Simulation-Based Excursion Statistics

By Gordon A. Fenton[1] and Erik H. Vanmarcke,[2] Members, ASCE

ABSTRACT: Failure of an engineered system is typically associated with extremes in properties and/or loads characterizing the system. Since detailed field studies are usually unavailable, system properties are best modeled as random functions during the design or analysis process. To assess system reliability, the challenge is to establish relationships between parameters of the random function model and occurrence of threshold excursions or extremes. This paper estimates level excursion statistics of the two-dimensional Gauss-Markov model through simulations of an associated local average process. Among the statistics obtained are the mean number of excursions, their areas, and a new cluster measure that reflects the spatial distribution of the excursion regions. Motivated by the lack of analytical results for processes that are not mean-square differentiable and by the limitation of the few available theories to high threshold levels, the methodology yields empirical results that can be used directly in reliability analyses and that can be easily extended to higher dimensions and nonstandard excursion measures.

INTRODUCTION

Many physical systems of interest to designers and researchers have properties and/or states varying both in space and time. Attempts to realistically model such systems lead naturally to their description as random processes. In a reliability context, it is the excursions of a system's attributes that are typically most relevant; many reliability problems are defined in terms of threshold excursions—when load exceeds strength, for example. Thus, a thorough understanding of the nature of excursion regions associated with the stochastic model is essential.

Most theories governing extremal statistics of random fields deal with excursion regions, regions in which the process Z exceeds some threshold. The few exact results that exist usually only apply asymptotically for high thresholds and a large class of random functions are not amenable to existing extrema theory at all. For such processes the analysis of a sequence of realizations is currently the only way to obtain their excursion statistics.

Recent work by the writers (Fenton and Vanmarcke 1991) clearly illustrates the need for excursion statistics at low thresholds as well. The reference deals with the assessment of liquefaction risk in a soil mass whose properties are considered to be spatially random variables. Analysis of individual realizations typically produce microzonation maps similar to that shown in Fig. 1, where the dark regions, or excursions, represent soil volumes that have liquefied. Clearly, the risk of global liquefaction or damage at the site is dependent on both the total volume of liquefied material within the soil mass as well as the degree of clustering of the liquefied pockets (if the pockets are well distributed, with intervening stable regions, then the risk of global liquefaction is reduced). For such a problem, the size and

---

[1]Asst. Prof., Dept. of Appl. Mathematics, Tech. Univ. of Nova Scotia, Halifax, Nova Scotia, B3J 2X4.

[2]Prof., Dept. of Civ. Engrg. and Operations Res., Princeton Univ., NJ 08544.
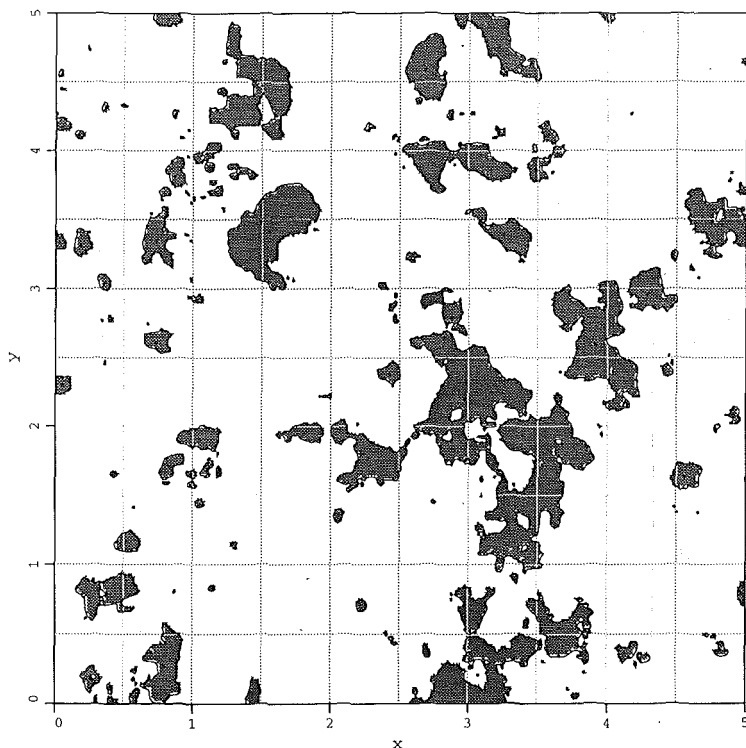
1129

FIG. 1. Sample Function of Binary Field Y [(26)]; Regions Shown in Gray Represent Regions of $\mathfrak{X}$ that Exceed Threshold $b = 1\sigma_D$; $\mathfrak{X}$ Is Generated via 2-D LAS Algorithm According to (21) with $\theta = 1/2$

spatial distribution of excursions at relatively low thresholds, where existing theories do not apply, are of considerable interest.

In this paper a methodology for simulation-based estimation of the statistics of level excursions and extrema will be developed. The treatment herein is limited to the two-dimensional case although the procedure is easily extended to higher dimensions. Six quantities having to do with level excursions and extrema of two-dimensional random fields are examined (Fenton 1990).

1. The total area of excursion regions within a given domain (A_b).
2. The number of isolated excursion regions (N_b).
3. The area of isolated excursion regions (A_e).
4. An integral geometric characteristic defined by Adler (1981) (Γ).
5. A measure of "clustering" defined herein (Ψ).
6. The distribution of the global maxima.

These quantities will be estimated for a single class of random functions, namely homogeneous isotropic scalar Gaussian processes with Markovian covariance structure (Gauss-Markov processes), over a limited range of scales of fluctuation and threshold levels. The study is by no means complete

1130

and should be viewed primarily as a new approach to the determination of these statistics. It is hoped that the appearance of empirical results will aid and guide the formulation of exact analytical expressions.

BACKGROUND THEORY

Within a given domain $\mathcal{V} = (0, R_1) \times (0, R_2)$ of area $A_T = R_1 R_2$, the total excursion area, A_b, where the process $Z(\mathbf{x})$ exceeds some threshold, can be defined by

$$A_b = \int_{\mathcal{V}} I_{\mathcal{V}}(Z(\mathbf{x}) - b\sigma) \, d\mathbf{x} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (1)$$

where $b\sigma$ = the threshold of interest, σ^2 being the variance of the process, and $I_{\mathcal{V}}(\cdot)$ = the indicator function defined on $\mathcal{V}$ (taken to be zero outside the domain $\mathcal{V}$)

$$I_{\mathcal{V}}(t) = 1 \quad \text{if } t \geq 0 \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (2a)$$

$$I_{\mathcal{V}}(t) = 0 \quad \text{if } t < 0 \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (2b)$$

For a homogeneous process, the expected value of A_b is simply

$$E[A_b] = A_T P[Z(\mathbf{0}) \geq b\sigma] \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (3)$$

which, for a zero-mean Gaussian process yields

$$E[A_b] = A_T[1 - \Phi(b)] \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (4)$$

where Φ = the standard normal distribution function. The total excursion area A_b is made up of the N_b areas of isolated (disjoint) excursions A_e as follows

$$A_b = \sum_{i=1}^{N_b} A_{ei} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (5)$$

for which the isolated excursion regions can be defined using a point set representation

$$\mathcal{A}_{ei} = \{\mathbf{x} \in \mathcal{V} : Z(\mathbf{x}) \geq b\sigma, \mathbf{x} \notin \mathcal{A}_{ej} \forall j \neq i\} \quad \dots\dots\dots\dots\dots\dots \quad (6a)$$

$$A_{ei} = \mathcal{L}(\mathcal{A}_{ei}) \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (6b)$$

where $\mathcal{L}(\mathcal{A}_{ei})$ = the Lebesque measure (or area) of the point set $\mathcal{A}_{ei}$. Given this definition, Vanmarcke (1984) expresses the expected area of isolated excursions as a function of the second-order spectral moments

$$E[A_{ei}] = 2\pi \left[\frac{F^c(b\sigma)}{f(b\sigma)}\right]^2 |\Lambda_{11}|^{-1/2} \quad \dots\dots\dots\dots\dots\dots\dots\dots \quad (7)$$

in which F^c = the complementary distribution function [for a Gaussian process, $F^c(b\sigma) = 1 - \Phi(b)$], f = the corresponding probability density function, and Λ_{11} = the matrix of second-order spectral moments

$$\Lambda_{11} = \begin{bmatrix} \lambda_{20} & \lambda_{11} \\ \lambda_{11} & \lambda_{02} \end{bmatrix} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (8)$$

Eq. (7) assumes that the threshold level is sufficiently high so that the pattern

1131

of occurrence of excursions tends toward a two-dimensional Poisson point process. The joint spectral moments λ_{kl} can be obtained either by integrating the spectral density function

$$\lambda_{kl} = \int_\infty^\infty \int_\infty^\infty \omega_1^k \omega_2^l S(\omega_1, \omega_2) \, d\omega_1 \, d\omega_2 \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (9)$$

or through the partial derivatives of the covariance function evaluated at the origin

$$\lambda_{kl} = -\left[\frac{\partial^{k+l} B(\tau)}{\partial \tau_1^k \partial \tau_2^l}\right]_{\tau=0} \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (10)$$

These relations presume the existence of the second-order spectral moments of $Z(\mathbf{x})$, which is a feature of a mean-square differentiable process. A necessary and sufficient condition for mean-square differentiability is (Vanmarcke 1984)

$$\left[\frac{\partial B(\tau)}{\partial \tau_1}\right]_{\tau=0} = \left[\frac{\partial B(\tau)}{\partial \tau_2}\right]_{\tau=0} = 0 \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (11)$$

A quick check of the Gauss-Markov process whose covariance function is given by

$$B(\tau) = \sigma^2 \exp\left(-\frac{2}{\theta}|\tau|\right) \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (12)$$

verifies that it is not mean-square differentiable. Most of the existing theories governing extrema or excursion regions of random fields depend on this property. Other popular models that are not mean-square differentiable and so remain intractable in this respect are: (1) Ideal white noise; (2) the moving average of ideal white noise (uniformly weighted window); and (3) fractal processes.

Local Average Processes

One of the major motivations for the development of local average theory for random processes is to convert random functions that are not mean-square differentiable into processes that are. Vanmarcke (1984) shows that even a very small amount of local averaging over an area of dimension $T_1 \times T_2$ will produce finite covariances of the derivative process. For a two-dimensional local average process we can define $Z_D(\mathbf{x})$ as

$$Z_D(\mathbf{x}) = \frac{1}{D}\int_D Z(\xi) \, d\xi \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (13)$$

formed by averaging Z over a domain $D = T_1 T_2$ centered at $\mathbf{x}$. Vanmarcke presents the following relationships for the variance of the derivative process $\dot{Z}_D$ in the two coordinate directions

$$\mathrm{var}[\dot{Z}_D^{(1)}] = \frac{2}{T_1^2}\sigma^2\gamma(0, T_2)[1 - \rho(T_1|T_2)] \quad\dots\dots\dots\dots\dots\dots\dots \quad (14)$$

$$\mathrm{var}[\dot{Z}_D^{(2)}] = \frac{2}{T_2^2}\sigma^2\gamma(T_1, 0)[1 - \rho(T_2|T_1)] \quad\dots\dots\dots\dots\dots\dots\dots \quad (15)$$

1132

where

$$\dot{Z}_D^{(i)} = \frac{\partial}{\partial x_i} Z_D(\mathbf{x}) \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (16)$$

$$\rho(T_1|T_2) = \frac{1}{T_2^2\sigma^2\gamma(0, T_2)} \int_{-T_2}^{T_2} (T_2 - |\tau_2|) B(T_1, \tau_2) \, d\tau_2 \quad \dots\dots\dots\dots \quad (17)$$

$$\rho(T_2|T_1) = \frac{1}{T_1^2\sigma^2\gamma(T_1, 0)} \int_{-T_1}^{T_1} (T_1 - |\tau_1|) B(\tau_1, T_2) \, d\tau_1 \quad \dots\dots\dots\dots \quad (18)$$

and $\gamma(\cdot)$ = the variance function to be defined later. Furthermore, Vanmarcke shows that the joint second-order spectral moment of the local average process is always zero for $D > 0$, i.e.

$$\mathrm{cov}[\dot{Z}_D^{(1)}, \dot{Z}_D^{(2)}] = 0 \quad (\forall D > 0) \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (19)$$

This result implies that the determinant of the second-order spectral moment matrix for the local average process can be expressed as the product of the two directional derivative process variances

$$|\Lambda_{11,D}|^{1/2} = \sigma_{\dot{Z}_D}^2 = \{\mathrm{var}[\dot{Z}_D^{(1)}]\mathrm{var}[\dot{Z}_D^{(2)}]\}^{1/2} \quad \dots\dots\dots\dots\dots\dots \quad (20)$$

Since the theory governing statistics of level excursions and extrema for mean-square differentiable random functions is reasonably well established for high thresholds [see, for example, Cramer and Leadbetter (1967), Adler (1981), and Faber (1989)] attention will now be focused on an empirical and theoretical determination of similar measures for processes that are not directly mean-square differentiable. This will be accomplished through the use of a small amount of local averaging employing the results just stated. In particular, the six quantities specified in the introduction will be evaluated for the two-dimensional Gauss-Markov process having covariance function

$$B(\tau_1, \tau_2) = \sigma^2\exp\left(-\frac{2}{\theta}\sqrt{\tau_1^2 + \tau_2^2}\right) \quad \dots\dots\dots\dots\dots\dots\dots\dots \quad (21)$$

realizations of which will be generated using the two-dimensional local average subdivision (LAS) method described by Fenton and Vanmarcke (1990). The variance function $\gamma(T_1, T_2)$ corresponding to (21), as defined by Vanmarcke (1984), is approximated by

$$\gamma(T_1, T_2) = \frac{1}{2} [\gamma(T_2)\gamma(T_1|T_2) + \gamma(T_1)\gamma(T_2|T_1)] \quad \dots\dots\dots\dots\dots \quad (22)$$

where

$$\gamma(T_i) = \left[1 + \left(\frac{T_i}{\theta}\right)^{3/2}\right]^{-2/3} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (23)$$

$$\gamma(T_i|T_j) = \left[1 + \left(\frac{T_i}{\theta_j^i}\right)^{3/2}\right]^{-2/3} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (24)$$

$$\theta_j^i = \theta\left\{c_\alpha + (1 - c_\alpha)\exp\left[-\left(\frac{T_j}{\theta c_\alpha}\right)^2\right]\right\} \quad \dots\dots\dots\dots\dots\dots \quad (25)$$

1133

For the exponential covariance function, (21), the value of c_α should be taken as $\pi/2$. Other forms of this approximate 2-D variance function are given by Vanmarcke (1984).

Since the LAS approach automatically involves local averaging of the nonmean-square differentiable point process [(21)], the realizations will in fact be drawn from a mean-square differentiable process. The subscript D will be used to stress the fact that the results will be for the local average process.

ANALYSIS OF REALIZATIONS

Two-dimensional LAS generated realizations of homogeneous, zero-mean, isotropic, Gaussian processes are to be analyzed individually to determine various properties of the discrete binary field, Y, defined by

$$Y_{jk,D} = I_V(\mathcal{Z}_{jk,D} - b\sigma_D) \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (26)$$

where σ_D^2 = the variance of the local average process and $\mathcal{Z}_{jk,D}$ denotes its realization (indexed by j, k). The indicator function I_V is given by (2) and so $Y_D(x)$ has value 1 where the realization $\mathcal{Z}_D$ exceeds the threshold and 0 elsewhere. Fig. 1 shows a typical realization of the binary field Y obtained by determining the $b = 1\sigma_D$ excursion regions of $\mathcal{Z}$ for a scale of fluctuation $\theta = 1/2$. Also shown in Fig. 1 are the contours obtained by linear interpolation. The centroid of each excursion is marked with a darker pixel.

A space-filling algorithm was devised and implemented both to determine the area of each isolated excursion region, $A_{ei,D}$, according to (6), as well as to find the number of "holes" in these regions. In this case, the Lebesque measure is simply

$$A_{ei,D} = \mathcal{L}(\mathcal{A}_{ei,D}) = \sum_j \Delta A_{eij,D} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (27)$$

where

$$\Delta A_{eij,D} = I_{\mathcal{A}_{ei,D}}(\mathcal{Z}_D(x_j) - b\sigma_D)D \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (28)$$

is just the incremental area of each pixel within the discrete set of points $\mathcal{A}_{ei,D}$ constituting the ith isolated excursion. In practice, the sum is performed only over those pixels that are elements of the set $\mathcal{A}_{ei,D}$. Note that the area determined in this fashion is typically slightly less than that obtained by computing the area within a smooth contour obtained by linear interpolation. The difference, however, is expected to be minor at a suitably fine level of resolution.

A comment deserves to be made about the extrema statistics obtained from realizations of the LAS algorithm. The LAS method produces a local average process and thus the statistics obtained are, strictly speaking, those of a local average process and will be affected by the size of the averaging domain. Noting that as the resolution of the field is increased the local average process approaches that of the point process, we will restrict ourselves herein to an analysis of a high-resolution field. Our concentration will be primarily on the variation of extrema statistics with scale of fluctuation and threshold level and the dependence on the size of the averaging domain left for later work.

The fields to be generated will have resolution 128×128 and physical size 5×5. This gives a fairly small averaging domain having edge sizes of

$T_1 = T_2 = 5/128$, for which the variance function defined by (22)–(25) corresponding to (21) ranges in value from 0.971 to 0.999 for $\theta = 0.5$ to $\theta = 4$. In all cases, the variance of the governing (21) will be taken as unity and so σ_D^2 equals the variance function. The following statistics, with the exception of the distribution of the global maxima, are based on 400 realizations of the field.

Total Area of Excursion Regions

Since an exact relationship for the expected total area of excursion regions within a given domain, (4), is known for a Gaussian process, an estimation of this quantity from a series of realizations represents a further check on the accuracy of the simulation method. Fig. 2 shows the normalized average total area of excursions, $\bar{A}_{b,D}/A_T$, for $A_T = 25$. Here and to follow, the
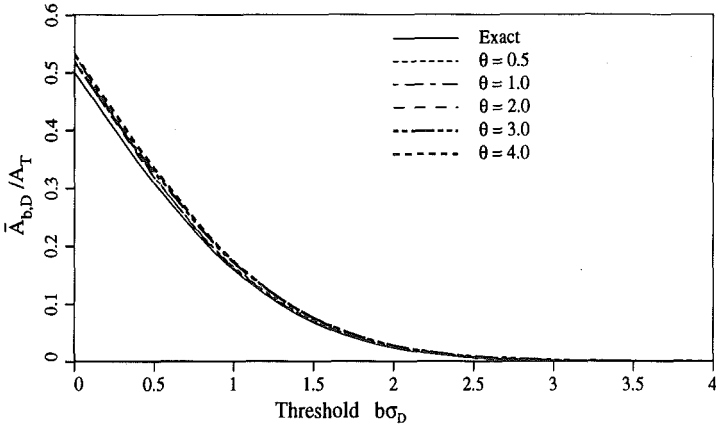

FIG. 2. Average Total Excursion Area Ratio, $\bar{A}_{b,D}/A_T$, as Function of Threshold $b\sigma_D$
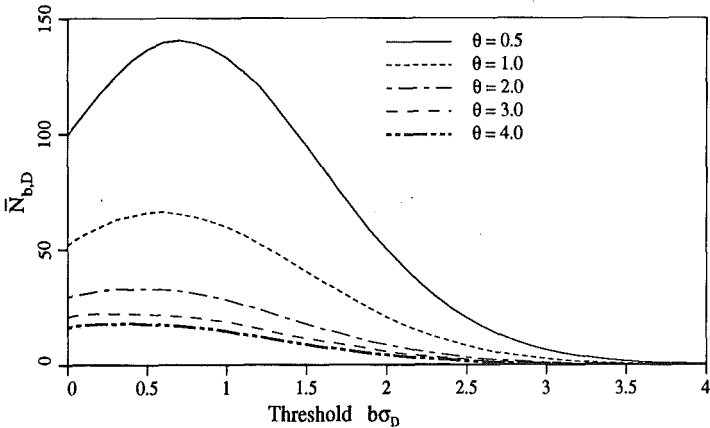

FIG. 3. Average Number of Isolated Excursions, $\bar{N}_{b,D}$, Estimated from 400 Realizations of Locally Averaged Two-Dimensional Gauss-Markov Process [(21)]

1135

overbar denotes the quantity obtained by averaging over the realizations. The estimated area ratios show excellent agreement with the exact.

Expected Number of Isolated Excursions

Fig. 3 shows the average number of isolated excursion regions observed within the domain, $\bar{N}_{b,D}$, as a function of scale and threshold. Here the word *observed* will be used to denote the average number of excursion regions seen in the individual realizations. A similar definition will apply to other quantities of interest in the remainder of the paper. The observed $\bar{N}_{b,D}$ is seen in Fig. 3 to be a relatively smooth function defined all the way out to thresholds in excess of $3\sigma_D$.

An attempt will be made to fit the theoretical results that describe the mean number of excursions of a local average process above a relatively high threshold to the data shown in Fig. 3. This expectation is just the ratio of (3) and (7) (Vanmarcke 1984)

$$E[N_{b,D}] = \frac{E[A_{b,D}]}{E[A_{ei,D}]} = \frac{A_T f_D^2(b\sigma_D)}{2\pi F_D^c(b\sigma_D)} \sigma_{\dot{Z}_D}^2 \quad \dots\dots\dots\dots\dots\dots\dots\dots (29)$$

in which f_D and F_D^c = the probability density function and complementary cumulative distribution function of the local average process respectively. $\sigma_{\dot{Z}_D}^2$ = the geometric mean of the directional variances of the derivative process as defined by (20). For the Gaussian process, (29) becomes

$$E[N_{b,D}] = \frac{A_T e^{-b^2}}{4\pi^2 \sigma_D^2 [1 - \Phi(b)]} \sigma_{\dot{Z}_D}^2 \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots (30)$$

To determine $\sigma_{\dot{Z}_D}^2$ the functions $\rho(T_1|T_2)$ and $\rho(T_2|T_1)$ must first be calculated using (17) and (18). Consider $\rho(T_1|T_2)$ for the quadrant symmetric Gauss-Markov process

$$\rho(T_1|T_2) = \frac{2}{T_2^2 \sigma^2 \gamma(0, T_2)} \int_0^{T_2} (T_2 - \tau_2) B(T_1, \tau_2) \, d\tau_2$$

$$= \frac{2}{T_2^2 \gamma(0, T_2)} \int_0^{T_2} (T_2 - \tau_2) \exp\left(-\frac{2}{\theta} \sqrt{T_1^2 + \tau_2^2}\right) d\tau_2 \quad \dots\dots\dots\dots (31)$$

Making the substitution $r^2 = T_1^2 + \tau_2^2$ gives

$$\rho(T_1|T_2) = \frac{2}{T_2^2 \sigma^2 \gamma(0, T_2)} \int_{T_1}^{\sqrt{T_1^2+T_2^2}} \left(\frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} - r e^{-2r/\theta}\right) dr \quad \dots\dots (32)$$

To avoid trying to numerically integrate a function with a singularity at its lower bound, the first term in the integrand can be evaluated as follows

$$\int_{T_1}^{\sqrt{T_1^2+T_2^2}} \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} \, dr = \int_{T_1}^\infty \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} \, dr - \int_{\sqrt{T_1^2+T_2^2}}^\infty \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} \, dr$$

$$= T_2 T_1 K_1\left(\frac{2T_1}{\theta}\right) - \int_{\sqrt{T_1^2+T_2^2}}^a \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} \, dr - \int_a^\infty \frac{T_2 r e^{-2r/\theta}}{\sqrt{r^2 - T_1^2}} \, dr \quad \dots\dots (33)$$

The second integral on the right-hand side of (33) can now be evaluated numerically and for a chosen sufficiently large, the last integral has the

1136

simple approximation $1/2\theta T_2 \exp(-2a/\theta)$. The function $K_1 = $ the modified Bessel function of order 1.

Unfortunately, for small T_1, the evaluation of (33) is extremely delicate as it involves the small differences of very large numbers. An error of only 0.1% in the estimation of either K_1 or the integrals on the right hand side can result in a drastic change in the value of $\sigma^2_{\dot{z}_D}$ particularly at larger scales of fluctuation. The results shown in Table 1 were obtained using $T_1 = T_2 = 5/128$ [for which $\rho(T_1|T_2) = \rho(T_2|T_1)$] and a 20 point Gaussian quadrature-integration scheme.

Using these variances, (29) exceeded the observed $\bar{N}_{b,D}$ at thresholds below $3\sigma_D$ by a factor of about two (Fenton 1990). This may be as a result of the combination of the difficulty in accurately determining $\sigma^2_{\dot{z}_D}$ for small averaging dimensions and the fact that (29) is an asymptotic relationship, valid only for $b \to \infty$.

An alternative approach to the description of $\bar{N}_{b,D}$ involves selecting a trial function and determining its parameters. A trial function of the form

$$\bar{N}_{b,D} = A_T(a_1 + a_2 b\sigma_D)\exp\left[-\frac{1}{2}(b\sigma_D)^2\right] \quad \dots\dots\dots\dots\dots\dots (34)$$

was chosen and an excellent fit to the observed data, as shown in Fig. 4 for three different scales of fluctuation, was obtained using the coefficients shown in Table 1. The functional form of (34) was selected so that it exhibits the correct trends beyond the range of thresholds for which its coefficients were derived.

Expected Area of Isolated Excursions

Within each realization, the average area of isolated excursions, $\bar{A}_{e,D}$, is obtained by dividing the total excursion area by the number of isolated areas. Further averaging over the 400 realizations leads to the mean excursion areas shown in Fig. 5, which are again referred to as the "observed" results. The empirical relationship of the previous section, (34), can be used along with the theoretically expected total excursion area [(4)] to obtain the semi-empirical relationship

$$\bar{A}_{e,D} = \frac{[1 - \Phi(b)]\exp\left[\frac{1}{2}(b\sigma_D)^2\right]}{a_1 + a_2 b\sigma_D} \quad \dots\dots\dots\dots\dots\dots\dots (35)$$

which is compared to the relationship in Fig. 5 and is seen to show very good agreement. Note that the vertical scales in this and subsequent figures are different from plot to plot to illustrate the degree of fit.

TABLE 1. Computed Variances of Local Average Derivative Process

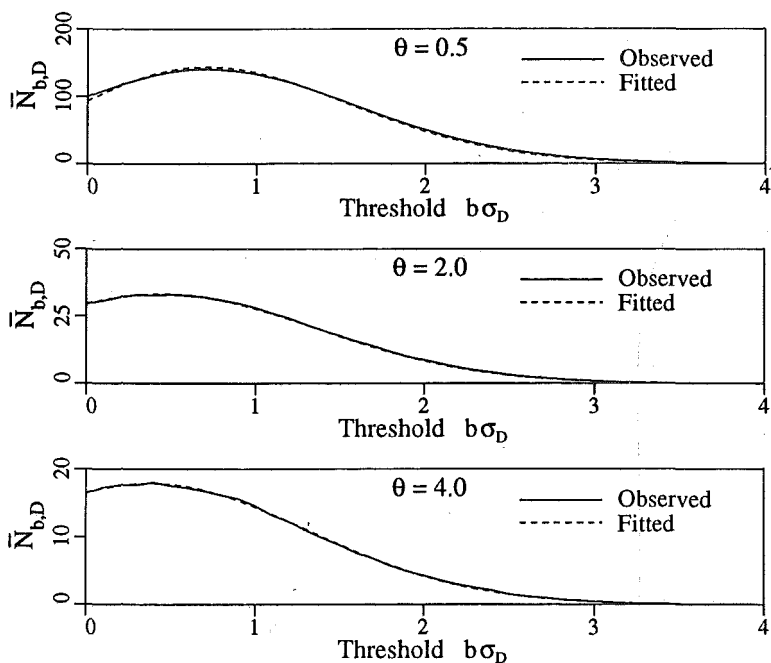| Scale (1)	$\rho(T_1|T_2)$ (2)	$\sigma^2_{\dot{z}_D}$ (3)	a_1 (4)	a_2 (5)
0.5	0.8482	196.18	3.70	5.20
1.0	0.9193	105.18	2.05	1.90
2.0	0.9592	53.32	1.18	0.65
3.0	0.9741	33.95	0.81	0.41
4.0	0.9822	23.30	0.66	0.29

1137

FIG. 4. Comparison of Empirical Fit by (34) with Observed Average Number of Isolated Excursions Obtained by Simulation

Integral Geometric Characteristic of 2-D Random Fields

In his thorough treatment of the geometrical properties of random fields, Adler (1981) develops an integral geometric (IG) characteristic, $\Gamma(\mathcal{A}_{b,D})$, as a statistical measure of two-dimensional random fields. It is considered in this paper mainly to compare existing theory with simulation-based observations. The definition of $\Gamma(\mathcal{A}_{b,D})$ will be shown here specifically for the two-dimensional case although a much more general definition is given by Adler. First, using a point-set representation, the excursion set $\mathcal{A}_{b,D}$ can be defined as the set of points in $\mathcal{V} = [0, T_1] \times [0, T_2]$ for which $Z_D(\mathbf{x}) \geq b\sigma_D$

$$\mathcal{A}_{b,D} = \{\mathbf{x} \in \mathcal{V} : Z_D(\mathbf{x}) \geq b\sigma_D\} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (36)$$

The Hadwiger characteristic of $\mathcal{A}_{b,D}$, $\varphi(\mathcal{A}_{b,D})$, is equal to the number of connected components of $\mathcal{A}_{b,D}$ (the number of isolated excursion regions) minus the number of holes in $\mathcal{A}_{b,D}$. Finally, if $\hat{\mathcal{V}}$ is defined as the edges of $\mathcal{V}$ that pass through the origin (the coordinate axes), then the IG characteristic is formally defined as

$$\Gamma(\mathcal{A}_{b,D}) = \varphi(\mathcal{A}_{b,D}) - \varphi(\mathcal{A}_{b,D} \cap \hat{\mathcal{V}}) \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (37)$$

Essentially, $\Gamma(\mathcal{A}_{b,D})$ is equal to the number of isolated excursion areas that do not intersect the coordinate axes minus the number of holes in them. Fig. 6 shows the average value of the IG characteristic, $\bar{\Gamma}(\mathcal{A}_{b,D})$, obtained from the locally averaged Gauss-Markov process realizations.
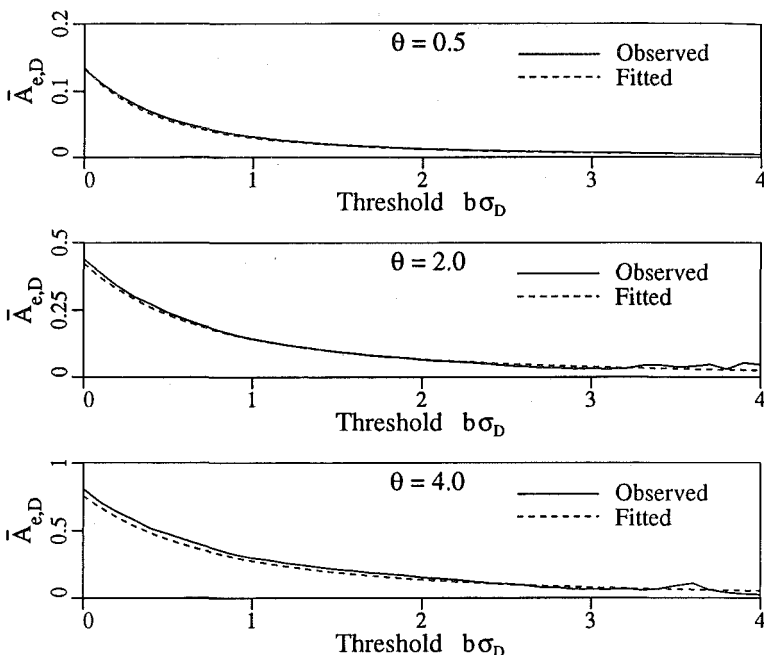
1138

FIG. 5. Comparison of Semi-Empirical Fit by (35) with Observed Average Area of Isolated Excursions Obtained by Simulation

Adler presents an analytic result for the expected value of $\Gamma(\mathscr{A}_{b,D})$ that has been modified here to account for local averaging of a Gaussian process

$$E[\Gamma(\mathscr{A}_{b,D})] = \frac{bA_T}{(2\pi)^{3/2}\sigma_D^2} \exp\left(-\frac{1}{2}b^2\right)\sigma_{\dot{Z}_D}^2 \quad \dots\dots\dots\dots\dots\dots\dots (38)$$

The presence of b in the numerator implies that this function goes to zero when $b = 0$, which does not correspond with Fig. 6. At higher thresholds, errors in the estimation of $\sigma_{\dot{Z}_D}^2$ lead also to poor agreement.

Using a function of the same form as (34)

$$\bar{\Gamma}(\mathscr{A}_{b,D}) = A_T(g_1 + g_2 b\sigma_D)\exp\left[-\frac{1}{2}(b\sigma_D)^2\right] \quad \dots\dots\dots\dots\dots\dots (39)$$

yields a much closer fit over the entire range of thresholds by using the empirically determined parameters shown in Table 2. Fig. 6 illustrates the comparison for three scales of fluctuation.

Clustering of Excursion Regions

Once the total area of an excursion and the number of components that make it up have been determined, a natural question to ask is how the components are distributed: do they tend to be clustered together or are they more uniformly distributed throughout the domain? Questions of this nature arise in many reliability contexts such as the liquefaction risk study discussed in the introduction. Another example might be strength degra-
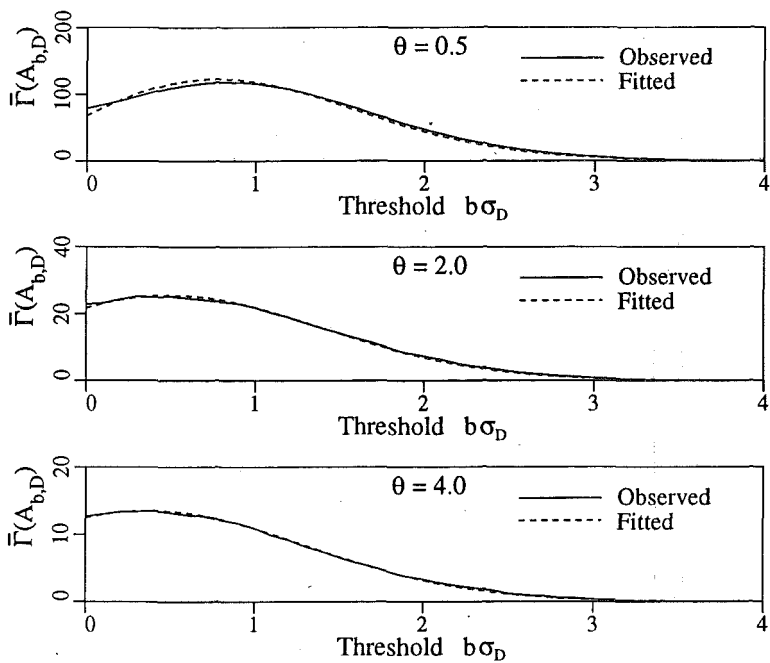
1139

FIG. 6. Comparison of Empirically Predicted IG Characteristic [(39)] with Observed Average Values Obtained by Simulation

TABLE 2. Empirically Determined Parameters of (39) Based on Observed Average IG Characteristic $\bar{\Gamma}$ Obtained by Simulation

Scale (1)	g_1 (2)	g_2 (3)
0.5	2.70	5.10
1.0	1.50	1.80
2.0	0.87	0.58
3.0	0.61	0.32
4.0	0.50	0.22

dation in a concrete bridge deck: do the low strength regions tend to be clustered? How does this affect expected repair costs?

It would be useful to define a measure, herein called Ψ, that varies from 0 to 1 and denotes the degree of clustering; 0 corresponding to a uniform distribution and larger values corresponding to denser clustering. The determination of such a measure involves first defining a reference domain within which the measure will be calculated. This is necessary since a homogeneous process over infinite space always has excursion regions throughout the space. On such a scale, the regions will always appear uniformly distributed (unless the scale of fluctuation also approaches infinity). For example, at scales approaching the boundaries of the known universe, the distribution of galaxies appears very uniform. It is only when attention is restricted to smaller volumes of space that one begins to see the local

1140

clustering of stars. Thus an examination of the tendency of excursions to occur in groups must involve a comparison, within the reference domain, of the existing pattern of excursions against the two extremes of uniform distribution and perfect clustering.

A definition for Ψ that satisfies these criteria can be stated as follows

$$\Psi = \frac{J_u - J_b}{J_u - J_c} \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (40)$$

where J_b = the polar moment of inertia of the excursion areas about their combined centroid, J_c = the polar moment of inertia of all the excursion areas concentrated within a circle, and J_u = the polar moment of inertia about the same centroid if the excursion area were distributed uniformly throughout the domain. Specifically

$$J_b = \sum_i^{N_{b,D}} J_{ei} + A_{ei,D}|\bar{\mathbf{x}}_b - \bar{\mathbf{x}}_i|^2 \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (41)$$

$$J_{ei} = \sum_j \Delta A_{eij,D}|\bar{\mathbf{x}}_i - \mathbf{x}_j|^2 \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (42)$$

$$J_u = \frac{A_{b,D}}{A_T} \int_\gamma |\bar{\mathbf{x}}_b - \mathbf{x}|^2 \, d\mathbf{x} \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (43)$$

$$J_c = \frac{A_{b,D}^2}{2\pi} \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (44)$$

where J_{ei} = the polar moment of inertia of the ith excursion region of area A_{ei} about its own centroid, $\bar{\mathbf{x}}_i$. $\Delta A_{eij,D}$ is as defined by Eq. 28 and $\bar{\mathbf{x}}_b$ = the centroid of all the excursion regions. The choice of the second moment of area is largely arbitrary, it only being important that the measure be invariant under rotations. It can be easily seen that this definition will result in $\Psi = 0$ when the excursion regions are uniformly distributed over the space ($J_b \to J_u$) and $\Psi \to 1$ when the excursion regions are clustered within a small region ($J_b \to J_c$). It is also possible for Ψ to take negative values, indicating the occurrence of two local clusters at opposite sides of the domain. This information is just as valuable as positive values for Ψ, but in practice has not been observed to occur on average.

All that remains is to define Ψ in the limiting cases. Eq. (40) ensures that Ψ will be quite close to 1 in the case of only a single excursion region. It seems natural then to take $\Psi = 1$ if no excursions occur. At the other extreme, as $A_{b,D} \to A_T$, both the denominator and numerator of (40) become very small. Although the limit for noncircular domains is zero, it appears that the measure becomes somewhat unstable as $A_{b,D} \to A_T$. This situation is of little interest since the cluster measure of a domain that entirely exceeds a threshold has no meaning. It is primarily a measure of the scatter of isolated excursions and so its use should be restricted to thresholds greater than the process mean, for example.

Individual realizations were analyzed to determine the cluster measure Ψ and then averaged over 200 realizations to obtain the results shown in Fig. 7. Definite, relatively smooth trends both with scale of fluctuation and threshold level are evident, indicating that the measure might be useful to categorize the degree of clustering.
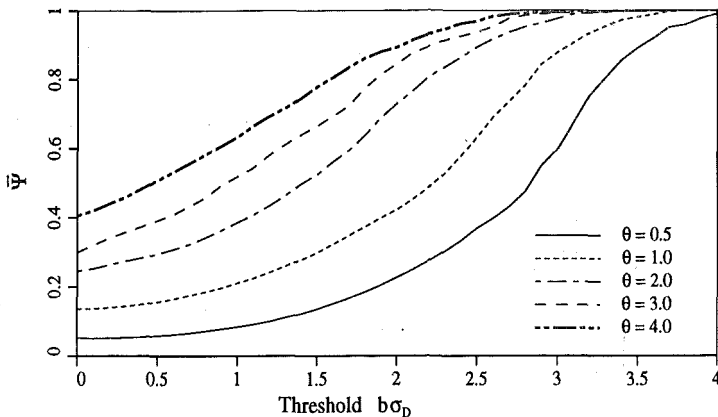
1141

FIG. 7. Average Values of Cluster Measure $\bar{\Psi}$ Estimated from 200 Realizations of Locally Averaged Gauss-Markov Process

TABLE 3. Empirically Determined Effective Number of Independent Samples n_{eff} and Parameters of Type 1 Extreme Distribution (46)

Scale (1)	n_{eff} (2)	α (3)	μ (4)
0.5	2,900	3.14	3.41
1.0	900	2.49	3.05
2.0	800	2.05	2.52
3.0	70	1.78	2.15
4.0	35	1.62	1.86

Distribution of Global Maxima

Extracting the maximum value of $\mathcal{X}_D$ from each realization allows the estimation of its corresponding probability density function (or equivalently the cumulative distribution) with reasonable accuracy given a sufficient number of realizations. A total of 2,200 realizations of the locally averaged Gauss-Markov process were generated for each scale of fluctuation considered. Conceptually, it is not unreasonable to expect the cumulative distribution of the global maximum $F_{max}(b)$ to have the form of an extreme value distribution for a Gaussian process

$$F_{max}(b) = [\Phi(b)]^{n_{eff}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (45)$$

where n_{eff} = the *effective* number of independent samples in each realization. As the scale of fluctuation approaches zero, n_{eff} should approach the total number of field points (128×128) and as the scale becomes much larger than the field size, n_{eff} is expected to approach 1 (the field becomes totally correlated). Except at the shortest scale of fluctuation considered, $\theta = 0.5$, the function defined by (45) was disappointing in its match with the cdf obtained from the realizations. For values of n_{eff} obtained by fitting (45) to the data at $F_{max}(b) = 0.5$ (shown in Table 3), it was found that (45) increased too sharply at all but the smallest scale of fluctuation. The better fit at $\theta = 0.5$ is to be expected since at very small scales, the field consists of a set of (almost) independent random samples, thus satisfying the con-
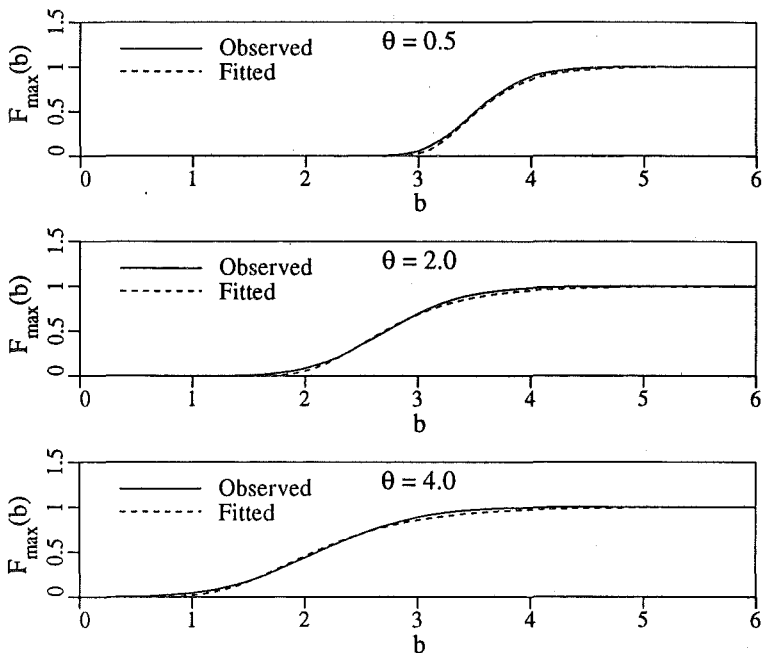
1142

FIG. 8. Observed Cumulative Distribution of Global Maximum of Each Realization Compared to Type 1 Distribution Given by (46)

ditions under which (45) applies. Not surprisingly, an improved match is obtained using a *two*-parameter type 1 extreme value distribution having the double exponential form

$$F_{max}(b) = \exp[-e^{-\alpha(b-\mu)}] \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (46)$$

where the parameters α and μ, estimated by an order-statistics method developed by Leiblein as cited in Ang and Tang (1984) using the simulation data, are presented in Table 3 for each scale of fluctuation. The comparison between the simulation-based cumulative distribution and that predicted by the type 1 extreme value distribution is shown in Fig. 8.

SUMMARY AND CONCLUSIONS

Simulation-based estimates of the mean total excursion area, mean number, and area of isolated excursions, and an integral geometric characteristic of the excursion field are compared with existing theories and matched to empirical relationships. For the Gauss-Markov process considered, which is rendered mean-square differentiable by a small amount of local averaging, the existing theories are found to be poor models. This is due in part to difficulties in numerical evaluation and to their asymptotic nature, applicable only at high thresholds. Simulation-based estimates, on the other hand, provide useful results over all threshold levels and can be easily extended to higher dimensions and to other measures (for example, the cluster measure) not amenable to closed-form solution. Such results may be meaning-

fully employed in a variety of applications, most notably reliability or decision analyses.

The study presented in this paper is not intended to be definitive, since the ranges in threshold levels, scales of fluctuation, and averaging domains are limited. Rather, the methodology paves the way for much more detailed studies and as a guide to the discovery of exact theories governing this and other common processes.

ACKNOWLEDGMENTS

APPENDIX I.   REFERENCES

Adler, R. J. (1981). *The geometry of random fields*. John Wiley and Sons, New York, N.Y.
Ang, A. H.-S., and Tang, W. H. (1984). Probability concepts in engineering planning and design, Vol. 2, John Wiley and Sons, New York, N.Y.
Cramer, H., and Leadbetter, M. R. (1967). *Stationary and related stochastic processes*. John Wiley and Sons, New York, N.Y.
Faber, M. H. (1989). "Excursions of Gaussian random fields in structural reliability," PhD thesis, University of Aalborg, Aalborg, Denmark.
Fenton, G. A. (1990). "Simulation and analysis of random fields," PhD thesis, Princeton University, Princeton, N.J.
Fenton, G. A., and Vanmarcke, E. H. (1990). "Simulation of random fields via local average subdivision." *J. Engrg. Mech.*, ASCE, 116(8), 1733–1749.
Fenton, G. A., and Vanmarcke, E. H. (1991). "Spatial variation in liquefaction risk assessment." *Proc. ASCE Geotech. Engrg. Congress*, ASCE, 594–607.
Vanmarcke, E. H. (1984). *Random fields: analysis and synthesis*, MIT Press, Cambridge, Mass.

APPENDIX II.   NOTATIONS

The following symbols are used in this paper:

A_b = total excursion area above threshold $b\sigma$;
$A_{b,D}$ = excursion point set above threshold $b\sigma_D$;
A_e = area of isolated excursions above threshold $b\sigma$;
$\mathscr{A}_e$ = isolated excursion point set above threshold $b\sigma$;
A_T = total domain area;
B = covariance function;
b = threshold level;
D = area of local averaging domain;
$E[\cdot]$ = expectation operator;
F = cumulative distribution function;
F^c = complementary cumulative distribution function;
f = probability density function;
$I_\mathcal{V}$ = indicator function defined on point set $\mathcal{V}$;

1144

$\mathcal{L}$ = Lebesque measure;
K_1 = modified Bessel function of order 1;
N_b = number of isolated excursion regions above threshold $b\sigma$;
n_{eff} = effective number of independent samples;
$P[\cdot]$ = probability operator;
S = spectral density function;
T = dimension of the local averaging domain;
$\mathcal{V}$ = domain point set;
$\hat{\mathcal{V}}$ = boundaries of domain that pass through origin;
$\mathbf{x}$ = (x_1, x_2) coordinates in domain;
Y = discrete binary field formed from excursion set;
Z = random function;
Z_D = local average process;
$\mathcal{Z}_D$ = realization of local average process Z_D;
$\dot{Z}_D$ = mean-square derivative of local average process Z_D;
Γ = integral geometric characteristic;
γ = variance function;
θ = scale of fluctuation;
Λ = matrix of second-order spectral moments;
λ = second-order spectral moment;
ρ = correlation function;
σ = standard deviation;
σ_D = standard deviation of local average process;
τ = physical lag;
Φ = standard normal distribution function;
φ = Hadwiger characteristic of excursion field;
Ψ = cluster measure; and
ω = frequency or wave number.