



Recommendation of RILEM TC249-ISC on non destructive in situ strength assessment of concrete

Denys Breyse · Jean-Paul Balayssac · Samuele Biondi · David Corbett ·
Arlindo Goncalves · Michael Grantham · Vincenza A. M. Luprano ·
Angelo Masi · Andre Valente Monteiro · Zoubir Mehdi Sbartai

Received: 6 December 2018 / Accepted: 5 June 2019 / Published online: 25 June 2019
© RILEM 2019

Abstract This recommendation is written to improve the assessment of the in situ. Compressive strength of concrete in existing structures by combining core strength values and non-destructive measurements. Both average strength and its scatter are considered. Deriving a characteristic strength from the assessment results is not considered here. The recommendation applies for most common techniques (ultrasonic pulse velocity, rebound hammer, pull-out) but also for less common techniques (penetration test,

etc.). The recommendation does not apply to situations in which no core has been taken from the existing structure and is limited to situations where NDT is combined with cores. The recommendation introduces the concept of Estimation Quality Level, corresponding to the target of assessment, and which is put in relation with the means and strategy developed for assessing concrete. The text details all steps that must be followed from the data gathering to the checking of the quality of the final estimations. For more clarity, an

Contributors Maitham Al Wash, Elena Candigliota, Leonardo Chiauzzi, Vincent Garnier, Andrzej Moczko, Xavier Romao, Emilia Vasaneli, Adorjan Borosnyoi, Katalyn Szilagyi, Said Kenai, Valerio Pfister, Hisham Qasrawi.

This Recommendation was developed by RILEM TC 249-ISC consisting of Denys Breyse, Jean-Paul Balayssac, Samuele Biondi, David Corbett, Arlindo Goncalves, Michael Grantham, Vincenza Luprano, Angelo Masi, Andre Valente Monteiro, Zoubir Mehdi Sbartai. It results from the active contribution of all contributors listed above. The draft recommendation was submitted for approval to the full TC and subsequently approved by RILEM TC 249-ISC.

TC Chairman: Pr. Denys Breyse (University Bordeaux, I2M-UMR CNRS 5295, France).

TC Secretary: Pr. Jean-Paul Balayssac (LMDC, INSA/UPS, Toulouse, France).

TC Members: Samuele Biondi (Italy), David Corbett (CH), Arlindo Goncalves (Portugal), Michael Grantham (UK), Vincenza Luprano (Italy), Angelo Masi (Italy), Andre Valente Monteiro (Portugal), Zoubir Mehdi Sbartai (France).

D. Breyse (✉) · Z. M. Sbartai
University Bordeaux, I2M-UMR CNRS 5295, Talence,
France
e-mail: denis.breyse@u-bordeaux.fr

J.-P. Balayssac
INSA Toulouse, LMDC, Toulouse, France

S. Biondi
University Chieti-Pescara, Chieti, Italy

D. Corbett
Proceq, Schwerzenbach, Switzerland

A. Goncalves · A. V. Monteiro
LNEC, Lisbon, Portugal

M. Grantham
Leeds University, Leeds, UK

M. Grantham
Sandberg LLP, London, UK



illustrative example is described for each step of the assessment process.

Keywords Concrete strength · Model calibration · Non-destructive techniques · In-situ investigation · Rebound hammer · Strength assessment · Ultrasonic pulse velocity

List of symbols

a, b	Conversion model parameters
C	Calibration factor
C_M	Calibration factor (scaling calibration)
COV	Coefficient of variation (= standard deviation/average value)
COV_{rep}	Coefficient of variation of test results (for test result precision)
C_S	Calibration factor (shift calibration)
EQL	Estimation quality level
est	Index for estimated values
f_c	Concrete compressive strength
$f_{c,est}$	Estimated concrete compressive strength
$f_{c,est,cal}$	Estimated and calibrated concrete compressive strength
$f_{c,est,o}$	Prior estimated concrete compressive strength
i	Index for measured values (from 1 to N_{TL} for NDT test results, from 1 to N_c for core strength) or assessed values
j	Index for conversion model parameters (from 1 to N_{par})
k	Index for NDT techniques
M	Conversion model ($f_{c,est} = M(\text{Tr})$)
mean	Index designating the mean value of measured (test results, core strengths) or estimated properties (strength)
N_c	Number of cores
NDT	Non-destructive testing
N_{par}	Number of parameters of the conversion model
N_{read}	Number of readings (repetitions of individual measurements) in order to derive a test result

N_{rep}	Number of repetitions of a test in order to derive the test repeatability
N_{TL}	Number of test locations (= number of NDT results)
par_j	Value of the j -th conversion model parameter
r^2	Determination coefficient
RH	Rebound hammer (for technique or test result)
RMSE	Root mean square error
$RMSE_{fit}$	Fitting root mean square error, estimated on the identification set
$RMSE_{pred}$	Prediction root mean square error, estimated when the conversion model is applied to new data
sd	Standard deviation
sd_{rep}	Standard deviation of test results (for test result precision)
TL	Test location
TR	Test region
Tr	Test result
Tr_i	Value of the i -th NDT test result (when a single NDT is used)
$Tr_{k,i}$	Value of the i -th test result of the k -th NDT method (when several NDT are combined)
TRP	Test result precision
UPV	Ultrasonic pulse velocity (for technique or test result)

1 Definitions

Accuracy	Accuracy corresponds to the closeness of agreement between one measured or assessed value and the reference value of the evaluated property, e.g. strength measured on cores.
Conditional coring	Conditional coring corresponds to a way of defining the location of the N cores on the basis of prior NDT test results obtained at a larger number of test locations.
Conversion model	Mathematical function used to convert a NDT test result into a strength value at a test location.

V. A. M. Luprano
ENEA, Brindisi, Italy

A. Masi
University Basilicata, Potenza, Italy



Estimation quality level	Estimation quality level (EQL) is a target accuracy level for the concrete strength estimation. Three different EQLs are defined, corresponding to more or less severe requirements.	Test reading	A test reading is a value of one single measure (i.e. one value of rebound number, UPV, pull-out force or any other NDT method).
Investigation domain	The investigation domain (ID) corresponds to the structure, part of the structure or set of components whose strength is to be assessed. It may contain various test regions.	Test region	The test region (TR) is a given volume of concrete which is known, or assumed, to belong to a same strength population. A test region contains several test locations. It may be a part of a component, a whole component, a set of components or a larger zone, like a whole floor in a building or a whole structure. A test region can also denote one or several structural elements (or precast concrete components) assumed or known to be from the same strength population or equivalent to the defined volume associated with identity testing for compressive strength.
Precision	Precision corresponds to the closeness of agreement between measured values obtained by replication. It is related to repeatability and reproducibility within a single test location TL (including a displacement in the immediate vicinity, for instance, inside a rebar mesh).		
Predefined coring	Predefined coring corresponds to a way of defining the location of cores within a test region, independently of the NDT test results.	Test result	A test result (Tr) is a value representative of a test location. It can be a single test reading, or a mean or a median of N values of test readings.
Risk level	The risk level quantifies the probability that the exact value of a quantity does not belong to the target tolerance interval centered on the estimated value of this quantity.		
Target tolerance interval	A target tolerance interval corresponds to the range of values that is supposed to contain the true value of the target. The target tolerance interval is indexed to a quantity to be assessed, to an EQL and to a risk level.		
Test location	A test location (TL) is a limited area selected for measures used to obtain one test result that is used in the estimation of in situ compressive strength. Several test readings can be carried out in the same test location.		

2 Scope

This document aims to give recommendation to assess the in situ compressive strength of concrete in existing structures by combining core strength values and non-destructive measurements. Both mean strength and its scatter are considered.

This recommendation primarily focuses on the processes for estimating compressive strength (local value, mean strength, strength standard deviation) from measurements performed on site. Deriving a characteristic strength from the assessment results is not considered in this recommendation.

The focus is on existing reinforced concrete structures where both aging effects and reinforcing steel bars may have a great influence on NDT test results and on subsequent strength estimations [1]. Aging effects include possible mechanical or physical deterioration (e.g. cracking, concrete delamination

due to reinforcement corrosion). Another specificity of existing structures, in practice, is linked to a lack of detailed information about the concrete used, and moreover there are no companion specimens which could be used for comparison.

This recommendation applies for most common techniques (ultrasonic pulse velocity, rebound hammer, pull-out) but also for less common techniques (penetration test, etc.). The recommendation does not apply to situations in which no core has been taken from the existing structure and is limited to situations where NDT is combined with cores.

An example is provided in this recommendation to simply illustrate how this approach can be used at each step

3 Previous research and state-of-the-art of practice

A significant amount of research has been carried out, both in the laboratory and on real structures, to analyze the existing relations between concrete core strength and NDT test results. Despite decades of practice and research, the interest of using NDT methods for an efficient and reliable assessment of concrete strength remains a controversial issue. There are even some cases in which the analysis and processing of NDT test results on existing structures has not improved the information gained by cores.

NDT methods are commonly used in every day practice. However, all problems are not always correctly addressed, either because lack of knowledge, training or experience regarding this complex issue. Examples of common bad practices that can be found both in laboratory studies and in situ are [2]:

- The derivation of a strength value from an NDT test result by simply using a “universal” conversion model, assumed to be true;
- The estimation of the quality of a conversion model through the value of a correlation factor between measured and estimated strengths, on the same dataset that has been used for building or calibrating this model;
- The combination of several non-destructive techniques to improve the quality of estimation, without balancing the apparent improvement by

an analysis of the assessment uncertainty, which is commonly larger when techniques are combined;

- The lack of consideration for the effects of external factors (e.g. carbonation and moisture content of concrete).

The main consequence is that NDT methods are often perceived as unreliable when it comes the on-site strength assessment challenge. RILEM has promoted NDT methods for many years and two RILEM recommendations have been issued, respectively in 1976 [3] and in 1993 [4]. The former compared several recommendations existing in Romania and other eastern European countries and promoted the use of surface hardness methods (i.e. rebound hammer measurements, RH). The latter used a combination of rebound hammer and ultrasonic pulse velocity (UPV) measurements to build a conversion model between these test results and concrete strength. The resulting approach was named SonReb and is still used in many cases, mostly in a research context.

The RILEM TC7 recommendation [3] intended to establish a correlation (conversion model) to derive strength from rebound hammer test results. The practicality of the correlation was discussed and the uncertainty on strength was estimated. In engineering practice, the use of correcting factors makes the result so imprecise that this approach was quickly abandoned.

The basic idea of RILEM TC43 recommendation [4] was similar to that of [3] and the need of calibration induced the same drawbacks which prevented to reach an acceptable precision. However, the idea of combining UPV and rebound hammer survived and was applied by many researchers and engineers, with a variety of mathematical expressions as tentative conversion models. However, the quality of the estimated strength remains unknown. The interest of combining UPV and RH test results which are not equally sensitive to all physical and material properties is still a matter of discussion. Some studies indicate that it improves the assessment whereas other studies reach the opposite conclusion.

Standards have been published in Europe [5] and in the US [6–9] which show how NDT measurements must be carried out for the most common NDT methods but they do not explain how strength is derived from the test results. The relationship between NDT test results and strength is discussed in [10]

which provides meta-analyses on test result repeatability.

The European Standard [11] is devoted to strength evaluation in concrete buildings and components, with a main focus on characteristic strength assessment and on concrete strength class determination. It allows assessment either with cores only or through a combination of NDT test results and cores. The revised version of this standard (expected to be published in 2020) will give more possibilities for using NDT test results. A correlation between strength and NDT test results can be established from 9 pairs of strength and NDT test results (the number of cores can be slightly larger in order to consider for possible outliers). However, it will not provide any formal guidance on how to establish the correlation, and how to estimate the quality of the strength estimation.

Research has been very active in this field and several interesting concepts have been recently promoted [12]:

- The fact that it is impossible to find any universal conversion model is now well accepted and any efficient approach must consider the construction of a specific model, or the calibration (adaptation) of a previous model. The challenge is therefore to optimize the trade-off between the residual model error and the number of cores required by the calibration process;
- Non-destructive techniques are adapted to a quick and extensive coverage of the investigation domain, thus paving the way for a screening that would indicate in which areas the strength is expected to reach extreme values. This makes it possible to identify homogeneous regions with possibly different concrete properties [13] and to identify locations that would be preferable for taking cores [14];
- The importance of NDT measurement repeatability has been highlighted as a major factor limiting any further step of the assessment process [2], which justifies that a more careful attention has to be paid to this factor;
- An innovative method has been introduced during the test results processing (designated as “bi-objective approach”), which theoretically enables capture of both mean strength and strength variability without increasing the number of cores [14];
- Thanks to the development of synthetic simulations [15], it became possible to compare different investigation and assessment approaches, in order to derive practical conclusions regarding their respective efficiency [16].

The purpose of this recommendation is to provide the methodology and means to assess in situ concrete strength in the context of existing structures, while benefiting from the most recent research advances.

4 Key steps of the concrete strength assessment methodology

4.1 Reliability of concrete assessment: the role of uncertainties

In this recommendation, attention is paid not only to strength assessment, but also to the estimation of the quality of this assessment. The common concrete strength assessment process can be subdivided into three main stages (Fig. 1):

- (a) Data collection, which includes both NDT measurements and core strength measurements. It should be noted that the option of estimating strength directly from NDT test results without any core is excluded from this recommendation;
- (b) Conversion model identification, which will derive strength estimates from NDT test results alone;
- (c) Model use and concrete strength estimation.

Several tasks that must be carried out before the on-site investigation (including getting preliminary information about the structure) are not detailed in this recommendation. The reader is invited to read TC-ISC Guidelines (Sect. 2) for more details.

Limited information, uncertainty and errors at each stage influence the assessment process. The recommended process aims at estimating concrete strength with an accuracy consistent with predefined objectives. Research has shown that the uncertainty on strength estimates is the final result of a complex process in which uncertainties play a major role at several steps (Fig. 2) [12].

The uncertainty attached to the derived conversion model (i.e. the uncertainty on the values of the model parameters) results from:

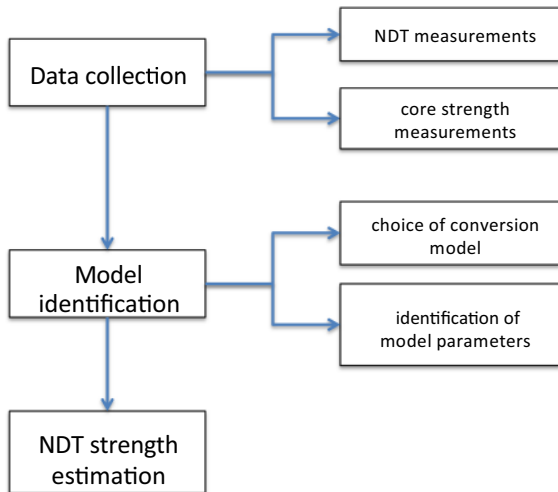


Fig. 1 Three main steps of the usual concrete strength assessment process

- (a) The sampling uncertainties, as the model is derived from a limited dataset. Let us define N_{core} as the number of cores, which is also the number of (f_{ci}, Tr_i) pairs where f_{ci} is the i -th strength value measured on a core and Tr_i is the i -th nondestructive test result. One must also consider that, depending on how core locations are selected, the core strength values can provide a more or less representative picture of the whole population.
- (b) The measurement uncertainties, since the f_{ci} and Tr_i values are obtained from experimental tests and do not exactly correspond to the “real” value of the same quantity at the test location. Repeating the same ND measurement at the same test location, which is often possible at low

expense, would provide a more precise ND test result. It must be noted that the test location is not defined as a single point but as a limited area, which enables the consideration of repeatability for rebound test results.

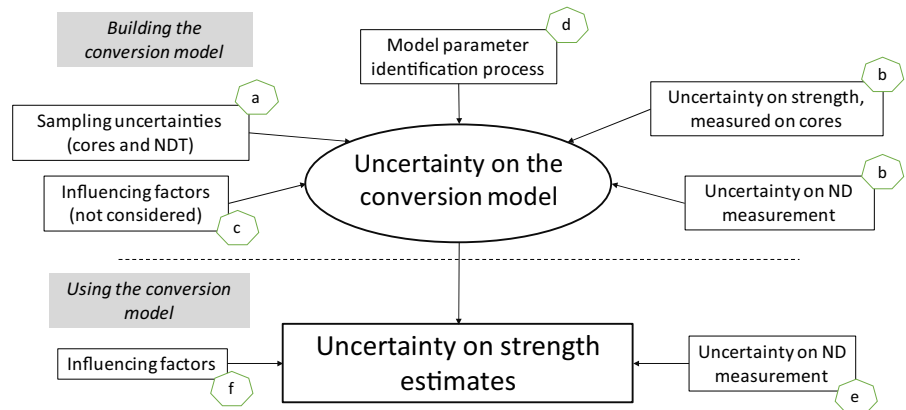
- (c) Any other influencing factor that can affect the test result but is not considered explicitly in the conversion model. Many such influencing factors are known for concrete, like the moisture content, the carbonation depth, the type of aggregates, etc.
- (d) The estimation process of the model parameters also has a small influence: different identification approaches can be used, like fitting a specific model by the method of least square errors, or calibrating a prior curve with a drift or a scaling factor. Each option would lead to slightly different uncertainties.

Once the conversion model has been identified, it is used in a second stage in order to estimate strength values from new nondestructive test results (Tr). This stage induces

- (e) Measurement errors on new ND test results,
- (f) Influencing factors that can further increase the evaluation uncertainty, for instance if the measurement conditions differ from those prevailing when the first series of measurements was carried out.

The influence of all these sources of uncertainty must be considered in the process. The challenge is therefore revised: instead of estimating the real strength value, it becomes that of guaranteeing that

Fig. 2 Influence of uncertainties at all steps of the strength assessment process on the uncertainty of strength estimates



the real strength value lays inside a target tolerance interval indexed to a predefined EQL and risk level.

4.2 Recommended organization of the investigation and assessment

The assessment of concrete strength provides $f_{c,est}$ values. It is based on: (a) available NDT test results at test locations, (b) the use of a conversion model to convert them into estimated strength values at the local (component) scale.

Available information at the beginning of the assessment process is provided, as shown on Fig. 3, by a number N_{TL} of NDT test results and a number N_c of core strengths. The interest in NDT based assessment is based on the fact that $N_{TL} \gg N_c$. Strength values, $f_{c,est}$, can thus be derived at many test locations where strength has not been assessed by cores.

Any investigation strategy must consider the following items:

- Choice of test locations for NDT measurements,
- Choice of points (rules for choice, number, location) where cores are to be taken. These two first decision steps interact because of constraints (cost, time, geometry, accessibility...) on the whole process,
- Identification of the conversion model,
- Use of the conversion model to estimate strength from NDT measurements,
- Evaluation of the quality of estimation.

A preliminary stage is that of defining the requirements in terms of accuracy of the estimated concrete properties.

The flowchart in Fig. 4 describes how the process must be organized and emphasizes (with a bold frame) important tasks (T1 to T9) that are necessary in order

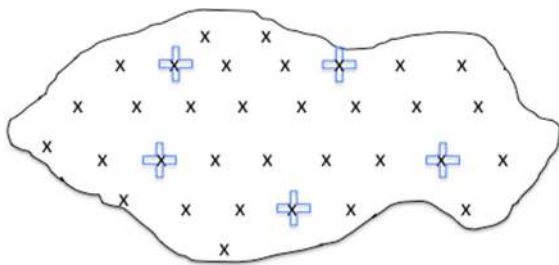


Fig. 3 Illustration of available test results (black “x” = NDT test results, blue crosses = core strength test results, here $N_c = 5$). (Color figure online)

to obtain a reliable estimation of concrete properties. All of these tasks are described in the following and illustrated with an example.

5 Definition of the estimation quality level

Task 1: Define the estimation quality level

Three quantities which can be estimated following this recommendation are respectively the mean value of local strength (or concrete mean strength), the standard deviation of local strength (or concrete strength standard deviation) and the local error on local strength. The estimation quality level (EQL) concept corresponds to a target quality regarding assessment, or residual uncertainty, on each of these estimated quantities.”

Three EQLs are defined which correspond to:

- A target uncertainty that is supposed to be obtained on three quantities defined: mean strength, standard deviation of strength and local strength. Mean strength and strength standard deviation are estimated over the test region.
- The process (including data collection and methods for model identification at the first two steps of Fig. 1) used to reach this target.

The difference between the three EQLs (from EQL1 to EQL3) is a decreasing tolerance interval on each estimated quantity. For each EQL, Table 1 indicates the target tolerance interval that is expected to be obtained regarding the evaluation of compressive strength. For each estimated quantity, these numbers define the uncertainty range to which it is supposed to belong, with a confidence level $(1 - \alpha)$. The α value quantifies the risk level, i.e. the probability that the true value will be outside the tolerance interval. The uncertainty range of the strength standard deviation can be expressed either in relative terms (as a percentage of its true value) or in absolute terms (in MPa). The error on local strength (root mean square error, or simply RMSE) is always positive and can be expressed either in relative terms (as a percentage of the mean concrete strength) or in absolute terms.

The values given in the Table indicate what accuracy will be obtained on the strength estimates by following the Recommendation. The only thing to do at this step is to choose the targeted EQL. Since EQL3 (resp. EQL2) is more ambitious than EQL2

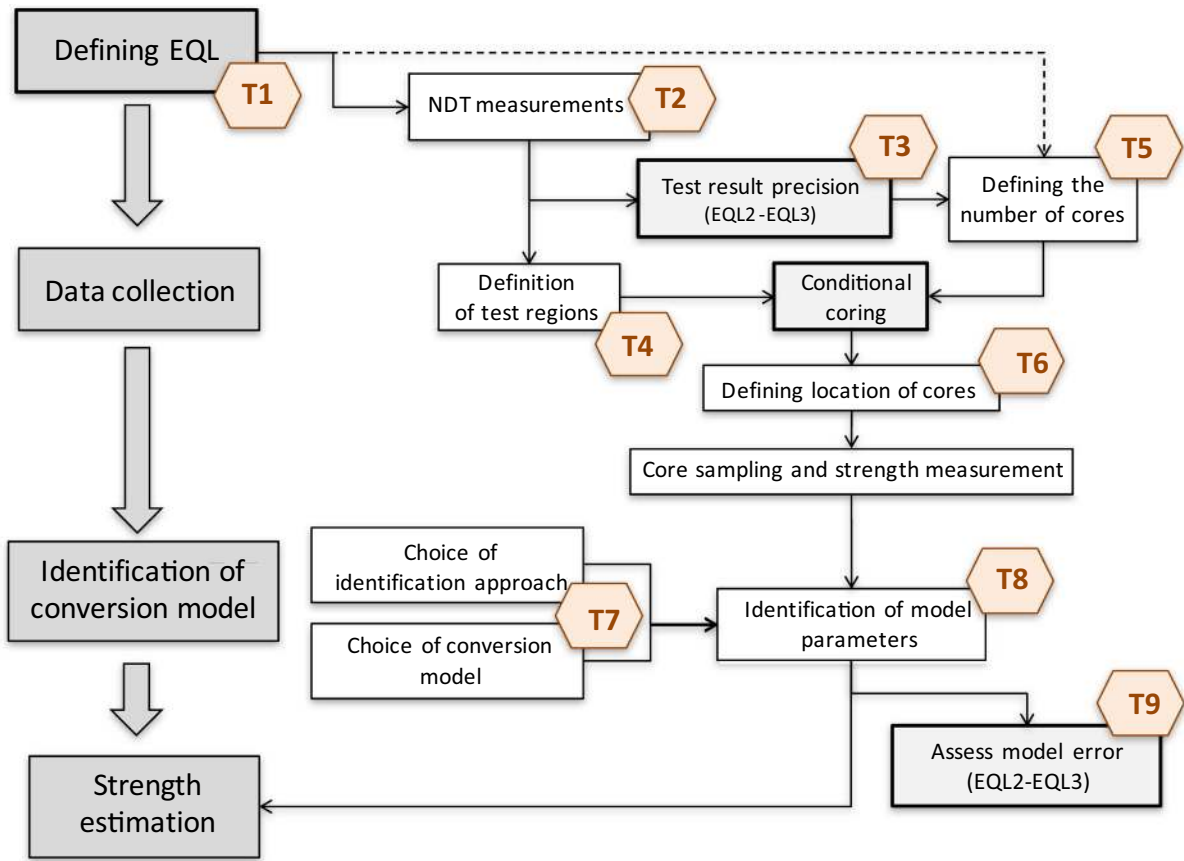


Fig. 4 Detailed steps of the recommended concrete assessment process

Table 1 Indicative relation between estimation quality levels (EQL) and target tolerance intervals quantifying the accuracy on strength estimates

Estimated quantity	EQL1	EQL2	EQL3
Mean value of local strengths	± 15%	± 15%	± 10%
Standard deviation of local strengths			
Relative	Not addressed	± 50%	± 30%
Absolute		± 4 MPa	± 2 MPa
Error on local strength			
Relative	Not addressed	20%	15%
Absolute		6 MPa	4.5 MPa

(resp. EQL1) it thus requires a higher amount of resources: a larger number of cores and NDT test results, data of a better quality, more reliable models, more robust data analysis, etc.

Illustrative example—TASK 1: Define the EQL

Let us consider a simple 4-story reinforced concrete frame. At each level, there are 20 columns (5 × 4) and 16 longitudinal beams (4 × 4), which yields a total of 144 structural

components. All structural components have the same dimensions and three possible test locations are identified on each structural component. Nothing is known about the building history, except that it was built in the 1970s with ordinary concrete

The choice of the target estimation quality level is EQL2, as defined in Table 1, with an absolute tolerance interval on standard deviation and local error. This means that the expected result is ± 15% on mean strength, ± 4 MPa on the concrete standard deviation and ± 6 MPa on any local strength estimation



6 Data collection

6.1 Before going on-site

Before beginning an on-site investigation, much information should be gathered and analyzed, which can impact the non-destructive concrete strength assessment process. EN13791 and other standards provide details on what must be done before going on-site. These recommendations focus on the sampling, testing and analysis of results steps.

The following questions must be answered:

- What is the objective of the investigation: assessment of load capacity, detection of poor concrete areas, refurbishment of the structure or simply the necessity to check the structural condition as a routine check-up, etc.?
- What questions have been asked by the engineers or by the structure managers?
- What are the specifications for the investigation?
 - What is the scale of the assessment (component, whole structure, part of the structure)?
 - What assessment uncertainty level is required, i.e. what is the corresponding EQL?

All relevant information about the structure under investigation should be collected.

The investigation strategy is constrained by the resources available, whether they are time constraints (number of hours/days devoted to the investigation), budget constraints or technical constraints (limitations for taking cores). Thus, the density of the investigation is a direct result of these constraints. It is expected that increasing the number of tests (and data) would naturally improve the quality of the assessment, but there is no simple way to quantify the quality of the assessment as a function of the investigation density.

6.2 Analyzing the investigation domain

The Investigation Domain (ID) corresponds to the structure, part of the structure or set of components whose strength has to be assessed. The delimitation of the ID can be done by a visual inspection of the structure and/or by a priori considerations provided by former investigations or a good knowledge of the structure. Structural considerations cannot be neglected during the definition of the ID.

6.3 Beginning with NDT measurements

Task 2: Carry out the NDT measurements

Both the number and location of measurements, and the type of NDT methods determine the sampling plan of NDT investigation.

Many methods can be proposed for estimating the compressive strength of concrete. The most common test methods are: ultrasonic pulse velocity (UPV), rebound hammer (RH), pull-out test (Capo-test), micro-core testing (cores with a diameter smaller than 50 mm), penetration resistance (pin test or Windsor probe). All NDT methods provide test results from which a strength estimate can be derived only through a conversion model.

A test location (TL) is a limited area selected for measures from which a single test result is obtained and used for the estimation of in situ compressive strength.

A test result (Tr) is a value representative of a test location. It can be a single test reading, a mean or a median of several values of test readings, where a test reading is a value of one single measure (one value of rebound number, UPV, pull-out force or any NDT method).

Several test readings can be carried out in the same test location, N_{read} being the minimum number of replicates to be sure that the test result is inside a predefined confidence interval defined from the measure uncertainty. Recommended values of the number of test readings required for getting one representative test result can be derived from specific standards and recommendations for each NDT method. For instance, for rebound hammer a large number of readings (about ten) are required in order to derive a representative test result at a given test location.

The number of NDT measurements depends first on the amount of resources (time, money, material) which can be devoted to the investigation. Its choice must account for the fact that each ND test result will finally lead to an estimation of the concrete strength at the same location, but that some resource must be kept aside for taking cores and getting some additional reference strength values (see Task 5). The core location must be chosen in order to cover all areas of interest in the investigated domain:

- To try to cover the full range of strengths by taking measurements in different areas (i.e. “weak areas” as well as “good areas” as they are expected or appear from a visual inspection). This is true at all scales, that of the structure (e.g. between floors) as well as that of a component (e.g. different heights in a column);
- If some parts of the structure or components have a stronger influence on the structural safety, a more reliable result may be required there, and that may be provided by densifying NDT measurements in these areas, without however compromising the structural safety;
- If several NDT methods are used and will be combined for strength estimation, the same test locations must be used for the measurements with the different methods.

Several additional items have to be considered in order to avoid NDT test results being influenced by factors other than concrete strength, such as:

- Prior to any investigation, the position of the reinforcement must be determined. It is strongly recommended to carry out all measurements in such a way as to avoid any adverse effects due to the proximity of rebar (e.g. it is known that the presence of rebar increases the wave velocity).
- If cracks are present in the zone planned to be investigated, care must be taken to perform the tests far enough away from the cracks.
- The effect of carbonation on the rebound hammer test results, and possibly other NDT measurements, can be very significant since the hardness of a carbonated concrete is larger than its original value. Since the inner strength of the concrete is not modified by carbonation, the NDT test result may lead to wrong conclusions.

Table 2 synthesizes the three different scales that must be considered within the investigation domain and to what types of data they correspond.

Illustrative example—TASK 2: Carry out the NDT measurements

The building has 144 structural elements, with three possible test locations on each component, which yields 432 possible test locations

For cost reasons, rebound hammer measurements are chosen for screening the building properties. Sixty (60) test locations are

chosen for NDT measurements, with 6 beams and 9 columns for each floor which respectively corresponds to 37% and 45% of the components. The larger number of test locations on columns is justified by their higher influence on the structural capacity of the building. These test locations are equally shared between the four levels

The cumulative distribution function of RH test results is given in Figure Ex-1 where the respective distributions for beams and columns are plotted. Each curve is built by simply ranking all test results of a dataset (here for two datasets, namely beams and columns) from the lowest value to the highest value. The values corresponding to the different floors are identified by different markers. The RH test results range between 27 and 43 with median values of 37 and 34 for beams and columns, respectively

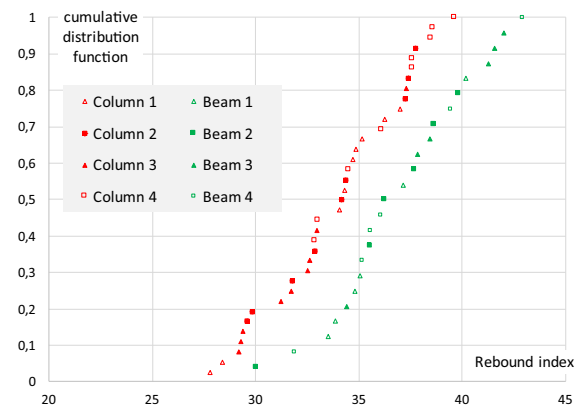


Figure Ex-1. Cumulative distribution function of rebound hammer test results

6.4 Assessing test result precision

Task 3: Assess the NDT test result precision

Test results are the only data that are used in the following, in order to estimate strength properties and the quantities corresponding to each EQL. The precision of these input data has a major effect on the uncertainty of the strength estimate.

It is recommended, during the investigation stage, to devote a specific part of the program to evaluate the magnitude of the test result precision (TRP), which is a key factor in the next stages of compressive strength assessment. This is mandatory if the target accuracy of strength estimate is EQL2 or EQL3 (Table 3).

TRP is evaluated through several repetitions of the measurement process at some test locations over the investigation domain in similar conditions. A TRP value is derived which is representative of the repeatability of the test result over the investigation

Table 2 Definition of scales and corresponding experimental data

Area	Measurement point	Test location TL	Test region TR
Type of data	Test reading	Test result	Distribution of test results

domain. It will become a key information at a later stage of the process (see Sect. 6.7—Task 5).

The repeatability of the test result can be the same over the investigation domain, or it can vary in some test regions. Finally, a cautious estimate of the test result precision must be derived. Table 4 provides indicative values of TRP for the most common NDT methods.

Values in Table 4 are derived from the level of repeatability that can be obtained in practice for each NDT method and from how sensitive variations in NDT test results are to variations in concrete strength. It must be pointed out that this variability is not that of the concrete strength which can be much larger.

The comparison between sd_{rep} or COV_{rep} obtained during the tests with values of Table 4 leads to

qualifying the NDT precision as being high, medium or poor.

Note that some precision levels may be difficult to obtain in practice due to the intrinsic nature of the test and material response. For instance, a high precision level is difficult to obtain for rebound hammer test results or micro-cores. Poor test result precision may prevent the identification of a reliable conversion model and, as a consequence, can drastically limit the precision of strength estimate.

If it is planned to combine several NDT methods, the precision must be quantified for each technique, since any method with a poorer precision constitutes a handicap for an efficient combination.

Illustrative example—TASK 3: Assess the NDT test result precision (TRP)

One simple way to assess TRP is:

- to select a few (2 to 5) test locations,
- at each test location:

to repeat N_{rep} times the measurement process (note that the whole measurement process to obtain a test result is repeated, which requires a larger number of test readings). The minimum value for N_{rep} is 3, but 4 or 5 repetitions are preferable, as they provide a more representative assessment;

to quantify the standard deviation of test results sd_{rep} or COV_{rep} between the N_{rep} test results at this test location,

- from these few values of sd_{rep} or COV_{rep} , to derive one single value which is representative of the repeatability of the test result over the investigation domain.

In practice, four specific test locations are selected to assess the TRP value. Table Ex-1 provides their location and the results of 5 repetitions of NDT measurements.

Table Ex-1. Rebound Hammer test results for assessing TRP

Element	floor	TL number	1	2	3	4	5	mean	st. dev.	COV (%)
Beam	1	242	34.9	33.6	36.0	34.3	38.5	35.5	1.9	5.4
Beam	3	317	41.6	39.0	41.7	41.0	38.3	40.3	1.6	4.0
Column	2	78	32.4	35.8	33.4	33.2	33.0	33.5	1.3	3.9
Column	4	184	31.7	31.6	31.2	31.9	30.4	31.4	0.6	1.9

The average standard deviation (sd_{rep}) on the four test locations is equal to 1.4 rebound number units, for a mean rebound number equal to 35.2, which also corresponds to a mean coefficient of variation (COV_{rep}) equal to 4.0%. The two last columns provide the standard deviation and coefficient of variation. Referring to Table 4, one can conclude that these NDT test results correspond to a medium precision TRP.

Table 3 Recommendation regarding the Test Result Precision evaluation for the three EQLs

EQL1	EQL2	EQL3
Recommended	Mandatory (TRP value must be given)	Mandatory (TRP value must be given, and must correspond to medium or high precision)

Table 4 Indicative order of magnitude of test result precision (TRP)

	High precision	Medium precision	Poor precision
Usual NDT			
RH			
Coefficient of variation COV_{rep}	$COV_{rep} \leq 3\%$	$3\% < COV_{rep} \leq 7\%$	$COV_{rep} > 7\%$
Standard deviation sd_{rep}	$sd_{rep} \leq 1$	$1 < sd_{rep} \leq 3$	$sd_{rep} > 3$
UPV			
Coefficient of variation COV_{rep}	$COV_{rep} \leq 1\%$	$1\% < COV_{rep} \leq 3\%$	$COV_{rep} > 3\%$
Standard deviation sd_{rep}	$sd_{rep} \leq 50$ m/s	50 m/s $< sd_{rep} \leq 125$ m/s	$sd_{rep} > 125$ m/s
Other non-destructive methods			
Coefficient of variation of pull-out force	$COV_{rep} \leq 5\%$	$5\% < COV_{rep} \leq 8\%$	$COV_{rep} > 8\%$
Coefficient of variation of Windsor probe penetration test	$COV_{rep} \leq 2\%$	$2\% < COV_{rep} \leq 5\%$	$COV_{rep} > 5\%$
Coefficient of variation of strength on micro-cores	$COV_{rep} \leq 15\%$	$15\% < COV_{rep} \leq 25\%$	$COV_{rep} > 25\%$

References for COV_{rep} values in Table 4

Regarding rebound and UPV, synthetic information is available in [2] and in

Lee M.H., Choi C.W., A study on the minimum number of rebound number test and pulse velocity method for estimating compressive strength of concrete, J. of Korea Concrete Institute, 16, 6, 833–840, 2004

It can also be found for UPV in

Machado M., Curvas de correlacao para caracterizar concretos usados no Rio de Janeiro por meio de ensaios nao destrutivos, PhD report, Univ. Rio de Janeiro, 2005

Rojas-Henao L.M., Ensayos de informacion y extraccion de probetas testigo en hormigones autocompactantes, PhD report, Univ. Politecnica Madrid, 2012

and for rebound in

Szilagyi K., Rebound surface hardness and related properties of concrete, PhD report, Budapest Univ. of Technology and Economics, 2013

Regarding pull-out tests, values are derived from comparison of results obtained by various researchers [8]. Regarding penetration depth, and test results obtained with 6 readings [8]. Regarding micro-cores, the variability is much larger than that for cores and may be highly influenced by the size of the largest aggregates. See

Celik A.O., Kilinc K., Tuncan M., Tuncan A., Distributions of compressive strength obtained from various diameter cores, ACI Mat. J., pp. 597–606, 2012

6.5 Dataset consolidation: outliers

When analyzing real datasets, it is often found that some observations are different from the majority of the data. Such observations are usually called outliers and can be defined as individual data values that are numerically distant from the rest of the sample, thus masking its probability distribution.

Outliers always deserve a careful attention, either because they may have a significant impact in concrete strength estimation or, in some cases, they may in fact represent a different concrete population that deserves a separate assessment.

An outlier analysis can usually be seen as a two-step process. In the first step, several methods can be used to identify potential candidate outliers of the data. In



the second step, the candidate outliers need to be handled to account for their effect in subsequent statistical analyses of the data.

Full guidelines [20] provide more detailed information about the techniques that can be used to deal with outliers.

Illustrative example: Dataset consolidation

From the analysis of the two distribution curves plotted on Figure Ex-1, there are no apparent outliers, i.e. a value of a test result that would appear as departing from the rest of the data. Therefore, no specific test is carried out and all data are kept.

6.6 Defining the test regions

Task 4: Define the Test Regions

If the investigation domain covers a whole structure or a large part of a structure, a specific procedure can be applied in order to discriminate it into test regions (TR). This procedure is aimed at reducing the variability inside each TR, test regions being identified such that the concrete is assumed to be homogeneous in each of them.

The test region (TR) may be chosen as a part of a component, as a whole component, as a set of components or as a larger zone, like a whole floor in a building or a whole structure. A test region can also denote one or several structural elements (or precast concrete components) assumed or known to belong to the same strength population.

One pair of values (mean and standard deviation) is considered as representative of the concrete strength population in a single test region. This standard deviation (also called measure uncertainty) includes the precision of the measure and the variability of the material.

The identification of test regions can be helpful to properly identify the conversion models (see Sect. 7). It is also relevant in the case of structural assessment by identifying zones with different concrete strengths.

Illustrative example—TASK 4: Defining the test region

The issue is that of either dividing the dataset into several subsets, i.e. into test regions which have to be analyzed separately, or to keep them all together.

One option could be to consider beams and columns separately. Another option could be to consider a specific population either for columns of the third and fourth floors that respectively have lower and higher values than the rest of the population.

However, the differences remain small, and one must be aware that considering a new population may involve the identification of a specific conversion model, which would require more cores. For these reasons, in this example, all data are considered as belonging to the same population and to represent a unique distribution of concrete properties (i.e. one test region).

6.7 Defining number and location of cores

Task 5: Define the number of cores

The strengths measured on cores cannot be considered as the true values of concrete strength but they are seen in this recommendation as the reference values, and which are used for identifying the conversion model (Task 8) and for checking the model error (Task 9).

The number of cores is among the main influencing factors regarding the strength estimation quality. This is illustrated through risk curves that can be calculated through simulations [15]. Thus, for a given accepted risk, the recommended number of cores used to estimate the compressive strength in conformity with a prescribed EQL must be in agreement with the magnitude of the target tolerance intervals as defined in Table 1 (Fig. 5).

There is no way to formally derive a recommended number of cores covering all possible situations from a theoretical model or from real engineering practice. The numbers provided in this section are derived from calculations carried out on synthetic data [12, 15]. Such data allow for a much more extensive analysis of possible scenarios. They are also the only way to address the assessment issue in terms of risk. Details regarding these calculations are given in full Guidelines [20].

The recommended numbers of cores are supposed to apply in most real situations for in situ investigation of concrete. They must be taken as a general guidance, in line with the three EQLs and the corresponding tolerance intervals defined in Table 1.

For each EQL these numbers depend on four inputs:

- The non-destructive test result precision (TRP) as previously estimated (“high precision”, “medium precision” or “poor precision”. If TRP has not been assessed, one must consider the numbers for “poor precision”);
- How the target tolerance interval is expressed in Table 1 (i.e. absolute or relative);
- The mean compressive strength;
- The coefficient of variation (COV) of compressive strength.

Since these last two parameters are usually unknown, the most unfavorable combination (regarding possible mean strength and COV of compressive strength) should be considered, for the relevant TRP. If the investigator has some reliable information, including that provided by NDT screening, which allows him to reduce the possible range of variation of mean strength and strength standard deviation, he can use this information to define the recommended number of cores. Such information can be based on current practice at the time of construction [17] or on the available NDT test results data set [18, 19]. This information must be documented.

This recommendation does not provide values for the recommended number of cores. The reader is invited to refer to exhaustive information that will be published in the Guidelines [20] for covering all possible situations.

Illustrative example—TASK 5: Defining the number of cores

The target EQL is EQL2 with an absolute target tolerance on standard deviation and local error (see Table 1), and the TRP is medium, with rebound hammer measurements.

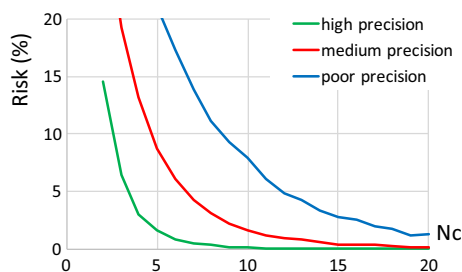


Fig. 5 Risk curves for the estimation of mean strength with a $\pm 10\%$ tolerance interval, as a function of number of cores N_c and TRP level (case of mean strength = 30 MPa and COV = 20%)



Guidelines [20] provide values for the recommended number of cores. For the specific case of EQL2 and medium precision TRP, these numbers are given in Table Ex-2 below.

Table Ex-2. Recommended number of cores for TRP2 and EQL2, with rebound hammer measurements ($f_{c, \text{mean}}$ in MPa, COV in %)

Medium TRP					
fc mean	COV				
	10	15	20	25	30
10	4	5	6	6	7
15	3	4	5	5	5
20	3	3	5	5	5
25	3	3	5	5	5
30	3	4	5	6	6
35	4	5	6	7	8
40	5	6	7	9	10
45	6	7	9	10	11
50	6	8	10	12	13

If we exclude the cases where the mean strength is above 40 MPa and the coefficient of variation of strength is above 25%, this table indicates that 9 cores cover all possible configurations. Therefore, it is recommended to take 9 cores for the conversion model identification.

Task 6: Define the location of cores and take cores

Cores must be chosen in order to provide values which make it possible to build a representative picture of the investigated domain, without compromising the structural safety.

Once the number of cores is decided, the core locations must be defined either independently of NDT measurements (predefined coring) or based on prior information provided by NDT measurements (conditional coring). Figure 6 illustrates these two options. If predefined coring is chosen (left option), the location of cores and that of NDT measurements are chosen independently. With conditional coring (right option), the location of cores is defined after an analysis of NDT test results, aiming to ensure a more correct coverage of the whole range of concrete strength.

Conditional coring does not induce any additional cost, and only requires that NDT test results have been made available prior to taking cores (it may however be only a part of the full NDT program). Conditional coring is recommended, since it is expected to lead to a

Fig. 6 Predefined (“no” option) or conditional (“yes” option”) coring

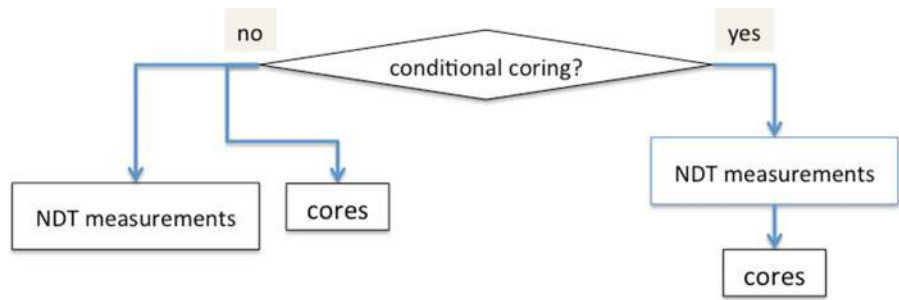


Table 5 Recommendation regarding the Conditional Coring, for the three EQLs

EQL1	EQL2	EQL3
Recommended	Recommended	Mandatory

more reliable conversion model and is mandatory for EQL3 (Table 5).

Conditional coring is even more preferable when the number of cores is low (3 to 8 cores), since it ensures a better coverage of the concrete condition/strength range. With such a small number of cores, the risk with predefined coring is that most cores (or even all) are selected at locations where the strength is higher (or lower) than the average strength. However, the correct definition of core location based on NDT test results requires that these data are themselves reliable. A consequence is that conditional coring cannot be based on poor test result precision (TRP).

7 Identification of the conversion model

7.1 Methodology

The model identification stage consists in processing the information at locations where both core strength f_{ci} and NDT test results Tr_i are available, which can be written as a series of pairs (Tr_i, f_{ci}) , with $i = 1, N_c$. The f_{ci} strengths are “reference strength” values.

The model identification allows the values of model parameters that minimize the discrepancies between estimated strength $f_{c,est i}$ and measured strength $f_{c,i}$ at the same locations to be found. To do that, the user must:

- (a) Choose a model type. Many possible conversion models exist, corresponding to different mathematical expressions and a different number N_{par} of model parameters,

Illustrative example—TASK 6: Defining the location of cores and take cores

As recommended for EQL2, the conditional coring option is chosen. The core location is thus defined on the basis of an evenly distributed coverage of concrete properties, obtained by combining a correct coverage of all floors and the two types of structural elements (beams and columns).

This definition only requires that attention is paid to the rebound number distribution of Figure Ex-1 in order to select a correct set of cores. The choice is to take two cores for each floor (one in a beam and one in a column). The ninth core is chosen in a column of the first level. The test results are summarized in Table Ex-3 with the location of cores, the NDT test result obtained during the screening stage and the measured compressive strength.

Table Ex-3. Experimental results on the 9 cores (C = column, B = beam)

Core number	1	2	3	4	5	6	7	8	9
NDT test nr	20	59	96	138	203	256	309	331	350
Floor	1	2	2	3	4	1	2	3	4
Element type	C	C	C	C	C	B	B	B	B
Rebound number	37.0	28.4	35.1	31.7	33.0	33.5	38.6	42.0	36.1
Strength (MPa)	32.1	21.1	32.3	24.8	26.7	27.8	29.8	42.4	32.5

The processing of the 9 core strengths leads to a mean estimated strength from cores equal to 29.9 MPa and to a standard deviation equal to 6.1 MPa, which corresponds to a COV equal to 20%.

If we come back to Table-Ex2, it indicates that for this concrete mean strength and COV, five cores could be enough. The consequence is probably that with 9 cores, our final estimation may have a reduced tolerance interval than prescribed by the EQL requirement. This will be checked at the end of the assessment process.

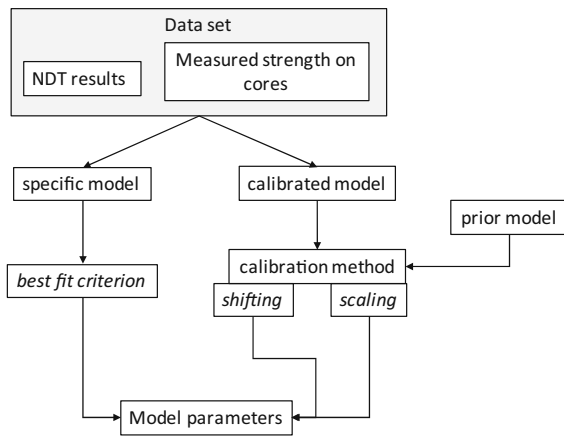


Fig. 7 The two main options for identifying model parameters

- (b) Choose and carry out a model identification approach.

Once the model parameters are identified, the model can be used to estimate the compressive strength at any point where an NDT test result is available.

A conversion model can be identified and applied either with a specific model for each test region or with the same model covering several or all test regions. There is no general guidance regarding this choice. The reader is invited to refer to the full guidelines [20] for more information.

The application of the conversion model is limited to the test regions where it has been identified.

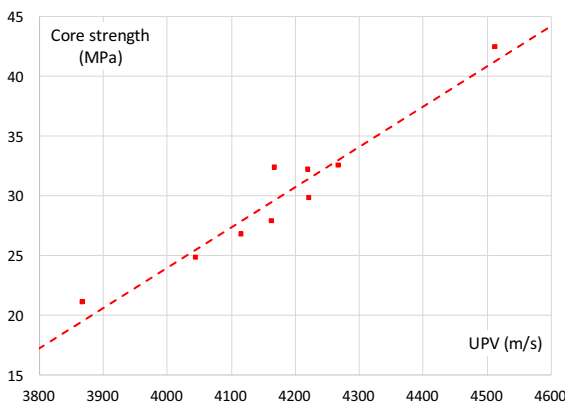


Fig. 8 Principle of identification of a conversion model, here from 9 (UPV_{*i*}, *f_{ci}*) pairs, illustrative example with a linear conversion model ($f_{c,est} = -110.8 + 0.0337 \text{ UPV}$)

Task 7: Choose a conversion model and a model identification approach

7.2 Choice of the conversion model

The general expression of the mathematical model *M* used as a conversion model can be written as:

$$f_{c,est,i} = M(\text{par}_j, \text{Tr}_{k,i})$$

where $\text{Tr}_{k,i}$ (with $i = 1, N_{\text{TL}}$) is the i -th test result of the k -th NDT method and par_j (with $j = 1, N_{\text{par}}$) is the value of the j -th model parameter. This expression works both for mono-variate models (single NDT, $k = 1$) and bi-variate models (combined NDT, $k > 1$).

Once the model identification parameter stage is completed, the value of the N_{par} model parameters par_j is known, and the model can be written as:

$$f_{c,est} = M(\text{Tr}_k)$$

The input data of a conversion model are exclusively non-destructive test results *Tr*, and explicit mathematical models must be preferred, as they make peer-review and control easier.

Since the model must be described explicitly, alternative conversion models such as neural network-based models are excluded. Furthermore, models that would require additional material data (like for instance properties of the concrete mix such as maximum aggregate size or water to cement ratio that cannot be assessed on site by nondestructive techniques) are also excluded.

The general expression of a conversion model makes possible the use of several NDT methods and thus the estimation of concrete strength by combining the values of several NDT test results (from different techniques) in a single equation. The use of combined methods requires that:

- These methods are affected in different ways by influential factor(s), as it is the case regarding the influence of concrete moisture on rebound hammer test results and ultrasonic velocity test results,
- The precision of each of the methods to be combined is similar: combining a high precision method with a poor precision method is useless,

- Since the number of parameters of the conversion model is larger for combined methods, the practical minimum number of cores increases accordingly.

It is recommended to limit the number of model parameters to 2 for mono-variate models and to three parameters when two methods are combined.

The type of mathematical model (linear, power-law, exponential types are the most common ones) is not the most influencing factor regarding the accuracy of the concrete strength assessment. When the fitting quality of various mathematical models is compared on a given dataset, it may slightly change between models, but this is not decisive.

- Linear models are preferable as long as the dataset does not exhibit a clear nonlinear tendency.
- Nonlinear models may be better when the strength range covered by the model is very large.

Whatever the model, extrapolation, i.e. using the model outside the range of values used for fitting, must be avoided.

Illustrative example—TASK 7: Choosing a conversion model

The comparison of strength values and ND test results (see Table Ex-3) does not reveal any major inconsistency and the 9 pairs of values are kept for the model parameter identification. The simplest model is chosen, with a linear relationship between the rebound hammer test result R and the estimated strength.

$$f_{c,est} = a + b \times R$$

The two parameters a and b have to be identified.

Task 8: Identify the parameters of the conversion model

7.3 Conversion model identification approach

An existing conversion model must not be used for directly deriving strength estimates for NDT test results. The identification of a specific model is mandatory in all cases. Identifying a specific model or calibrating a prior model are the two main options for identifying model parameters (Fig. 7).

7.3.1 Specific model identification: regression and bi-objective approach

The most common way of identifying the conversion model is to directly identify the N_{par} model parameters by using regression analysis. This only requires choosing a specific mathematical shape for the model expression and identifying its parameters, as illustrated in Fig. 8.

With the number of (Tr_i, f_{ci}) pairs being larger than the number of model parameters, the gap between estimated and measured strengths can be minimized. Regression analysis based on a least-square criterion is the most commonly used approach.

An alternative approach for mono-variate models is aimed at addressing simultaneously the experimental mean strength $f_{c,mean}$ and the strength standard deviation $sd(f_c)$ [15]. This leads to solving a problem with two equations and two unknowns, whose solution is straight forward. The detailed process depends on the mathematical shape of the model.

For instance, in the case of a linear model ($f_{c,est} = a + b \times Tr$), the two conditions can be respectively written as:

$$f_{c,est,mean} = a + b \times Tr_{mean} = f_{c,mean}$$

and

$$sd(f_{c,est}) = b \times sd(Tr) = sd(f_c)$$

where $f_{c,mean}$ and Tr_{mean} are respectively the mean values of measured core strength and NDT test results, and where $sd(f_c)$ and $sd(Tr)$ are respectively the standard deviations of measured core strength and NDT test results.

Consequently, the values of the two model parameters become:

$$b = sd(f_c) / sd(Tr)$$

and

$$a = f_{c,mean} - b \times Tr_{mean}$$

The same principles apply to a bi-objective power law model or a bi-objective exponential model (detailed information is available in full Guidelines [20]).

7.3.2 Model identification through calibration

An alternative approach is the identification of the conversion model using a prior model which is just calibrated to fit test results. The prior model can be written as:

$$f_{c,est,o} = M(\text{par}_j, \text{Tr})$$

where the value of all model parameters (par_j) is fixed.

The prior model must be documented and must have been established in a similar context, including the same range of strengths and the same type of concrete (e.g. fiber reinforced concrete, precast concrete, etc.).

The prior estimated strength is calculated on all points of the identification set from the Tr_i test results:

$$f_{c,est,o,i} = M(\text{par}_j, \text{Tr}_i)$$

The calibrated model can be written:

$$f_{c,est,cal} = M(\text{par}_j, \text{Tr}, C)$$

where C is a calibration factor, e.g. an additional parameter that is identified during the calibration stage. The calibration stage consists of identifying the C value (C is thus the unique degree-of-freedom of the model) in order to get the best fit between the estimated calibrated strengths and measured strengths.

In practice, there are two ways for performing this calibration, both comparing the mean value of prior strength $f_{c,est,o,mean}$ with the mean value of the measured strengths on cores $f_{c,mean}$.

- *Shift calibration*, where the difference between the two means is quantified:

$$C_S = f_{c,mean} - f_{c,est,o,mean}$$

leading after calibration to:

$$f_{c,est,cal} = f_{c,est,o,mean} + C_S$$

- *Scaling calibration*, where the ratio between the two mean values is quantified:

$$C_M = f_{c,mean}/f_{c,est,o,mean}$$

leading after calibration to:

$$f_{c,est,cal} = C_M \times f_{c,est,o,mean}$$

7.3.3 Recommendation regarding the model identification approach

There is no general guidance for choosing the best model identification approach, which may depend on the situation.

An important criterion is the number of cores N_c , since the lower N_c is, the lower the number of free model parameters must be. This can lead to choosing a calibration approach (only one free parameter, C) when N_c is very low (namely equal to 3 or 4), while a specific model is preferable when N_c is larger.

When concrete variability has to be assessed (i.e. for EQL2 and EQL3), the bi-objective approach is recommended, since it is the only one to correctly address it.

Illustrative example—TASK 8: Choosing a model identification approach and identifying the model parameters

As EQL2 aims at quantifying the concrete standard deviation, the bi-objective approach, which aims at capturing both mean strength and strength variability, is chosen.

The two equations defining the a and b conversion model parameters are:

$$b = \text{sd}(f_{c,est})/\text{sd}(R) \text{ and } a = f_{c,mean} - b \times R_{mean}$$

With the data of Table Ex-3, $f_{c,mean} = 29.9$ MPa,

$\text{sd}(f_{c,est}) = 6.1$ MPa, $R_{mean} = 35.1$ and $\text{sd}(R) = 4.0$, it is found that $b = 1.52$ and $a = -23.2$. The conversion model can finally be written as:

$$f_{c,est} = -23.2 + 1.52 \times R.$$

This is plotted in figure Ex-2 where the experimental results on the nine cores are also plotted.

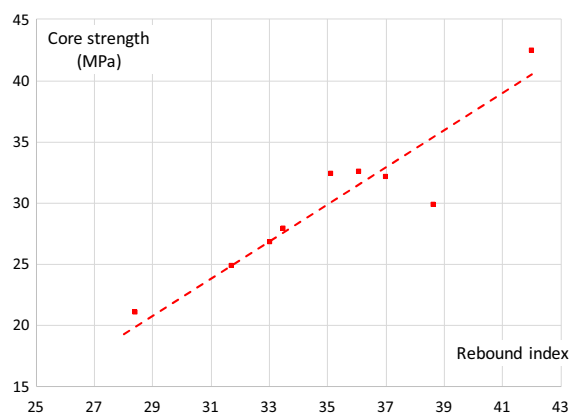


Figure Ex-2. Core data and conversion model identified with the bi-objective approach.

8 Strength estimation: using the conversion model and quantifying the final uncertainty

8.1 Estimating local strength, mean strength and strength standard deviation with the conversion model

Local strength is the first parameter that needs to be estimated, with a target accuracy which depends on the estimation quality level EQL, as defined in Table 1. Once the conversion model is identified, it can be used for estimating strength at all test locations where only one NDT test result is available (see Fig. 3). This leads to a series of local strength estimates that can be combined with strength values directly measured on the N_c cores.

Illustrative example—Using the conversion model for strength estimation

From the 60 rebound numbers that have been obtained at the screening stage (36 on columns and 24 on beams), 60 estimated strength values can be computed and the pattern of strength distribution can be derived. Statistical properties for each floor and component type can also be calculated. These results are provided in Table Ex-4.

Table Ex-4. Estimated mean strength and estimated strength standard deviation in the building (MPa)

Structural element	Floor	N_{TL}	Mean	sd
Beam	1	6	31.0	3.8
Beam	2	6	31.9	5.2
Beam	3	6	36.3	4.4
Beam	4	6	32.7	5.8
Column	1	9	27.8	5.0
Column	2	9	28.2	4.8
Column	3	9	25.0	3.9
Column	4	9	32.1	3.8
Beams (all)		24	33.0	5.0
Columns (all)		36	28.3	4.9
Building (all elements)		60	30.2	5.4

The overall average concrete strength and the standard deviation were the two targets of the investigation. The final estimation is respectively 30.2 MPa and 5.4 MPa.

These numbers are close to those provided by the 9 cores, which were respectively 29.9 MPa and 6.1 MPa, with only a small reduction of the estimated variability.

As an alternative to overall strength values, the strength distribution within the building can be used by engineers for refining the assumption of design strength values in the structural assessment.

8.2 Model uncertainty: fitting error and prediction error

Task 9: Check the model error

Whatever the option used for identifying the model parameters (specific or calibrated), it is recommended to check the accuracy of strength estimation. This is mandatory for EQL3.

The recommended measure of model error is the root mean square error (RMSE) which provides the statistical lack of fit, expressed in MPa, i.e. the “mean distance” between the real value and the value which is estimated by using the model. The RMSE values directly quantify (in MPa) the mean statistical error in the estimation of the local strength value and can be obtained by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (f_{c,est\ i} - f_{ci})^2}{n}}$$

where n is the number of points where $f_{c,est\ i}$ is calculated. RMSE is an adequate measure for comparing forecasting errors of different models.

One must understand the difference between the fitting error ($RMSE_{fit}$) and the prediction error ($RMSE_{pred}$). The former relates to the dataset used for model identification while the latter relates to a new dataset. The fitting error is calculated on the identification dataset, by comparing f_{ci} with $f_{c,est,i}$ for all N_c points at which pair data (Tr_i, f_{ci}) were used in order to calibrate the model. The prediction error is calculated by comparing measured values and strength estimates at points which have not been used to calibrate the model.

Prediction error is always larger than fitting error, which is due to the random nature of the estimated parameters of the conversion model. Any estimation of the model error based only on the fitting error, as is common in practice, is meaningless.

8.3 Quantifying the prediction error

Two main possibilities exist for quantifying $RMSE_{pred}$: (a) direct assessment on a control dataset, (b) cross-validation analysis.

Direct checking consists of keeping a number of cores for validation, i.e. not using the corresponding strength data to fit the conversion model and, once the model has been identified, to estimate the concrete

strength from NDT test results of this new dataset, and to compare the estimated strength to the measured strength on this control set. The main drawback is that cores being expensive, their number is limited, and it is usually more useful to use all core data to improve the model identification.

Cross-checking offers an alternative approach. In most real applications, only a limited amount of data is available, which leads to the idea of splitting the available dataset: part of data, the training set (TS), is used for fitting the model while the remaining part of data, the validation set (VS), is used for evaluating the performance of the model. A single data split yields a validation estimate of the risk, and averaging over several splits yields a cross-validation estimate. There are several ways of dividing the dataset into TS and VS by varying the size of the two sets, which is illustrated in the RILEM TC-ISC 249 Guidelines.

Illustrative example—TASK 9: Quantifying the model error

The conversion model was identified from the 9 pairs of strengths and rebound numbers, which lead to:

$$f_{c,est} = -23.2 + 1.52 \times R$$

The calculated value of the fitting error is

$$RMSE_{fit} = 2.2 \text{ MPa.}$$

As we have no additional cores for testing the model on new data, a leave-one-out cross validation is carried out. It consists in identifying nine models, each of them from eight pairs of strengths and rebound numbers, where only one pair is removed from the original set.

For each model, the identified equation is used to estimate the strength on the ninth core, and the difference between the estimated strength and the core strength is calculated. These nine differences are finally processed to derive the $RMSE_{pred}$ value.

The details of these calculations are not developed here. The nine models lead to average values of a and b coefficients equal to—23.0 and 1.51 respectively with standard deviations of 3.3 and 0.10 respectively. One finally gets:

$$RMSE_{pred} = 2.9 \text{ MPa}$$

This value is significantly larger than the fitting error. It is a good estimate of the predictive error of the conversion model, and corresponds to the accuracy for the estimation of any local strength from the measured rebound number. It is in full agreement with the indicative values of Table 1 for the EQL2 requirement.

Illustrative example—SYNTHESIS

The conversion model was identified from the 9 pairs of strengths and rebound numbers, which lead to:

$$f_{c,est} = -23.2 + 1.52 \times R$$

As we have no additional cores for testing the model on new data, a leave-one-out cross validation is carried out. It consists of identifying conversion models on nine 8-core subsets.

The final results of the investigation can be summarized:

- conversion model: $f_{c,est} = -23.2 + 1.52 R$
- estimated mean strength $f_{c,est,mean} = 30.2 \text{ MPa}$
- estimated strength standard deviation $sd(f_{c,est}) = 5.4 \text{ MPa}$
- prediction error on local strength values
 $RMSE_{pred} = 2.9 \text{ MPa}$

The mapping of strength within the building from the 80 rebound numbers is an additional result.

Acknowledgements The authors of this RILEM TC 249-ISC gratefully acknowledge all contributors who are cited above and do not appear as full authors of the recommendation for administrative reasons.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Masi A, Chiauzzi L (2013) An experimental study on the within-member variability of in situ concrete strength in RC building structures. *Constr Build Mater* 47:951–961
- Balayssac JP, Laurens S, Arliguie G, Breyse D, Garnier V, Derobert V, Piwakowski B (2012) Description of the general outlines of the French project SENSO—quality assessment and limits of different NDT methods. *Constr Build Mater* 35:131–138
- RILEM TC7 (1976) Non-destructive testing. Comparison between recommendations existing in some East European countries concerning the determination of concrete strength by surface hardness methods, Chair. I. Facaoaru. *Mater Struct* 9(51):207–210
- RILEM TC43 (1993) Draft recommendation for in situ concrete strength determination by combined non-destructive methods, Chair. I. Facaoaru. *Mater Struct* 26(155):43–49
- EN 12504-2, Testing concrete in structures—Part 2: Non-destructive testing—Determination of rebound number, ISSN 0335-3931, March 2013; EN 12504-3, Testing concrete in structures—Part 3: Determination of pull-out force, ISSN 0335-3931, June 2005; EN 12504-4, Testing concrete—Part 4: Determination of ultrasonic pulse velocity, ISSN 0335-3931, May 2005
- ASTM C803 (2010) Standard test method for penetration resistance of hardened concrete. American Society for testing Materials (ASTM)
- ASTM C805 (2013) Standard test method for rebound number of hardened concrete. American Society for testing Materials (ASTM)



8. ASTM C823 (2017) Standard practice for examination and sampling of hardened concrete in constructions. American Society for testing Materials (ASTM)
9. ASTM C900 (2015) Standard practice for pullout strength of hardened concrete. American Society for testing Materials (ASTM)
10. ACI228.1R-03 (2003) In-place methods to estimate concrete strengths, chair. N.J. Carino. American Concrete Institute, Detroit
11. EN 13791 (2007) European Standard, Assessment of in situ compressive strength in structures and precast concrete components, ISSN 0335-3931 (ongoing revision—updated version EN 13791 expected in 2020)
12. Breyse D, Balayssac JP (2018) Strength assessment in reinforced concrete structures: from research to improved practices. *Constr Build Mater* 182:1–9
13. Masi A, Chiauuzzi L, Manfredi V (2016) Criteria for identifying concrete homogeneous areas for the estimation of in situ strength in RC buildings. *Constr Build Mater* 121:586–597
14. Alwash M, Sbartai ZM, Breyse D (2016) Nondestructive assessment of both mean strength and variability of concrete: a new bi-objective approach. *Constr Build Mater* 113:880–889
15. Alwash M, Breyse D, Sbartai ZM (2017) Using Monte-Carlo simulations to evaluate the efficiency of different strategies for nondestructive assessment of concrete strength. *Mater Struct* 50:1
16. Breyse D, Balayssac JP, Biondi S, Borosnyói A, Candiogliota E, Chiauuzzi L, Garnier G, Grantham M, Gunes O, Luprano V, Masi A, Pfister V, Sbartai ZM, Szilagyi K (2017) Nondestructive assessment of in situ concrete strength: comparison of approaches through an international benchmark. *Mater Struct* 50:133
17. Masi A, Digrisolo A, Santarsiero G (2014) Concrete strength variability in Italian RC buildings: analysis of a large database of core tests. *Appl Mech Mater* 597:283–290
18. Pereira N, Romão X (2018) Assessing concrete strength variability in existing structures based on the results of NDTs. *Constr Build Mater* 173:786–800
19. Pfister V, Tundo A, Luprano VAM (2014) Evaluation of concrete strength by means of ultrasonic waves: a method for the selection of coring position. *Constr Build Mater* 61:278–284
20. Breyse D, Balayssac JP et al (2019) Non-destructive in situ strength assessment of concrete: practical application of 249-ISC RILEM TC recommendations. Springer, Berlin

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

