*XIII*èmes *Journées Nationales Génie Côtier – Génie Civil*
Dunkerque, 2-4 juillet 2014
DOI:10.5150/jngcgc.2014.077 *© Editions Paralia CFL*
disponible en ligne – http://www.paralia.fr – available online

Optimal configuration for identification of parameters for chloride ingress models using Bayesian networks

Thanh Binh TRAN [1], Emilio BASTIDAS-ARTEAGA [1], Franck SCHOEFS [1], Stéphanie BONNET [1]

1. LUNAM Université, Université de Nantes, Institute for Research in Civil and Mechanical Engineering, CNRS UMR 6183, Nantes, France.
thanh-binh.tran@etu.univ-nantes.fr

Abstract:
Chloride ingress into concrete is one of the major causes leading to the degradation of reinforced concrete structures. Its modelling is an important task to plan and quantify maintenance operations of structures. Relevant material and environmental parameters for modelling could be determined from inspection data that is very limited due to time-consuming and expensive tests. The main objective of this paper is to develop a method based on Bayesian updating for selecting appropriate inspection configuration that can provide an optimal balance between accuracy and cost. The results indicate that Bayesian approach could be a useful tool to identify model parameters even from insufficient inspection data.
Keywords: Chloride ingress, Corrosion, Bayesian network, Parameter identification.

1. Introduction

Chloride penetration into reinforced concrete (RC) is one of the main factors responsible for generating corrosion in reinforcing bars, which may shorten the lifetime of RC structures (BONNET *et al*., 2009). Hence, inspection of chloride ingress in RC structures is an important task to determine the level of chloride inside concrete that is useful for optimizing maintenance costs of structures (BASTIDAS-ARTEAGA & SCHOEFS., 2012). Under natural exposure, chloride ingress is related to an important number of uncertainties (BASTIDAS-ARTEAGA *et al*., 2011; SAASSOUH & LOUNIS, 2012) such as: chloride surface concentration, chloride diffusion coefficient, etc. These uncertainties accompany with the variability in space of chloride concentration in concrete require a large number of inspection points. However, in real practice, the inspection just can carried out with a limited number of points due to time-consuming, the expensive of the tests and the difficulties to implement in practice. Therefore, it is necessary to use the available information in the best way for uncertainty quantification by using statistic and/or probabilistic methods. The Bayesian method is a reasonable choice to deal with this problem.
The Bayesian network (BN) is an effective tool for the identification of parameters. Some studies (BASTIDAS-ARTEAGA *et al*., 2012; RICHARD *et al*., 2012) proposed

an approach based on the use of BN allowing the parameter identification from real data and showing an agreement between numerical prediction and experimental measurements. In this study, BN will be also used as a tool for identification of parameters in chloride penetration model. Different configurations and schedules of inspection will be taken into consideration to determine an improved inspection scheme.

2. Bayesian identification and its application to chloride ingress

2.1 Introduction to Bayesian network

Generally, a BN is a specific type of graphical model that is represented as a Directed Acyclic Graph (DAG). Nodes in DAG are graphical representation of objects and events that exists in real world, and they are used to represent variables or states. Causal relations between nodes are represented by drawing an arc (edge) between them. If there is a causal relationship between the variables (nodes), there will be a directional edge, leading from the cause variable to the effect variable. Each variable in the DAG has a Probability Density Function (PDF), which dimension and definition depends on the edges leading into the variables. Figure 1 describes a simple BN that consists of three nodes corresponding to three random variables X_1, X_2 and X_3 in which X_2 and X_3 are children of the parent node X_1. The children nodes have conditional probability distributions that depend on their parent node. The parent node has a marginal probability distribution. The Bayes' rule allows for computing the posterior probability $p(X_1|X_2)$, given the prior and the conditional probabilities $p(X_1)$ and $p(X_2|X_1)$:

$$p(X_1 \mid X_2) = \frac{p(X_2 \mid X_1)\, p(X_1)}{p(X_2)} \tag{1}$$

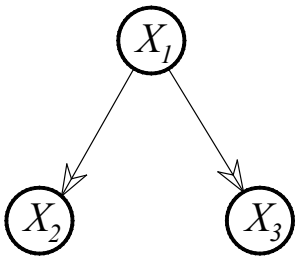

Figure 1. A simple Bayesian network.

2.2 Application to chloride ingress

2.2.1 Chloride ingress and modelling

In saturated concrete, the Fick's diffusion equation (TUUTTI, 1982) is usually used to predict the unidirectional diffusion (in x-direction):

$$\frac{\partial C_{f_c}}{\partial t} = D_c \frac{\partial^2 C_{f_c}}{\partial x^2} \quad (2)$$

where C_{fc} (kg/m^3) is the concentration of chloride dissolved in pore solution, t (year) is the time and D_c (m/s^2) is the effective chloride diffusion coefficient. Assuming that concrete is a homogeneous and isotropic material with the following initial conditions: (1) the concentration is zero at time $t = 0$ and (2) the chloride surface concentration is constant during the exposure time; the free chloride ion concentration *C(x,t)* at depth x after time t for a semi-infinite medium is:

$$C(x,t) = C_s \left[1 - erf\left(\frac{x}{2\sqrt{D_c t}} \right) \right] \quad (3)$$

where C_s (kg/m^3) is the chloride surface concentration and *erf(.)* is the error function.
Equation (3) is just valid when RC structures are saturated and subjected to constant concentration of chlorides on the exposure surfaces. In real structures, these conditions rarely appear because concrete is a heterogeneous material and the chloride concentration in the exposed surfaces could be time-variant. Besides, this solution does not consider chloride binding capacity, concrete aging and other environmental factor such as temperature and humidity (BASTIDAS-ARTEAGA *et al*., 2011). Although this solution neglects some important physical phenomena, this model will be used herein to illustrate the proposed methodology for the identification of random variables using BN. The methodology can be after extended to more realistic chloride ingress models.

2.2.2 Bayesian model of chloride ingress

The chloride ingress could be modelled by the BN described in Figure 2 where C_s and D are the two parent nodes (random variables to identify) labelled number 1 and 2. There are n child nodes $C(x_i, t_j)$ representing the chloride concentration at depth x_i and at inspection time t_j labelled number from 3 to k. The number of child nodes is computed as:

$$n = n_x n_t \quad (4)$$

where n_x is the total number of points in depth and n_t is the total number of inspection times. Assuming that C_s and D are two independent random variables, the values of $C(x_i, t_j)$ could be easily estimated from equation (3). In this BN, the probability of chloride concentration $p(C(x_i, t_j))$ can be calculated as follows (BASTIDAS-ARTEAGA *et al*., 2012; NGUYEN, 2007):

$$p\left(C\left(x_i,t_j\right)\right) = \sum_{D,C_s} p\left(C\left(x_i,t_j\right) \mid D,C_s\right) p\left(D,C_s\right) \quad \text{with} \quad p\left(D,C_s\right) = p\left(D\right) p\left(C_s\right) \quad (5)$$

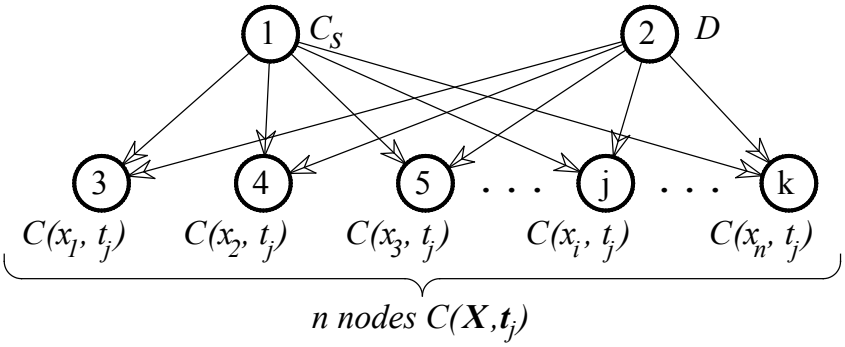

Figure 2. The BN modeling chloride ingress.

To estimate $p(C(x_i,t_j))$, the conditional probability $p(C(x_i,t_j)|D,C_s)$ must be already known in equation (5). This conditional probability is computed based on the Conditional Probability Table (CPT) of the BN. The BN allows entering evidences into the network and then updating the probabilities in the network. In this study, the evidences correspond to measures of chloride concentration at given points and times. Then, the term $p(C(x_i,t_j)|o)$ represents the probability distribution of $C(x_i,t_j)$ given evidence o and a posterior distribution can be computed by applying the Bayes' theorem:

$$p\left(D \mid o\right) = p\left(D \mid C\left(x_i,t_j\right)\right) p\left(C\left(x_i,t_j\right) \mid o\right) \tag{6a}$$

$$\text{with } p\left(D \mid C\left(x_i,t_j\right)\right) = \frac{p\left(C\left(x_i,t_j\right) \mid D\right) p\left(D\right)}{p\left(C\left(x_i,t_j\right)\right)} \tag{6b}$$

and

$$p\left(C_s \mid o\right) = p\left(C_s \mid C\left(x_i,t_j\right)\right) p\left(C\left(x_i,t_j\right) \mid o\right) \tag{7a}$$

$$\text{with } p\left(C_s \mid C\left(x_i,t_j\right)\right) = \frac{p\left(C\left(x_i,t_j\right) \mid C_s\right) p\left(C_s\right)}{p\left(C\left(x_i,t_j\right)\right)} \tag{7b}$$

Determination of these conditional probabilities is carried out by the BN Tool Box which is built on the Matlab® Software.

2.2.3 Parameter identification using BN

We aim at identifying C_s and D from chloride profiles. We assume that C_s and D follow *a priori* uniform distributions defined in given intervals. The intervals for each parameter should contain all possible values and can be defined on the basis of existing databases, similar study cases, or expert knowledge. The assumption of uniform distribution for unknown parameter could avoid making any assumption about distribution shape (BASTIDAS-ARTEAGA *et al*., 2012). Most of parameters in chloride ingress are defined in continuous space. However in order to avoid using approximate inference algorithms which will be a disadvantage when working with

continuous variables, continuous variables must be replaced by discrete random variables (STRAUB, 2009). The discretization of each parameter is described in Table 1.

Data from inspections will be introduced to the BN as evidences. This data could be obtained both from experimental measurements or expert knowledge. In this study, for the purpose of generalization the optimization, the numerical evidences obtained from known input random variables will be generated with a sufficient number of simulations. The probability that $C(x_i,t_j)$ belongs to a given interval for different depth is then computed for the identification of the term $p(C(x_i,t_j)|o)$.

In this study, numerical evidences were generated using Monte Carlo methods with parameters given in Table 2. The mean values for each parameter were taken from (BASTIDAS-ARTEAGA *et al*., 2009). However, the COV for each parameter were reduced to 20% and 15% for C_s and D, respectively. This is due to the fact that within one type of concrete, the variation was narrowed. The assumption that C_s and D follow lognormal distributions is also in agreement with some other researches (DURACRETE, 2000; VU & STEWART, 2000). These assumptions were used to generate 10000 random values for C_s and D corresponding to 10000 independent inspections points.

Table 1. Discretization of parameters.

Parameters	***Number of intervals***	***A priori distribution***	***Range***
C_s (kg/m^3)	16	Uniform	(0.5; 8)
D (m/s²)	20	Uniform	(6×10^{-13}; 3×10^{-12})
$C(x_i,t_j)$ (kg/m^3)	-	-	(0 ; 8)

Table 2. Theoretical values of parameters to identify

Parameters	***Distribution***	***Mean***	***COV***	***Standard deviation***
C_s	Lognormal	2.95 (kg/m²)	20%	0.59
D	Lognormal	1.33×10^{-12} (m/s)	15%	0.2×10^{-12}

3. Selection of configuration in Bayesian network

In this section, different configurations of the BN corresponding to different inspection schemes will be analysed for selecting inspection schemes that provide the best estimation for parameters. Each configuration will be evaluated by the error of the identified parameter $Z_{identified}$ with respect to the theoretical value Z_{theory} as:

$$Error(Z) = \frac{Z_{identified} - Z_{theory}}{Z_{theory}}.100\% \quad (8)$$

where Z represents the mean or the standard deviation of the parameter to identify.

3.1 Identification using one point in depth of inspection

In this part, the estimation of the chloride surface concentration (C_s) and chloride diffusion coefficient (D) will be analysed from evidences obtained at one depth point. Figure 3a shows the chloride profile computed from equation (3) where C_s and D have the values as shown in Table 2 at an inspection time t_{ins} = 10 years. The total inspection depth is assumed at 12cm. At deeper point, the chloride content is almost zero. The BN now consists of three nodes: two parent nodes are C_s and D, one child node $C(x_i,t_j)$ representing for chloride concentration at depth x_i and time $t_{ins} = t_j = 10$ years.

3.1.1 Convergence of the BN with the number of intervals

As previously mentioned in 2.2.3, continuous variables need to be discretised into equivalent intervals. The number of intervals could be adjusted to obtain the balance between accuracy of results and time used for computations. When a more accurate result is expected, a high value of number of intervals is often chosen. Figure 3b describes the estimations of the error of the mean value of C_s with different discretisation and inspection depths. It is clear that, no fluctuation is recorded in the case in which each node $C(x_i,t_j)$ in the BN is divided into 200 intervals. This means that a high number of intervals could lead to a convergence in BN. Consequently, we will keep 200 intervals for node $C(x_i,t_j)$ for all BNs in this part. This numerical implementation can be seen as a suboptimal for the estimation.

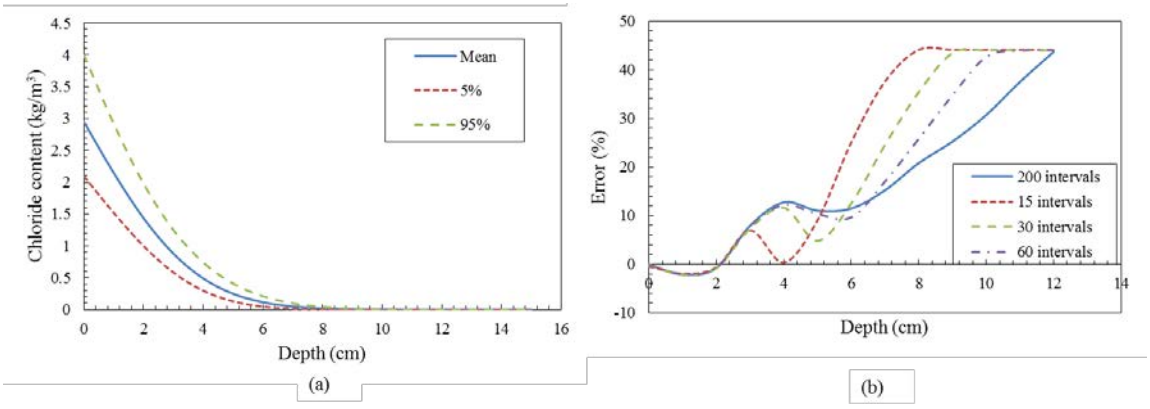

Figure 3. (a) Chloride profile at t=10 years - (b) The convergence of BN with different intervals.

3.1.2 Analysis of the results

Figure 4 shows the error in the identification of the mean and standard deviation of C_s and D. For C_s, evolution of the error of both mean and standard increases with the depth. These estimations are corresponding to the evolution of the chloride profile in Figure 3a where the chloride content reduces as depth increasing. This means that, data from chloride profiles near the surface will provide more information for the updating,

whereas less information in the deeper parts increases the error in the identification. When the chloride content is closed to zero, these errors will reach the highest value closing to 40% for the mean. On the contrary, with the evidences near the surface ($x \approx 0$), we can obtain the best estimation for the mean and standard deviation of C_s, with errors are 1% and 3% respectively. This might due to the fact that in equation (3), when we set $x \approx 0$, $C(x_i,t_j) \approx C_s$. Consequently, the chloride concentration at the surface is most valuable in the identification of C_s and the BN will put more weight on the evidence obtained at $x = 0$ cm.

It is also observed in Figure 4 that the error in the identification of D decreases when the depth $x < 8$cm and after increases. This behaviour corresponds to the fact that chloride content at deeper parts is more useful in predicting the diffusion coefficient. However, at deep points where the chloride contents are close to zero for t_{ins} =10 years, the errors will increase. The errors in the identification of the standard deviation of D followed similar behaviours, however their values are very far from the theoretical values with important errors (more than 200%). Therefore, it can be concluded that it is impossible to perform a good identification of D using evidences obtained from only one point in depth.

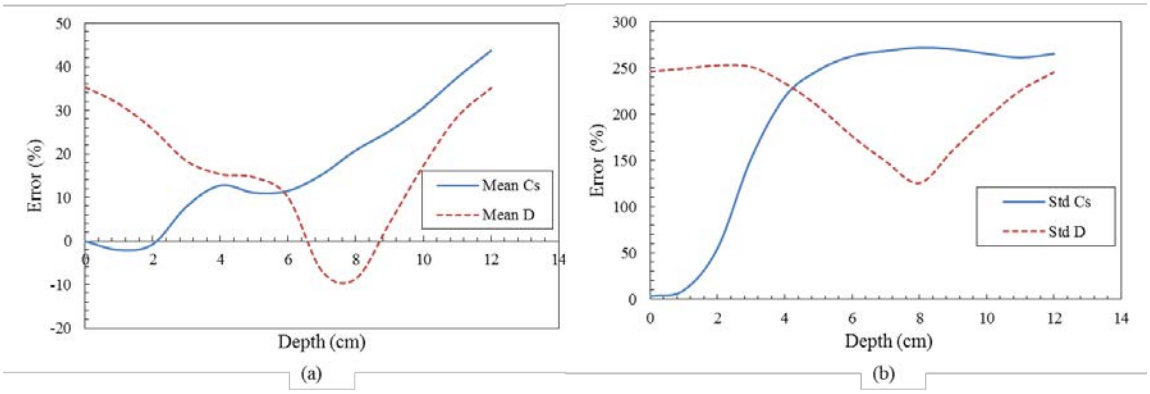

Figure 4. Error in identification using one depth point: (a) Mean – (b) Standard deviation.

3.2 Identification using full inspection depth

3.2.1 Using the same boundary for all the child nodes

In this section, the identification in BN will use data from total inspection depth. The total inspection depth (12 cm) is divided into intervals to select several points for updating the BNs. The intervals length should not be smaller than 0.3cm due to the accuracy of the equipment for determining chloride profiles. The BNs will now have the number of child nodes equal to the number of measurements in depth. Table 3 describes different cases considered in this part.

Figure 5 shows the error in the identification using full inspection depth with the same boundaries for the child nodes. It is noted that there is no remarkable change in the identification of the mean of C_s (Figure 5a) because the errors in 5 surveyed cases are under estimated at approximate -5%. Meanwhile it seems that increasing the number of measurements in the inspection depth might produce more errors for the standard deviation of C_s (Figure 5b).

For chloride diffusion coefficient, the gap between identified values and theoretical values for D are reduced significantly when the size of the discretization intervals is smaller. The errors in the estimation of the mean of D are less than 5% when the intervals are smaller than 0.5 cm. The standard deviation of D also reveals a better evolution when the error decreases from more than 200% with $\Delta x = 3cm$ to about 20% with $\Delta x = 0.3cm$. This behaviour is expected because when the inspection depth is divided into small intervals, we could obtain more information describing the level of chloride ingress that is useful for characterizing the diffusion coefficient. Hence, data from full inspection depth could be more useful in the identification of D.

Table 3. Different discretization cases and number of points in depth.

Case	***Δx (cm)***	***Discretization***	***Number of points in depth***
1	*0.3*	*0:0.3:12*	*41*
2	*0.5*	*0:0.5:12*	*25*
3	*1*	*0:1:12*	*13*
4	*2*	*0:2:12*	*7*
5	*3*	*0:3:12*	*5*

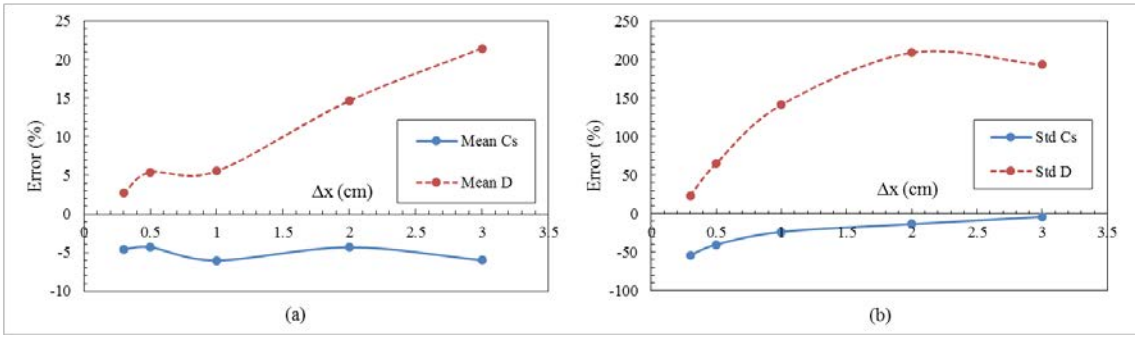

Figure 5. Error in identification using full inspection depth: (a) Mean – (b) Standard deviation.

3.2.2 Using different boundaries for child nodes

As discussed in section 3.2.1., when the inspection depth is divided into large intervals, the errors in the estimation of chloride diffusion coefficient increase significantly. In this case, a high value for the number of intervals will be required to obtain a better estimation. Nevertheless, increasing the number of intervals will increase the size of the

CPTs, which might also increase the computational time for the BN. TRAN *et al* (2014) proposed an algorithm to use different ranges for child nodes and keep the number of intervals in the discretization at a low value to reduce the computational time. By this means, we can optimize the information in deep points for building the evidences.
Figure 6 shows the comparison in the estimation of the mean and standard deviation of *D*. It is clear that, in both cases: using a large number of intervals and using different ranges, the errors are reduced remarkably despite of the inspection length was divided into large intervals. This proposed approach could be very useful in practice when the number of measured points is limited or when the chloride content penetrated into concrete is low.

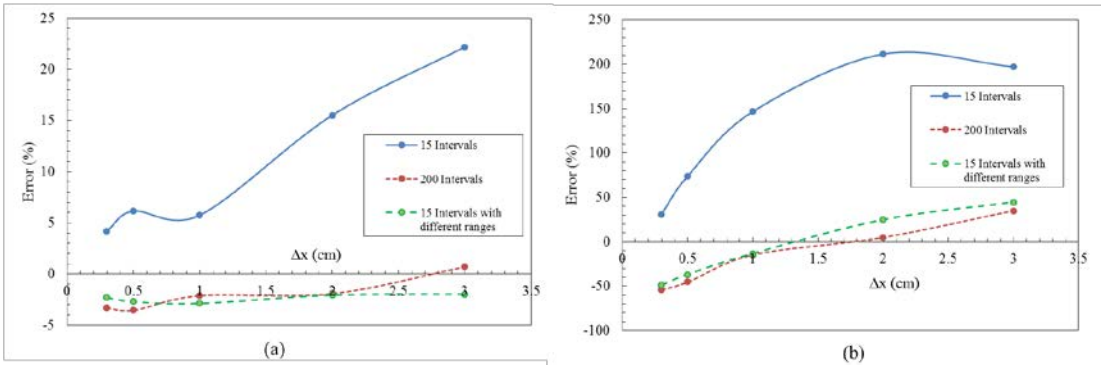

Figure 6. Error estimation between 3 cases for D: (a) Mean – (b) Standard deviation.

3.2.3 Using evidences from different inspection times
In this section, evidences obtained from various inspection times will be introduced in the BN for the identification process. According to section 3.2.2, various boundaries were used for each child node in the BN with a sufficient number of intervals to minimize the fluctuation effects/errors in the results.
From Figure 7, it can be seen that the inspection time t_{ins} influences the estimation of both the mean and standard deviation of *D*. The identification is improved when t_{ins} increases until arriving at an optimal inspection value $t_{ins,opt}$ that varies between 30 and 40 years for the identification of the mean and standard deviation. This phenomenon can be explained by the fact that when $t_{ins} \approx 35$ years the chloride concentration in the total inspection length is sufficient for describing the chloride ingress process – *i.e.*, there is sufficient chloride content at each point in the space to improve the identification. When $t_{ins} > 40$ years, the chloride content at $x = 12$ cm is larger than zero and therefore the identification errors increase because the inspection length is not large enough to describe the problem.

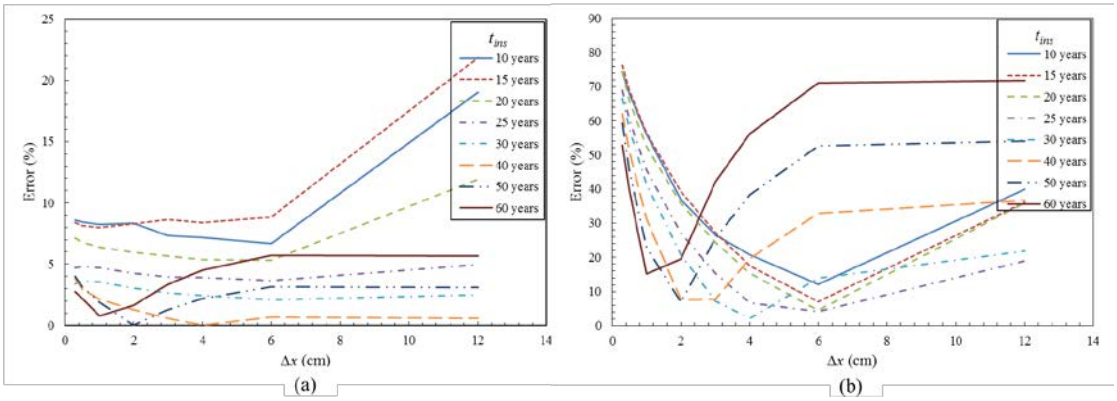

Figure 7. Error estimation for D with evidences from different inspection times: (a) Mean - (b) Standard deviation.

It is also worth noticing that there is an optimal value Δx_{opt} for each inspection time. The optimum value decreases when t_{ins} increases. This is related to the fact that for larger t_{ins} the chloride content inside the total inspection length is larger. Consequently, it is necessary to add more information to improve the representation of the chloride profile. It is also noted that the error is larger for smaller values of Δx in comparison with the Δx_{opt}. There is no remarkable change in the estimation the mean value of D when Δx vary from 0.3 cm to the optimal value. However, the variation is more important for the identification of the standard deviation of D. This is due to the small size of Δx could provide the identification errors between the two adjacent points (TRAN *et al*., 2014). When Δx is larger than the optimal value, the errors for both mean and standard deviation increase because the information becomes poor for describing the chloride ingress process.

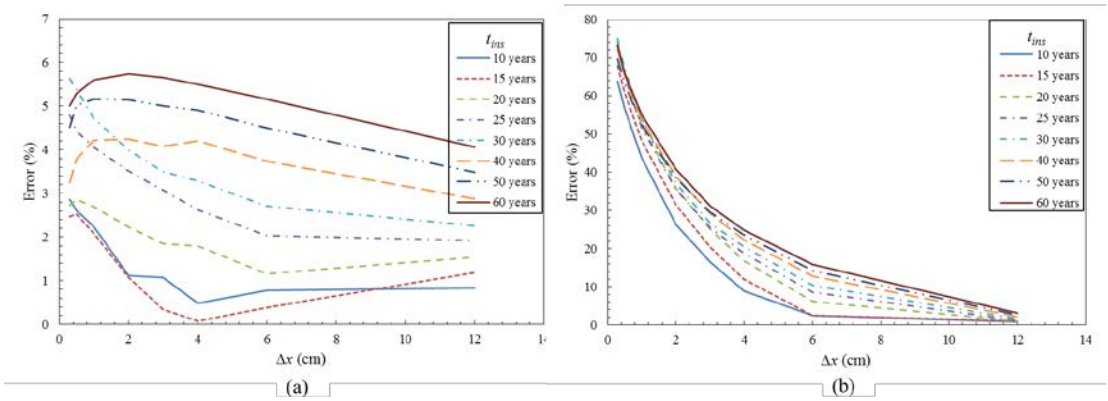

Figure 8. Error estimation for C_s with evidences from different inspection times: (a) Mean - (b) Standard deviation.

For C_s, the results presented in Figure 8 reveal that to obtain a good estimation of C_s, it is better to use the evidences at early inspection times. This is because the chloride

surface concentration does not depend on the time of inspection and at early inspections times the chloride concentration in the neighbouring of the concrete surface will be close to C_s.

4. Conclusions

The penetration of chloride is one of the main causes inducing corrosion of RC structures. The identification of parameters in chloride ingress modelling is crucial in predicting chloride ingress into concrete that will help to optimise the maintenance of structures exposed to chloride-contaminated environments. Inspection data used for the identification is very limited due to time-consuming and expensive tests. Therefore, it is necessary to use these data in an optimal scheme. Within this framework, the BN could provide a possibility to identify model parameters with different information. In this study, results based on numerical evidences revealed that there are optimal configurations of BN for the identification of each parameter (C_s or D). For C_s, an early inspection with one point close to the surface could provide a good identification. For D, to obtain identified values close to the theoretical values, the identification should use the evidences from full inspection depth. At a specific inspection time, there is an optimal discretization for the inspection length that could provide the best estimation for D. These optimal configurations could be combined to improve the identification of the model parameters.

5. References

BASTIDAS-ARTEAGA E., BRESSOLETTE P., CHATEAUNEUF A., SANCHEZ-SILVA M. (2009). *Probabilistic lifetime assessment of RC structures under coupled corrosion–fatigue deterioration processes*. Structural Safety, Vol. 31(1), pp 84–96. http://dx.doi.org/10.1016/j.strusafe.2008.04.001

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, Vol. 33, pp 720–730. http://dx.doi.org/10.1016/j.engstruct.2010.11.008

BASTIDAS-ARTEAGA E., SCHOEFS F. (2012). *Stochastic improvement of inspection and maintenance of corroding reinforced concrete structures placed in unsaturated environments*. Engineering Structures, Vol. 41, pp 50–62. http://dx.doi.org/10.1016/j.engstruct.2012.03.011

BASTIDAS-ARTEAGA E., SCHOEFS F., BONNET S. (2012). *Bayesian identification of uncertainties in chloride ingress modeling into reinforced concrete structures*. In Third International Symposium on Life-Cycle Civil Engineering. Vienna, Austria.

BONNET S., SCHOEFS F., RICARDO J., SALTA M. (2009). *Effect of error measurements of chloride profiles on reliability assessmen*t. In: 10th International Conference on Structural Safety and Reliability. Osaka, Japan.

DURACRETE (2000). *Probabilistic calculations. DuraCrete—probabilistic performance based durability design of concrete structures. EU—brite EuRam III*. Contract BRPR-CT95-0132. Project BE95-1347/R12-13.

NGUYEN X.S. (2007). *Algorithmes probabilistes appliqués à la mécanique des ouvrages de génie civil*. PhD Thesis, INSA de Toulouse.

RICHARD B., ADELAIDE L., CREMONA C. (2012). *A Bayesian approach to estimate material properties from global statistical data*. European Journal of Environmental and Civil Engineering, (March 2013), pp 37–41.

SAASSOUH B., LOUNIS Z. (2012). *Probabilistic modeling of chloride-induced corrosion in concrete structures using first- and second-order reliability methods*. Cement and Concrete Composites, Vol. 34(9), pp 1082–1093. http://dx.doi.org/10.1016/j.cemconcomp.2012.05.001

STRAUB D. (2009). *Stochastic Modeling of Deterioration Processes through Dynamic Bayesian Networks*, (October), Journal of Engineering Mechanics, Vol. 135(10), pp 1089-1099 http://dx.doi.org/10.1061/(ASCE)EM.1943-7889.0000024

TRAN T.B, BASTIDAS-ARTEAGA E., SCHOEFS F., BONNET S. (2014). *Bayesian updating for optimization of inspection schedules of chloride ingress into concrete*. In: Second International Symposium on Uncertainty Quantification and Stochastic Modeling. Rouen, France.

TUUTTI K. (1982). *Corrosion of steel in concrete*. Swedish Cement and Concrete Institute.

VU K.A.T., STEWART M.G. (2000). *Structural reliability of concrete bridges including improved chloride-induced corrosion*. Structural Safety, Vol. 22, pp 313–333.