A Condition-Based Deterioration Model for the Stochastic Dependency of Corrosion Rate and Crack Propagation in Corroded Concrete Structures

Boutros El Hajj, Franck Schoefs, Bruno Castanier, Thomas G. Yeung

A Condition-Based Deterioration Model for the Stochastic Dependency of Corrosion Rate and Crack Propagation in Corroded Concrete Structures

Boutros El Hajj & Franck Schoefs

University of Nantes, CNRS, GeM, Institute for Research in Civil and Mechanical Engineering, CNRS, Nantes, France

Bruno Castanier

Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS)-EA CNRS 7315, LUNAM Université, University of Angers, Angers, France

&

Thomas Yeung

Research Institute of Communication and Cybernetics of Nantes (IRCCyN) - UMR CNRS 6597, LUNAM Université, École des Mines of Nantes, Nantes, France

Abstract: *Physics-based models are intensively studied in mechanical and civil engineering but their constant increase in complexity makes them harder to use in a maintenance context, especially when degradation model can/should be updated from new inspection data. On the other hand, Markovian cumulative damage approaches such as Gamma processes seem promising; however, they suffer from lack of acceptability by the civil engineering community due to poor physics considerations. In this article, we want to promote an approach for modeling the degradation of structures and infrastructures for maintenance purposes which can be seen as an intermediate approach between physical models and probabilistic models. A new statistical, data-driven state-dependent model is proposed. The construction of the degradation model will be discussed within an application to the cracking of concrete due to chloride-induced corrosion. Numerical experiments will later be conducted to identify preliminary properties of the model in terms of statistical inferences. An estimation algorithm is proposed to estimate the parameters of the model in cases where databases suffer from irregularities.*

*To whom correspondence should be addressed. E-mail: *franck. schoefs@univ-nantes.fr*.

1 INTRODUCTION

Throughout history, the success and progress of society and mankind were always secured by the ability to communicate, produce, and exchange goods and knowledge. Nowadays, we call it infrastructure: roads, power lines, ports, dams, bridges, wind farms, etc.

All infrastructures suffer from degradation and need to be maintained. Taking into consideration the consequences of failure, maintenance has become a major challenge with economic and safety issues. In Europe, a large number of infrastructures were built after the Second World War; thus, it is very common to find structures requiring repairs and rehabilitation. Some repaired structures exhibit poor repair performance and need to be assessed and given a reliable safety level. A third of the steel structures in the Atlantic area were built more than 100 years ago (Duratinet; Boéro et al., 2009).

Furthermore, the search for a substitute source of energy, motivated research to develop new means of production as well as new locations for this production (e.g., exploitation wharves and offshore wind farms). One of the most targeted sites in maritime countries, like France, is coastal areas where exist very aggressive environmental conditions for materials and for inspections to be carried out easily. In addition, the investment cost of these means of production makes the costs of exploitation of this energy not profitable compared to the cost of nuclear energy, for example.

Even though currently we dispose of developed inspection and repairing techniques to control structures on one hand, and of highly complex mathematical models that aims at decision making and prognostics on the other hand, the ability to use all of these advancements consistently for one maintenance problem remains extremely hard for both mathematicians and engineers.

This is mainly because maintenance decisions are complex problems affected by many uncertain factors such as material properties, influenced-environment factors (behavior and impacts on the structure; Bastidas-Arteaga et al., 2012), quality of inspection (e.g., imperfect inspections; Sheils et al., 2010), and risk expressed with the degradation level (O'Connor et al., 2013; O'Connor and Kenshel, 2013). In this study, we deal with the latter one.

A literature review of degradation models of structures and infrastructure in reliability and maintenance contexts shows two major trends (Frangopol et al., 2004): on one hand, we have models based on the simulation of the physical laws of degradations; and on the other hand, we have probabilistic models based on statistical quantities (e.g., time-life models where the relation between time and failure is described).

In maintenance contexts, the quality of a model is no longer limited to its ability of modeling the pathology, but also for its quality to predict future performances of the structure, its ability to integrate new observations (especially from non-destructive testing [NDT]), and its ability to be implemented in complex dynamic decision platforms. Moreover, some authors have strongly stated the importance of incorporating physical meanings into statistical models, especially for data-driven models that are frequently used in condition-based maintenance (Si et al., 2011).

Since the end of the 1990s, physics-based degradation models have increasingly become complex, incorporating physical–chemical and mechanical couplings, leading to an explosion in the number of parameters: for example, it is common nowadays to find a 10-parameters deterministic model of chloride ions propagation whereas in the 1990s models included only two

parameters (Rakotovao Ravahatra et al., 2014). The use of these approaches in a reliability context raises several problems: how "to randomize" models whose parameters are usually correlated with random variables having no prior information on them and how to perform sensitivity studies in the absence of these trends? Rakotovao Ravahatra et al. (2014) have shown that the scatter between models is in the same order of magnitude as the one induced by uncertainty propagation. Another common approach relies on the numerical solving of coupled physical equations (Bastidas-Arteaga et al., 2011): the computation time is then incompatible with maintenance optimization algorithms and the authors suggested the building of mono-variate meta-models using Markov Chains (Bastidas-Arteaga et al., 2012). An alternative for the model's calibration is to associate the integration of expertise in the model and the multiplication of the associated experiments, which grows in number as a function of the parameters number. Furthermore, health-monitoring data usually issued from NDT is not directly linked to these physics-based models and their associated parameters.

One approach that seems promising for maintenance optimization in civil engineering is the construction of data-driven degradation models based on stochastic processes such as the gamma process (Van Noortwijk, 2009) or Brownian motion (Si et al., 2011). It allows modeling the evolution of the degradation using observations (here via NDT) while maintaining the most critical aspects of the degradation mechanism in the model for decision making and an ease of integration in increasingly complex maintenance decision schemes. We may further highlight modeling difficulties when the selected pathologies have nonstationary behaviors over time (acceleration or deceleration effects of degradation). To model the nonstationarity, an interesting approach is to use state-based or condition-based extensions. This choice is motivated by the idea that a variation of the speed of degradation is dependent on the current condition or state of degradation rather than on the time of initiation of degradation (Nicolai et al., 2007). In the case of the gamma processes, conditional or state-dependent extensions have been proposed in the case of mono-variate models (Vatn, 2012), and multivariate models (Zouch et al., 2012). Multivariate models offer more advantages in maintenance and decision-making contexts such as the ability to model and take into account unobservable degradations and model the effects of imperfect maintenance actions on the degradation model. Furthermore, multivariate models allow distinguishing between similar structures; for example, let us consider two structures that suffer from corrosion-cracking and share the same maximum crack width, but one has a higher corrosion current, hence, the two

structures will give different prognostics and as a consequence might produce different decisions.

However, for both multivariate and mono-variate models, the authors have failed to find a robust procedure for the identification of input parameters, as well as a lack of application procedure, as a consequence, making their use difficult in an operation context (Riahi et al., 2010; Schoefs et al., 2011). In reply to these last limitations lies the main contribution of this article.

In this article, we detail the construction process of a state-dependent statistical degradation meta-model, which is defined as models based on:

1. A small number of parameters.
2. Probabilistic relevance and physical expertise.
3. Indicators of degradation and durability directly accessible from NDT.

The construction process starts with the degradation mechanism analysis, then with the identification of degradation indicators that are accessible through NDT and important for decision making (representative of the pathology), and finally with the proposal of the adequate mathematical expressions of the model.

Furthermore, a general estimation algorithm that aims to estimate the parameters of the model, especially for cases where we have irregular databases, is proposed. By irregular databases we mean databases suffering from missing, censored, or truncated data.

The construction of the state-dependent degradation model will be discussed within an application of the cracking of a submerged concrete structure subject to corrosion. Preliminary works have been done in El Hajj et al. (2014).

We find an advantage in using state-based degradation meta-models when addressing two main stakes:

1. First, the ageing model description with a physical meaning of the main probabilistic trends and couplings of the inputs (NDT assessment) and outputs (decision parameter). With this approach, we tackle a key issue: the scourge between more and more complex physical models and the increasing complexity of NDT results modeling and assessment (decoupling, fusion, …) with heterogeneous developments between these two scientific fields.
2. Second, the simple description, flexibility, calibration, and statistical calculation make this model easy to implement and beneficial to utilize in a risk management framework. The evaluation of these meta-models is done through state-dependent stochastic processes using information given by

NDT. The idea is to facilitate the transfer between available information and the model.

The remainder of this article is as follows. In Section 2, we will introduce and analyze the cracking process of the corroded reinforced concrete structure. In Section 3, we will detail the construction of the degradation meta-model. In Section 4, the model parameter estimation algorithm is presented and discussed in case of missing data. Section 5 is devoted to illustrating an application of the model in a risk management context. And finally, conclusions and perspectives are drawn in Section 6.

2 DEGRADATION ANALYSIS

One of the main reasons of failure of reinforced concrete structures is due to cracking. Several techniques are devoted to measure and monitor the evolution of cracks in concrete. In the case of a reinforced concrete structure, the steel reinforcement has indeed strengthened the material property, but can also be a cause of new cracks because of its own corrosion.

The cracking process of a reinforced concrete structure can be divided into three phases: the diffusion of the aggressive agent (chlorides in this case), the corrosion of the reinforcement steel, and the crack propagation from the reinforcement steel toward the surface. The first phase is characterized by the diffusion of chlorides in the concrete.

When the chloride concentration on the surface of the reinforcement steel exceeds a threshold concentration, we reach steel depassivation followed by the initiation of the corrosion. The second phase is dominated by the expansion of corrosion products in which they slowly fill the pores surrounding the reinforcement steel. Finally, the propagation phase lies in the accumulation of corrosion. This generates tensile stress and results in an excessive crack propagation of concrete cover till rupture. Note that crack opening may appear after a fatigue loading too. In this study propagation is assumed to result from the corrosion solely.

The construction of the degradation meta-model is based on state-dependent processes that represent "physical" indicators of the pathology. Therefore, in the construction of the model, as it will be detailed in the paragraph 3, the degradation is based on carefully chosen physical indicators that are suitable to be used in a maintenance-aimed degradation model. A great effort is made for the selection of these indicators in a way to represent the best the degradation process. Here, we discuss the choice of the criteria.

Degradation is a complex process that includes numerous factors, some of them varying with time

(e.g., crack width), and some of them are quasi unvarying with time (e.g., concrete cover). All factors have an impact on the degradation; however, only the changing ones can be associated with methods that aim to track the evolution of degradation. And eventually, the unchanging factors' effects are indirectly represented through the rest of the indicators.

For every phase, there are m potential indicators to represent it. However, we want to consider only the important parameters in terms of decision and degradation tracking; thus, we need to restrain the choice to the best ones. Here, we propose to choose two indicators per phase. We want the two indicators to be the most adequate to represent a degradation phase; between them, the two indicators must keep information about the deterioration level, and the potential of evolution. Hence, we ask the question, what makes an indicator a good choice?

An indicator's value in a maintenance-aimed degradation model is in its observable character and in the value of the information that this indicator can give us, especially in terms of state and growth of degradation.

Therefore, the choice of an indicator is based on two things:

1. Observable character and accessibility through NDT methods.
2. Significance or weight of indicator in representing the degradation process.

It is important to point out a limit of this approach. In the proposed selection of indicators, we took into consideration only two parameters out of a potential m parameters to model the degradation process ($m > 2$). As a consequence, some indicators are left out of the model (environmental parameters, e.g., humidity). From a classic mechanist point of view, this approach may be criticized as it leaves out information issued from the indicators. However, the proposed model is not a physics-based model (no physical laws are modeled). The proposed model can be seen as a data-driven state-dependent Markovian process (has the Markovian property); consequently, we are propagating both the degradation process and its history. Therefore, the left-out parameters are in fact indirectly included in the process (O'Connor et al., 2013; O'Connor and Kenshel, 2013).

A degradation model must cover the three phases of the degradation process. However, we restrict the study and the construction to the third phase, that is, crack propagation. The same approach can be carried out to the other phases of this pathology. This construction can serve as a pedagogical illustration of the construction and tools of the meta-model (El Hajj et al., 2015).
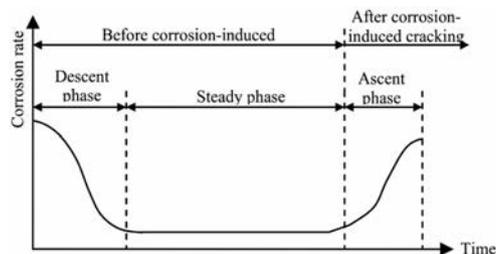

Fig. 1. Tendency of the corrosion rate (Yuan et al., 2009).

2.1 Third phase

Parameters of importance during this phase are the corrosion current density and the width of the crack (Li et al., 2006).

The corrosion rate, V_{corr}, represents the volumetric loss of metal per unit of area and unit of time (mm^2/year); it can be obtained nondestructively from the corrosion current density, i_{corr} ($\mu A/cm^2$) through Faraday's law and the density of the metal. i_{corr} is an instantaneous rate of corrosion measured using NDT, highly sensitive to external conditions (temperature, humidity, etc.; Breysse et al., 2009). Thus, it is important to note that for modeling and decision, support calibration curves depending on the environmental conditions must be available or that inspections must always be carried out under similar environmental conditions and for a given preceding period. We assume the second condition is verified, the first one has not been established for all situations and it is only available in few study cases (Breysse et al., 2009). Figure 1 represents the tendency of the corrosion rate on all three phases of the cracking. According to the corrosion electrochemical principle, the corrosion rate is proportional to the corrosion current density (Yuan et al., 2009).

The crack is considered reachable and its width is easily measured (Gauge block or image analysis: O'Byrne et al., 2013; O'Byrne et al., 2014).

In Figure 2, Vu et al. (2006) draw the shapes of the variation of the width of a crack versus time for two cases of corrosion rate (invariant or time-varying) and two cover thicknesses. The importance of modeling the corrosion rate by a cumulative process is clear. The hypothesis of an invariant corrosion rate is not conservative: in 20 years, it leads to an overestimation of 50% of the crack width, for example, triggering an early repairing decision.

As far as we know, no published work has aimed to study the mutual dependencies between the two processes (corrosion current intensity and cracking) described previously. One main reason can be attributed to the lack of experiments for this particular phase; another good reason is that it is virtually impossible
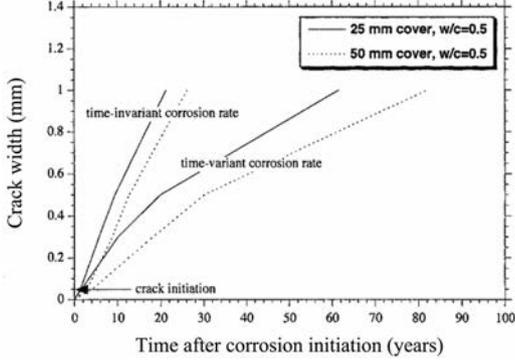
Fig. 2. Tendency of the width of crack—phase 3 (Vu et al., 2006).

to integrate the mutual dependencies in the available physics-based models.

The extension of the degradation model to two processes can be very rewarding in terms of maintenance and inspection optimization, especially in terms of reliability of the data because of the diversification of the NDT techniques (Ploix et al., 2011; El Hajj et al., 2014) and flexibility in the inspection policy when costs or information quality are quite different (Schoefs et al., 2012).

The two indicators will be modeled using nonstationary state-dependent stochastic processes: the evolution laws depend on the current level of degradation. Finally, these two processes form a bivariate cumulative process where the two subprocesses are mutually dependent.

In this document, we argue that the use of two indicators gives a better understanding of the degradation process, provides an additional level to risk quantification of structures, and finally presents the decision maker with more potential maintenance decisions, at every time of the lifetime, based on information coming from two distinct yet dependent physical indicators assessed through NDT.

3 CONSTRUCTION OF THE DEGRADATION META-MODEL

We propose here to define the degradation for each phase as a bivariate process, where each process is a state-dependent stochastic process similar to the approach presented in (Zouch et al., 2012).

On the one hand, the Eurocode associates failure to a crack width exceeding a limit crack width $w_{\lim}$ (3-mm crack width for this reliability case); therefore, the crack width is a direct condition indicator and its value defines the range of serviceability. On the other hand, i_{corr} reveals the speed of filling of rust, that is the motive be-

hind the propagation of cracking, hence is seen as a potential of evolution.

Let the bivariate process be written, $(\rho_t, \theta_t)_{\forall t \geq 0}$, where $(\rho_t)_{\forall t \geq 0}$ describes the condition state and $(\theta_t)_{\forall t \geq 0}$ is the potential of its evolution.

The two processes $(\rho_t)_{\forall t \geq 0}$ and $(\theta_t)_{\forall t \geq 0}$, hereafter written ρ and θ, are both dependent and observable.

The evolution of degradation over a period of time is given by positive increments for the degradation processes respectively $(\Delta \rho, \Delta \theta)$, which are continuous random variables. We assume that the degradation increments in a given time interval τ are random variables which are a function of the present degradation state (ρ, θ).

The degradation process is therefore assumed to be a Markov process. A suitable candidate for the distribution laws of each increment $(\Delta \rho, \Delta \theta)$ is the gamma distribution (Van Noortwijk, 2009) defined by two parameters (α and β, where α is the shape parameter and β is the scale parameter). In the described bivariate state-dependent model, we will consider that only the shape parameter α is a function of the current state (ρ, θ) and τ, but independent of time, however, β is considered as constant.

To simplify the identification step, we will consider that the state dependency is exclusively governed through the shape functions, the scale functions β_θ and β_ρ are considered constant throughout the life cycle. Therefore, we have to model and identify the shape functions α_{θ_i} and α_{ρ_i} which are respectively, in the case of the gamma distribution, proportional to the expected value of the increments for θ and ρ on the time interval τ.

The cracking phase of the corroded concrete structure is then characterized by

1. $(\rho_t)_{\forall t \geq 0}$ modeling the width of the crack "a" (mm);
2. $(\theta_t)_{\forall t \geq 0}$ modeling the corrosion current density "i_{corr}" ($\mu A/cm^2$).

To be able to simulate the model we need to choose a sequence of simulation. First, we seek to characterize the evolution in terms of one process, then for the other.

The choice of the sequence is motivated by mechanical expert judgments; there is a *cause-effect* relationship between the two indicators. The corrosion current density is the *cause*, and the width of the crack is the *effect*. The deterioration starts as an electrochemical process expressed by the corrosion current density than is translated as a cracking on surface expressed by the crack width.

Furthermore, the corrosion current density has an effect on the propagation of cracks, and vice versa (mutual-dependencies). This correlation is modeled

in terms of mutual acceleration effects directly in each of the shape functions of the gamma distributions.

The probability density function of the bi-process $(\Delta\rho, \Delta\theta)$ is given by

$$f_{\Delta\rho,\Delta\theta}\ (x, y; \tau, \rho, \theta) = f_{\Delta\theta}\ (y; \tau, \rho, \theta)$$
$$\times f_{\Delta\rho}\ (x; \tau, \rho, \theta, y) \quad (1)$$

Therefore, next we introduce the conditional probability density functions of each individual increment $f_{\Delta\theta}(y; \tau, \rho, \theta)$ and $f_{\Delta\rho}(x; \tau, \rho, \theta, y)$, used to simulate the process. For every current condition state (ρ, θ), we have the one-step transition density function of the process over the next inspection interval τ.

Finally, to simulate, we first seek to characterize the evolution in terms of the causal process (corrosion current density, Equation (2)), then doing so for the respective effect process (crack width, Equation (3)).

We can then write, $\forall(\rho, \theta) > 0$:

$$\Delta\theta\ (\tau;\ \rho, \theta)\ \sim\ \mathrm{gamma}\left(\alpha_\theta\ (\ \rho, \theta)\times\tau,\ \beta_\theta\right) \quad (2)$$

$$\Delta\rho\ (\tau;\ \rho, \theta, \Delta\theta)\ \sim\ \mathrm{gamma}\left(\alpha_\rho\ (\rho, \theta, \Delta\theta)\times\tau,\ \beta_\rho\right) \quad (3)$$

The choice of each shape function is motivated by the evolution of its respective process in time (Figures 1 and 2). In other terms, the S-shaped condition state evolution of the corrosion current density (Figure 1) requires a bell-shaped shape function, similarly, the L-shaped condition state evolution of the crack width (Figure 2) requires an akin shape function. As a result, we propose as shape functions, $\forall(\rho, \theta) > 0$:

$$\alpha_\theta\ (\rho, \theta) = (a_3 \times \rho + a_4)\ \times e^{\dfrac{-(\theta - a_1)^2}{a_2}} \quad (4)$$

$$\alpha_\rho\ (\rho, \theta, \Delta\theta) = \left(a_6 \times \left(\theta + \frac{\Delta\theta}{2}\right) + a_7\right)\ \times e^{-a_5.\rho} \quad (5)$$

The exponential parts of the shape functions ensure the required shape of the shape-function (e.g., bell shaped). The linear functions before the exponential plays an acceleration role, allowing by that to model the dependencies of the two processes.

One of the main motives in using degradation meta-models is to minimize the number of parameters, explaining the simple linear form of the acceleration function. However, if further knowledge on the correlation of the two physical indicators is available, it is possible to complexify these functions to account for the suitable acceleration and deceleration effects between the two indicators.

We notice in Equation (5) the expression $(\theta + \frac{\Delta\theta}{2})$. In fact, this is a direct consequence of the sequential simulation where we simulate $\Delta\theta$ and use it to estimate $\Delta\rho$.

Table 1
Definition of parameters

| β_θ, β_ρ	*Proportionality factors common to different structures (materials, etc.)*
a_1	Abscissa of the inflection point of the corrosion current density
a_2	Reflects the dispersion around the inflection point
a_3, a_6	Acceleration coefficients
a_4	Speed at the origin of the corrosion current density
a_5	Reflects the kinetics of the process ρ
a_7	Crack growth rate at the origin
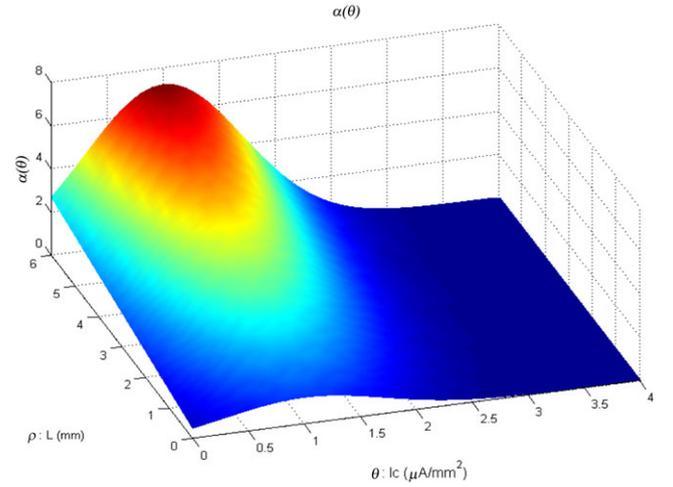
Fig. 3. $\alpha_\theta(\rho, \theta)$: shape function of the θ process.

This expression accounts for the mean value of the corrosion current density over the interval τ.

Now that the model has been defined, a physical meaning can be given for each parameter. In fact, the mathematical formulation of the model allows us to identify physical tendencies (or responsibilities) associated with each parameter. In Table 1, physical meanings of each parameter are summarized.

To further illustrate the model, we propose in Figures 3 and 4 to plot the state-based shape functions of both processes. The following parameters (called hereafter original parameters) are used in this study:

$$a_1 = 1,\ a_2 = 1,\ a_3 = 1,\ a_4 = 1.2,\ a_5 = 0.8,\ a_6 = 1.8,$$
$$a_7 = 2,\ \beta_\rho = 0.3,\ \beta_\theta = 0.3$$

Nonetheless, a sensitivity test is carried out next to illustrate the meanings of the parameters, as well as the validity of the model for different sets of parameters.
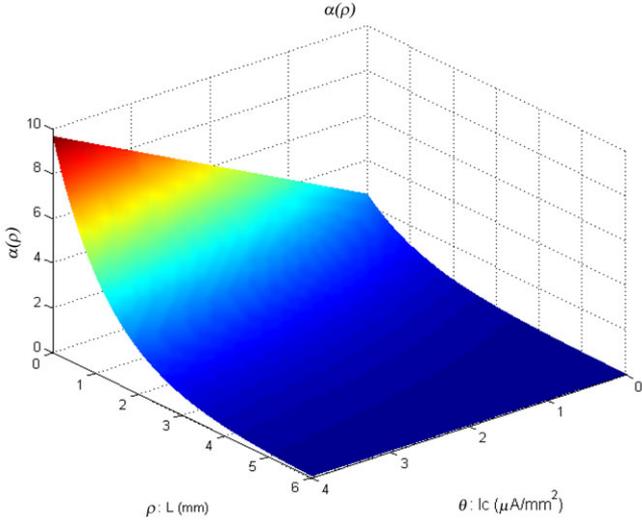
Fig. 4. $\alpha_\rho(\rho, \theta, \Delta\theta)$: shape function of the ρ process for a constant $\Delta\theta$.

3.1 Numerical experiments for the parameters identification

In Table 1, we defined the parameters of the model. In this section, we propose to numerically illustrate their influence on the overall degradation behavior all along the propagation phase.

To that purpose, we will vary the original parameters (chosen in the last section). Finally, in Figure 5, the means of the simulated degradation evolution of 100 structures for 5 selected cases are illustrated.

The 5 cases are summarized by

1. Original parameters
2. Add 0.5 to the a_1
3. Double a_6 and a_7
4. Subtract 0.2 from a_5
5. Double a_3 and a_4

The first thing to point out is that all simulations respect the tendencies found in Figures 1 and 2 (conservation of the S-shape and L-shape).

For the first step, we add 0.5 to a_1. From Table 1, a_1 is defined as the abscissa of the inflection point of the corrosion current density. Using lines L1 and L2 in Figure 5, we see the move of 0.5 on the inflection point.

The second step, we double the values of a_6 and a_7. These parameters account for the acceleration effect of θ and $\Delta\theta$ on ρ. Therefore, we can see the effects on the corrosion current density where it becomes faster.

Note: We note that for every step all the variations from before are kept. For example, in the second step, we keep the 0.5 added to a_1.

Third, we subtract 0.2 from a_5—the parameter responsible for reflecting the kinetics of the process ρ (Table 1). The subtracted 0.2 will raise the shape function (Equation (5)), hence, generating higher increments. This effect is seen clearly in the increasing of the speed of cracking and the corrosion current density (mutual dependency).

Finally, we double a_3 and a_4—parameters responsible for the acceleration effect of ρ on θ. We can see the big effect that this variation has on the increase of speed of θ, not as much for ρ.

To conclude, one aim of using these approaches is to introduce physical meanings to main probabilistic trends. Here, this property is made apparent. Furthermore, the fact that the parameters have physical meanings will allow us to identify effects on the model, such as a maintenance action's effect. For example, a cathodic protection slows the corrosion current density. To take into account this decrease of speed, we can multiply α_θ (Equation (4)) by a parameter that is lower than 1 to generate smaller increments (El Hajj et al., 2015).

4 ESTIMATION PROCESS

4.1 Description of the database

The database is considered to be constructed from periodic inspections at a fixed time step τ on statistically independent but identical structures. Two values define the size of the database N, these are: n, the number of structures, and T, the number of inspections carried out on each structure. We note that in this study, for simplicity, no spatial correlation is accounted for: that implies, for example, that n represents a set of structures (piles or beams of a given bridge, quay,...) in the same environment built with the same materials or a set of independent components (beams) on a given structure (O'Connor et al., 2013; O'Connor and Kenshel, 2013; Schoefs et al., 2009, 2016).

The database is assumed to contain measurements of the crack width and the corrosion current density of each j structure denoted $\{(\rho_t^{(j)}, \theta_t^{(j)}), t \geq 0, j \in 1, n\}$. Increments are directly computed using simple subtraction, and the resulting database group used in the estimation of the parameters of the two state-dependent stochastic processes are: $(\rho_t^{(j)}, \theta_t^{(j)}, \Delta\rho_t^{(j)}, \Delta\theta_t^{(j)}, t \geq 0, j \in 1, n-1)$.

When a database contains N values over N, the database is complete. In this case, the estimation process is founded on a heuristic based on the classical maximum likelihood estimates (MLE) method, and can be found in El Hajj et al. (2014).

For the other cases where the available database's size is smaller than N, we talk about incomplete
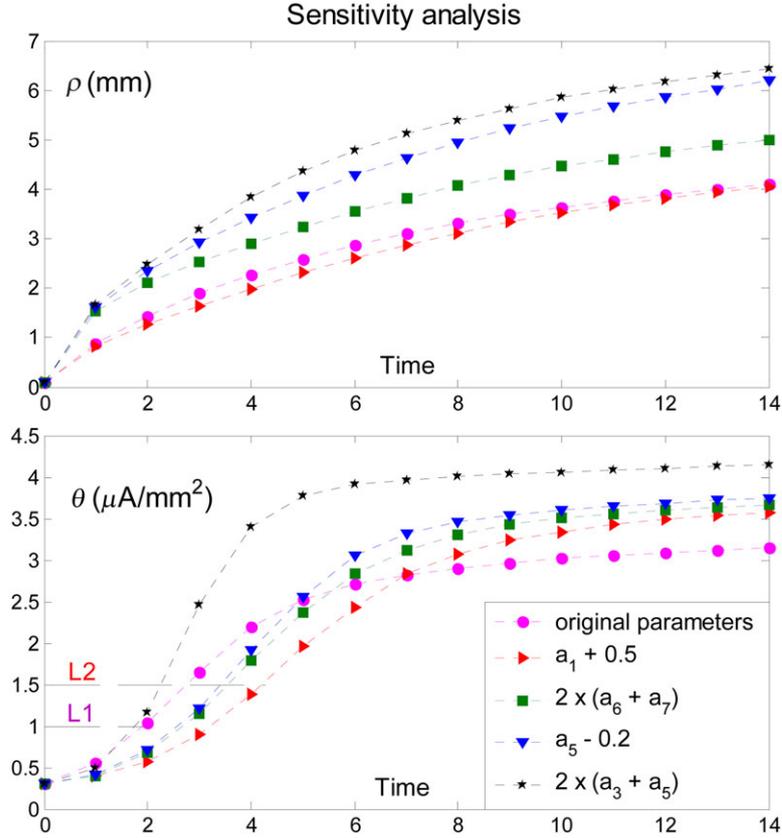
7

Fig. 5. Different patterns of the average degradation on the overall crack propagation phase.

databases. An estimation algorithm based on the stochastic estimation maximization algorithm (SEM) is proposed here for such cases. The algorithm aims to simulate lost information based on known ones to improve the estimation efficiency.

Incompleteness of databases is common in civil engineering. Causes of incompleteness may vary from inspections techniques, measurement error (Schoefs el al., 2012; Torres-Luque et al., 2014), accuracy of the machines, and cancelled inspections (due to security concerns, weather, costs, no available technicians, etc). Data that are lost can be classed into three categories; those are truncated, censored, and missing:

1. Censored values are those reported as less than some value—left censored (e.g., <5 cm), greater than some value—right censored (e.g., >0.1 μA/cm^2), or as an interval—interval censored (e.g., a value between 67 and 75 days).
2. Truncated values are those that are not reported if the value exceeds some limit—right truncated (e.g., if the crack width is above 3 cm we stop recording) or if values exceed the physical under-

standing (e.g., corrosion depth of steel more than original thickness).
3. Missing data are when values are lost due to recording interruptions related to field data measurement or missed inspections.

All three categories occur frequently in civil engineering. As missing data can have a significant effect on the drawn conclusions, it is important to take such limitations into consideration. In a condition-based maintenance context, every decision is based on the degradation level and, therefore, it is really important to give the best available prediction.

Essentially, the way we deal with the three types of incompleteness in databases is the same. Next, we will present the imputing and estimation algorithm used in the estimation process in the presence of missing data or incomplete database.

4.2 Modified stochastic expectation maximization algorithm

A high level of censoring or missing data increases the numerical instability of the optimization problem in the

estimation process, especially in the maximization of the likelihood. The Expectation–Maximization (EM) algorithm, introduced by Dempster et al. (2007), is an iterative procedure designed to find MLE in the context of parametric models where the observed data can be viewed as incomplete. The EM algorithm is a simple approach; unfortunately it has its limitations. As noted in Dempster et al. (2007), the convergence rate of EM is linear and governed by the fraction of missing information meaning that the EM algorithm can be extremely slow when the proportion of missing data is high. Moreover, the EM is proved to converge to a stationary point of the log-likelihood function, but when several stationary points are present, the algorithm does not necessarily converge to a significant local maximum of the log-likelihood function.

To answer to the limitations of the EM algorithm, the SEM algorithm has been proposed by Celeux et al. (1985). The S in SEM stands for stochastic. The SEM algorithm incorporates a stochastic step between the E-step and the M-step of the EM algorithm. The stochastic step is based on the Random Imputation Principal (RIP), meaning to replace each missing quantity with a value drawn at random from the conditional density $q(y|x, \Theta^{(n)})$, where $\Theta^{(n)}$ are the parameters estimates at the nth iteration, y denotes the missing data, x denotes the observed data, and z denotes the complete database: $z = x \cup y$. $\Theta^{(n)}$, not to be confused with θ, that is, the physical indicator in the model.

The SEM nth iterations are as follows:

1. E-step: compute the conditional density $f_{\Delta\rho, \Delta\theta}(x, y; \tau, \rho, \theta)$ using $\Theta^{(n)}$.
2. S-step: using RIP, we Simulate the unobserved data to draw a complete sample $z^{(n)}$.
3. M-step: find the ML estimates of $\Theta^{(n+1)}$ using the complete sample $z^{(n)}$.

The program starts by scanning the provided database for potentially missing data, and indexes them. Then, using the observed data, initial parameters are estimated and used to start the iterative modified SEM algorithm.

In the S-step, we include a test that verifies if the generated data respect the monotonicity of the gamma process. For the M-step, and because of the high number of parameters in $\Theta^{(n+1)}$, an iterative two-step heuristic algorithm is used for the maximization of the log-likelihood. In fact the selected shape functions lead to numerical instability problems with conventional optimization procedures. To circumvent this problem, we constructed a heuristic based on the fixed point theorem. This heuristic is applied iteratively to provide
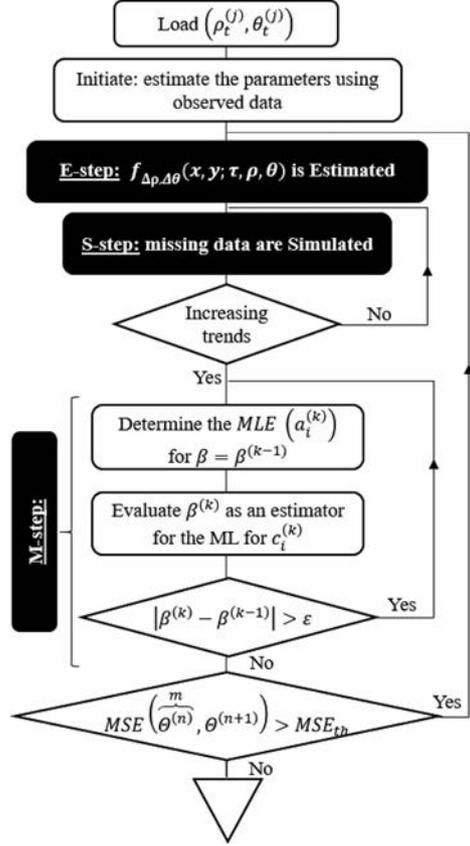

Fig. 6. Stochastic Estimation–Maximization algorithm.

estimates of the parameters of the model (Equations (2) and (3)).

In Figure 6, the SEM algorithm is presented.

The SEM iterative algorithm stops for a desired accuracy of the estimated parameters: this accuracy is expressed in terms of target mean square error between the last $m + 1$ sets of parameters, given by

$$\text{MSE}\left(\overbrace{\Theta^{(n)}}^{m} - \Theta^{(n+1)}\right) =$$

$$\text{trace}\left\{ E\left\{ \left(\overbrace{\Theta^{(n)}}^{m} - \Theta^{(n+1)}\right)\left(\overbrace{\Theta^{(n)}}^{m} - \Theta^{(n+1)}\right)^{\text{T}} \right\} \right\}$$

where $\overbrace{\Theta^{(n)}}^{m}$ is the mean of the last m estimations to limit variations related to one simulation.

We define here a threshold $\text{MSE}_{\text{th}} = 0.1$, that is, for when the estimated parameters are not changing anymore.

Note: We will not demonstrate the convergence of this fixed-point type algorithm. However, the large

number of numerical experiments that we describe below portray the good properties of this algorithm.

4.3 Numerical experiments for illustrating the convergence of the process

For the third phase of degradation, no real field databases were available for us to use. Therefore, we propose to simulate a set of virtual databases. In each case, we will detail the method used to simulate a database.

First, we start by illustrating the use of the SEM algorithm on a general case where values are censored and missing from the generated database. Then we study the effect of missing data and censored data on the estimation process.

The mean squared error (MSE) is used as the performance criteria for the estimation process. To illustrate the performance of the estimation process and its convergence, we propose to evaluate the MSE of the estimated parameter vector $\hat{\Theta}$ based on the simulated database given the true parameter values θ^*:

$$\text{MSE} = \text{trace} \left\{ E \left\{ \left(\hat{\Theta} - \theta^* \right) \left(\hat{\Theta} - \theta^* \right)^{\mathrm{T}} \right\} \right\} \quad (6)$$

where the original parameters are grouped in $\Theta^* = \{a_1{}^*, a_2{}^*, a_3{}^*, a_4{}^*, a_5{}^*, a_6{}^*, a_7{}^*, \beta_\theta{}^*, \beta_\rho{}^*\}$, and the estimated parameters in $\hat{\Theta} = \{\widehat{a_1}, \widehat{a_2}, \widehat{a_3}, \widehat{a_4}, \widehat{a_5}, \widehat{a_6}, \widehat{a_7}, \widehat{\beta_\theta}, \widehat{\beta_\rho}\}$.

All calculations are carried out under Matlab$^©$.

4.3.1 General case: random missing and censored data. We distinguish between two types of missing data:

1. Simultaneously or "total" missing data: If a data point is missing from one process ($\rho_t^{(j)}$ or $\theta_t^{(j)}$), certainly a data point will be missing, at the same inspection, in the other process (respectively $\theta_t^{(j)}$ or $\rho_t^{(j)}$).
2. Nonsimultaneously or "partial" missing data: A more realistic case where the missing data can be unavailable from one process, but not necessarily from the second one.

Furthermore, a right censoring (Type 1) is also considered, meaning that inspections will cease to be recorded after a certain time.

In this first example, we will illustrate the general case where we consider randomly censored and *partially* missing data.

We consider the case of $n = 15$ structures, we start by simulating full n simulations with no missing nor censored data using the two state-dependent stochastic processes. For each simulation, 10 inspections are
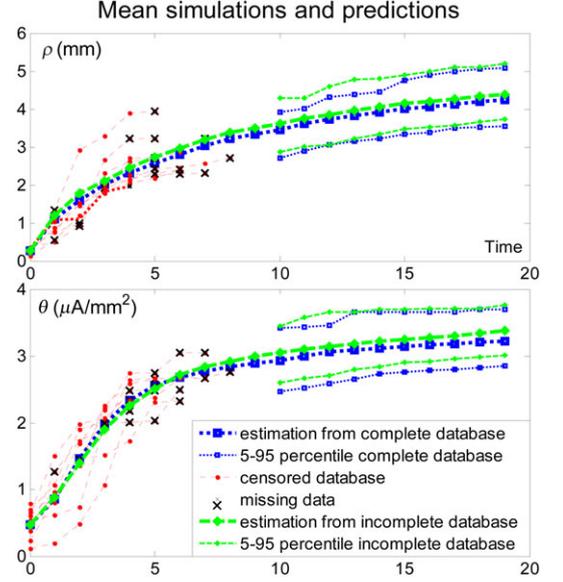

Fig. 7. Mean simulations estimated from randomly missing and censored database.

considered. Therefore, the full database is formed of a total of 150 data points.

Now, we randomly delete 25% of the data from each simulation. To do so, we use a uniform distribution over the recorded time (10 inspections in this case).

Then, to model the censorship, we consider a normal distribution, centered on the 10th inspection with a variance = 5 and truncated to the right. Then, we randomly choose 15 numbers and these numbers will be the censorship times after which all data points are removed from the databases.

Figure 7 compares the average estimated behavior of the structures from the incomplete database. In this case, the estimated parameters are estimated after 5 iterations of the SEM (under 20 seconds using an office PC), stopping the algorithm for a MSE = 0.0905 < 0.1.

In Figure 7, we can see the estimation from incomplete and complete databases re-joins the experimental trends in Figures 1 and 2. Furthermore, we can see that the curves representing the means and the variances on the long term are close to each other.

The aim of Figure 7 being to exemplify the estimation and prediction process, we propose now to further investigate the quality of the estimation. To this aim, we will use a more quantitative criterion to compare the two estimations that is the relative error committed on each parameter.

In Table 2, the estimated parameters and the relative errors are summarized.

Table 2
Numerical example parameters estimate

Database parameters		Complete database estimates		Incomplete database estimates	
	Θ^*	$\hat{\Theta}_c$	$\bar{e}_c$	$\hat{\Theta}_{inc}$	$\bar{e}_{inc}$
a_1	1	1.04	−0.04	0.88	0.12
a_2	1	1.17	−0.17	1.65	−0.65
a_3	1	0.58	0.42	0.87	0.13
a_4	1.2	1.49	−0.24	1.41	−0.18
β_ρ	0.3	0.29	0.03	0.26	0.13
a_5	0.8	1.03	−0.29	0.96	−0.20
a_6	1.8	4.25	−1.36	3.46	−0.92
a_7	2	1.13	0.44	0.78	0.61
β_θ	0.3	0.31	−0.03	0.39	−0.30

The relative error here is given by the equation:

$$\bar{e} = \frac{a_i^* - \widehat{a_i}}{a_i^*} \qquad (7)$$

In Table 2 $\hat{\Theta}_c$ are the estimates of the complete databases, $\bar{e}_c$ is the relative error associated with these estimates; and $\hat{\Theta}_{inc}$ are the estimates of the incomplete databases, $\bar{e}_{inc}$ is the relative error associated with these estimates.

In compliance to Figure 7, Table 2 shows that the estimated parameters from the incomplete and complete databases are close to the original parameters. Except for a_6 where the errors are higher, this is mainly due to the fact that this parameter represents a mutual dependency between processes (Table 1), hence requires more data points to be captured better.

To further quantify the quality of the estimation, we propose to illustrate the effect in terms of relative errors on the shape functions.

In Figures 8 and 9, we illustrate the relative errors committed on the shape functions for both processes (ρ and θ) and for both cases (i.e., complete and incomplete databases) using the estimated parameters $\hat{\Theta}_c$ and $\hat{\Theta}_{inc}$.

To plot these surfaces, the relative errors between the shape functions are calculated using the following expressions, $\forall\, 0 \leq \rho \leq 6$ and $0 \leq \theta \leq 4$:

$$e_{\alpha(\rho)} = \left(\frac{\alpha_\rho^* - \widehat{\alpha_\rho}}{\alpha_\rho^*} \right) \quad \text{and} \quad e_{\alpha(\theta)} = \left(\frac{\alpha_\theta^* - \widehat{\alpha_\theta}}{\alpha_\theta^*} \right) (8)$$

where α_ρ^* and α_θ^* are the shape functions associated with the original parameters Θ^*, $\widehat{\alpha_\rho}$ and $\widehat{\alpha_\theta}$ are the shape functions associated with the estimated parameters $\hat{\Theta}$.

In Figure 8, we plotted the relative errors committed on the shape functions of the ρ process. "A" is the point with the highest crack width (3.9) and corrosion cur-
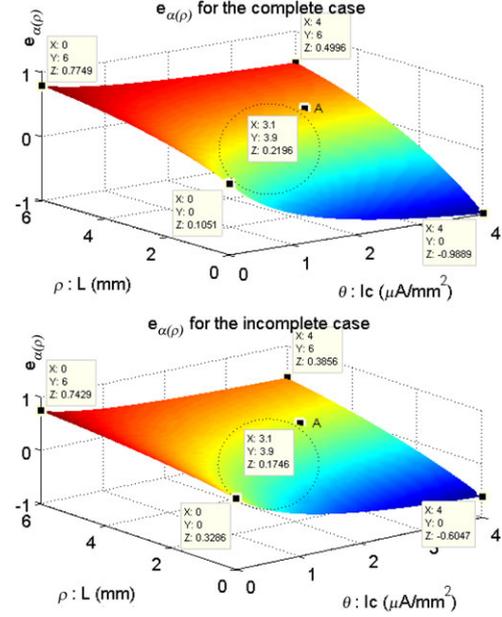

Fig. 8. Relative error committed on the α_ρ for complete and incomplete databases.
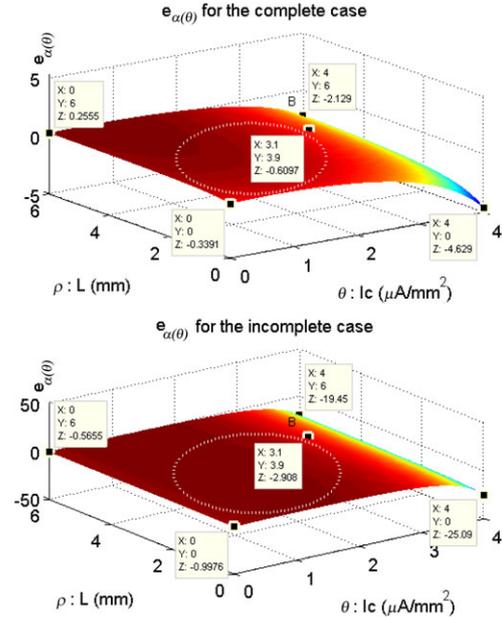

Fig. 9. Relative error committed on α_θ for complete and incomplete databases.

rent density (3.1) from Figure 7 to better position our example.

The surface of the relative errors is a plane tilted to the right towards the θ axis.

When we compare the errors at the five points shown on the plots, we can see that they are close.

Table 3
Mean relative errors on the estimated shape parameters

	$\overline{e_{\alpha(\rho)}}$	$\overline{e_{\alpha(\theta)}}$
Complete	0.2773	0.2560
Incomplete	0.2133	0.6046

Table 4
MSE for missing data

Missing rate	n	5	10	20
75%	Total	2.2	0.878	0.423
	Partial	1.78	0.850	0.389
50%	Total	1.756	0.737	0.422
	Partial	1.702	0.668	0.376
25%	Total	1.527	0.628	0.305
	Partial	1.469	0.591	0.301
0%	Full database	0.96	0.46	0.24

The region inside of the dotted circle formed between "A" and the origin is where the majority of the data points (ρ, θ) are. This might explain the good estimation in these regions and how the error increases away from it.

In Figure 9, we plotted the relative errors committed on the shape functions of the θ process. "B" is the point with the highest crack width (3.9) and corrosion current density (3.1) from Figure 7 to better position our example.

The surface of the relative errors illustrates an increase of the error with θ.

When we compare the errors at the five points shown on the plots, we can see that for the incomplete case the errors are a lot higher for high values of corrosion current density.

For the region inside of the dotted circle formed between "B" and the origin, where the majority of the data points (ρ, θ) are, we see that the errors for the incomplete data points are slightly higher than those of the complete.

To further compare the two cases, we propose to calculate the mean relative errors committed on the shape parameters. The mean relative error is calculated for the range between the "A" or "B" and the origin and given by

$$\overline{e_{\alpha(\rho)}} = \overline{\left(\frac{|\alpha_\rho{}^* - \widehat{\alpha_\rho}|}{\alpha_\rho{}^*}\right)} \text{ and } \overline{e_{\alpha(\theta)}} = \overline{\left(\frac{|\alpha_\theta{}^* - \widehat{\alpha_\theta}|}{\alpha_\theta{}^*}\right)} \quad (9)$$

where $\rho \in [0, 3.9]$ and $\theta \in [0, 3.1]$.

The mean relative error are summarized in Table 3.

In Table 3, we see that the mean relative error for ρ are close for both cases, but for θ the complete case is lower by half almost.

To conclude, this example illustrates the ability of the SEM model to estimate the parameters from a database that suffers from missing and censored data points.

4.3.1.1 Effect of the level of missing data. In this section, we investigate the effects of the percentage of missing data on the estimates by means of MSE by independently considering the two cases: (i) the total missing data and (ii) the partial missing data.

We want to study the accuracy of our estimation algorithm as a function of the level and type of missing data over the lifetime.

We use a series of 10 inspections on n structures using the same model as in the previous section, then we randomly remove data from the database using a uniform distribution between the first and the tenth inspection. We propose to vary this missing rate from 75% to 0%, where 0% corresponds to the use of the complete data set for the estimation.

The results are shown in Table 4. Each cell is the mean MSE for the estimated parameters (for both processes) based on 400 scenarios generated for the two types. The number of structures and the percentage of the missing data are the variables of this study. No censored data are considered in this study.

We notice that the MSE and n are inversely proportional, and that the estimation process is convergent. We can also notice that the MSE in the partial case is slightly smaller than the MSE in the total case. The most significant difference is in the worst case scenario of five structures and 75% missing data rate where the MSE drops 0.4.

An aim of this section is to study the effect of the missing data on the accuracy of the estimation process; therefore, it is important to explain for instance the significance of an MSE equal to 2.2, and if it is a good or a bad estimate. To this aim, we propose to plot six random results of estimations on the same graph with their MSE. We consider the following cases:

1. Three results, $n = 10$ with 50% missing data rate;
2. Two results, $n = 10$ with 75% missing data rate;
3. One result, $n = 5$ and 75% (worst case scenario).

The results are illustrated in the Figure 10.

From Figure 10, all plots came close on the plot except for the last result MSE $= 2.3227$ that overestimates the degradation on the long term (dashed and starred lines). It is important to remind that this result is for five structures with a large 75% missing data rate,
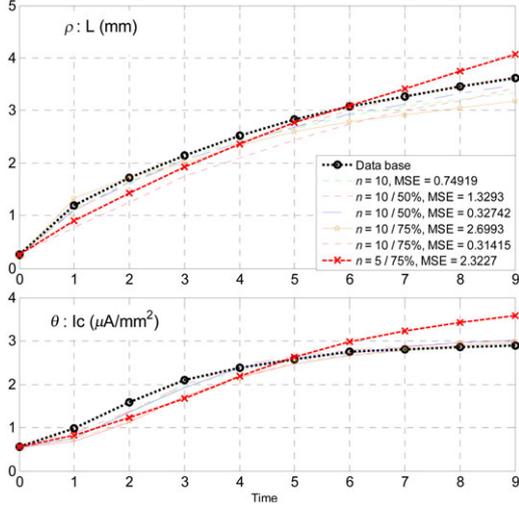
Fig. 10. Six random estimations with their MSE.

meaning that approximately 12 points are used to estimate nine parameters.

On another hand, this is a multiparametric problem (nine parameters), and clearly a complex one; it is possible to have different sets of parameters giving the same simulations. For example, from Table 2 we compute the MSE for the two estimated sets and we get:

$$\text{MSE}_c = 0.79 > \text{MSE}_{inc} = 0.53$$

However, the simulation using the full database estimates will be better because it maximizes the likelihood function. Consequently, one might question the efficiency of MSE as criterion to assess the quality of estimation, but by carrying out 400 simulations the results will balance out allowing us to compare between the cases.

4.3.1.2 Effect of censored data.   Here, we aim to study the effect of censored data on the estimation process. For this purpose, we consider 10 inspections carried out on n structures, then we apply a censorship to the right to these data sets like it was introduced at the start of Section 4.3.1.

We propose a 0% to 80% censorship on the data, where 80% means that it is possible that only two inspections are left, etc. Then we vary the variances respectively with the desired censorship and we randomly choose the censorship times.

We propose to calculate the MSE of the estimated parameters based on the simulation of 400 scenarios. The number of structures (n) and the censorship rate are the variables in this study. Results are summarized in Table 5.

Table 5
Effect of censored data and number of structures on MSE

Censorship rate	80%	40%	0%	Full database
n				
5	3.07	2.52	1.47	0.96
10	1.37	0.81	0.59	0.46
20	0.32	0.32	0.30	0.24

By increasing the number of inspected structures, we offer a valuable assistance to the estimation process even in cases of high levels of censored data. We may point out also the impact of the first inspections (80% case) where for a fair number of structures they can give good estimates.

5 RISK MANAGEMENT EXAMPLE

As defined in the Eurocode—EN 1990 (CEN, 2002), structural reliability is the ability of a structure or a structural member to fulfill the specified requirements for which it has been designed; it includes structural safety, serviceability, and durability. Given the random nature of the quantities involved in structural design (environment, material properties, etc.), the assessment of structural reliability cannot be done by deterministic means, but requires a probabilistic analysis.

In this context, the objective of a safety verification is to keep the probability of failure P_f lower than a given threshold. This threshold is basically a function of the consequences that the failure of the structure can have. In this example, we will consider that the corroded structure is a deck, that is, an important element with medium consequences. The Eurocode refers to this as a reliability case 2 or RC2 (not to be confused with Reinforced Concrete, which is also referred to as RC in this document). The code demands that the reliability index β be equal to 4.7 (equivalent to $P_f = 10^{-7}$) in the case of a 1-year reference period. In this example, we will consider an inter-inspection time of 1 year.

As a consequence, we propose here to define the decision criterion as the probability of having a failure before the next inspection.

The probability of a failure before the next inspection symbolized by $P_f(\rho, \theta)$ is a function of the current state (ρ, θ), and given by the following equation:

$$P_f(\rho, \theta) = P(\Delta\rho + \rho > L)$$

For the durability of RC structures, Eurocode 2 expresses the failure by comparing the crack width of the concrete cover with a cracking threshold $L_{\max}$. The
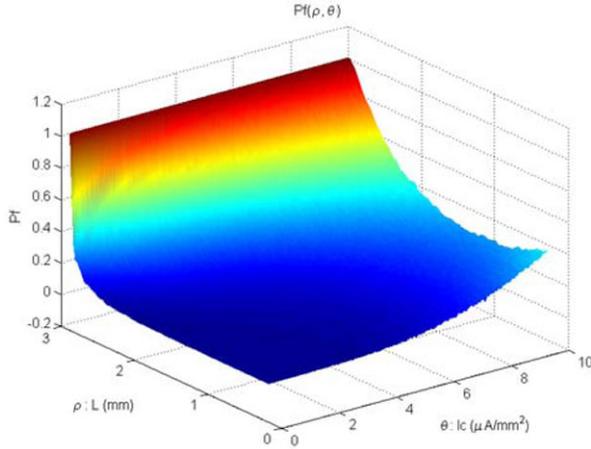
Fig. 11. Probability of failure based on degradation level (ρ_i, θ_i).
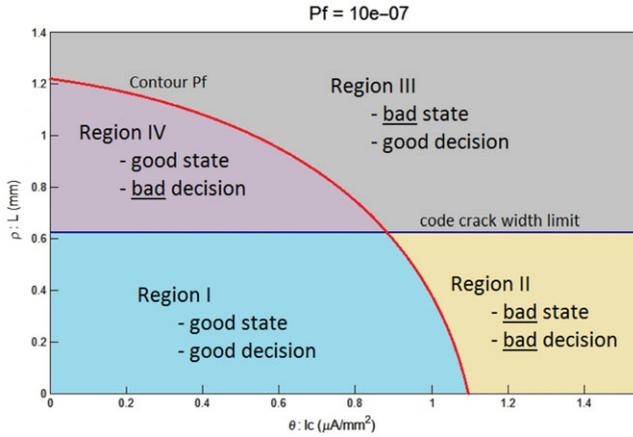

Fig. 12. Bivariate decision plan and a comparison with a mono-variate plan.

latter depends on both the characteristic of the structure and its environmental conditions. According to Table 7.101N from Eurocode 2 (Eurocode 2, 2005), for an exposure class XS3 (corrosion of the reinforcement induced by chlorides from sea water), the performance of the reinforced concrete structure are assumed modified when the width of a crack is greater than or equal to 3 mm.

For a selected range of ρ ($0 < \rho < L_{\max} = 3$ mm) and θ, we estimated the probability of failure for every combination (ρ, θ). Results are plotted in Figure 11.

The use of this curve in reliability-based risk management could be in a classical way where a P_f threshold defines an acceptance and a critical areas. Therefore, we define an iso-curve as the line joining all observations (ρ, θ) having the same $P_f = 10^{-7}$. This iso-curve is then drawn in Figure 12.

The iso-curve divides the plot in to two areas: an acceptance reliability area where $P_f < 10^{-7}$ (regions I and IV), and a critical reliability area where $P_f > 10^{-7}$ (regions II and III).

The system is said to be safe for an inspection (ρ, θ) in the acceptance reliability area and unsafe for an inspection in the critical reliability area.

In this manner, decisions are based on information coming from two sources, or two physical indicators. Next, we propose to illustrate the benefit of using a bivariate approach instead of a classical mono-variate approach.

To this aim, we propose to calculate the value of the crack width associated with a $P_f = 10^{-7}$. Then, the calculated value, $L' = 0.62$, is represented as a line on Figure 12. For simplification purposes, let us call this line the L-line.

The contour plot and the L-line divide the plane into four regions (Figure 12). For regions I and III, the mono-variate approach and the bivariate approach generate the same decision, whereas for regions II and IV decision are opposite.

According to a mono-variate approach, if the crack width is below the L-line, that is, crack width $< L'$, the structure is considered to be safe. And if the crack width is higher than, then the structure is in a critical zone.

Such formulation might cause the same decision for two structures with the same crack width, but with two very different corrosion current density (θ), that can be seen as an acceleration factor of the cracking.

From an L-line's point of view (or a mono-variate approach), an observed small crack width with an extremely high corrosion current density can lead to an assessment as "safe structure," which in reality might be misleading, if not dangerous. For example, a state (ρ, θ) in region II is safe according to the L-line. However, it is logical to consider that for high values of corrosion current density we might be in a critical zone because corrosion is intensively active and a maintenance action can prevent an earlier failure. In such cases, a decision made according to the L-line only can be faulty compared to a bivariate approach where information issued from additional physical indicators are considered in the decisional process.

Also for region IV, a bad decision is made. However, in this case the decision is not as dangerous as for region II because a mono-variate approach assesses the structure as critical and demands a maintenance. But in reality, and according to the bivariate approach, we are in a safe zone; the mono-variate maintenance decision will generate unnecessary maintenance costs.

To conclude, in this application we illustrated a potential benefit of using a bivariate degradation metamodel for risk-management applications, where the two

physical indicators of the model are at the same time the inputs (NDT assessments) and the outputs (decision parameters).

6 CONCLUSIONS AND PERSPECTIVE

The gap between the sophistication of physical degradation models and complexity of NDT results is currently a huge challenge: one of the consequences is that model inputs and NDT outputs are less and less related. This article suggests a solution based on simplified models whose inputs are directly the NDT outputs.

A degradation meta-modeling approach based on data-driven bivariate nonstationary stochastic processes with a gamma trend is discussed within and application of chloride induced cracking. The degradation model is based on two correlated state-dependent stochastic processes. The construction and calibration of the processes are done via NDT data, including visual inspection or image processing (see O'Byrne et al., 2013, and Yeum et al., 2015, for crack for instance).

Expert knowledge is introduced to reflect the main useful degradation properties that the model should tackle for decision making. This allows both to simplify the construction and the fitness of the model to the data and to overcome some limits in the current practices in civil engineering (here the correlation between the crack width and the corrosion current density). This approach is shown to be robust in case of data missing and allows reformulating risk based decision aid tools.

This model shows a lot of potential, especially in decision making: we provide a set of values for corrosion current density and crack width that leads to a target probability of failure before the next inspection. Starting from this potential, other applications can be suggested. An important point to address is the modeling of the correlation of these two processes, most importantly the statistical correlation.

ACKNOWLEDGMENTS

REFERENCES

Bastidas-Arteaga, E., Chateauneuf, A., Sánchez-Silva, M., Bressolette, P. & Schoefs, F. (2011), A comprehensive probabilistic model of chloride ingress in unsaturated concrete, *Engineering Structures*, **33**(3), 720–30.

Bastidas-Arteaga, E. & Schoefs, F. (2012), Stochastic improvement of inspection and maintenance of corroding reinforced concrete structures placed in unsaturated environments, *Engineering Structures*, **41**, 50–62.

Boéro, J., Schoefs, F., Capra, B. & Rouxel, N. (2009), Risk management of French harbour structures. Part 2: Current practices, needs—experience feedback of owners, *PARALIA*, **2**, 6.13–6.24.

Breysse, D., Yotte, S., Salta, M., Schoefs, F., Ricardo, J. & Chaplain, M. (2009), Accounting for variability and uncertainties in NDT condition assessment of corroded RC-structures, *European Journal of Environmental and Civil Engineering*, **13**(5), 573–91.

Celeux, G. & Diebolt, J. (1985), The SEM algorithm: a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem, *Computational Statistics Quarterly*, **2**(1), 73–82.

CEN. (2002), Eurocode - Basis of structural design, EN 89.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (2007), Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)*, **39**(1), 1–38.

El Hajj, B., Castanier, B., Schoefs, F., Bastidas-Arteaga, E. & Yeung, T. (2015), A condition-based maintenance policy based on a probabilistic meta-model in the case of chloride-induced corrosion, in T. Haukaas (ed.), *Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12)*, Vancouver, Canada.

El Hajj, B., Castanier, B., Schoefs, F. & Yeung, T. (2014), A condition-based deterioration model for the crack propagation in a submerged concrete structure, in *Proceedings of the Annual European Safety and Reliability Conference (ESREL 2014), Wroclaw, Poland*, Vol. **103**.

Eurocode 2. (2005), BS EN 1992-2:2005 - Eurocode 2: Design of concrete structures. Part 2: Concrete bridges—design and detailing rules, Brussels, Belgium, 97 pp.

Frangopol, D. M., Kallen, M.-J., Van Noortwijk, J. M. & Van Noortwijk, J. M. (2004), Probabilistic models for life-cycle performance of deteriorating structures: review and future directions, *Progress in Structural Engineering and Materials*, **6**(4), 197–212.

Li, C.-Q., Melchers, R. E. & Zheng, J.-J. (2006), Analytical model for corrosion-induced crack width in reinforced concrete structures, *Structural Journal*, **103**(4), 479–87.

Nicolai, R. P., Dekker, R. & Van Noortwijk, J. M. (2007), A comparison of models for measurable deterioration: an application to coatings on steel structures, *Reliability Engineering and System Safety*, **92**(12), 1635–50.

O'Byrne, M., Ghosh, B., Schoefs, F. & Pakrashi, V. (2014), Regionally enhanced multiphase segmentation technique for damaged surfaces, *Computer-Aided Civil and Infrastructure Engineering*, **29**(9), 644–58.

O'Byrne, M., Schoefs, F., Ghosh, B. & Pakrashi, V. (2013), Texture analysis based damage detection of ageing infrastructural elements, *Computer-Aided Civil and Infrastructure Engineering*, **28**(3), 162–77.

O'Connor, A. & Kenshel, O. (2013), Experimental evaluation of the scale of fluctuation for spatial variability modeling of chloride-induced reinforced concrete corrosion, *Journal of Bridge Engineering*, **18**(1), 3–14. Available at: http://ascelibrary.org/doi/10.1061/%28ASCE%29BE.1943-5592.0000370, accessed May 2016.

O'Connor, A., Sheils, E., Breysse, D. & Schoefs, F. (2013), Markovian bridge maintenance planning incorporating

corrosion initiation and non-linear deterioration, *ASCE Journal of Bridge Engineering*, **18**(3), 189–99.

Ploix, M. A., Garnier, V., Breysse, D. & Moysan, J. (2011), NDE data fusion to improve the evaluation of concrete structures, *NDT&E International*, **44**(5), 442–48.

Rakotovao Ravahatra, N., de Larrard, T., Duprat, F., Bastidas-Arteaga, E. & Schoefs, F. (2014), Sensitivity analysis of simplified models of carbonation—extension in spatial variability—updating through Bayesian network, *Proceedings of the 2nd International Symposium on Uncertainty Quantification and Stochastic Modeling, Uncertainties 2014*, Rouen, France.

Riahi, H., Bressolette, P. & Chateauneuf, A. (2010), Random fatigue crack growth in mixed mode by stochastic collocation method, *Engineering Fracture Mechanics*, **77**(16), 3292–309.

Schoefs, F., Abraham, O. & Popovics, J. S. (2012), Quantitative evaluation of contactless impact echo for non-destructive assessment of void detection within tendon ducts, *Construction and Building Materials*, **37**, 885–92.

Schoefs, F., Bastidas-Arteaga, E., Tran, T. V., Villain, G. & Derobert, X. (2016), Characterization of random fields from NDT measurements: a two stages procedure, *Engineering Structures*, **111**, 312–22. Available at: http://www.sciencedirect.com/science/article/pii/S0141029615007476, accessed May 2016.

Schoefs, F., Clément, A. & Nouy, A. (2009), Assessment of ROC curves for inspection of random fields, *Structural Safety*, **31**(5), 409–19.

Schoefs, F., Yáñez-Godoy, H. & Lanata, F. (2011), Polynomial chaos representation for identification of mechanical characteristics of instrumented structures, *Computer-Aided Civil and Infrastructure Engineering*, **26**(3), 173–89.

Sheils, E., O'Connor, A., Breysse, D., Schoefs, F. & Yotte, S. (2010), Development of a two-stage inspection process for the assessment of deteriorating infrastructure, *Reliability Engineering and System Safety*, **95**(3), 182–94.

Si, X. S., Wang, W., Hu, C. H. & Zhou, D. H. (2011), Remaining useful life estimation—a review on the statistical data driven approaches, *European Journal of Operational Research*, **213**(1), 1–14.

Torres-Luque, M., Bastidas-Arteaga, E., Schoefs, F., Sánchez-Silva, M. & Osma, J. F. (2014), Non-destructive methods for measuring chloride ingress into concrete: state-of-the-art and future challenges, *Construction and Building Materials*, **68**, 68–81.

Van Noortwijk, J. M. (2009), A survey of the application of gamma processes in maintenance, *Reliability Engineering and System Safety*, **94**(1), 2–21.

Vatn, J. (2012), A state based model for opportunity based maintenance, in *11th International Probabilistic Safety Assessment and Management Conference (PSAM 11) and the Annual European Safety and Reliability Conference (ESREL 2012), Helsinki, Finland*, **1**, 1–4.

Vu, K. A. T., Stewart, M. G. & Mullard, J. (2006), Corrosion-induced cracking: experimental data and predictive models, *Structural Journal*, **102**(5), 719–26.

Yeum, C. M. & Dyke, S. J. (2015), Vision-based automated crack detection for bridge inspection, *Computer-Aided Civil and Infrastructure Engineering*, **30**(10), 759–70.

Yuan, Y., Ji, Y. & Jiang, J. (2009), Effect of corrosion layer of steel bar in concrete on time-variant corrosion rate, *Materials and Structures*, **42**(10), 1443–50.

Zouch, M., Yeung, T. & Castanier, B. (2012), Optimizing road milling and resurfacing actions, *Proceedings of the Institution of Mechanical Engineers, Par O: Journal of Risk and Reliability*, **226**(2), 156–68.