *denotes the joint bivariate distributions of* $\boldsymbol{Z}(\cdot)$, $i, j \in \{1, 2\}$, $z_1, z_2 \in I\!R$, *and* $\boldsymbol{s} \in I\!R^d$. *Moreover,* $\mathcal{W}(\cdot)$ *has continuous sample paths a.s. and* $\mathcal{W}(-\infty) = \mathcal{W}(\infty) = 0$ *a.s.*

Now, for a Borel-measurable function $g : I\!R \to I\!R$, we define the weighted ℓ^p-norm of g by

$$\|g\|_p \equiv \begin{cases} (\int_{I\!R} |g(z)|^p u(z) dz)^{1/p} & ; \quad p \in [1, \infty), \\ \sup\{|g(z)| u(z) : z \in I\!R\} & ; \quad p = \infty, \end{cases}$$

where $u(\cdot)$ is a nonnegative weight function. We use this form of the norm to quantify the magnitude of ξ_n. By the Continuous Mapping Theorem (see, e.g., Theorem 4.2.12 in Pollard (1984)), the null asymptotic distribution of $\|\xi_n\|_p$ can be determined as an immediate consequence of Theorem 2.1, as follows.

Corollary 2.1. *Under the same conditions as in Theorem* 2.1*, as* $n \to \infty$, $\|\xi_n\|_p \stackrel{\mathcal{D}}{\to} \|\mathcal{W}\|_p$, $p \in [1, \infty]$.

Let $H(\cdot; p)$ denote the CDF of $\|\mathcal{W}\|_p$. Let $H_n(\cdot; p)$ denote the CDF of $\|\xi_n\|_p$, and let the subsampling estimator of $H_n(\cdot; p)$ be denoted as $\hat{H}_n(z; p) \equiv K_n^{-1} \sum_{i=1}^{K_n} I(\|\xi_l^{*i}\|_p \le z)$, $z \in I\!R$.

Theorem 2.2. *Suppose the conditions in Theorem* 2.1 *hold. Furthermore, assume the subsampling design given in Section* 1*, and suppose that* $H(\cdot; p)$ *is con-*

tinuous and that $\lambda_l/\lambda_n \to 0$ *as* $n \to \infty$. *Then, for any* $p \in [1, \infty]$ *as* $n \to \infty$, $\sup_{z \in I\!R} |\hat{H}_n(z;p) - H_n(z;p)| \xrightarrow{\mathcal{P}} 0$.

Let $q_\alpha(p)$ denote the α-th quantile of $\|\mathcal{W}\|_p$, $q_\alpha(p) \equiv \inf\{z : H(z;p) \geq \alpha\}$, where $0 < \alpha < 1$. Because of Corollary 2.1. and $P(\|\mathcal{W}\|_p > q_{1-\alpha}(p)) = \alpha$, the test that rejects the null hypothesis H_o if $\|\xi_n\|_p > q_{1-\alpha}(p)$ has an asymptotic significance level of α. However, the quantiles of $\|\mathcal{W}\|_p$ depend on the joint bivariate distribution functions and are not known in practice. Let $\hat{q}_\alpha(p)$ denote the α-th quantile of $\hat{H}_n(z;p)$, $\hat{q}_\alpha(p) \equiv \inf\{z : \hat{H}_n(z;p) \geq \alpha\}$, which is the $[K_n\alpha]$-th order statistic of $\|\xi_l^{*1}\|_p, \ldots, \|\xi_l^{*K_n}\|_p$. Then the approximate level-α hypothesis test, based on the subsampling method, is to reject the null hypothesis H_o if $\|\xi_n\|_p > \hat{q}_{1-\alpha}(p)$. By Theorem 2.2, under H_o, $P(\|\xi_n\|_p > \hat{q}_{1-\alpha}(p)) \to \alpha$ as $n \to \infty$. Hence, this test has an asymptotic significance level of α.

3. Prediction of Changes over Time

In Section 2 we developed a test to detect changes in the SCDFs for a given region at two different time points. In this section, we develop a statistical method to *quantify* changes by a weighted integrated squared difference (WISD) between the two SCDFs, as defined in (5). The WISD is an unknown random quantity whose realizations measure the discrepancy between sample paths of the two SCDFs. Further, for fixed $z_1, z_2 \in I\!R$, if the weight function $w(\cdot)$ is the indicator function on (z_1, z_2), then the WISD is a measure of the SCDF discrepancy for z values ranging between z_1 and z_2. Hence the WISD can be used to assess the differences of the SCDFs for z in different regions of $I\!R$, for instance when z takes values that are below or above certain cut-off levels.

Based on the finite samples $\{Z_1(\boldsymbol{s}_1), \ldots, Z_1(\boldsymbol{s}_{N_n})\}$ and $\{Z_2(\boldsymbol{s}_1), \ldots, Z_2(\boldsymbol{s}_{N_n})\}$, collected at sampling sites $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N_n}\}$, a reasonable predictor for $\mathcal{X}_\infty$ is the finite-sample version $\mathcal{X}_n$. Recall from the asymptotic structure and the sampling design that the center of the sampling grid is $\boldsymbol{c}h_n$, where $\boldsymbol{c} \in (0,1]^d$. As it turns out, in the non-centered design as in Lahiri et al. (1999), an appropriate normalizing constant for the predictor $\mathcal{X}_n$ is

$$b_n \equiv \begin{cases} \lambda_n^{d/2} h_n^{-1} & ; \quad G_1 \neq G_2, \\ \lambda_n^d h_n^{-1} & ; \quad G_1 = G_2. \end{cases} \tag{9}$$

That is, the limiting distribution of $\mathcal{X}_n$ when the two invariant CDFs coincide is different from that of $\mathcal{X}_n$ when the two invariant CDFs differ. We have a statistical test for the null hypothesis $H_o : G_1 = G_2$. If the data provide strong evidence that there are differences in the SCDFs over time, we can use the asymptotic distribution of $\mathcal{X}_n$ for $G_1 \neq G_2$ to quantify these differences. Hence, we have a procedure that first detects, and then quantifies the changes over time. On the other hand, if the data do not show strong evidence against the null hypothesis, it is still of interest to predict $\mathcal{X}_\infty$.

In this section, we establish the limit distribution of the centered and scaled WISD:

$$\mathcal{Y}_n \equiv b_n(\mathcal{X}_n - \mathcal{X}_\infty) = \int_{I\!R} b_n \left[(F_{n,1}(z) - F_{n,2}(z))^2 - (F_{\infty,1}(z) - F_{\infty,2}(z))^2\right] w(z)dz, \quad (10)$$

where the normalizing constant b_n is defined in (9). Using the subsampling described in Section 1, we define the subsampling versions of $\mathcal{X}_n$ and $\mathcal{X}_\infty$ as $\mathcal{X}_n^{*i} \equiv \int_{I\!R}(F_{n,1}^{*i}(z) - F_{n,2}^{*i}(z))^2 w(z)dz$ and $\mathcal{X}_\infty^{*i} \equiv \int_{I\!R}(F_{\infty,1}^{*i}(z) - F_{\infty,2}^{*i}(z))^2 w(z)dz$, where $F_{n,t}^{*i}$ and $F_{\infty,t}^{*i}$ are defined in (6) and (7), $i = 1, \ldots, K_n$. Then we establish the asymptotic distribution for the subsampling versions of $\mathcal{Y}_n$, namely $\mathcal{Y}_l^{*i} \equiv b_l(\mathcal{X}_n^{*i} - \mathcal{X}_\infty^{*i})$, $i = 1, \ldots, K_n$, where the normalizing constant for $\mathcal{Y}_l^{*i}$ is $b_l \equiv \lambda_l^{d/2} h_l^{-1}$ if $G_1 \neq G_2$, and $b_l \equiv \lambda_l^d h_l^{-1}$ if $G_1 = G_2$.

Let $Z\!\!Z^+$ denote the set of all nonnegative integers. For a vector $\boldsymbol{x} = (x_1, \ldots, x_d)' \in I\!R^d$, $d \geq 1$, let $|\boldsymbol{x}| \equiv \sum_{i=1}^d |x_i|$ and $\|\boldsymbol{x}\| \equiv \sum_{i=1}^d x_i^2$ denote the ℓ^1- and ℓ^2-norms of $\boldsymbol{x}$. For $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)' \in (Z\!\!Z^+)^d$, write $\boldsymbol{x}^{\boldsymbol{\alpha}} \equiv \prod_{i=1}^d x_i^{\alpha_i}$, $\boldsymbol{\alpha}! \equiv \prod_{i=1}^d \alpha_i!$, and let $D^{\boldsymbol{\alpha}}$ denote the differential operator $D_1^{\alpha_1} \ldots D_d^{\alpha_d}$ on $I\!R^d$, where $D_l^{\alpha_l} \equiv \partial^{\alpha_l}/\partial x_l^{\alpha_l}$ for $1 \leq l \leq d$. The following assumptions are made.

(B.1) For given $z_1, z_2 \in I\!R$, $i, j \in \{1, 2\}$,
   (i) $G_{ij}(z_1, z_2; \cdot)$ has bounded, Lebesgue-integrable partial derivatives of order 2 on $I\!R^d$;
   (ii) for $|\boldsymbol{\alpha}| = 2$,there exist nonnegative integrable functions $\mathcal{H}_{\boldsymbol{\alpha},ij}(z_1, z_2; \cdot)$ such that for all $\boldsymbol{x}, \boldsymbol{s} \in I\!R^d$ with $\|\boldsymbol{x}\| \leq 1$, $|D^{\boldsymbol{\alpha}} G_{ij}(z_1, z_2; \boldsymbol{s} + \boldsymbol{x}) - D^{\boldsymbol{\alpha}} G_{ij}(z_1, z_2; \boldsymbol{s})| \leq \|\boldsymbol{x}\|^\eta \mathcal{H}_{\boldsymbol{\alpha},ij}(z_1, z_2; \boldsymbol{s})$ for some $\eta > 0$.

(B.2) There exist constants $C > 0, 1/2 < \gamma \leq 1$, such that $\sum_{|\boldsymbol{\alpha}|=2} |D^{\boldsymbol{\alpha}} \mathcal{G}_{tt}(z_1, z_2 ; \boldsymbol{s})| \leq C|G_t(z_2) - G_t(z_1)|^\gamma$, where $\mathcal{G}_{tt}(z_1, z_2; \boldsymbol{s}) \equiv P(z_1 < Z_t(\boldsymbol{0}) \leq z_2, z_1 < Z_t(\boldsymbol{s}) \leq z_2)$ for all $z_1, z_2 \in I\!R$, $\boldsymbol{s} \in I\!R^d$, $t = 1, 2$.

(B.3) $(h_n^{2(2\gamma+1)} \lambda_n^d)^{-1} + (h_n \lambda_n / \log \lambda_n)^{-1} \to 0$ as $n \to \infty$, γ as in (B.2).

(B.4) $\int_{I\!R} |G_1(z) - G_2(z)| \varsigma(z, z)^{1/2} w(z)dz < \infty$ and $\int_{I\!R} [\sigma(z, z)\varsigma(z, z)]^{1/2} w(z)dz < \infty$, where $\sigma(\cdot, \cdot)$ is defined in Theorem 2.1 and $\varsigma(z_1, z_2) \equiv |R_0|^{-1} \sum_{|\boldsymbol{\alpha}|=2} a(\boldsymbol{\alpha})(\boldsymbol{\alpha}!)^{-1} \int_{I\!R^d} [D^{\boldsymbol{\alpha}} G_{11}(z_1, z_2; \boldsymbol{s}) - D^{\boldsymbol{\alpha}} G_{12}(z_1, z_2; \boldsymbol{s}) - D^{\boldsymbol{\alpha}} G_{12}(z_2, z_1; \boldsymbol{s}) + D^{\boldsymbol{\alpha}} G_{22}(z_1, z_2; \boldsymbol{s})] d\boldsymbol{s}$, with z_1, $z_2 \in I\!R$, and $a(\boldsymbol{\alpha}) \equiv \int_{(0,1]^d} \int_{(0,1]^d} [(\boldsymbol{x} - \boldsymbol{s})^{\boldsymbol{\alpha}} - (\boldsymbol{x} - \boldsymbol{c})^{\boldsymbol{\alpha}} - (\boldsymbol{c} - \boldsymbol{s})^{\boldsymbol{\alpha}}] d\boldsymbol{s} d\boldsymbol{x}$.

The assumptions in (B.1) are smoothness conditions for the joint bivariate distributions, viewed as functions of $\boldsymbol{s}$ at fixed z_1 and z_2. They are almost minimal for proving FCLTs. Assumption (B.2) is needed in proving the tightness of a process as a random element of the space $D[-\infty, \infty]$ and is a type of Lipschitz condition. Assumption (B.3) specifies the relationship between the growth rate

λ_n and the infill rate h_n in the asymptotic framework. Finally, assumption (B.4) is a condition needed to ensure weak convergence after the integration in (10).

The first main result for this section establishes weak convergence of $\mathcal{Y}_n$ as follows.

Theorem 3.1. *Assume the mixed asymptotic framework described in Section* 1 *and assume that* (*A*.1)–(*A*.3)*,* (*B*.1)–(*B*.4) *hold. Then, as* $n \to \infty$*,*

$$\mathcal{Y}_n = b_n(\mathcal{X}_n - \mathcal{X}_\infty) \xrightarrow{\mathcal{D}} \mathcal{Y}_\infty \equiv \begin{cases} \int_{I\!R} 2(G_1(z) - G_2(z))\mathcal{V}(z)w(z)dz & ; \quad G_1 \neq G_2, \\ \int_{I\!R} 2\mathcal{W}(z)\mathcal{V}(z)w(z)dz & ; \quad G_1 = G_2, \end{cases}$$

where b_n *is given in* (9)*. Here* $(\mathcal{W}(\cdot), \mathcal{V}(\cdot))'$ *is a vector Gaussian process with mean* $(0,0)'$*. The covariance function for* $\mathcal{V}(\cdot)$ *is* ς *defined in* (*B*.4)*, that for* $\mathcal{W}(\cdot)$ *is* σ *defined in Theorem* 2.1*, and the cross-covariance function is*

$$\upsilon(z_1, z_2) \equiv |R_0|^{-1} \sum_{|\boldsymbol{\alpha}|=1} a^*(\boldsymbol{\alpha})(\boldsymbol{\alpha}!)^{-1} \int_{I\!R^d} D^{\boldsymbol{\alpha}} G_{11}(z_1, z_2; \boldsymbol{s}) - D^{\boldsymbol{\alpha}} G_{12}(z_1, z_2; \boldsymbol{s}) - D^{\boldsymbol{\alpha}} G_{12}(z_2, z_1; \boldsymbol{s}) + D^{\boldsymbol{\alpha}} G_{22}(z_1, z_2; \boldsymbol{s}) d\boldsymbol{s}, \tag{11}$$

where $a^*(\boldsymbol{\alpha}) \equiv \int_{(0,1]^d} (\boldsymbol{s} - \boldsymbol{c})^{\boldsymbol{\alpha}} d\boldsymbol{s}$*.*

When the two invariant CDFs are not the same, the limiting random variable $\mathcal{Y}_\infty$ is a weighted product of the invariant CDF difference and the Gaussian process $\mathcal{V}(\cdot)$. Otherwise, $\mathcal{Y}_\infty$ is a weighted product of the two Gaussian processes $\mathcal{W}(\cdot)$ and $\mathcal{V}(\cdot)$. The normalizing constant b_n is the same as that given by Lahiri (1999) when $G_1 \neq G_2$, but it is scaled up by a factor of $\lambda_n^{d/2}$ when $G_1 = G_2$.

We denote the CDF of $\mathcal{Y}_\infty$ by H, the CDF of $\mathcal{Y}_n$ by H_n, and recall that the subsampling estimator of $H_n(\cdot)$ is $\hat{H}_n(z) \equiv K_n^{-1} \sum_{i=1}^{K_n} I(\mathcal{Y}_l^{*i} \leq z)$, $z \in I\!R$. Note that the CDFs $H(\cdot)$, $H_n(\cdot)$ and $\hat{H}_n(\cdot)$ depend on whether $G_1 = G_2$.

Theorem 3.2. *Suppose that the conditions in Theorem* 3.1 *hold. Assume the subsampling design given in Section* 1*, and suppose that* $H(\cdot)$ *is continuous, that* $\lambda_l/\lambda_n \to 0$*,* $h_n/h_l \to 0$ *as* $n \to \infty$*, and that* (*B*.3) *holds with* λ_n *replaced by* λ_l*. Then, as* $n \to \infty$*,* $\sup_{z \in I\!R} |\hat{H}_n(z) - H_n(z)| \xrightarrow{\mathcal{P}} 0$*.*

Let π_α denote the α-th quantile of the random variable $\mathcal{Y}_\infty$. By Theorem 3.1, the prediction interval $I_{1-\alpha} \equiv \{X : \pi_{\alpha/2} < b_n(\mathcal{X}_n - X) < \pi_{1-\alpha/2}\}$ attains a prediction probability of $1 - \alpha$, asymptotically. Because the asymptotic distributions depend on population parameters that are not known in practice, we use the subsampling method to estimate $\pi_{\alpha/2}$ and $\pi_{1-\alpha/2}$. Under the conditions of Theorem 3.2, the subsampling estimator $\hat{H}_n(\cdot)$ of the sampling distribution $H_n(\cdot)$ of $\mathcal{Y}_n$ yields asymptotically valid prediction intervals for $\mathcal{Y}_\infty$. Specifically, let $\hat{\pi}_\alpha$ denote the α-th quantile of $\hat{H}_n(\cdot)$, which is the $[K_n\alpha]$-th order statistic of

$\mathcal{Y}_l^{*1}, \ldots, \mathcal{Y}_l^{*K_n}$. Then $P(X : \hat{\pi}_{\alpha/2} < b_n(\mathcal{X}_n - X) < \hat{\pi}_{1-\alpha/2}) \to (1-\alpha)$ as $n \to \infty$. That is, the interval $\hat{I}_{1-\alpha} \equiv (\mathcal{X}_n - \hat{\pi}_{1-\alpha/2}/b_n, \mathcal{X}_n - \hat{\pi}_{\alpha/2}/b_n)$ is an asymptotically valid $100(1-\alpha)\%$ prediction interval.

4. An Example

In the early 1990s, the United States Environmental Protection Agency and the United States Forest Service conducted an annual forest-health monitoring program in the New England states. A selected indicator of forest health is the crown defoliation index (CDI) of red-maple trees − it has a range of $[0, 100]$ and is constructed from measures of visible injury to the tree crowns, adjusted by the sizes of the trees. Estimation and prediction of the current status and the changes over time of the CDI values can be used to assess the overall forest health and to detect change.

Lahiri et al. (1999) analyze CDI values for red-maple trees in the state of Maine at one time point ($t = 1992$) by constructing simultaneous prediction intervals for the SCDF. The spatial domain of interest $R \subset I\!R^2$ is the state of Maine. The 77 sampling sites where actual observations were made form an incomplete hexagonal grid as in Figure 4(a). The hexagonal sampling grid has a spacing of 27 km, which we set equal to one unit $= h_n$. The total size of the grid is determined by the growth rate $\lambda_n = 10h_n = 10$. At each sampling site, measurements of visible injury were made on individual tree crowns, and then a weighted average was taken over the number of trees within a given sampling site. Hence, each datum of CDI characterizes the corresponding sampling site as a spatial support unit, and can be modeled by a r.f. with continuous spatial index (Lahiri et al. (1999)).

In order to apply the sampling and subsampling procedure described in Section 1, the sampling grid is first completed with imputed observations (Figure 4(b)) as follows. Consider a sampling site $\boldsymbol{s}_0$ at the center of a complete hexagonal grid $h(\boldsymbol{s}_0) \equiv \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_6\}$ (Figure 4(b)), where there is no observed value at $\boldsymbol{s}_0$; the "most similar" hexagon of $\boldsymbol{s}_0$, say the hexagon $h(\boldsymbol{s}_0^*) \equiv \{\boldsymbol{s}_1^*, \ldots, \boldsymbol{s}_6^*\}$ centered at $\boldsymbol{s}_0^*$, is selected and the observed value at site $\boldsymbol{s}_0^*$ is used as the imputed value for site $\boldsymbol{s}_0$. The "similarity" between $h(\boldsymbol{s}_0)$ and $h(\boldsymbol{s}_0^*)$ is measured by the Euclidean distance between the observed CDI vectors $(z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_6))'$ and $(z_t(\boldsymbol{s}_1^*), \ldots, z_t(\boldsymbol{s}_6^*))'$. This procedure is aimed at preserving the spatial structure of the CDI process. After the imputation, no distinction is made between the imputed values and the observed data in the analysis. Next, sampling sites in the hexagonal grid (Figure 4(b)) are transformed to a square grid (Figure 4(c)) through a nonsingular linear transformation. Because of the one-to-one relationship between the hexagonal and square grids of sampling sites, the analysis of the data on the square grid is equivalent to that on the hexagonal grid. More details can be found in Lahiri et al. (1999).

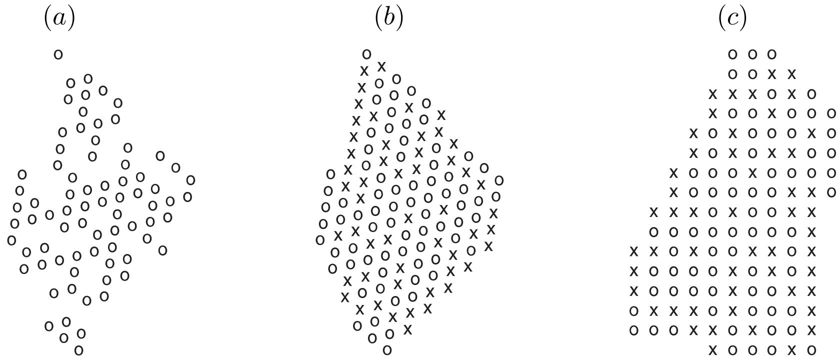

Figure 4. Locations in Maine for sampled data (shown as $\circ$) and imputed data (shown as $\times$): (a) the incomplete hexagonal grid; (b) the hexagonal grid completed with imputed data locations; (c) the transformed square grid.

Now we apply the statistical methods developed in Section 2 and 3 to compare foliage conditions in the state of Maine at three different time points ($t = 1991, 1992, 1993$). We conduct hypothesis testing and construct prediction intervals for the WISD of the CDI values of red-maple trees, using the complete square grid of the actual and imputed data. Figure 5 shows smoothed CDI surfaces (using S-Plus functions `persp` and `interp`) over the square sampling grids, based on the observed and imputed values, and Figure 1 superimposes the ECDF curves from the three years. There seem to be no major differences in the CDI values over these years.

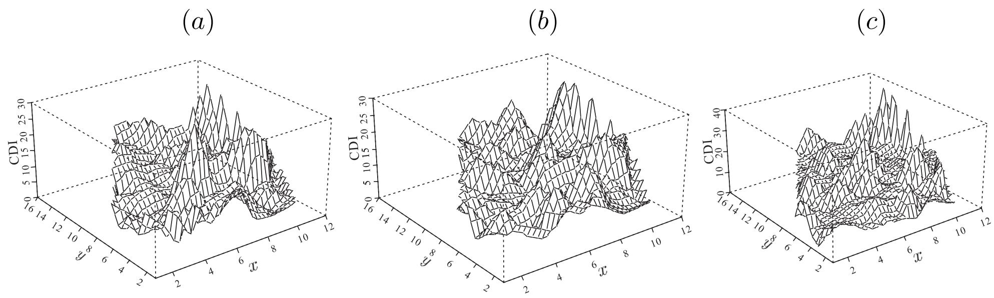

Figure 5. Smoothed surfaces of the crown defoliation index (CDI) values over the square grid of sampling sites in different years: (a) 1991; (b) 1992; (c) 1993.

For the subsampling design, the subsampling grid is defined to have spacing h_l and the size of each subsampling region is determined by λ_l. Each of the resulting K_n subsampling regions has $(\lambda_l/h_n)^2$ points for use in computing $F^{*i}_{\infty,t}$ in (7), and $(\lambda_l/h_l)^2$ points for use in computing $F^{*i}_{n,t}$ in (6). Let L_i denote the set of indices for the $(\lambda_l/h_l)^2$ sampling sites and N_i denote the set of indices for the $(\lambda_l/h_n)^2$ sampling sites, each in the i-th subsampling region R_{*i}, $i = 1,\ldots,K_n$. Then the subsampling versions of the ECDF $F_{n,t}$ and the SCDF $F_{\infty,t}$ are $F^{*i}_{n,t}(z) \equiv |L_i|^{-1}\sum_{j\in L_i} I(Z_t(\boldsymbol{s}_j) \le z)$ and $F^{*i}_{\infty,t}(z) \equiv |N_i|^{-1}\sum_{j\in N_i} I(Z_t(\boldsymbol{s}_j) \le z)$, $z \in I\!R$, as in (6) and (7). Here we consider all possible combinations of λ_l and h_l. Because of the sampling configuration, as shown in Figure 4, the only choices that would result in $K_n > 2$ subsamples are $\lambda_l = 4$ or 6, and $h_l = 2$ or 3. Hence we present the results for three cases: $\lambda_l = 4, h_l = 2$; $\lambda_l = 6, h_l = 2$; and $\lambda_l = 6, h_l = 3$. Moreover, to account for multiple comparisons, we use a Bonferroni correction. That is, we perform the test for all pairs at 1/3 of the given significance level to ensure an overall level that is $\le \alpha$.

Based on the method described in Section 2, we conduct a simultaneous test of $H_o : G_1 = G_2$ versus $H_a : G_1 \ne G_2$, for possible pairs of the three years, at significance level $\alpha = 0.1$. The normalizing constant in ξ_n is $\lambda_n = 10$, hence $\xi_n = 10(F_{n,1} - F_{n,2})$. We use the ℓ^2- and ℓ^∞-norms and equal weights for the test statistic $\|\xi_n\|_p$, that is, $p = 2$ or ∞, and $u(z) = 1$ for all $z \in I\!R$. The scaling constant λ_l is found in $\xi^{*i}_l = \lambda_l(F^{*i}_{n,1} - F^{*i}_{n,2})$, $i = 1,\ldots,K_n$. We use $\lambda_l = 4$ or 6, resulting in $K_n = 56$ or 21. For an overall significance level α, by the Bonferroni correction, the critical value for each pairwise comparison is the quantile at the $1-\alpha/3$ level and is estimated by $\hat{q}_{1-\alpha/3}(p)$, the $[K_n(1-\alpha/3)]$-th order statistic of $\|\xi^{*i}_l\|_p$. Table 1 shows the observed test statistics, $\|\xi_n\|_p$, and the critical values, $\hat{q}_{1-\alpha/3}(p)$, for rejection. For both $p = 2$ and ∞, and both $\lambda_l = 4$ and 6, none of the observed values exceed the corresponding critical value. Hence we fail to reject the null hypothesis and conclude that there is no strong evidence that foliage conditions have changed from 1991 to 1993.

Table 1. Results of the hypothesis test $H_o : G_1 = G_2$, at significance level $\alpha = 0.1$.

λ_l	years	$\|\xi_n\|_2$	$\hat{q}_{0.97}(2)$	$\|\xi_n\|_\infty$	$\hat{q}_{0.97}(\infty)$
4	'91 vs '92	0.16	0.60	0.13	0.38
	'91 vs '93	0.41	0.97	0.26	0.56
	'92 vs '93	0.38	1.10	0.19	0.56
6	'91 vs '92	0.16	0.38	0.13	0.22
	'91 vs '93	0.41	0.78	0.26	0.47
	'92 vs '93	0.38	0.77	0.19	0.39

We construct 90% prediction intervals for the WISD $\mathcal{X}_\infty$, comparing the foliage conditions in the years 1991-1993, pairwise. The normalizing constant in

$\mathcal{Y}_n$ is $b_n = \lambda_n/h_n = 10$ if $G_1 \neq G_2$, and $b_n = \lambda_n^2/h_n = 100$ if $G_1 = G_2$. The scaling constant in $\mathcal{Y}_l^{*i}$ is $b_l = \lambda_l/h_l$ if $G_1 \neq G_2$, and $b_l = \lambda_l^2/h_l$ if $G_1 = G_2$. Because there is no strong evidence that the invariant CDFs have changed over time, we might predict as if $G_1 = G_2$ applies here. Hence $\mathcal{Y}_n = 100(\mathcal{X}_n - \mathcal{X}_\infty)$ and $\mathcal{Y}_l^{*i} = (\lambda_l^2/h_l)(\mathcal{X}_n^{*i} - \mathcal{X}_\infty^{*i})$, $i = 1, \ldots, K_n$. Two types of weight functions are considered: one with equal weights $w(z) = 1$ for all z, and the other with a point mass at z_0. In the latter case, the WISD is the squared difference of the two SCDFs evaluated at the cut-off level z_0. Following Lahiri et al. (1999), a CDI value below 12.5 indicates good tree health and one above 12.5 indicates poor tree health. Note that the results in Section 3 hold for a point-mass weight, as a direct consequence of the weak convergence of the integrand in (10) evaluated at z_0. We use $\hat{\pi}_\alpha$, the $[K_n\alpha]$-th order statistic among $\mathcal{Y}_l^{*i}$s, to estimate the α-th quantile of $\mathcal{Y}_\infty$. Again we use $\lambda_l = 4$ or 6, giving $K_n = 56$ or 21. In addition, we set $h_l = 2$ for $\lambda_l = 4$ and $h_l = 2$ or 3 for $\lambda_l = 6$. Table 2 summarizes the 90% simultaneous prediction interval $\hat{I}_{0.9}$ of the WISD $\mathcal{X}_\infty$.

From the statistical comparisons between the SCDFs given in Table 2, we conclude that there are no significant changes in red-maple forest health in Maine from 1991 to 1993. This coincides with our informal assessment from observing the surface plots in Figure 5 and the ECDF plots in Figure 1.

Table 2. 90% prediction intervals for $\mathcal{X}_\infty$, the weighted integrated square distance (WISD), with equal weights $w(z) = 1$ for all z, and with point-mass weight $w(z) = I(z = 12.5)$.

λ_l	h_l	years	$w(z) = 1;\ \forall z$	$w(z) = I(z = 12.5)$
4	2	’91 vs ’92	[0, 0.023)	[0, 0.0051)
		’91 vs ’93	[0, 0.180)	[0, 0.0018)
		’92 vs ’93	[0, 0.170)	[0, 0.0043)
6	2	’91 vs ’92	[0, 0.014)	[0, 0.0051)
		’91 vs ’93	[0, 0.190)	[0, 0.0018)
		’92 vs ’93	[0, 0.180)	[0, 0.0028)
6	3	’91 vs ’92	[0, 0.007)	[0, 0.0047)
		’91 vs ’93	[0, 0.170)	[0, 0.0029)
		’92 vs ’93	[0, 0.150)	[0, 0.0023)

5. Conclusions

We have developed inference procedures to detect and quantify changes in the SCDFs at two different points in time, for a given spatial domain of interest, using a fairly flexible asymptotic structure and a methodology based on subsampling. Consequently, we have moved beyond exploratory data analysis of finite samples, such as can be found in Majure, Cook, Cressie, Kaiser, Lahiri and Symanzik

(1995), to making probability statements about the changes in the underlying spatial random processes.

A referee has suggested an investigation to account for the effect of imputation on statistical inference. The method of imputation chosen is a type of empirical Gibbs sampler, that is, the imputed values have a conditional mean and a conditional variance that match empirically the local behavior of the spatial process. Thus we believe that the effect of imputation will be small, although a more complete investigation would require a simulation study.

Both the r.f. model and the subsampling method follow a nonparametric approach. Several alternatives have been discussed in Lahiri et al. (1999). In particular, Bühlmann and Künsch (1999) proposed a spatial block bootstrap. Even though more computational time is generally required, the block bootstrap seems to handle missing data with greater flexibility. Moreover, Handcock (1999) constructed a parametric model to make inferences about the SCDF and compared his results with those in Lahiri et al. (1999). Despite the flexibility such parametric models possess, the underlying distributions need to be specified explicitly, in contrast to the distribution-free nature of the approach here. Further comparative studies will be needed to help determine which method is more suitable, and under what conditions.

6. Proofs

Proof of Theorem 2.1. Since the limiting result is under the null hypothesis that $G_1 = G_2$, we omit the subscript and let G denote the invariant CDF for both r.f.s $Z_1(\cdot)$ and $Z_2(\cdot)$. Let $G^{-1}(\cdot)$ denote the inverse of $G(\cdot)$. Since $G(\cdot)$ is continuous, by standard arguments (see, e.g., Pollard (1984, p.155)), it is enough to show that the rescaled process $\tilde{\xi}_n(\cdot) \equiv \xi_n(G^{-1}(\cdot))$ converges in distribution to $\tilde{\mathcal{W}}(\cdot) \equiv \mathcal{W}(G^{-1}(\cdot))$ as random elements of $D[0,1]$, where $D[0,1]$ denotes the space of all right-continuous functions on $[0,1]$ with left-hand limits equipped with the Skorohod metric. We use Theorem 15.1 in Billingsley (1968) to prove the FCLT of $\tilde{\xi}_n(\cdot)$ by showing weak convergence of its finite-dimensional distributions and then establishing the tightness of $\tilde{\xi}_n(\cdot)$ and almost-sure continuous sample paths of $\tilde{\mathcal{W}}(\cdot)$. The weak convergence of finite-dimensional distributions follows by Theorem 3.1 of Lahiri (1998) and the Cramér-Wold device. The arguments for tightness are similar to those of Theorem 22.1 of Billingsley (1968). More details are given in the proof of Theorem 3.1 of Zhu, Lahiri and Cressie (2000).

Proof of Theorem 2.2. The proof is similar to, but slightly more general than the proof of Theorem A.2 in Lahiri et al. (1999), with a generalization from dimension $d = 2$ to $d \geq 1$. More details are given in the proof of Theorem 3.2 of Zhu et al. (2000).

Proof of Theorem 3.1. Let $\upsilon_n(z) \equiv b_n[(F_{n,1}(z) - F_{n,2}(z))^2 - (F_{\infty,1}(z) - F_{\infty,2}(z))^2]$, $\upsilon_\infty(z) \equiv 2(G_1(z)-G_2(z))\mathcal{V}(z)$ if $G_1 \neq G_2$, and $\upsilon_\infty(z) \equiv 2\mathcal{W}(z)\mathcal{V}(z)$ if $G_1 = G_2$, where $z \in I\!R$. Then $\mathcal{Y}_n = \int_{I\!R} \upsilon_n(z)w(z)dz$ and $\mathcal{Y}_\infty = \int_{I\!R} \upsilon_\infty(z)w(z)dz$, where $\mathcal{Y}_n$ and $\mathcal{Y}_\infty$ are defined in (10) and Theorem 3.1. In addition, we separate each of the integrals of $\mathcal{Y}_n$ and $\mathcal{Y}_\infty$ into two parts. For $M \in (0,\infty]$, let $\mathcal{Y}_n = \int_{[-M,M]} \upsilon_n(z)w(z)dz + \int_{[-M,M]^c} \upsilon_n(z)w(z)dz \equiv \mathcal{Y}_{n,M} + \mathcal{Y}_n^{(M)}$. Similarly, $\mathcal{Y}_\infty = \int_{[-M,M]} \upsilon_\infty(z)w(z)dz + \int_{[-M,M]^c} \upsilon_\infty(z)w(z)dz \equiv \mathcal{Y}_{\infty,M} + \mathcal{Y}_\infty^{(M)}$. Let $\phi_{n,t}(z) \equiv \lambda_n^{d/2} h_n^{-1}(F_{n,t}(z) - F_{\infty,t}(z))$, $z \in I\!R$, $t = 1,2$. Under (A.1)−(A.3) and (B.1)−(B.3), $\phi_{n,t} \xrightarrow{\mathcal{D}} \mathcal{V}_t$, as $n \to \infty$, by Theorem 2.1 of Lahiri (1999).

In the case of $G_1 \neq G_2$, $b_n = \lambda_n^{d/2} h_n^{-1}$. We first show the convergence of $\upsilon_n(\cdot)$. Rewrite $\upsilon_n(\cdot)$ as $\upsilon_n(z) = -b_n^{-1}(\phi_{n,1}(z) - \phi_{n,2}(z))^2 + 2(F_{n,1}(z) - F_{n,2}(z))(\phi_{n,1}(z) - \phi_{n,2}(z))$, $z \in I\!R$. By Lemma 4 of Zhu et al. (2000), a finite-dimensional weak-convergence result holds for $\phi_n \equiv \phi_{n,1} - \phi_{n,2}$. Next, the tightness of $\phi_n(\cdot)$ and almost-sure continuous sample paths of $\mathcal{V}(\cdot)$ hold, since they hold for both $\phi_{n,1}(\cdot)$ and $\phi_{n,2}(\cdot)$. Hence, $\phi_n \xrightarrow{\mathcal{D}} \mathcal{V}$. Now, by similar arguments as in the proof of Theorem 2.1 in this paper, the FCLT holds for $\lambda_n^{d/2}(F_{n,t} - G_t)$, $t = 1,2$. Hence, for a given $z \in I\!R$, $E(F_{n,t}(z) - G_t(z))^2 = O(\lambda_n^{-d})$ and $F_{n,t}(z) \xrightarrow{\mathcal{P}} G_t(z)$, as $n \to \infty$. Note that the second term in $\upsilon_n(\cdot)$ dominates the first term. By Slutsky's Theorem, $\upsilon_n \xrightarrow{\mathcal{D}} \upsilon_\infty = 2(G_1 - G_2)\mathcal{V}$.

Then we show convergence of $\int_{I\!R} \upsilon_n(z)w(z)dz$. We check the conditions (C.1)−(C.3) in Lemma 5 of Zhu et al. (2000). Since $\upsilon_n \xrightarrow{\mathcal{D}} \upsilon_\infty$ and by the Continuous Mapping Theorem, $\mathcal{Y}_{n,M} \xrightarrow{\mathcal{D}} \mathcal{Y}_{\infty,M}$ for all $M > 0$. By the FCLT of υ_n and Lemma 3 of Zhu et al. (2000), $E(\upsilon_\infty(z)^2) = 4(G_1(z) - G_2(z))^2 \varsigma(z,z)$ and $E(\upsilon_n(z)^2) = E(\upsilon_\infty(z)^2)(1 + o(1))$. Hence given an $M > 0$, by the Cauchy-Schwarz Inequality, $P(\sup_n |\mathcal{Y}_n^{(M)}| > \epsilon) \leq \epsilon^{-1} \int_{|z|>M} C\Big(E(\upsilon_\infty(z)^2)\Big)^{1/2} w(z)dz$, where C is a positive constant. By (B.4), the upper bound tends to zero as $M \to \infty$. Similarly, as $M \to \infty$, $P(|\mathcal{Y}_\infty^{(M)}| > \epsilon) \to 0$. Hence we can apply Lemma 5 of Zhu et al. (2000) and conclude that $\mathcal{Y}_n \xrightarrow{\mathcal{D}} \mathcal{Y}_\infty$ as $n \to \infty$.

In case $G_1 = G_2$, $b_n = \lambda_n^d h_n^{-1}$. Again, we first show convergence of $\upsilon_n(\cdot)$. Rewrite $\upsilon_n(\cdot)$ as $\upsilon_n(z) = -h_n(\phi_{n,1}(z) - \phi_{n,2}(z))^2 + 2\lambda_n^{d/2}(F_{n,1}(z) - F_{n,2}(z))(\phi_{n,1}(z) - \phi_{n,2}(z))$, $z \in I\!R$. We have shown that $\phi_n = \phi_{n,1} - \phi_{n,2} \xrightarrow{\mathcal{D}} \mathcal{V}$ as $n \to \infty$. Recall from Theorem 2.1, when $G_1 = G_2$, $\xi_n = \lambda_n^{d/2}(F_{n,1} - F_{n,2}) \xrightarrow{\mathcal{D}} \mathcal{W}$ as $n \to \infty$. By similar arguments as in Lemma 4 of Zhu et al. (2000), the finite-dimensional distributions of the vector process $(\xi_n, \phi_n)'$ converges weakly to those of the Gaussian process $(\mathcal{W}, \mathcal{V})'$, where $\mathcal{W}(\cdot)$ and $\mathcal{V}(\cdot)$ have cross-covariance function $\upsilon(\cdot,\cdot)$ defined in (11). Now define the metric d_2 on the product space $D[-\infty,\infty] \times D[-\infty,\infty]$ by $d_2((f_1,g_1)',(f_2,g_2)') \equiv \max\{d_1(f_1,f_2), d_1(g_1,g_2)\}$, where $(f_1,g_1)',(f_2,g_2)' \in D[-\infty,\infty] \times D[-\infty,\infty]$. Since both $\xi_n(\cdot)$ and $\phi_n(\cdot)$

are tight, $(\xi_n(\cdot), \phi_n(\cdot))'$ is also tight. This, together with the convergence of the finite-dimensional distributions of $(\xi_n, \phi_n)'$, implies that $(\xi_n, \phi_n)' \xrightarrow{\mathcal{D}} (\mathcal{W}, \mathcal{V})'$. Hence, by the Continuous Mapping Theorem, $\xi_n \phi_n \xrightarrow{\mathcal{D}} \mathcal{W}\mathcal{V}$ as $n \to \infty$. Since the process $(\phi_{n,1}(z) - \phi_{n,2}(z))^2$ is tight and $h_n \to 0$, we have $\upsilon_n \xrightarrow{\mathcal{D}} 2\mathcal{W}\mathcal{V}$.

The proof of convergence of $\int_{I\!R} \upsilon_n(z) w(z) dz$ is similar to that in the case $G_1 \neq G_2$. Since $\upsilon_n \xrightarrow{\mathcal{D}} \upsilon_\infty$, $\xi_n \xrightarrow{\mathcal{D}} \mathcal{W}$, $\phi_n \xrightarrow{\mathcal{D}} \mathcal{V}$, and $\mathcal{W}(\cdot)$ and $\mathcal{V}(\cdot)$ are Gaussian processes, by the Cauchy-Schwarz Inequality $P(\sup_n |\mathcal{Y}_n^{(M)}| > \epsilon) \leq \epsilon^{-1} \int_{|z|>M} C^*(\varsigma(z,z)\sigma(z,z))^{1/2} w(z) dz$, for some positive constant C^*. By (B.4), the upper bound tends to zero as $M \to \infty$. Hence we can apply Lemma 5 of Zhu et al. (2000) and conclude that $\mathcal{Y}_n \xrightarrow{\mathcal{D}} \mathcal{Y}_\infty$ as $n \to \infty$.

Proof of Theorem 3.2.

Let $\mathcal{X}_\infty(R_{*i})$ denote $\mathcal{X}_\infty$ restricted to the region R_{*i}, and let $\tilde{H}_n(z) = K_n^{-1} \sum_{i=1}^{K_n} I(b_l(\mathcal{X}_n^{*i} - \mathcal{X}_\infty(R_{*i})) \leq z)$, $z \in I\!R$. Following the same arguments as in the proof of Theorem 2.2, we obtain for all $z \in I\!R$, $E(\tilde{H}_n(z) - H_n(z))^2 \to 0$ as $n \to \infty$.

Fix $\epsilon > 0$. By the continuity of $H(\cdot)$, there exists an $\eta > 0$ such that $\sup_{z \in I\!R} |H(z+\eta) - H(z-\eta)| \leq \epsilon$. Now let $D_n \equiv K_n^{-1} \sum_{i=1}^{K_n} I(b_l |\mathcal{X}_\infty^{*i} - \mathcal{X}_\infty(R_{*i})| > \eta)$. Then for $z \in I\!R$, $P(|\hat{H}_n(z) - \tilde{H}_n(z)| > 4\epsilon) \leq P(|\tilde{H}_n(z+\eta) - \tilde{H}_n(z-\eta)| > 3\epsilon) + P(D_n > \epsilon) \leq C[\epsilon^{-2} E(\tilde{H}_n(z+\eta) - H_n(z+\eta))^2 + \epsilon^{-2} E(\tilde{H}_n(z-\eta) - H_n(z-\eta))^2] + \epsilon^{-1} P(b_l |\mathcal{X}_\infty^{*1} - \mathcal{X}_\infty(R_{*1})| > \eta)$, for some constant C. Note that the first two terms on the right-hand side tend to 0. Let $b_l^* \equiv \lambda_l^{d/2} h_n^{-1}$ if $G_1 \neq G_2$, and $b_l^* \equiv \lambda_l^d h_n^{-1}$ if $G_1 = G_2$. Because (B.3) holds with λ_l, Theorem 3.1 holds with λ_n replaced by λ_l. Hence $b_l^*(\mathcal{X}_\infty^{*1} - \mathcal{X}_\infty(R_{*1})) \xrightarrow{\mathcal{D}} \mathcal{Y}_\infty$. Since $b_l = b_l^*(h_n/h_l)$ and $h_n/h_l \to 0$, we have $b_l(\mathcal{X}_\infty^{*1} - \mathcal{X}_\infty(R_{*1})) \xrightarrow{\mathcal{P}} 0$ as $n \to \infty$. Hence, the last term on the right-hand side tends to 0, and consequently, $\hat{H}_n(z) - H_n(z) \xrightarrow{\mathcal{P}} 0$ as $n \to \infty$ for all $z \in I\!R$. Because $\mathcal{Y}_\infty$ has a continuous CDF $H(\cdot)$ on $I\!R$ and, by Theorem 3.1, $H_n(\cdot)$ converges to it, Theorem 3.2 holds.

Acknowledgement

This research was partially supported by the U.S. Environmental Protection Agency under Cooperative Agreement CR 822919-01-0 with Iowa State University and Assistance Agreement R827527-01-0 with The Ohio State University, and by the National Science Foundation grant DMS98-16630. The authors would like to thank an anonymous referee and an associate editor for their helpful comments.

References

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Bühlmann, P. and Künsch, H. R. (1999). Comment on "Prediction of spatial cumulative distribution functions using subsampling" by Lahiri et al. *J. Amer. Statist. Assoc.* **94**, 97-99.

Cressie, N. (1993). *Statistics for Spatial Data*, Revised Edition. Wiley, New York.

Doukhan, P. (1994). *Mixing: Properties and Examples*. Springer, New York.

Handcock, M. S. (1999). Comment on "Prediction of spatial cumulative distribution functions using subsampling" by Lahiri et al. *J. Amer. Statist. Assoc.* **94**, 100-102.

Lahiri, S. N. (1998). Central limit theorems for weighted sums under some spatial sampling schemes. Preprint, Department of Statistics, Iowa State University.

Lahiri, S. N. (1999). Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method. *Probab. Theory Related Fields* **114**, 55-84.

Lahiri, S. N., Kaiser, M. S. and Cressie, N. and Hsu, N. J. (1999). Prediction of spatial cumulative distribution functions using subsampling. *J. Amer. Statist. Assoc.* **94**, 86-110.

Majure, J. J., Cook, D. and Cressie, N., Kaiser, M. S., Lahiri, S. N. and Symanzik, J. (1995). Spatial CDF estimation and visualization with applications to forest health monitoring. *Comput. Sci. Statist.* **27**, 93-101.

Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

Zhu, J., Lahiri, S. N. and Cressie, N. (2000). Asymptotic inference for spatial CDFs over time. Technical Report 666, Department of Statistics, The Ohio State University.

Department of Statistics, University of Wisconsin − Madison, 1210 West Dayton Street, Madison, WI 53706-1685.

E-mail: jzhu@stat.wisc.edu.

Department of Statistics, Iowa State University, Ames, IA 50011-1210, U.S.A.

E-mail: snlahiri@iastate.edu

Department of Statistics, The Ohio State University, Columbus, OH 43210-1247, U.S.A.

E-mail: ncressie@stat.ohio-state.edu

(Received June 2000; accepted January 2002)