Spatial variability assessment of structures from adaptive NDT measurements

Mestapha Oumouni, Franck Schoefs

Version of Record: https://www.sciencedirect.com/science/article/pii/S0167473020301302
Manuscript_08fc379c0c34766a681d7aee24bf3a86

Spatial variability assessment of structures from adaptive NDT measurements

M. Oumouni[a,*], F. Schoefs[a]

[a] *University of Bretagne Loire, Université de Nantes, Research Institute in Civil and Mechanical Engineering (GeM), UMR CNRS 6183, 2 rue de l'Houssinire, BP 92208, 44322 Nantes, France*

Abstract

Inspection by non-destructive testing (NDT) techniques is an effective way for assessing structures' condition state, their pathologies and locating potential critical areas. However, establishing accurate diagnoses requires numerous measurements while the available budget is limited. Nevertheless, determining the spatial variability of the material properties leads fewer of the required measurements, as it characterizes zones with almost identical properties and helps identify weak areas. But assessing this spatial variability still requires numerous measurements. It should be limited by a rational criterion that could be expressed in terms of accuracy and cost minimization. In the present work, we develop an adaptive approach to characterize spatial variability of structures' material properties modelled by transformed Gaussian random fields. The methodology is performed in two steps. Firstly, the correlation length of the Gaussian random field is computed for a given accuracy. Secondly, the number of measurements is refined to estimate the first two moments of the random field with a given accuracy too. These two steps result in a minimization of the number of measurements (i.e the cost) for a target accuracy. The proposed methodology and theoretical results are illustrated using synthetic and real data.

Keywords: Non-Destructive Testing, Spatial variability, Correlation length,

*Corresponding author
Email addresses: mestapha.oumouni@univ-nantes.fr (M. Oumouni), franck.schoefs@univ-nantes.fr (F. Schoefs)

Gaussian random field, Maximum Likelihood Estimate, Confidence region.
2010 MSC: 00-01, 99-00

1. Introduction

In current practice, non-destructive testing (NDT) techniques (ultrasonic, resistivity and electromagnetic methods) are crucial tools for the management and the evaluation of the condition state of infrastructures. Structural damages can lead to major human, environmental or economic consequences. Therefore, adapting repair and maintenance strategies and improving diagnosis from inspection for evaluating the risk are a permanent challenge. Bayesian methods are appropriate for condition state and reliability updating when a prior distribution is known and data are periodically collected to derive the posterior distribution. Further, a generalised Bayesian framework has shown its efficiency to reduce uncertainties in the case of imperfect inspections [17, 24, 26, 35, 36]. The common challenge of these generalized frameworks is to address the issue of maintenance cost reduction. These procedures are convenient when there is no significant spatial variability of the deterioration process; in that case provide the size of the sample (number of uncorrelated measurements).

A spatial variability model of material properties is the key to describing the non-homogeneity of mechanical and physical properties of large structural components. It is derived from three main factors: the intrinsic randomness of material composition (e.g. scale and nature of aggregates and particles), the severity of the environmental conditions during the life of the structure (e.g. carbonation, chemical composition of water, temperature and humidity) and the construction conditions and anisotropy (e.g. workmanship). Thus, assessing the spatial variability allows to locate potential damaged areas on infrastructures and to propagate this property with time to predict the time-variant reliability [25, 31, 33].

On the other hand, recent studies have shown that this spatial variability has an important and direct impact on structural reliability [28, 33, 34] that requires

its characterization and modeling. Non-Destructive Testing or Structural Health Monitoring are efficient and cost effective tools that provide useful information for spatial variability characterization of material properties. Several researches have focused on spatial variability of quantities of interest [8, 12, 19, 28, 29]. However, to our knowledge, none of them has focused on sampling optimization for optimal assessment of the correlation length, which governs the spatial correlation, roughness and anisotropy of the random field. For example, Nguyen et al [19] combined several NDT techniques, variogram fitting and kriging prediction, to describe the spatial variability of concrete and using the simple confidence region of the mean which is not accurate since it doesn't consider the correlation between measures. In [8] an adaptive method uses the kriging prediction based on the spatial correlation from prior information. The kriging procedure is progressively performed in view to select the spatial locations with minimum variance. However, the estimated parameters can be poorly estimated due to the bad prediction of the prior spatial variability.

Schoefs et al, [29] proposed a two-step procedure to assess the minimal number of measurements under some assumptions. Firstly, the spatial variability is evaluated using a fast NDT. Secondly, after introducing a threshold distance of negligible correlation, Monte-Carlo simulations combined with Karhunen-Loève discretization are performed to compute the confidence region of the mean and the variance. However, a poor estimation of the correlation length at the first stage of the method can provide inaccurate reliability estimates. Further, the simulated confidence region through the Bootstrapping approach is not accurate since the correlation between measures is not considered.

The main objective of this paper is to propose an adaptive methodology that computes the number and locations of measurements needed to assess spatial variability of quantities within a required confidence level. The adaptive approach is based on two estimates. The first one is an error indicator that estimates an accurate scale of fluctuation [5, 39] based on Maximum Likelihood Estimate (MLE) [4, 16, 37]. Second, the statistical indicator quantifies the error on the mean; its decay allows to revise the accuracy of the estimated

moments of the field at the first stage. The advantage of this approach is that the spatial correlation is not neglected and the parameters of the model are provided within the target accuracy. Hence, our contribution provides a new approach to improve cost/benefit ratio of NDT techniques on site in terms of number of measurements given a target accuracy.

The paper starts with a review of key concepts about spatial random field modeling with a focus on stationary and ergodic random fields. Then we describe the iterative MLE and properties of the estimates. Section 3 develops the procedure to adapt the number and locations of the measurements during the inspection protocol. Finally, Section 4 presents numerical examples with synthetic and real data illustrating and validating the proposed methodology.

2. Spatial random field for inspection modeling

The random fields theory is a basic probabilistic theory for modeling uncertainties of spatially dependent quantities (material property, geometry, deterioration process, load, etc). Spatial uncertainty is statistically quantified by its mean, variance, and spatial correlation. In most applications, it is given by a continuous spatial function and depends on a scale of fluctuation. It describes how local properties are statistically similar in comparison with others (close or far). The scale of fluctuation quantifies how properties vary in space (strong for small scale or weak for large one) .

Random fields could take several forms of the stationary and non-stationary properties, depending on the intrinsic characteristics of the material, external factors, and the quantities of interest to be estimated. They are usually described by their spatial correlation or other statistical estimates like mean, variance, variogram or scale of fluctuation. Spatial or spatio-temporal random fields are extensively used in structural engineering for modelling the degradation process and reliability assessment [23, 28, 34]. Second order (with finite variance) stationary random fields have been used to model the spatial variability of material properties; for instance surface chloride concentration on concrete

structures, concrete or soil properties [9, 11, 20]. The Gaussian Random Field (GRF) or a transformation of GRF (log-normal random field) is widely used in many applications, because it is completely defined by its mean and its covariance function.

2.1. Stationary and ergodic random field

In classical statistics, the repetition and the independence of events are essential to derive parameter estimates or testing hypotheses. The equivalence of these two properties in spatial statistics is the stationarity (spatial repetition) and ergodicity (asymptotic independence where the correlation between events vanishes at infinity). Therefore, we focus on these two key properties in the following. We consider a spatial set D of the d-dimensional space of coordinates x in $\mathbb{R}^d$ (with $d = 1, 2\, or\, 3$).

Definition 2.1. Z is a second order stationary random field (in the weak sense), if its moments are invariant in space under translation, for each points x and $x + h$ in D, h being the Euclidean distance (lag distance),

$$\mathbb{E}[Z(x)] = \mu, \tag{1}$$

$$\mathbb{E}[(Z(x) - \mu)(Z(x + h) - \mu)] = C(h) = C(0)c(h); \tag{2}$$

where $C(\cdot)$ is the stationary covariance, $\sigma^2 := C(0)$ is the variance of Z and $c(h) = \frac{C(h)}{C(0)}$ is the stationary correlation function of Z. In addition, if $C(h)$ depends only on the distance $\|h\|$, Z is an isotopic field.

We define a strong stationarity of Z when the distribution of each vector $(Z(x_1 + h), \ldots, Z(x_N + h))$ is independent on h. Both concepts, weak and strong stationarity are equivalent for a Gaussian field. There is a subclass of stationary fields which have the interesting property of ergodicity, where the first and second moment of Z can be estimated from a single path of Z by spatial average. In practice, the ergodic model is convenient for purely spatial phenomena [5]. The ergodicity means basically a mixture of stationarity with asymptotic independence where the stationary correlation vanishes at infinity.

Model	$c(h)$	Parameters
Matérn	$\frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}h}{\theta}\right)^{\nu}\mathcal{K}_{\nu}\left(\frac{\sqrt{2\nu}h}{\theta}\right)$	$\theta>0,\ \nu>0$
Powered exponential	$\exp\left(-\left[\frac{h}{\theta}\right]^{\nu}\right)$	$0<\nu\leq 2,\ \theta>0$
Triangular	$\begin{cases} 1-\frac{h}{\theta} & h\leq\theta \\ 0 & h>\theta \end{cases}$	$\theta>0$
Cosine exponential	$\exp\left(-\frac{h}{\theta}\right)\cos\left(\frac{h}{\theta}\right)$	$\theta>0$
Cardinal Sine	$\frac{\sin\left(\frac{h}{\theta}\right)}{\frac{h}{\theta}}$	$\theta>0$

Table 1: Some parametric classes of stationary correlation functions for a Gaussian field

A second order stationary random field Z is ergodic (Slutsky's ergodic theorem [5]) if,

$$\lim_{D\to\mathbb{R}^d}\frac{1}{|D|^2}\int_D\int_D C(x,y)dxdy=0. \quad \text{(in quadratic mean)} \tag{3}$$

2.2. Correlation function model

We focus in this subsection on the covariance function. From real data (path), the objective is to select a parametric model for this function. Table 1 provides examples of covariance functions for modeling stationary random fields. Many models are considered in the literature; several criteria impact their selection. For example the smoothness of trajectories (paths of Z i.e the function $x\to z(x)$) or the memory of the process (the decay of c to zero with respect to the lag h). They are characterized by the scale of fluctuation $\theta>0$ and the parameter of the smoothness $\nu>0$. The Matérn covariance function generalizes the most commonly used in engineering applications (Exponential and Gaussian covariance) [10]. $\mathcal{K}_{\nu}$ denotes the modified Bessel function of the second kind and Γ is the classical gamma function. The parameter ν characterizes the roughness of Z when $0<\nu<1$ or the smoothness when $\nu\geq 1$. When $\nu=\frac{1}{2}$, c coincides with the classical exponential correlation, $c(h):=e^{-h/\theta}$, and when

$\nu \longrightarrow \infty$, c approaches the Gaussian correlation, $c(h) := e^{-h^2/(2\theta^2)}$ for which the trajectories of Z are infinitely differentiable. These last two models are the most used in structural engineering applications [8, 11, 19, 29]. However, neither model takes into account into the roughness or the nugget effect of the spatial data which characterizes the behaviour of the covariance model at the origin. In order to simplify the presentation of the proposed adaptive approach, we consider a model correlation function with a fixed $\nu > 0$ and unknown scale of fluctuation θ. Further we assume following properties for modelling random fields and inspection protocols.

1. Quantities of interest are assumed to be modelled by a known transformation $\mathcal{T}(Z)$ of a Gaussian field Z.
2. The GRF Z is second order stationary, and ergodic. Then, we focus on the characterization of its mean, its variance and the spatial scale of fluctuation.
3. Inspections are carried out on a (quasi)-regular grid (constant lag h) where the number of measurements N_s can be large.

2.3. Spatial parameters estimation

We introduce some notations, let D be a spatial set (structural component) and Z a Gaussian random filed with mean μ, variance σ^2 and a stationary covariance C. We get a realization of Z at N_s locations $x_1, \ldots, x_{N_s}$, hence the random vector $\mathbf{Z} := (Z(x_1), \ldots, Z(x_{N_s}))'$ has a multivariate Gaussian distribution. We note by R the correlation matrix of vector $\mathbf{Z}$ with entries $R_{i,j} := \dfrac{C(|x_i - x_j|)}{\sigma^2} = c(|x_i - x_j|)$; $\; i, j = 1, \ldots, , N_s$, where $(x_{i+1} - x_i) = h$ is the lag distance.

There are two widely used approaches for estimating θ. One approach is the Maximum Likelihood Estimate MLE where estimates $\hat{\theta}$, $\hat{\mu}$ and $\hat{\sigma^2}$ maximize the joint probability density of $\mathbf{Z}$. Another approach relies on the Least Square Estimate (LSE) where an estimate of the model parameter (see Table 1) provides the best fit for the empirical covariance or variogram. This latter is faster

and more tractable when data are not sampled from a Gaussian or log-normal distribution. Because the knowledge of the joint probability density of $\mathbf{Z}$ is not required in the case of LSE. This method simply considers the covariance model. The natural estimate of C is the experimental covariance at the lag distance h [4, 14]:

$$\hat{C}(h) = \frac{1}{N_h} \sum_{i=1}^{N_h} \left(Z(x_i) - \hat{\mu}\right) \left(Z(x_i + h) - \hat{\mu}\right), \tag{4}$$

where $\hat{\mu} := \frac{1}{N_s} \sum_{i=1}^{N_s} Z(x_i)$ is the unbiased estimate of μ, N_h being the number of points distant with h from the considered locations. The fitted parameters from (4) are biased since $\hat{C}$ is a biased estimator of C. Sometimes, it is preferred to perform the best fitting of the experimental variogram $\hat{V}(h)$ [4] defined by:

$$\hat{V}(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} \left(Z(x_i) - Z(x_i + h)\right)^2, \tag{5}$$

The estimate $\hat{V}(h)$ is unbiased because it does not require any prior estimate of μ; this can increase the precision of parameters since they are also unbiased (for more details; see [4]). Under some regularity conditions on the continuous variogram with respect to the parameters and the spatial locations, the least square estimate converges to the Normal distribution [4, 7]. However, the covariance matrix of the asymptotic distribution depends on the eigenvalues of the unknown matrix R, so it is difficult to compute efficiently the confidence region of the parameters. Consequently, we propose to use the MLE for estimating parameters where the covariance matrix of the asymptotic distribution can be computed from the inverse of the Fisher information matrix. The Fisher information-matrix gives the amount of information that an observable random vector $\mathbf{Z}$ carries about unknown parameters θ. This information-matrix gives an asymptotic distribution of the MLE of θ. Formally, it is given by the variance of the derivative of the logarithm of the likelihood with respect to θ.

2.4. Maximum Likelihood Estimate

When data are assumed to be realizations of Gaussian or log-normal distribution, it is often preferable to construct the maximum likelihood estimate

by maximizing the likelihood of the data. The idea of the method is to construct an estimator which ensures that the probability of the observed values is maximal. Statistical tests can be performed to check whether data originate from a multivariate Normal distribution; the generalized multivariate skewness and kurtosis test is proposed by [15]. Some transformations like the Box-Cox transformation [1] can be performed to obtained Normal distributed data.

Let $\mathbf{Z}$ be a Gaussian vector discretized from a stationary Gaussian field with mean μ, variance σ^2, and correlation matrix $R(\theta)$. We note $\eta := (\mu, \sigma^2, \theta)$ and $\widehat{\eta}$ its MLE which is given by maximizing the likelihood,

$$L(\eta, \mathbf{Z}) = (2\pi\sigma^2)^{-\frac{N_s}{2}} |R|^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{Z}-\mu)' R^{-1}(\mathbf{Z}-\mu)\right); \tag{6}$$

where $|R|$ is the determinant of R and ' refers to transpose. In practice, we minimize the negative log-likelihood given by the formula:

$$\ell(\eta, \mathbf{Z}) = \frac{1}{2}\left(N_s \log(\sigma^2) + \log|R| + \frac{1}{\sigma^2}(\mathbf{Z}-\mu)' R^{-1}(\mathbf{Z}-\mu)\right) \tag{7}$$

An iterative resolution is used to solve this optimization problem according to the following steps:

1. First, we choose an initial estimate θ_0, and we compute the estimates of μ and σ^2, as follows:
   - MLE of the mean μ:
   $$\hat{\mu} = \frac{\mathbf{Z}' R^{-1}\mathbf{1}}{\mathbf{1}' R^{-1}\mathbf{1}}, \tag{8}$$
   where we note by $\mathbf{1}$ the vector with N_s entries all equal 1.
   - MLE of the variance σ^2:
   $$\hat{\sigma^2} = \frac{(\mathbf{Z}-\hat{\mu})' R^{-1}(\mathbf{Z}-\hat{\mu})}{N_s} \tag{9}$$
2. Next, we compute $\hat{\theta}$ by minimizing $\ell(\hat{\eta}, \mathbf{Z})$ knowing $\hat{\mu}$ and $\hat{\sigma^2}$ from step 1.
3. We repeat this step until convergence.

The convergence criterion is reached when two successive approximations of the scale $\hat{\theta}$ (for instance) are close to a fixed threshold.

To compute $\hat{\eta}$, some authors propose to use the Cholesky decomposition of R (at each iteration) which enables us to write $R = QQ'$ where Q is a lower triangular matrix and Q' its transposed. Thus, by performing in step 2 the transformation $\xi = \dfrac{Q^{-1}(Z-\hat{\mu})}{\hat{\sigma}}$ we have to minimize the negative log-likelihood:

$$\ell(\eta, \xi) = 2\log|Q| + \sum_{n=1}^{N_s} \xi_n^2, \tag{10}$$

Where $|Q|$ is the determinant of Q and $\log|Q| = \sum_{i=1}^{N_s} \log(Q_{i,i})$ with $(Q_{i,i})_{i=1}^{N_s}$ are diagonal entries of Q. This technique can be numerically more efficient because of the efficiency and stability of the Cholesky decomposition in comparison with others decomposition method. The MLE is more attractive and computationally feasible since it has the usual asymptotic convergence of consistency and asymptotic Normality. Under some regularity conditions and the decay of the covariance function of the underlying Gaussian field, authors in [37] show the efficiency of the method on the case of a stationary Gaussian field such that:

$$\{\widehat{\eta} - \eta\} \overset{distribution}{\longrightarrow} \mathcal{N}(0, \Sigma^{-1}); \tag{11}$$

where Σ denotes the Fisher matrix information. The inverse of the Fisher matrix is the covariance matrix of the vector η. The entries of Fisher matrix are given by the second partial derivative of the log-likelihood:

$$\Sigma_{i,j} := -\mathbb{E}\left[\frac{\partial^2 \log(L(\eta, Z))}{\partial\eta_i \partial\eta_j}\right]. \tag{12}$$

Mardia [16] extended the same result in the case of Gaussian vector with a linear regression. The convergence result (11) is very useful since it allows building the confidence regions of $\hat{\eta}$ by computing the covariance matrix of the asymptotic distribution.

Remark 2.1. Authors in [13, 12] propose to use an iterative algorithm to compute the estimate of θ by performing the transformation $\xi = Q^{-1}(Z-\mu)/\sigma$ and minimize the negative log-likelihood of ξ:

$$\ell(\eta, \xi) = \sum_{n=1}^{N_s} \xi_n^2, \tag{13}$$

However, this estimate is much different than the estimate which maximizes (10), because the likelihood in (13) approaches the true likelihood in (10) only for a small scale $\theta \approx 0$.

2.5. Asymptotic distribution of estimated parameters with MLE

The correlation matrix is assumed to depend only on the scale of fluctuation θ , so $\eta = (\mu, \sigma^2, \theta)$. The covariance matrix Σ^{-1} is computed by substituting $\hat{\eta}$ in the Fisher matrix information. The MLE of $\hat{\mu}$ is unbiased since $\mathbb{E}[\hat{\mu}] = \mu$ with minimal variance which is computed explicitly from (12) by taking $\eta_i = \eta_j = \mu$. Then, the MLE of $\hat{\mu}$ follows a Normal distribution with mean μ and variance $\text{var}[\hat{\mu}]$:

$$\text{var}[\hat{\mu}] = -\left(\mathbb{E}\left[\frac{\partial^2 \log(L(\eta, Z))}{\partial^2 \mu}\right]\right)^{-1} = \frac{\sigma^2}{\mathbf{1}R^{-1}\mathbf{1}'}. \tag{14}$$

The Fisher matrix of the estimate $(\hat{\sigma^2}, \hat{\theta})$ is explicitly computed using the classical derivative formulas $\frac{\partial \log(|R(\theta)|)}{\partial \theta} = Tr(R^{-1}\frac{\partial R}{\partial \theta})$ and $\frac{\partial R^{-1}}{\partial \theta} = -R^{-1}\frac{\partial R}{\partial \theta}R^{-1}$ where $Tr(B)$ designates the trace of a matrix B. Further, since the estimates $\hat{\sigma^2}$ and $\hat{\theta}$ are asymptotically independent, their variance are estimated by:

$$\text{var}[\hat{\theta}] \approx \left(\mathbb{E}\left[\frac{\partial^2 \log(L(\hat{\eta}, Z))}{\partial^2 \theta}\right]\right)^{-1} = \frac{\sqrt{2}q_\alpha}{\sqrt{Tr\left((R^{-1}(\hat{\theta})\partial_\theta R(\hat{\theta}))^2\right)}} \tag{15}$$

$$\text{var}[\hat{\sigma^2}] \approx \left(\mathbb{E}\left[\frac{\partial^2 \log(L(\hat{\eta}, Z))}{\partial^2 (\sigma^2)}\right]\right)^{-1} = \frac{2\sigma^4}{N_s}. \tag{16}$$

Therefore, we estimate the asymptotic law of $\hat{\theta}$ as a Gaussian distribution with mean θ and variance $\text{var}[\hat{\theta}]$. Similarly, we estimate the asymptotic law of $\hat{\sigma^2}$ as a Gaussian distribution with mean σ^2 and variance $\text{var}[\hat{\sigma^2}]$. However, from the expression of $\hat{\sigma^2}$ given in (9), this estimate of the variance is biased (only asymptotically unbiased). Thus, we consider in what follows an unbiased estimate S^2 defined by the following transformation:

$$S^2 = \frac{N_s}{N_s - 1}\hat{\sigma^2} = \frac{(Z - \hat{\mu})' R^{-1}(Z - \hat{\mu})}{N_s - 1} \tag{17}$$

By neglecting the error on the covariance matrix R we estimate the law of $\frac{N_s - 1}{\sigma^2}S^2$ as a chi-squared distribution $\chi^2(N_s - 1)$ with $N_s - 1$ degrees of free-

dom.

$$S^2 \sim \frac{\sigma^2}{N-1}\chi^2(N_s - 1) \tag{18}$$

3. Practical implementation for a target accuracy

When focusing on a practical application on a real database, the aim is to optimize (reduce the number of measurements and respect a target accuracy) the total number of measurements $N = N_p \times N_s$ where N_p is the number of repetitive tests to improve the quality of the inspection, N_s is the total number of inspected locations (Fig 1).

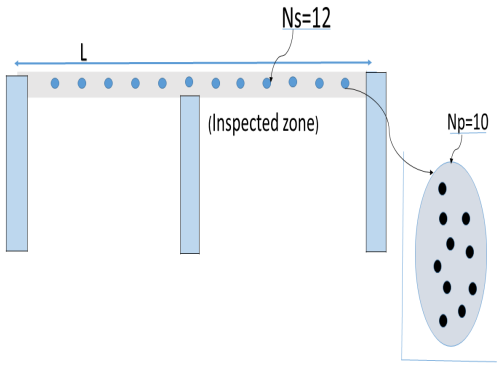

Figure 1: Example of NDT measurements

The quality of inspection can be characterized by several approaches where the consequence of measurement uncertainty on the decision is analysed by using a risk based approach [2, 27, 32]. The quality of measurements can be improved by performing reasonable repetitions of measures [29]. We assume in this paper that repetitive measurements are sufficient and the measurement uncertainty is neglected. Therefore, under the ergodicity assumption, we have to select the total number of correlated inspections N_s for one structural component. The first step consists in focusing on the assessment of the scale of fluctuation. It is estimated by making its confidence region as small as possible. For a relatively large N_s, the MLE $\hat{\theta}$ is approximated by a Gaussian vector with mean θ and variance $var[\hat{\theta}]$ given in (15). Thus, the asymptotic confidence region I_θ of θ is

given by:

$$I_\theta \approx \left[\hat{\theta} - \frac{\sqrt{2}q_\alpha}{\sqrt{Tr\left((R^{-1}\partial_\theta R)^2\right)}}; \hat{\theta} + \frac{\sqrt{2}q_\alpha}{\sqrt{Tr\left((R^{-1}\partial_\theta R)^2\right)}}\right], \tag{19}$$

where q_α denotes the α-quantile of the Normal distribution. This confidence interval I_θ satisfies approximately the probability: $\mathbb{P}(\theta \in I_\theta) \approx 1 - \alpha$.

The aim is to determine the number of measurements N_s for which the size of I_θ reaches the target accuracy level. Thus, we define the following error $\beta(\hat{\theta})$ which depends on the current estimate (MLE) of θ:

$$\beta(\hat{\theta}) = \frac{\sqrt{2}q_\alpha}{\sqrt{Tr\left((R^{-1}(\hat{\theta})\partial_\theta R(\hat{\theta}))^2\right)}} \tag{20}$$

where $\partial_\theta R(\hat{\theta})$ represents the derivative of each entries of R with respect to θ. It is approximately computed from the parametric model of the correlation by the finite difference approximation:

$$\partial_\theta R(\hat{\theta}) \approx \frac{R(\hat{\theta} + \delta\theta) - R(\hat{\theta})}{\delta\theta}, \tag{21}$$

where $\delta\theta$ is a small variation of θ (for instance $\delta\theta \approx \frac{\theta}{10^r}$ for some $r > 2$). The definition (20) of the error $\beta(\hat{\theta})$ is independent on σ^2, since σ^2 is a multiplicative parameter. Therefore, $\beta(\hat{\theta})$ is used for determining a suitable scale θ independently on the first and second moments of Z. The accuracy on the estimate is specified by the target accuracy $\epsilon > 0$ which is defined by the asset owner. It can be defined as a small percentage of the length L of the structure (for example $\epsilon \approx L/100$). Thus, we define the following constraint:

$$\beta(\hat{\theta}) \leq \epsilon. \tag{22}$$

The practical implementation follows 3 key steps:

- start with initial and small spatial locations with N_{s_0} measurements;
- evaluate $\hat{\theta}_0$ the MLE of θ and its corresponding indicator $\beta(\hat{\theta}_0)$;
- increase progressively the number of measurements until $\beta(\hat{\theta})$ satisfies the shut off conditions (22).

The final number N_{s_n} ensures that the final estimation of $\hat{\theta}_n$ evaluated with N_{s_n} measurements satisfies approximately,

$$\mathbb{P}(\hat{\theta}_n \in [\theta - \epsilon, \theta + \epsilon]) \approx 1 - \alpha. \tag{23}$$

Once the scale of fluctuation is estimated from N_{s_n} spatial measurements, the second step consists of checking the accuracy of the approximation of the mean $\hat{\mu}$ and variance $\hat{\sigma^2}$ of the field Z. The MLE of the mean follows a Gaussian distribution with mean μ and variance defined in (14). Therefore, by considering the α-quantile of a Normal distribution with order α, we exhibit the following confidence interval of the mean μ using the current value of $\hat{\theta}$:

$$I_\mu = \left[\hat{\mu} - \frac{q_\alpha S}{\sqrt{\mathbf{1}' R^{-1} \mathbf{1}}}, \hat{\mu} + \frac{q_\alpha S}{\sqrt{\mathbf{1}' R^{-1} \mathbf{1}}}\right], \tag{24}$$

where S^2 is the unbiased estimate of σ^2 in (17). The vector $\mathbf{1}$ contains N_{s_n} entries all equal to 1. Similarly, under the current estimation of θ from N_{s_n} measures, we exhibit from (18) the following confidence interval of σ^2:

$$I_{\sigma^2} = \left[\frac{(N_{s_n} - 1)S^2}{\vartheta_{1-\alpha/2}}; \frac{(N_{s_n} - 1)S^2}{\vartheta_{\alpha/2}}\right] \tag{25}$$

where $\vartheta_{1-\alpha/2}$ and $\vartheta_{\alpha/2}$ are the α-quantile of the $\chi^2(N_{s_n} - 1)$ distribution with order $1 - \alpha$ and α respectively.

In practice, depending on the quantity of interest, it is important to note that the precision on the confidence region for θ can be different from the precision on the confidence region μ. This is because the accuracy on θ can require more measurements than the accuracy on μ. Furthermore, the width of I_μ does not decrease quickly with respect to the number of measurements, since its decay depends on the ratio L/θ and levels out for a relatively large measurements N_s. Therefore, the statistical indicator in (26) which quantifies the width of the confidence region I_μ is used to revise the current number of measurements N_{s_n}. Thus, when this bound decreases numerically from N_s, a new number of

measurements $\hat{N}_s$ is provided within the condition:

$$\frac{q_\alpha S}{\sqrt{\mathbf{1}' R(\hat{\theta})^{-1}\mathbf{1}}} \leq \epsilon. \tag{26}$$

where the fitted $\hat{N}_s \times \hat{N}_s$ correlation matrix $R(\hat{\theta})$ is computed from the parameter model with the approximation of $\hat{\theta}$ provided in the first stage. From this new number of measurements, we compute the new MLE to obtain the target assessment on the parameters θ, μ and σ^2. Sometimes the statistical bound in (26) decreases slowly, or it levels out from N_s (strong correlation between measurements). In that case, the current number of measurements N_{s_n} and its corresponding MLE estimates are maintained. The flowchart of the method is illustrated in Figure 2.

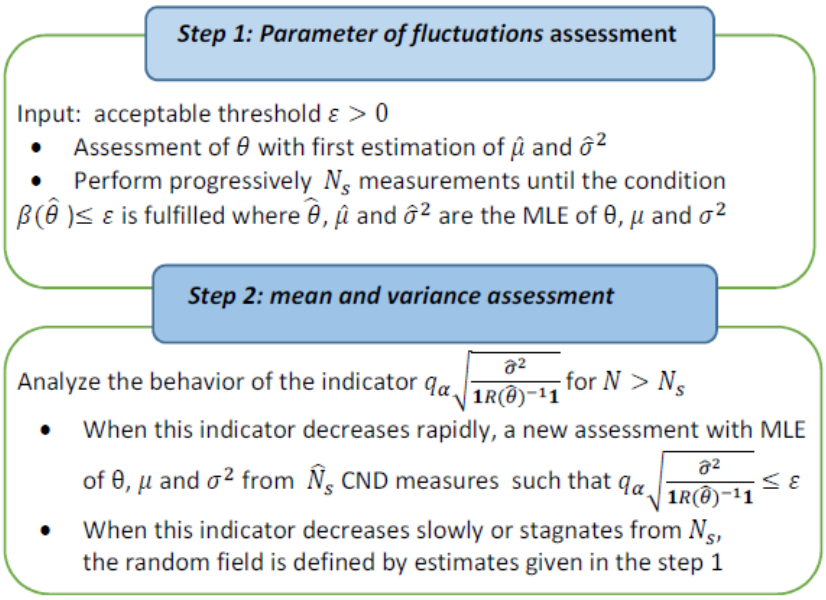

Figure 2: Flowchart of the proposed approach

4. Numerical application

In this section, we illustrate the potential of the proposed approach with a set of experimental and synthetic data. The synthetic data (subsection 4.1) are generated from simulations of a Gaussian random field Z in one and two dimensions. The data are simulated through the circulant embedding method [5, 6] which gives an exact simulation through the Fourier Decomposition of a positive circulant matrix. The field Z is generated with a stationary Matern covariance function with known regularity parameter $\nu = 1/2$ (exponential covariance) because it is widely used in the literature. Two examples of experimental data are considered. The first example (Section 4.3.1) with $N_s = 80$ measurements of water content allows analysing the error on parameters, especially, θ. In the second one (Section 4.3.2), data consists of surface chloride content and the coefficients of diffusion both are given with a small number of measurements $N_s = 16$. The accuracy is conducted using the statistical indicator (26) under the current estimation of the scale θ.

4.1. Simulated data

4.1.1. One dimensional variability

In order to illustrate the proposed adaptive approach, we consider a set of one-dimensional components (beams, indexed by spatial coordinate x) with large length L in comparison with the scale of fluctuation θ ($\frac{L}{\theta} \gg 1$). The GRF is then defined on the interval $[0, L]$ where $L = 300$; it is discretized in $N = 257$ equidistant spatial locations. The parameters of Z are $\mu = 1$, $\sigma^2 = 1$ and $\theta = 15$. The sample path of Z is then a Gaussian vector $\mathbf{Z}$ with length $N = 257$. Figure 3 (left) plots the considered path of Z, all following results are obtained from this sample path. Figure 3 (right) compares the exact stationary covariance C with the simulated one and the fitted covariance (with $N = 257$). It shows a good agreement between the exact and the fitted one. The latter is computed using parameters $\hat{\theta}$ and S^2, both estimated from the MLE of θ and σ^2; the MLE of θ is illustrated in the Figure 4 (left). The plot of the

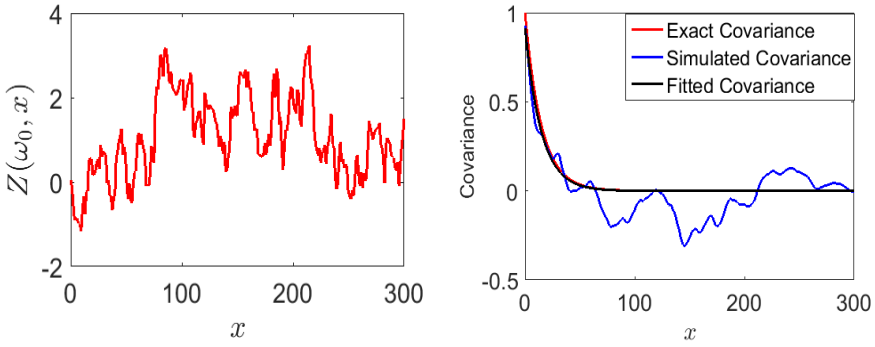

Figure 3: path of Z (left); (right) Stationary Covariance ($\theta = 15, \nu = 1/2, \mu = 1,\ \sigma^2 = 1$).

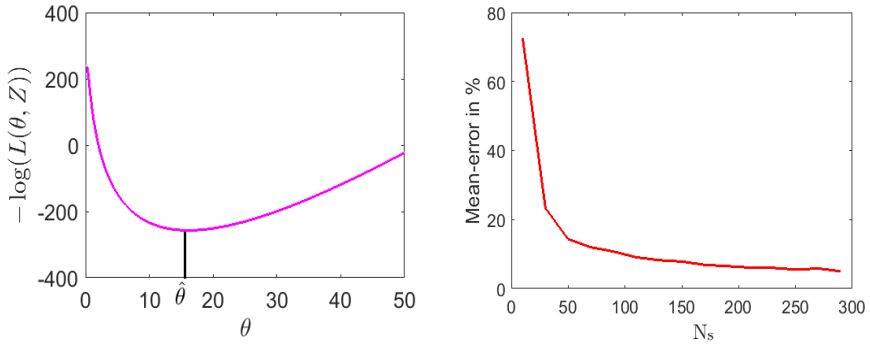

Figure 4: (left) Negative log-likelihood of Z; (right) Mean-Error estimate of the MLE .

negative logarithm of likelihood of Z as a function of θ shows a clear minimum value that approaches $\theta = 15$. The convergence of MLE is illustrated in Figure 4 (right) where the mean-error in (%) is computed as a function of N_s the number of equidistant points in the path Z: $N_s \in (20 : N)$. The mean-error is computed with 500 MC simulations of $\hat{\theta}$. It shows that the MLE converges similar to $O(\frac{1}{\sqrt{N}})\, in\, \%$. We divide $\mathbf{Z}$ on six nested vectors with several lengths $N_s = 9, 17, 33, 65, 129, 257$. First we assume that μ and σ^2 are both known in view to distinguish the effect of this knowledge on the accuracy of the MLE of θ and then on the error $\beta(\theta)$. Figure 5 (right) plots for each sample of Z the error defined by (20) where the α-quantile $q_\alpha = 1.96$ ($\alpha = 5\%$). The exact error is computed from an exact correlation matrix and its derivative with respect

to θ, while the numerical indicator $\beta(\hat{\theta})$ is computed from the fitted correlation using MLE of θ as defined in (20). The exact error is the absolute value $|\hat{\theta} - \theta|$.

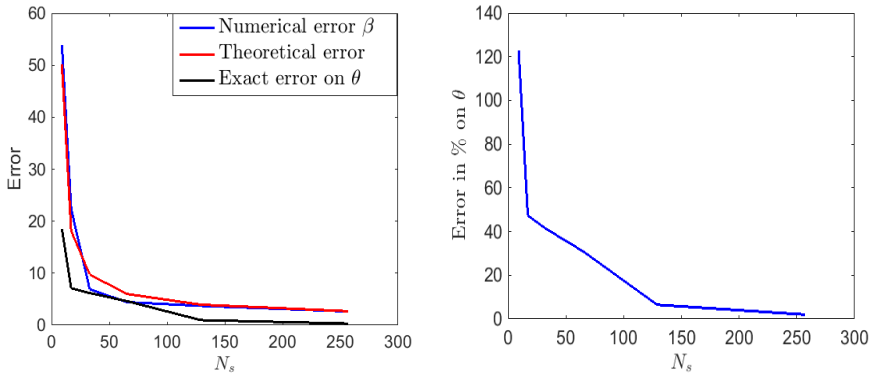

Figure 5: Error bounds on θ (left), Error in (%) on θ (right); μ, σ known.

We observe a strong similarity between the exact and numerical error indicator. It shows that indicator $\beta(\hat{\theta})$ decreases to zero when the number of spatial locations N_s increases.

Figure 5 (right) presents the error in (%) on θ using MLE with respect to N_s. It shows that the error is large for small number of measurements N_s. It exceeds 40% when N_s is less than 33 measurements. We noted that for some other numerical test cases, the numerical error is very large with a very small number of measurements $N_s < 9$; such inaccuracy is explained by a poor estimation of $\hat{\theta}$ when the fitted correlation matrix is nearly singular or ill-conditioned. Conversely, a very large number N_s does not necessarily lead to an accurate estimation of θ when the ratio $\frac{L}{\theta} \approx 2$ is small since, the correlation matrix is ill-conditioned.

Let us consider now that the mean μ and the variance σ^2 are unknown. Figure 6 (left) illustrates numerical and theoretical indicators and the error on θ when we estimate all the components of the vector $(\hat{\mu}, \hat{\sigma^2}, \hat{\theta})$ following the iterative method in Section 3. Figure 6 (right) plots the error in (%) on the scale of fluctuation θ with respect to the number of measurements N_s. It shows

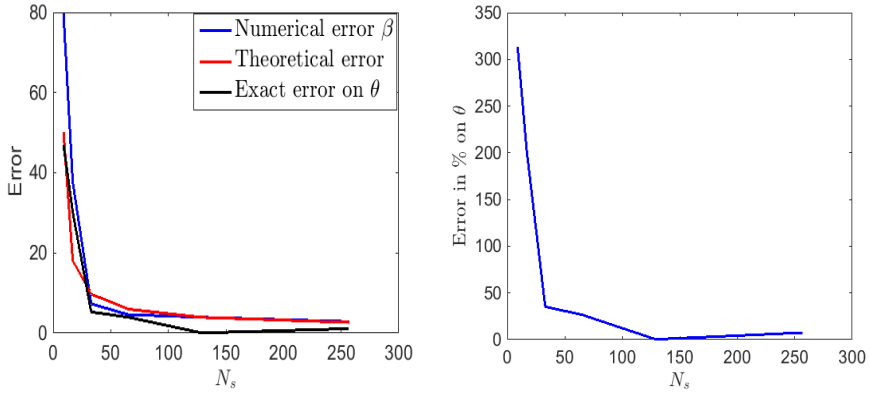

Figure 6: Error bounds (left) and error in (%) on θ (right); μ and σ^2 are unknown.

that the approximation with MLE of the mean and variance has no significant effect (except when N_s is small < 9) on the accuracy of the scale θ. The error indicator has a similar order of magnitude with the previous results where μ and σ^2 are known.

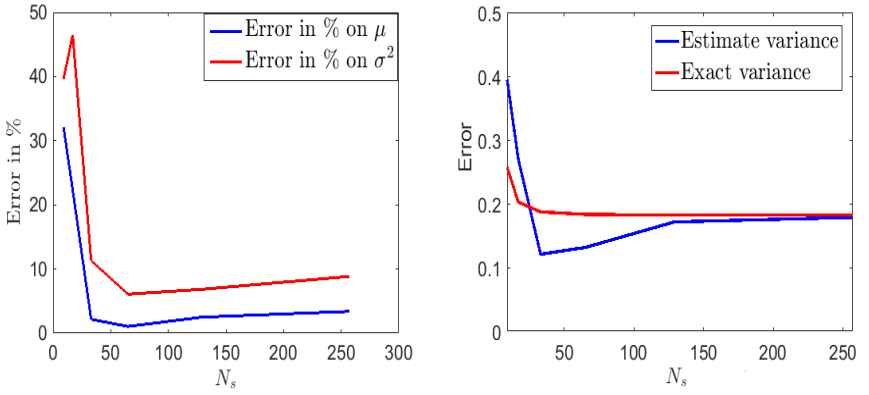

Figure 7: Error in % on μ and σ^2 (left); Upper bound on μ (right)

Figure 7 (left) presents the error in % on μ and σ^2 given by MLE. It shows that for a large number of measurements ($N_s > 100$), the error levels out for both parameters σ^2 and μ. Figure 7 (right) plots the upper bound of μ given by the upper bound of its confidence region (24) with α-quantile $q_\alpha = 1.96$. It

shows that this latter levels out as N_s becomes large. This results in a poor improvement of the precision on μ and σ^2 for $N_s > 100$; more measurements are not useful for this length L. We investigate this relationship in the next subsection.

4.1.2. Effect of the size of the structure on the accuracy

Identification of spatial random field from a set of NDT measurements usually relies on spatial average. It is in fact difficult to find several structural components with the same spatial variability and then, getting statistically independent events is rare. Using spatial average yields to an accurate estimate when the ergodic property is satisfied; thus, a constraint is required on the dimension of structure. An accuracy on the mean and variance can be reached when the variance of the estimates $\hat{\mu}$ and $\hat{\sigma^2}$ are as small as possible within a target confidence. This accuracy is reached when the upper bound in (26) satisfies the target precision. However, this bound decreases very slowly with respect to the number of measurements. Thus in order to analyse its behavior, we introduce the following simple upper bound of the variance of $\hat{\mu}$ (MLE of μ):

$$\hat{s} = \sigma^2 \sum_{i,j=1}^{N_s} \frac{R_{i,j}}{N_s^2}, \tag{27}$$

where R is the correlation matrix. The positive quantity $\hat{s}$ is the variance of the estimator $\dfrac{\sum_{j=1}^{N_s} Z_j}{N_s}$, the unbiased estimate of the mean μ. On the other hand, $\hat{s}$ is seen as a discrete formula of the integral in (3). Therefore, the condition $\hat{s} \approx 0$, which implies an accuracy on the spatial statistics, can hold when the ratio $\frac{L}{\theta}$ is large. Otherwise, when the ratio $\frac{L}{\theta}$ is small, the numerator $\sum_{i,j=1}^{N_s} R_{i,j}$ in (27) has a behavior similar to $O(N^2)$, then $\hat{s} \approx O(1)$, that explains why the parameters of the field cannot be accurately estimated when $\frac{L}{\theta}$ is small. Figure 8 compares the variance of $\hat{\mu}$ from (14) with $\hat{s}$ defined in (27) where the fitted correlation matrix R is computed for both upper bounds with MLE $\hat{\theta} = 14.03$. The figure confirms that MLE has a minimal variance and shows that this bound decreases very slowly when $N_s > 100$ which explains the stagnation of

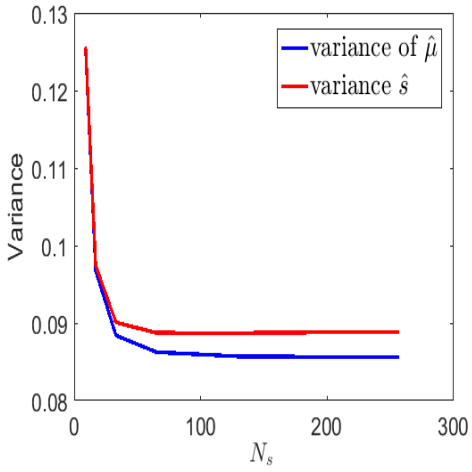

Figure 8: Variance of $\hat{\mu}$ and $\hat{s}$.

the error on μ and σ^2. Indeed, the numerator $\sum_{i,j=1}^{N_s} R_{i,j}$ in (27) increases when N_s increases because the presence of large correlated points. Physically, this stagnation means that there are zones with almost similar properties in the average. To illustrate the effect of the size of the component on the inference

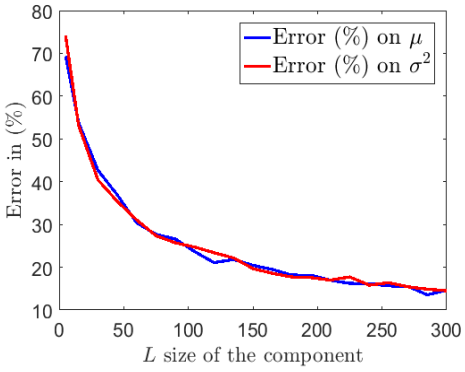

Figure 9: Error on μ and σ^2 as function of L

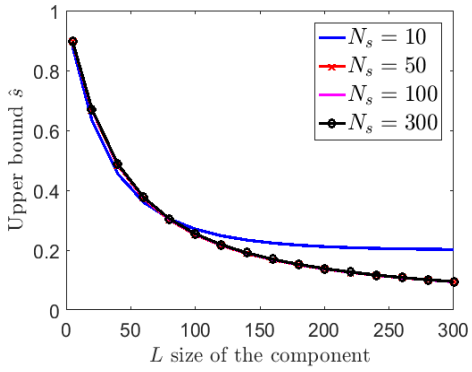

Figure 10: behavior of the bound $\hat{s}$.

accuracy, we consider the previous Gaussian random field described above in 4.1.1. discretized on components of size $5m$ to $300m$. Figure 9 illustrates the effect of L on the estimated parameters μ and σ^2 where the number of locations

is $N_s = 129$. It shows that the error decreases (in the mean sense) with respect to the size of the structure for the same number of measurements N_s. Like in Section 4.1, we note that for a fixed length L, a large number of N_s does not increase the precision on μ and σ^2 because the numerator increases as N_s increases in (27); when the size L is large the numerator is small since R contains entries close to zero. Figure 10 plots four curves of the upper bound $\hat{s}$ computed with four numbers $N_s = 10, 50, 100, 300$. It illustrates how the upper bound $\hat{s}$ of the variance of $\hat{\mu}$ decreases strongly with L. It is also important to note that the upper bound $\hat{s}$ can blow up either if L is small or if σ^2 is too large.

4.2. Two dimensional variability

4.2.1. Isotropic example

The Gaussian random field $Z(x, y)$ is herein defined on the rectangle $[0, L_1] \times [0, L_2]$ where $L_1 = 200, L_2 = 80$ with mean $\mu = 1$, unit variance $\sigma^2 = 1$ and the scale of fluctuation $\theta = 10$. The field Z is discretized on $N = N_1 \times N_2$ equidistant spatial locations with $N_1 = N_2 = 65$ and defined by an isotropic and non-separable exponential covariance (Matèrn covariance with $\nu = 1/2$).

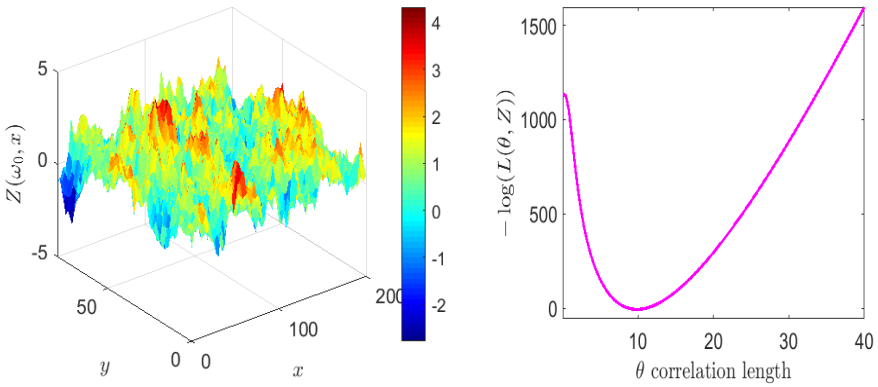

Figure 11: Path of Z ($\theta = 10, \nu = 1/2, \mu = 1$ and $\sigma^2 = 1$), (left); Negative log-likelihood of Z (right).

The sample path of the field Z is given by a Gaussian matrix $\mathbf{Z}$ with size $N = 65^2 = 4225$. We extract from the matrix $\mathbf{Z}$ five nested matrix with sizes $N_s =$

$5^2, 9^2, 17^2, 33^2, 65^2$. Each matrix represents a realisation of Z on N_s equidistant points of rectangle $[0, L_1] \times [0, L_2]$. Figure 11 (left) plots the considered path of Z. Figure 11 (right) plots the curve of the negative logarithm of the likelihood of Z as a function of θ given in (10). Figure 12 (right) presents the error in (%)

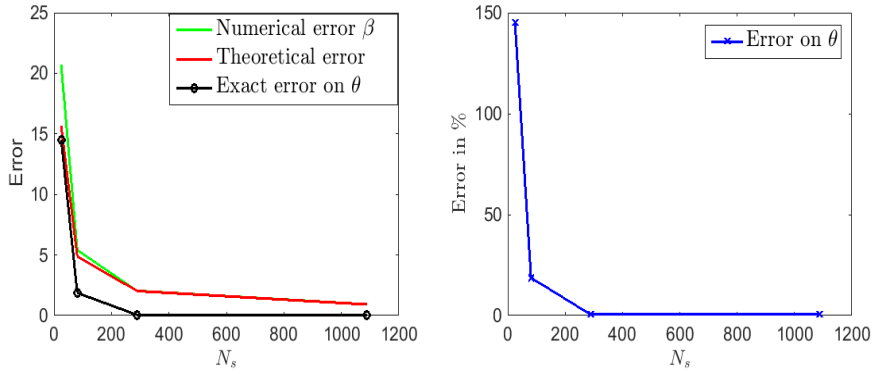

Figure 12: Error bounds and error in (%) on θ.

on θ using the MLE with respect to N_s. It illustrates the decreasing of such error to zero when the number of spatial locations N_s increases. Figure 12 (left) plots for each of the 4 sample paths the error defined by (20) for the α-quantile $q_\alpha = 1.96$. It shows a strong similarity between the exact and numerical errors. The numerical indicator shows that the indicator $\beta(\hat{\theta})$ decreases to zero when the number of spatial locations N_s increases. Thus, it is considered a practical bound of the error on θ. Note that the error $\beta(\hat{\theta})$ in Figure 12 (left) is very large when the sample path of Z has the size $N_s < 5^2$ since the moments μ and σ^2 are badly estimated and the correlation matrix is nearly singular. Figure 13 plots the error in % on μ and σ^2 given by MLE. Figure 14 plots the upper bound (26) of the error on μ. As in one dimensional case, this bound of the error decreases very slowly when $N_s > 400$. Thus, when the statistical error levels out as $N > N_s$ and satisfies the condition (26), the current number of measurements N_s that gives $\hat{\theta}$ on the first stage also provides an accurate estimation of μ and σ^2.

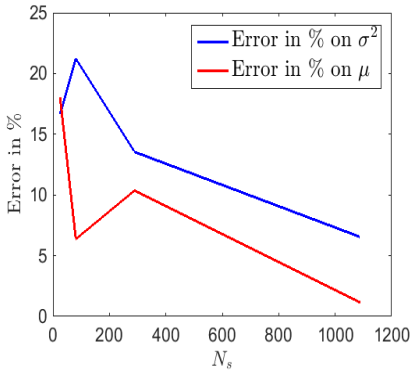

Figure 13: Error in % on μ and σ^2.

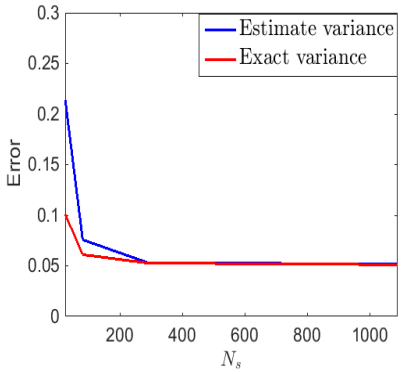

Figure 14: Upper bound of μ.

4.2.2. Anisotropic example

Here, the Gaussian random field Z is anisotropic (orthotropic) with the same characteristics as the stochastic field described above in 4.2.1 on $N = N_1 \times N_2$ equidistant spatial positions with $N_1 = N_2 = 33^2 = 1089$. The scale of fluctuations in x-direction is $\theta_x = 10$ and $\theta_y = 3$ in y-direction. The upper bounds of θ_x and θ_y are computed as in (20) respectively by:

$$\beta_x(\hat{\theta}_x) = \frac{\sqrt{2}q_\alpha}{\sqrt{Tr\left((R^{-1}(\hat{\theta})\partial_{\theta_x}R(\hat{\theta}))^2\right)}} \tag{28}$$

$$\beta_x(\hat{\theta}_y) = \frac{\sqrt{2}q_\alpha}{\sqrt{Tr\left((R^{-1}(\hat{\theta})\partial_{\theta_y}R(\hat{\theta}))^2\right)}}; \tag{29}$$

where $\partial_{\theta_x}R(\hat{\theta}))$ and $\partial_{\theta_x}R(\hat{\theta}))$ are the derivative of the matrix $R(\theta)$ with respect to the parameter θ_x and θ_y.

Figure 15 illustrates numerical and theoretical error indicator for $\theta = (\theta_x, \theta_y)$. It shows that the error is large when the number of measurements N_s is small ($N_s = 9$ or $N_s = 25$) and both errors in (28, 29) decrease to zero when $N_s > 100$. Thus, to determine a suitable number of measurements with NDT protocol according to a target precision; we can use the precision $\epsilon \approx L/100$ where L is the length of the structure.

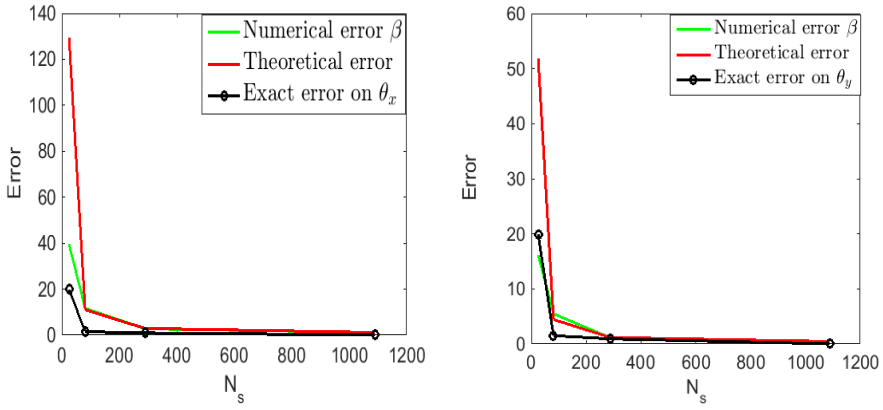

Figure 15: Error in (%) on θ_x (left) and θ_y (right).

4.3. Validation of the method based on an experimental data base

4.3.1. Water content data base

The objective of this example is to illustrate the proposed methodology using CND measurements. They were performed by a capacitive technique for quantifying the water content W (in %) of a reinforced concrete beam exposed to natural environmental conditions in Bouguenais (region of Pays de la Loire, France). These data were analysed by Schoefs et al [29] without focusing on the accuracy on the estimation of the scale of fluctuation. The set of measurements was performed on $N_s = 80$ horizontal and equidistant spatial locations of measurements along the beam with length $L = 16m$. At each location, $N_p = 30$ repetitions of measures were carried out in view to neglect and reduce the epistemic measurement uncertainties (see [29] for more details). Therefore, the final trajectory W is obtained by taking the expectation over all N_p unbiased measurements at each location. Figure 16 plots the considered path of the water content. Here, the water content is modelled by a second order random field with log-normal marginal distribution. Therefore, its logarithm $Z = \log(W)$ is seen as a Gaussian random field. The MLE of the parameters for Z, are given by; $\hat{\mu}_Z = 2.24$, $\hat{\sigma}_Z^2 = 0.009$ and $\hat{\theta} = 0.38$. From these estimates, estimates of

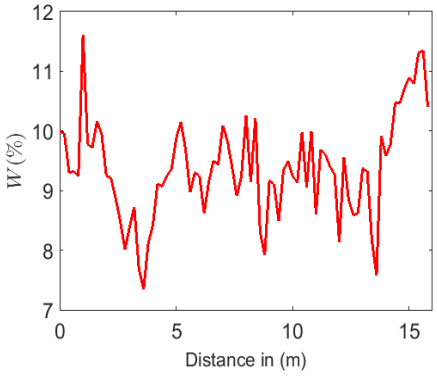

Figure 16: Trajectory of $W(\%)$

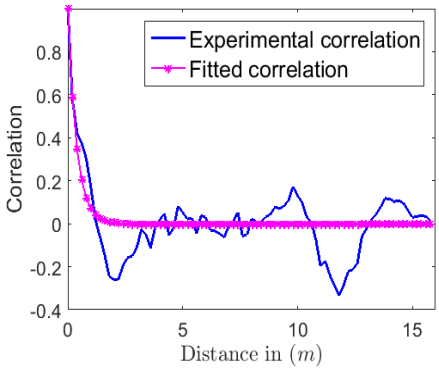

Figure 17: Experimental and fitted correlation of $W(\%)$

the mean and variance of the field W are:

$$\begin{aligned} \hat{\mu}_W &= \exp(\hat{\mu}_Z + \hat{\sigma}_Z^2/2) \\ \hat{\sigma}_W^2 &= \exp(2\hat{\mu}_Z + \hat{\sigma}_Z^2)(\exp(\hat{\sigma}_Z^2) - 1) \end{aligned}$$

The uncertainty on the assessment of θ comes from its estimation with MLE performed with limited number of measurements N_s. The confidence region of θ quantifies this error. The latter can be reduced by using a suitable set of measurements.

In order to illustrate the proposed adaptive method, we consider various nested sets with several lengths $N_s = 4, 9, 17, 33, 65$. For each set of measurements, we estimate θ with MLE and we compute the error from equation (20) where $q_\alpha = 1.96$ ($\alpha = 5\%$) and $q_\alpha = 1.645$ ($\alpha = 10\%$). Figure 18 illustrates the confidence region of the scale θ for these two quantiles. It shows that the indicator $\beta(\hat{\theta})$ decreases regardless to the number of measurements N_s and the quantile q_α.

Table 2 gives the MLE of $(\theta, \mu_Z, \sigma_Z^2)$ with respect to N_s. It shows how the relative upper bound $\beta(\theta)/L$ decreases with N_s. At the given relative precision level $\epsilon \approx 0.04L$, the needed number of inspections is $N_s = 17$; and it is $N_s = 33$ if we set the relative precision level $\epsilon \approx 0.0067L$. Further, these two examples

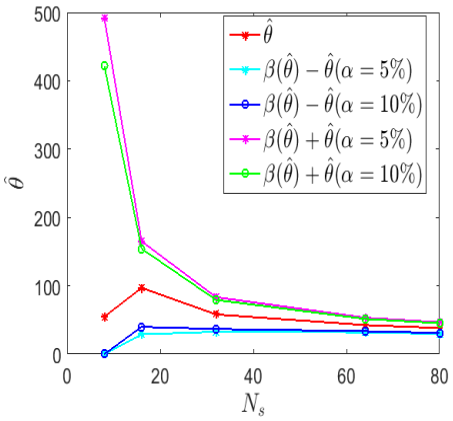

Figure 18: Confidence region of θ

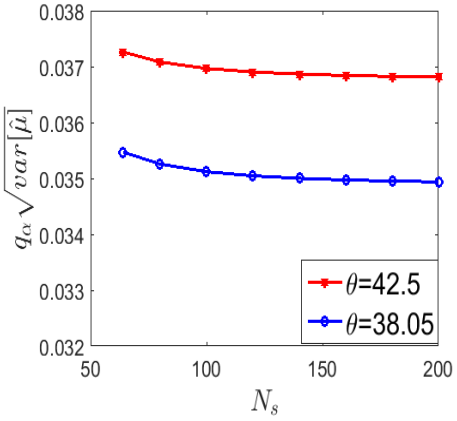

Figure 19: Evolution of $q_\alpha \sqrt{var[\hat{\mu}_Z]}$ with N_s.

N_s	9	17	33	65	80
$\beta(\theta)/L$	0.2732	0.0426	0.0158	0.0067	0.0053
$\hat{\theta}$ (m)	0.96	0.54	0.57	0.42	0.38
$\hat{\mu}_Z$	2.237	2.256	2.242	2.236	2.236
$\hat{\sigma}_Z^2$	0.007	0.091	0.01	0.009	0.009

Table 2: Relative error $\beta(\theta)/L$ and MLE of $(\theta, \mu_Z, \sigma_Z^2)$

show that by considering a convenient relative level ϵ, the adaptive NDT allows saving a considerable cost of inspection per component.

The accuracy on μ and σ^2 levels out when the number N_s increases. This level is illustrated in Figure 19 by the statistical indicator in (26). First, it decreases and then levels out as the number of measurements increases for both estimation of the scale $\hat{\theta} = 42.5$ (value estimated from $N_s = 64$) and $\hat{\theta} = 38.05$ (value estimated from the whole data). Thus, performing more inspections does not improve the accuracy of μ and σ^2 estimates.

Figure 20 plots one simulation of the water content on $N = 245$ positions which coincides with the measured path in Figure (16) on the same positions. We use this path to predict the behaviour of the relative error $\beta(\theta)/L$ for $N_s >$

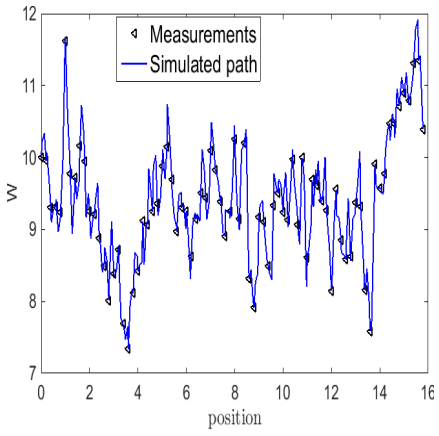

Figure 20: Simulated W conditioning on the measured path

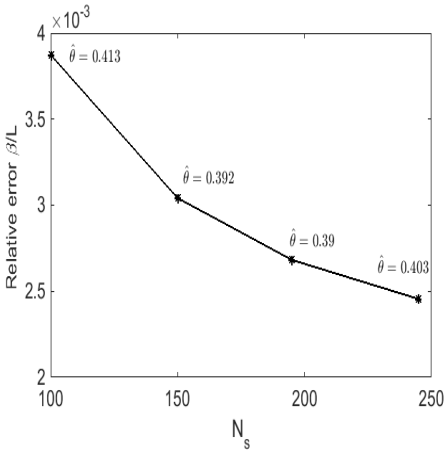

Figure 21: Relative error on θ

80. Figure 21 shows that this error decreases gradually with respect to N_s. However, it decreases much slower than the decay given in Table 2.

The target relative accuracy level is the key to selecting the required number of measurements. It is defined by an asset owner or an expert judgment, depending on the length of the structure and how it affects the reliability assessment [34].

We note that for a short illustration, the behaviour of the relative error was estimated in Table 2 on the nested grid of positions with numbers of measurements ($N_s = 9, 17, 33, 65$). In practice, we start with a small grid and add positions of measurements for refining the grid by analysing the relative error at each position.

4.3.2. Chloride content measurements data base

This data set consists of two paths of the main parameters used in modelling chloride-induced corrosion for reinforced concrete; i.e., the surface chloride content C_s and the apparent diffusion coefficient D_{app}. Both parameters are computed through the curve fitting method from Fick's second law of diffusion using experimental chloride profile [21, 30]. The set of measurements was performed

using a semi-destructive test along the beam with length $L = 9m$ (region of Pays de la Loire, France).

Position (m)	$C_s \times 10^{-4}$)	$D_{app} \times 10^{-12}$
0.2	44.5836	1.3611
0.5	65.1834	1.1616
1.1	74.2033	1.1688
1.4	41.3765	3.6419
2.9	44.5428	2.0330
3.8	74.6240	0.9099
4.1	77.8895	1.0631
4.7	54.9014	1.4133
5.3	52.9310	1.3102
6.2	33.9131	0.7662
6.8	67.2010	0.6475
7.1	34.3982	0.6831
7.4	44.4848	1.5020
8.0	38.4594	2.5143
8.3	36.6141	1.5706
8.9	36.6042	1.1876

Table 3: Estimate of C_s and D_{app} by curve fitting method from Fick's second law.

N_s	$var[\log(C_s)]$	$var[\log(D_{app})]$
$N_s = 2$	0.0004	0.0074
$N_s = 4$	0.0297	0.0056
$N_s = 8$	0.0991	0.0681
$N_s = 16$	0.1967	0.0818

N_s	$\mathbb{E}[\log(C_s)]$	$\mathbb{E}[\log(D_{app})]$
$N_s = 2$	-27.3418	-5.3272
$N_s = 4$	-27.2073	-5.3709
$N_s = 8$	-27.3899	-5.2318
$N_s = 16$	-27.3576	-5.3167

Table 4: Variance and mean estimates of $\log(C_s)$ and $\log(D_{app})$

Table 3 gives the considered path of C_s and D_{app} on $N_s = 16$ positions. Both parameters are modelled by a second order random field with log-normal marginal distribution and both are defined by the exponential correlation. Table 4 gives estimates of the mean and variance of both fields $\log(C_s)$ and $\log(D_{app})$ with four nested grid points. The first stage of the inspection methodology is devoted to the assessment of θ on the whole path of C_s and D_{app}. Using the MLE, we obtain $\hat{\theta}_{C_s} = 18.65\,cm$ as an estimate of the scale θ for the random field $log(C_s)$ and $\hat{\theta}_{D_{app}} = 29.42\,cm$ an estimate of θ for $log(D_{app})$. The error on the assessment of θ is quantified according to (20). For these random fields, we get the relative error estimate: $\beta_{C_s}/L = 0.038$ and $\beta_{D_{app}}/L = 0.04$. In this case

of small number of measured values ($N \leq 16$), the error on θ can be inaccurate because the normal distribution of the estimate $\hat{\theta}$ is only asymptotic. Therefore, the accuracy level will be conducted using the statistical indicator (26) under an estimate of θ and ignoring the error on θ.

The statistical indicator (26) quantifies the accuracy on the mean and partially on the variance of the random field. Figure 22 (left) illustrates the behaviour of such indicator for four nested grid of positions ($Ns = 2, 4, 8, 16$) where $q_\alpha = 1.96$ ($\alpha = 5\%$) for both parameters $log(C_s)$ and $log(D_{app})$. Figure

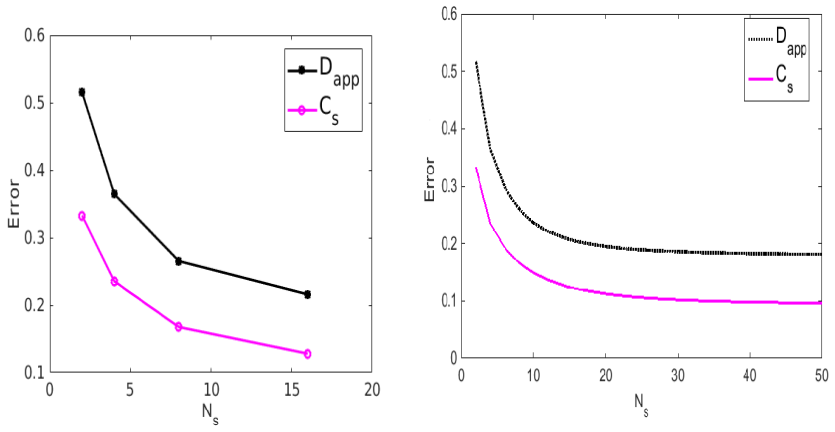

Figure 22: Evolution of the errors with N_s.

22 (right) illustrates the numerical behaviour of the statistical indicator (26) as a function of N_s, where N_s is the number of equidistant positions in the interval $(0, 9\,m)$. It shows that this statistical indicator declines gradually from $N_s > 8$ and it levels out for both random fields when $N_s > 16$. Thus, including further inspections slowly improves the accuracy of μ and variance σ^2. For example, in Figure 22 (left) at the given precision level $\epsilon = 0.265$, ($\epsilon \approx 0.97\%|\mathbb{E}[\log(C_s)]|$ and $\epsilon \approx 5\%\mathbb{E}[\log(D)]$), the required minimal number of inspections is $N_s = 8$, while the spatial variability of $log(C_s)$ evaluated with $N_s = 16$ measures ensures the accuracy level $\epsilon = 0.215$.

5. Conclusion

This paper proposes an approach for characterizing the spatial variability of structural components through discrete and limited measurements. It allows to assesses the parameters of a stationary random field given a target accuracy on the estimates. The adaptive approach is based on two errors to determine the spatial correlation and the statistics of the model. The method relies on two steps. First, an error is used to estimate an accurate scale of fluctuation of the Gaussian random field using the Maximum Likelihood Estimate. Second, the accuracy of the estimated moments of the field with the current number of measurements is revised by analysing the decay of the statistical indicator within a target precision level. The advantage of this approach is that the spatial correlation is not neglected when assessing mean and variance of the model, and the range of the correlation is properly computed within a target accuracy. Further, the statistical errors on the parameters estimation are perfectly known. Therefore, our contribution offers a new approach to carry out assessment of structural condition state with a good cost/benefit ratio.

For future work, it is appropriate to improve this adaptive method including a local adaptation and considering a non-stationary model with trend or variance-stationarity and piece-wise stationary fields.

Acknowledgements

The work reported in this article was supported by the Region Pays de la Loire de France through the regional project SI3M (Strategies of Identification of Meta-models for Structural Maintenance). The authors would like to thank the region for the support.

References

[1] D. F. Andrews, *A note on the selection of data transformations*, Biometrika; 58, 249-254, 1971.

[2] Berens AP, *NDE reliability data analysis.* In: Metals handbook, vol. 17. American Society for Metals; p. 689-701, 1978.

[3] Bonnet S, Schoefs F, Ricardo J, Salta M., *Effect of error measurement of chloride profiles on reliability assessment. The 10th International Conference on Structural Safety and Reliability*, Osaka, Japan, 2009.

[4] Cressie N, . *Statistics for Spatial Data*, Revised Edition, John Wiley & Sons, New York; 1999.

[5] Chilès, J.-P., and Delfiner, P., *Geostatistics: Modeling Spatial Uncertainty*, New York: Wiley; 1999.

[6] Dietrich, C. et G. Newsam, *Fast and exact simulation of stationary Gaussian processes throught circulant embedding of the covariance matrix.* SIAM J. SCI. COMPUT. Vol. 18, No. 4, pp. 1088-1107; 1997.

[7] Gaetan. C, · X. Guyon, *Spatial Statistics and Modeling.* Springer Series in Statistics, New York; 2010.

[8] Gomez-Cardenas C, Sbartaï ZM, Balayssac JP, Garnier V, Breysse D., *New optimization algorithm for optimal spatial sampling during non-destructive testing of concrete structures.* Eng Struct; 88:92-99; 2015.

[9] Griffiths DV, Fenton GA, *Influence of soil strength spatial variability on the stability of an undrained clay slope by finite elements.* Slope Stability, ASCE 184-193; 2000.

[10] Guttorp, P. and Gneiting, T. *Studies in the history of probability and statistics. XLIX. On the Mat érn correlation family.* Biometrika 93 989-995; 2006.

[11] Kenshel O, O'Connor A. *Assessing chloride induced deterioration in condition and safety of concrete structures in marine environments.* European Journal of Environmental and Civil Engineering 13, 593-613; 2009.

[12] Kenshel O. *Influence of spatial variability on whole life management of reinforced concrete bridges.* University of Dublin, Trinity College, Dublin, Ireland; 2009.

[13] Li Y. *Effect of spatial variability on maintenance and repair decisions for concrete structures.* Delft University; Netherlands. 2004.

[14] Matheron, G. *"The intrinsic random function and their applications".* Advances in Applied Probability 5, 439-468; 1973.

[15] Mardia, K. V. *Measures of multivariate skewness and kurtosis with applications.* Biometrika, 57, 519-530; 1970.

[16] K. V. Mardia, R. J. Marshall, *Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression.* Biometrika, 71, No. 1, 135-146; 1984.

[17] Malioka, V. and Faber, M.H. *Modeling of the Spatial Variability for Concrete Structures.* Proceedings, Second International Conference on Bridge Maintenance, Safety and Management, IABMAS 2004, Kyoto, Japan, October 18-22, 2004, pp. 825-826.

[18] Malioka, V. and Faber, M.H. *Condition Indicators for Inspection and Maintenance Planning* Proceedings, 9th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP9, San Francisco, USA, July 6-9, 2003, Volume 2, pp. 1101-1108.

[19] Nguyen NC, Sbartaï ZM, Lataste JF, Breysse D, Bos F. *Assessing the spatial variability of concrete structures using NDT techniques-laboratory tests and case study.* Constr Build Mater, 9:240-50; 2013.

[20] O'Connor A, Kenshel O. *Experimental Evaluation of the Scale of Fluctuation for Spatial Variability Modeling of Chloride-Induced Reinforced Concrete Corrosion.* Journal Of Bridge Engineering; 18:3-14, 2013.

[21] I. Othmen, S. Bonnet, F. Schoefs, *Statistical investigation of different analysis methods for chloride profiles within a real structure in a marine environment*, Ocean Engineering 157 (1) (2018) 96-107 (June 2018).

[22] M. Oumouni, F. Schoefs, B. Castanier. *Spatio-Temporal modelling of degradation processes through stochastic Gamma and Gaussian processes.* Safety and Reliability – Theory and Applications. CRC Press Taylor & Francis Group; 120–130; 2017.

[23] M. Oumouni, F. Schoefs, B. Castanier. *Modeling time and spatial variability of degradation through gamma processes for structural reliability assessment*, Structural Safety, V 76, January 2019, Pages 162-173.

[24] M. Oumouni, F. Schoefs. *A Perturbed Markovian process with state-dependent increments and measurement uncertainty in degradation modeling*, Comp.-Aided Civil and Infrastruct. Engineering; under revision 2020.

[25] Papakonstantinou KG, Shinozuka M. *Spatial stochastic direct and inverse analysis for the extent of damage in deteriorated RC structures.* J Comput Struct; 128, 286-296; 2013.

[26] M. I. Rafiq, M. K. Chryssanthopoulos, and T. Onoufriou, *Performance updating of concrete bridges using proactive health monitoring methods,* Reliability Engineering and System Safety, vol. 86, no. 3, pp. 247-256, 2004.

[27] Rouhan A, Schoefs F. *Probabilistic modeling of inspection results for offshore structures.* Struct Saf, 25:379-99; 2003.

[28] Schoefs F, Clement A, Nouy A. *Assessment of spatially dependent ROC curves for inspection of random fields of defects.* Struct Saf; 31:409-19, 2009.

[29] F. Schoefs, E. Bastidas-Arteaga, T.V. Tran, G. Villain, X. Derobert. *Characterization of random fields from NDT measurements: A two stages procedure.* Eng Struct 111; 312-322; 2016.

[30] F. Schoefs, M. Oumouni, R. Clerc, I. Othmen, S. BONNET, *Statistical analysis and probabilistic modeling of chloride ingress spatial variability in concrete coastal infrastructures*, in: EDITION 4, SPLIT, Croatia, Coastal and Maritime Mediterranean Conference, 2017.

[31] B. Shafei, A. Alipour, M. Shinozuka*A Stochastic Computational Framework to Investigate the Initial Stage of Corrosion in Reinforced Concrete Superstructures*; Comp.-Aided Civil and Infrastruct. Engineering; Volume 28, 482-494; 2013.

[32] Sheils E, O'Connor A, Breysse D, Schoefs F, Yotte S. *Development of a two stage inspection process for the assessment of deteriorating infrastructure*. Reliab Eng Syst Saf;95:182-94; 2010.

[33] Stewart MG, Mullard JA. *Spatial time-dependent reliability analysis of corrosion damage and the timing of first repair for RC structures*. Eng. Struct.; 29:1457-64; 2007.

[34] Stewart MG, Suo Q. *Extent of spatially variable corrosion damage as an indicator of strength and time-dependent reliability of RC beams*. Eng Struct, 31: 198-207; 2009.

[35] M. G. Stewart, M. H. Faber, and C. Gehlen, *Temporal and spatial aspects of probabilistic corrosion models,* in Proceedings of the 3rd International IABMAS Workshop on Life-cYcle Cost Analysis And Design of Civil Infrastructure Systems and The JCSS Workshop on Probabilistic Modelling of Deterioration Processes in Concrete Structures, vol. 20.

[36] Straub D., Malioka V., Faber M.H. *A framework for the asset integrity management of large deteriorating concrete structures*. Structure and Infrastructure Engineering, 5(3) pp. 199 - 213. 2009.

[37] Sweeting, T.J. *Uniform asymptotic normality of the maximum likelihood estimator*. Annals of Statistics 8, 1375-1381; 1980.

[38] Pasqualini O., Schoefs F., Chevreuil M., Cazuguel M., *Measurements and statistical analysis of fillet welded joints geometrical parameters for probabilistic modelling of the fatigue capacity*, Marine Structures, 34/dec. 2013, 226-248, doi: 10.1016/j.marstruc.2013.10.002

[39] Vanmarcke, E. H., 1998. *Random Fields: Analysis and Synthesis*. Cambridge, MA: The MIT Press. 1998.