UNIVERSITÀ DEGLI STUDI DI PALERMO

Meccanica dei Materiali e delle Strutture
Vol. VI (2016), no.1, pp. 227-234
ISSN: 2035-679X
Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, Dei Materiali
DICAM

TOLERANCE LIMITS OF SPATIALLY CORRELATED SAMPLES FOR THE ASSESSMENT OF CONCRETE STRENGTH

Massimo F. Bonfigli, Marco Breccolotti e Annibale L. Materazzi

Dipartimento di Ingegneria Civile e Ambientale (DICA)
Università degli Studi di Perugia
Via G. Duranti 93, 06125 Perugia, Italy
e-mail: federico.bonfigli@strutture.unipg.it

Key words: In-situ concrete strength, Random fields, Tolerance limits, Correlated samples.

Parole chiave: Resistenza in-situ di calcestruzzi, Campi Stocastici, Limiti di tolleranza, campioni correlati.

Abstract. *The assessment of in-situ concrete strength is a key step for the evaluation of the safety of any existing RC building. The distribution of concrete strength in a structure can be reasonably assumed to be a realization of a random field with a given correlation function. Nevertheless, all current approaches used in practice for the assessment of in-situ concrete strength typically neglect this aspect. This work presents a probabilistic tool to estimate lower tolerance limits for the evaluation of a given percentile of the population of concrete strength with a pre-established confidence level taking into account the spatial correlation of the samples. The basic assumption is that concrete strength is modelled as a Gaussian random field with a known correlation function. The results are a generalization of the traditional tolerance intervals for uncorrelated samples. Two examples are furthermore presented to illustrate the potential loss of confidence if correlation of core test values is neglected.*

Sommario. *La valutazione della resistenza in-situ di calcestruzzi è un'operazione cruciale per qualsiasi valutazione di sicurezza strutturale di edifici in CA esistenti. La distribuzione della resistenza del calcestruzzo in una struttura può essere vista come una realizzazione di un campo stocastico con una propria funzione di correlazione. Tale aspetto viene tuttavia ignorato da tutti gli approcci attualmente adottati per la valutazione della resistenza dei calcestruzzi in-situ. Questo documento fornice uno strumento probabilistico per la stima di limiti di tolleranza inferiori per la valutazione di un fissato percentile della popolazione delle resistenze in-situ del calcestruzzo, dato un fissato livello di confidenza e tenendo in considerazione la correlazione spaziale dei campioni. Le ipotesi di base sono quelle di resistenze del calcestruzzo distribuite secondo un campo stocastico gaussiano con una legge di correlazione nota. I risultati ottenuti sono una generalizzazione dei tradizionali intervalli di tolleranza per campioni non correlati. Due esempi applicativi sono inoltre forniti per illustrare la potenziale perdita di confidenza nelle stime qualora la correlazione spaziale delle resistenze sia trascurata.*

1 INTRODUCTION

In many countries there is an increasing need of assessing the structural capacity of existing buildings, either to allow them to host new functions or to simply check their safety against load cases that were originally neglected during the design stage (typically the seismic action). The evaluation of reinforced concrete structures must always begin with an assessment of the geometry and of the material properties. Often the original design documents are not available anymore, so that there is no information on the materials that were used. To overcome this difficulty destructive tests must be carried out on the structures, typically consisting in the extraction and testing material samples. For what it concerns concrete, the samples consist in cores drilled from the structural elements in various locations and then tested in compression testing machines. The fundamental assumption is that the measured compressive strength of the extracted core is representative of the uniaxial compressive strength of the material within the structure itself. Afterwards, the measured values have to be statistically interpreted so to obtain a single estimation of in-situ compressive strength that is suitable for the structural assessment according to the relevant building codes. Several national and international standards are today available for the interpretation of measured concrete strength values, and for such a task they propose empirical, theoretical or mixed approaches. All of them, however, somewhat assume that the measured samples are independently drawn from a single distribution, whereas in reality it is much more reasonable to think of concrete strength in a structure as a realization of a random field with a given correlation structure. This implies that the samples are somewhat correlated and thus their relative spatial location may reduce (if the samples are quite close) the total amount of information (and consequently increasing the uncertainties) they provide. The most basic statistical theory on which are based several criteria for the assessment of concrete strength is that of tolerance intervals. In the following sections an extension of this approach will be derived for the case of correlated samples with a known correlation function. This theory is intended to be a first step to develop assessment tools and criteria that are aware of the importance of selecting proper spatial locations for the concrete cores to be extracted so to reduce the risk of overestimating the material strength.

2 TOLERANCE LIMITS

2.1 Basic definitions

Several international codes (see e.g. ACI 214.4R-10[1], EN 13791:2007[2]) estimate the in-situ concrete strength by means of an estimator of the type:

$$f_{p,\alpha} = \mathbf{x}_m - k\mathbf{s} \tag{1}$$

where $\boldsymbol{x}_m$ is the sample mean, $\boldsymbol{s}$ is the sample standard deviation, k is a proper coefficient. The value $f_{p,\alpha}$ is the estimation of the p^{th} percentile of concrete strength f_p with a desired (low) probability of overestimation (1-α), i.e. is such that:

$$P(\mathbf{x}_m - k\mathbf{s} \le f_p) = 1 - \alpha \tag{2}$$

A comprehensive description of the statistical theory of tolerance intervals has been provided by Guttman[3]. The core of the problem is the determination of k as a function of the number of samples n, of the percentile p and of the desired confidence α. For the case of normally distributed and independent samples the well-known solution is given by:

$$k(n,p,\alpha) = t^{-1}_{n-1,1-\alpha}(z_{1-p}\sqrt{n})/\sqrt{n} \tag{3}$$

where $z_x = \Phi^{-1}(x)$ is the inverse cumulative distribution function of a standard normal random variable evaluated in x and n is the number of samples. If the samples are correlated it is expected that k additionally depends on the correlation matrix of the samples.

2.2 Notation

Some elementary notation must be introduced before proceeding into the theoretical derivation of the theory. The set of measured core strengths can be expressed by a column vector X as:

$$X = (\mathbf{x}_1,...,\mathbf{x}_n)^T \tag{4}$$

By means of row vector W of equal weights, the mean $\mathbf{x}_\mathrm{m}$ and standard deviation $\mathbf{s}$ of the sample strengths are expressed by:

$$\mathbf{x}_m = n^{-1}\sum \mathbf{x}_i = WX \tag{5}$$

$$\mathbf{s} = \sqrt{\sum (x_i - \mathbf{x}_m)^2/(n-1)} = \sqrt{X^T(I_n - \mathbf{1}W)X(n-1)} \tag{6}$$

where I_n is the nxn identity matrix and $\mathbf{1}$ is a column vector of all ones. The correlation matrix of the n samples at locations r_i is given by:

$$C = \begin{pmatrix} \rho(r_1 - r_1) & \cdots & \rho(r_1 - r_n) \\ \vdots & \ddots & \vdots \\ \rho(r_n - r_1) & \cdots & \rho(r_n - r_n) \end{pmatrix} \tag{7}$$

Furthermore, the average of the out of diagonal elements of C is expressed as:

$$\rho_m = [n(n-1)]^{-1}\sum_{i\neq j}\rho(\mathrm{r_i} - r_j) \tag{8}$$

2.3 Derivation of a tolerance limit for correlated samples

The objective of this section is to derive a general expression for the coefficient k of equation (1) that takes into consideration the correlation matrix so that $f_{p,\alpha}$ satisfies equation (2) even if the samples are no more uncorrelated. The fundamental assumptions are that in-situ concrete strength distribution can be reasonably modeled as a homogeneous isotropic Gaussian random field and that the correlation law of the process is available. This latter hypothesis is rather strong because up to now very little is known about the spatial structure of the concrete strength distribution. Many authors in literature[4,5] have modeled concrete strength using Gaussian or exponential correlation laws, but these assumptions are not based on solid experimental investigations. Additionally, in the very few experimental campaigns that can be found in literature discordant results are found. Rackwitz[6] showed that the spatial correlation disappears at a distance of 10m, whereas others[7] recorded a quite lower distance. Nevertheless, the aim of the work is to provide a basic theoretical framework to be used as a starting point to address the problem of the spatial correlation of the strength measurements to improve the confidence in the estimates; further experimental studies are necessary to investigate the specific properties of concrete strength distribution.

The derivation begins from the definition given by equation (2). If from both terms the unknown mean μ of the field is subtracted and then they are divided by the unknown standard deviation of the field σ it results:

$$P\left(\frac{\mathbf{x}_m - \mu}{\sigma} - \frac{k\mathbf{s}}{\sigma} \le z_p\right) = 1 - \alpha \tag{9}$$

If a normal random variable $\mathbf{z}$ is defined as**:**

$$\mathbf{z} = (\mathbf{x}_m - \mu)/\sigma = (WX - \mu)/\sigma \tag{10}$$

then eq. (9) can be rewritten as:

$$P\left(\frac{\mathbf{z} + z_{1-p}}{\mathbf{s}/\sigma} \le k\right) = 1 - \alpha \tag{11}$$

Now it is recalled that a vector of correlated normal random variables can always be expressed, by means of a proper decomposition of the correlation matrix, as a linear combination of a vector Y of independent standard normal random variables, so that:

$$X = \sigma C^{1/2} Y + \mathbf{1}\mu \tag{12}$$

In the preceding equation the C matrix has been decomposed using the principal square root matrix. Combining eq. (12) with eq. (6) the following is obtained:

$$\mathbf{s} = \sigma \sqrt{\frac{Y^T C^{1/2} (I_n - \mathbf{1}W) C^{1/2} Y}{n-1}} = \sigma \sqrt{\frac{Y^T B Y}{n-1}} \tag{13}$$

It can be easily noticed that the matrix $B=C^{1/2}(I_n\text{-}\mathbf{1}W)C^{1/2}$ has real eigenvalues and eigenvectors, and its spectral decomposition will thus be denoted by:

$$B = K^T \Lambda K \tag{14}$$

Replacing eq. (12) and eq. (14) into, respectively, equations (10) and (13), inserting these latter into eq. (11) and finally performing some basic mathematical manipulations the following expression is obtained:

$$P\left(\frac{\left(WC^{1/2}Y + z_{1-p}\right)\sqrt{n^2/c^*}}{\sqrt{(KY)^T \Lambda K Y}} \le k\sqrt{\frac{n^2}{(n-1)c^*}}\right) = 1 - \alpha \tag{15}$$

where c^* simply denotes the sum of the correlation matrix entries. It is important to notice that eq. (15) describes a pivotal quantity, i.e. a quantity that does not depend anymore on the unknown parameters of the field, namely its mean and standard deviation. The only value that remains unknown is k, that is the quantity that needs to be determined. A random variable $\mathbf{u}$ is then defined as follows:

$$\mathbf{u} = \frac{\left(WC^{1/2}Y + z_{1-p}\right)\sqrt{n^2/c^*}}{\sqrt{(KY)^T \Lambda K Y}} \tag{16}$$

Its CDF is in general rather complex, but it can very easily be determined using Monte-Carlo simulations. If its inverse distribution evaluated in x is denoted by $F^{-1}(x)$, then the value of k to be used is given by:

$$k = k(n, p, \alpha, C) = \sqrt{\frac{(n-1)[(n-1)\rho_m + 1]}{n}} F^{-1}_{1-\alpha} \tag{17}$$

where the term c^* has been replaced by the out-of-diagonal average of the terms of the correlation matrix ρ_m.

According to this theory, the evaluation of a lower tolerance limit for correlated samples is carried out first estabilishing the percentile of interest p and the target confidence in the estimate α, then determining numerically the CDF of $\boldsymbol{u}$ evaluated in $(1-\alpha)$ using eq. (16) and replacing it into eq. (17) to obtain k. The estimate is finally given by eq. (1).

2.4 Theoretical considerations

The introduction of the parameter ρ_m is very convenient, because it is a single valued term that ranges between 0 and 1 and it gives an overall view on the average level of spatial correlation of the samples.

However this indicator, while very useful, does not provide a complete view on the overall amount of uncertainty involved in the sampling scheme. In order to explain this fact the following example is presented. Let's assume that it is required to estimate the 5% percentile of concrete strength testing a 7.5m by 7.5m RC slab using three samples with a target confidence of 90%. Infinite different spatial layouts for the samples could be chosen: in this case layout (a) and (b) of Figure 1 are being compared. If it is assumed that the correlation law follows a Gaussian correlation of the type $\rho(x)=exp(-x^2/4)$, then both layout (a) and (b) have practically the same level of average spatial correlation $\rho_m=0.31$.

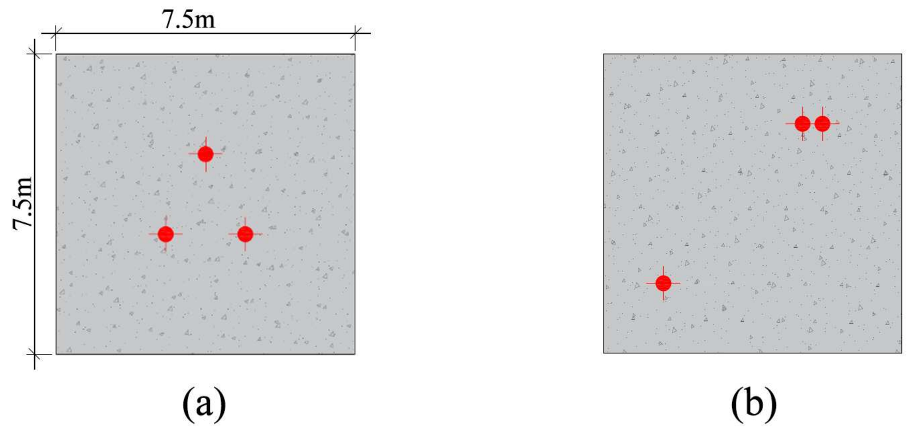

Figure 1: Two different sampling layouts having the same level of average spatial correlation ρ_m =0.31.

The required estimate is performed in both cases using equation (1), but the value of the coefficient k to be used in the estimations must be evaluated according to the aforementioned procedure, which yields significantly different values for the two cases. In the case of layout (a) it is obtained k=6.59, whereas in case (b) k=9.63 must be used. The higher value of k is due to the fact that layout (b) yields less information on the properties of the underlying field (the two very close samples practically provide the same data), and as a result a higher value of k has to be used to reach the same target confidence.

This example is useful to highlight the fact that the spatial layout has a deep influence on the outcome of the evaluation, and simply selecting random locations for the samples (as it is suggested in several current codes) may produce suboptimal situations similar to that of layout (b). In particular, according to the theoretical framework here presented, the main idea that should be always kept into consideration when choosing a sampling scheme is to select locations that are as spread out as possible (so to lower the value of ρ_m) and also so that the minimum distance between any two samples is as high as possible (so to increase the amount of unique information on the field provided by the sample). It is thus desirable that current codes in future will provide more directions on this topic.

3 CASE STUDIES

In this section two examples are presented to illustrate the potential loss of confidence in the estimates if the spatial correlation of samples is neglected.

3.1 Example 1

In this first case a 12m x 16m reinforced concrete slab is considered. The objective is to obtain an estimation for the 5% percentile of concrete strength with a probability of overestimation of 10% (i.e. a confidence of 90%). The number of concrete cores that will be used is 12 and it is assumed that they are disposed in a grid layout as in Figure 2.

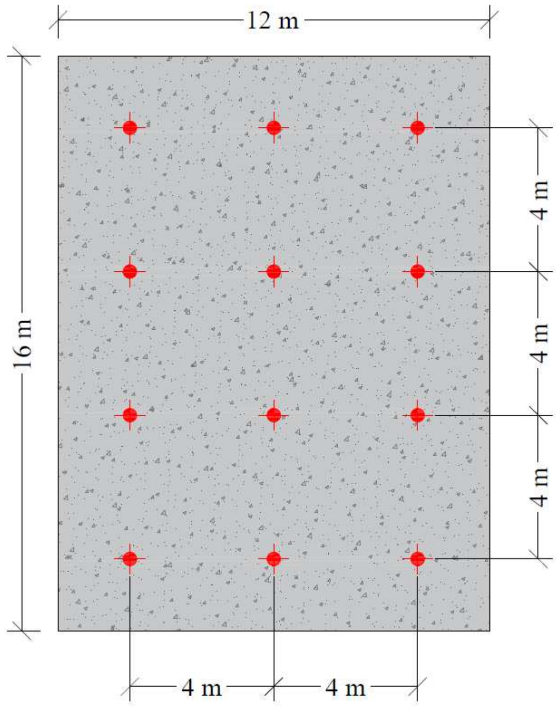

Figure 2: Sampling layout of example 1. Twelve cores are to be extracted from a 12m x 16m slab. The samples are placed in a 4x3 grid with 4m spacing between rows and columns.

The estimate is carried out first by neglecting the spatial correlation and thus evaluating k using the eq. (3), for which it is obtained $k=2.44$. Afterwards, the target confidence that should have been chosen so to obtain the same value of k considering the spatial correlation of the samples is numerically evaluated. This analysis is useful since it allows to estimate the loss of confidence that results from having neglected the dependence of the measured sample values. In the investigations it is assumed that the correlation follows a Gaussian law of the type $\rho(x)=exp(-x^2/d^2)$ or an exponential correlation law $\rho(x)=exp(-|x|/b)$ and several different values for the parameters d and b are evaluated so to cover situations with a modest spatial correlation and situations where the dependence between the cores is rather strong. The results of the analyses are summarized in tables 1 and 2.

Gaussian correlation $\rho(x)=exp(-x^2/d^2)$		Exponential correlation $\rho(x)=exp(-\|x\|/b)$	
d (m)	Actual confidence (1-α)	b (m)	Actual confidence (1-α)
2 m	0.90	1.5 m	**0.85**
3.5 m	**0.83**	2.5 m	**0.64**
5 m	**0.71**	4 m	**0.40**
10 m	**0.42**	7 m	**0.22**

Table 1: Example 1 – Achieved confidence if the spatial correlation of samples is neglected. The boldface numbers highlight situations in which the actual confidence is significantly lower than the target one (90%).

Observing the values it is clear how the correlation of the samples can potentially reduce the achieved confidence depending on the actual correlation structure. A value of $d=3.5m$ for a Gaussian correlation law yields a confidence that is slightly lower of the target one, whereas a value of d equal or greater than $5m$ (which results in a correlation that is consistent with some observations by Rackwitz[6]) yields a confidence of 71%, that is 29% lower than the desired one. With the exponential law results are even more dramatic since the decay in correlation is slower.

3.2 Example 2

Differently from the classical theory of tolerance intervals for independent samples, now the spatial configuration of these has a deep impact on the results. Because of this, in the second example a slightly different spatial configuration of the twelve cores is being investigated. The new configuration is shown in Figure 3.

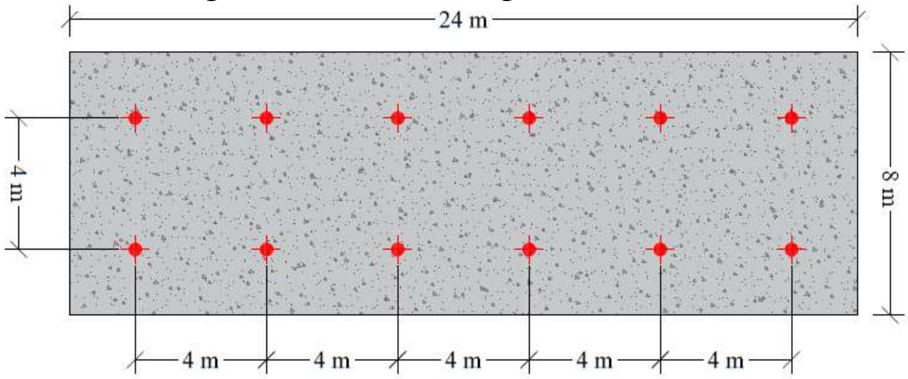

Figure 3: Sampling layout of example 2. Twelve cores are to be extracted from a 24m x 8m slab. The samples are placed in a 6x2 grid with 4m spacing between rows and columns.

The same kind of analysis described in example 1 has been carried out, and its results are summarized in Table 2. These figures are very similar to those of the previous example. Again, if the spatial correlation is neglected and the classical theory of tolerance intervals is being used a sensible loss of confidence may recorded depending on the actual correlation structure of concrete.

Gaussian correlation $\rho(x)=exp(-x^2/d^2)$		Exponential correlation $\rho(x)=exp(-\|x\|/b)$	
d (m)	Actual confidence (1-α)	b (m)	Actual confidence (1-α)
2 m	0.90	1.5 m	**0.86**
3.5 m	**0.84**	2.5 m	**0.66**
5 m	**0.73**	4 m	**0.43**
10 m	**0.46**	7 m	**0.24**

Table 2: Example 2 – Achieved confidence if the spatial correlation of samples is neglected. The boldface numbers highlight situations in which the actual confidence is significantly lower than the target one (90%).

4 CONCLUSIONS

In this work the influence of the spatial correlation of concrete strength on the estimation of the in-situ concrete strength using cores extracted from a structure has been investigated. The classical formulas for tolerance limits of normally distributed independend samples has been extended to the case of correlated samples. This basic framework has been proposed as a starting point to further investigate the problem and to develop feasible workarounds for practical applications.

Furthermore, it has been shown that neglecting the spatial correlation of values may potentially lead to a level of confidence in the estimates that is significantly lower than the target one. Additional extensive experimental investigations are however required to propertly describe the spatial probabilistic structure of strength of concrete cores.

BIBLIOGRAFIA

[1] American Concrete Institute, *Guide for Obtaining Cores and Interpreting Compressive Strength Results*. ACI 214.4R-10.

[2] CEN, *Assessment of in-situ compressive strength in structures and precast concrete components*. EN 13791:2007.

[3] I. Guttman, *Statistical Tolerance Regions: classical and Bayesian,* Charles Griffin & Company, 1970.

[4] M. G. Stewart, "Extent of spatially variable corrosion damage as an indicator of strength and time-dependent reliability of RC beams", *Engineering Structures*, **21**,198-207 (2009).

[5] A. Firouzi and A. R: Rahai, "Prediction of extent and likelihood of corrosion-induced cracking in reinforced concrete bridge decks", *Int. J. of Civ. Eng.*, **9**(3), 183-192 (2010).

[6] R. Rackwitz and K. F: Müller, "On the correlation and autocorrelation of in situ strength and standard strength test results", *Proc. Quality Control of Concrete Structures,* Stockholm (1979).

[7] K. Vu and M. G: Stewart, "Service life prediction of reinforced concrete structures exposed to aggressive environments", *9th Int. Conf. on Durability of Building Materials and Components*, Brisbane (2002).