



HAL
open science

Assessment of concrete strength in existing structures using nondestructive tests and cores : analysis of current methodology and recommendations for more reliable assessment

Maitham Fadhil Abbas Alwash

► To cite this version:

Maitham Fadhil Abbas Alwash. Assessment of concrete strength in existing structures using non-destructive tests and cores : analysis of current methodology and recommendations for more reliable assessment. Mechanics [physics]. Université de Bordeaux, 2017. English. NNT : 2017BORD0587 . tel-01531241

HAL Id: tel-01531241

<https://theses.hal.science/tel-01531241>

Submitted on 1 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE
DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE SCIENCES PHYSIQUES ET DE L'INGENIEUR
SPÉCIALITÉ : MÉCANIQUE

Par Maitham Fadhil Abbas ALWASH

**ASSESSMENT OF CONCRETE STRENGTH IN EXISTING STRUCTURES
USING NONDESTRUCTIVE TESTS AND CORES: ANALYSIS OF
CURRENT METHODOLOGY AND RECOMMENDATIONS FOR MORE
RELIABLE ASSESSMENT**

Sous la direction de : Denys BREYSSE
(co-directeur : Zoubir Mehdi SBARTAÏ)

Soutenue le 10 mai 2017

Membres du jury :

M. BALAYSSAC, Jean-Paul	Professeur, Université Paul Sabatier	Rapporteur
Mme. YOTTE, Sylvie	Professeur, Université de Limoges	Rapporteur
M. GONÇALVES, Arlindo	Directeur de recherche, LNEC, Portugal	Président
Mme. VILLAIN, Géraldine	Chercheur-Ingénieur, HDR, IFSTTAR	Examineur
M. BREYSSE, Denys	Professeur, Université de Bordeaux	Directeur de thèse
M. SBARTAÏ, Zoubir Mehdi	MCF, HDR, Université de Bordeaux	Co-directeur de thèse
M. Vincent GARNIER	MCF, HDR, Université Aix-Marseille	Invité

Évaluation de la résistance mécanique du béton dans les ouvrages existants en utilisant les tests non-destructifs et carottes: analyse de la méthodologie courante et recommandations pour évaluation plus fiable

Résumé

Pour évaluer la résistance mécanique du béton dans un ouvrage existant, la méthodologie courante combine des mesures non destructives (CND) comme le rebond ou/et la vitesse des ondes ultrasoniques avec la technique destructive (carottes) afin de produire une relation “modèle de conversion” entre la résistance mécanique et les mesures CND. Le modèle de conversion est utilisé pour estimer la valeur locale de résistance mécanique à chaque emplacement de test en utilisant la valeur CND correspondante. Ensuite, on calcule les estimations de la résistance moyenne et/ou de l'écart-type de la résistance (variabilité de la résistance du béton). Cependant, la fiabilité d'estimation est toujours discutable en raison des incertitudes associées à l'évaluation de la résistance basée sur les mesures CND. Pour améliorer la fiabilité, les incertitudes doivent être réduites en spécifiant et en contrôlant leurs facteurs d'influence. Par conséquent, l'objectif de cette thèse est d'analyser la méthodologie d'évaluation courante afin de fournir des recommandations pratiques qui peuvent améliorer la fiabilité de l'évaluation de la résistance in-situ du béton dans les ouvrages existantes par des tests non destructifs et des carottes.

Pour ce but, un simulateur a été construit afin d'analyser les effets des facteurs les plus influents en utilisant une vaste campagne de données provenant de sources différentes (études in situ ou en laboratoire et données synthétiques générées).

La première contribution de ce travail est le développement d'une nouvelle approche d'identification du modèle “bi-objectif” qui peut efficacement capturer la variabilité de la résistance mécanique en plus de la moyenne. Après avoir étudié l'effet du mode de sélection des emplacements pour les carottes, une méthode a été proposée pour sélectionner ces emplacements en fonction des mesures CND “sélection conditionnelle” qui améliore la qualité de l'évaluation sans coût supplémentaire. Une dernière innovation est l'établissement de courbes de risque qui quantifient la relation entre le nombre de carottes et la précision de l'estimation. Enfin, des recommandations ont été formulées afin de fournir des estimations plus fiables.

Mots-clés: résistance in-situ du béton, résistance mécanique moyenne, variabilité du béton, vitesse des ondes ultrasoniques, rebond, CND, combinaison, simulation, incertitude, stratégie d'évaluation, fiabilité d'évaluation

Unité de recherche

Institut de Mécanique et d'Ingénierie, UMR 5295
Département Génie Civil Environnemental
Bâtiment B18, Allée Geoffroy Saint Hilaire, CS 50023, 33615 Pessac Cedex, France

Assessment of concrete strength in existing structures using nondestructive tests and cores: analysis of current methodology and recommendations for more reliable assessment

Abstract

To assess concrete strength in an existing structure, the current methodology combines nondestructive measurements (NDT) like rebound hammer or/and pulse velocity with destructive technique (cores) in order to implement a relationship “conversion model” between the compressive strength and NDT measurements. The conversion model is used to estimate the local strength value at each test location using the corresponding NDT value. Then the estimated mean strength and/or estimated strength standard deviation (concrete strength variability) values are calculated. However, the reliability of these estimated values is always a questionable issue because of the uncertainties associated with the strength assessment based upon NDT measurements. To improve the reliability, the uncertainties must be reduced by specifying and controlling their influencing factors. Therefore, the objective of this thesis is to analyze the current assessment methodology in order to provide practical recommendations that can improve the reliability of assessing the in-situ strength in existing concrete structures by nondestructive tests and cores.

To this end, a simulator was built in order to analyze the effects of the most influencing factors using a large campaign of datasets from different sources (in-situ or laboratory studies, and generated synthetic data).

The first contribution of this work is the development of a new model identification approach “bi-objective” that can efficiently capture the strength variability in addition to the mean strength. After studying the effect of the way of selection the core locations, a method was proposed to select these locations depending on the NDT measurements “conditional selection” that improves the quality of assessment without additional cost. A third innovation was the development of a procedure to identify the relation between the number of cores and the accuracy of the estimation. Finally recommendations were derived in order to provide more reliable estimated values.

Keywords : in-situ concrete strength, mean strength, concrete variability, ultrasonic pulse velocity, rebound hammer, NDT, combination, simulation, uncertainty, assessment strategy, assessment reliability

Unité de recherche

Institut de Mécanique et d'Ingénierie, UMR 5295
Département Génie Civil Environnemental
Bâtiment B18, Allée Geoffroy Saint Hilaire, CS 50023, 33615 Pessac Cedex, France

ACKNOWLEDGEMENT

Many people were involved in this project. It seems almost impossible to express my gratitude to everyone who helped during its completion. To all who should be mentioned for their generous assistance, my warmest thanks.

In particular, I would like to express my sincerest gratitude to my supervisor Prof. Denys Breyse and my co-supervisor Dr. Zoubir Mehdi Sbartaï for their guidance, patience, motivation, enthusiasm, immense knowledge and for three years of inspiring discussions.

I would like to express my deepest gratitude to all the members of the jury for their interest and the quality of their remarks.

I have the chance to participate several meetings of the RILEM committee technique TC 249-ISC, so I would like to acknowledge the chairman Prof. Denys Breyse and the members of this committee for allowing me attending their meetings, presenting some of my results and for their valuable discussions. Special thanks are due to S. Biondi, M. Fontan, V. Luprano S. and K. Szilágyi for sharing the information about their strategies in the benchmark that was carried out by the member of this committee technique.

During this work I had the opportunity to collaborate with two teams of researchers: Prof. Said Kenai and Dr. Khoudja Ali Benyahia from University of Blida and University of Chlef (Algeria), and Prof. Adorján Borosnyói and Dr. Katalin Szilágyi from Budapest University of Technology And Economics (Hungary), so I would like to acknowledge them for sharing their valuable data.

I am also grateful to the staff at the Civil and Environmental Engineering Department (GCE) in the laboratory I2M at the University of Bordeaux for their motivation and assistance.

I am equally grateful to all my PhD colleagues, in particular: Nicolas Piegay, Nicolas Houillon, Elodie Rodney, Abdelhammid Mokeddem, Cécile Verdet, Ghina Darwich and Thi Thuy Phuong Tran.

Last but not the least, I am thanking my parents. They brought me forth, raised me, supported me, taught me, loved me and still loving me. I would like also to express my heartfelt appreciation to my uncle Khudhur Al-Falooji and my aunt Saadiyah Alwash for their help and spiritual support. I would like extend my appreciation to my wife for her patient, support and concern toward completing this work. Finally, special thanks to my angles: Deenh, Mustafa, Ahmed and Mohammed. To them I whole-heartedly dedicate this thesis.

Maitham ALWASH

May 2017

LIST OF CONTENTS

LIST OF CONTENTS	9
LIST OF FIGURES	13
LIST OF TABLES	17
LISTS OF ABBREVIATIONS / NOTATIONS / GLOSSARY OF TERMS	19
LIST OF PUBLICATIONS	21
RÉSUMÉ SUBSTANTIEL	23
CHAPTER 1: INTRODUCTION	29
1.1 Problem statement.....	29
1.2 Scope of thesis.....	30
1.3 Objective and general research methodology.....	31
CHAPTER 2: CONCRETE STRENGTH ASSESSMENT IN THE EXISTING STRUCTURES: LITERATURE REVIEW	33
2.1 Introduction.....	33
2.2 Strength assessment using cores only.....	33
2.2.1 Planning an investigation program: number and location of cores.....	34
2.2.2 Drilling cores.....	35
2.2.3 Testing cores.....	36
2.2.4 Interpreting the core strengths.....	37
2.2.5 Conclusion of the assessment methodology when using cores only.....	38
2.3 Strength assessment using nondestructive tests only.....	39
2.3.1 Rebound hammer method.....	39
2.3.2 Ultrasonic pulse velocity method.....	43
2.4 Strength assessment using cores and single nondestructive technique.....	47
2.4.1 Model identification approaches: Regression approach.....	48
2.4.2 Model identification approaches: Calibration approach.....	50
2.4.3 Types of models.....	52
2.4.4 Factors affecting the quality of assessment.....	52
2.4.5 Sources of uncertainty.....	53
2.4.6 Quality of assessment.....	54
2.4.7 Conclusion of assessing concrete strength using cores and single nondestructive technique.....	54
2.5 Strength assessment using cores and combination of nondestructive techniques.....	55
2.5.1 Model identification approaches.....	56
2.5.2 Types of models.....	57

2.5.3	Factors affecting the quality of assessment	58
2.5.4	Sources of uncertainty.....	58
2.5.5	Quality of assessment	58
2.5.6	Efficiency of combination.....	58
2.5.7	Conclusion of assessing concrete strength using cores and combination of nondestructive techniques.....	59
2.6	Conclusions	60
CHAPTER 3: MEANS AND TOOLS.....		61
3.1	Introduction	61
3.2	Definition of the assessment strategy.....	61
3.3	Sources of data.....	63
3.3.1	In-situ on structure and laboratory studies data	63
3.3.2	Synthetic data	67
3.4	The simulator developed in the present thesis.....	69
CHAPTER 4: ANALYSIS OF CURRENT METHODOLOGY FOR CONCRETE STRENGTH ASSESSMENT		75
4.1	Introduction	75
4.2	Studying the effect of several key influencing factors.....	75
4.2.1	Effect of number of test locations for cores.....	75
4.2.2	Effect of quality of measurements	77
4.2.3	Effect of in-situ concrete strength variability	78
4.2.4	Effect of type of model (linear or nonlinear).....	79
4.2.5	Effect of combining NDT techniques	81
4.2.6	Conclusions.....	83
4.3	Analyzing several assessment strategies presented in an international benchmark.....	84
4.3.1	The benchmark in brief.....	84
4.3.2	Simulation of the assessment strategies	85
4.3.3	Analysis of simulation results.....	86
4.3.4	Conclusions.....	90
4.4	Analyzing the existing model identification approaches	91
4.4.1	Comparing the prediction capacity of the existing model identification approaches	91
4.4.2	Conclusions.....	96
CHAPTER 5: DEVELOPING NEW MODEL IDENTIFICATION APPROACH: BI-OBJECTIVE APPROACH.....		97
5.1	Introduction	97
5.2	Development of the bi-objective approach	97
5.2.1	Derivation of the model parameters for the case of linear model	98
5.2.2	Derivation of the model parameters for the case of nonlinear model.....	98

5.3	Validation of the bi-objective approach.....	99
5.3.1	Case of linear model	99
5.3.2	Case of power model	103
5.3.3	Failure of the model identified by each approach	106
5.4	Conclusions	108
CHAPTER 6: QUALITY OF ASSESSMENT AND RECOMMENDATIONS FOR BETTER PRACTICE		109
6.1	Introduction	109
6.2	Analysis of quality of assessment using synthetic datasets.....	109
6.2.1	Datasets.....	109
6.2.2	Assessing mean strength and strength standard deviation (concrete variability) and developing the cumulative distribution functions (CDF)	111
6.2.3	Assessing the quality of estimation by developing “Risk Curves”	113
6.2.4	Studying the effect of the quality of measurements on the quality of assessment	119
6.2.5	Studying the effect of the way of selection the <i>NC</i> test locations on the quality of assessment ..	122
6.2.6	What is the minimum number of test locations for cores that can ensure a specific quality of assessment?	124
6.3	Analysis of quality of assessment using real datasets.....	125
6.4	Conclusions	127
6.5	Recommendations for better practice.....	128
CHAPTER 7: CONCLUSIONS AND PERSPECTIVES		131
INTRODUCTION GÉNÉRALE.....		135
CONCLUSION GÉNÉRALE.....		139
REFERENCES		143
APPENDIX (A).....		153
APPENDIX (B).....		155
APPENDIX (C).....		157
APPENDIX (D).....		159

LIST OF FIGURES

Figure 1.1 General research methodology.....	32
Figure 2.1 Error of the estimated mean strength as a function of number of cores, (ACI 214.4R, 2003).....	35
Figure 2.2 Core drilling rig.....	36
Figure 2.3 Schematic sectional view of rebound hammer illustrating the operation process.....	40
Figure 2.4 Conversion curves for estimating the average compressive strength of a cylinder from rebound number, (Proceq SA, 2016).....	42
Figure 2.5 Some conversion curves identified by different researchers for estimating the cube compressive strength from rebound number.....	43
Figure 2.6 Ultrasonic pulse velocity device (schematic diagram), (ASTM C597, 2002).....	44
Figure 2.7 Transducers arrangements in ultrasonic pulse velocity test.....	45
Figure 2.8 Some conversion curves identified by different researchers for estimating the cube compressive strength from ultrasonic pulse velocity.....	47
Figure 2.9 Scatter diagram of the (x, f_c) pairs with a regression model.....	49
Figure 2.10 Scatter diagram of the (x, f_c) pairs with models obtained by the calibration approach.....	51
Figure 2.11 Iso-strength curves for a reference concrete in SONREB method (IAEA, 2002).....	56
Figure 3.1 Definition of the assessment strategy.....	62
Figure 3.2 Algorithm for generating dataset using synthetic simulation.....	69
Figure 3.3 Main algorithm of the simulator prepared for the analysis in the present study.....	71
Figure 3.4 Two cases identified wrong conversion models obtained from the analysis of real dataset using $NC=4$ cores selected randomly from the whole dataset size=60.....	72
Figure 4.1 Fitting error, FEC, as a function of NC resulted from analyzing the dataset S1.....	76
Figure 4.2 Fitting and prediction errors (FEC and PEC) as a function of NC resulted from analyzing the in-situ dataset S1.....	77
Figure 4.3 Effect of quality of measurements on the Predictive RMSE, PEC, after analyzing a synthetic dataset ($\bar{f}_{ist} = 45$ MPa, $s(f_{ist}) = 5$ MPa) with single NDT technique (pulse velocity).....	78

Figure 4.4 Effect of true in-situ strength variability on the predictive <i>RMSE</i> for single and combined NDT techniques with average qualities and for $NC=10$	79
Figure 4.5 Effect of model type on the predictive <i>RMSE</i> for single and combined techniques when analyzing in-situ dataset S1.....	80
Figure 4.6 Effect of model type on the predictive <i>RMSE</i> for single and combined techniques when analyzing laboratory study dataset S12	80
Figure 4.7 Effect of model type on the predictive <i>RMSE</i> for single and combined techniques when analyzing laboratory study dataset S10	81
Figure 4.8 Effect of model type on the predictive <i>RMSE</i> for single and combined techniques when analyzing laboratory study dataset S11	81
Figure 4.9 Estimation of mean strength and strength standard deviation (concrete variability) for 16 strategies at KL3 (Breysse, et al., 2017)	85
Figure 4.10 Comparison of estimated (mean strength and strength standard deviation) to the true values; results from 1000-simulations of contributor's strategy B at KL3.....	86
Figure 4.11 Predicting mean strength using the regression and calibration approaches	93
Figure 4.12 Predicting strength variability using the regression and calibration approaches	94
Figure 5.1 Predicting mean strength using the regression, calibration, and bi-objective approaches, all linear models.....	100
Figure 5.2 Predicting strength variability using the regression, calibration, and bi-objective approaches, all linear models.....	101
Figure 5.3 Predicting mean strength and strength standard deviation (strength variability) using the regression, calibration, and bi-objective approaches, linear and power models.....	104
Figure 6.1 Clouds of 1000- (V, f_c) for several generated datasets.....	111
Figure 6.2 CDF curves of mean strength and concrete strength variability estimations for dataset D20-20-AQ and when the NDT technique is the pulse velocity.....	112
Figure 6.3 Proposed concept to derive risk value corresponding to a specific uncertainty level using CDF curve: graph (a) illustrates the proposed concept, while graph (b) displays how the change in the scatter of CDF curve changes the risk value.....	113
Figure 6.4 Risk curves of all datasets with (AQ) for mean strength estimation with: $U=\pm 10\%$, $NT=100$, and random core selection.....	115
Figure 6.5 Risk curves of all datasets with (AQ) for strength variability estimation with: $U=\pm 25\%$, $NT=100$, and random core selection.....	117

Figure 6.6 Comparing the risk curves obtained from absolute and relative uncertainty levels for the case of strength variability estimation with: average quality measurements, $NT=100$, and random core selection.....	118
Figure 6.7 Studying the effect of quality of measurements on the risk curves corresponding to each of mean strength and strength variability estimations with $NT=100$, and random core selection.....	120
Figure 6.8 Risk curves corresponding to the uncertainty level $U=\pm 35\%$ for the strength variability estimation with: two different quality of measurements (HQ and LQ), $NT=100$, and random core selection	121
Figure 6.9 Studying the effect of the way of core location selection on the risk curves corresponding to each of mean strength and strength variability estimations with $NT=100$, and for the three quality of measurements: HQ, AQ, and LQ.....	123
Figure 6.10 Assessment methodology for evaluating concrete strength in the existing structures that proposed in the present study.....	129

LIST OF TABLES

Table 2.1 Strength correction factors values as provided by ACI 214.4R-03, (f_{cs} in MPa)	38
Table 2.2 Several models derived by different researchers in order to estimate strength using single NDT technique	52
Table 2.3 Several models derived by different researchers in order to estimate strength using combination of rebound hammer and pulse velocity techniques.....	57
Table 2.4 r^2 values as provided by different researchers for models derived for: single rebound hammer technique, single ultrasonic pulse velocity technique and their combination	59
Table 3.1 Summary of the collected datasets.....	65
Table 4.1 Studying the efficiency of combination, bold numbers indicate the minimum <i>RMSE</i> values corresponding to each dataset.....	82
Table 4.2 Characteristics of the five strategies at KL3 and values of the quality indicators....	87
Table 4.3 Characteristics of the selected strategies for the parametric study and values of the quality indicators (in bold: characteristics differ from those in Table 4.2)	89
Table 4.4 Uncalibrated models selected from literature for the present analysis when using the calibration approaches	92
Table 4.5 Standard deviation values of NI-values of mean strength and concrete variability estimated by regression and calibration approaches (all linear models) for the datasets S18 and S39 – all results are in MPa	95
Table 5.1 Standard deviation values from <i>NI</i> values of mean strength and concrete variability estimated by regression and bi-objective approaches (linear models) for the datasets S18 and S39 – all results are in MPa	102
Table 5.2 Standard deviation values from <i>NI</i> values of mean strength and concrete variability estimated by regression and bi-objective approaches (linear and power models) for the datasets S1 – all results are in MPa.....	104
Table 5.3 Model failure percentages when identifying linear models using regression and bi-objective approaches for datasets S1, S18, and S39.	107
Table 6.1 Characteristics of 36 synthetic datasets considered for analyzing the quality of assessment.....	110
Table 6.2 Minimum NC as a function of true in-situ concrete variability, quality of measurements and the way of core locations selection for sample size $NT=100$, and $NC \leq 20$, (N/A means value is not available)	125
Table 6.3 Summary of the effects of factors considered in this study on the quality of assessment analyzed using real datasets.....	126

LISTS OF ABBREVIATIONS / NOTATIONS / GLOSSARY OF TERMS

Glossary of terms

<i>Strength variability</i>	The standard deviation of compressive strengths corresponding to one test region
<i>Mean strength</i>	The average value of compressive strengths corresponding to one test region
<i>Test location</i>	limited area selected for measurements used to provide one test result for each type of measurements (DT or NDT)
<i>Test result</i>	value of NDT or DT that corresponds to one test location, the test result represents the average or median of several repetitions of the test method at this location
<i>Test region</i>	one member or several members (or the whole structure) assumed or known to belong to the same concrete (i.e. same quality or composition). It includes several test locations

Abbreviations / Notations

a, b, b_1, b_2, \dots	parameters of the conversion models
$CV(f_c)$	the concrete strength variability (in terms of coefficient of variation) for the whole population (dataset), %
DT	destructive test method (cores)
f_c	core compressive strength corresponding to one test location, or the strength corresponding to a standard specimen in the case of laboratory study dataset. For the case of real data this strength will consider as a true (reference) value
f_{cest}	estimated compressive strength corresponding to one test location
$f_{c\ uncal.}$	estimated compressive strength calculated from an uncalibrated prior model that selected from literature or standards
FEC	fitting error curve
f_{ist}	true in-situ compressive strength corresponding to one test location. It is known only in the case of synthetic data
NC	number of test locations for cores (or number of standard specimens for the case of laboratory dataset) used to identify the model. Considering that one core is extracted from each location, NC also refers to the

	number of cores
<i>NDT</i>	nondestructive test method
<i>NRMSE</i>	normalized root mean square error
<i>NRMSE_{av}</i>	normalized root mean square error for the mean strength
<i>NRMSE_{sd}</i>	normalized root mean square error for the strength standard deviation
<i>NT</i>	total number of test locations for NDT measurements
<i>r²</i>	coefficient of determination
<i>R</i>	rebound hammer test result corresponding to one test location (represents the average of several replicates within the test location)
<i>RMSE</i>	root mean square error
<i>PEC</i>	prediction error curve
<i>superscript</i> ⁻	mean value of the variable under consideration
<i>s()</i>	Standard deviation of the variable under consideration
<i>s(f_{cest})</i>	Estimated concrete strength variability
<i>s(f_{ist})</i>	True in-situ concrete strength variability (known only in the case of synthetic data)
<i>s(\bar{f}_{cest})</i>	the standard deviation of the estimated mean strengths, it means the variability of the estimated mean strength
<i>s(s(f_{cest}))</i>	the standard deviation of the estimated strength standard deviations, it means the variability of the estimated concrete strength variability
<i>V</i>	ultrasonic pulse velocity test result corresponding to one test location (represents the average of several replicates within the test location)
<i>x</i>	test result, <i>V</i> or <i>R</i> , of the NDT measurement corresponding to one test location

LIST OF PUBLICATIONS

Alwash, M., Breysse, D., Sbartai, ZM., 2015. Non-destructive strength evaluation of concrete: Analysis of some key factors using synthetic simulations. *Construction and Building Materials*, 99, pp. 235-245.

Alwash, M., Sbartai, ZM., Breysse D., 2016. Non-destructive assessment of both mean strength and variability of concrete: a new bi-objective approach. *Construction and Building Materials*, 113, pp. 880-889.

Alwash, M., Breysse, D., Sbartai, ZM., 2017. Using Monte-Carlo simulations to evaluate the efficiency of different strategies for nondestructive assessment of concrete strength. *Materials and Structure*, 50, 14p.

Alwash, M., Breysse, D., Sbartai, ZM., Szilágyi, K., Borosnyói, A., 2017. Factors affecting the reliability of assessing the concrete strength by rebound hammer and cores. *Construction and Building Materials*, 140, pp. 354-363.

Alwash, M., Breysse, D., Sbartai, ZM., 2014. *A parametric study using synthetic simulation for improving the evaluation of concrete strength in existing buildings by NDT techniques*. Orléans, 32^{èmes} Rencontres Universitaires de Génie Civil.

Alwash, M., Breysse, D., Sbartai, ZM., 2015. *Amélioration de la méthodologie d'évaluation non destructive de la résistance du béton dans les structures existantes*. Bayonne, 33^{èmes} Rencontres Universitaires de Génie Civil.

Alwash, M., Sbartai, ZM., Breysse D., 2015. *Evaluation of Concrete Strength by Non-destructive Techniques: a New approach for Calibration*. Berlin, International Symposium Non-Destructive Testing in Civil Engineering, NDT-CE.

Alwash, M., Breysse, D., Sbartai, ZM., 2015. *Analysis of Key Influent Factors of Investigation for NDT Concrete Strength Estimation in Relation with an International Benchmark*. Berlin, International Symposium Non-Destructive Testing in Civil Engineering, NDT-CE. (poster)

Alwash, M., Breysse, D., Sbartai, ZM., 2016. *Comment fiabiliser l'estimation de la résistance en place du béton dans les ouvrages existants par méthodes non destructives?*. Nancy, Journées Fiabilité Matériaux et Structures (JFMS).

Alwash, M., 2016. *Improving the methodology of nondestructive evaluation of concrete strength in the existing concrete structures*. Liège, Belgium, 34^{èmes} Rencontres Universitaires de Génie Civil. (young researchers competition)

RÉSUMÉ SUBSTANTIEL

Introduction

La méthodologie actuelle pour évaluer la résistance mécanique du béton d'un ouvrage existant combine des mesures non destructives (CND) comme la mesure du rebond au marteau d'impact (R) et/ou la mesure de la vitesse de propagation des ondes ultrasonores (V) avec des mesures destructives (sur carottes) de manière à développer une relation, nommée "modèle de conversion", entre la résistance mécanique (f_c) et les mesures non destructives. Le modèle de conversion est utilisé pour estimer ensuite la valeur locale de la résistance en chaque emplacement de test à partir de la valeur de la propriété ND qui a été mesurée. On peut alors calculer la résistance moyenne estimée et/ou l'écart-type de la résistance estimée (variabilité de la résistance du béton). Toutefois, la fiabilité de ces estimations reste une question ouverte. Améliorer cette fiabilité impose de réduire les incertitudes, de préciser et de contrôler les facteurs influents. L'objectif de cette thèse est d'analyser en détail les différentes étapes de la méthodologie d'évaluation afin de proposer des recommandations pratiques pour améliorer la fiabilité de l'estimation sur site de la résistance mécanique du béton, en combinant carottes et mesures non destructives.

Méthodes et Outils

Les analyses qui seront conduites dans ce travail s'appuient sur une logique et des outils que l'on retrouvera tout au long du mémoire.

Dans un premier temps, on définira les éléments principaux de la stratégie d'évaluation dans laquelle on distinguera deux étapes : le programme d'investigation et l'estimation de la résistance.

Le programme d'investigation est conditionné par les ressources disponibles (budget, délais) et couvre :

- (a) la sélection de la technique CND utilisée (mesure de rebond, mesure de vitesse ou combinaison des deux),
- (b) le choix du nombre de mesures de chaque type, non destructives et destructives, et du nombre de répétitions pour les essais ND (c'est à dire du nombre de mesures que l'on fait pour en déduire un résultat d'essai),
- (c) la manière dont on choisit les emplacements de tests, soit de manière prédéfinie ou arbitrairement pour les mesures ND et soit de manière prédéfinie ou arbitrairement ou sur la base des résultats des mesures ND pour les carottes (on parle alors de "carottage conditionnel").

L'estimation de la résistance a trait à l'analyse des résultats d'essais et à leur traitement pour en déduire et utiliser un modèle de conversion. Elle passe par :

- (a) le choix de la forme de modèle de conversion (les modèles explicites peuvent avoir une forme linéaire, polynomiale, exponentielle...) et de l'approche d'identification de ses paramètres (on peut utiliser un modèle a priori sans calibration, ou bien utiliser une approche de calibration, ou bien encore identifier un modèle directement par l'approche de régression),
- (b) l'utilisation du modèle retenu pour estimer la résistance mécanique locale aux emplacements où l'on ne dispose que des résultats des essais ND. Ces valeurs estimées peuvent enfin être utilisées pour estimer la résistance moyenne et la variabilité de la résistance.

Le travail de recherche reposera concrètement sur l'analyse de ressources, sous la forme de séries de données (résultats d'essais ND, résistance mécanique sur carottes). Trois familles de séries de données seront exploitées au fil de ce travail : des données issues de travaux de laboratoire, des données issues de structures réelles, des données synthétiques.

Dans les deux premiers cas, les données seront issues de la littérature existante ou de partenariats avec d'autres chercheurs. Nous exploiterons au total 39 jeux de données réelles qui couvrent une large gamme de bétons. Ces données correspondent à 3922 emplacements de tests, où les informations disponibles peuvent se présenter sous la forme d'une mesure ND et de la résistance ((V, f_c) , ou (R, f_c)) ou de deux mesures ND et de la résistance (V, R, f_c) . Sur l'ensemble de ces emplacements de tests, 358 sont issues de mesures sur ouvrages. Une collaboration avec le Pr. S. Kenai et K. Ali-Benyahia (Algérie) nous a donné accès aux résultats d'un programme d'investigation systématique conduit sur un bâtiment avant leur destruction. Une collaboration avec l'Université de Budapest (Pr. A. Borosnyói, K. Szilágyi) nous a donné accès à plus de 2900 résultats d'essais (R, f_c) issus de différents travaux de laboratoire. Ces jeux de données montrent une grande diversité des caractéristiques du béton (composition, âge, conditions de cure...) et des conditions de mesure. Tous ces jeux contiennent aussi bien les valeurs des résultats d'essai (valeur du rebond) que celle des mesures individuelles (il faut 10 répétitions de la mesure du rebond pour déduire un résultat d'essai). Cette information est essentielle pour étudier la qualité des mesures (répétabilité des mesures).

Tous les jeux de données existants, aussi exhaustifs soient-ils, ont des limites, ne serait-ce que par l'impossibilité qu'il y a à étudier séparément chacun des facteurs pouvant avoir une influence sur la qualité finale de l'évaluation. C'est la raison pour laquelle nous avons aussi travaillé avec des données synthétiques. Nous avons développé une procédure de simulation, sur la base des propositions de (Breysse D. , 2012; Breysse & Martínez-Fernández, 2014). Les simulations synthétiques offrent de multiples avantages:

- (a) la capacité de modifier chacun des facteurs influents (c'est à dire chacun des degrés de liberté cités ci-dessus en définissant la stratégie d'évaluation) pour étudier comment la qualité de l'évaluation de la résistance est affectée,
- (b) la possibilité de générer un nombre illimité de mesures (résultats ND, résistance) alors que les limites budgétaires et opérationnelles rendent ceci impossible sur ouvrage réel,
- (c) la possibilité de quantifier l'écart entre la résistance estimée à l'issue de la procédure et la résistance simulée ("vraie" pour les données synthétiques), ce qui est bien entendu impossible avec des données réelles.

L'ensemble de ces avantages confère une grande souplesse à l'emploi de données synthétiques, et les conclusions tirées des données synthétiques seront confrontés, autant que possible, à des résultats obtenus sur des jeux de données réelles.

La procédure de simulation s'appuie sur un simulateur développé en langage VBA (Visual Basic for Applications) sous Excel. Le simulateur a deux fonctions principales : la génération des données synthétiques, et la simulation et l'analyse des stratégies d'évaluation. Cette seconde fonction sera utilisée pour les jeux de données synthétiques et réelles.

Analyse de la méthodologie d'évaluation actuelle

Cette analyse est divisée en trois parties. Dans un premier temps, on étudie l'effet de plusieurs facteurs influents sur la qualité de l'évaluation. Ces facteurs sont :

- le nombre NC d'emplacements où l'on prélève des carottes qui seront utilisées pour identifier le modèle de conversion,
- la qualité des mesures ND, via leur répétabilité,
- la variabilité en place de la résistance du béton,
- la forme mathématique du modèle de conversion,
- le choix d'une estimation sur la base d'un seul type de mesures ND ou de mesures combinées.

Pour analyser ces effets, les différents types de données (réelles de labo ou sur ouvrage, synthétiques) seront considérés. Pour quantifier la qualité du modèle de conversion et son aptitude à évaluer la résistance, on calculera les coefficients de détermination r^2 et l'erreur quadratique moyenne $RMSE$. Ces valeurs seront calculées de deux manières : en considérant uniquement les emplacements de tests avec des carottes servant à identifier le modèle (erreur d'ajustement) et en considérant toutes les emplacements de tests avec des mesures ND mais sans carottes (erreur de prédiction).

Les résultats de ces analyses montrent l'effet majeur du nombre d'emplacements de tests pour les carottes utilisées pour identifier le modèle, NC , sur l'erreur. Pour de petites valeurs de NC , l'erreur d'ajustement est faible (voire nulle) mais l'erreur de prédiction est forte. Cependant, quand NC augmente, l'erreur de prédiction se réduit progressivement, même si elle demeure toujours strictement supérieure à l'erreur d'ajustement. On en déduit que seule l'erreur de prédiction doit être utilisée comme une mesure représentative de la qualité de l'évaluation.

La variabilité in-situ de la résistance (via l'écart-type vrai $s(f_{ist})$) joue aussi un rôle : augmenter $s(f_{ist})$ conduit à une augmentation de $RMSE$, mais à un taux tel que l'erreur relative $RMSE/s(f_{ist})$ diminue. Cela conduit à être vigilant quant à la manière d'exploiter les résultats, puisque l'on peut parvenir à des conclusions contradictoires selon que l'on considère une erreur absolue ou une erreur relative.

L'analyse montre aussi l'impact significatif de la qualité des mesures sur la qualité de l'évaluation finale, alors que le choix de la forme du modèle de conversion a peu d'effet : il est limité aux faibles valeurs de NC , pour lesquelles un modèle linéaire réduit l'erreur. Le dernier facteur étudié est l'efficacité de combiner deux techniques ND (mesure de vitesse et mesure de rebond) avec les mesures de résistance sur carottes. Les résultats montrent que la combinaison n'est pas toujours efficace, et que son efficacité dépend du NC et de la qualité des mesures.

La seconde partie de cette étude repose sur l'analyse de différentes stratégies qui ont été définies et employées par des experts dans le cadre d'un benchmark international mené au sein du TC ISC 249 de la RILEM. L'objectif du benchmark est de comprendre quels sont les facteurs qui déterminent l'efficacité de ces stratégies et de quantifier cette efficacité pour l'estimation de la résistance moyenne aussi bien que pour celle de la variabilité. Cette efficacité est quantifiée en répétant des séries de 1000 simulations de structures synthétiques possédant les mêmes propriétés statistiques et en définissant des indicateurs pertinents pour l'efficacité des estimations.

L'analyse des stratégies illustre que les valeurs estimées montrent une dispersion significative au fil des simulations, qui ne peut être négligée : répéter une même stratégie sur un bâtiment différent, mais possédant les mêmes caractéristiques d'un point de vue statistique peut conduire à des estimations qui s'écartent plus ou moins des valeurs vraies. L'intérêt des simulations synthétiques est de révéler cette propriété, et d'en permettre la quantification. On montre en particulier qu'aucune des stratégies ne permet d'estimer de manière fiable la

variabilité du béton (écart-type de la résistance). Après avoir étudié l'effet du mode de sélection des emplacements pour les carottes, une méthode a été proposée pour sélectionner ces emplacements en fonction des mesures CND "sélection conditionnelle" qui améliore la qualité de l'évaluation sans coût supplémentaire.

La troisième partie s'intéresse aux approches existantes pour identifier le modèle de conversion : identification par régression ou par calibration d'un modèle a priori, avec deux variantes, par translation (méthode- Δ) ou affinité (méthode-k)). Ces approches sont appliquées à huit bases de données réelles (quatre bases avec des mesures de rebond et quatre bases avec des mesures de vitesse) pour estimer la résistance moyenne et l'écart-type des résistances. Les résultats obtenus montrent que toutes les approches (régression ou calibration) peuvent estimer de manière efficace la contrainte moyenne. Dans tous les cas, augmenter NC améliore la précision de l'estimation. Cependant ces approches sont toutes limitées ou défaillantes pour l'estimation de la variabilité, que l'on dispose de mesures de rebond ou de vitesse. Cette conclusion revêt un caractère général et impose une approche originale.

Développement d'une nouvelle approche d'identification de modèle: approche bi-objectif

L'inaptitude des approches existantes d'identification à estimer la variabilité du béton a été confirmée. Nous proposons dans cette partie une procédure innovante, nommée "approche bi-objectif" qui permet d'estimer à partir des mêmes données et de manière simultanée la résistance moyenne et son écart-type. L'idée de départ repose sur le constat que tout programme d'investigation combinant des mesures non destructives et des carottes fournit un jeu de données de NC paires (x, f_c) , où x et f_c sont respectivement les valeurs des résultats d'essais et des résistances mécaniques au même emplacement de test. Ce jeu de données est employé pour identifier les paramètres du modèle de conversion entre les valeurs de mesures ND et les résistances en compression. Les formes usuelles des expressions mathématiques des modèles (linéaire, exponentiel, polynomial du second degré) ont deux paramètres. Analytiquement, le respect de deux conditions explicites (égalité entre valeurs mesurées et valeurs dérivées du modèle de conversion de la résistance moyenne ET de l'écart-type des résistances) permet d'identifier les deux paramètres du modèle de conversion. Ainsi, d'un point de vue mathématique, la méthode bi-objectif peut être utilisée avec n'importe quelle technique ND pour identifier les paramètres d'un modèle de conversion à deux paramètres. Cette méthode innovante d'identification a été mise en œuvre dans le cas d'un modèle linéaire et dans le cas d'un modèle puissance.

La validité de cette approche est confirmée par son application à plusieurs jeux de données pour estimer la résistance moyenne et son écart-type. Les résultats obtenus montrent que l'approche bi-objectif est aussi efficace que les autres approches pour identifier la résistance moyenne et qu'en outre, elle permet d'identifier la variabilité du béton, et ce d'une manière d'autant plus fiable que NC augmente. Cette méthode peut être employée aussi bien avec des mesures de rebond que de vitesse, et pour des modèles linéaires aussi bien que pour des modèles puissance, quoique ces derniers génèrent plus d'incertitudes si NC est faible.

Analyse de la qualité de l'évaluation

Dans ce qui précède, les effets des facteurs influençant le plus la qualité de l'estimation de la résistance ont été confirmés. L'objectif est dans cette partie d'avancer vers une quantification de ces effets, afin de parvenir à établir des liens explicites entre la précision des estimations et ces facteurs : nombre d'emplacements de carottage NC , qualité des mesures ND (répétabilité des mesures), manière de sélectionner les emplacements de carottage, type de technique ND,

utilisation d'une seule technique ou combinaison de plusieurs, approche d'identification du modèle de conversion. Dans ce but, une vaste base de données synthétiques a été générée. Elle contient 36000 triplets (V, R, f_c) et couvre une large gamme de bétons en termes de résistance moyenne et de variabilité. Les résultats obtenus sur cette base synthétique ont ensuite été confirmés par une étude, plus limitée, sur une base de données réelles contenant environ 2500 paires (R, f_c) .

Le concept de précision de l'estimation a été introduit, qui repose sur la définition d'une marge d'incertitude (absolue ou relative) sur chacune des deux cibles : résistance moyenne et écart-type de la résistance. L'ensemble des facteurs influents étant défini, les simulations de Monte-Carlo permettent de quantifier dans quelle mesure la valeur estimée se trouve (ou pas) dans l'intervalle défini par la marge d'incertitude et, en conséquence, le risque d'erreur, c'est à dire d'estimation hors tolérance. On peut ainsi tracer des courbes quantifiant le risque d'erreur, pour une tolérance acceptée prédéfinie, correspondant à chaque situation spécifique, et en faisant varier le NC . Ces courbes sont qualifiées de "courbes de risque".

L'utilisation des courbes de risque permet de déduire le NC minimum qu'il est nécessaire d'avoir pour une précision requise de l'estimation, avec un degré de confiance donné. On a pu ainsi vérifier que la valeur minimale de NC ne peut pas se réduire à une prescription unique, comme l'indiquent les codes (ACI 228.1R, 2003; EN 13791, 2007), et que ce nombre dépend du contexte. Nous pensons que ce constat implique une révision sérieuse des recommandations, qui conduira à indiquer un nombre minimal (via des formules ou des tables) qui dépende de la situation : qualité des mesures, variabilité du béton, méthode d'identification du modèle... Nous montrons par exemple que pour une tolérance relative donnée sur les valeurs à estimer ($\pm U\%$) la variabilité du béton joue un rôle négatif pour estimer la résistance moyenne, mais un rôle positif pour estimer cette variabilité.

Recommandations pour de meilleures pratiques

A partir de l'ensemble de ces travaux et résultats, nous proposons quelques recommandations pour améliorer la pratique des ingénieurs:

- a. Une attention particulière doit être portée à la qualité des mesures (répétabilité de l'essai) sur site et autant que possible à son amélioration, vu sa très forte influence sur la qualité de l'évaluation de la résistance.
- b. Le nombre approprié d'emplacements de carottage NC peut être estimé en fonction de : la marge d'incertitude acceptée sur les valeurs estimées (et du niveau de confiance attaché à cet intervalle), la variabilité in-situ du béton et la qualité des mesures. En l'absence actuelle de prescriptions par les standards, le Tableau 6.2 peut servir de première base pour quantifier ce nombre.
- c. La méthode de sélection conditionnelle des emplacements de prélèvement des carottes, sur la base des résultats des mesures ND préalables, est recommandée dans la mesure où elle améliore la qualité des estimations sans induire de coût.
- d. A contrario, le choix d'une forme particulière de l'expression mathématique du modèle d'inversion a peu d'importance sur la qualité des estimations.
- e. Le recours à l'emploi de techniques combinées (vitesse ultrasonique / rebond) doit être limité au cas d'un nombre suffisant de carottes et quand les deux techniques ont une qualité de mesure proche.
- f. Pour l'estimation de la résistance moyenne, toutes les approches (régression, calibration, bi-objectif) peuvent être employées pour identifier le modèle de

conversion. Si l'on utilise une approche de calibration, il faut cependant veiller à ce que le modèle a priori ait été établi sur un béton comparable et pour les mêmes gammes de valeurs que celui en cours d'étude.

- g. La procédure bi-objectif est la seule qui puisse être employée pour estimer la variabilité du béton.
- h. Pour estimer la qualité des estimations de la résistance, l'erreur quadratique moyenne (*RMSE*) peut être employée. Elle doit être calculée à partir de valeurs qui n'ont pas été exploitées en phase d'identification du modèle de conversion. Si l'on souhaite comparer des modèles employés sur différents bétons, il convient d'utiliser l'erreur quadratique moyenne relative ($RMSE/s(f_c)$).

CHAPTER ONE

INTRODUCTION

1.1 Problem statement

Concrete is a widely used composite material in the field of constructions due to its strength and durability, the availability and cheapness of the raw materials, and the ability to be formed in the desired architectural shape. Consequently, nowadays, there are a huge number of existing concrete structures and day by day this number is in evolution. Among the concrete mechanical and physical properties, the concrete compressive strength is the most important property because it is essential for designing a structural member or calculating its load bearing capacity. The compressive strength is not a fixed value during the concrete life due to several effects like curing at the earlier ages and internal cracks developed (due to environmental or loading effects) at latter ages. Therefore, assessing the in-situ concrete compressive strength is required in many situations such as:

- the quality control of concrete to ensure the compliance with specification for the case of new construction,
- contractual disputes after the non-compliance of standard specimens,
- when there is a doubt about the quality of concrete in a new construction,
- the change in the use (or function) of a structure requires the assessment of the concrete strength in order to accurately calculate the structural capacity,
- damage resulting from fire, fatigue, overload, or environmental degradation,
- seismic retrofitting issue arising nowadays in several countries (like Italy and Turkey) also emphasizes the need for an accurate in-situ assessment of concrete strength in existing structures (Biondi & Candigliota, 2008; Gunes, 2015).

It is clear that the first three situations concern the new structures while the last three belong to in-situ assessment in existing structures. Many test methods are available to estimate the in-situ strength. The core test is the most direct method used for this purpose. However, it has many drawbacks since it is expensive, time consuming, sometimes limited due to difficult access of coring machine, only representative of small volume of concrete and it has locally destructive effect on the structure (Meynink & Samarin, 1979).

Other testing methods that are used to assess the in-situ concrete strength are the nondestructive methods, such as: rebound hammer, ultrasonic pulse velocity, probe penetration, pullout, and break off. The main features of these methods (as compared with the core test) are their simplicity, rapidness and low cost. However, their main drawback is that they are indirect methods since they cannot measure the in-situ strength directly. They measure other properties and strength is derived from a correlation “conversion model” between the strength and measured properties. Due to the fact that concrete is a largely heterogeneous material, a unique model for all concretes does not exist (Breysse, 2012).

Therefore, the current methodology is based on establishing a conversion model for each particular case using both destructive and nondestructive tests. Then, the derived model is used to estimate the concrete strength at the test locations where the nondestructive measured

values are available for the concrete under investigation. This methodology has many degrees of freedom:

- How many test locations are required for cores?
- Which nondestructive method or methods?
- How many test locations are required for nondestructive method?
- What is the way for selecting these test locations?
- What is the mathematical type (or form) of the conversion model?
- Which model identification approach?
- Using single or combination of nondestructive methods?
- What is the quality of assessment of the predicted strength value?
- How can the quality of assessment vary when any answer to the above questions changes?

Standards (ACI 228.1R, 2003; EN 13791, 2007) try to provide answers to some of these questions, however, they fail to supply a general methodology which provides the answers to all above questions.

1.2 Scope of thesis

The general problem presented in the above section cannot be the focus of a single PhD thesis, therefore several refinements are considered. Firstly, this research deals with the **assessment of in-situ strength in existing structures** because it is more essential nowadays with the arising needs for the seismic retrofitting. For example in Turkey, there are several millions of buildings that were classified as risky buildings and consequently reliable information about their strength is required for any decision about the retrofitting or renewing (Gunes, 2015). To distinguish between the new and existing structures, the following definition is adopted herein for the existing structures, in agreement with the RILEM TC 249 ISC guidelines: the existing reinforced concrete structures are defined as structures where age effects (crack, concrete delamination due to reinforcement corrosion, etc.) and reinforcing steel bars have a great influence on the NDT results and the predicted strength. Another feature of existing structures is the common lacks in the detailed information about the concrete used and the missing of companion specimens which could be used for comparison.

The actual in-situ strength at any test location in a concrete structure is always unknown, therefore in the present work **the reference value for this unknown strength will be the core strength at this test location.**

As it was stated in the previous section, there are several nondestructive techniques that can be used to assess the concrete compressive strength. However, in this PhD work, **only the rebound hammer and ultrasonic pulse velocity methods are considered** because they are the widely used methods on existing reinforced concrete structures.

In the real practice, the mean compressive strength and the characteristic compressive strength are the most common assessed values. The latter being an important input for structural computation. The assessment of characteristic strength depends on the mean strength and on the standard deviation of the compressive strengths (concrete strength variability), thus the concrete variability is also a required value. The (ACI 214.4R, 2003) reported that the coefficient of variation, $CV(f_c)$, due to in-situ concrete strength variation within a structure (i.e. concrete variability/mean strength) is 13%. However (Masi & Chiauzzi, 2013) found a $CV(f_c)$ value of 21% within one member of a structure. (Masi & Vona, 2009) studied the concrete variability in many buildings in Italy and they observed that the probable values of $CV(f_c)$ range between 15-35%. (Pucinotti, 2013) also stated that in many cases the $CV(f_c)$ reaches 35%. That is why the assessment of concrete variability is needed in addition to mean

strength value. The present PhD work focuses on the estimation of **local strength values, mean strength and concrete strength variability**. The characteristic strength can be derived once these properties are known, but it is out of our focus.

When the concrete strength variability is high or when available information indicates that the concrete in whole structure (or the parts under consideration) having different composition or quality, standards (ACI 228.1R, 2003; EN 13791, 2007) recommend to divide the investigated structure into several test regions. To this end, researchers (Giannini, et al., 2014; Masi, et al., 2016) have been proposed criteria for dividing the investigated structure into concrete homogenous zones (test regions). In the present work, **each dataset that will be used in the analysis is considered to be belonged to one test region** i.e. the subdivision issue is out of our scope.

As stated in Section 1.1, in the current methodology, a conversion model between concrete strength and nondestructive measurements is identified. In the common form of this model, the concrete strength represents the dependent (response) variable while the nondestructive measured properties represent the independent variables (inputs). However, a possibility promoted by some researchers, see for examples (Kheder, 1999; Atici, 2011; Szilágyi, et al., 2011; Martínez-Molina, et al., 2014), is to use (in addition to nondestructive measurements) several concrete characteristics (water-to-cement ratio, aggregate-to-cement ratio, admixture content, concrete density, age, etc.) as independent variables in the model. However the main drawback of this type of models is the need to know the concrete characteristics as inputs into the model while they usually remain unknowns in old structures. Therefore, in this study, **models with only nondestructive measurements as the inputs will be considered**.

1.3 Objective and general research methodology

This thesis aims to study the current assessment methodology in order to provide practical recommendations that can improve the reliability of assessment of in-situ strength in existing concrete structures by nondestructive tests and cores.

In order to achieve this objective, the present thesis will follow the research methodology shown in Figure 1.1. After synthesizing the problem and its influencing factors, we will define the main elements of the assessment strategy. Then the tools required for the analysis through this thesis will be provided. The first tool is the construction of a simulator using VBA code. This simulator has two objectives: the generation of synthetic data, and the simulation of the assessment strategy, that can be repeated many times. The second tool is regarding the datasets that will be used in the analysis. Many datasets will be made available; they were collected from collaboration with other researches and from the scientific literature. The details will be presented in Chapter 3.

The study in this thesis can be classified into three main axes. The first axis (Chapter 4) will be the analysis of the existing assessment methodology. This analysis will be subdivided into three main parts: the analysis of the most influencing factors of the assessment methodology, the analysis and comparison of several assessment strategies from an international benchmark carried out by RILEM TC 249 ISC, and the analysis of the existing model identification approaches.

Our work will also cover original contributions in the second and third axes. We will first develop a new model identification approach to capture both the mean strength and concrete strength variability in the second axis (Chapter 5). Then, in the third axis, the quality of assessment will be analyzed in relation with the main factors of the assessment methodology (Chapter 6). Finally we will provide practical recommendations for engineers (Chapter 6).

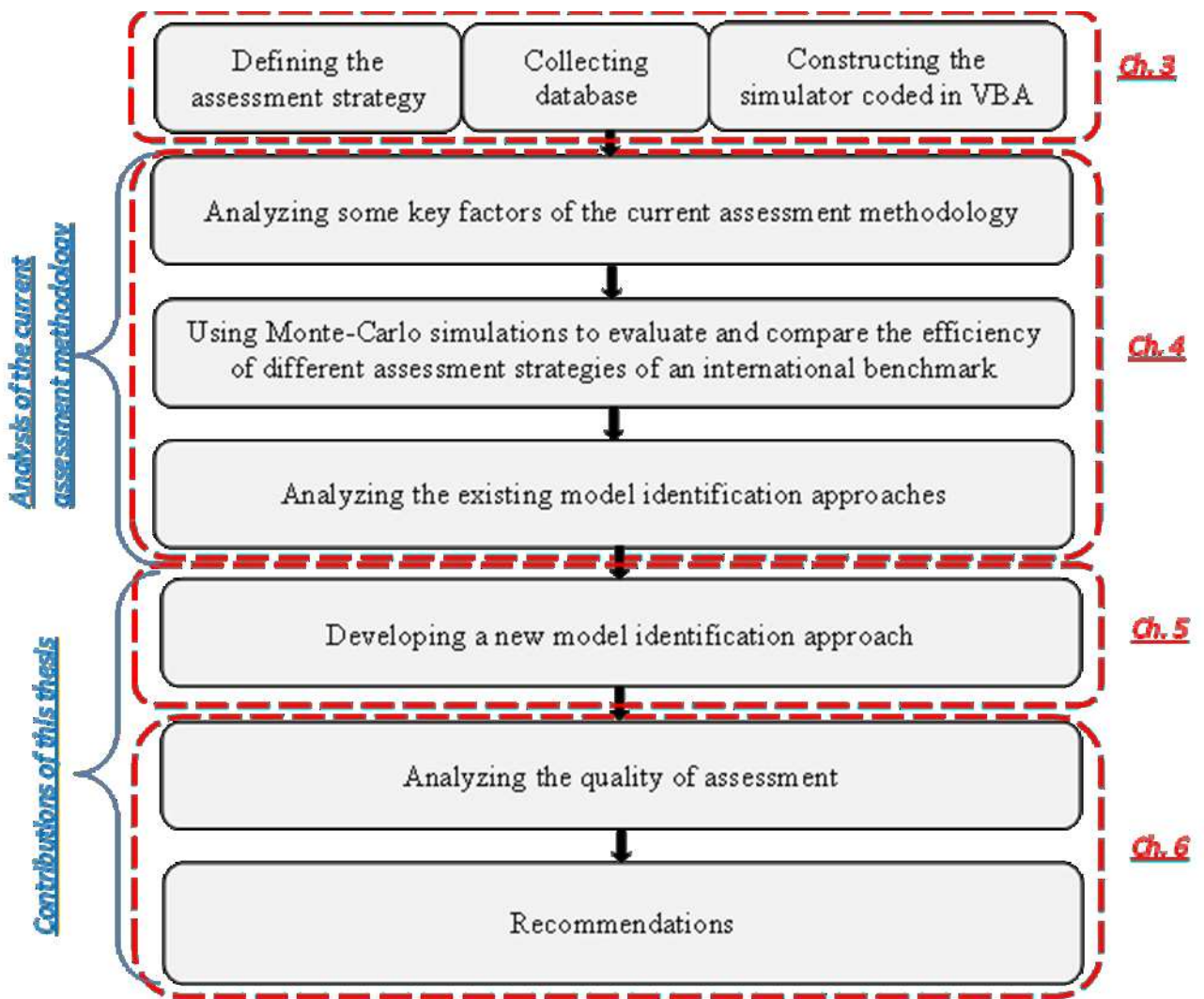


Figure 1.1 General research methodology

CHAPTER TWO

CONCRETE STRENGTH ASSESSMENT IN THE EXISTING STRUCTURES: LITERATURE REVIEW

2.1 Introduction

The diagnosis of existing structures is an important issue in order to ensure the proper safety and serviceability conditions during their life. Depending on the objective of the diagnosis, the in-situ assessment of reinforced concrete existing structure may include one or more activities, such as: checking the concrete homogeneity, checking the steel reinforcement corrosion, detecting the cracks and measuring their depth, detecting the concrete deterioration (due to environmental effects, chemical attack, fire damage, fatigue and overloading), detecting the surface carbonation and measuring its depth, estimating the elastic and mechanical properties of concrete (modulus of elasticity, flexural strength and the compressive strength).

Among these activities, the assessment of compressive strength of concrete is our concern in this study since it is essential for the structural computation. To achieve this target, different destructive and nondestructive methods are applied. Core test (as destructive method) and rebound hammer and ultrasonic pulse velocity (as nondestructive methods) are considered herein due to their widespread use in real practice.

The assessment of concrete compressive strength includes the estimation of one or more of the following properties: local strengths, mean strength, concrete strength variability and the characteristic strength. In the present work, the first three properties are considered while the characteristic strength (which can calculate if the mean and concrete variability values are knowns) will be out of our focus.

This chapter presents the current methodologies in the real practice for assessing the in-situ strength in existing structures with their usefulness, limitations and standards requirements. It is subdivided into four main sections, where each section is dedicated to a specific option for the assessment of in-situ strength: using cores only, using nondestructive methods only, using cores with a single nondestructive method, and using cores with a combination of nondestructive methods.

2.2 Strength assessment using cores only

Core test is the most direct technique to assess the concrete compressive strength in existing structures. Many standards and guidelines are available that guide engineers and investigators in real practice. The widely used ones are: (EN 12504-1, 2000; EN 13791, 2007) in Europe and (ACI 214.4R, 2003; ASTM C42, 2012) in the USA.

The assessment methodology for assessing the concrete strength using cores can be summarized in the following main steps:

- Planning an investigation program

- Drilling cores
- Testing cores
- Interpreting the core strengths

In this section, each of these steps is discussed before drawing some conclusions.

2.2.1 Planning an investigation program: number and location of cores

Planning a testing program includes specifying the number of cores and selecting their locations in the structure. In order to select a suitable number of cores it is necessary to bear in mind two questions: what is the size of the test region covered by the investigation (one member, several members in one floor, one floor or the whole building)? and what is the accepted uncertainty level associated with the measured strength? i.e. what is the reliability of the final results?

Standard (EN 12504-1, 2000) tells nothing about the number of cores and it focuses on the process of taking cores, examining them and testing cores in compression. Standard (EN 13791, 2007) requires at least 15 cores (with minimum nominal diameter 100 mm) for strength assessment using an approach (named approach A in this standard, this approach is devoted to assess the characteristics in-situ strength using cores only). Nevertheless, it accepts using 3 to 14 cores when applying another approach (named approach B which is also provided in this standard in order to assess the characteristics in-situ strength using cores only). However, (EN 13791, 2007) cautions that the reliability of approach B is less than that of approach A due to the uncertainty associated with small number of cores. In any case for each test region these numbers of cores should be multiplied by three when using 50 mm cores.

(ACI 214.4R, 2003) is more detailed regarding this issue since it provides a formula for calculating the number of cores corresponding to a specific concrete strength variability and an accepted predetermined error between the estimated mean strength and population mean strength as given in Equation (2.1) (for 95% confidence):

$$NC = [2 CV(f_c)/e]^2 \quad (2.1)$$

NC : the recommended number of cores,

e : the predetermined maximum error expressed as a percentage of the population mean, and

$CV(f_c)$: the concrete strength variability (in terms of coefficient of variation) for the whole population, in percent.

In Figure 2.1, the error curves as a function of number of cores for different values of $CV(f_c)$ were plotted. These curves show that, for a specific value of error, the concrete having a larger variability requires more cores for estimating the mean strength. Using more cores leads to more accurate assessment however the cost and destructive effect are the main constraints against the excessive use of cores.

The selection of core locations is governed by the purpose of the in-situ strength assessment. For example, if the purpose is the assessment of load bearing capacity of a structure then the test locations are concentrated on the most stressed parts of the structure (coring in these elements should be selected with care to avoid weakening them). However if the purpose is the assessment of damaged area, the test locations are selected within this area and the results may be compared with those from a comparable undamaged area. Generally, the parts of the structure considered by the investigation program should be divided into test regions (populations) and within each test region the test locations should be selected randomly (EN 13791, 2007).

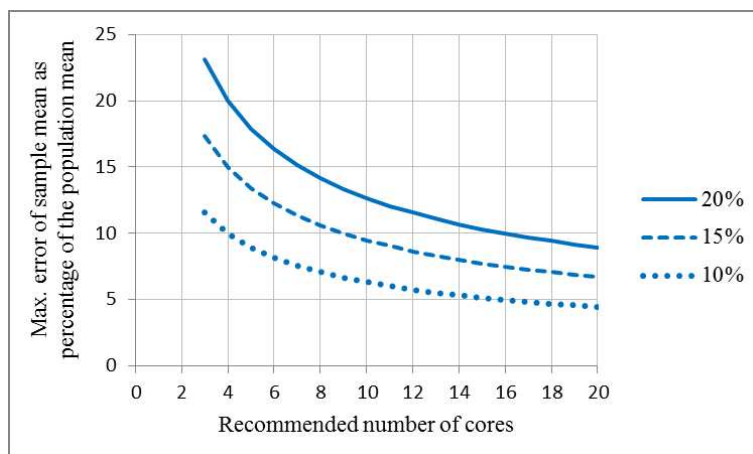


Figure 2.1 Error of the estimated mean strength as a function of number of cores, (ACI 214.4R, 2003)

Moreover, when selecting a test location within a structural member, several factors should be considered due to their effects on the measured compressive strength. These factors are:

- the core orientation with respect to concrete casting, with lower strength obtained from horizontally drilled core,
- the concrete at the bottom of member is generally stronger than concrete at the top,
- the location of steel reinforcement in the member which needs to be avoided (ASTM C42, 2012). It is also necessary to avoid coring near edges or joints of the structural member.

2.2.2 Drilling cores

Before drilling a core it is necessary to choose its dimensions (diameter and length) and as a result the preferable length-to-diameter ratio (L/D). The selection of core diameter is governed by several effects. The most important one is the core diameter-to-maximum aggregate size ratio (D/A). The uncertainty increases as D/A decreases. (EN 12504-1, 2000) requires the core diameter-to-maximum aggregate size ratio to be equal or more than 3.0 while (ASTM C42, 2012) accepts a value of 2.0 or more for this ratio. The second effect is the member dimensions which can control the core diameter when the member has small dimensions. The spacing between the steel reinforcement can also lead to lower core diameter. In many countries a minimum diameter of 100 mm is used, with 150 mm preferred, although in Australia 75 mm is considered to be generally acceptable (Bungey, et al., 2006).

Core length is controlled by the L/D ratio and core diameter. L/D ratio should be between 1.0 and 2.0 (Neville & Brooks, 2010). Standards (EN 12504-1, 2000) and (ASTM C42, 2012) prefer L/D ratio of 1.0 when the resulted strength is to be compared with cube strength and 2.0 for the comparison with cylinder strength. For cores having L/D ratio less than 2 and the resulted strength is to be compared with cylinder strength, a suitable correction factor should be applied. L/D ratio less than 1.0 is not allowed to be used for compressive test whilst the ratio more than 2.0 can be reduced to 2.0 by trimming the core length.

Core specimen is extracted from the structural member using a cutting tool having diamond bits attached to the core barrel, see a core drilling equipment in Figure 2.2. Since any movement during drilling may lead to damaged core so the rig should be firmly anchored to the concrete member. Moreover it should be perpendicular to the surface from which the core is taken off in order to avoid cutting a distorted core. The drill bit is usually lubricated or cooled with water.

(EN 12504-1, 2000) and (ASTM C42, 2012) require that the core specimens used for determining compressive strength shall not contain embedded reinforcement whenever possible. Whilst (BS 1881: Part 120, 1983) and the Concrete Society (Concrete Society Technical Report No.11, 1987) suggest a correction factor to be applied to core strength in order to account for the presence of bar reinforcement perpendicular to the axis of the core. Cores containing bar reinforcement in or close to the longitudinal axis of core are not allowed for strength assessment (EN 12504-1, 2000; ASTM C42, 2012). Since the correction factor is derived from an empirical formula so applying this factor to the core strength will add some uncertainty. Consequently it is better to avoid using the correction factor by using nondestructive tools like covermeter or GPR (Ground Penetrating Radar) to locate suitable position among the reinforcements before drilling core. The other possibility to avoid using the correction factor is to drill a core with sufficient length and to remove the portion containing the reinforcement (of course if the remaining portion has accepted length to diameter ratio). Finally, the produced hole in the member after the drilling process should be filled with concrete or epoxy or other suitable fillers.

From above it is obvious that a skilled operator is mandatory in order to obtain a core specimen that is undamaged and representative of the in-situ concrete.



Figure 2.2 Core drilling rig

2.2.3 Testing cores

In the laboratory, cores should be visually examined in order to check their condition and to have information that can assist the interpretation of core strength obtained from compressive testing later. The visual inspection can provide information about: the aggregate size and shape, presence of voids, depth of carbonation, presence of crack and its depth and width, presence of reinforcement, drilling damage or defects and any abnormalities in the specimen (True, 2003). The calculation of approximate core density can also be helpful in the interpretation of the measured core strength. In addition, core testing by measuring the ultrasonic pulse velocity can provide more insight on the concrete uniformity and highlight any internal invisible defects.

Before testing core by compression, each core specimen should be prepared by sawing its ends so as to produce a specimen having suitable length, flat ends and perpendicular to the longitudinal axis of core. When the specimen does not conform to flatness and

perpendicularity requirements of the standards (EN 12390-1, 2000; ASTM C39, 2014), it should be prepared by grinding or capping the ends. Thereafter, core dimensions (diameter and length) are measured so as to use these values for calculating L/D ratio and the cross sectional area of core.

During the drilling and the preparation of specimen, water is used to lubricate the cutting tools. Consequently, this wetting water changes the moisture condition of core (as compared to the original moisture in the structure) and also induces a moisture gradient within the specimen. The measured core strength at saturated condition is lower than that of a comparable dry specimen by 10-15% (Bungey, et al., 2006). Therefore, testing core having moisture condition other than that in the structure will produce core strength divergence from the true in-situ strength. There is no standard method to ensure an identical moisture condition between the core, at compression test time, and the structure. However (ASTM C42, 2012) provides a procedure to minimize the moistening effects of drilling and preparation processes. (EN 12504-1, 2000) requires testing core at saturated condition (according to this standard the saturation is obtained after at least 40 hours of soaking in water $(20 \pm 2)^\circ$ C. (EN 13791, 2007) recommends testing core at a moisture condition depending on that in the structure i.e. testing core in dry condition if the condition in the structure is dry; however where the condition is wet in the structure, core is tested in saturation condition. It is obvious that there is no clear consensus regarding this issue which induces some additional uncertainty on the assessment of true in-situ strength.

Compression test of core is carried out in accordance with standards (EN 12390-3, 2009; ASTM C39, 2014) and the core compressive strength is obtained by dividing the maximum load by the core cross sectional area based on average core diameter.

2.2.4 Interpreting the core strengths

Core compressive strength obtained from compressive test differs from the corresponding in-situ strength in the structure at the location where the core was extracted. This difference is due to the effects of several factors such as: core moisture condition, excess voids percentage in core specimen, length-to-diameter ratio, core diameter, drilling direction relative to concrete casting direction, presence of reinforcement, and drilling effect. Therefore, suitable correction factors can be used in order to obtain the equivalent in-situ strength. Defining these factors was the objective of many research works (Concrete Society Technical Report No.11, 1987; Bartlett & MacGregor, 1994; Ergün & Kürklü, 2012; Masi, et al., 2013; Uva, et al., 2013). The most popular formula for converting the compressive strength measured on core specimen f_{cs} into the equivalent estimated in-situ cube strength is that presented in (BS 1881: Part 120, 1983) and shown in Equation (2.2):

$$\text{estimated in - situ cube strength} = \frac{DD}{1.5+D/L} \times F_{reinf} \times f_{cs} \quad (2.2)$$

where DD represents the drilling direction parameter: DD is 2.5 for cores drilled horizontally, and 2.3 for cores drilled vertically.

F_{reinf} is a correction factor which accounts for the presence of reinforcing bar in the core specimen in the direction perpendicular to the core axis, with $F_{reinf} = 1$ for the case of core free of reinforcement, otherwise $F_{reinf} = 1.0 + 1.5 (\phi_r h)/(D L)$ where: ϕ_r is the bar diameter, h is the distance of bar axis from the nearer end of core, D is core diameter, L is the core length (after end preparation). The term $(\phi_r h)$ is replaced by $(\sum \phi_r h)$ when there are more than one bar. Close bars with spacing less than the diameter of the larger one, only the bar having the larger value of $(\phi_r h)$ should be considered.

(ACI 214.4R, 2003) provides an alternative formula in order to derive the equivalent estimated in-situ cylinder strength from the compressive strength of core specimen f_{cs} , see Equation (2.3):

$$\text{estimated in – situ cylinder strength} = (F_{L/D}F_{dia}F_{mc}F_d) \times f_{cs} \quad (2.3)$$

Where $F_{L/D}$, F_{dia} , F_{mc} and F_d are correction factors for the effects of respectively: length-to-diameter ratio, core diameter, moisture condition of core, and sustained damage due to the drilling process. The values of these factors are given in Table 2.1 as provided by (ACI 214.4R, 2003).

It is necessary to underline here that no one knows the true in-situ concrete strength of a structure and the values provided by Equations (2.3) and (2.4) are only an estimation obtained from multiplying the strength of core specimen by several correction factors.

If the target is the concrete strength value used in the structural computation (i.e. the characteristic strength in the European standards or specified compressive strength f_c' in American standards), then each standard provides methods for getting these values using the estimated in-situ strengths obtained from Equation (2.2) or Equation (2.3).

Table 2.1 Strength correction factors values as provided by ACI 214.4R-03, (f_{cs} in MPa)

Factor	Condition	Mean value
$F_{L/D}$	As-received	$1 - (0.130 - 0.00043f_{cs})(2 - L/D)^2$
	Soaked 48h	$1 - (0.117 - 0.00043f_{cs})(2 - L/D)^2$
	Air dried	$1 - (0.144 - 0.00043f_{cs})(2 - L/D)^2$
F_{dia}	50 mm	1.06
	100 mm	1.00
	150 mm	0.98
F_{mc}	As-received	1.00
	Soaked 48h	1.09
	Air dried	0.96
F_d	-	1.06

2.2.5 Conclusion of the assessment methodology when using cores only

Core test is the most direct method to evaluate the concrete strength in the existing structure. However, it is obvious that this method suffers from many drawbacks:

- a) The number of cores required by the standards is large which leads to high investigation cost,
- b) When planning the core locations in the structures, we are not fully free in selecting these locations due to: structural considerations because drilling excessive cores in the high stressed members can weaken the structure and consequently affect its capacity to bear the applied loads; drilling considerations because the access of large and heavy drilling machine may be difficult (or impossible) in several situations,

- c) Drilling is a complicated process because it includes setting up the machine, checking the perpendicularity and extracting core and for each step the standards provide several requirements to be respected. Consequently, a skill operator is mandatory in order to obtain a core specimen that is undamaged and representative of the in-situ concrete.
- d) Before the compressive testing, it is required to follow several preparation and storage conditions and the shortening in any of these conditions leads to misleading core strength. Therefore this is another reason for the expensive cost of core test.
- e) Core test takes a lot of time as compared with the other tests due to drilling, preparation, testing processes,
- f) The obtained strength of core specimen differs from the true in-situ value existing in the structure at location where the core was extracted.

2.3 Strength assessment using nondestructive tests only

Nondestructive techniques are widely used as indirect methods for assessing the concrete properties (surface hardness, strength, dynamic modulus of elasticity, homogeneity, etc.). Among many existing nondestructive methods, rebound hammer and ultrasonic pulse velocity are the most common ones. Therefore, this section focuses on these two methods. For each method, firstly, the main principle behind the method is illustrated. Next, a brief description of the testing procedure and its limitations is provided. Then, the factors affecting the test results are pointed out. After that, the correlation between the concrete compressive strength with the test result from nondestructive method is discussed. Finally, the main conclusions that summarize the advantages and limitations are presented.

2.3.1 Rebound hammer method

The rebound hammer (or Schmidt hammer) was developed by the Swiss engineer Ernst Schmidt in 1948 as a surface hardness method. The earlier developed surface hardness methods (like Testing Pistol by Williams, Spring Hammer by Frank and Pendulum Hammer by Einbeck) were based on measuring the indentation which resulted from the impact on the concrete surface by a specified mass having specified kinetic energy (Malhotra, 2004). However, the principle of the rebound method mainly differs from that of the indentation type surface hardness methods because it measures the rebound of a specified mass having specified kinetic energy when it impacts the concrete surface. Nowadays, the rebound hammer is the most popular NDT surface hardness technique in the concrete testing practices.

2.3.1.1 Operating principle of rebound hammer instrument and its types

As mentioned above, the instrument is based on the rebound principle that depends on the surface hardness of concrete. Figure 2.3a shows the main components of the instrument: outer body, specified mass or hammer, plunger, spring, rebound indicator and the latching system. Figures 2.3(b, c, d) illustrate how the instrument works. After positioning the plunger perpendicular to the concrete surface at the required test location, the instrument is pushed toward the concrete surface and as a result the mass moves away from the concrete surface producing the spring stretching as shown in Figure 2.3b. When the mass reaches the maximum displacement, it is released to impact the plunger shoulders (Figure 2.3c), however due to the concrete hardness the mass rebounds and it takes with it a slide indicator as shown in Figure 2.3d (Carino, 2008). On an arbitrary scale ranged from 10 to 100, the rebound distance is indicated as a value so-called “rebound number” which represents the rebound distance as a percentage of maximum mass displacement (before impact) (Fischli & Moczko, 2012).

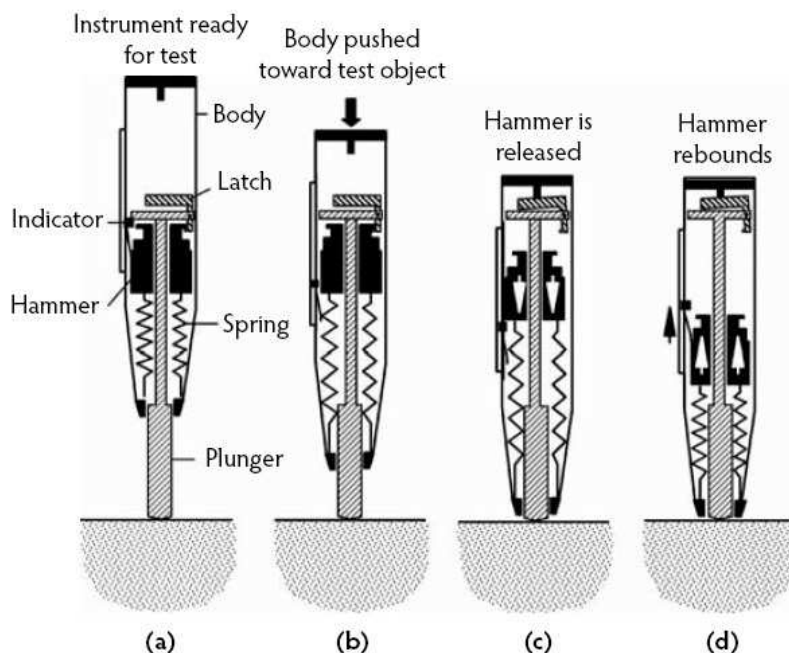


Figure 2.3 Schematic sectional view of rebound hammer illustrating the operation process

Original Schmidt hammer (usually known as type N model) has a weight of 1.7 kg and impact energy of 2.207 Nm and can be used for testing concrete having a compressive strength range 10-70 MPa (Proceq SA, 2016). Currently, several other models of the device are manufactured with different characteristics depending on the purpose for which it is designed. For testing the thin walled concrete elements with a thickness (50-100) mm, type L device is manufactured with impact energy of one-third of type N and less weight (1.4 kg). Another model is type P device, or pendulum type, which is recommended for testing low strength concrete. In 2007, the Silver Schmidt hammer was manufactured. The main feature in this new device is the ability to record, beside the rebound number, the ratio of the kinetic energy of the hammer mass after impact to that before the impact (Szilágyi & Borosnyói, 2009). Through the present report, the considered rebound hammer device is the standard type i.e. type N.

2.3.1.2 Testing procedure and its limitations

Testing hardened concrete using the rebound hammer is a rapid and simple process. Several standards (EN 12504-2, 2012; ASTM C805, 1997) and recommendations reports (RILEM TC 7- NDT, 1977; RILEM TC 43-CND, 1983) were established in order to manage the testing procedure.

Firstly, a suitable test location, usually 300×300 mm (150mm in diameter according to (ASTM C805, 1997)), should be selected. The test locations should be located in the structure on members 100 mm thick or more. If testing smaller thickness member is unavoidable then it must be rigidly supported. Rebound hammer reading is significantly affected by the concrete surface condition and its finishing method (trowelled or formed). Therefore, the tested surface should be clean, smooth and dry with preference to formed finish. Moreover, surface grinding using abrasive stone should be applied to heavily textured, soft, or loose mortar surfaces in order to prepare these surfaces before testing.

Secondly, the test is performed by holding the device firmly and perpendicularly to the concrete surface with gradually pushing it toward the surface until the hammer impacts, then the rebound number reading is recorded. The reading is affected by the surface layer condition

so applying the plunger on subsurface void or a coarse aggregate particle can lead to misleading reading (Carino, 2008). To account for these effects, the test is repeated several times (9 replicates according to (EN 12504-2, 2012) or 10 replicates according to (ASTM C805, 1997)) within the test location with a minimum spacing of 25 mm between each two testing points and a minimum edge distance of 25 mm. The rebound number value that corresponds to a test location, usually so-called test result, is calculated as the average of the readings within this test location (the median of the readings according to EN 12504-2). Abnormal readings should be discarded according to the standards provisions. Another essential value calculated from the readings corresponding to a test location is their standard deviation (or coefficient of variation) which represents the within-test variability. This value is also known as test result precision or repeatability.

It is necessary to check the device just before and after carrying out the tests campaign using a steel anvil provided for this purpose. Also, the ambient temperature during testing should be within the interval (0-50) °C.

2.3.1.3 Factors affecting the test results

As mentioned above, the rebound number depends on the hardness of the surface layer of the concrete member. However, several other factors can also affect the rebound number and its correlation with strength. Some of them are summarized below:

- the effect of gravity: this effect presents when applying the hammer vertically (upward or downward), while it can be avoided if the hammer is applied in the horizontal direction,
- test surface smoothness: as discussed above, smooth and clean test surface being mandatory,
- the effect of concrete moisture condition: dry surface gives higher rebound number as compared to wet surface. The difference in rebound numbers can be up to 10-12 points (Malhotra, 2004),
- the surface carbonation effect: this effect significantly exists in old structures. Several millimeters of carbonation depth (usually no more than 20 mm) can lead to an increase in the rebound number up to 50% (IAEA, 2002),
- the effect of coarse aggregate type: for example the rebound number for siliceous gravel concrete may be significantly higher than that resulted from crushed limestone concrete. Moreover, for the same type of aggregate the difference in the sources can lead to different rebound number/strength correlations (Bungey, et al., 2006),
- the effect of type of cement: high alumina cement can have a compressive strength 100% higher than that resulted from using a correlation curve based on ordinary Portland cement (IAEA, 2002),
- the effect of cement content: the error in the estimation of strength resulting from a change in the cement content is unlikely to exceed $\pm 10\%$ (RILEM TC 7- NDT, 1977),
- member rigidity effect: as discussed above, slender member should be rigidly supported to avoid any effect on the rebound number value,
- other factors may have some effects like the curing and compaction conditions.

2.3.1.4 Correlation between concrete strength and rebound number

There is a little apparent theoretical relationship between the rebound number and the concrete compressive strength (Malhotra, 2004). Therefore, empirical conversion curves (or models) that correlate the concrete strength with the rebound number are widely used as an alternative. Since 1954, Proceq SA has been manufacturing the rebound hammer device that earlier invented by Ernst Schmidt. Proceq SA provides several conversion curves corresponding to different devices (type N, type L, etc.) and specimen shape (cube or cylinder), Figure 2.4 shows the conversion curves corresponding to hammer type N and cylinder specimens (Proceq SA, 2016). As shown, the effect of direction of testing (horizontal or vertical up/downward) is taken into account. Moreover, (Proceq SA, 2016) provides correction factors for the shape and carbonation effects. However, as illustrated above, it may exist other factors that affect significantly the rebound number values. Therefore, using these curves for assessing the concrete strength in structures produces seriously unreliable estimations.

For the last 60 years, a lot of research works have been devoted to build empirical models between the concrete compressive strength and rebound number. After a survey regarding the existing models, more than 60 empirical models were listed by (Szilágyi & Borosnyói, 2009), while more than 80 different models were found by (Breyse, 2012). Several of these models are shown in Figure 2.5. Clearly, these curves are widely scattered, which emphasizes the important effects of the influencing factors mentioned above and indicates that using models for cases other than that from which they were derived can lead to misleading estimations. In other wording, general empirical model correlating the concrete strength and rebound value does not exist and the model calibration for the case study under consideration is a mandatory requirement.

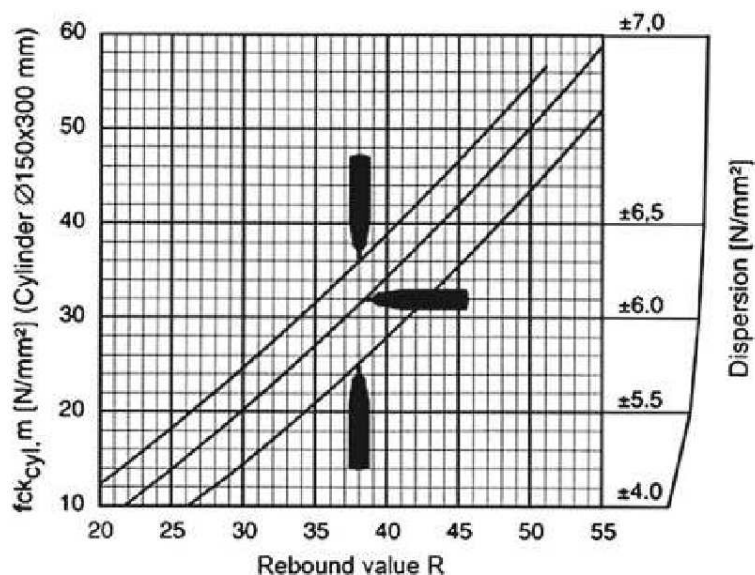


Figure 2.4 Conversion curves for estimating the average compressive strength of a cylinder from rebound number, (Proceq SA, 2016)

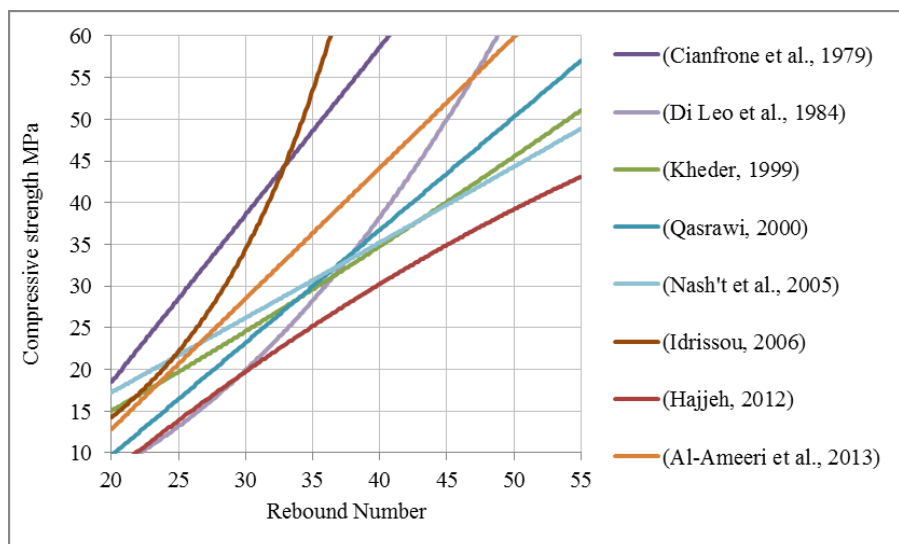


Figure 2.5 Some conversion curves identified by different researchers for estimating the cube compressive strength from rebound number

2.3.1.5 Conclusion of assessing concrete strength using rebound hammer only

Rebound hammer technique is the most simple, rapid and cheapest nondestructive method for assessing the concrete strength both in-situ and in laboratory. Identifying a theoretical correlation between the concrete compressive strength and the rebound value is difficult so alternatively empirical correlation is commonly used. However, the value derived from this test, rebound value, is affected by several mix and member characteristics in addition to the concrete strength. Consequently, using a general model between concrete strength and rebound number for assessing a real structure leads to unreliable estimations and calibration is mandatory. Therefore, rebound hammer technique cannot be used alone in assessing the concrete strength and it should be combined with the destructive tests in order to derive/calibrate a model for the case under consideration.

2.3.2 Ultrasonic pulse velocity method

This technique is one of the stress wave propagation methods that is based on the determination of the longitudinal wave pulse velocity where pulses are generated in concrete by an electro-acoustical transducer. After the Second World War, the earlier devices (Soniscope in Canada and Ultrasonic Concrete Tester in England) were developed (Carino, 1994). Since that, the method and its device have been progressively developed to become a widespread use nondestructive method for testing concrete in-situ as well as in laboratory.

2.3.2.1 Physical principle of the ultrasonic pulse velocity method

The application of an impulse to a large solid elastic media generates propagating stress waves that are classified, according to the particles motion with respect to propagation direction, in three main types: longitudinal waves (P-waves), transverse waves (S-waves) and Rayleigh waves (R-waves) (Naik, et al., 2004). These waves travel at different velocities with the longitudinal waves being the quicker.

The principle of ultrasonic pulse velocity method is to introduce generated pulses into concrete using transmitting transducer which is held on the concrete surface and to measure the arrival time, transit time t , of the first waves (longitudinal waves) to a receiving transducer. A schematic diagram of the device is shown in Figure 2.6. After measuring the distance between the two transducers, path length l , the ultrasonic pulse velocity of longitudinal waves, V_p , can be simply calculated using Equation (2.4):

$$V_p = l/t \quad (2.4)$$

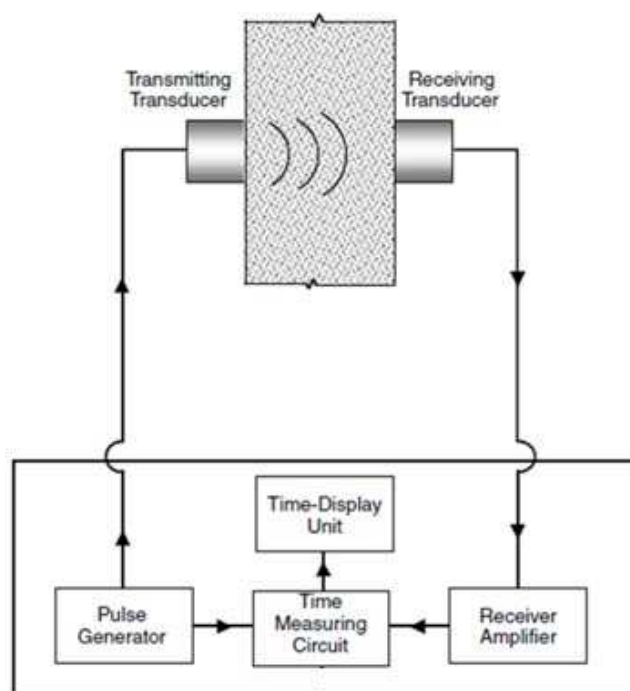


Figure 2.6 Ultrasonic pulse velocity device (schematic diagram), (ASTM C597, 2002)

2.3.2.2 Testing procedure

Many countries have standardized procedures for testing concrete using ultrasonic pulse velocity. (Komlos, et al., 1996) listed more than 20 standards published between 1972 and 1994. The most popular standards are the European standard (EN 12504-4, 2004) and the American standard (ASTM C597, 2002).

The testing procedure should start by verifying that the device measures the transit time, t , properly. A reference bar is usually provided by the manufacturer for this purpose. At the selected test location, it is necessary to ensure a perfect coupling between the transducers and the concrete surface because the presence of air pockets leads to misleading reading of the transit time. Many existing viscous materials can be used as a coupling agent like: grease, petroleum jelly, soft soap and oil. A thin layer of coupling agent should be applied to the transducer face or the concrete surface. For rough surfaces, the preferred coupling agent is thick grease. However, for very rough surface, a surface preparation by grinding or using a quick setting mortar is recommended (Bungey, et al., 2006).

After the application of the coupling agent, the transducers are pressed firmly against the concrete surface and the minimum transit time reading is recorded from several repeated readings (transducers are removed then re-applied at the same points) in order to avoid reading resulted from insufficient coupling. The distance between the transducers should be accurately measured and consequently the ultrasonic pulse velocity is calculated.

At each test location, in order to reduce the measurement uncertainty, the test is repeated several times (replicates) i.e. at different points within the small area of test location and consequently the test result for this test location is the mean value of these replicates. According to (ACI 228.1R, 2003) five replicates are required for the case of existing construction while three are sufficient for new construction. On contrary, the (EN 12504-4, 2004) and (ASTM C597, 2002) standards say nothings about this important issue and this is one of their shortcomings.

Most standards describe three possible ways of transducers arrangements (Komlos, et al., 1996), these are: direct transmission, semi-direct (or diagonal) transmission and indirect (or surface) transmission, see Figure 2.7. The direct transmission is the most accurate and should be used when it is possible (Garnier, 2012). When the access to opposite faces of concrete member is impossible or when there is reinforcement congestion, semi-direct transmission can be used. However, when only one surface is accessible, the transducers are applied on the same surface (indirect transmission). Due to uncertain measured path length and the fact that the calculated velocity represents the surface layer, the indirect configuration should be avoided when it is possible (ASTM C597, 2002).

In order to carry out the testing procedure efficiently and to produce accurate test results, an experienced operator is necessary.

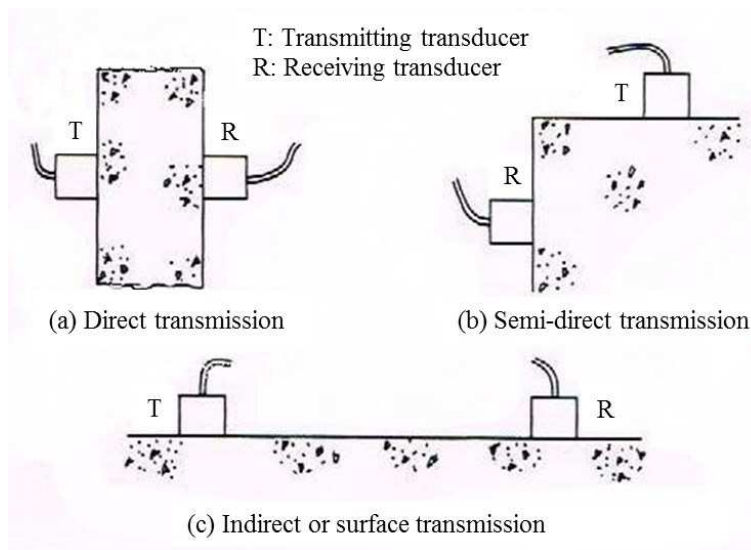


Figure 2.7 Transducers arrangements in ultrasonic pulse velocity test

2.3.2.3 Factors affecting the test results

The calculated pulse velocity at each test location is affected by several factors:

- coarse aggregate: since the wave velocity is larger in the aggregates than in the cement paste, the aggregate-to-cement ratio affects significantly the pulse velocity versus concrete compressive strength relationship. For a given strength, the pulse velocity increases as the aggregate-to-cement ratio increases (Wheen, 1974),
- moisture content: for a given strength, wet concrete shows higher pulse velocity as compared with dry concrete (Bungey, 1980),
- water-to-cement ratio: as the water-to-cement ratio decreases, the compressive strength and the pulse velocity increase (Kaplan, 1959). However, the ratio of increases in compressive strength and pulse velocity are not the same (Lin, et al., 2007),
- concrete age: the velocity increases with the concrete age but with a decreasing rate (Popovics, et al., 1990). This effect is significant at earlier ages, therefore it can be neglected in the case of existing old structures,
- concrete temperature: when the concrete temperature varies within the interval (10-30)°C, no significant changes take place in the pulse velocity unless the occurrence of changes in elastic properties or strength (EN 12504-4, 2004),

- cracks and voids: the presence of cracks and voids leads to longer travel path of the propagated wave and as a result the transit time increases (Carino, 2008),
- path length: generally, the path length has no effect on the pulse velocity, however for small path lengths, the pulse velocity may be significantly affected by the heterogeneous nature of concrete (Jones & Façaoaru , 1969). Therefore, (EN 12504-4, 2004) recommends that the minimum path length is 100 mm for concrete having maximum aggregate size ≤ 20 mm. While for maximum aggregate size within the interval (20-40) mm, the minimum path length should be 150 mm,
- transducer frequency: the frequency of the commonly used transducer is 54 kHz. However, for laboratory specimens or in-situ concrete members that have small lateral dimensions, the frequency should be selected carefully in order to be sure that the lateral dimension (path length) is equal or greater than the wavelength (wavelength = pulse velocity/frequency) (Bungey, et al., 2006),
- reinforcing bars: the pulse velocity of reinforced concrete in the vicinity of reinforcement is higher as compared with that of plain concrete. Consequently, wherever possible, reinforcing bars parallel and close to the path between transducers (or transverse bars that intersect this path) should be avoided when selecting the transducers testing positions (BS 1881: Part 203, 1986).

2.3.2.4 Correlation between concrete strength and pulse velocity

From the concepts of wave propagation in solids, there is a direct theoretical relation between the pulse velocity of longitudinal waves, V_p , and the elastic properties of the solid where they propagate through (Oixian & Bungey, 1996), as given in Equation (2.5):

$$V_p = \sqrt{\frac{E_d(1-\nu_d)}{\rho(1+\nu_d)(1-2\nu_d)}} \quad (2.5)$$

where E_d , ν_d and ρ are respectively the dynamic modulus of elasticity, dynamic Poisson's ratio and density of the solid. However for composite material like concrete (which have two main constituents: cement paste and aggregate), the attempts to relate theoretically the modulus of elasticity with strength and consequently the pulse with strength faced difficulties. This is because of the complexity to define this relationship for composite material by considering the individual constituent properties (elastic and strength properties) in relation to their proportions (Bungey, 1980). Other researchers (Popovics, 2001; Naik, et al., 2004) stated that there is no physical relationship between the strength and pulse velocity. Therefore empirical relationship (or model) is usually established between the pulse velocity and concrete compressive strength.

Based on their laboratory studies, many researchers have identified models correlating the concrete strength with pulse velocity in order to be used in estimating the in-situ concrete strength from the ultrasonic pulse measurements. In a state of art paper (Breysse, 2012) Breysse found more than 70 models of different types (linear, power, exponential, polynomial and miscellaneous). Some conversion models are shown in Figure 2.8. These curves are widely dispersed and consequently there is no general concrete strength-pulse velocity model which can be applied anywhere. This observation confirms the significant effects of the influencing factors listed above and points out that models should be used only for the cases from which they derived. In other wording, the model calibration for the concrete under investigation is a mandatory requirement.

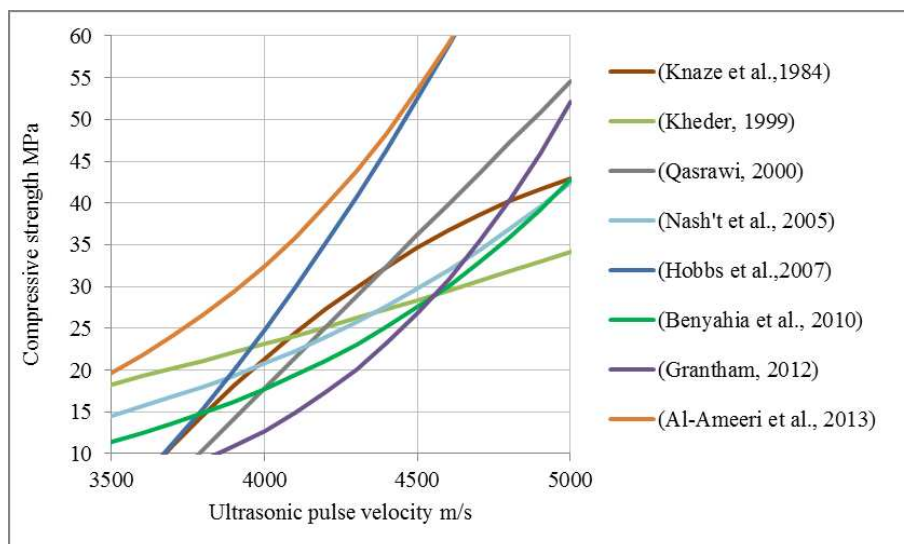


Figure 2.8 Some conversion curves identified by different researchers for estimating the cube compressive strength from ultrasonic pulse velocity

2.3.2.5 Conclusion of assessing concrete strength using pulse velocity only

Ultrasonic pulse velocity technique is a relatively simple and cheap nondestructive method for assessing the concrete strength both in-situ and in laboratory. It is a fully nondestructive method which enables retesting at the same point while this is impossible with rebound hammer due to local damage that may be produced by the impact energy. The output of this test, pulse velocity, has indirect correlation with the compressive strength. However, identifying this correlation theoretically is complicated (if not impossible) so alternatively empirical correlation is used. Unfortunately, the pulse velocity is affected by several mix, member and testing characteristics factors in addition to the concrete strength. Consequently, using a general concrete strength -pulse velocity model for assessing a real structure leads to unreliable estimations and calibration remains mandatory. Therefore, Ultrasonic pulse velocity technique cannot be used alone in assessing the concrete strength and it should be combined with the destructive tests in order to derive/calibrate a model for the case under investigation.

2.4 Strength assessment using cores and single nondestructive technique

The two previous Sections (2.2 and 2.3) have identified several drawbacks associated with the use of either destructive (DT) or nondestructive (NDT) techniques alone for assessing the concrete strength. Therefore, to overcome these drawbacks the cores are used together with rebound hammer or/and pulse velocity for the strength assessment in the existing structures.

The idea is based on carrying out nondestructive tests, according to procedures illustrated above in Section 2.3, at test locations selected carefully in order to cover the part of the structure under investigation. Next, cores are extracted, in accordance with Section 2.2, from *NC* test locations that were already tested by NDT measurements and then compressive strengths of these cores are obtained in laboratory using the compressive test. *NC* represents the number of test locations with cores (or the number of cores considering one core is extracted from each test location). The pairs of (NDT measurement, core strength) are used to establish an empirical model using model identification approaches (two approaches are available: regression and calibration that will be presented in the following subsections). The produced model can then be used to estimate the concrete compressive strengths at the test locations where only NDT measured values are available (i.e. no core).

This methodology is widely adopted by engineers in real practice for assessing strength in existing structures. Many case studies have been presented in scientific papers (Biondi & Candigliota, 2008; Luprano, et al., 2008; Pucinotti, 2015). The widespread standards that deal with the assessment of concrete strength using combination of DT and NDT techniques are (EN 13791, 2007) in Europe and (ACI 228.1R, 2003) in USA.

The number of cores is the main controlling factor in this methodology. According to the requirements of the European standard (EN 13791, 2007), the minimum number of cores is respectively 18 for regression analysis approach (Alternative 1 as so-called in this standard) and 9 for calibration approach (Alternative 2). (ACI 228.1R, 2003) standard also requires at least 12 cores (six test locations with two cores at each location) to develop an adequate strength relationship.

The way of selection the NC test locations for cores may also play a role on the reliability of assessment. The common way for selecting core locations within the NDT test locations is independent of the NDT test results. However studies, like (Pfister, et al., 2014; Breyse, et al., 2017), propose to define the core locations depending on the NDT test results. Through the present study, the NDT results based selection will be called “conditional selection”.

The existing model identification approaches will be presented in this section with an overview on the possible types of models. The factors affecting the assessment methodology will also be discussed. Then, the sources of uncertainty and the assessment reliability of the estimated strengths are pointed out.

2.4.1 Model identification approaches: Regression approach

Regression approach is the most popular statistical approach that is used to identify the conversion model between the core strengths and NDT measurements, test results, and it is recommended by almost all the standards and guide reports (RILEM TC 7- NDT, 1977; IAEA, 2002; ACI 228.1R, 2003; EN 13791, 2007). The basic principles of this approach will be illustrated herein.

2.4.1.1 Simple linear regression model

As stated above, after an investigation program, a set of NC -(x, f_c) pairs are obtained, where x represents the NDT measurement and f_c is the core strength that corresponding to one test location. Considering f_c as the dependent variable (response variable) and x as independent variable, the (x, f_c) pairs can be plotted on a scatter diagram as shown, as an example, in Figure 2.9. Assuming a linear model between the strength and NDT measurement, Equation (2.6):

$$f_{c i} = \alpha + \beta x_i + \varepsilon_i \quad i = 1, 2, \dots, NC \quad (2.6)$$

where α and β are the unknown parameters of the true regression model, and ε_i is a random error with mean value of zero and unknown variance (Montgomery & Runger, 2003). There are many possibilities for the estimators of α and β that can satisfy Equation (2.6). Least squares method is used to provide the estimators a and b for the unknowns α and β respectively. This method is based on minimizing the sum of squares of the vertical distance between the true regression line and data points (i.e. ε_i , see Figure 2.9). Consequently the values of a and b , corresponding to this method, are as given by Equations (2.7) and (2.8) (Walpole, et al., 2012):

$$b = [NC \sum_{i=1}^{NC} x_i f_{c i} - (\sum_{i=1}^{NC} x_i)(\sum_{i=1}^{NC} f_{c i})] / [NC \sum_{i=1}^{NC} x_i^2 - (\sum_{i=1}^{NC} x_i)^2] \quad (2.7)$$

$$a = \bar{f}_c - b \bar{x} \tag{2.8}$$

where \bar{f}_c and \bar{x} are respectively the average values of f_c and x values. Therefore, the fitted regression model that can be used to estimate strength from NDT measurement is:

$$f_{cest} = a + b x \tag{2.9}$$

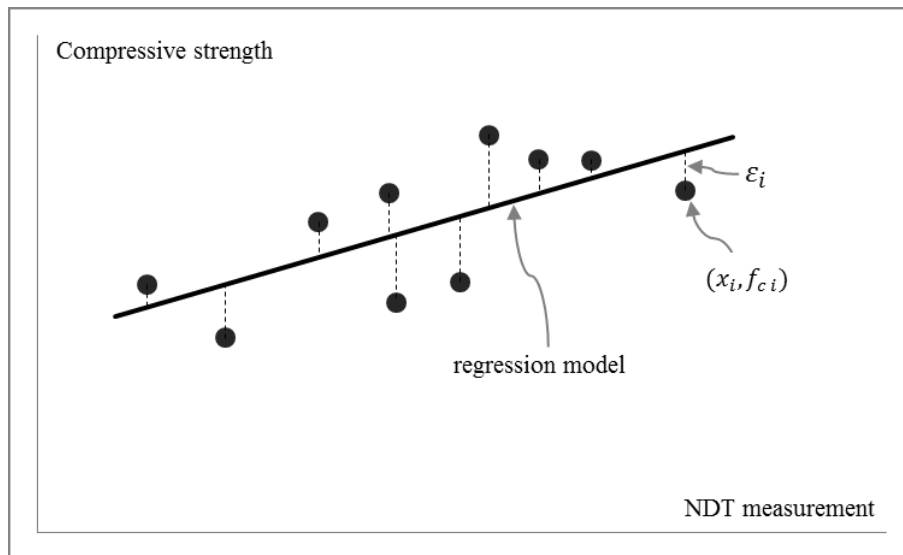


Figure 2.9 Scatter diagram of the (x, f_c) pairs with a regression model

2.4.1.2 Nonlinear regression model

The above discussion for linear regression model can be extended to nonlinear models (specific nonlinear forms, for details see (DeCoursey, 2003)) using logarithmic transformation (Montgomery & Runger, 2003). For example, the power model in Equation (2.10) can be transformed as follows:

$$f_{c i} = \alpha x_i^\beta \varepsilon_i \tag{2.10}$$

$$\ln f_{c i} = \ln \alpha + \beta \ln x_i + \ln \varepsilon_i \tag{2.11}$$

Clearly, Equation (2.11) is linear in its parameters and looks similar to Equation (2.6). Consequently estimators a_1 and b_1 for the unknown coefficients α and β can be calculated as follows:

$$a_1 = e^a, \quad b_1 = b \tag{2.12}$$

where a and b are calculated from Equations (2.7) and (2.8) after replacing x_i with $\ln x_i$ and $f_{c i}$ with $\ln f_{c i}$. Therefore, the fitted regression model that can be used to estimate strength from NDT measurement is:

$$f_{cest} = a_1 x^{b_1} \tag{2.13}$$

For nonlinear models that cannot be transformed into linear form, iterative numerical methods can be used for the least squares minimizations in order to derive the values of the unknown coefficients (DeCoursey, 2003).

2.4.1.3 Fitting error

The statisticians usually use the coefficient of determination r^2 as an indication about the adequacy of the fitted regression model. It represents the ratio of the variation in the response variable that is explained by the fitted regression model to the total variation in this variable. r^2 is calculated as shown in Equation (2.14) (Devore & Berk, 2007):

$$r^2 = 1 - \frac{SSE}{SST} \quad (2.14)$$

where SST is the total sum of squares and it represents the total variation in values of the response variable. While SSE is the error sum of squares (or sum of squares of residuals) and it measures the amount of variation unexplained by the fitted regression model. SST and SSE are calculated from the following equations:

$$SST = \sum_{i=1}^{NC} (f_{c i} - \bar{f}_c)^2 \quad (2.15)$$

$$SSE = \sum_{i=1}^{NC} (f_{c i} - f_{cest i})^2 \quad (2.16)$$

The values of r^2 are between 0 and 1. While $r^2=1$ means that the model fits all $NC-(x, f_c)$ pairs perfectly, it does not guarantee a good prediction capacity of the model. The coefficient of determination is widely used by NDT researchers as an indicator of the quality of fitting. However, other researchers (Meynink & Samarin, 1979; Brozovsky, et al., 2013; Martínez-Molina, et al., 2014) use also the correlation coefficient, r , which measures the strength and direction of the linear association between the dependent (response) and independent variables. In fact, the coefficient of determination is the square of the correlation coefficient.

Another indicator that may be used for this purpose by NDT experts (Liu, et al., 2009) is the Root Mean Square Error ($RMSE$):

$$RMSE = \sqrt{\sum_{i=1}^{NC} (f_{c i} - f_{cest i})^2 / NC} \quad (2.17)$$

As shown in Equation (2.17), $RMSE$ has the strength units and it should be as small as possible. Its practical interest is that it more directly provides the magnitude of error associated with the estimated strengths.

2.4.2 Model identification approaches: Calibration approach

In real practice, the calibration approach can also be used to identify model between the strength and NDT measurement (EN 13791, 2007; Kheder, 1999; Breysse, 2012; Breysse, et al., 2017; Pucinotti, 2015). The principle is to calibrate an existing prior model selected from literature or standards using the $NC-(x, f_c)$ pairs. Two possibilities for calibration are often used in real practice:

- multiplying factor method, and
- shifting factor method.

2.4.2.1 Multiplying factor method (k-method)

The principle comes to update an uncalibrated prior model by a coefficient k to produce a calibrated model,

$$f_{cest}(x) = k * f_{c uncal.}(x) \quad (2.18)$$

Where $f_{c\ uncal.}$ is the estimated compressive strength calculated from the selected uncalibrated prior model. The coefficient k is calculated as in the following steps:

- a) Calculate the mean value of core strengths \bar{f}_c ,
- b) Use the uncalibrated prior model to calculate the estimated strengths at core locations using the corresponding NDT measurements and then take the mean of these values, $\bar{f}_{c\ uncal.}$
- c) Calculate the calibration factor

$$k = \bar{f}_c / \bar{f}_{c\ uncal.} \quad (2.19)$$

2.4.2.2 Shifting factor method (Δ -method)

The concept here is to shift the uncalibrated prior model by a coefficient Δ ,

$$f_{cest}(x) = f_{c\ uncal.}(x) + \Delta \quad (2.20)$$

The coefficient Δ is calculated as in the following steps:

- a) Use the uncalibrated prior model to calculate the estimated strength at each core location $f_{c\ uncal. i}$ then,
- b) Calculate the shifting factor Δ

$$\Delta = \sum_{i=1}^{NC} (f_{c i} - f_{c\ uncal.i}) / NC = (\bar{f}_c - \bar{f}_{c\ uncal.}) \quad (2.21)$$

Figure 2.10 shows the scatter diagram of the dataset presented in Figure 2.9 with an uncalibrated model selected from literature and the calibrated models that result from using the two calibration methods (k-method and Δ -method).

Regarding the quality of the identified model using the calibration approach, the same indicators (i.e. r^2 , r and $RMSE$) that presented above in Subsection 2.4.1.3 can be applied herein.

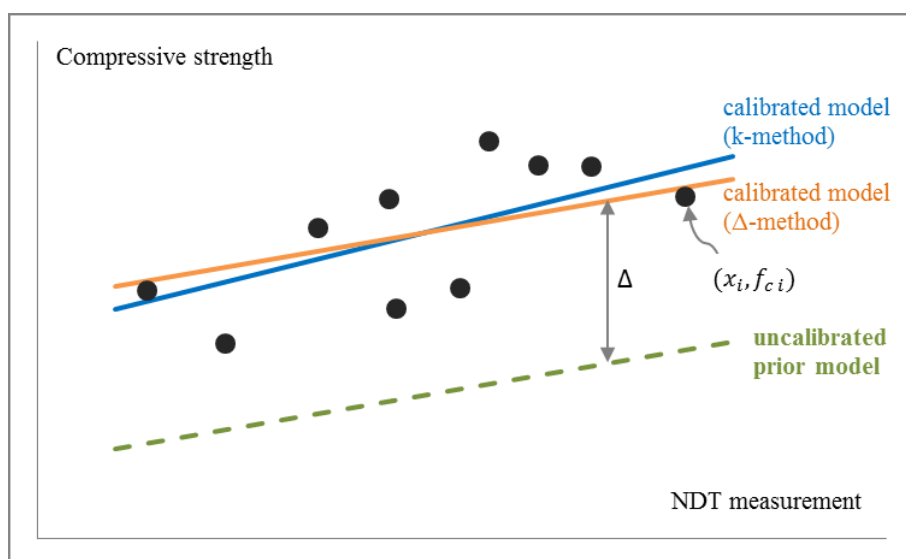


Figure 2.10 Scatter diagram of the (x, f_c) pairs with models obtained by the calibration approach

2.4.3 Types of models

As stated above, many factors can affect the relationship between the compressive strength of concrete and the NDT measurement (rebound number or the pulse velocity). Additionally many sources of uncertainty prevent a perfect fit of a model. Consequently, the problem of model identification has not a unique solution (or curve shape). The scientific literature offers a large variety of models that have been identified by different researchers. This variety indicates that nobody known the form of the true model and that all model types are less or more equivalent. The model forms are linear, power, exponential, polynomial or miscellaneous. Table 2.2 shows several models derived by different researchers.

Table 2.2 Several models derived by different researchers in order to estimate strength using single NDT technique

Model (f_{cest} in MPa, V in km/s)	Reference
$f_{cest} = 1.25R - 23 \quad 20 \leq R \leq 24$ $f_{cest} = 1.73R - 34.5 \quad 24 \leq R \leq 50$	(EN 13791, 2007)
$f_{cest} = 0.0946 R^{1.6484}$	(Brozovsky, et al., 2013)
$f_{cest} = 5.1698 e^{0.0341R}$	(D'Ambrisi, et al., 2008)
$f_{cest} = -0.00186R^2 + 2.0449R - 46.426$	(Knaze & Beno, 1984)
$f_{cest} = -0.0003 R^3 + 0.0399R^2 - 0.1525R + 3.9976$	(Schmidt, 1950) after (Szilágyi & Borosnyói, 2009)
$f_{cest} = 36.72V - 129.077$	(Qasrawi, 2000)
$f_{cest} = 0.038 V^{4.85}$	(Rojas-Henao, et al., 2012)
$f_{cest} = 0.008 e^{2V}$	(Demirboğa, et al., 2004)
$f_{cest} = 62.5V^2 - 497.5V + 990 \quad 4 \leq V \leq 4.8$	(EN 13791, 2007)

2.4.4 Factors affecting the quality of assessment

The assessment methodology is affected by all factors that influence the core strengths (Section 2.2) and NDT measurements (Section 2.3). In addition to these factors, the characteristics of the assessment methodology can also play a significant role. Therefore, we classify the influencing factors in accordance with the possibility to control them in the methodology into two groups:

- Uncontrolled factors: they include all factors that may affect the quality of assessment and that cannot be controlled because they are not included in the conversion model (due to the difficulty to assess accurately these factors in the existing structures) like: concrete mix characteristics (aggregate type/size/percentage, cement type/percentage, water-to-cement ratio, and additives), concrete moisture condition, concrete surface carbonation, concrete temperature, and voids.
- Controlled factors: they include all factors that can affect the quality of assessment, but that can be controlled (changed) in order to modify the quality of assessment, such as: number of test locations for cores used to identify the conversion model, number of test locations for NDT measurements, the way of selecting the test locations, the quality of measurements (within-test variability), the type of NDT technique, using single or combination of NDT techniques, the model identification approach and the model type (or shape). Since these factors represent the inherent characteristics of the

assessment methodology, they will be also so-called “methodology inherent characteristics” through this study.

Our concern in the present work will be the controlled factors, and any uncontrolled factor will be considered as a source of uncertainty in the assessment methodology.

2.4.5 Sources of uncertainty

The uncertainty is a quantity intended to characterize a range of values which contains the reference value, where the latter may be either the true value or the expectation (EUROLAB Technical Report 1/2006, 2006). In other wording, the range (estimated value \pm uncertainty) represents the interval within which the true value will be assured to lie with a specified confidence level. Uncertainty differs from error which represents the difference between the estimated and true values (Bell, 2001). Since the true strength or true NDT measurement are unknowns in the real structures, the error cannot be determined and the uncertainty is the only way to measure the quality of assessment.

For the strength assessment issue, many sources of uncertainty are expected to feed the global uncertainty of the estimated strength. These are:

- in-situ concrete strength variability (intrinsic variability); the concrete strength variability is induced in the structure due to: quality control (mixing, casting, compacting and curing processes), the uneven environmental effects on the structure, and the cracks and deterioration development in the structure (Szilágyi, et al., 2014). These elements can induce strength variability in three scales: batch-to-batch variability, member-to-member variability, and within member variability (Pereira & Romao, 2016). To understand the effect of the concrete variability on the uncertainty, let's consider, as an example, the assessment of the mean strength in a structure using a sample of *NC* cores shown in Figure 2.1. From this figure, it is clear that, for a specific *NC* value, the percentage error (or uncertainty) of the estimated mean strength increases with the increase in the concrete strength variability.
- sampling uncertainty: it is associated with sample size i.e. in any investigation program, tests are carried out at limited number of test locations while many other possible test locations within the investigated structure (or the investigated part of structure) are left untested. The lack of knowledge associated with non-surveyed test locations generates uncertainty (Pereira & Romao, 2016). In any investigation program, this uncertainty arises from two sources: the first one concerns the selection of test locations for NDT measurements (*NT* test locations) from all possible test locations, and the second relates to the selection of *NC* test locations for cores from the *NT* test locations that already selected for NDT,
- measurement uncertainty: it is a parameter associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to a measured quantity (rebound number, pulse velocity, etc.) (JCGM 100, 2008). At any test location, the uncertainty of a test result (average of several replicates) is a function of number of replicates and within-test variability (ACI 228.1R, 2003),
- model uncertainty: as discussed in Subsection 2.4.3, the true shape for the curve (model) that correlates the concrete strength with the NDT measurements is unknown, therefore any proposed model will already have an error due to this uncertainty. Moreover, the uncertainty associated with the model identification approach also feeds the model uncertainty. Another source of the model uncertainty is the influencing factors that are not considered in the model due to difficulty or impossibility to measure these factors (uncontrolled factors) (Der Kiureghian & Ditlevsen, 2009). For

example, including the mix characteristics in the model is usually impossible in the existing structures due to the missing information from the construction period.

2.4.6 Quality of assessment

For decades, studying the quality of assessment (i.e. how close the estimated value to the true one) of the concrete strength by NDT measurements has been the objective of many scientific researches. However, this issue is quite controversial (Proverbio & Venturi, 2005; Fischli & Moczko, 2012).

Regarding rebound hammer technique, some researchers (Carino, 2008; Brencich, et al., 2013; Pucinotti, 2015) are pessimistic, considering that rebound hammer is unable to provide a reliable estimate of the concrete strength. On contrast, other researchers like (Malhotra, 2004) consider that the accuracy of estimation of compressive strength of test specimens cast, cured, and tested under laboratory conditions by a properly calibrated hammer lies between ± 15 and $\pm 20\%$. Furthermore, the probable accuracy of estimation of concrete strength in a structure is $\pm 25\%$ (Malhotra, 2004; Ministry of railways-India, 2009). Szilágyi and Borosnyói (Szilágyi & Borosnyói, 2009) indicate that the expected error of the strength estimation by the Schmidt rebound hammer under general service circumstances is about $\pm 30\%$. FHWA (FHWA, 1997) states that the accuracy of rebound hammer for estimating in-situ compressive strength is between $\pm 30\%$ and $\pm 40\%$.

Regarding ultrasonic pulse velocity technique, using model established/calibrated for the case under consideration, the strength can be estimated with a $\pm 20\%$ accuracy (Bungey, 1980; Popovics, et al., 1990; Komlos, et al., 1996).

Some other studies (Kheder, 1999; Qasrawi, 2000; Hobbs & Kebir, 2007) have reported that the combination of rebound hammer with the ultrasonic pulse velocity can improve the quality of assessment.

All these numbers illustrate that: there is no consensus between the specialists, and the quality of concrete strength assessment using NDT measurements is still low even after the model identification/calibration with the core strengths. Consequently the quality of assessment remains an open question and needs to be improved or at least really known. The only way to do this is by controlling the sources of uncertainty stated above in Subsection 2.4.5. To this end, it is necessary to correlate the uncertainty with the methodology inherent characteristics in order to study how the quality of assessment can vary with any change in these factors.

2.4.7 Conclusion of assessing concrete strength using cores and single nondestructive technique

- a) From the basics of the existing approaches (regression and calibration) illustrated above, it is obvious that none of these approaches has the objective to capture the concrete variability although the standards recommend the estimation of the concrete strength variability since it is an essential parameter in the calculation of the characteristic strength of concrete.
- b) The minimum number of cores required by standards (18 for regression approach and 9 for calibration approach according to (EN 13791, 2007), and (ACI 228.1R, 2003) requires at least 12 cores) is variable and generally high. A consequence is the slow development of strength assessment using NDT and cores.
- c) The assessment methodology is affected by all factors that influence the core strengths and NDT measurements. In the existing structures, influencing factors (like concrete mix characteristics, concrete moisture condition, concrete surface carbonation, concrete temperature, and voids) are difficult to assess accurately. Therefore, these

factors are usually not included in the conversion model and consequently they are considered as uncontrolled factors. Therefore, these factors will be considered as a source of uncertainty in the assessment methodology.

- d) The methodology inherent characteristics (number of test locations for cores used to identify the conversion model, number of test locations for NDT measurements, the way of selecting the test locations, the quality of measurements, the type of NDT technique, using single or combination of NDT techniques, the model identification approach and the model type) can also affect the uncertainty of assessment. Since the methodology inherent characteristics factors can be controlled in the assessment methodology, therefore these factors deserve a more comprehensive analysis.
- e) The quality of assessment of concrete strength using NDT measurements remains low even after the model identification/calibration with the core strengths. Consequently the quality of assessment remains an open question and needs to be improved or at least really known. The only way to do this is by controlling the sources of uncertainty (in-situ concrete strength variability, sampling uncertainty, measurement uncertainty, model uncertainty). To this end, it is necessary to correlate the uncertainty with the methodology inherent characteristics (controlled factors) in order to study how the quality of assessment can vary with any change in these factors. Therefore this issue needs farther studies.

2.5 Strength assessment using cores and combination of nondestructive techniques

Instead of using a single NDT technique with cores for assessing the concrete, the NDT techniques can be used in combination (in addition to the cores). The theoretical principle of combination is that when two or more NDT techniques are affected inversely by an influencing factor, combining these techniques can reduce or eliminate this effect and as a result improve the reliability of strength estimation (Soutsos, et al., 2012; Sbartai, et al., 2012). As an example the effect of the concrete moisture condition which produces an increase in pulse velocity and decrease in rebound number when it increases. However, the benefit of this improvement in reliability resulted from using the combination of NDT techniques should be assessed against the additional time, cost, and complexity of using this combination (Samarin, 2004).

Combining the ultrasonic pulse velocity and rebound hammer techniques is the most popular combination which is known as SONREB. RILEM Technical Committee (TC 43) played a major role in the development of the SONREB method. Its recommendation (RILEM TC 43-CND, 1993) provided a procedure to establish iso-strength curves for a reference concrete (concrete has the materials and composition from a particular region or country for which the curves are devoted), as an example of these curves, see Figure 2.11. For different concrete compositions, correction factors are used for this purpose. When the composition is unknown (as it is the case for old structures), the correction factor should be estimated using cores extracted from the structure under investigation (RILEM TC 43-CND, 1993). In fact, the iso-strength curves represent specific conversion models that correlate the concrete strength with NDT values (pulse velocity and rebound number) and the correction factor looks like the calibration factor that was illustrated in Subsection 2.4.2.1. The nomogram shown in Figure 2.11 is not unique and many other versions were developed by researchers all around the world [see examples (Cianfrone & Facaoaru, 1979; Knaze & Beno, 1984; Schickert, 1984) for iso-strength curves, (Qasrawi, 2000; IAEA, 2002) for iso-rebound number curves, and (Galan, 1984) for iso-pulse velocity curves].

The variety of these iso-curves emphasizes that they can give a correct strength prediction only for the particular cases that they derived for it. Therefore, to assess concrete on site, the strengths should be predicted using model derived for the concrete under consideration. The assessment methodology is similar to that illustrated in Section 2.4. It includes: carrying out the NDT measurements, extracting cores, establishing a model using the dataset of (pulse velocity V , rebound number R , compressive strength f_c) values, and estimating strength at any test location from applying V and R values (corresponding to this test location) in the model.

In this section, an extension to the discussion of using single NDT technique detailed in Section 2.4 will be provided with a focus on the modifications for the case of combined NDT techniques. Finally, the efficiency of combination will be discussed.

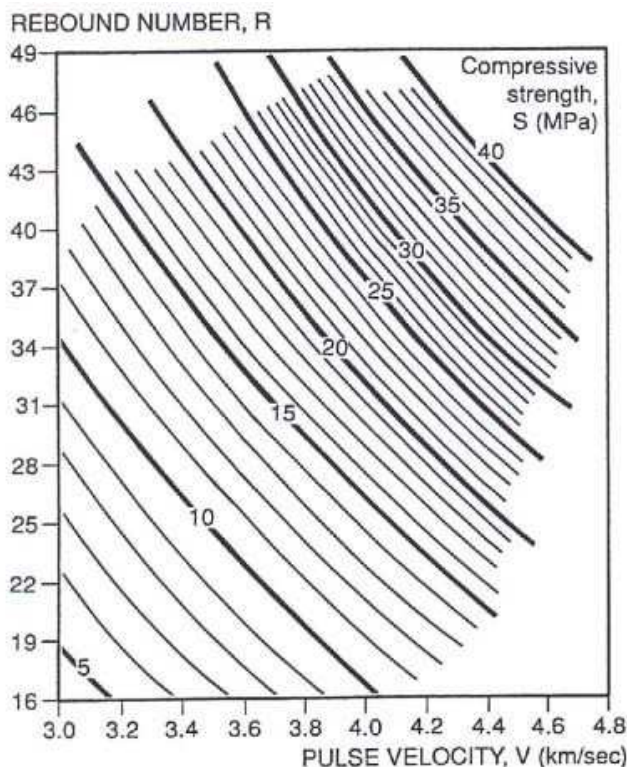


Figure 2.11 Iso-strength curves for a reference concrete in SONREB method (IAEA, 2002)

2.5.1 Model identification approaches

Regression approach illustrated above for the case of single NDT technique as an independent variable can be extended to the case of more than one NDT technique as independent variables. Therefore Equation (2.6) is modified as shown:

$$f_{c i} = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_q x_{iq} + \varepsilon_i \quad i = 1, 2, \dots, NC \quad (2.22)$$

where q is the number of independent variables ($x_1, x_2 \dots$). Using the least squares minimizations (Ross, 2009), the estimators (a, b_1, b_2, \dots, b_q) of the true regression model parameters ($\alpha, \beta_1, \beta_2, \dots, \beta_q$) are obtained and consequently the fitted regression model that can be used to estimate strength is:

$$f_{cest} = a + b_1 x_1 + b_2 x_2 + \dots + b_q x_q \quad (2.23)$$

Statistical software (Excel, Minitab, Matlab, etc.) is available for performing the least squares minimizations and providing the estimated model parameters.

Regarding the case of nonlinear regression model, the same principle of logarithmic transformation (Montgomery & Runger, 2003) illustrated above in Subsection 2.4.1.2 can be applied here. Moreover, the calibration methods (k-method and Δ -method) can be used directly for the case of NDT combination because the only difference is the uncalibrated model.

2.5.2 Types of models

For combining the pulse velocities and rebound numbers with cores, different model forms have been considered by the researchers, such as: bilinear, double power, exponential, polynomial and other miscellaneous form. In the state of art paper (Breysse, 2012), Breysse has gathered about 69 models from literature. Table 2.3 shows several models derived by different researchers.

A possibility promoted by some researchers, see for examples (Kheder, 1999; Atici, 2011), is to combine (in addition to NDT measurements) several concrete characteristics (water-to-cement ratio, aggregate-to-cement ratio, admixture content, concrete density, age, etc.) in the model. However the main drawback of this type of models is the need to know the concrete characteristics as inputs in the model while they usually remain unknowns in old structures.

Another possibility is to combine more than two NDT techniques. For example, the French National Project SENSO (Balayssac, et al., 2012; Sbartaï , et al., 2012) developed a methodology for combining ultrasonic pulse velocity with electrical resistivity and ground penetrating radar (GPR) for the evaluation of compressive strength and other concrete characteristics as porosity, elasticity and saturation rate. However, it is out of our scope. In this study, models with only NDT measurements (V and R) as inputs will be considered.

Table 2.3 Several models derived by different researchers in order to estimate strength using combination of rebound hammer and pulse velocity techniques

Model	units	Reference
$f_{cest} = 8.630 V + 1.416R - 51.581$	f_{cest} in MPa V in km/s	(Soshiroda, et al., 2006)
$f_{cest} = 7.695 \times 10^{-11} V^{2.6} R^{1.4}$	f_{cest} in MPa V in m/s	(RILEM NDT4, 1993) after (Fiore, et al., 2013)
$f_{cest} = e^{0.446V+0.048R}$	f_{cest} in MPa V in km/s	(Machado , et al., 2009)
$f_{cest} = 0.67e^{0.72V+0.04R}$	f_{cest} in MPa V in km/s	(Sravindrajah, et al., 1988)
$f_{cest} = 0.42 R^{0.63} e^{0.58V}$	f_{cest} in MPa V in km/s	(Al-Ameeri, et al., 2013)
$f_{cest} = -173.04 + 4.07V^2 + 57.96 V + 1.31R$	f_{cest} in MPa V in km/s	(Shariati , et al., 2011)
$f_{cest} = -24.1 + 1.24R + 0.058V^4$	f_{cest} in MPa V in km/s	(Meynink & Samarin, 1979)
$f_{cest} = -76.30 + 0.17 V^{0.46} R^{0.7}$	f_{cest} in MPa V in m/s	(Proverbio & Venturi , 2005)
$f_{cest} = (R/(3.64 + 0.023R - 0.56V))^2$	f_{cest} in kg/cm ² V in km/s	(Postacioglu, 1985)

2.5.3 Factors affecting the quality of assessment

The assessment methodology using combination of NDT techniques is affected by all factors illustrated above in Subsection 2.4.4 (controlled and uncontrolled factors) i.e. the factors that influence the core strengths, NDT measurements and the methodology inherent characteristics.

2.5.4 Sources of uncertainty

As discussed above in Subsection 2.4.5, the main sources of uncertainty are: in-situ concrete strength variability, sampling uncertainty, measurement uncertainty and the model uncertainty.

2.5.5 Quality of assessment

According to RILEM TC 43 (RILEM TC 43-CND, 1993), the accuracy of SONREB method is 12-15% (for a 90% confidence level) when SONREB curves (or models) are calibrated according to the core strengths while the concrete composition is unknown. This accuracy level is derived from laboratory studies, therefore it cannot be assured for the case of in-situ assessment due to the effect of many uncontrolled factors that presence on site.

2.5.6 Efficiency of combination

Combination of NDT techniques has received a lot of attention during the last decades. However, there is no general consensus about the efficiency of combination. Some researchers found that combining ultrasonic pulse velocity and rebound hammer yields more reliable and results closer to the true strengths (Wheen, 1974; Qasrawi, 2000; Hobbs & Kebir, 2007), while others did not find a significant improvement in the concrete assessment by combining methods (Komlos, et al., 1996; Luprano, et al., 2008). In any case, the combination is not an end in itself and it should be applied in the cases where such combination is the most beneficial way of assessing the concrete strength (Leshchinsky, 1991).

In common practice, in order to study the efficiency of combination, the coefficient of determination r^2 for the model derived for the combination of NDT techniques is usually compared with those corresponding to the models established from using these techniques separately. Table 2.4 gives the r^2 values provided by different research works for models derived for the cases of using rebound hammer technique, ultrasonic pulse velocity technique and their combination. It is clear that r^2 value for the combination case is always the best (i.e. r^2 combined $>$ r^2 single) and consequently one could conclude that the combination gives a more reliable assessment as compared with the use of the techniques separately. In fact this is a hasty conclusion because, from the statistical viewpoint, adding new term (new independent variable) to the model decreases error degrees of freedom and, as a result, increases r^2 or at least does not decrease it (Walpole, et al., 2012; Montgomery & Runger, 2003). Therefore, the increase in r^2 cannot confirm that the model produced from combination is better than that obtained from using single technique. We think that r^2 can be a misleading indicator about the efficiency of combination and may lead to wrong conclusions. Thus, assessing the model prediction capability is the only way to decide whether combination is better than single technique.

Table 2.4 r^2 values as provided by different researchers for models derived for: single rebound hammer technique, single ultrasonic pulse velocity technique and their combination

Reference	Single R	Single V	Combined $V+R$
(Proverbio & Venturi , 2005)	0.82	0.46	0.86
(Mahmoudipour, 2009)	0.25	0.61	0.67
(Pucinotti, 2015)	0.24	0.82	0.89
(Shariati , et al., 2011)	0.94	0.92	0.95
(Kheder, 1999)	0.67	0.18	0.85
(Soshiroda, et al., 2006)	0.93	0.81	0.94
(Hobbs & Kebir, 2007)	0.93	0.90	0.95
(Nash't, et al., 2005)	0.77	0.59	0.80
(Al-Ameeri, et al., 2013)	0.96	0.98	0.99
(Machado , et al., 2009)	0.90	0.64	0.99

2.5.7 Conclusion of assessing concrete strength using cores and combination of nondestructive techniques

- a) Using combination of rebound hammer and ultrasonic pulse velocity techniques with cores for assessing the concrete strength is a common practice. However, there is no general consensus about the efficiency of this combination. Some researchers found that it yields more reliable results while others did not find a significant improvement in the concrete assessment by combining methods. Moreover, information about combination from standards is very poor. Therefore, the efficiency of combination needs to be studied. Also, the effect of the methodology inherent characteristics (controlled factors) on the efficiency of combination has to be studied.
- b) In real practice, in order to study the efficiency of combination, the coefficient of determination r^2 for the model derived for the combination of NDT techniques is compared with those corresponding to the models established from using these techniques separately. In fact, r^2 can be a misleading indicator about the efficiency of combination and may lead to wrong conclusions. Thus, assessing the model prediction capability is the only way to decide whether using combined techniques is better than using single NDT technique.

2.6 Conclusions

- a) The methodology inherent characteristics (number of test locations for cores used to identify the conversion model, number of test locations for NDT measurements, the way of selecting the test locations, the quality of measurements (within-test variability), the type of NDT technique, using single or combination of NDT techniques, the model identification approach and the model type) can affect the uncertainty of assessment. Since the methodology inherent characteristics factors can be controlled in the assessment methodology, therefore these factors deserve a more comprehensive analysis.
- b) Using combination of rebound hammer and ultrasonic pulse velocity techniques with cores for assessing the concrete strength is a common practice. However, there is no general consensus about the efficiency of this combination. Some researchers found that it yields more reliable results while others did not find a significant improvement in the concrete assessment by combining methods. Moreover, information about combination from standards is very poor. Therefore, the efficiency of combination needs to be studied. Also, the effect of the methodology inherent characteristics factors (controlled factors) on the efficiency of combination has to be studied.
- c) In real practice, in order to study the efficiency of combination, the coefficient of determination r^2 for the model derived for the combination of NDT techniques is compared with those corresponding to the models established from using these techniques separately. In fact, r^2 can be a misleading indicator about the efficiency of combination and may lead to wrong conclusions. Thus, assessing the model prediction capability is the only way to decide whether using combined techniques is better than using single NDT technique. This issue needs more focusing on.
- d) From the basics of the existing approaches (regression and calibration) illustrated above, it is obvious that none of these approaches has the objective to capture the concrete variability although the standards recommend the estimation of the concrete strength variability since it is an essential parameter in the calculation of the characteristic strength of concrete. Thus the improvement of the model identification process is necessary.
- e) The quality of assessment of concrete strength using NDT measurements remains low even after the model identification/calibration with the core strengths. Consequently the quality of assessment remains an open question and needs to be improved or at least really known. The only way to do this is by controlling the sources of uncertainty (in-situ concrete strength variability, sampling uncertainty, measurement uncertainty, model uncertainty). To this end, it necessary to correlate the uncertainty with the methodology inherent characteristics in order to study how the quality of assessment can vary with any change in these factors. Therefore this issue needs a farther study.

CHAPTER THREE

MEANS AND TOOLS

3.1 Introduction

This chapter deals with the scientific tools that are adopted in the present thesis in order to analyze and improve the strength assessment methodology. It begins by defining the assessment strategy and its elements. Then, sources of datasets used through the present work are detailed. Finally, the simulator that has been built for this study is illustrated showing how its main algorithm works.

3.2 Definition of the assessment strategy

From the previous chapter, there is evidence that the methodology of strength assessment using destructive and nondestructive tests has many degrees of freedom (type of test method, number of measurements, type of model, model identification approach, etc.). Therefore, there are many possible combinations of the values of these elements. In the present thesis, each combination of these values is so-called “assessment strategy”. For example, if one defines his strategy to have specified characteristics (number of test locations for ultrasonic pulse velocity measurements =40, number of test locations for cores $NC=5$, cores locations are randomly selected, using regression approach to identify a power model). Then any change in these characteristics produces a new assessment strategy. Consequently, the assessment methodology can be achieved using many possible strategies.

Therefore, through the different analysis presented in this thesis, any assessment strategy follows the general definition presented in Figure 3.1. As shown, the assessment strategy is defined to consist in two main parts: the investigation program and the strength assessment. The investigation program includes:

- Selecting the suitable NDT method such as using rebound hammer or ultrasonic pulse velocity or a combination of NDT methods. This selection depends on many factors, among them is the investigation budget. Due to the difference in the cost between the test methods (or between using single and combination of NDT techniques), for fixed budget, the number of possible tests using rebound hammer will be larger than that when using the pulse velocity method. It is necessary also to select the number of replicates for each test method (i.e. number of times the measurement is repeated within a test location in order to derive a test result) which depends on the within-test variability of the test method and the investigation budget.
- Selecting the number of test locations for each type of measurements (DT and NDT). Regarding the number of test locations for cores, NC , as discussed in Chapter 2, standards provide the minimum limits. However concerning the number of test locations for NDT measurements, NT , for assessing the existing structures, standards do not give any specified values (only for new construction, the ACI 228.1R recommends specified values for testing different structural members

using different test methods). Nevertheless, the number of NDT test locations generally depends on the investigation budget, the in-situ concrete strength variability, and the reliability of the final results of assessment (ACI 228.1R, 2003).

- The way of defining the test locations is also an essential issue in the investigation program. On site, the investigation program usually starts by selecting the *NT* test locations for carrying out NDT measurements. The selection of these locations can be “random” or “predefined” by the investigator according to specific constraints (like, for example, focusing on exterior columns or choosing to test all the selected columns at mid-height). Regarding the selection of *NC* core locations from the *NT* test locations, it is usually either random or predefined selection (i.e. selection independent of NDT test results). However, studies, like (Pfister, et al., 2014; Breyse, et al., 2017), propose to define core locations depending on the NDT test results i.e. “conditional selection”.
- The last stage in the investigation program is carrying out on site measurements and laboratory tests in order to provide the final investigation results.

The second part of the assessment strategy “strength assessment” deals with the analysis and the interpretation of the test results obtained during the investigation stage. It consists in the following steps:

- Selecting the model type and the model identification approach. As discussed in Chapter 2, several model types are used by the investigators in the real practice and there is no general consensus on a specific type. Regarding the model identification approaches, it was discussed that people usually uses either prior model without any calibration, or prior model after calibration using one of the calibration methods described in Chapter 2, or identifying models using the regression analysis.

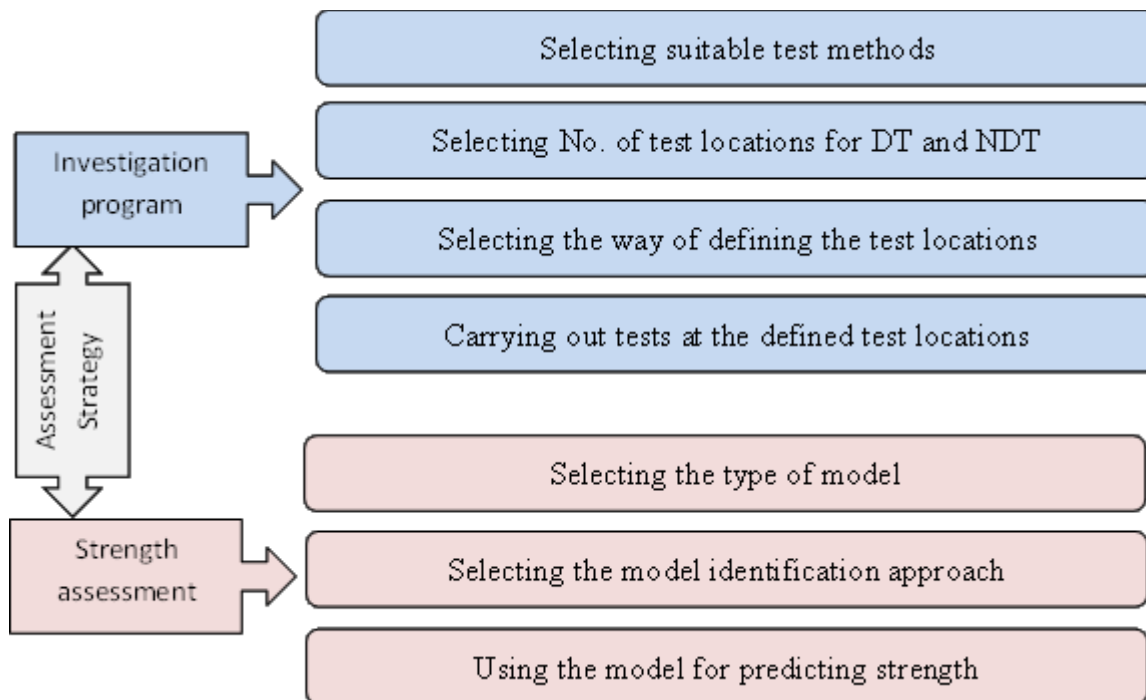


Figure 3.1 Definition of the assessment strategy

- Using the established model to predict the strength at the test locations where there are the NDT measured values. These estimated local strengths can be used to determine the estimated mean strength and the estimated concrete strength variability.

3.3 Sources of data

3.3.1 In-situ on structure and laboratory studies data

In order to analyze the current assessment methodology and the most influencing factors, datasets in form of (NDT measurement, compressive strength) are required. Through the different phases of the present study, two sources of experimental datasets will be used: laboratory studies and in-situ investigation on structures. Table 3.1 gives a summary about 39 datasets (S1, S2, S3, ..., S39) used in the present PhD work. These datasets include 3922 test locations. The test results at each test location may be (V, f_c) , (R, f_c) or (V, R, f_c) . Among these test locations, the in-situ datasets have 358 test locations that cover concretes with mean strength, \bar{f}_c , varying from 13.4 to 32.6 MPa and concrete strength variability, $s(f_c)$, ranging from 2.1 to 13.3 MPa. Regarding the laboratory studies datasets, they cover a wider range of concrete strengths (mean strength ranges from 21 to 77.4 MPa and strength variability varies from 4.8 to 18.3 MPa). Some of the datasets presented in Table 3.1 are collected from scientific literature while the others are from research collaborations with Algerian researchers (Said Kenai and Khoudja Ali Benyahia) and with Hungarian researchers (Katalin Szilágyi, Adorján Borosnyói).

The Algerian data originate from a full investigation program carried out on an existing building before the demolition. It was situated in Blida (70 km from the capital of Algeria). The building consisted of two blocks; block 1 was a 3-story reinforced concrete frame while block 2 had a 2-story frame. The in-situ tests consisted in NDT measurements (pulse velocity and rebound hammer) that were carried out at 205 test locations distributed over the members (beams and columns) of each story in the two blocks (120 test locations for block 1 and 85 for block 2). 75mm diameter cores were also extracted from these 205 test locations for compressive strength test.

The collaboration with the Hungarian team provided the test result pairs (R, f_c) for more than 2900 test locations produced from different laboratory studies. These laboratory studies have a large variety of concrete characteristics (mix properties, age, curing, and admixture) and testing conditions. The dataset sizes vary from 100 to 216 test result pairs. They cover a wide range of concrete mean strength \bar{f}_c (34.5-77.4 MPa) and concrete strength variability (in terms of strength standard deviation, $s(f_c)$, from 6.7 to 18.3 MPa or in terms of strength coefficient of variation, $CV(f_c)$, from 11 to 33%). Regarding the rebound number, each dataset has a mean value, \bar{R} , within the range (32-48) and the variability in rebound number within each dataset, $s(R)$, falls within the range (2.1-7.5). Moreover, these datasets include the values of individual readings at each test location (10 rebound hammer readings on the same surface of a concrete specimen during the laboratory tests and R , test result value, represents the average value of these 10 replicates). Therefore, within-test variability (or repeatability) of rebound measurements at each test location (in terms of the coefficient of variation) is known for each test result. This information is necessary for studying the effect of quality of measurements.

To evaluate the quality of assessment, the strength estimations are necessary to be compared with the true in-situ strengths. However, as discussed in Chapter 2, the true in-situ strengths are always unknowns in the real structures and core strengths can only be considered as true reference values to these strengths. Therefore, through this study, the strengths (local strengths, mean strengths, and strength variabilities) provided by the datasets given in Table

3.1 will be called “true reference” values that will be used as references (i.e. the strength estimations will be compared with these true reference strengths).

Of course, in-situ dataset is preferable for analyzing the assessment methodology because it is more representative of the situation in real structures. However, due to the cost, time, structural and aesthetic considerations, the availability of the well-documented and reliable in-situ datasets remains limited. In addition, the sizes of these datasets are usually small (except for the cases of demolition) as it is shown in Table 3.1. For the statistical stability reason, larger datasets are required. Laboratory studies datasets can meet this requirement, however they are less representative of the situation in real structures due to several reasons:

- Due to the effect of quality control conditions, the in-situ strength in a structure is usually less than the strength of the standard specimens produced from the same concrete and compacted and cured in a standard laboratory control conditions,
- The difference between the sources of strength variability for laboratory dataset and the sources for in-situ dataset. The in-situ strength variability in the structure results from within-member variability (for instance, strength at lower part of the column is higher than that at higher part), batch to batch variability, and/or due to strength variation between the weak and normal undamaged zones in the structure. While the strength variability in a laboratory dataset is resulted from using different mix characteristics and/or from testing specimens at different ages.

Therefore, the application of in-situ dataset or laboratory study dataset has its own limitations. It is the reason why, besides these two types of data, synthetic data derived from synthetic simulations will be also used in the present PhD work.

Table 3.1 Summary of the collected datasets

Dataset Symbol	Dataset size	Source of dataset	Dataset characteristic						Reference
			\bar{R}	$s(R)$	\bar{V} m/s	$s(V)$ m/s	\bar{f}_c MPa	$s(f_c)$ MPa	
S1	205	in-situ	34.27	5.80	3671	434	18.81	6.50	(Ali-Benyahia, et al., 2017)
S2	23	in-situ	41.41	4.50	3769	280	14.09	4.72	(Pucinotti, 2015)
S3	22	in-situ	—	—	3095	421	16.21	5.29	(Giannini, et al., 2014)
S4	21	in-situ	—	—	3739	170	32.57	5.51	(Giannini, et al., 2014)
S5	21	in-situ	36.33	2.74	—	—	19.74	3.76	(Monteiro & Gonçalves, 2009)
S6	18	in-situ	25.61	1.75	3840	73	13.41	2.14	(Hannachi & Guetteche, 2012)
S7	18	in-situ	26.33	7.45	2765	928	18.07	8.50	(Masi & Vona, 2008)
S8	16	in-situ	39.21	4.39	3225	467	26.26	13.33	(Nobile, 2015)
S9	14	in-situ	28.07	8.01	3309	673	23.99	6.23	(Masi, et al., 2016)
S10	80	Lab.	36.46	5.47	5079	179	51.53	11.89	(Cianfrone & Facaoaru, 1979)
S11	63	Lab.	37.09	2.26	4206	156	27.17	4.76	(Knaze & Beno, 1984)
S12	60	Lab.	30.52	4.47	4461	416	23.61	7.73	(Oktar, et al., 1996)
S13	40	Lab.	39.41	8.41	4721	319	37.70	12.04	(Jain, et al., 2013)
S14	30	Lab.	43.74	4.67	4450	200	47.94	10.76	(Rojas-Henao, et al., 2012)
S15	20	Lab.	30.48	4.01	4037	168	27.35	7.42	(Nikhil, et al., 2015)
S16	16	Lab.	31.38	5.25	4656	184	37.33	12.63	(Sbartai, et al., 2012)
S17	120	Lab.	—	—	4409	223	31.37	11.16	(Musmar & Abedalhadi, 2008)
S18	60	Lab.	—	—	4436	96	61.83	9.36	(El Mir & Nehme, 2016)
S19	24	Lab.	—	—	4513	181	20.96	7.20	(del Rio, et al., 2004)

S20	120	Lab.	33.68	5.26	—	—	22.45	5.88	(Hajjeh, 2012)
S21	144	Lab.	39.30	2.95	—	—	35.95	6.69	(Szilágyi , 2013)
S22	118	Lab.	48.14	3.64	—	—	63.30	11.74	(Szilágyi , 2013)
S23	114	Lab.	32.34	2.62	—	—	41.73	9.35	(Szilágyi , 2013)
S24	144	Lab.	48.38	2.12	—	—	77.41	11.70	(Szilágyi , 2013)
S25	100	Lab.	47.29	3.39	—	—	67.19	7.65	(Szilágyi , 2013)
S26	136	Lab.	37.11	4.68	—	—	45.78	9.71	(Szilágyi , 2013)
S27	120	Lab.	36.96	5.21	—	—	44.50	11.29	(Szilágyi , 2013)
S28	120	Lab.	37.84	4.24	—	—	42.62	10.26	(Szilágyi , 2013)
S29	118	Lab.	46.53	6.29	—	—	65.35	12.77	(Szilágyi , 2013)
S30	172	Lab.	46.04	3.64	—	—	69.20	11.88	(Szilágyi , 2013)
S31	208	Lab.	46.44	4.68	—	—	71.12	14.34	(Szilágyi , 2013)
S32	216	Lab.	46.81	4.84	—	—	70.25	16.38	(Szilágyi , 2013)
S33	160	Lab.	44.18	3.90	—	—	60.72	11.87	(Szilágyi , 2013)
S34	212	Lab.	39.07	5.64	—	—	54.71	15.13	(Szilágyi , 2013)
S35	212	Lab.	37.39	5.77	—	—	45.82	14.71	(Szilágyi , 2013)
S36	204	Lab.	40.07	5.64	—	—	63.55	18.30	(Szilágyi , 2013)
S37	136	Lab.	39.48	5.38	—	—	55.16	13.57	(Szilágyi , 2013)
S38	167	Lab.	39.95	7.49	—	—	34.69	11.08	(Szilágyi , 2013)
S39	130	Lab.	48.79	2.57	—	—	66.74	10.78	(Szilágyi , 2013)

3.3.2 Synthetic data

Another data type adopted in this study is the synthetic data. These data are generated by Monte-Carlo simulation using the procedure proposed by (Breyse, 2012; Breyse & Martínez-Fernández, 2014). The main interests of synthetic simulation are: (a) its ability to consider many influencing factors and its flexibility in changing the values of these factors, (b) its ability to generate a huge number of (NDT measurement, compressive strength) pairs within a synthetic building while it is impossible with real building due to the cost and time constraints, (c) the ability to repeat the simulations many times in order to reduce the influence of chance, and (d) the easiness to quantify the error in the estimated strength since the “true in-situ strength” is known at each test location in the synthetic building and consequently the difference between this value and estimated value can be calculated. However, as stated in the previous subsection, for real experimental datasets given in Table 3.1, the core strength (or standard specimen strength for the laboratory datasets) is used as a true reference to calculate the error.

The basic idea herein is to simulate statistically a synthetic building having the following information at each test location:

- True in-situ strength and NDT values,
- Measured values for core strength and NDT measurements,

Generally, the main steps in the simulation are: defining the simulation input variables and their domains, generating random values for the input variables using the suitable probability distributions, calculating the values of the output variables using the synthetic relationships between the inputs and outputs and finally repeating the simulation many times.

(Breyse, 2012) considered the true in-situ strength, f_{ist} , and concrete moisture content, S_r , as simulation inputs. Of course, other possible influencing factors (i.e. carbonation, cracking, aggregate type and size, etc.) may also exist. However, it is very complex (if not impossible) to consider all the influencing factors.

Regarding the generation of the input variables, there is no general consensus about the statistical distribution of concrete strength in a structure, (ACI 214.4R, 2003; Bungey, et al., 2006) stated that the distribution is normal when the control is excellent while for poor quality concrete (high concrete variability) the distribution is lognormal. However, (Celik, et al., 2012) studied the strength distribution obtained from various diameter cores and they indicated that Weibull distribution for almost all of the core sample data were found to characterize the compressive strength well. In the present study, the true in-situ concrete strength f_{ist} is generated by assuming a normal distribution $N(\bar{f}_{ist}, s(f_{ist}))$ while a truncated normal distribution, $N(\bar{S}_r, s(S_r))$ with $0 \leq S_r \leq 100\%$, is used to generate the values for the degree of saturation S_r .

After an in-depth literature review of available experimental results, (Breyse, 2012) developed synthetic models in order to correlate the inputs with the true in-situ ultrasonic pulse velocity V_t and true in-situ rebound number R_t :

$$V_t = V_{ref} (f_{ist} / f_{cref})^{1/bf} (S_r / S_{rref})^{1/b_s} \quad (3.1)$$

$$R_t = R_{ref} (f_{ist} / f_{cref})^{1/cf} (S_r / S_{rref})^{1/c_s} \quad (3.2)$$

It must be pointed that these models (Equations 3.1 and 3.2) do not pretend to be “general models” and they can be considered as models that only represent a case study from many case studies existing in the real practice.

The reference values ($V_{ref}, R_{ref}, f_{cref}, S_{rref}$) in Equations (3.1) and (3.2) are arbitrary values introduced in order to normalize the equations, and have no influence on the general behavior. The reference values are $R_{ref} = 40$, $V_{ref} = 4000$ m/s, $S_{rref} = 85\%$ and $f_{cref} = 40$ MPa. The exponents quantify the relative sensitivity of V_t and R_t to the variations in strength and humidity. Their values have been carefully chosen in order to accurately describe what is observed in practice. The strength sensitivity exponents bf and cf have been respectively taken equal to 4.90 and 2.10. The humidity sensitivity exponents bs and cs have been respectively taken equal to 7.14 and -3.33 (Breysse & Martínez-Fernández, 2014). The bs and cs values respectively correspond to an increase of 6% in pulse velocity and a decrease of 12% in rebound number as the concrete humidity changes from air-dry specimen (assumed at $S_r=65\%$) to fully saturated condition ($S_r=100\%$). These values are in agreement with what we have found in literature. For example, for pulse velocity, (ASTM C 597, 2002) stated that the pulse velocity in saturated concrete may be up to 5% higher than in dry concrete. (Lencis, et al., 2013) found an increase of 19% in pulse velocity between totally dry and maximum saturated conditions i.e. 6.65% between $S_r=65\%$ and saturated conditions. Experimental results from (Kheder, 1999) showed an increase in pulse velocity of (200–400) m/s between air dry and wet conditions. Also for rebound number, the results from (Kheder, 1999) showed a decrease of (3–4) points in rebound values i.e. less than 10%. (Malhotra, 2004) stated that well-cured, air-dried specimens, when soaked in water and tested in the saturated surface-dried condition, show rebound readings 5 points lower than when tested dry.

As it was illustrated in Chapter 2, measurement uncertainties always exist. Therefore, random errors (\mathcal{E}_V , \mathcal{E}_R and \mathcal{E}_f) are added to the generated true in-situ values (V_t , R_t , and f_{ist}) in order to produce what will be measured values (V , R , and f_c) that corresponding to what would be obtained in the real practice after a routine investigation program. These last values represent the simulation outputs.

$$V = V_t + \mathcal{E}_V \quad (3.3)$$

$$R = R_t + \mathcal{E}_R \quad (3.4)$$

$$f_c = f_{ist} + \mathcal{E}_f \quad (3.5)$$

The magnitude of these errors are obtained by assuming a normal distributions $N(0, sd_V)$, $N(0, sd_R)$ and $N(0, sd_f)$ with zero average and standard deviations, sd_V , sd_R and sd_f , represent the within-test variability of the measurements. The possible range of within-test variability has been widely documented in the literature (ACI 228.1R, 2003; Breysse, 2012; Szilágyi, et al., 2014). Depending on the values of sd_V , sd_R and sd_f (Breysse, et al., 2017) classify the measurements to three quality levels: high (HQ), average (AQ) and low (LQ). The values of sd_V , sd_R and sd_f are respectively 50 m/s, 1 unit, and 1 MPa for high quality measurements while the values corresponding to average quality are 100 m/s, 2 units, and 1.5 MPa. For low quality the values are 200 m/s, 4 units, and 2 MPa.

Each data simulation produces the outputs (V , R , and f_c) corresponding to one test location, therefore it should be repeated a number of times equal to the required total number of test locations (dataset size) in the synthetic building. Figure 3.2 shows the steps of data simulation used to generate a synthetic dataset.

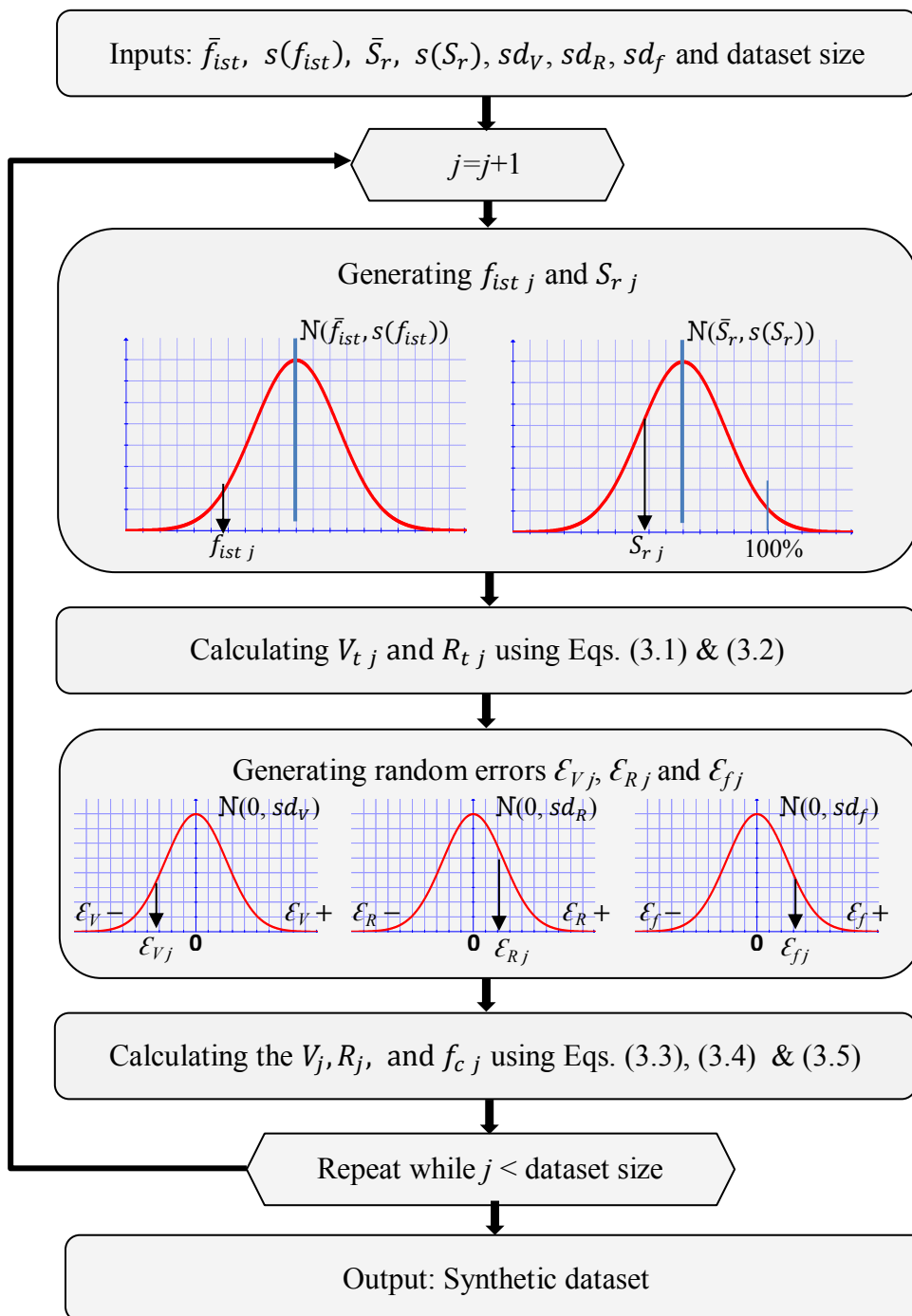


Figure 3.2 Algorithm for generating dataset using synthetic simulation

3.4 The simulator developed in the present thesis

To perform the analysis in the present study, a simulator was built. It was coded using the programming language VBA (Visual Basic for Applications) that integrated with Excel. This simulator has two main objectives:

- Generating the synthetic dataset following the algorithm illustrated in Figure 3.2,
- Simulating and analyzing the assessment strategy defined in Figure 3.1 using synthetic or real experimental datasets.

The flowchart in Figure 3.3 shows the main algorithm in this simulator. It begins with the selection of the dataset (NDT measurements, strengths from destructive tests DT) that will be used through the following steps of the simulation. As detailed in Section 3.3, two main sources of data are used in the present study: experimental (either from laboratory studies or from in-situ investigation) and synthetic datasets. Therefore, the simulator has been designed to deal with any of these options. If the synthetic dataset option is selected, the simulator follows the algorithm in Figure 3.2 to generate the dataset. Regarding this issue, two choices were considered, see Figure 3.3: (a) generating only one dataset and using it for all simulation repetitions, (b) generating a new dataset for each new simulation repetition. The first choice represents the simulation of one given building, while the second choice is devoted to simulate many buildings having the same statistical properties. The second choice will be applied to Section 4.3 where the efficiency of specific assessment strategy (i.e. all the elements of strategy remain unchanged through all simulation repetitions) is required to be analyzed. Whilst, the first choice will be used through all other sections and chapters of this thesis. When the experimental dataset option is selected, an Excel sheet is used to enter the dataset into the simulator. In this case, the same dataset will be used through all simulation repetitions (i.e. similar to the first choice for the case of synthetic data).

As discussed in Section 3.2, planning any on site investigation program requires defining the NT test locations for carrying out NDT measurements. The selection of these locations can be “random” or “predefined” by the investigator to follow specific constraints. To reproduce this feature, these two options are considered in the simulator. Regarding the further selection of NC core locations from the NT test locations, the simulator can deal with the three options presented in Section 3.2: random, predefined, and conditional.

In the real practice, after the in-situ measurements and strength testing in laboratory, these results are used to identify a conversion model between the NDT measurements and concrete strengths. This stage is reproduced in the simulator with several options regarding the choice of the model type and that of the model identification approach used to identify the model. In each simulation, the simulator can identify simultaneously several models using different identification approaches. Each identified model is used to estimate the local strengths which are used to calculate the estimated mean strength, \bar{f}_{cest} , estimated strength standard deviation (strength variability), $s(f_{cest})$, and errors. The error is calculated from the differences between the estimated local strengths, f_{cest} , and the corresponding “true in-situ strengths, f_{ist} , for the case of synthetic dataset. For the experimental dataset, the true in-situ strengths remain unknown and consequently the true reference strengths, f_c , provided by the dataset will be used for the calculation of error.

Due to the measurement uncertainties and/or the random components of the process illustrated in Figure 3.3 (i.e. the random selection of the test locations), the estimated values are not unique (i.e. repeating the simulation leads to new estimated values). Therefore, to reduce the influence of chance and obtain the relevant statistical information, the simulator is designed to repeat the simulation a certain number of times. The number of simulation repetitions (NI) is an input data for the simulator and the value is chosen in each analysis in order to ensure the stability of the results within a specified confidence level (see Appendix C).

The process can be applied for a single value of NC that is provided to the simulator or for a series of NC values and in the latter case the whole process is repeated for each value of NC as shown in Figure 3.3, where NC_{max} is an input value that represents the upper limit of the NC values.

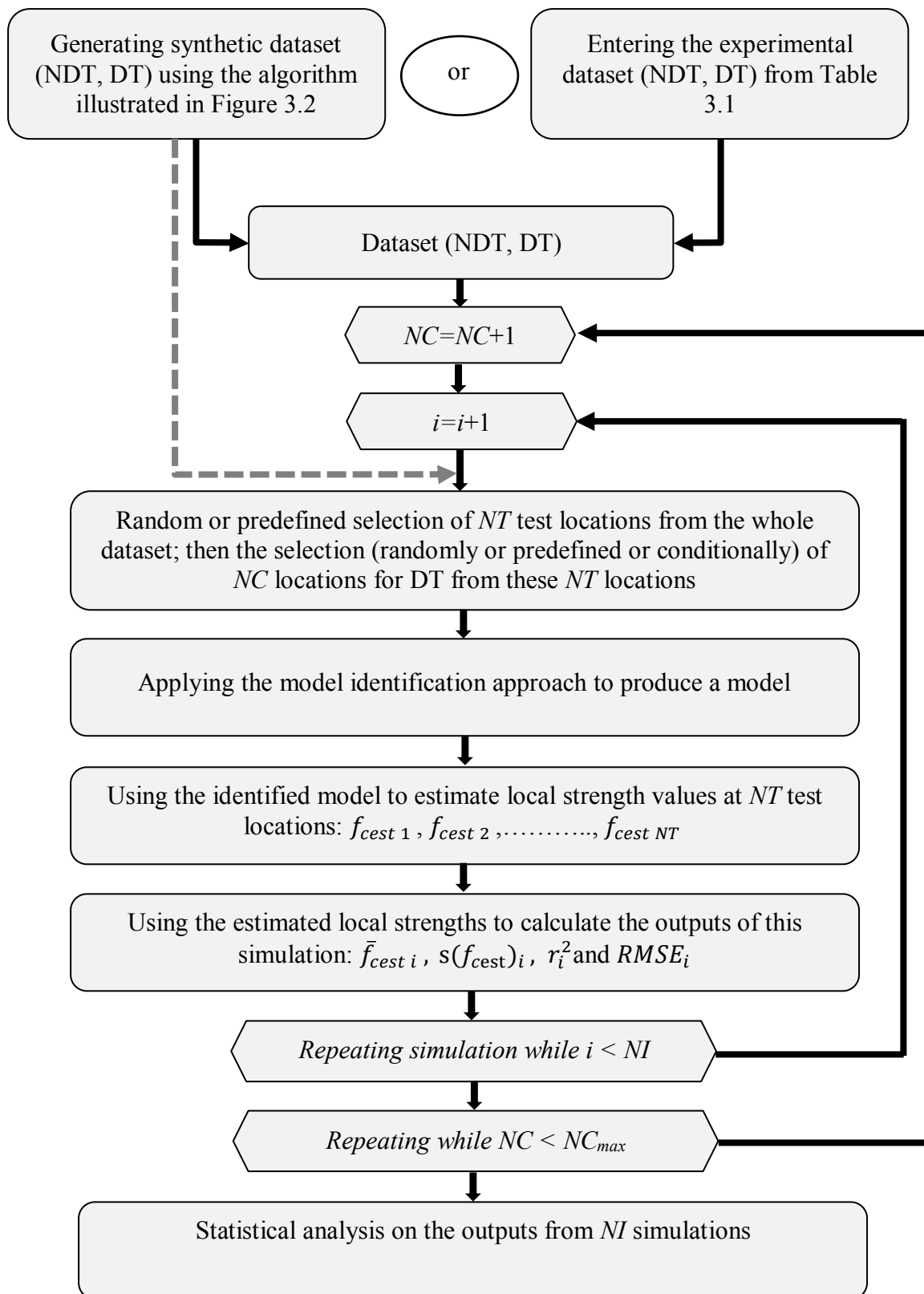


Figure 3.3 Main algorithm of the simulator prepared for the analysis in the present study

For each NC value, the simulator stores the results from each simulation, therefore, after NI simulations, the simulator outputs (for each used model identification approach) will be NI values of \bar{f}_{cest} , $s(f_{cest})$, and errors. These outputs are post-processed in different ways through the chapters of this thesis. In Section 4.2, the focus is on the average values of errors in the estimated local strengths. In Section 4.3 all the estimated mean strength and the estimated strength standard deviation (strength variability) values resulting from each strategy are used for the comparison with the stored true in-situ values and with the results from other strategies. In Section 4.4 and Chapter 5, the average and standard deviation of the NI -values of the estimated mean strength and the estimated strength standard deviation are derived. Finally, for the reliability analysis (Chapter 6), the outputs are used to plot the cumulative distribution curves of the estimated mean strength and the estimated strength standard deviation. These curves are used for constructing what we will call “risk curves”.

The analysis results of Sections 4.2 and 4.3 obtained from using the first version of the simulator show that for several simulation repetitions one of or both the situations shown in Figure 3.4 are occurred. Figure 3.4a shows that, for dataset having the true strengths and pulse velocities (shown in gray), when selecting randomly four cores ($NC=4$) from this population one of the possible choices is the four points shown in black. After identifying a conversion model using these NC pairs of (V, f_c) , this model predicts a zero value and several negative values of the estimated local strengths when it is used to assess the strength values of the dataset under consideration. While Figure 3.4b shows another possible choice for four cores that leads to identify a model has a negative slope for the case of linear model or (negative exponent for power model). This means that the strength decreases while the NDT value increases, which generally has no physical meaning with ultrasonic pulse velocity and rebound hammer measurements. Therefore, the NC pairs of (x, f_c) which are used for the model identification may produce a very bad or even physically impossible conversion model.

The risk of having the above situations is larger if NC is small, and if the measurement uncertainties are relatively large as compared to the range of variation of concrete strength in the domain of investigation.

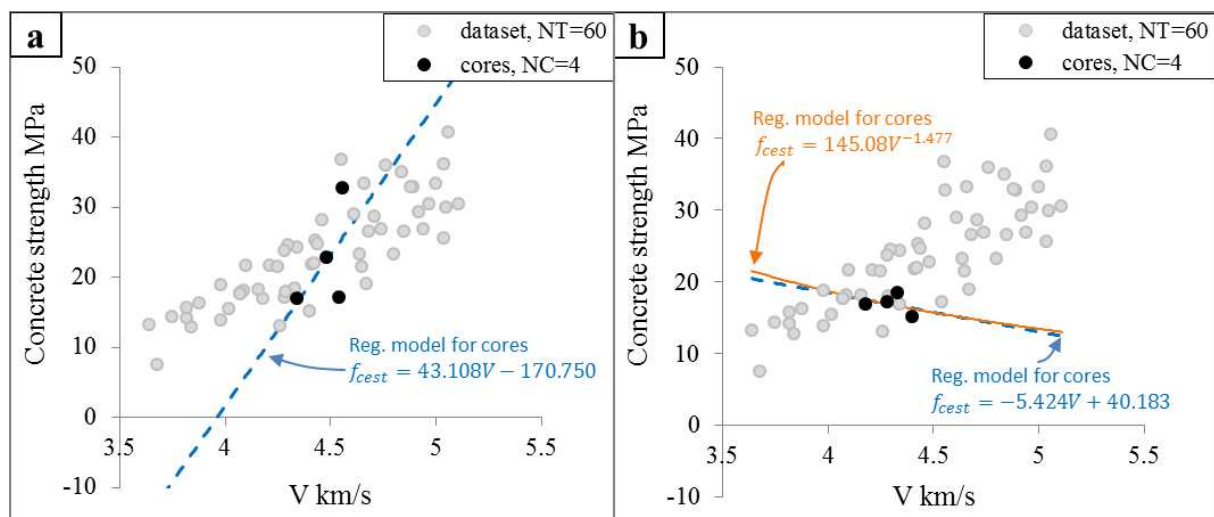


Figure 3.4 Two cases identified wrong conversion models obtained from the analysis of real dataset using $NC=4$ cores selected randomly from the whole dataset size=60

In this study, another version of the simulator has been developed. It includes (in addition to the characteristics described for the first version) two conditions (if $f_{cest} \leq 0$ or if slope < 0) for linear model and (if $f_{cest} \leq 0$ or if exponent < 0) for power model. For each model type, when one or both of these conditions is satisfied, the simulator will reject this model and return to select new test locations for cores in order to identify another model. Consequently, this version of the simulator will repeat the simulation more than NI repetitions in order to provide the results for NI repetitions all having correct models. Counters are added to the simulator in order to quantify the number of failed repetitions (i.e. the repetitions identify wrong models) for each approach. The percentage of failed repetitions for each approach with respect to the total number of repetitions will be one of the simulator outputs. This issue will be discussed in detailed in Chapter 5.

Since the first version of the simulator is used to obtain the analysis results of Sections 4.2 and 4.3, the second version of the simulator is used for carrying out the analyses through the other parts of this thesis (i.e. Section 4.4, Chapter 5, and Chapter 6).

CHAPTER FOUR

ANALYSIS OF CURRENT METHODOLOGY FOR CONCRETE STRENGTH ASSESSMENT

4.1 Introduction

This chapter is devoted to analyze the current methodology for assessing the concrete strength in the existing structures presented in Chapter 2. It includes three main parts. The first one focuses on studying the effects of several methodology inherent characteristics (controlled factors, see Subsection 2.4.4). The second part of this chapter analyzes and compares several assessment strategies presented in an international benchmark in order to provide some remarks about the efficient strategy when the investigation budget is fixed. The third part analyzes the existing model identification approaches (regression and calibration) and compares their efficiency for assessing the mean strength and concrete strength variability.

4.2 Studying the effect of several key influencing factors

In this section, the effects of several methodology inherent characteristics (controlled factors, see Subsection 2.4.4) are studied. The considered factors are: the number of test locations for cores, NC , which are used for the regression/calibration of the model, the quality of NDT measurements (or within-test variability), the type of model, and the combination of NDT techniques. In addition to these factors, the uncontrolled factor of the in-situ concrete strength variability is also considered. For analyzing all these effects, all sources of data illustrated in Chapter 3 (real in-situ data, laboratory studies data, and generated synthetic data) are considered. In order to assess the quality of fitted model and its ability to estimate strength, $RMSE$ and r^2 errors are calculated for the estimated local strengths. We have published this study in “Construction and building Materials” Journal with an analysis based on synthetic data only (the paper is available in Appendix A). We will synthesize here the simulation procedure, the main results of the analysis (for all sources of data considered in this study), and the final conclusions.

4.2.1 Effect of number of test locations for cores

As stated in Chapter 2, the fitting error is commonly used as an indicator regarding the adequacy of the fitted model. To study the effect of number of test locations for cores, NC , on the fitting error in the local strengths, the in-situ dataset S1 (selected from Table 3.1) was analyzed. To this end, the process illustrated in Figure 3.3 was followed with $NT=205$ and NC varies from 2 to 20. After the random selection of NC from NT test locations, the regression analysis was used to identify three linear models corresponding to three cases: using pulse velocity method V , using rebound hammer method R , and using combined method ($V+R$). Then, r^2 and $RMSE$ were calculated using respectively Equations (2.14) and (2.17). This process was repeated NI times ($NI=1000$). Therefore, for each NC , the final results were the average of NI values of r^2 and the average of NI values of $RMSE$. Figure 4.1a shows the average of NI values of r^2 as a function of NC and the averages of NI values of $RMSE$ were

plotted on Figure 4.1b. Each figure shows three Fitting Error Curves (FEC) corresponding to three cases (single technique V , single technique R and combination of $V+R$). From Figures 4.1a and 4.1b, the following observations can be noted:

- the general features are identical for the three cases (V and R measurements alone or in combination),
- it is clear that the r^2 values decrease (i.e. fitting error increases) as NC increases. The same observation is drawn with the fitting error $RMSE$ where $RMSE$ increases as NC increases. This is rational because the number of points to be fitted using a model having a fixed number of parameters is increased.
- the other interesting observation is about the points with $r^2=1$ or $RMSE=0$ when $NC=2$ for the case of single technique and $NC=3$ for the case of combination of NDT techniques. This means that, when NC equal to the number of model parameters, the model parameters can be identified without any fitting error.

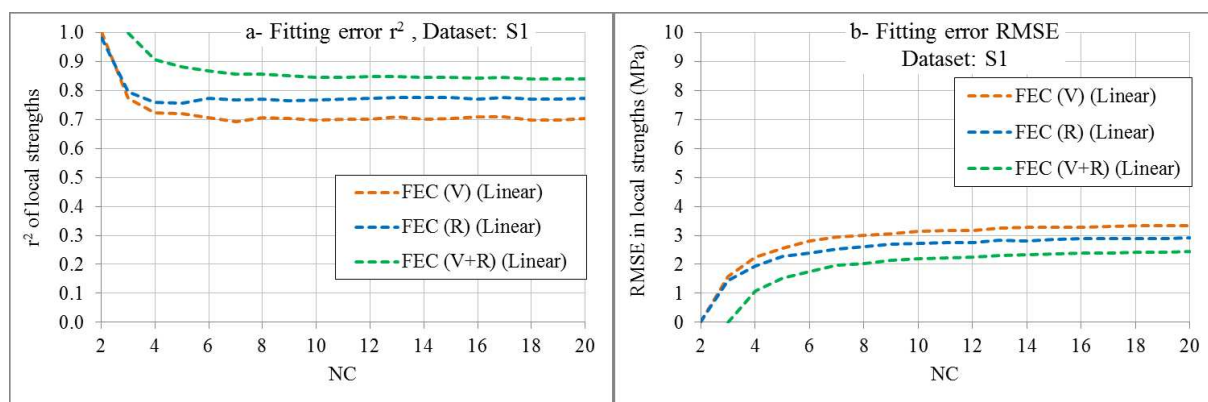


Figure 4.1 Fitting error, FEC, as a function of NC resulted from analyzing the dataset S1

In real practice, the model identified for the test locations with cores is used to predict the strength at test locations where there are only NDT measurements. Therefore, the prediction ability of the model is an important issue. Statistic provides many measures of forecasting, see (Hyndman & Koehler, 2006; Shcherbakov, et al., 2013). Among these measures, $RMSE$ is used by the NDT experts for evaluating the prediction error (Huang, et al., 2011; Nobile, 2015). The prediction error ($RMSE$) can be calculated for $(NT-NC)$ test locations using the following modified form of Equation (2.17):

$$RMSE = \sqrt{\sum_{i=1}^{NT-NC} (f_{cest i} - f_{c i})^2 / (NT - NC)} \quad (4.1)$$

Following the same process used to calculate and plot the FEC, the Prediction Error Curves (PEC) obtained from analyzing the dataset S1 were identified and plotted on Figure 4.2. For comparison, the FEC presented in Figure 4.1b were also plotted on Figure 4.2. From Figure 4.2, the following observations can be noted:

- the general features are identical for the three cases (V and R measurements alone or in combination),
- the prediction error exhibits an adverse pattern to that of fitting error, since it decreases while NC increases: by increasing NC , despite of increasing fitting error, the model provides a better picture of true strengths. Thus, by considering only the fitting error, one would obtain a totally wrong picture,

- one must point here that when NDT strength assessment is studied in the scientific literature, authors often limit their analysis to the first (fitting) stage, considering the r^2 and $RMSE$ from fitting as a criterion of the assessment quality. However it is clear that the values of PEC are always larger than the corresponding values of FEC, therefore using the fitting error is misleading,
- for each of the three cases (V and R measurements alone or in combination), the difference between the FEC and PEC decreases as NC increases,
- for the cases where the fitting errors = 0 ($NC=2$ for the case of single technique and $NC=3$ for the case of combination of NDT techniques) the prediction errors are large i.e. the models have poor predictive ability. This discrepancy must be pointed, since a very low number of cores is a common practice. Scientific literature offers many cases in which the model in such a situation is said to be good, this statement being just based on r^2 fitting value, when the predictive ability would in fact be very poor, but has not been quantified.

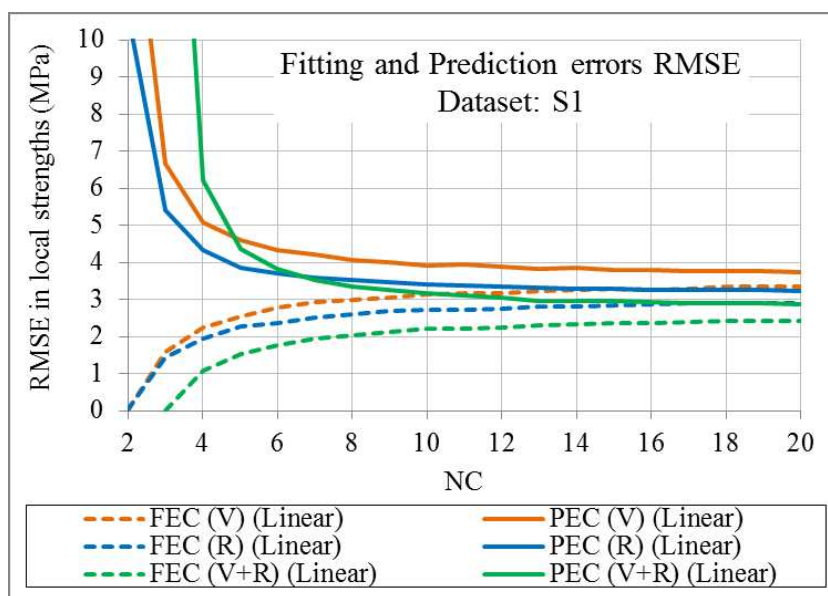


Figure 4.2 Fitting and prediction errors (FEC and PEC) as a function of NC resulted from analyzing the in-situ dataset S1

4.2.2 Effect of quality of measurements

To study the effect of within-test variability (or repeatability) on the prediction error in the estimated local strengths, a synthetic dataset was considered. Following the algorithm illustrated in Figure 3.2, the simulator generated a synthetic dataset having the following characteristics: $\bar{f}_{ist} = 45$ MPa, $s(f_{ist}) = 5$ MPa, $\bar{S}_r = 80\%$, $s(S_r) = 5\%$. Then, the algorithm in Figure 3.3 was followed in order to produce the final results (at each NC value, the average of NI -values of $RMSE$ from NI -repetitions), for more details about this simulation see Appendix A. As stated in Chapter 3, according to the amount of within-test variability, (Breyse, 2012) has classified the quality of measurements into three levels: high (HQ), average (AQ) and low (LQ). Figure 4.3 shows the prediction error curves, PEC, that correspond to these three quality levels obtained from the case of using the pulse velocity method with cores for identifying linear regression models. It is clear that for the same concrete, using low quality measurements (higher within-test variability) produces larger prediction error in the estimated strengths. This confirms the need to control the measurement uncertainty in order to improve the quality of assessment.

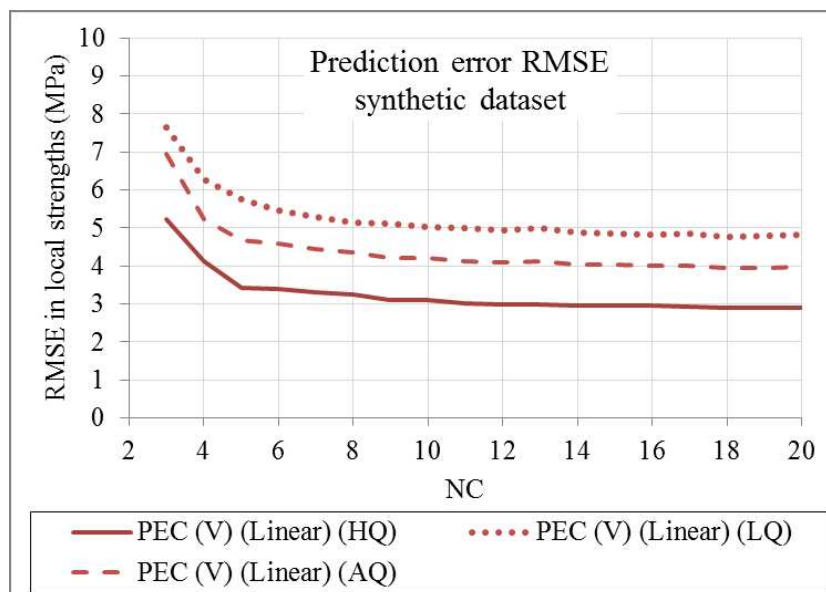


Figure 4.3 Effect of quality of measurements on the Predictive RMSE, PEC, after analyzing a synthetic dataset ($\bar{f}_{ist} = 45 \text{ MPa}$, $s(f_{ist}) = 5 \text{ MPa}$) with single NDT technique (pulse velocity)

4.2.3 Effect of in-situ concrete strength variability

The in-situ concrete strength variability is one of the inherent properties of concrete that depends on many factors like the mixing procedure, curing, and compacting processes in addition to the environmental effects. To study its effect on the final prediction error resulted in the estimated strengths, the simulator was used to generate (following the algorithm in Figure 3.2) and analyze (following the algorithm in Figure 3.3) ten synthetic datasets. These datasets have the same characteristics ($\bar{f}_{ist} = 45 \text{ MPa}$, $\bar{S}_r = 80\%$, $s(S_r) = 5\%$, and AQ measurements) except the $s(f_{ist})$ value which was varied from $s(f_{ist}) = 1 \text{ MPa}$ for the first dataset to $s(f_{ist}) = 10 \text{ MPa}$ for the last one. For all datasets, the core locations were selected randomly, while NC remains constant with ($NC=10$), and the regression approaches were used to identify linear models. After the NI repetitions for each dataset, the final predictive $RMSE$ values (the average of NI values of $RMSE$ from NI repetitions) were plotted on Figure 4.4a for the ten datasets (for more details about this simulation, see Appendix A). Three cases for NDT measurements were considered in Figure 4.4: V and R measurements alone or in combination.

It is obvious that, for the three cases, the increase in the in-situ strength variability leads to a simultaneous increase in predictive $RMSE$. However, the behavior appears to be different when the $RMSE/s(f_{ist})$ values (i.e. relative error) are plotted for the ten datasets since the relative error decreases as $s(f_{ist})$ increases (Figure 4.4b). This indicates that the predictive $RMSE$ consists in two parts: the error due to the in-situ concrete variability, and the error due to the other sources of uncertainties. This observation coincides with our earlier discussion about the sources of uncertainty, see Subsection 2.4.5. In short wording, one must pay attention to how the quality of the assessment is quantified: looking at absolute error $RMSE$ or relative error $RMSE/s(f_{ist})$ would lead to different conclusions.

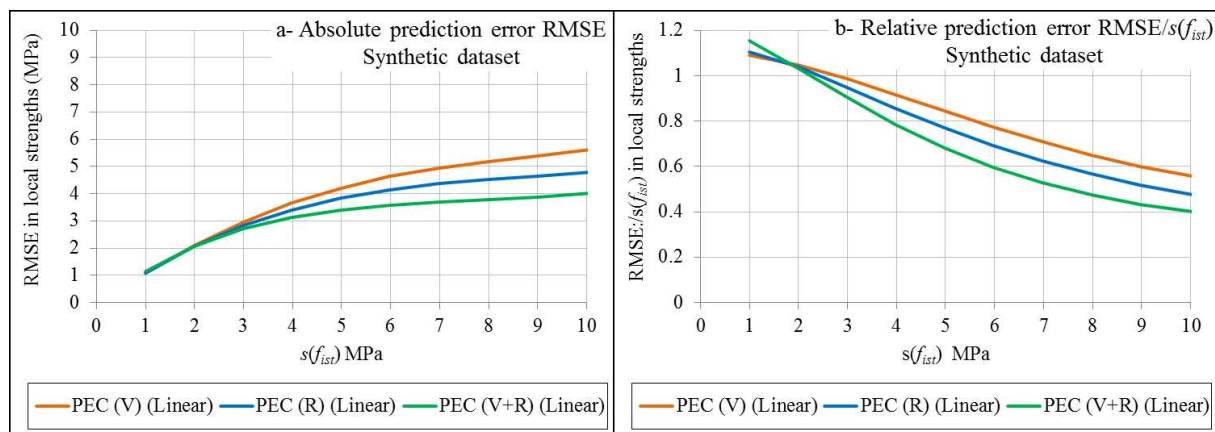


Figure 4.4 Effect of true in-situ strength variability on the predictive *RMSE* for single and combined NDT techniques with average qualities and for $NC=10$

4.2.4 Effect of type of model (linear or nonlinear)

In the above subsections the regression approach was used to identify linear models. This section is devoted to study the effect of model type by comparing the predictive *RMSE* curves, PEC, obtained from nonlinear models (power model) with those corresponding to linear models. In order to analyze this issue comprehensively and provide more robust conclusions, four real datasets (S1, S10, S11, and S12) were selected from Table 3.1. They were selected due to the variety in their characteristics: they are from different sources of data (S1 obtained from in-situ study while the others belong to laboratory studies), they have a wide range of mean strengths (ranging from 18.8 to 51.5 MPa) and different concrete strength variabilities (coefficient of variation ranging from 17.5 to 34.5 %), see Table 3.1. Moreover, these datasets provide the test results for both NDT techniques (rebound hammer and pulse velocity) in addition to the core strengths. Consequently, the effects of the type of NDT technique and of the use of single or combination of NDT techniques can be considered in this analysis.

The methodology illustrated in Figure 3.3 was applied for each dataset and for three cases (single technique *V*, single technique *R*, and combination of *V+R*) with *NT* equals to the dataset size, as given in Table 3.1. *NC* varies from 4 to 20 with random selection of these locations. For each dataset, for each *NC*, the regression approach was used to identify simultaneously linear and power models. After *NI* repetitions, the final predictive *RMSE* curves that correspond to each dataset were plotted on Figures 4.5-4.8 respectively. Each point on these curves represents the average of *NI* values of *RMSE* from *NI* repetitions. From these figures, the following observations can be noted:

- for the case of single NDT technique, for each technique, the curves for linear and power models are very close. However, there is generally a significant divergence between these curves when $NC < 5$. Furthermore, for the curves corresponding to the case of pulse velocity technique shown in Figure 4.7, some divergence exists even when $NC \geq 5$. This behavior is due to the effect of measurement uncertainty (or quality of measurement) associated of the NDT technique because in the same figure the PEC curves (for linear and power models) for rebound hammer show no divergence.
- for the case of combination of NDT techniques, it is clear that, when $NC > 5$, there is no significant effect of the model type. However, when $NC < 5$, the PEC curves of power models diverge from the curves corresponding to linear models.

A more detailed focus on the case where the curves diverge (typically at low NC , $NC < 5$) shows that the power models lead to higher $RMSE$ values. This means that, when $NC < 5$, the probability to identify models that are not well representative to the concrete under consideration is higher for the case of power model. This is due to the effect of the uncertainty in the values of power model parameters and more specifically in the exponent of the power model. This uncertainty can lead, at the prediction stage, to strength estimations significantly far from true values. However, as NC increases, the difference between linear and power models vanishes and when $NC > 5$, the type of model has generally no significant effect on the prediction error in the local strengths.

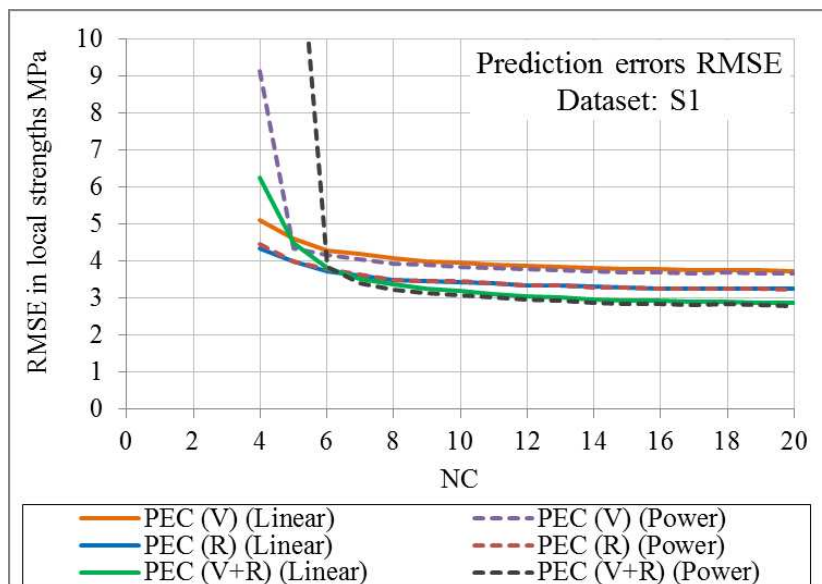


Figure 4.5 Effect of model type on the predictive RMSE for single and combined techniques when analyzing in-situ dataset S1

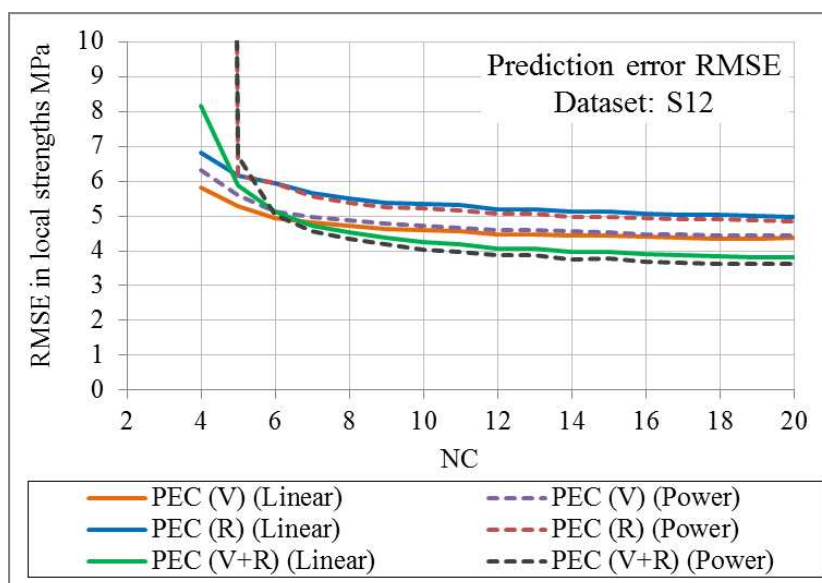


Figure 4.6 Effect of model type on the predictive RMSE for single and combined techniques when analyzing laboratory study dataset S12

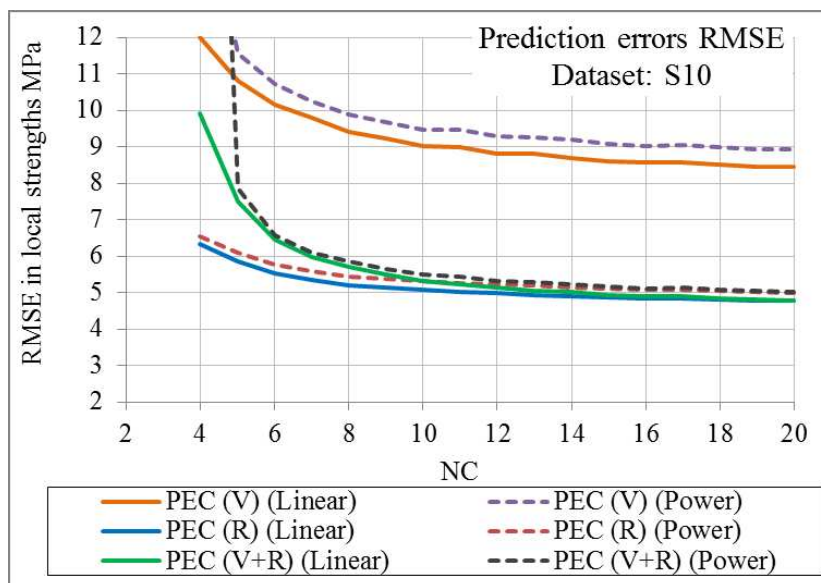


Figure 4.7 Effect of model type on the predictive RMSE for single and combined techniques when analyzing laboratory study dataset S10

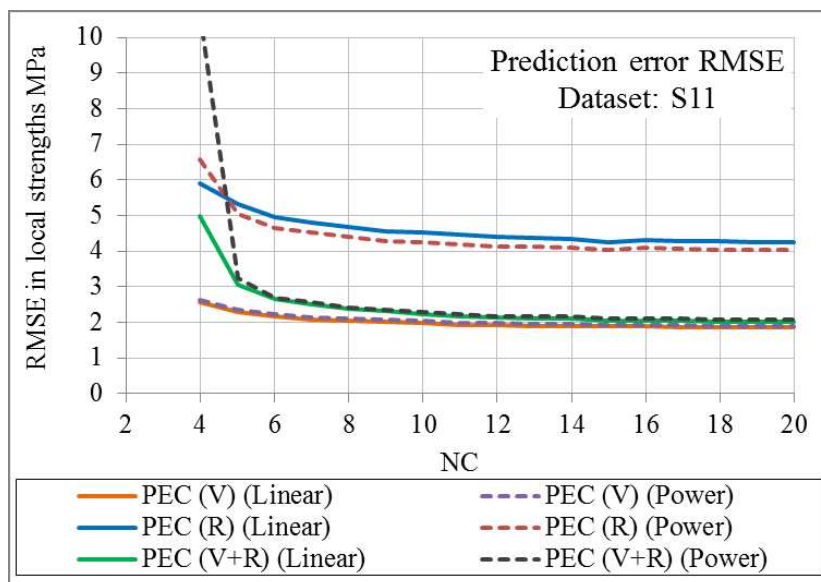


Figure 4.8 Effect of model type on the predictive RMSE for single and combined techniques when analyzing laboratory study dataset S11

4.2.5 Effect of combining NDT techniques

The aim of combining NDT techniques is to improve the quality of assessment. The efficiency of combination points out whether using combined techniques is more efficient than using a single NDT technique or not. The efficiency of combination studied herein is limited to the case of combining the rebound hammer and pulse velocity and any other combination is out of our scope. In real practice, in order to study the efficiency of combination of (V+R), the coefficient of determination r^2 for the model derived from the combination of NDT techniques is usually compared with those corresponding to the models established from using these techniques separately. As discussed in Chapter 2, due to statistical aspects, r^2 for combination is always, at the fitting stage, higher than that resulting from using the NDT techniques alone. Therefore r^2 can be a misleading indicator about the efficiency of combination and may lead to wrong conclusions. Moreover, from the above

discussion in Subsection 4.2.1, it is clear that when r^2 is large (i.e. small fitting error), the prediction error is large, which also highlights the shortcoming of r^2 as an indicator about the efficiency of combination. Consequently, comparing the prediction capacity (prediction error) of the model identified for combination with that identified for each NDT technique separately is the only way to decide whether using combination of NDT techniques can improve the quality of assessment or not.

To emphasize on this issue, thirteen in-situ and laboratory studies datasets (S1, S2, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, and S16) were selected from Table 3.1 to be analyzed here. The methodology illustrated in Figure 3.3 was followed with NT equals to the dataset size and a fixed value of NC ($NC=6$). For each dataset, for each simulation, the NC was selected randomly in order to identify a linear regression model for each of three cases: single technique pulse velocity, single technique rebound hammer, and their combination. At the end of each simulation, the fitting errors, r^2 , and the prediction errors, $RMSE$, were calculated for these three cases. The simulation was repeated NI repetitions ($NI=1000$) and the average of NI values of r^2 and the average of NI values of $RMSE$ corresponding to each case are given in Table 4.1.

The results at fitting stage show that the r^2 for combined method is, as expected, larger than that for single techniques. However, the analysis of $RMSE$ values at predictive stage confirms that this first statement is clearly meaningless. Depending on the considered datasets, $RMSE$ can be either better or poorer for combination than for single technique. It is only in three cases (datasets S8, S9, and S16) that the combination has a positive effect on the quality of assessment (i.e. it is efficient). Consequently, the prediction error must be used as the only indicator about the efficiency of combination.

Table 4.1 Studying the efficiency of combination, bold numbers indicate the minimum $RMSE$ values corresponding to each dataset

Dataset symbol	NT	NC	Fitting error in terms of r^2			Prediction error in terms of $RMSE$ (MPa)		
			Single V	Single R	Combined $V+R$	Single V	Single R	Combined $V+R$
S1	205	6	0.71	0.74	0.86	4.3	3.8	3.8
S2	23	6	0.64	0.23	0.70	3.9	5.3	4.6
S6	18	6	0.27	0.31	0.43	2.5	2.5	3.5
S7	18	6	0.57	0.78	0.84	6.8	4.9	6.2
S8	16	6	0.79	0.68	0.92	8.2	9.2	5.8
S9	14	6	0.81	0.88	0.97	3.6	2.2	1.6
S10	80	6	0.51	0.83	0.87	10.1	5.6	6.5
S11	63	6	0.84	0.54	0.89	2.2	5.1	2.7
S12	60	6	0.69	0.64	0.83	5.0	6.0	5.2
S13	40	6	0.79	0.97	0.98	8.0	2.5	3.1
S14	30	6	0.24	0.20	0.41	13.5	13.3	16.3
S15	20	6	0.73	0.77	0.89	5.7	4.0	4.9
S16	16	6	0.68	0.76	0.93	9.0	7.3	5.6

To study the efficiency of combination, Figures 4.5-4.8 can be good examples in which the prediction errors corresponding to single and combination of NDT techniques are plotted. The following observations can be noted:

- the prediction errors due to the case of combination reduce as the NC increases, however, when NC is larger than 10 or 12 the curves stabilize and there is no significant reduction in the prediction errors,
- from Figures 4.5 and 4.6, when $NC > 6$, the case of combination gives the minimum error as compared with the cases of single NDT techniques. In contrast, when $NC < 6$, using single NDT technique (rebound hammer in Figure 4.5 and pulse velocity in Figure 4.6) is the best,
- from Figures 4.7 and 4.8, it is clear that the combination cannot assess the strengths better than the single NDT techniques for all values of NC considered in this analysis. From Figures 4.7, it is obvious that PEC of pulse velocity technique is significantly higher than that corresponding to rebound hammer technique. Since for each dataset (each figure) the effects of the sources of uncertainty other than the measurement uncertainty are the same (the same concrete, sample size, NC , etc.) on both techniques. Therefore the difference between PEC of pulse velocity technique and that of rebound hammer technique is mainly due to measurements uncertainty (or the quality of measurements). For Figure 4.8 the same issue can be observed but with higher PEC is provided by the rebound hammer technique. Consequently, the reason of the inefficient behavior of combination shown in Figures 4.7 and 4.8 is due to the difference in quality of measurements of the combined two techniques,
- Some researchers find that the pulse velocity method is more efficient than the rebound hammer method or versus versa. However, the comparison of the curves of single technique show that the rebound hammer technique is the best in some cases (Figures 4.5 and 4.7) while in other cases the pulse velocity is the best (Figures 4.6 and 4.8). Therefore, we think that the superiority of any of the two techniques is affected considerably by the quality of measurements.

In summary, the efficiency of combination must be studied using the prediction error. The combination of pulse velocity and rebound hammer techniques is not always efficient and this efficiency depends on number of test locations for cores and the quality of measurements. This explains the reasons why in the literature the efficiency of combining NDT is controversial (combination is said to bring some added-value in some cases but not in others).

4.2.6 Conclusions

In order to understand the effects of several factors which affect the assessment of concrete strength in existing structures by NDT measurements a parametric study has been performed using both synthetic and real datasets. Through this study several factors remained unchanged, such as: all models were identified using regression approach and cores were selected randomly. From this analysis, the following conclusions can be drawn:

- errors in the estimated strengths are affected by the number of test locations for cores, NC . For a small NC , the fitting error is small while the prediction error is large, but increasing NC leads to an increase in the fitting error while the prediction error is progressively reduced. These errors depend on the quality of measurement as, for the same NC , they decrease with better quality of measurements. The prediction error is always larger than the fitting error, therefore fitting error is a misleading indicator

about quality of assessment and prediction error must be used for this purpose. In practice, the prediction error can be estimated by using additional cores.

- as the true in-situ concrete strength variability increases, $s(f_{ist})$, the predictive *RMSE* increases while the ratio $RMSE/s(f_{ist})$ decreases. Consequently, one must pay attention to how the quality of the assessment is quantified: looking at the absolute error, *RMSE*, or relative error, $RMSE/s(f_{ist})$, would lead to different conclusions.
- the effect of the model type (linear or power) depends mainly on *NC*. It can be significant for a small *NC* but it reduces as *NC* increases,
- the efficiency of combination must be studied using the prediction error (and not the fitting error). The combination of pulse velocity and rebound hammer techniques is not always efficient and the efficiency depends on *NC* and the quality of measurements. This explains the reasons why in the literature the efficiency of combining *V* and *R* is a controversial issue, combination being said to bring some added-values in some cases but not in others. Finally, it is important to point out that this conclusion is related to the combination of *V* and *R* and cannot be generalized to others combination methods.

4.3 Analyzing several assessment strategies presented in an international benchmark

In common practice, the real challenge for engineers is the selection of an assessment strategy among a large number of alternative strategies using the same investigation budget. Therefore, we will focus here on this issue by analyzing several strategies proposed by NDT experts, presented in an international benchmark, in order to assess concrete strength in a synthetic building (synthetic dataset). The properties targeted in this simulation are the concrete mean strength and concrete strength variability. This work has been published in “Materials and Structures” Journal (see the paper in Appendix B). Therefore the present section synthesizes the simulation methodology, the main results, and the final conclusions while more detailed information is available in Appendix B.

4.3.1 The benchmark in brief

A RILEM Technical Committee TC-249 ISC (nondestructive in-situ strength assessment of concrete) was created with the aim of establishing guidelines for an efficient use of nondestructive techniques for in-situ concrete strength assessment. TC members have decided to carry out a benchmark whose objectives were to identify and compare the expert practices regarding NDT assessment of concrete strength. The main elements of this benchmark have been detailed in (Breysse, et al., 2017). In this benchmark, depending on the amount of available resources, three knowledge levels (KL1, KL2 and KL3) were considered. Sixteen different strategies have been proposed by the benchmark contributors in order to assess the concrete strength of a synthetic building. Each strategy has provided two values for each knowledge level:

- the estimated mean strength,
- the estimated strength standard deviation (concrete strength variability).

For knowledge level KL3, these two values, as provided by each strategy, are plotted on Figure 4.9.

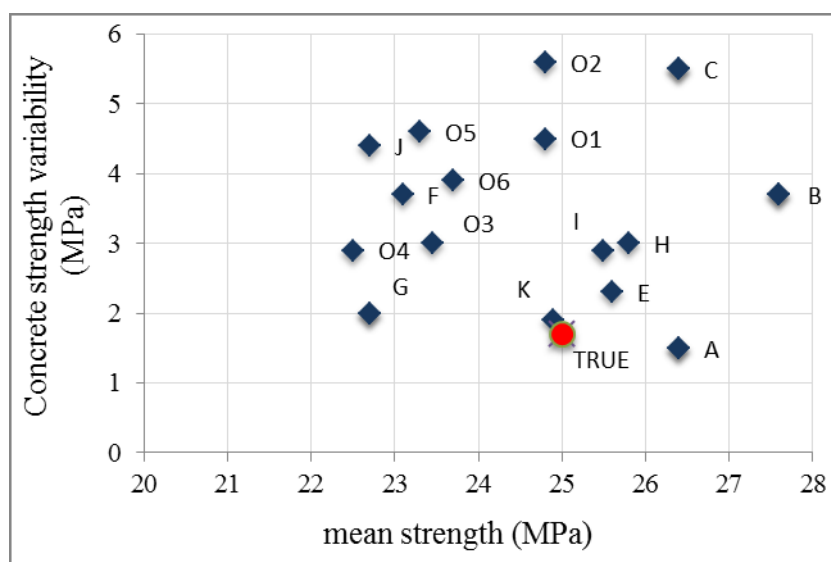


Figure 4.9 Estimation of mean strength and strength standard deviation (concrete variability) for 16 strategies at KL3 (Breysse, et al., 2017)

4.3.2 Simulation of the assessment strategies

In fact, if the same strategy is applied to another building having the same properties (statistically speaking) different estimated values of mean strength and standard deviation would be derived. They correspond to a different point on Figure 4.9. Hence, a general conclusion about the efficiency of any strategy cannot be simply drawn from this single result because the effect of chance. Therefore, using Monte-Carlo simulations appears to be an adequate approach in order to assess and compare the efficiency of several strategies proposed in the benchmark. Each simulation corresponds to a different building (but having the same statistical properties) on which the strategies are carried on.

Five strategies have been selected among those proposed by the contributors to the benchmark. The selection criterion was the exhaustiveness of information they had provided about the process (i.e. about how the measured values are exactly processed and how estimations are derived), and the fact that the process would be fully automatic without any expert interaction. These five strategies (namely B, E, F, J and O5 in (Breysse, et al., 2017)) have the advantage of covering the three main possibilities regarding the conversion model identification:

- using a prior model without any calibration (strategy F),
- calibrating a prior model by using NDT test results and core strengths (strategies E, J and O5),
- fitting a specific model between NDT test results and core strengths using the regression analysis (strategy B).

Three of these strategies use a single NDT (rebound R for strategies J and O5, ultrasonic velocity V for the strategy B), while strategies F and E use combined NDT. Each of the five selected strategies has been reproduced within the simulator so that it works as proposed by the contributor.

The simulation process is exactly the same as that presented and detailed in (Breysse, et al., 2017). The only difference is that the process is repeated herein NI times ($NI=1000$), thus generating NI synthetic buildings instead of a single one. This means that the algorithm presented in Figure 3.3 is applied with the option of generating a new synthetic dataset for

each simulation. Therefore, after NI simulations, for each strategy, the results are NI pairs of true in-situ values of mean strength and concrete strength variability, $(\bar{f}_{ist}, s(f_{ist}))$, and the corresponding estimated NI pairs, $(\bar{f}_{cest}, s(f_{cest}))$. Since for the synthetic building the true in-situ strengths are known, so the estimated values will be compared with the true in-situ strengths.

As an example, the results obtained from NI simulations of strategy B are shown in Figure 4.10, while the details of the five strategies with the final results are given in Appendix B. Looking only at the results provided by the contributor of strategy B shown in Figure 4.9, one could conclude that this strategy is not efficient because its point is far from the true point. However, when looking at Figure 4.10, different conclusion can be reached, since the center of the estimated cloud is close to the center of true cloud. An additional important issue is the scatters of both clouds (those of true values and those of estimated values) which cannot be captured when only a single simulation is available. This observation emphasizes our above statement that one single application of an assessment strategy is not enough to deliver a relevant indication about the efficiency of the strategy.

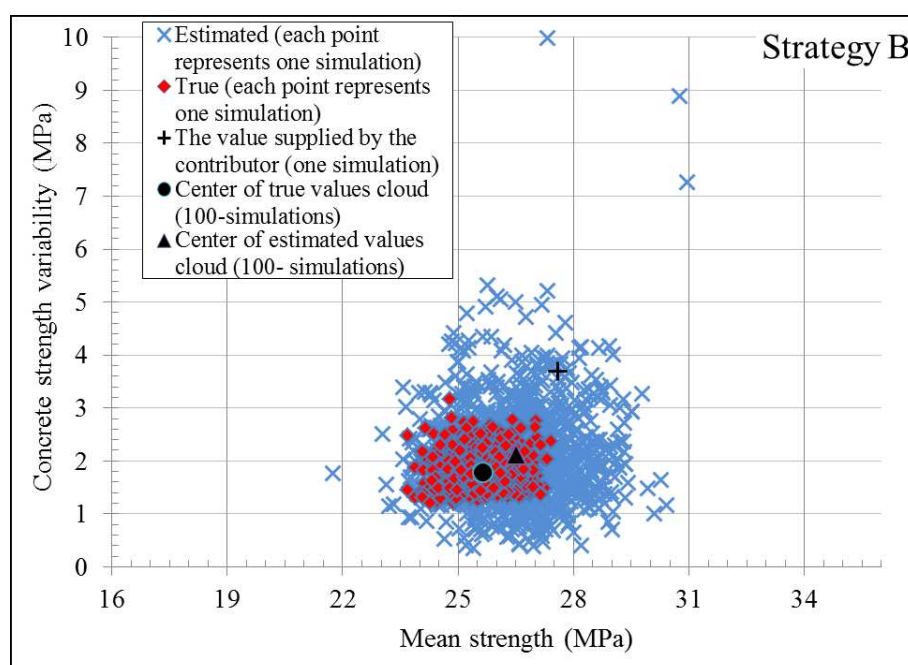


Figure 4.10 Comparison of estimated (mean strength and strength standard deviation) to the true values; results from 1000-simulations of contributor's strategy B at KL3

4.3.3 Analysis of simulation results

4.3.3.1 Indicators of estimation quality

In order to evaluate the quality of assessment (or the efficiency) of each strategy, two types of indicators are considered. They respectively correspond to (a) the difference between true in-situ values and estimated values, (b) the scatter of estimated values.

The first indicator is the *NRMSE* (Normalized Root Mean Square Error) (Ris, et al., 1999) which is devoted to quantify the accuracy of forecasting. For each estimated cloud, two values of *NRMSE* are calculated regarding mean strength and strength standard deviation, respectively:

$$NRMSE_{av} = \left(\sqrt{\sum_{i=1}^{NI} (\bar{f}_{cest\ i} - \bar{f}_{ist\ i})^2 / NI} \right) / \bar{f}_{ist} \quad (4.2)$$

$$NRMSE_{sd} = \left(\sqrt{\sum_{i=1}^{NI} [s(f_{cest})_i - s(f_{ist})_i]^2 / NI} \right) / \overline{s(f_{ist})} \quad (4.3)$$

where \bar{f}_{ist} and $\overline{s(f_{ist})}$ are respectively the averages of NI values of true in-situ mean strength and true in-situ strength standard deviation (i.e. the center of true cloud). For the presented case study, the values of \bar{f}_{ist} and $\overline{s(f_{ist})}$ are respectively equal to 25.6 MPa and 1.8 MPa.

The second indicator quantifies the horizontal and vertical scatters of the estimated cloud i.e. the standard deviation of the mean strengths $s(\bar{f}_{cest})$, and the standard deviation of the strength standard deviations $s(s(f_{cest}))$. The values corresponding to the true cloud are $s(\bar{f}_{ist}) = 0.66$ MPa and $s(s(f_{ist})) = 0.30$ MPa. The scatters of the true cloud correspond to the randomness of the simulation process, for a given set of concrete properties.

Table 4.2 summarizes the main features of the selected five strategies and the values of the four indicators. The efficiency of the strategy increases as normalized indicators decrease and as the two other indicators are closer to true in-situ values.

Table 4.2 Characteristics of the five strategies at KL3 and values of the quality indicators

Str.	Q	No. of measurements			Way of core select.	Model identification approach	Quality indicators			
		Core	V	R			$NRMSE_{av}$ %	$NRMSE_{sd}$ %	$s(\bar{f}_{cest})$ MPa	$s(s(f_{cest}))$ MPa
B	LQ	5	17	0	CC	Regression analysis	6	54	1.30	0.92
E	AQ	3	6	12	PC	calibration (k-method)	5	55	1.45	0.57
F	AQ	3	9	12	PC	No calibration	4	116	0.93	0.61
J	AQ	3	0	29	PC	calibration (k-method)	8	129	2.22	0.64
O5	AQ	2	0	40	CC	Calibration (Δ -method)	14	121	1.78	0.48

Q: quality of measurements (LQ: low quality, AQ: average quality, HQ: high quality)

PC: core locations predefined before the starting of the investigation program

CC: conditional core locations i.e. defined after the first series of NDT measurements

4.3.3.2 Comparison of the strategies efficiency

From the results synthesized in Table 4.2, regarding the estimation of mean concrete strength, all strategies can be considered efficient since the normalized error $NRMSE_{av}$ is less than 10% (strategies B, E, F, J) and 15% for the strategy O5. However, some strategies are not robust, with a significant probability of having a large over or under estimation, directly linked to the $s(\bar{f}_{cest})$ value. This problem is particularly relevant with Strategy J and, to a lower extent, with Strategy O5. Strategy F appears to be the most efficient because it has the minimum

values of $RMSE_{av}$ and $s(\bar{f}_{cest})$ but this strategy deserves further comments. It must be pointed that it uses an uncalibrated model. With such a method, the uncalibrated model can in some situations (by the effect of chance) more or less fit with the true properties, but it can also be very far from these true properties. As discussed in Chapter 2, the issue of calibration has been widely documented (IAEA, 2002; EN 13791, 2007) and uncalibrated models are risky. Therefore the most efficient strategies regarding mean strength estimation are strategies B and E.

Regarding the strength variability estimation, there is a general tendency to overestimate, with particularly very bad indicators for strategies F, J and O5, for which the overestimation exceeds 100% (notwithstanding with a large scatter). Strategies E and B get better results, with a lower overestimation. This is probably the beneficial consequence of options privileged in these approaches: (a) the combination of two NDT for strategy E, (b) the conditional coring, using 5 cores, and the weighted average of strength estimates for strategy B. However, the scatter $s(s(f_{cest}))$ is larger for strategy B due to the uncertainty arising from using low quality measurements. Nevertheless, with $NRMSE_{sd}$ values always larger than 50%, these strategies are not fully satisfactory regarding the estimation of concrete strength variability.

4.3.3.3 Analyzing the effects of several factors

The analysis of the five strategies highlights the effects of several influencing factors on the quality of assessment. We will analyze in more details these effects, by considering them in a more systematic way and focusing on three items: the quality of measurements, the way of selecting the location of cores and the amount of resources allocated to the investigation program. The general idea is to keep all main features of the strategies while simply varying one of these influencing factors and analyzing how the quality of the assessment is affected. Table 4.3 summarizes all results, which can be compared to those obtained with original strategies (Table 4.2).

- Effect of amount of resources

Three cost levels (KL1, KL2 and KL3) have been considered in the benchmark (Breyse, et al., 2017) corresponding to a progressively increasing amount of resources. All results discussed above (in this section) had been obtained at KL3. It seems interesting to see how the quality of assessment varies when KL changes. Therefore two strategies (B, E) have been selected to be analyzed at level KL1 (amount of resources at KL1=1/3 KL3). The numbers of measurements are globally divided by a factor 3 and are provided in Table 4.3, which can be compared to the numbers for the same strategies at KL3 in Table 4.2. All other characteristics (quality of measurements, the way of core selection and model identification approach) of these two strategies remain unchanged.

The final results presented in Table 4.3 show that reducing the amount of resources from KL3 to KL1 increases significantly all the metrics ($NRMSE_{av}$, $NRMSE_{sd}$, $s(\bar{f}_{cest})$ and $s(s(f_{cest}))$). This behavior is expected since reducing the number of measurements (DT and NDT) causes the uncertainties to grow up. For example, strategy B herein is affected by a dramatical increase in the model uncertainty due to the fact that the two parameters of the regression model are derived from information gathered on only two cores instead of three (Strategy B) or one core instead of two (Strategy E). The result of the identification is thus highly affected by the uncertainties.

Table 4.3 Characteristics of the selected strategies for the parametric study and values of the quality indicators (in bold: characteristics differ from those in Table 4.2)

Str.	Q	No. of measurements			Way of core select.	Model identification approach	indicators of quality			
		Core	<i>V</i>	<i>R</i>			$NRMSE_{av}$ %	$NRMSE_{sd}$ %	$s(\bar{f}_{cest})$ MPa	$s(s(f_{cest}))$ MPa
Studying the effect of amount of resources (KL1)										
B	LQ	2	4	0	CC	Regression analysis	8	114	2.01	2.03
E	AQ	1	2	6	PC	calibration (k-method)	10	59	2.71	0.86
Studying the effect of quality of measurements (KL3)										
B	AQ	3	11	0	CC	Regression analysis	7	82	1.60	1.47
E	HQ	2	4	8	PC	calibration (k-method)	4	28	1.24	0.54
J	HQ	2	0	20	PC	calibration (k-method)	7	54	1.59	0.48
Studying the effect of way of core selection (KL3)										
E	AQ	3	6	12	CCA	calibration (k-method)	5	55	1.40	0.57
J	AQ	3	0	29	CCA	calibration (k-method)	7	134	1.48	0.62
O5	AQ	2	0	40	CCA	Calibration (Δ -method)	6	121	1.68	0.48

- Effect of quality of measurements

In the benchmark three quality levels of measurements are considered. These levels are: high quality (HQ), average quality (AQ) and low quality (LQ) with the values of within-test variability as that provided in Chapter 3. In order to further study this effect, three strategies (B, E and J) have been selected among the five strategies considered in this section. According to the rules of the benchmark, the cost of measurements increases as the quality level improves. The available amount of resources being fixed, the number of measurements reduces when the quality level improves. For the selected three strategies, the quality level has been improved (from LQ to AQ for strategy B and from AQ to HQ for strategies E and J). Consequently the numbers of measurements have been reduced accordingly. These numbers are provided in Table 4.3 and can be compared with the initial values given in Table 4.2.

From Table 4.3, it is clear that improving the quality of measurements has a slightly positive effect on the indicators of the mean strength, $NRMSE_{av}$ and $s(\bar{f}_{cest})$, for strategies E and J, while this effect is negative for strategy B. This positive effect is more remarkable for the concrete strength variability: for strategies E and J, $NRMSE_{sd}$ reduces to one half their values in Table 4.2 while the reduction in $s(s(f_{cest}))$ is small. On the contrary, for strategy B, $NRMSE_{sd}$ and $s(s(f_{cest}))$ significantly increase in comparison with the corresponding values given in Table 4.2. These apparently inconsistent results can be explained by the controversial

effects of improving the quality of measurement on one hand while the number of tests is reduced on the other hand. Depending on which influence is the more important, the result may be either positive for strategies (E, J) or negative for strategy B. The measurement uncertainty has a predominant effect for strategies E and J in which the quality increased from average to high. However, for strategy B, improving the quality of measurement (from low to average) reduced the number of cores from 5 to 3 and this second factor had a predominant effect (it can be noted that this strategy is based on a regression model which requires to identify two model parameters). These results confirm that capturing the coupled effects of measurement uncertainty and number of cores will deserve a further interest.

- Effect of the way of core selection

Two ways for selecting the core locations have been considered by the benchmark contributors: predefined core locations “PC” and conditional core locations “CC”. The idea of the conditional cores is to select core locations that cover, as much as possible, the whole range of the NDT measurements distribution. Thus we propose the following rule for selecting core locations: (a) rank all NDT test results from the lowest to the highest, (b) subdivide the set into NC subsets, (c) take a core at the location which has the NDT value closest to the median value of the subset. This rule will be denominated “CCA” in this section. To study the effect of the way of core selection on the quality of estimation, three strategies (E, J and O5) have been selected. The contributors of strategies (E, J) used the PC option, while contributor O5 used CC but following a different rule (see Appendix B). The proposed CCA rule will be applied to these three strategies. The other characteristics of the strategies remain unchanged including the number of measurements, since CCA implies no additional cost.

It appears that conditional cores can improve both the accuracy and precision (reducing the scatter) of the mean strength estimation while it generally has no effect on the accuracy and precision of the estimated concrete variability. Moreover, for the strategy E that already obtained good results, it has effect neither on mean strengths nor on concrete variability estimations. The main effect is the decrease of the horizontal scatter of the cloud of the estimated strength, which is particularly visible for strategy J. The explanation is that when one has a small number of cores ($NC = 3$ here), the model uncertainties resulting from a too narrow range of variation of the NDT test results may be very large. Conditional cores, by ensuring a better coverage of the strength range reduce these uncertainties. Since conditional cores induce no additional cost and since it can have only positive effect, this process must be strongly recommended.

4.3.4 Conclusions

In this section, several strategies defined by experts in an international benchmark devoted to assess the concrete strength and its variability have been selected. The efficiency of these strategies has been studied by simulating their application to a series of 1000 synthetic buildings having the same statistical properties and by defining quality indicators of the assessment, in terms of accuracy and precision.

The comparison of the simulated results with those provided by the contributors to the benchmark confirmed that it is irrelevant to evaluate a strategy from a unique result (“one shot study”). This statement is derived from synthetic datasets but it can, for the same reasons, be extended to real in-situ datasets, since the effect of chance (or lack of chance) may prevent any valid statement. This work has confirmed that:

- it is very dangerous to consider only the estimated strengths (mean value, standard deviation) without considering the variability of these estimations,

- the quality of assessment obviously improves when both quantity and quality of test results increase. However when the total amount of resources is fixed, the optimal balance between quantity and quality of test results does not follow simple rules and deserves to be studied through a more comprehensive analysis,
- properly assessing the concrete strength variability is a much more difficult challenge than assessing the mean concrete strength. All strategies considered here have shown a handicap for assessing the concrete strength variability. Thus, correctly assessing concrete strength variability remains an ambitious challenge. We will come back to this issue in the next section and in Chapter 5 of this document,
- the quality of measurements, or measurement uncertainty, is a key factor regarding the quality of assessment,
- the proposed method for the conditional selection of core locations can improve the quality of assessment without any additional cost so it is strongly recommended.

4.4 Analyzing the existing model identification approaches

As discussed in Chapter 2, the existing model identification approaches can be classified into two main categories: regression approach when a specific model is identified using the least squares method, and calibration approaches when a prior model is modified for best agreement with a given dataset. We will analyze and compare in this section the estimation capacity of the existing model identification approaches regarding the assessment of mean strength and concrete strength variability.

4.4.1 Comparing the prediction capacity of the existing model identification approaches

For this purpose, eight real datasets (S3, S5, S17, S18, S19, S20, S38 and S39) were selected from Table 3.1 with four datasets with ultrasonic pulse velocity measurements and the others with rebound hammer measurements. Datasets S3 and S5 were obtained from in-situ studies while the other six datasets resulted from laboratory studies. These eight datasets cover a wide range of mean strength (from 16.21 MPa to 66.74 MPa) and concrete strength variability (with coefficients of variation from 15.1% to 35.6%).

In this section, a linear model was adopted for the three model identification approaches. For the calibration approaches, an uncalibrated prior model is required. For each of the eight datasets considered in this section, a linear uncalibrated model was selected from literature. The selection criterion was: the model has a range of NDT values close or larger than that of the dataset. Table 4.4 summarized the selected uncalibrated model corresponding to each dataset.

Each dataset was analyzed following the procedure illustrated in Figure 3.3. Firstly, all pairs (NDT, DT) of the dataset were considered in each simulation (i.e. NT = dataset size). Secondly, NC test locations (i.e. NC pairs of (NDT, DT)) were selected randomly from the total available test locations (NT) in order to identify models. The value of NC was varied from 1 to 20, however for the analysis of some datasets (S3, S5, and S19), NC was changed from 1 to 10 due to the limited size of these datasets. Thirdly, the model identification approaches (regression, calibration with Δ -method, and calibration with k-method) were used to identify three linear models. Each model was used to estimate the local strength values (for NT test locations) and consequently to calculate the values of estimated mean strength \bar{f}_{cest} , and estimated concrete strength variability (in terms of standard deviation) $s(f_{cest})$. The whole procedure was repeated NI times ($NI=1000$). At each repetition, new test locations were selected for NC which leads to a new conversion model for each approach and consequently to new values of estimated mean strength and estimated strength variability.

Table 4.4 Uncalibrated models selected from literature for the present analysis when using the calibration approaches

Dataset symbol	Uncalibrated model	Reference of model
S18	$f_{uncal.} = 36.72V - 129.08$	(Qasrawi, 2000)
S17	$f_{uncal.} = 36.72V - 129.08$	(Qasrawi, 2000)
S19	$f_{uncal.} = 41.10V - 149.84^*$	(Musmar & Abedalhadi, 2008)
S3	$f_{uncal.} = 8.30V - 3.34^*$	(Masi, et al., 2016)
S39	$f_{uncal.} = 1.57R - 18.54$	(Al-Ameeri, et al., 2013)
S38	$f_{uncal.} = 1.57R - 18.54$	(Al-Ameeri, et al., 2013)
S20	$f_{uncal.} = 0.93R - 1.12$	(Mahmoudipour, 2009)
S5	$f_{uncal.} = 1.73R - 34.50$	(EN 13791, 2007)

*models are not provided directly in the assigned reference, however, we derived them from the data provided in the corresponding reference

Therefore, after NI repetitions, the final results for each dataset, each identification approach and each NC value were:

- the mean value and standard deviation of the estimated mean concrete strengths, i.e. respectively $\overline{f_{cest}}$ and $s(\overline{f_{cest}})$,
- the mean value and standard deviation of the estimated standard deviation values of concrete strength (concrete strength variability), i.e. respectively $\overline{s(f_{cest})}$ and $s(\overline{s(f_{cest})})$,

For each one of the eight datasets considered in this section, the values of $\overline{f_{cest}}$ that corresponding to the three identification approaches were plotted on Figure 4.11 as a function of NC values. However, the $\overline{s(f_{cest})}$ values were plotted on Figure 4.12. In addition, for comparison, reference lines (line in red) were also plotted on these figures. The reference lines represent the mean strength or the concrete strength variability values that corresponding to each dataset as provided in Table 3.1. The variabilities in mean strengths, $s(\overline{f_{cest}})$, and concrete strength variability, $s(\overline{s(f_{cest})})$ are given in Table 4.5 for two datasets, S18 and S39, (i.e. one dataset with ultrasonic velocity measurements and the other with rebound hammer measurements).

Regarding the assessment of concrete mean strength, from Figure 4.11 and Table 4.5 the following observations can be noted:

- it is obvious that all approaches are able to predict the mean concrete strength with acceptable accuracy when $NC \geq 3$ for all datasets except for the dataset S38 (Figure 4.11f) when applying the Δ -method where a larger NC is required. In fact this confirms the significant effect of the prior uncalibrated model on the final results when applying the calibration approaches: the same model was used for the datasets S38 and S39, and it was efficient with S39 (Figure 4.11e) while being less efficient with S38. However, increasing NC has a significant effect on the prediction capability of the calibration approaches.

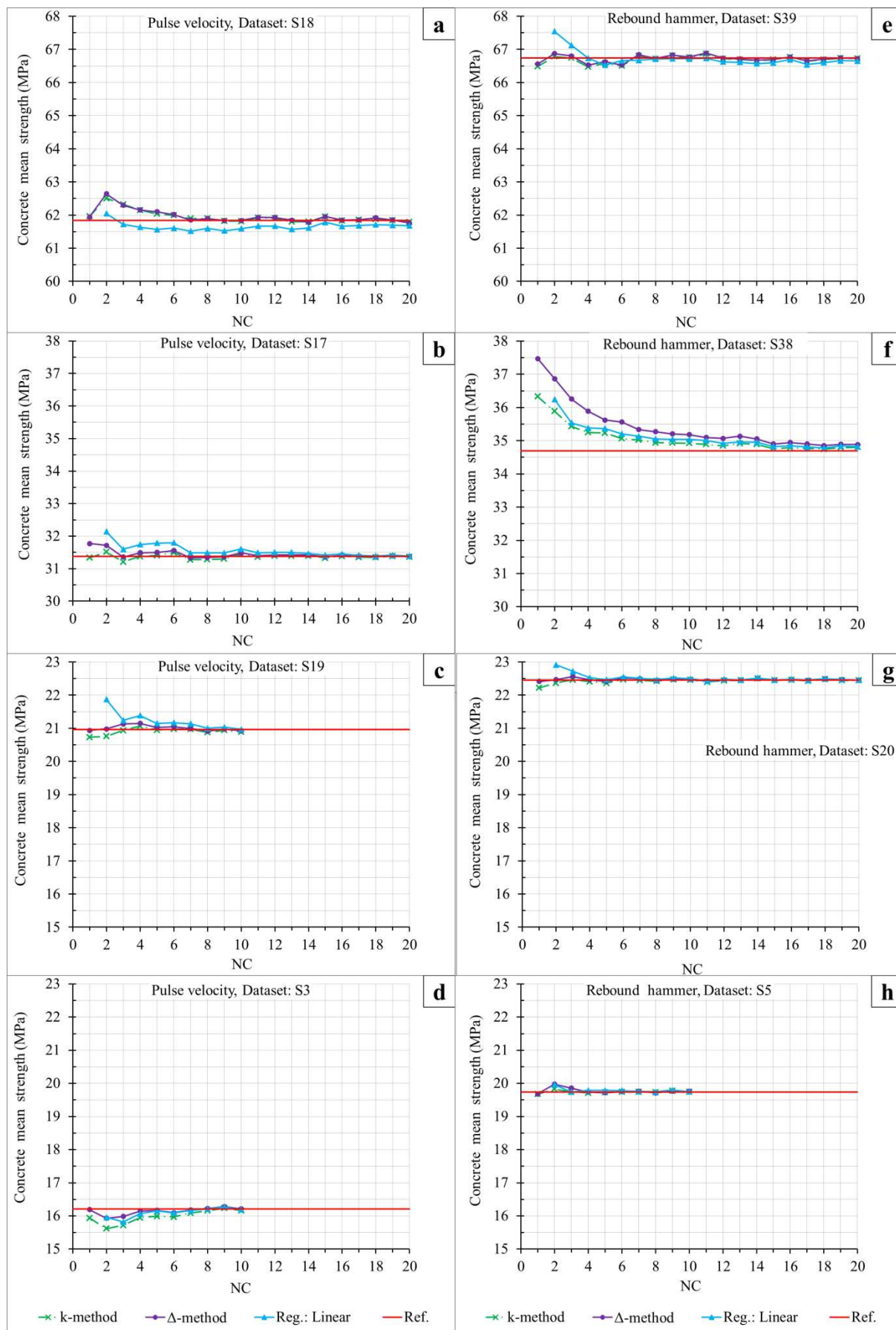


Figure 4.11 Predicting mean strength using the regression and calibration approaches

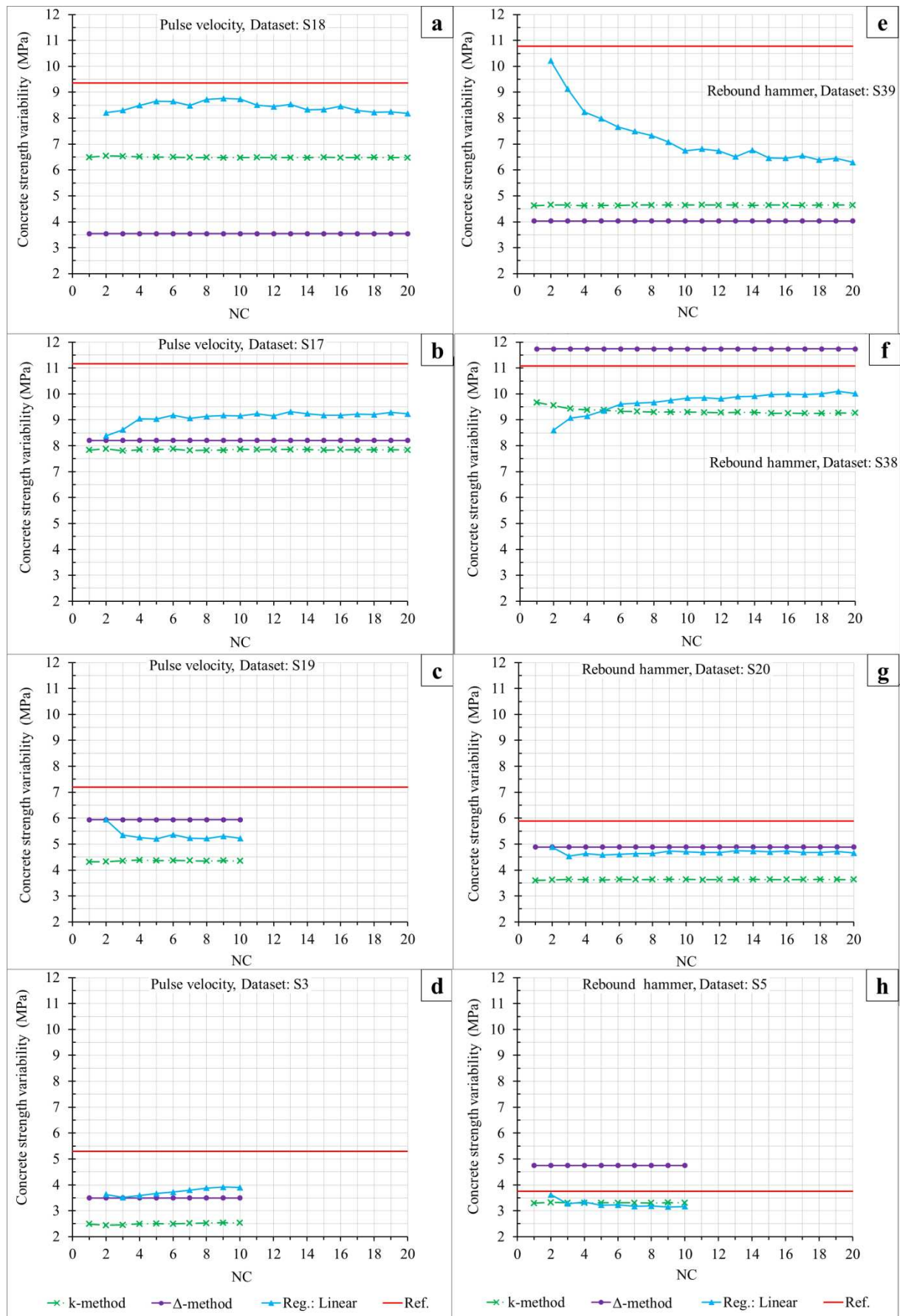


Figure 4.12 Predicting strength variability using the regression and calibration approaches

Table 4.5 Standard deviation values of NI values of mean strength and concrete variability estimated by regression and calibration approaches (all linear models) for the datasets S18 and S39 – all results are in MPa

NC	Dataset S18 (pulse velocity)						Dataset S39 (rebound hammer)					
	$s(\bar{f}_{cest})$			$s(s(f_{cest}))$			$s(\bar{f}_{cest})$			$s(s(f_{cest}))$		
	k-app.	Δ -app.	Reg. app.	k-app.	Δ -app.	Reg. app.	k-app.	Δ -app.	Reg. app.	k-app.	Δ -app.	Reg. app.
1	5.6	6.5	-	0.6	0.0	-	9.3	9.2	-	0.6	0.0	-
2	3.9	4.7	4.5	0.4	0.0	4.1	6.2	6.3	7.5	0.4	0.0	6.6
3	2.9	3.6	3.5	0.3	0.0	3.5	5.1	5.2	5.9	0.4	0.0	5.6
4	2.6	3.1	2.9	0.3	0.0	3.1	4.4	4.5	4.9	0.3	0.0	4.9
5	2.3	2.9	2.5	0.2	0.0	3.0	4.0	4.0	4.1	0.3	0.0	4.2
6	2.0	2.4	2.1	0.2	0.0	2.8	3.7	3.7	3.9	0.3	0.0	3.9
7	1.8	2.3	1.9	0.2	0.0	2.5	3.4	3.5	3.4	0.2	0.0	3.8
8	1.8	2.2	1.9	0.2	0.0	2.5	3.1	3.1	3.1	0.2	0.0	3.6
9	1.6	2.0	1.7	0.2	0.0	2.5	2.9	3.0	2.9	0.2	0.0	3.5
10	1.5	1.9	1.5	0.2	0.0	2.3	2.7	2.8	2.8	0.2	0.0	3.3
11	1.4	1.8	1.5	0.1	0.0	2.2	2.6	2.6	2.6	0.2	0.0	3.0
12	1.4	1.8	1.5	0.1	0.0	2.1	2.4	2.4	2.4	0.2	0.0	3.0
13	1.4	1.7	1.4	0.1	0.0	2.1	2.4	2.4	2.4	0.2	0.0	2.9
14	1.3	1.6	1.3	0.1	0.0	2.0	2.2	2.3	2.2	0.2	0.0	2.8
15	1.2	1.5	1.2	0.1	0.0	1.9	2.2	2.2	2.2	0.2	0.0	2.8
16	1.2	1.4	1.2	0.1	0.0	1.9	2.1	2.1	2.1	0.1	0.0	2.7
17	1.1	1.4	1.2	0.1	0.0	1.8	2.0	2.1	2.0	0.1	0.0	2.6
18	1.1	1.3	1.1	0.1	0.0	1.7	1.9	1.9	1.9	0.1	0.0	2.6
19	1.0	1.3	1.1	0.1	0.0	1.6	1.9	1.9	1.9	0.1	0.0	2.5
20	1.0	1.2	1.0	0.1	0.0	1.6	1.8	1.8	1.8	0.1	0.0	2.4

- from Table 4.5, for all approaches, for small NC , the values of $s(\bar{f}_{cest})$ are significantly high and they decrease as NC increases with more or less identical decreasing rate. Therefore, the decision about the suitable value of NC requires looking at both the values presented in Figure 4.11 and that given in Table 4.5.

Regarding the assessment of concrete strength variability, from Figure 4.12 and Table 4.5 the following observations can be highlighted:

- from the curves of the concrete strength variability estimations presented in Figure 4.12, it is clear that the regression approach underestimates the concrete variability by an amount that varies from 10% for S18 to 40% for S39, see Figures 4.12a and 4.12e respectively. It can also be noted that increasing NC has no significant effect on improving the prediction capability and may even have a negative effect as shown in Figure 4.11e,
- regarding the calibration approaches (k-method and Δ -method), it is obvious that they are unable to estimate correctly the concrete strength variability. Their behavior is

similar for the all datasets with nearly horizontal lines but with different values i.e. in some cases they overestimate the concrete variability while in the others they underestimate it. These differences are due to the prior model ability to represent the concrete under consideration. Since this ability cannot be evaluated in the real building, the efficiency of estimation by the calibration approaches mainly depends on chance,

- the $s(s(f_{cest}))$ values given in Table 4.5 decrease as the NC increases for all approaches except the Δ -method in which the values are constant (all have zero value) for all NC values. To understand this result for Δ -method, it is necessary to come back to Equation (2.20) and consider the variance of the estimated strength: $var(f_{cest}) = var(f_{c\ uncal.} + \Delta) = var(f_{c\ uncal.})$, this means that the variation of the estimated strengths is independent of Δ and consequently it is independent of cores (because in Δ -method the effect of cores is taken into account by Δ),

Finally, from Figures 4.11 and 4.12, the comparison of the assessment results obtained for the datasets with ultrasonic pulse velocity measurements (S3, S17, S18, and S19) with those produced from using rebound hammer measurements (S5, S20, S38, and S39), one can conclude that there is no significant effect of the type of NDT technique on the quality of assessment.

4.4.2 Conclusions

In this section, the prediction capacity of the calibration and regression approaches was studied. They were used for assessing the concrete mean strength and concrete strength variability (in terms of standard deviation). Eight real datasets were considered. From the final results, the following conclusions can be drawn:

- all approaches (regression, calibration) can efficiently estimate the mean concrete strength even with NC lower than the minimum number required by the standards. Furthermore, increasing NC can significantly improve the quality of assessment,
- the regression approach remains limited in capturing the concrete strength variability, it always underestimates the concrete variability and increasing NC has no significant effect on the quality of assessment. Regarding the calibration approaches, they produce unstable results (estimation efficiency varies as the prior model is changed) so their efficiency depends on chance. Moreover, increasing NC could not improve the results of calibration approaches especially for Δ -method which is proved its independence from the effect of NC ,
- there is no significant effect of the type of NDT technique (rebound hammer or pulse velocity) on the quality of assessment.

CHAPTER FIVE

DEVELOPING NEW MODEL IDENTIFICATION APPROACH: BI-OBJECTIVE APPROACH

5.1 Introduction

In this chapter, we propose a new model identification approach named “bi-objective approach” that is able to capture both the concrete mean strength and concrete strength variability. We have already published this work in “Construction and Building Materials” Journal with an analysis based on synthetic data, using a linear model only and for three targets of concrete strength characteristics (local strengths, mean strength, and strength variability). The paper is available in Appendix C. We will synthesize here the development of the proposed approach with the derivation of model parameters for both linear and power models. Then, the validity of the proposed approach will be studied by analyzing several in-situ and laboratory studies datasets. Some general conclusions will be finally derived.

5.2 Development of the bi-objective approach

From the analysis of the current assessment methodology presented in Chapter 4, it was possible to conclude that the existing model identification approaches are generally able to estimate properly the mean strength with an improving efficiency of estimation as NC increases. In contrast, regarding the estimation of concrete strength variability, the calibration approaches are unable to assess it while the regression approach remains limited.

We propose here an innovative approach, named “bi-objective”, which is devoted to capture the concrete strength variability in addition to the mean strength.

The basic idea is that any investigation program with NDT and DT (cores) techniques provides NC pairs of (x, f_c) , where x and f_c are respectively the NDT test result and core strength corresponding to one test location. These pairs are used to identify the parameters of a conversion model between the concrete strength and the NDT test results. As discussed in Chapter 2, usual mathematical shapes of such models have two parameters. It is the case for the most common ones:

- (a) linear models $f_{cest} = a + bx$,
- (b) power models $f_{cest} = a_1 x^{b_1}$, and
- (c) exponential models $f_{cest} = a_2 \exp(b_2 x)$.

Analytically, two conditions are required in order to derive the values of the two model parameters. For bi-objective approach, we consider our two objectives as the conditions for obtaining the unknown parameters, i.e. by ensuring that both mean strength and strength standard deviation are identical for true reference values and estimated ones:

$$\bar{f}_{cest} = \bar{f}_c \quad (5.1)$$

$$s(f_{cest}) = s(f_c) \quad (5.2)$$

where $\bar{f}_{cest}, \bar{f}_c$ are respectively the estimated and measured (on cores) mean strength values, while $s(f_{cest}), s(f_c)$ are respectively the estimated strength variability and the variability calculated from the core strengths. Therefore, the values for the two model parameters can be derived from applying simultaneously Equations (5.1) and (5.2). The derivation for the cases of linear and power models will be presented herein.

5.2.1 Derivation of the model parameters for the case of linear model

In order to derive the model parameters for the case of linear model the two conditions in Equations (5.1) and (5.2) are used. For a linear model:

$$f_{cest} = a + bx \quad (5.3)$$

the application of the first condition, Equation (5.1), leads to:

$$\bar{f}_{cest} = a + b \bar{x} = \bar{f}_c \quad (5.4)$$

while the second condition, Equation (5.2), is applied as follows:

$$s(f_{cest}) = s(f_c) \quad (5.5a)$$

$$s^2(f_{cest}) = s^2(f_c) \quad (5.5b)$$

$$var(f_{cest}) = var(f_c) \quad (5.5c)$$

$$var(a + b x) = var(f_c) \quad (5.5d)$$

$$b^2 var(x) = var(f_c) \quad (5.5e)$$

Consequently, the values of the unknown parameters are:

$$b = s(f_c)/s(x) \quad (5.6)$$

$$a = \bar{f}_c - b \bar{x} \quad (5.7)$$

where $\bar{x}, s(x)$ are respectively the mean and standard deviation values of x test results corresponding to the core locations.

5.2.2 Derivation of the model parameters for the case of nonlinear model

The above derivation can be extended to the case of nonlinear model using the idea of logarithmic transformation discussed in Chapter 2 for the development of the regression model. As stated in Subsection 2.4.1.2, logarithmic transformation can be applied with certain model forms like power and exponential models (for more details about these forms see (DeCoursey, 2003)). We will derive herein the parameters for the case of power model, however the same procedure can be applied for the case of exponential model. For a power model with two parameters:

$$f_{cest} = a_1 x^{b_1} \quad (5.8)$$

by taking the logarithm of both sides, we get:

$$\ln f_{cest} = \ln a_1 + b_1 \ln x \quad (5.9)$$

Clearly, Equation (5.9) is linear in its parameters and looks similar to Equation (5.3). Consequently, Equations (5.6) and (5.7) can be applied after replacing f_c and x with respectively $\ln f_c$ and $\ln x$. As a result the values of a_1 and b_1 can be calculated as follows:

$$b_1 = s(\ln f_c) / s(\ln x) \quad (5.10)$$

$$\ln a_1 = \overline{\ln f_c} - b_1 \overline{\ln x} \quad (5.11a)$$

$$a_1 = e^{\ln a_1} \quad (5.11b)$$

where $s(\ln f_c)$ and $s(\ln x)$ are the standard deviation of $\ln f_c$ and $\ln x$ values respectively, while $\overline{\ln f_c}$ and $\overline{\ln x}$ are the mean of $\ln f_c$ and $\ln x$ values respectively.

Since in the above derivation of the model parameters there are no constraints related to the type of NDT technique, the bi-objective approach can be used (from a mathematical viewpoint) with any NDT technique (rebound hammer, ultrasonic pulse velocity, etc.) which would be combined with DT technique (cores) in order to identify a conversion model.

5.3 Validation of the bi-objective approach

5.3.1 Case of linear model

In order to assess the validity of the bi-objective approach, the analysis of the existing model identification approaches (regression, calibration with Δ -method, and calibration with k-method) that was carried out in Section 4.4 was duplicated herein for the proposed bi-objective approach. The analysis has thus the following characteristics: eight real datasets (S3, S17, S18 and S19 with pulse velocity measurements while S5, S20, S38 and S39 with rebound hammer measurements), following the procedure illustrated in Figure 3.3, identifying linear models, $NI =$ dataset size, random selection of core locations, $NI=1000$, assessment of mean strength $\overline{f_{cest}}$ and concrete strength variability $s(f_{cest})$.

After NI repetitions, the final results obtained with the bi-objective approach for each dataset and each NC value are plotted on Figure 5.1 and Figure 5.2 for $\overline{f_{cest}}$ and $\overline{s(f_{cest})}$ respectively. In addition, for comparison purposes, the curves of $\overline{f_{cest}}$ corresponding to the existing identification approaches presented in Figure 4.11 are also shown in Figure 5.1 and the curves of $\overline{s(f_{cest})}$ from Figure 4.12 are plotted on Figure 5.2. Furthermore, the reference lines that represent the mean strength or the concrete strength variability values corresponding to each dataset as provided in Table 3.1 are drawn on Figures 5.1 and 5.2. The variabilities in mean strengths, $s(\overline{f_{cest}})$, and concrete strength variability, $s(s(f_{cest}))$ were given in Table 5.1 for datasets S18 and S39 (i.e. one dataset with ultrasonic measurements and the other with rebound hammer measurements). Beside the values of $s(\overline{f_{cest}})$ and $s(s(f_{cest}))$ resulted from using the bi-objective approach, the values that obtained from the regression approach (given in Table 4.5) are also provided in Table 5.1 for comparison purposes.

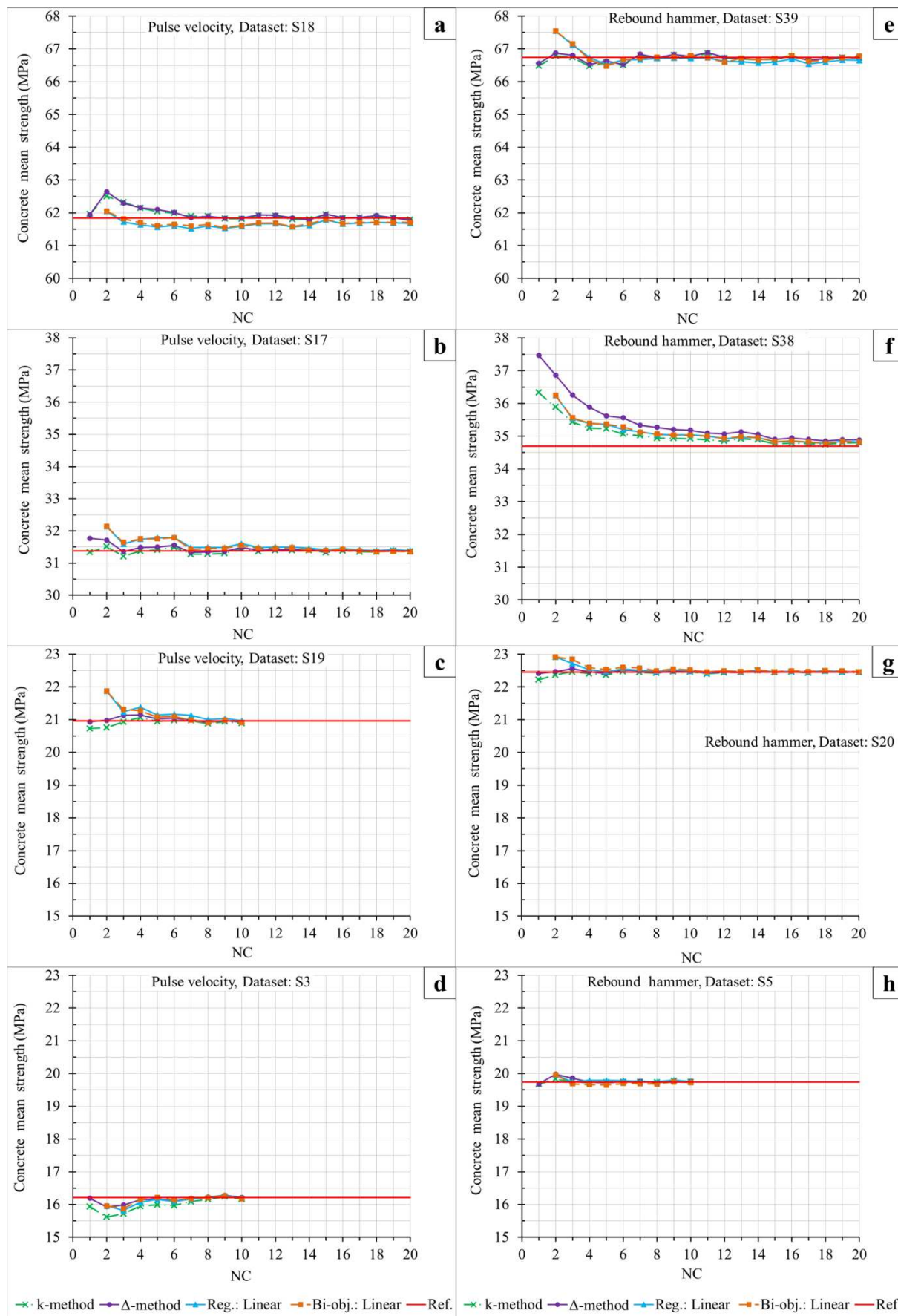


Figure 5.1 Predicting mean strength using the regression, calibration, and bi-objective approaches, all linear models

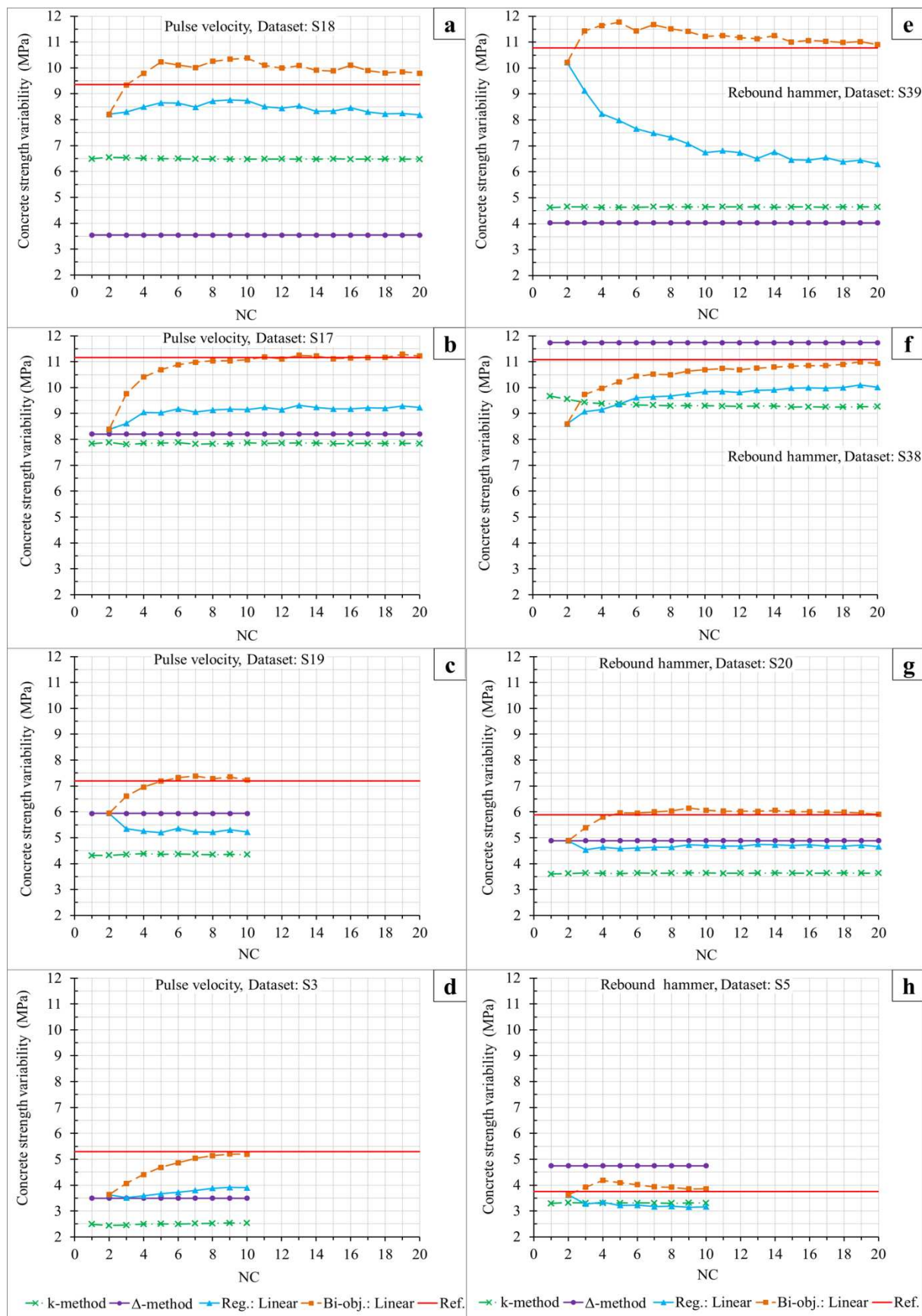


Figure 5.2 Predicting strength variability using the regression, calibration, and bi-objective approaches, all linear models

Table 5.1 Standard deviation values from NI values of mean strength and concrete variability estimated by regression and bi-objective approaches (linear models) for the datasets S18 and S39 – all results are in MPa

NC	Dataset S18 (pulse velocity)				Dataset S39 (rebound hammer)			
	$s(\bar{f}_{cest})$		$s(s(f_{cest}))$		$s(\bar{f}_{cest})$		$s(s(f_{cest}))$	
	Reg. app.	Bi-obj. app.	Reg. app.	Bi-obj. app.	Reg. app.	Bi-obj. app.	Reg. app.	Bi-obj. app.
1	-	-	-	-	-	-	-	-
2	4.5	4.5	4.1	4.1	7.5	7.5	6.6	6.6
3	3.5	3.4	3.5	3.5	5.9	6.4	5.6	5.7
4	2.9	2.8	3.1	3.1	4.9	5.5	4.9	4.9
5	2.5	2.4	3.0	3.1	4.1	4.7	4.2	4.4
6	2.1	2.1	2.8	2.9	3.9	4.3	3.9	3.9
7	1.9	1.9	2.5	2.7	3.4	3.9	3.8	3.7
8	1.9	2.0	2.5	2.8	3.1	3.6	3.6	3.3
9	1.7	1.7	2.5	2.8	2.9	3.4	3.5	3.1
10	1.5	1.6	2.3	2.6	2.8	3.2	3.3	2.8
11	1.5	1.5	2.2	2.5	2.6	2.9	3.0	2.5
12	1.5	1.5	2.1	2.4	2.4	2.8	3.0	2.4
13	1.4	1.4	2.1	2.3	2.4	2.7	2.9	2.4
14	1.3	1.4	2.0	2.2	2.2	2.6	2.8	2.3
15	1.2	1.3	1.9	2.1	2.2	2.5	2.8	2.2
16	1.2	1.3	1.9	2.1	2.1	2.4	2.7	2.1
17	1.2	1.2	1.8	2.0	2.0	2.3	2.6	2.0
18	1.1	1.1	1.7	1.9	1.9	2.2	2.6	1.9
19	1.1	1.1	1.6	1.9	1.9	2.1	2.5	1.9
20	1.0	1.0	1.6	1.9	1.8	2.1	2.4	1.7

From the results presented in Figures 5.1 and 5.2 and Table 5.1 the following observations can be highlighted:

- regarding the mean strength estimation, Figure 5.1, the bi-objective approach appears to be as efficient as other approaches for capturing the mean strength, with curves that are very close from those with the regression approach. Looking at Table 5.1, one can observe that the variability in the estimated mean strength, $s(\bar{f}_{cest})$, decreases as NC increases with close values obtained from the bi-objective and regression approaches.
- regarding the strength variability estimation, Figure 5.2, it is clear that the bi-objective approach is the only one which can efficiently capture the true reference value of concrete strength variability. The convergence of the bi-objective curve to the true reference line (red line) is from above for datasets S5, S18 and S39 while it is from bottom for the other datasets. Looking at the values of strength variability in terms of coefficient of variation, $CV(f_c)$, it is clear that these three datasets have smaller $CV(f_c)$ values as compared with the other datasets. Therefore, bi-objective approach overestimates the concrete strength variability for the less variable concrete and underestimates the concrete variability for the more variable concrete. Regarding the variability in the estimated strength variability, $s(s(f_{cest}))$ in Table 5.1, the behavior of bi-objective approach is similar to that of regression approach i.e. $s(s(f_{cest}))$ values decrease as NC increases.

- the bi-objective and regression approaches always have the same value for $NC=2$ (bold numbers in Table 5.1). This is logical due to the mathematical fact that for identifying a linear model with two unknown parameters and only two pairs of (x, f_c) , there is a unique solution whatever the approach, bi-objective or regression.
- the analysis results that based on datasets with either ultrasonic pulse velocity measurements (S3, S17, S18, and S19) or rebound hammer measurements (S5, S20, S38, and S39) confirm that the bi-objective approach can work well with these NDT techniques.

5.3.2 Case of power model

In section 5.2, the derivation of the bi-objective model parameters was extended to the case of nonlinear model (power model) so it is necessary to check the validity of the power model derived using the bi-objective approach. To this end, the in-situ dataset S1 was selected from Table 3.1. This dataset consists of 205 test locations with the test results of rebound hammer, pulse velocity and core measurements at each test location. Since bi-objective model has been developed for a case of single NDT technique, in this subsection only the pulse velocity and core measurements of dataset S1 were considered i.e. 205 pairs of (V, f_c) .

The analysis procedure follows the algorithm illustrated in Figure 3.3. From $NT=205$ test locations, the NC locations were selected randomly. The NC pairs of (V, f_c) were used to identify four models corresponding to four approaches (calibration with Δ -method, calibration with k-method, regression, and bi-objective). For the calibration methods the prior uncalibrated model that had been selected from the literature was the model $(f_{cest} = 2.0572 V^{1.7447})$, with V in km/s and f_{cest} in MPa) provided by (Kheder, 1999). For comparison purposes, two linear models corresponding to regression and bi-objective approaches were also derived. Consequently, six models were identified corresponding to six cases: Δ -method, k-method, regression (linear), regression (power), bi-objective (linear), and bi-objective (power). Each derived model was used to calculate the local strength values and consequently to provide the mean strength and strength standard deviation (strength variability). After NI repetitions of this process, for each case, we have NI values of \bar{f}_{cest} and NI values of $s(f_{cest})$. These results were post-processed to provide the $\overline{\bar{f}_{cest}}$, $s(\bar{f}_{cest})$ for the mean strength estimation and $\overline{s(f_{cest})}$, $s(s(f_{cest}))$ for strength variability estimation. These four quantities were calculated for each case for NC varies from 1 to 20.

Figure 5.3a shows \bar{f}_{cest} curves for the six cases as a function of NC while $\overline{s(f_{cest})}$ curves are shown in Figure 5.3b. Table 5.2 provides the values of $s(\bar{f}_{cest})$ and $s(s(f_{cest}))$ that correspond to regression and bi-objective approaches for both linear and power models.

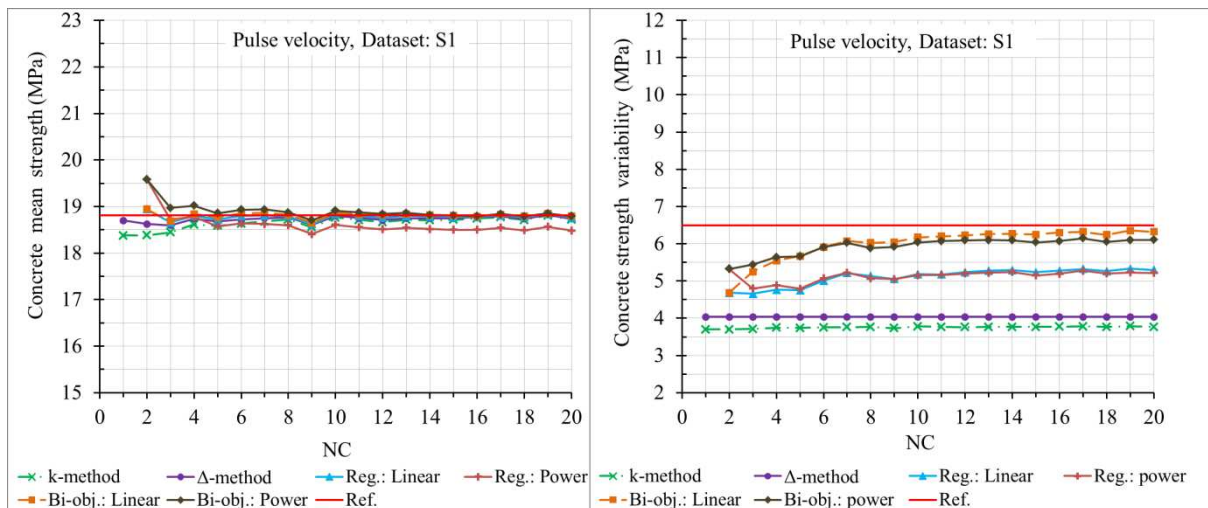


Figure 5.3 Predicting mean strength and strength standard deviation (strength variability) using the regression, calibration, and bi-objective approaches, linear and power models

Table 5.2 Standard deviation values from NI values of mean strength and concrete variability estimated by regression and bi-objective approaches (linear and power models) for the datasets S1 – all results are in MPa

NC	$s(\bar{f}_{cest})$				$s(s(f_{cest}))$			
	Linear		Power		Linear		Power	
	Reg. app.	Bi-obj. app.	Reg. app.	Bi-obj. app.	Reg. app.	Bi-obj. app.	Reg. app.	Bi-obj. app.
1	-	-	-	-	-	-	-	-
2	3.3	3.3	5.9	5.9	2.1	2.1	5.0	5.0
3	2.4	2.3	2.8	3.7	1.8	1.7	2.4	3.0
4	2.1	2.0	2.2	2.1	1.6	1.4	1.9	1.7
5	1.8	1.8	1.8	1.7	1.5	1.2	1.6	1.3
6	1.6	1.6	1.6	1.6	1.4	1.2	1.5	1.3
7	1.4	1.4	1.4	1.5	1.2	1.1	1.3	1.3
8	1.3	1.3	1.3	1.3	1.0	1.0	1.1	1.0
9	1.2	1.2	1.2	1.2	1.0	1.0	1.1	1.0
10	1.2	1.2	1.1	1.2	0.9	0.9	1.0	0.9
11	1.1	1.1	1.1	1.1	0.9	0.9	1.0	0.9
12	1.1	1.1	1.1	1.1	0.9	0.9	1.0	0.9
13	1.0	1.0	1.0	1.0	0.8	0.8	0.9	0.8
14	1.0	1.0	0.9	1.0	0.8	0.8	0.9	0.8
15	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.8
16	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.7
17	0.8	0.9	0.8	0.8	0.7	0.7	0.7	0.7
18	0.8	0.8	0.8	0.8	0.7	0.7	0.7	0.6
19	0.8	0.8	0.8	0.8	0.7	0.7	0.7	0.7
20	0.8	0.8	0.7	0.8	0.7	0.7	0.7	0.7

Regarding the assessment of mean strength, the observations from Figure 5.3a and Table 5.2 can be outlined as follows:

- bi-objective and the existing approaches appear to be able to predict the mean strength with acceptable accuracy as soon as $NC \geq 3$. However this decision about the suitable value of NC to predict the mean strength may be changed if we consider the variability in estimated mean strength, $s(\bar{f}_{cest})$, given in Table 5.2. The $s(\bar{f}_{cest})$ values decrease as NC increases with more or less identical decreasing rate for all approaches and for both types of models (linear and power shapes). As an example, $s(\bar{f}_{cest})$ is lower than 10% of the mean strength as soon as $NC \geq 5$,
- like the above observation in Subsection 5.3.1, the curves obtained for bi-objective and regression approaches are very close to each other for both cases: linear and power models,
- for $NC=2$, the bi-objective and regression approaches show estimation uncertainties (especially for power model) higher than that produced by calibration approaches. This indicates that the use of regression or bi-objective with NC equal to the number of parameters in the model will lead to an unreliable estimation (this issue was earlier discussed in Chapter 4 and will be developed in Chapter 6). It is necessary to emphasize here that using ($NC=2$) with the calibration approaches is also risky since the behavior of calibration approaches depends essentially on the prior model. However, increasing NC has a significant effect on the prediction capability of these approaches.

Regarding the estimation of concrete strength variability, from Figure 5.3b and Table 5.2 the following observations can be noted:

- the regression approach underestimates the concrete variability by about 20%, and increasing NC has no significant effect on the prediction capability. Regarding the calibration approaches (k-method and Δ -method), the behavior is similar, with nearly horizontal lines but with different values. Since this behavior mainly depends on the prior model and its ability to represent the concrete under consideration, calibration approaches are inadequate for capturing the concrete variability. It is only due to the effect of chance that such approaches could (sometimes) lead to the good estimation of concrete variability,
- it is thus clear that the bi-objective approach is the only one which can capture (with linear model is a little bit more efficient than power model in this case study) the concrete strength variability (the true reference value) even with NC lower than the minimum number required by the standard (18 for regression approach and 9 for calibration approach according to (EN 13791, 2007)). However, the selection of adequate NC value needs to have a look on the $s(s(f_{cest}))$ values accompanying this approach, see Table 5.2, which decrease with increasing NC . Moreover, the $s(s(f_{cest}))$ values are larger for the power model when NC is small and become nearly similar for linear and power models when $NC \geq 8$.

In summary, from the analyses that have been carried out in Subsections 5.3.1 and 5.3.2, the validity of the bi-objective approach has been approved for estimating the mean strength and concrete strength variability using linear or power models. Furthermore, one must add that this innovation is costless, since it only concerns the post-processing of test results, and has no impact on the investigation program.

5.3.3 Failure of the model identified by each approach

A last criterion that deserves to be analyzed when comparing the relative merits of the approaches is related to the fact that, in some situations, the NC pairs of (x, f_c) which are used for identifying a model may produce a very bad or even physically impossible conversion model (this issue has been already discussed in Section 3.4). It is the case for instance in the following situations:

- (a) when the estimated concrete strength at a test location obtained from applying the conversion model has a zero or negative value, see Figure 3.4a,
- (b) when the model has a negative slope for the case of linear model (or negative exponent for power model), see Figure 3.4b, which would mean that the strength decreases while the NDT value increases, which generally has no physical meaning with rebound hammer and ultrasonic velocity measurements. This situation can be encountered only with regression approach because the other approaches always guarantee a positive slope (or exponents).

The risk of having the above situations (i.e. identifying a “wrong model”) is larger if the NC is small, and if the measurement uncertainties are relatively large when compared to the range of variation of concrete strength in the domain of investigation.

As stated in Section 3.4, the simulator has included two conditions (if $f_{cest} \leq 0$ or if slope < 0) for linear model and (if $f_{cest} \leq 0$ or if exponent < 0) for power model. For each model type, when one or both of their conditions is satisfied, the simulator will reject this model and return to select other NC pairs of (x, f_c) in order to identify another model. Counters are inserted in the simulator to count the number of failed repetitions (repetitions identify wrong models) for each approach. The percentages of failed repetitions with respect to the total number of repetitions are given in Table 5.3 for the analyses of datasets S1, S18, and S39 presented in Section 5.3. For each dataset, this table provides three columns for the regression approach where the first and second columns represent the contribution of each one of the two conditions to the final failure percentages (i.e. the third column “sum”). However, for bi-objective approach the failure percentages that correspond to the condition (if $f_{cest} \leq 0$) are only provided since bi-objective approach identifies always model with positive slope.

This table shows large failure percentages when $NC=2$ with larger values for the regression approach. This is due to the fact that with $NC=2$ the probability of identifying a model with negative slope or obtaining an estimated strength with negative (or zero) value is high. However, the failure percentages reduce with increasing NC . From the comparison of the failure percentages for regression approach (i.e. the column “sum”) with those for bi-objective approach, it is clear that the decreasing rate is higher for the regression approach for dataset S1 and dataset S18. On contrary, the bi-objective approach has the higher decreasing rate for dataset S39.

Regarding the regression approach, for dataset S39, the contribution of the condition (slope < 0) to the final failure percentage (i.e. “sum”) is more dominating as compared with the condition ($f_{cest} \leq 0$) and the vice versa for the datasets S1 and S18. This indicates that, for regression approach, there is no general statement about the domination of any of the two conditions.

Regarding the bi-objective approach, for datasets S18 ($CV(f_c)=15\%$) and S39 ($CV(f_c)=16\%$) the failure percentages are relatively small (as compared with dataset S1) and they decrease rapidly as NC increases. While for the more variable concrete (dataset S1 which has $CV(f_c)=35\%$) the decreasing rate becomes slower and the failure percentages remain significant even with $NC=20$. Therefore, it appears that, for bi-objective approach, the

decreasing rate of the model failure percentages with increasing NC depends mainly on the strength variability of the concrete under consideration.

Table 5.3 Model failure percentages when identifying linear models using regression and bi-objective approaches for datasets S1, S18, and S39.

NC	Dataset S1				Dataset S18				Dataset S39			
	Reg. app.			Bi-obj. App.	Reg. app.			Bi-obj. App.	Reg. app.			Bi-obj. App.
	$Slope \leq 0$	$f_{cest} \leq 0$	sum	$f_{cest} \leq 0$	$Slope \leq 0$	$f_{cest} \leq 0$	sum	$f_{cest} \leq 0$	$Slope \leq 0$	$f_{cest} \leq 0$	sum	$f_{cest} \leq 0$
1	-	-	-	-	-	-	-	-	-	-	-	-
2	16.3	29.9	46.2	37.2	13.0	24.6	37.6	31.9	27.5	12.2	39.7	20.2
3	6.3	26.5	32.8	38.4	4.6	13.8	18.4	21.8	20.5	4.5	25.0	10.2
4	2.5	16.6	19.1	31.1	2.0	12.0	14.0	20.5	15.3	0.6	15.9	4.8
5	2.1	14.4	16.5	29.8	0.7	7.3	7.9	14.2	11.9	0.3	12.2	1.6
6	0.3	11.1	11.4	22.4	0.3	4.4	4.7	9.7	8.6	0.3	8.9	1.2
7	0.1	10.4	10.5	25.4	0.0	2.4	2.4	8.7	8.9	0.0	8.9	0.4
8	0.1	8.5	8.6	24.8	0.0	2.2	2.2	6.5	7.5	0.0	7.5	0.2
9	0.0	6.0	6.0	21.9	0.0	1.0	1.0	4.8	6.2	0.0	6.2	0.1
10	0.0	7.1	7.1	21.2	0.0	1.1	1.1	3.9	3.8	0.0	3.8	0.2
11	0.0	4.1	4.1	21.7	0.0	0.5	0.5	1.8	4.0	0.0	4.0	0.0
12	0.0	3.6	3.6	15.7	0.0	0.2	0.2	1.9	3.1	0.0	3.1	0.0
13	0.0	2.8	2.8	18.6	0.0	0.1	0.1	2.0	2.2	0.0	2.2	0.0
14	0.0	3.1	3.1	14.8	0.0	0.2	0.2	1.0	1.4	0.0	1.4	0.0
15	0.0	1.9	1.9	11.9	0.0	0.1	0.1	1.1	1.5	0.0	1.5	0.0
16	0.0	1.4	1.4	11.6	0.0	0.1	0.1	0.4	0.6	0.0	0.6	0.0
17	0.0	1.0	1.0	14.1	0.0	0.0	0.0	0.6	0.5	0.0	0.5	0.0
18	0.0	1.7	1.7	12.6	0.0	0.0	0.0	0.1	0.7	0.0	0.7	0.0
19	0.0	0.5	0.5	10.8	0.0	0.0	0.0	0.2	0.2	0.0	0.2	0.0
20	0.0	1.0	1.0	12.7	0.0	0.0	0.0	0.0	0.4	0.0	0.4	0.0

5.4 Conclusions

In this chapter, a new model identification approach, called “bi-objective”, has been developed in order to capture concrete strength variability in addition to mean strength. This innovation works at the test results post-processing stage, and has no effect on the cost of the investigation.

The validity of this approach was studied by applying it on several datasets from in-situ and laboratory studies in order to assess the mean strength and concrete strength variability. Moreover, the prediction ability of the bi-objective approach was compared with that of the existing approaches (calibration and regression). From the obtained results, the following conclusions can be drawn:

- the bi-objective approach and the existing approaches (regression and calibration) can efficiently estimate the mean concrete strength. Furthermore, increasing NC can significantly improve the estimation efficiency,
- when trying to quantify concrete strength variability, the bi-objective approach is the only one that can capture the true concrete variability. On the contrary, the regression approach remains limited in capturing the concrete variability: it always underestimates the concrete variability and increasing NC has no significant effect on the estimation efficiency. Regarding the calibration approaches, they produce unstable results (estimation efficiency varies as the prior uncalibrated model is changed) so their efficiency depends only on chance,
- too low number of test locations for cores is a source of problems. We had seen in previous chapter that it leads to large prediction error. This has been confirmed in this chapter when looking at the variability of estimates (i.e. $s(\bar{f}_{cest})$ and $s(s(f_{cest}))$), which may be very large when NC is only 2. In addition, the risk of getting an unphysical model is also high for low NC values. The suitable NC depends on the estimated concrete property (i.e. mean strength or concrete strength variability) and on the accepted level of uncertainty in the estimated value. This issue is complex and will be addressed in a more systematic way in the next Chapter,
- the bi-objective approach can be used (from a mathematical viewpoint) with any NDT technique (rebound hammer, ultrasonic pulse velocity, etc.) combined with DT technique (cores) in order to identify a conversion model. Furthermore, the results of analyzing datasets with ultrasonic pulse velocity measurements and that with rebound hammer measurements confirm the efficiency of the bi-objective approach with these two NDT techniques,
- the bi-objective approach can work with a power model as well as with the linear model but the uncertainty arising from using power model is larger than that of linear model when NC is small and we did not identify any convincing argument in favor of nonlinear models.

CHAPTER SIX

QUALITY OF ASSESSMENT AND RECOMMENDATIONS FOR BETTER PRACTICE

6.1 Introduction

Reliability of the estimated value is a real challenge for engineers when assessing the in-situ strength of concrete structures. From the analysis of inherent characteristics of the assessment methodology carried out in the previous chapters, there is evidence that the changes in the values of these factors have significant effect on the quality of assessment. Therefore, to have more detailed and accurate evaluation of the quality of assessment, it is necessary to analyze its correlation with its affecting factors: number of test location for cores (NC) used to identify the model, quality of measurements (within-test variability), the way of selection the core locations, true in-situ strength variability, the type of NDT technique, using single or combination of NDT techniques, and the model identification approach. This chapter is devoted to achieve this goal.

A large campaign of synthetic datasets (36000- (V, R, f_c) test results) having different characteristics were generated for the analysis in this chapter. The synthetic data have been selected due to their flexibility to provide large datasets with different characteristics that can be used to analyze the above listed factors simultaneously whereas this is difficult and expensive with the real datasets. However, in order to confirm the robustness of the findings with the synthetic datasets, the analysis of the factors listed above has also been done using real datasets but with rebound hammer only (2500- (R, f_c) test results selected from Table 3.1). We have published the analysis using real datasets in “Construction and Building Materials” Journal. The paper is available in Appendix D, however the final conclusions of this paper are given herein.

At the end of the present chapter, several recommendations for better practice will be provided for engineers on site. These recommendations have been derived from all the analyses that had been presented in this document.

6.2 Analysis of quality of assessment using synthetic datasets

6.2.1 Datasets

In order to derive more general conclusions about the reliability of assessment 36 synthetic datasets that cover a wide range of concrete mean strengths and concrete strength variabilities and three measurements quality levels were considered. Through this chapter, each synthetic dataset will be represented by the letter “D” followed by the true in-situ mean strength, \bar{f}_{ist} , then the true in-situ concrete strength variability (in terms of coefficient of variation, $CV(f_{ist})$) and the abbreviation of the measurement quality levels provided in Subsection 3.3.2 (i.e. HQ or AQ or LQ). The characteristics of these datasets are given in Table 6.1. These

datasets were generated using the synthetic simulation procedure detailed in Chapter 3 and illustrated graphically in Figure 3.2. Using the dataset characteristics, Table 6.1, the simulator generated 1000- (V, R, f_c) test results for each dataset (i.e. each dataset includes the test results for 1000 test locations).

Figure 6.1 shows some clouds of generated test results (each cloud includes 1000- (V, f_c) pairs) that correspond to several datasets selected from Table 6.1. These datasets show the changes in the dispersion of the points within each cloud due to the variety in the datasets characteristics. Figures 6.1a, 6.1b and 6.1c compare the dispersions within datasets having two different values of \bar{f}_{ist} (20 MPa and 50 MPa) and three different values of $CV(f_{ist})$ (10%, 20%, and 30%) while these datasets have the same quality of measurements (AQ). However, the comparison between the Figures 6.1b and 6.1d illustrates the change in the dispersion as the quality of measurements changes from AQ (Figure 6.1b) to LQ (Figure 6.1d) while the other characteristics of datasets remain unchanged. From Figure 6.1, it is clear that for a more variable concrete the dispersion increases in the direction of the longitudinal axis of the cloud while using bad quality measurements increases the dispersion in the direction of the transverse axis of the cloud.

Table 6.1 Characteristics of 36 synthetic datasets considered for analyzing the quality of assessment

Datasets symbols			Datasets characteristics			
			\bar{f}_{ist} MPa	$s(f_{ist})$ MPa	$CV(f_{ist})$ %	Measurements quality level
D20-10-HQ	D20-10-AQ	D20-10-LQ	20	2	10	high quality (HQ), or average quality (AQ), or low quality (LQ), according to the classification provided in Subsection 3.3.2 depending on the within-test variability of the measurements (sd_V, sd_R, sd_f)
D20-20-HQ	D20-20-AQ	D20-20-LQ	20	4	20	
D20-30-HQ	D20-30-AQ	D20-30-LQ	20	6	30	
D30-10-HQ	D30-10-AQ	D30-10-LQ	30	3	10	
D30-20-HQ	D30-20-AQ	D30-20-LQ	30	6	20	
D30-30-HQ	D30-30-AQ	D30-30-LQ	30	9	30	
D40-10-HQ	D40-10-AQ	D40-10-LQ	40	4	10	
D40-20-HQ	D40-20-AQ	D40-20-LQ	40	8	20	
D40-30-HQ	D40-30-AQ	D40-30-LQ	40	12	30	
D50-10-HQ	D50-10-AQ	D50-10-LQ	50	5	10	
D50-20-HQ	D50-20-AQ	D50-20-LQ	50	10	20	
D50-30-HQ	D50-30-AQ	D50-30-LQ	50	15	30	

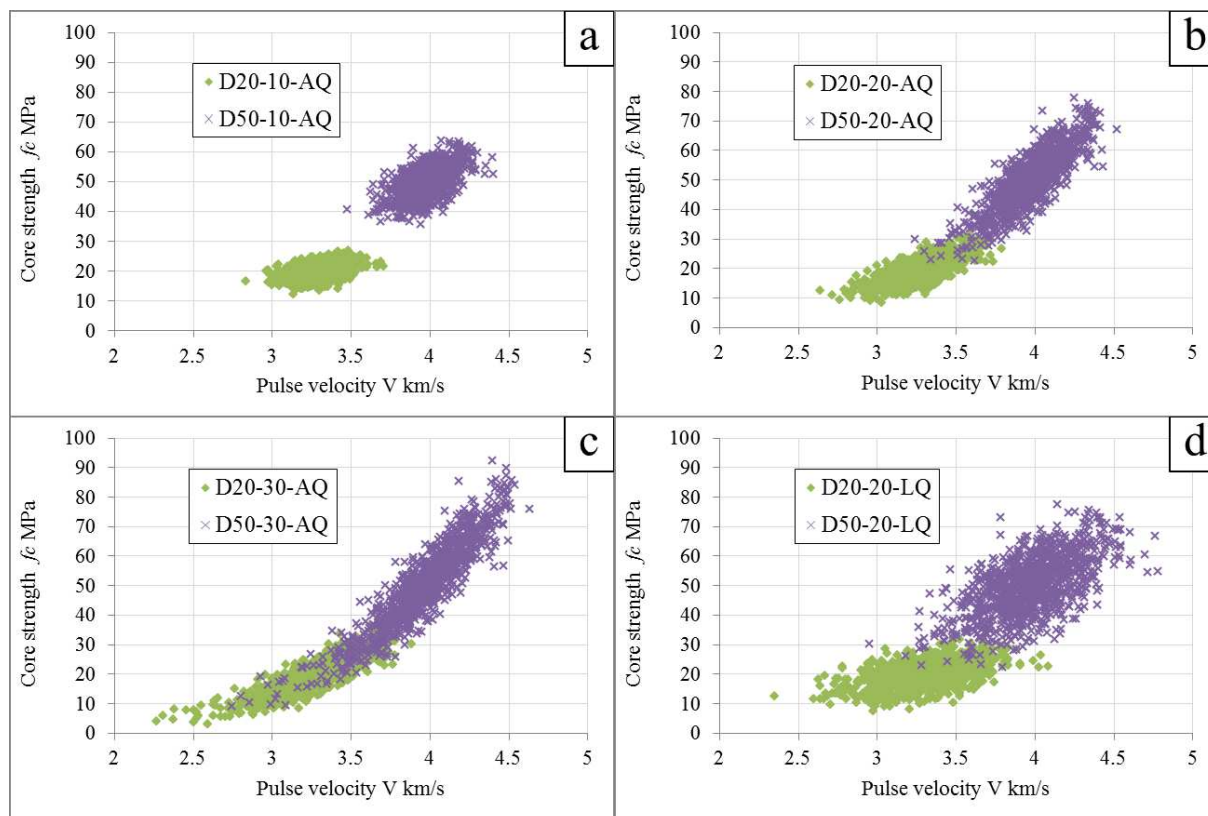


Figure 6.1 Clouds of 1000- (V, f_c) for several generated datasets

6.2.2 Assessing mean strength and strength standard deviation (concrete variability) and developing the cumulative distribution functions (CDF)

As shown in Chapters 4 and 5, the estimated mean strength, \bar{f}_{cest} , and estimated concrete strength variability, $s(f_{cest})$, scatter with standard deviations $s(\bar{f}_{cest})$ and $s(s(f_{cest}))$ respectively. These scatters are due to the sources of uncertainty discussed in Subsection 2.4.5 (in-situ strength variability, sampling uncertainty, measurements uncertainty, and model uncertainty). Consequently, to study the quality of assessment, it is necessary to analyze the scatter in the estimated quantity.

To this end, each of the datasets given in Table 6.1 was analyzed using the procedure illustrated in Figure 3.3. For the analysis of a dataset, the dataset size corresponds to 1000 test locations (the population size). From these test locations, NT test locations were selected randomly as a sample size ($NT=100$). Then, the NC test locations for cores were selected randomly from the NT test locations. The test results (V, R, f_c) of NC test locations were used to identify three linear models (for single V , single R , and combination of $V+R$) using the regression approach and two linear models (for single V and single R) using the bi-objective approach. Thus, there were five cases (models). For each case, the model was used to assess the local strengths at the NT test locations and consequently to calculate the mean strength, \bar{f}_{cest} , and concrete strength variability, $s(f_{cest})$, corresponding to this model. This process was repeated NI times ($NI=1000$) so the final outputs were, for a specific NC value and each case, a series of 1000 values of \bar{f}_{cest} and a series of 1000 values of $s(f_{cest})$. The whole process was repeated for values of NC varying from 2 to 20 in order to have the series of results corresponding to different NC values.

The adopted way to show these outputs is to draw each series in terms of cumulative distribution function CDF. Therefore, for each dataset, each NC value, each approach and

each NDT technique (single V or single R or combined $V+R$) a series of 1000 values of \bar{f}_{cest} and a series of 1000 values of $s(f_{cest})$ were plotted in terms of cumulative distribution function CDF. Since it is impractical to display herein the plots for all outputs, only four figures that correspond to dataset D20-20-AQ and for the case of using pulse velocity as NDT measurements are shown in Figure 6.2. Figures 6.2a and 6.2b refer respectively to the mean strength estimation, \bar{f}_{cest} , by regression and bi-objective approaches while Figures 6.2c and 6.2d refer to strength variability estimation, $s(f_{cest})$, with the same two approaches.

Figure 6.2 displays how the CDF curves vary with NC . In fact the analysis was performed for NC values (2-20), however to avoid the curves congestion, the figures show only the curves that correspond to some NC values (2, 6, 10, 14, and 20). On the four figures the red vertical line represents the true in-situ value for the dataset which is the target of assessment, given in Table 6.1, respectively 20 MPa for \bar{f}_{ist} and 4 MPa for $s(f_{ist})$. Therefore the quality of assessment is better as the CDF curve is closer to the red line (i.e. less scatter). In an ideal situation (no scatter), the CDF curve would coincide with the red line, but the sources of uncertainty always exist and as a result some scatter is unavoidable. However, as shown in Figures 6.2a and 6.2b, the scatter decreases as NC increases. Moreover, the reduction in the scatter (i.e. improvement in quality of assessment) appears remarkable for low values of NC while it becomes almost negligible (i.e. expensive without being effective) for higher NC values. The comparison of CDF curves of regression approach with those of bi-objective approach indicates that these two approaches are close in their reliability for assessing the mean strength.

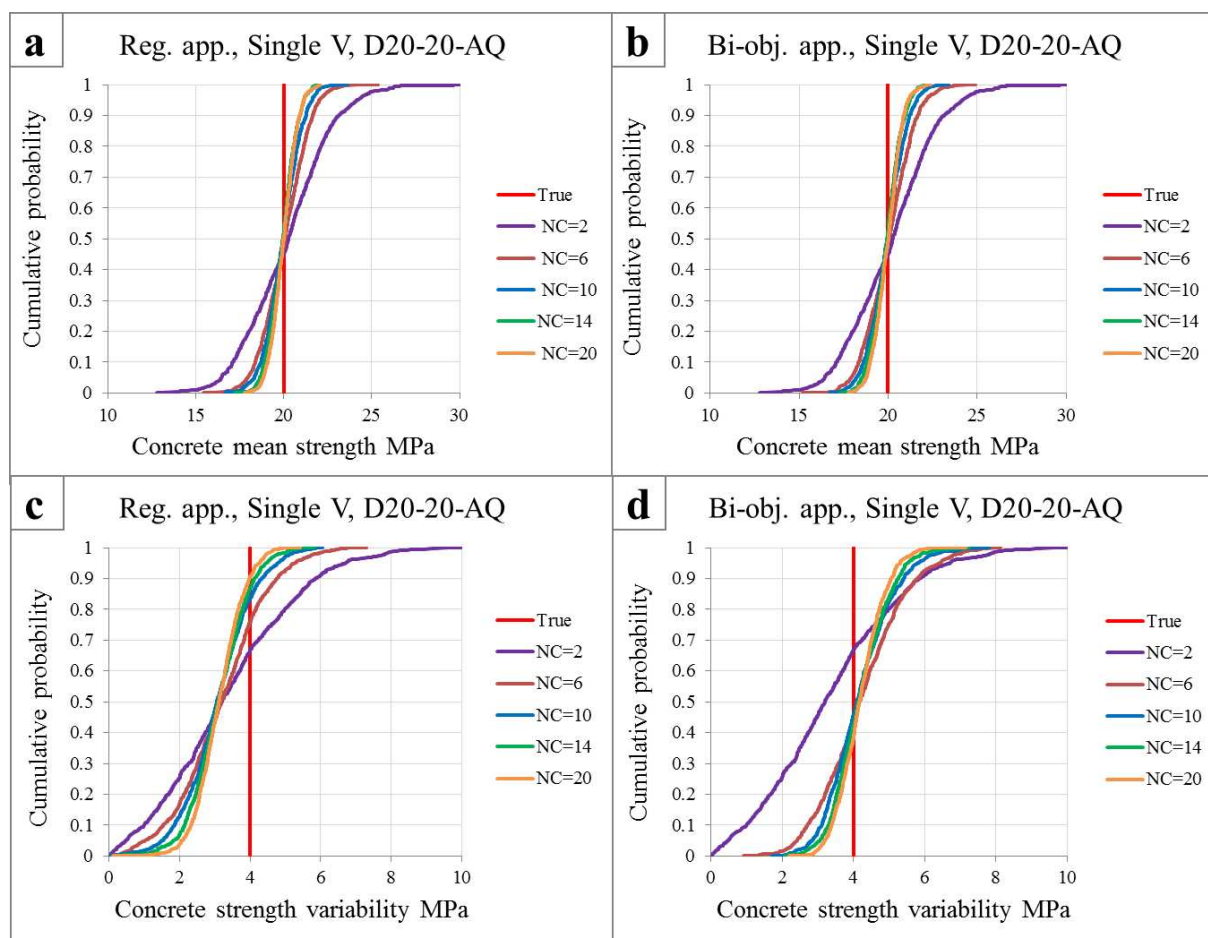


Figure 6.2 CDF curves of mean strength and concrete strength variability estimations for dataset D20-20-AQ and when the NDT technique is the pulse velocity

However, regarding strength variability assessment presented in Figures 6.2c and 6.2d, the result is different. For the regression approach, Figure 6.2c, increasing NC reduces the scatter but does not guarantee the convergence towards the true in-situ value. The CDF curves show, even with large NC , a left bias (meaning general underestimation of strength variability). In contrast, the CDF curves corresponding to bi-objective approach, Figure 6.2d, show a significant reduction in the scatter on both sides and a tendency to converge, however slower than for mean strength, towards the true in-situ value as NC increases. This finding confirms that the common regression approach cannot be used to estimate efficiently the strength variability and that the bi-objective approach answers this need as previously demonstrated in Chapter 5.

6.2.3 Assessing the quality of estimation by developing “Risk Curves”

As shown in the above subsection, due to many sources of uncertainties, the estimated values are scattered and their CDF can be plotted. The quality of estimation is improved as the scatter reduces. As shown in Subsection 2.4.6, a common indication about the quality of assessment is providing the estimated value with an error (uncertainty level). In other wording, providing the interval (estimated value \pm uncertainty level) with a specific confidence level that the true value will be assured to lie within lowest and highest bound of this interval. The uncertainty level is usually a relative value $\pm U\%$ as shown in Subsection 2.4.6, however an absolute value can also be used. To derive this information from the CDF curves, it is preferable to proceed in two steps:

- using the CDF curve, fix an interval $\pm(U \times T)$ around the true target value, T ,
- quantify the probability (or risk) that the estimated value lies outside the fixed interval.

Therefore, the idea is for a specific uncertainty level (or error) around the true value, what is the risk (probability) to have an estimated value which lies outside this target interval. The concept to derive the risk values (or the confidence level, confidence level = 1 - risk) from CDF curve is illustrated in Figure 6.3 (Figure 6.3a illustrates the proposed concept while Figure 6.3b demonstrates how the change in the scatter of CDF curve can change the risk value). Risk values can be finally plotted as a function of NC to establish new curves that we are named “Risk curves”.

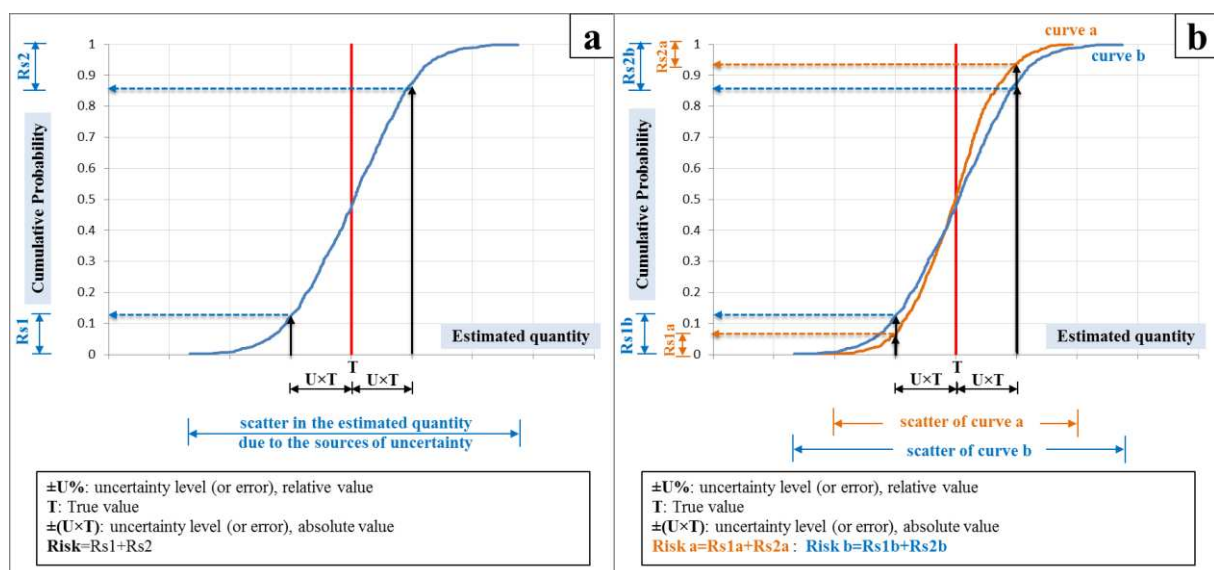


Figure 6.3 Proposed concept to derive risk value corresponding to a specific uncertainty level using CDF curve: graph (a) illustrates the proposed concept, while graph (b) displays how the change in the scatter of CDF curve changes the risk value

6.2.3.1 Analysis of risk curves corresponding to mean strength estimation

To construct the risk curves of mean strength estimation from the CDF curves, it is necessary to specify an uncertainty level, $U\%$. There is no unique value for U and the selection of U depends on how much uncertainty (or error) we accept in the final estimated value. Looking at Figure 6.3 it is obvious that selecting a large value for U , for the same value of T , reduces the segments of the CDF curve that are outside the interval $\pm(U \times T)$ (i.e. reduces the risk values) and vice versa. In the present study, ($U = \pm 10\%$) is adopted for all datasets used for the assessment of mean strength.

Figure 6.4 shows the risk curves that obtained from analyzing the 12 datasets with average quality of measurements (AQ), see Table 6.1. For clarity, these datasets were arranged in three groups according to the $CV(f_{ist})$ values: 10% (curves in green), 20% (curves in orange) and 30% (curves in blue). Furthermore, within each group the line style was varied in order to distinguish between curves having different values of \bar{f}_{ist} . Figures 6.4a, 6.4b, and 6.4c display the risk curves corresponding to the use of the regression approach to identify the models for three cases of NDT measurements (single V , single R , and combined $V+R$) combined with cores. Figures 6.4d, 6.4e show the curves corresponding to the use of bi-objective approach for two cases of NDT measurements (single V and single R) combined with cores. From these figures the following observations can be noted:

- for all datasets, risk values decrease as NC increases. This behavior can be explained using Figure 6.2a in which, for a specific dataset, increasing NC makes the CDF curve less scattering and consequently leads to lower risk values for a specific uncertainty level ($U \times T$), as shown in Figure 6.3b. Moreover, the reduction in risk values due to the increase in NC is significant for low values of NC and becomes negligible for higher NC ,
- in each figure, each curve in blue is above the corresponding curves (i.e. curves having the same \bar{f}_{ist} value) in orange and green, moreover each curve in orange is above the corresponding curve in green (see as an example the curves: D20-30-AQ, D20-20-AQ, and D20-10-AQ). This indicates that, for the datasets having the same \bar{f}_{ist} but with different $CV(f_{ist})$ values, the risk values increase with increasing the true in-situ concrete variability $s(f_{ist})$, $s(f_{ist}) = CV(f_{ist}) \times \bar{f}_{ist}$. This behavior can be explained with the aid of Figure 6.3. As discussed in Chapter 2, the true in-situ concrete variability $s(f_{ist})$ is a source of uncertainty and consequently its increase leads to more scattered CDF curve, see Figure 6.3b. Since the interval $\pm(U \times T)$, herein $T = \bar{f}_{ist}$, will be the same for datasets having the same \bar{f}_{ist} , more scatter in CDF curve gives higher risk values,
- in each figure, within each group of curves having the same $CV(f_{ist})$ but different values of \bar{f}_{ist} (for example the group: D20-30-AQ, D30-30-AQ, D40-30-AQ, and D50-30-AQ), the risk values decrease as \bar{f}_{ist} increases. To understand this behavior it is necessary to have in mind two issues: on one hand, although $CV(f_{ist})$ is the same for these datasets, $s(f_{ist})$ increases as \bar{f}_{ist} increases and consequently the CDF curves are more scattered. On the other hand, for these datasets the interval $\pm(U \times T)$ is not the same but it increases as \bar{f}_{ist} increases. Therefore, increasing \bar{f}_{ist} leads to a higher scatter of the CDF curve (i.e. leads to increase the risk if the interval of $\pm(U \times T)$ is fixed) and to a wider interval of $\pm(U \times T)$ (i.e. leads to decrease the risk if the scatter is fixed), however it appears that the later effect is dominating that leads to decrease the final risk values,
- the risk curves resulted from using the bi-objective approach and the corresponding curves obtained from the regression approach show the ability of the two approaches

to provide a reliable assessment of mean strength. This observation confirms what we have already concluded in Chapter 5,

- for both approaches (regression and bi-objective) the type of NDT technique (pulse velocity or rebound hammer) that is combined with DT (cores) to identify the model has no significant effect on the risk values,

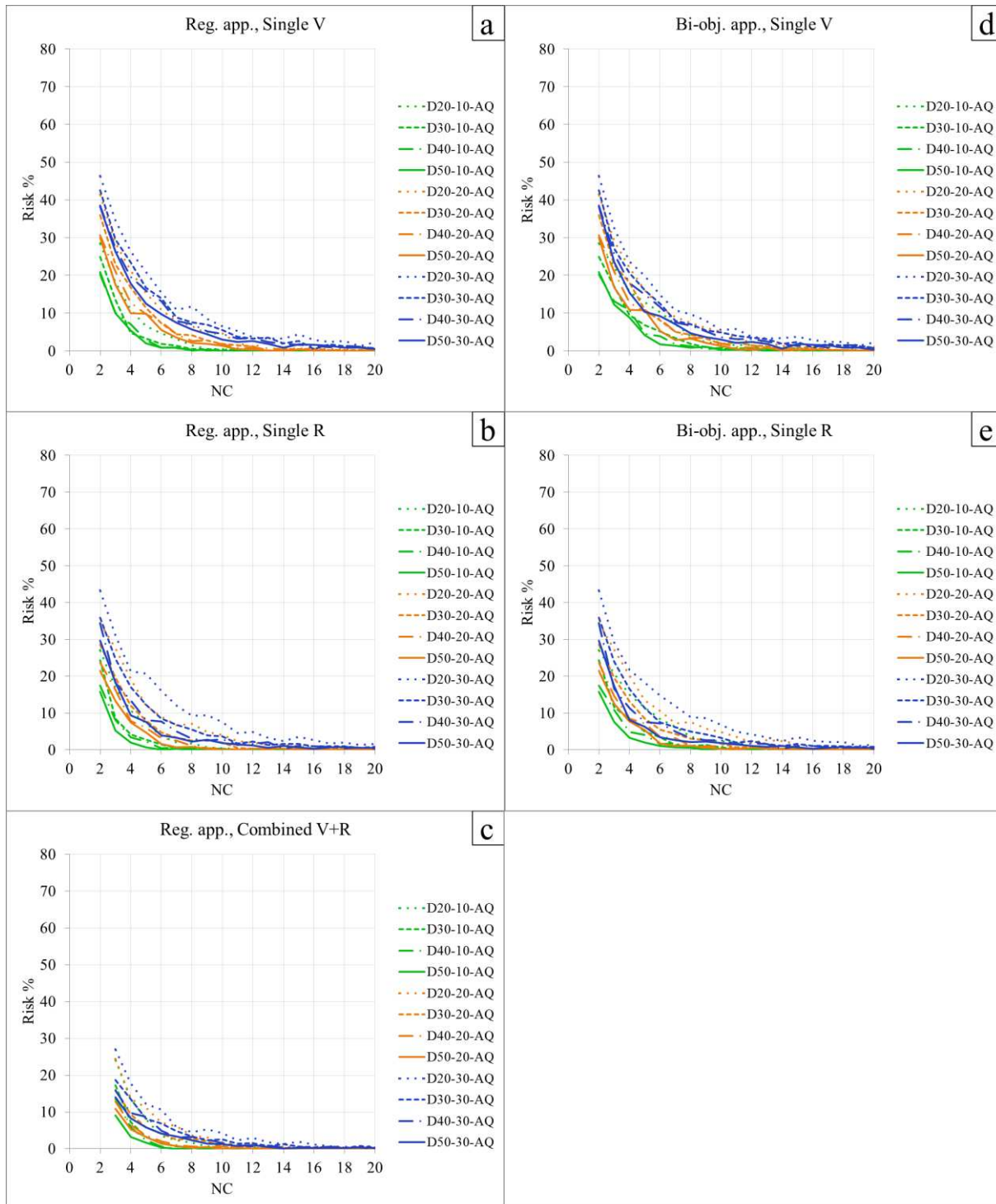


Figure 6.4 Risk curves of all datasets with (AQ) for mean strength estimation with: $U=\pm 10\%$, $NT=100$, and random core selection

- for all figures with single technique, it is clear that for risk=10% (i.e. confidence level 90%) the minimum NC required to assure the assessment of mean strength with a maximum relative error of 10% ($U=10\%$) is 8 cores for concrete with $CV(f_{ist}) = 30\%$. This number can be reduced to 6 cores for concrete with $CV(f_{ist}) = 20\%$. Furthermore, for the case of combination of NDT techniques ($V+R$), the minimum NC reduces to 6 cores for concrete with $CV(f_{ist}) = 30\%$ which indicates the efficiency of combination. However, to be fair in the evaluation of this efficiency, it is necessary to balance this reduction in NC on one hand and the increase in the cost of combination as compared with that of single technique on the other hand.

6.2.3.2 Analysis of risk curves corresponding to strength variability estimation

In order to study the quality of assessment of concrete strength variability, 12 datasets analyzed for mean strength assessment (i.e. datasets having average quality of measurements) were considered. For each dataset, the risk curve that corresponds to uncertainty level ($U=\pm 25\%$) was built. In fact, this uncertainty level ($U=\pm 25\%$) is greater than that selected for the case of mean strength estimation ($U=\pm 10\%$) due to the difficulty to assess the strength variability with reasonable confidence level and smaller uncertainty as it will be shown hereafter. Figures 6.5a, 6.5b, and 6.5c show the risk curves obtained from using the regression approach to identify models when the NDT techniques are respectively single V , single R , and combination of $V+R$. Figures 6.5d and 6.5e display the risk curves corresponding to the bi-objective approach for two cases of NDT techniques (single V , single R). For clarity, in each figure, the same curves colors and line styles convention that adopted in Figure 6.4 has been applied herein. From the graphs of Figure 6.5, the following observations can be noted:

- Figures 6.5a and 6.5b illustrate the limited ability of regression approach (with single NDT technique) in estimating the strength variability. They also show the handicap of regression approach when the datasets have low strength variability ($CV(f_{ist})=10\%$) since increasing NC offers no guarantee regarding the reduction of risk values (the convergence is slow or even lacking when the concrete variability is small). Moreover, using combination of NDT techniques ($V+R$), Figures 6.5c, can improve the quality of assessment for datasets having ($CV(f_{ist})=20\%$ and 30%) as compared with the case of single technique. However, this improvement in the quality of assessment needs to balance the additional cost due to the second added technique,
- on contrary, the bi-objective approach, Figure 6.5d and 6.5e, shows a good ability to assess the strength variability with an increasing efficiency when NC increases. Furthermore, the quality of assessment using bi-objective approach is better than that provided by regression approach even with the case of combination of NDT techniques,
- for both approaches (regression and bi-objective) the type of NDT technique (pulse velocity or rebound hammer) that combined with DT (cores) to identify the model has no significant effect on the risk values,
- in each figure, within each group of curves having the same $CV(f_{ist})$ but different values of \bar{f}_{ist} (for example the group: D20-30-AQ, D30-30-AQ, D40-30-AQ, and D50-30-AQ), the risk values decrease as \bar{f}_{ist} increases. This behavior is similar to that observed in Figure 6.4 for the case of mean strength estimation,
- in each figure, each curve in green ($CV(f_{ist})=10\%$) is above the corresponding curves (i.e. curves having the same \bar{f}_{ist} value) in orange ($CV(f_{ist})=20\%$) and in blue ($CV(f_{ist})=30\%$). Furthermore, each curve in orange is above the corresponding curve in blue (see as an example the curves: D20-10-AQ, D20-20-AQ, and D20-30-AQ). This indicates that, for the datasets having the same \bar{f}_{ist} but with different $CV(f_{ist})$ values, the risk values decrease with increasing the true in-situ concrete variability

$s(f_{ist})$. As discussed earlier, $s(f_{ist})$ is a source of uncertainty and consequently its increase leads to a more scattered CDF curve, see Figure 6.3b. However, for these datasets the interval $\pm(U \times T)$, herein $T = s(f_{ist})$, is not the same but it also increases as $s(f_{ist})$ increases. Therefore, increasing $s(f_{ist})$ leads to more scatter of the CDF curve and to wider interval of $\pm(U \times T)$, the second consequence has more impact and leads to a final decrease in the risk values when $s(f_{ist})$ increases.

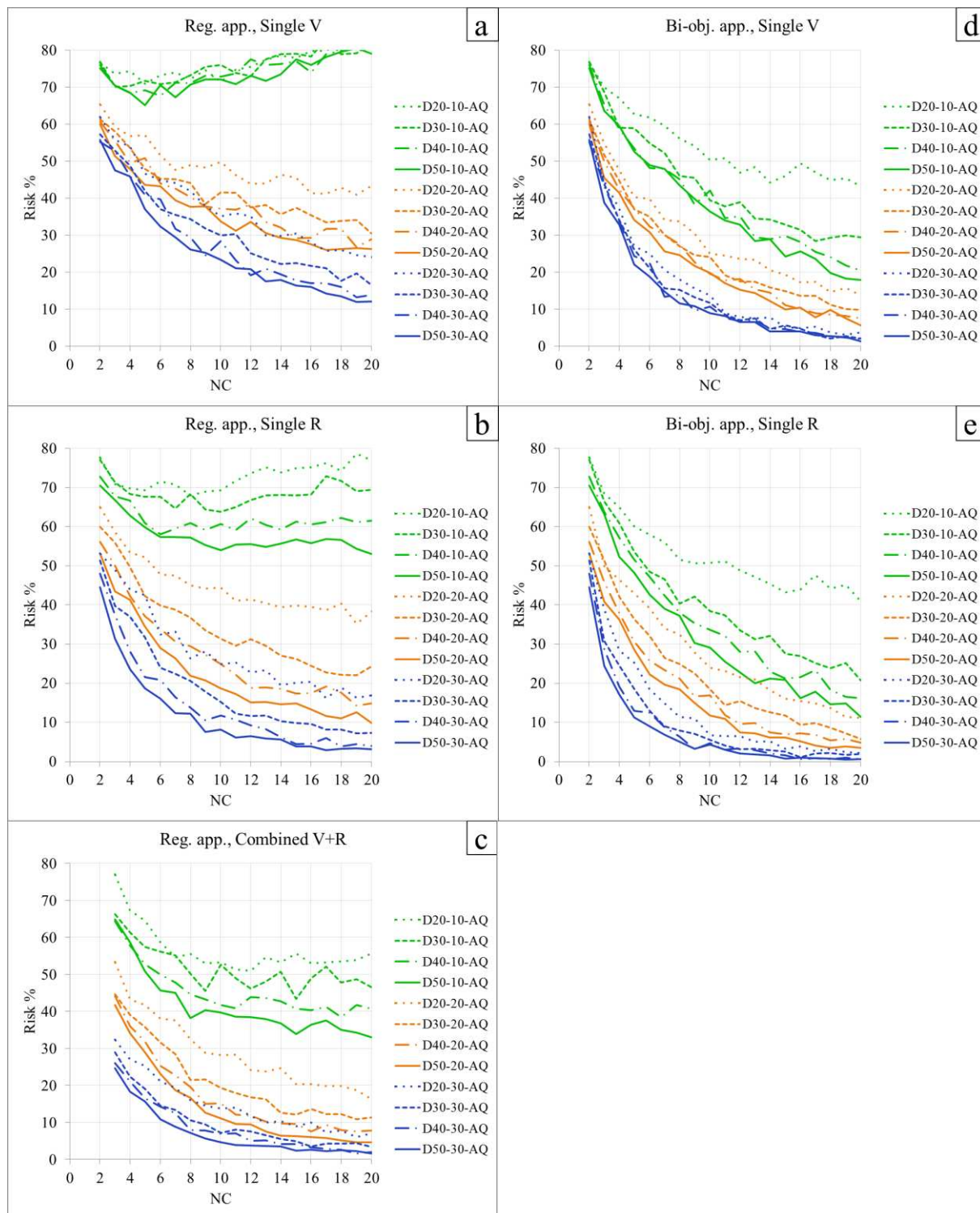


Figure 6.5 Risk curves of all datasets with (AQ) for strength variability estimation with: $U = \pm 25\%$, $NT = 100$, and random core selection

It can be noted that this behavior is adverse than that presented in Figure 6.4 for the case of mean strength estimation; i.e. decreasing the concrete variability makes the mean strength assessment more reliable, but the concrete variability assessment less reliable. For the case of mean strength estimation, the behavior is rational because capturing mean strength is easier in a less variable concrete. Whilst, for the case of strength variability estimation, the behavior (i.e. assessing the strength variability is easier in a more variable concrete) seems irrational. To understand the second behavior, one can consider the case study shown in Figure 6.6. Figure 6.6a shows three risk curves (D30-10-AQ, 30-20-AQ, and D30-30-AQ) selected from Figure 6.5d in which the results are obtained from fixing a relative uncertainty level ($U=\pm 25\%$). This means that the absolute value of uncertainty level, $\pm(U \times T)$, used in the analysis for each dataset was:

for D30-10-AQ: $T=3$ MPa (from Table 6.1), $(U \times T)=0.25 \times 3=0.75$ MPa

for D30-20-AQ: $T=6$ MPa (from Table 6.1), $(U \times T)=0.25 \times 6=1.50$ MPa

for D30-30-AQ: $T=9$ MPa (from Table 6.1), $(U \times T)=0.25 \times 9=2.25$ MPa

As shown in Figure 6.6a, the more variable dataset (i.e. D30-30-AQ) is more reliable (lower risk values), but it also corresponds to a larger absolute uncertainty level. If one chooses to fix the absolute uncertainty level for the three datasets by selecting an absolute uncertainty level (0.75 MPa for example) and to reanalyze the datasets, the resulting risk curves are different, as shown in Figure 6.6b. While in this case the curve for D30-10-AQ is the same as that on Figure 6.6a, the two other risk curves for D30-20-AQ and D30-30-AQ move upwards and change their order. According to this absolute uncertainty criterion, the more variable dataset D30-30-AQ becomes less reliable. Consequently, the statement about the effect of true in-situ concrete variability on the reliability of assessing the concrete strength variability must be done carefully, since it depends on how the uncertainty level is represented. In the present work, the relative uncertainty level is adopted because it is more common in the literature and in engineering practice.

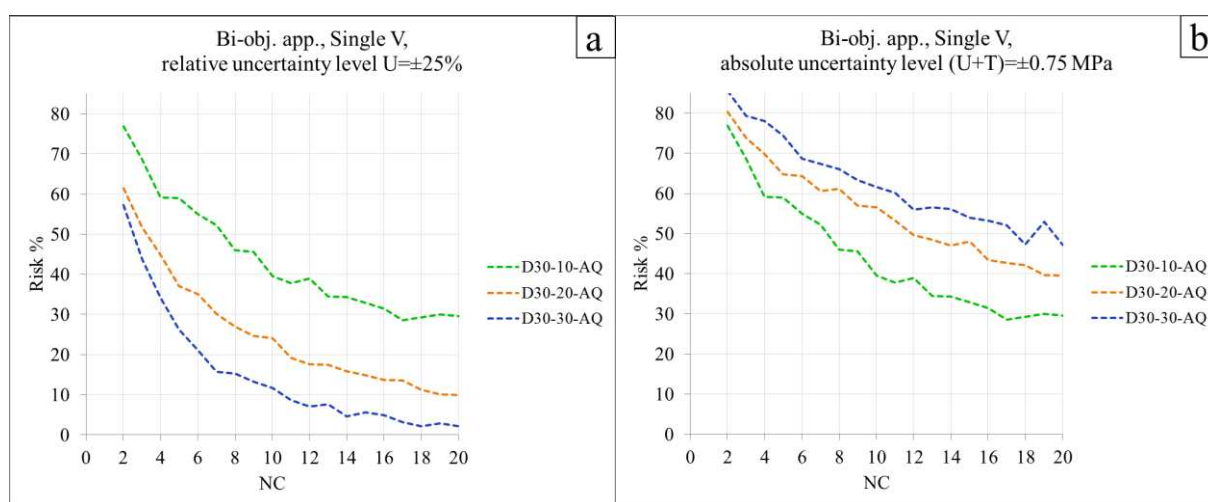


Figure 6.6 Comparing the risk curves obtained from absolute and relative uncertainty levels for the case of strength variability estimation with: average quality measurements, $NT=100$, and random core selection

6.2.4 Studying the effect of the quality of measurements on the quality of assessment

The risk curves for mean strength estimation and those for strength variability estimation shown in Figures 6.4, 6.5, and 6.6 are corresponding to datasets having average quality measurements (AQ). In order to study the effect of quality of measurements, it is necessary to construct the risk curves that correspond to other quality of measurements (i.e. HQ and LQ). To this end, the CDF curves obtained from the analysis of datasets with HQ and LQ, see Table 6.1, have been used to build the risk curves corresponding to these qualities of measurements with respectively $U=\pm 10\%$ for mean strength assessment and $U=\pm 25\%$ for strength variability assessment.

Regarding the mean strength assessment, the general behavior of these new risk curves are similar to that shown in Figures 6.4 for the case of average quality measurements, i.e. both approaches (regression and bi-objective) show good ability to provide a reliable assessment of mean strength. Furthermore, the type of NDT technique (pulse velocity or rebound hammer) that combined with DT (cores) to identify the model has no significant effect on the risk values. Consequently, Figure 6.7 shows the new risk curves that correspond to quality of measurements (HQ and LQ) for the regression approach and when NDT technique is single V . In addition, for comparison purposes, the risk curves corresponding to the same case (but with AQ) and presented in Figure 6.4a are also shown in Figure 6.7.

Regarding the strength variability assessment, the situation with HQ and LQ is similar to that shown in Figure 6.5 for the datasets with AQ, i.e. the bi-objective approach is significantly better than the regression approach. Furthermore, using ultrasonic pulse velocity or rebound hammer as NDT technique has no significant effect on the risk values. As a result, the risk curves for the datasets with HQ and LQ and corresponding to the bi-objective approach with the case of single V are shown in Figure 6.7 beside the curves corresponding to the AQ that already presented in Figure 6.5d.

In fact Figure 6.7 shows, for datasets with wide range of mean strength and strength variability, the relation between the risk values (or quality of assessment) and: the quality of measurements, the number of test locations for cores (for the case of random core selection), the true in-situ concrete variability, the approach used to identify the model, and the concrete property required to be assessed (mean strength or strength variability).

The risk curves in Figures 6.7a, 6.7b, and 6.7c (concerning the mean strength assessment with different quality of measurements) show the same tendencies i.e. the risk values decrease as NC increases and that the assessment of mean strength is easier for a less variable concrete. However, as shown, the quality of measurements has an important effect on the risk values which decrease dramatically as the quality of measurements is changed from LQ to HQ. As an example, for $NC=3$ for dataset D20-30-LQ the risk value is 48% however this value decreases to 35% for the dataset D20-30-AQ and reaches 19% for dataset D20-30-HQ. Furthermore, for each NC value, the differences between the risk values corresponding to various datasets having the same quality of measurements decrease as this quality is varied from LQ to HQ.

Figures 6.7d, 6.7e, and 6.7f display the risk curves for strength variability assessment that correspond to the three qualities of measurements (HQ, AQ, and LQ). Like Figure 6.5, the risk values decrease with increasing NC (but with a slower rate than those in Figures 6.7a, 6.7b, and 6.7c) and the assessment is more reliable for the more variable concrete. Moreover, the risk values decrease as the quality of measurements is varied from LQ to HQ but with less significant reduction as compared with the case of mean strength estimation. As an example for $NC=3$, the risk value for dataset D20-30-LQ is 50% and it becomes 45% for D20-30-AQ and reaches 33% for D20-30-HQ. As shown in Figures 6.7d, 6.7e, and 6.7f, the risk values for

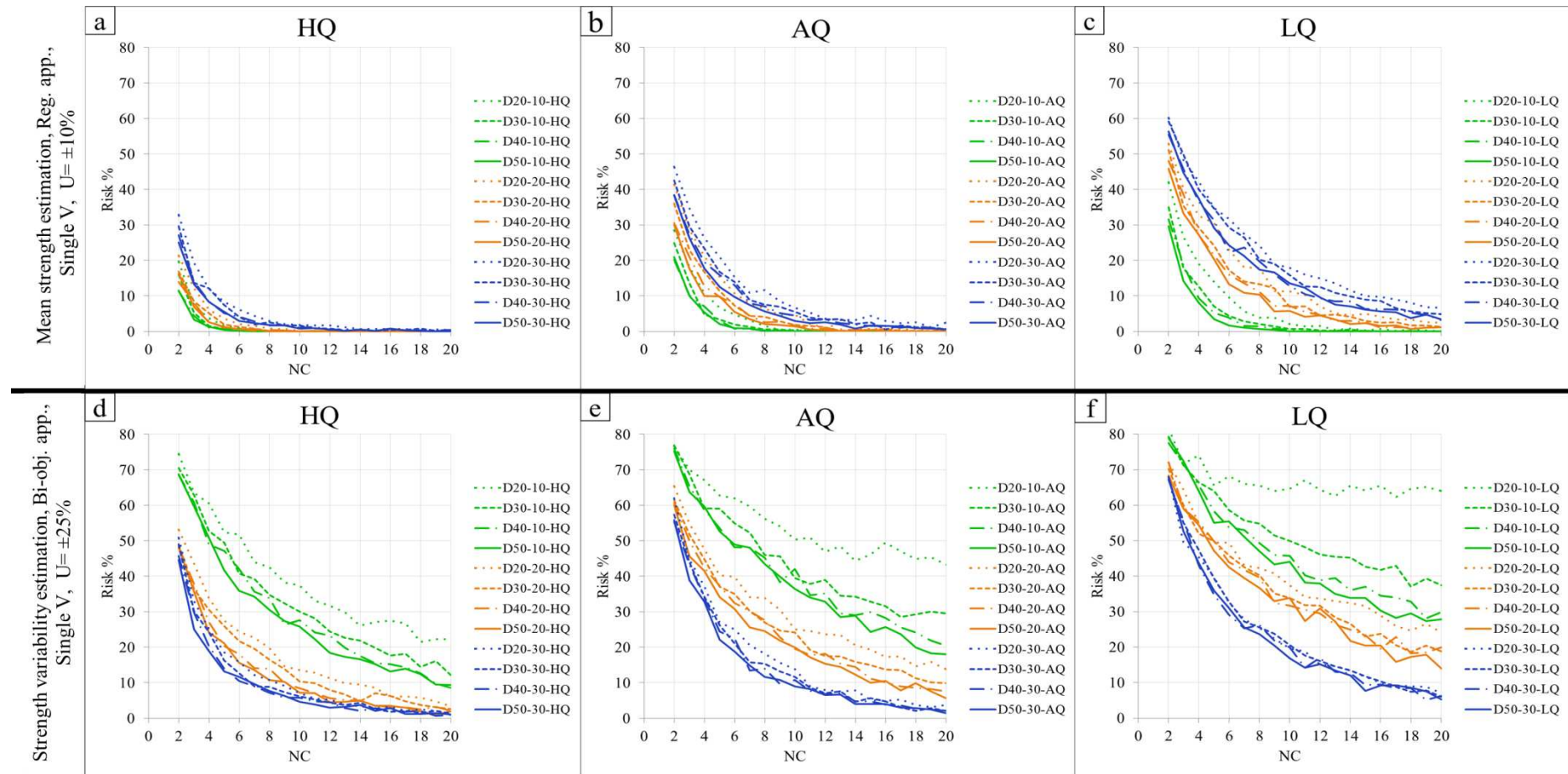


Figure 6.7 Studying the effect of quality of measurements on the risk curves corresponding to each of mean strength and strength variability estimations with $NT=100$, and random core selection

the case of strength variability assessment are high which indicates the difficulty to ensure a reasonable confidence level (confidence level=1 - risk) associated with the predicted value. As an example for risk value $\leq 10\%$ and $NC \leq 20$, we can assess the concrete variability for concrete having ($CV(f_{ist})=30\%$ or 20%) for HQ measurements while for AQ and LQ measurements the ability of assessment is restricted to concrete having ($CV(f_{ist})=30\%$). This means that to assess the strength variability in the concrete having ($CV(f_{ist})=10\%$ or 20%) it is necessary either to use $NC \gg 20$ or to increase the selected value of $U\%$ (i.e. the acceptance of wider error interval associated with the estimated value). Figure 6.8 shows the effect of selecting larger value for U ($U=\pm 35\%$) on the risk values for the case of HQ measurements, Figure 6.8a, and the case of LQ measurements, Figure 6.8b. As shown, the situation is improved for both cases in which the risk values are lower than the corresponding values in Figures 6.7d and 6.7f.

Regarding Figure 6.7e the risk curve of D20-10-AQ diverges slightly from the other curves and it seems less sensitive to the increase in NC . Whilst the curve of dataset D20-10-LQ, Figure 6.7f, is nearly horizontal when $NC \geq 5$ (i.e. its behavior is independent of NC) with large risk value of 65% . This situation corresponds to the case of low variability of concrete ($\bar{f}_{ist}=20$ MPa and $CV(f_{ist})=10\%$) and illustrates that, in such a case, low quality measurements are unable to capture the concrete strength variability. This behavior is due to the fact that the noise in these measurements makes them unable to detect the low variability of concrete ($s(f_{ist})=2$ MPa) and to provide a reliable estimation with a small margin of error ($(U \times T) = \pm 0.5$ MPa). Furthermore, the acceptance of larger value for the uncertainty level $U\%$ cannot improve the ability to capture the strength variability for D20-10-LQ as shown in Figure 6.8b.

From the present discussion, there is evidence that among the factors affecting the quality of assessment, the quality of measurements (or within-test variability) plays the most important role.

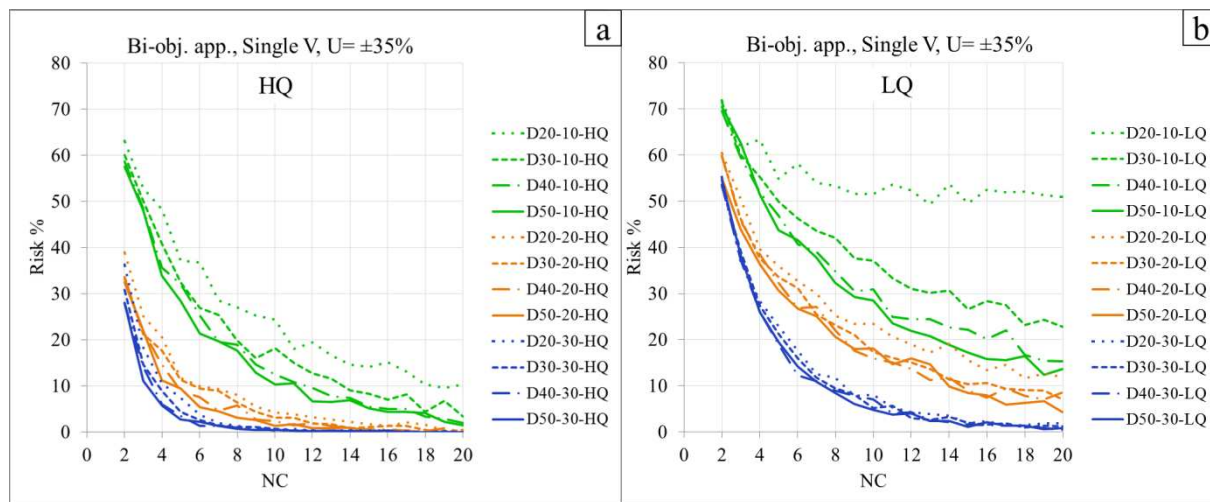


Figure 6.8 Risk curves corresponding to the uncertainty level $U=\pm 35\%$ for the strength variability estimation with: two different quality of measurements (HQ and LQ), $NT=100$, and random core selection

6.2.5 Studying the effect of the way of selection the NC test locations on the quality of assessment

All the previous results of this chapter are limited to the case of random selection of core locations. However these NC locations can be defined in accordance with certain conditions applied on the NDT results “conditional selection”. The idea of the conditional selection is to select the NC test locations that cover, as much as possible, the whole range of the NDT measurement distribution. The way of selecting the NC test locations for cores was studied in Chapter 4 (Subsection 4.3.3.3) when analyzing the results of an international benchmark carried out by RILEM TC-249 committee. It has been concluded that using the conditional selection has a beneficial effect on the quality of assessment. Since in Subsection 4.3.3.3 the analysis was limited to one dataset, we will deepen the analysis here by applying the conditional selection on other datasets having different characteristics.

To this end, the analysis procedure described in Subsection 6.2.2 was duplicated here for the case of conditional core selection. It means that the datasets from Table 6.1 was re-analyzed following the algorithm in Figure 3.3 but with the option of conditional cores. The same method proposed in Chapter 4 for selecting the cores conditionally has been applied herein. This method is as follows: (a) rank all NDT test results from the lowest to the highest value, (b) subdivide the set into NC subsets, (c) select a test location where the NDT value is closest to the median value of each subset. Except the way of core selection, the other analysis characteristics applied with random selection remain unchanged in the analysis with the conditional selection (i.e. $NT=100$ for NDT measurements, linear model, $NI=1000$, and NC varying from 2 to 20). The final resulting CDF curves were used to construct the risk curves corresponding to the case of conditional core selection.

For the assessment of mean strength for dataset with HQ, Figure 6.9a shows three selected risk curves corresponding to the conditional selection of cores beside the risk curves for the same datasets obtained from using the random selection which were earlier provided in Figure 6.7a (for the case of pulse velocity). These datasets are D20-10-HQ, D20-20-HQ, and D20-30-HQ which are selected because each one represents the extreme curve within its group (i.e. the group that has the same $CV(f_{ist})$ value), see Figure 6.7a. For the same reason, the risk curves of the corresponding datasets but with AQ and LQ are plotted on Figures 6.9b and 6.9c respectively for mean strength estimation. From these three figures, there is evidence that the conditional selection of cores can considerably reduce the risk values (or improve the quality of assessment) for all datasets when the quality of measurements is HQ. However, the significant of this effect is reduced as the quality of measurement reduces to AQ or LQ. Furthermore, the effect of using conditional selection is significant only for low NC values ($NC \leq 10$ or 12) and it becomes negligible for the high NC values.

Regarding the strength variability estimation, the risk curves of conditional selection that correspond to the same datasets are plotted on Figures 6.9d, 6.9e and 6.9f besides the risk curves of random selection from Figures 6.7d, 6.7e and 6.7f. As shown, the use of conditional selection reduces the risk values with amounts of reduction increase as the quality of measurements changes from LQ to HQ. On contrary to the situation in Figures 6.9a, 6.9b and 6.9c, in general, the effect of using conditional selection (for strength variability estimation) is more significant for high NC values than for low NC values.

Consequently, since proposed method for conditional selection induces no additional cost and since it can improve the reliability of assessment for both estimated quantities (mean strength and strength variability), this process must be strongly recommended. It must be noted that this method is fully compatible with in-situ concrete assessment: it only requires that the preliminary NDT investigation is carried out before the core location is easily decided.

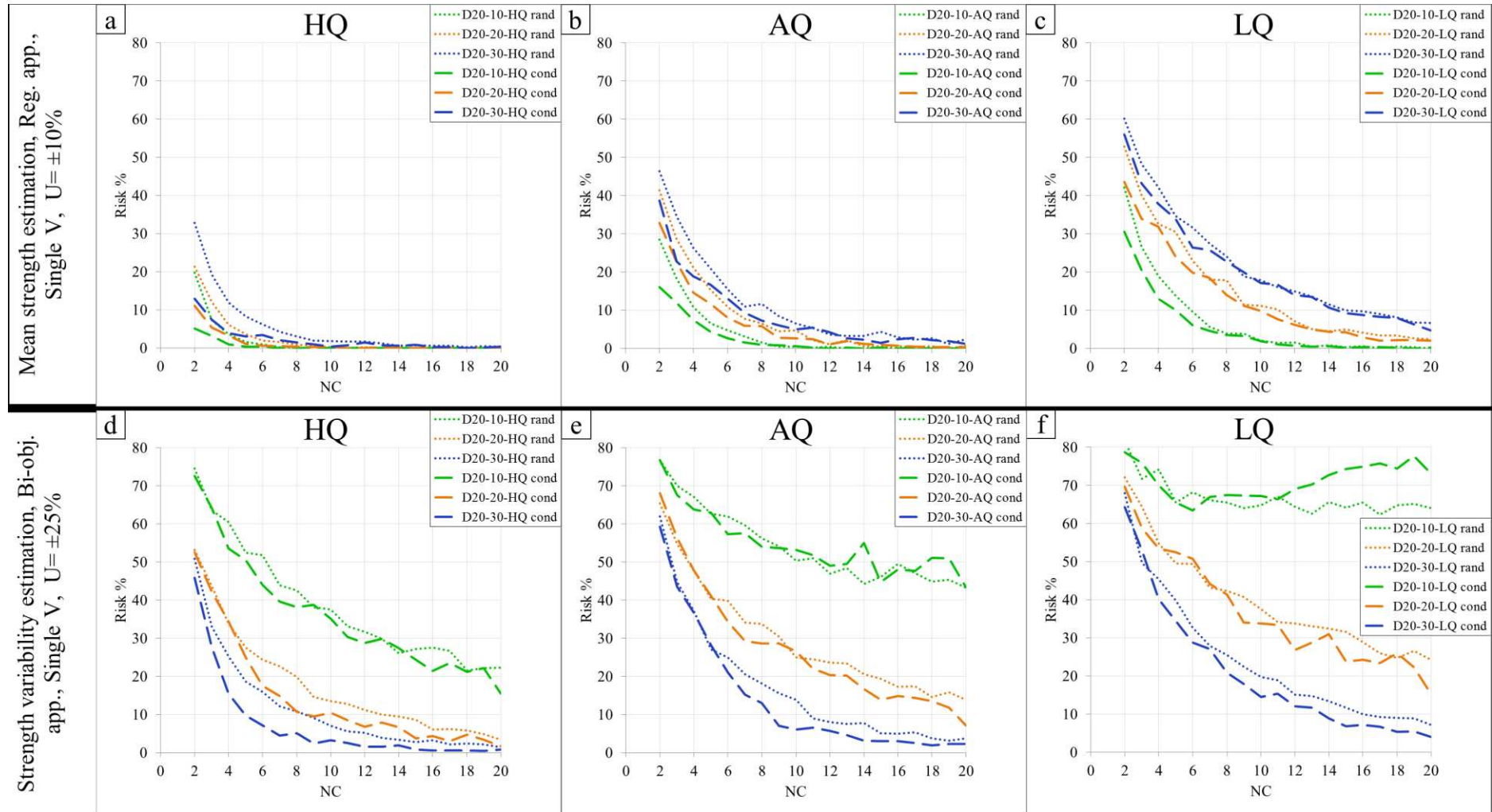


Figure 6.9 Studying the effect of the way of core location selection on the risk curves corresponding to each of mean strength and strength variability estimations with $NT=100$, and for the three quality of measurements: HQ, AQ, and LQ

6.2.6 What is the minimum number of test locations for cores that can ensure a specific quality of assessment?

In real practice, to carry out an investigation program, engineer needs to specify the number of test locations for cores, NC , which is necessary to identify adequate model between concrete strength and NDT measurements. For many reasons (economical, destructive, etc.) this number should be as small as possible. (ACI 214.4R, 2003) provides a formula for calculating the NC (with 95% confidence) corresponding to a specific concrete strength variability and an accepted predetermined error between the estimated mean strength and population mean strength as given in Eq. (2.1). The shortcomings of this formula are that: it is restricted to the case of mean strength assessment, it does not consider the quality of measurements, it provides high NC values because it was essentially derived for the case of assessment using cores only.

According to the requirements of the European standard (EN 13791, 2007), the minimum number of cores is 18 for regression analysis approach (Alternative 1 as so-called in this standard). (ACI 228.1R, 2003) requires a minimum of six to nine test locations for cores (with two cores that should be drilled from each location). As shown, each of these standards provides a single value for NC without any consideration about the type of concrete or the estimated quantity and without any information about the error associated with the estimated value.

From the analysis of quality of assessment that has been carried out in this chapter, there is evidence that the minimum NC is related to many influencing factors: true in-situ concrete variability, quality of measurements, the prescribed error (or uncertainty level), the way of selecting the core locations, the quantity to be assessed, and the confidence level in the estimated value. Using Figure 6.9, the minimum values that related to these limitations can be derived as shown in Table 6.2. Any value of NC in this table can be interpreted as the NC value that leads to estimated values (for example: the mean strength) which we are sure that 90% (confidence level = 90% as an example) of the estimated values will have an error $\leq \pm 0.10\bar{f}_{cest}$. In order to show the effect of the confidence level on the NC values, Table 6.2 provides the values corresponding to two different confidence levels: confidence level $\geq 90\%$ and confidence level $\geq 95\%$.

From this table, regarding the mean strength estimation, it is clear that the required minimum NC increases in two directions: as the true in-situ concrete variability varies from 10% to 30% and as the quality of measurements changes from HQ to LQ. However, it decreases as the conditional core selection method is applied. Therefore, for confidence level $\geq 90\%$ (or risk $\leq 10\%$), the NC values range from 3 to 15 which are all less than that required by the (EN 13791, 2007) standard ($NC=18$).

Regarding the strength variability assessment, the NC increases as the quality of measurements changes from HQ to LQ and, on contrary to the case of mean strength, as the in-situ concrete variability reduces from 30% to 10%. However, NC decreases as the conditional core selection method is applied. As shown, for many cases, the value is marked by “N/A” i.e. “not available” which means that the required NC value is either greater than 20 or that we are unable to provide an estimation within the prescribed error for this quality of measurements and confidence level. For confidence level $\geq 90\%$, the lower limit of NC range is 5 while the upper limit is unknown.

The wide range of NC values provided in Table 6.2 for each estimated quantity indicates that the minimum NC is not a simple unique value that can be applied in all situations, as provided by the standards (ACI 228.1R, 2003; EN 13791, 2007). Our conclusion is that this issue needs to be seriously revised in the current standards in order to provide minimum NC values (in

tables or formulas), like Table 6.2, with detailed information about their application limitations.

Table 6.2 Minimum NC as a function of true in-situ concrete variability, quality of measurements and the way of core locations selection for sample size NT=100, and NC ≤ 20, (N/A means value is not available)

Confidence level ≥ 90%								
Estimated quantity	Error (uncertainty level)	CV(f_{ist})	HQ		AQ		LQ	
			rand	cond	rand	cond	rand	cond
\bar{f}_{cest}	$\pm 0.10\bar{f}_{cest}$	10%	3	3	5	4	6	5
		20%	4	3	7	6	11	10
		30%	5	3	9	7	15	15
$s(f_{cest})$	$\pm 0.25s(f_{cest})$	10%	N/A	N/A	N/A	N/A	N/A	N/A
		20%	13	9	N/A	20	N/A	N/A
		30%	9	5	11	9	16	14
Confidence level ≥ 95%								
Estimated quantity	Error (uncertainty level)	CV(f_{ist})	HQ		AQ		LQ	
			rand	cond	rand	cond	rand	cond
\bar{f}_{cest}	$\pm 0.10\bar{f}_{cest}$	10%	4	3	6	5	8	7
		20%	5	4	10	9	14	13
		30%	7	4	12	10	N/A	20
$s(f_{cest})$	$\pm 0.25s(f_{cest})$	10%	N/A	N/A	N/A	N/A	N/A	N/A
		20%	19	15	N/A	N/A	N/A	N/A
		30%	13	9	16	13	N/A	20

6.3 Analysis of quality of assessment using real datasets

In this section, in order to confirm the robustness of the analysis using synthetic datasets and to emphasize our findings, the quality of assessment was studied using real datasets of NDT measurements (rebound hammer technique) and destructive tests. More than 2500 test results pairs (R, f_c) from seventeen real datasets (S21 to S37 in Table 3.1) were considered. These datasets cover a wide range of true mean strength \bar{f}_c (36-77 MPa) and true strength variability (in terms of strength coefficient of variation) $CV(f_c)$ (11-33%).

Similar to the case of synthetic datasets, two estimated quantities (mean strength and strength variability) had been the target of the assessment using two model identification approaches: regression and bi-objective. The factors analyzed with synthetic datasets, in Section 6.2, were considered herein for analyzing the quality of assessment with the real datasets. These factors are: quality of measurements (or within-test variability of rebound measurements), true concrete strength variability, NC, way of selection the NC test locations (random or conditional), and the model identification approach (regression or bi-objective). We have published this analysis using real datasets in “Construction and building Materials” Journal. The paper is available in Appendix D. However, its most important results are summarized in Table 6.3, these results confirm that:

- the assessment of concrete strength variability using the regression approach provides limited or unreliable assessment,
- regarding the effect of *NC*, reducing *NC* leads to less reliable estimations of the mean strength and concrete strength variability whatever the model identification approach (regression or bi-objective),
- reducing the within-test variability of the rebound measurements (improving the quality of measurements) leads to more reliable assessment,
- regarding the true strength variability of the concrete under investigation, less variable concrete gives more reliable estimation of concrete mean strength. However the estimation of the concrete strength variability is less reliable for a less variable concrete (this statement is based on considering relative uncertainty level or error),
- regarding the method of conditional selection of core locations, it improves the reliability of assessment for both estimated quantities (mean strength and strength variability).

All these results are in full agreement with our findings with synthetic data, which confirms the ability of synthetic datasets to draw conclusions similar to that derived from real datasets.

Table 6.3 Summary of the effects of factors considered in this study on the quality of assessment analyzed using real datasets

factor	mean strength estimation		strength variability estimation	
	regression	bi-objective	regression	bi-objective
Reducing <i>NC</i>	worsening	worsening	approach provides unreliable assessment	worsening
Reducing within-test variability of rebound measurements (improving quality of measurements)	improvement	improvement		improvement
Reducing the true concrete strength variability	improvement	improvement		worsening
Using conditional selection method for <i>NC</i> test locations instead of random selection	improvement	improvement		improvement

6.4 Conclusions

Reliability of the assessment is a real challenge for engineers when assessing the in-situ strength of concrete structures. Therefore, in this chapter, the quality of assessment was analyzed and it was correlated with its affecting factors: number of test locations for cores (NC) used to identify the conversion model, quality of measurements (within-test variability), the way of selecting the core locations, the in-situ concrete variability, the type of NDT technique, using single or combination of NDT techniques, and the model identification approach.

To this end, a large campaign of synthetic datasets (36000- (V, R, f_c) test results) having different characteristics were generated for the analysis. The analysis findings with the synthetic datasets have been confirmed by using additional real datasets but with rebound hammer only (2500- (R, f_c) test results selected from Table 3.1).

The following conclusions can be drawn:

- the bi-objective and the regression approaches show a good ability to provide a reliable assessment of mean strength. This reliability is improved as NC increases with more significant improvement at low values of NC . Furthermore, the combination of NDT techniques ($V+R$) improves also the reliability as compared with that of single technique, however this benefit should be compared with the growing in the cost due to the combination of techniques,
- regression approach is limited in some cases and handicapped in others for assessing the strength variability. Moreover, using combination of NDT techniques ($V+R$), can improve the quality of assessment for datasets having ($CV(f_{ist})=20\%$ and 30%) as compared with the case of single technique. However, this improvement in the quality of assessment should be compared with the additional cost due to the combination of techniques. On contrary, the bi-objective approach shows a good ability to assess the strength variability with more significant effect of the increase in NC on the quality of assessment. Furthermore, the quality of assessment using bi-objective approach is better than that provided by regression approach even in the case of combination of NDT techniques,
- Assessing the mean strength in less variable concrete is more reliable. On contrary, the assessment of concrete strength variability in less variable concrete is less reliable (this statement is based on considering a relative uncertainty level or error),
- there is evidence that among the factors affecting the quality of assessment, the quality of measurements (or within-test variability) plays the most important role,
- using the method of conditional selection of core locations instead of the random selection improves the quality of assessment for both estimated quantities (mean strength and strength variability) while it induces no additional cost. Therefore, this process is strongly recommended,
- the minimum NC is not a simple unique value that can be applied in all situations, as recommended by the standards (ACI 228.1R, 2003; EN 13791, 2007). Consequently, this issue needs to be seriously revised in the current standards in order to provide minimum NC values (in tables or formulas), with detailed information about their application limitations. Our risk curve concept can be used to provide more detailed practical information regarding the recommended NC values.

6.5 Recommendations for better practice

In Chapter 4, the inherent characteristics of the current assessment methodology were analyzed while in Chapter 5 a new model identification approach was proposed in order to improve the reliability of assessment. Chapter 6 has been devoted to analyze the quality of assessment and to relate it with the inherent characteristics of the assessment methodology. From the analyses carried out in these three chapters, we can provide the following recommendations for improving the engineering practice:

- a. Quality of measurements (or within-test variability) should receive special attention on site by improving it as far as possible because this improvement has an important effect on the quality of assessment. In any case, it can be quantified through the simple repetition of the measurement process within the test location.
- b. Number of test locations for cores should be selected depending on: the in-situ concrete variability, the quality of measurements, the quantity to be assessed, the accepted uncertainty level (error) in the estimated value, and the confidence level. Since detailed recommendations are not available in the current standards, Table 6.2 can be used to provide the first estimate.
- c. The method of conditional selection of core locations (NDT based selection) should be used to define the core locations since it can improve the quality of assessment without any additional cost.
- d. The selection of a model type (linear or nonlinear) does not deserve a lot of attention since the other sources of uncertainty have a more significant effect on the quality of assessment than this source of uncertainty.
- e. For the assessment of mean strength the regression and bi-objective approaches can be applied to identify model between concrete strength and NDT measurements. Calibration approaches can be used but with attention that should be paid to the selection of the prior model which should be already derived from a concrete similar to that under consideration and has the range of NDT values close to that obtained with the concrete under consideration.
- f. For the assessment of concrete strength variability, the bi-objective approach should be applied for this purpose.
- g. To assess the quality of the identified model, the prediction error ($RMSE$) should be calculated using additional cores extracted for this purpose since the fitting error leads to misleading conclusions. Furthermore, for comparing the qualities of models identified from concretes with different variabilities, the relative prediction error ($RMSE/s(f_c)$) should be used instead of the absolute error (i.e. $RMSE$).
- h. The use of combination of ultrasonic pulse velocity and rebound hammer should be restricted to the case of high number of cores and when the two techniques have close quality of measurements.

Consequently, according to these recommendations, the assessment methodology for evaluating concrete strength in the existing structures can be improved to be as shown in Figure 6.10.

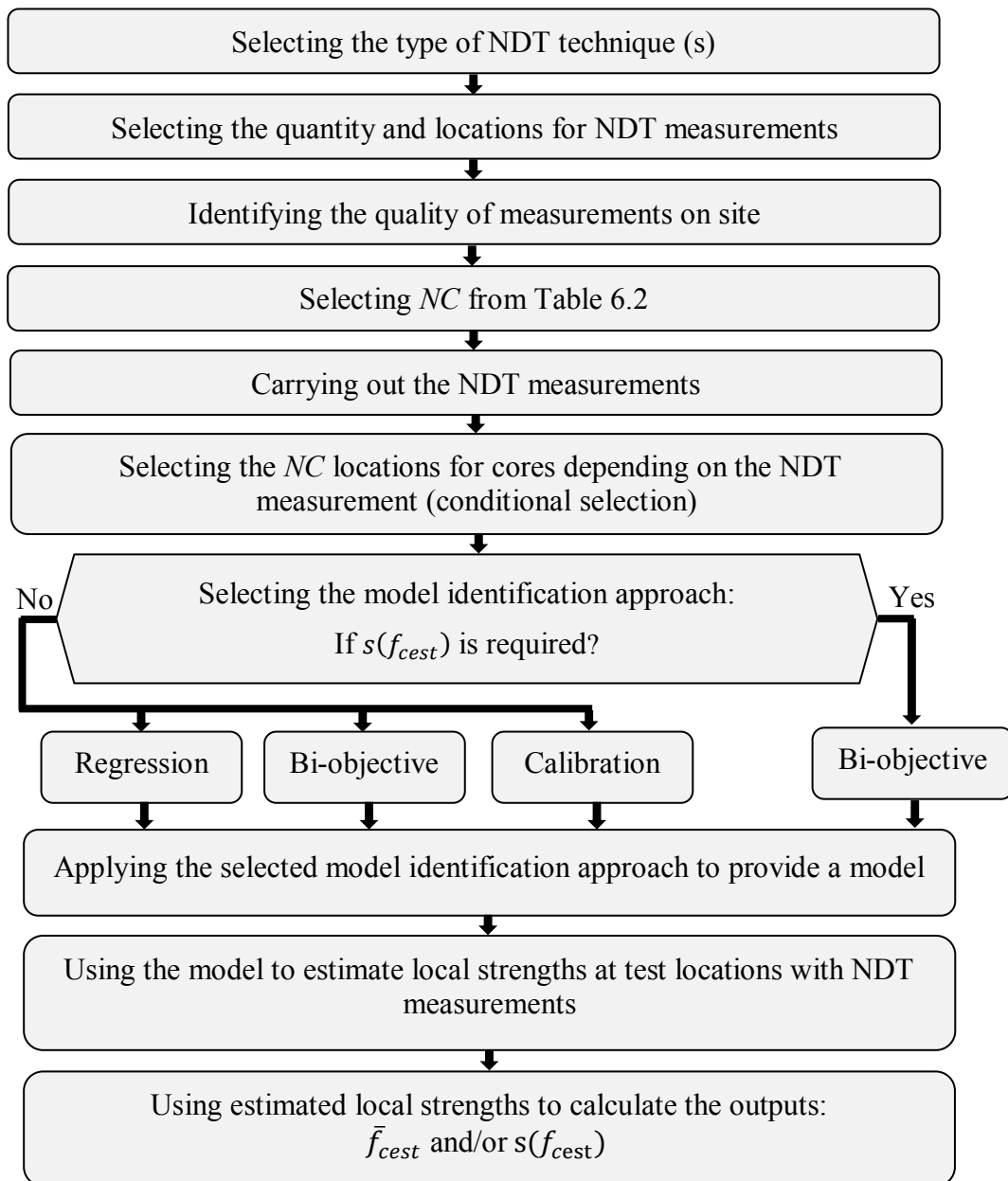


Figure 6.10 Assessment methodology for evaluating concrete strength in the existing structures that proposed in the present study

CHAPTER SEVEN

CONCLUSIONS AND PERSPECTIVES

Assessment of in-situ strength in the existing structures is always a challenge for engineers. The current methodology consists in carrying out nondestructive measurements (NDT) like rebound hammer or/and pulse velocity on several test locations in the structure. Then, cores are extracted from some of these locations for compressive tests in order to implement a relationship “conversion model” between the compressive strength and NDT measurements. The conversion model is used to estimate the local strength value at each test location using the corresponding NDT value. Then the estimated mean strength and/or estimated strength standard deviation (concrete strength variability) values are calculated. However, the reliability of these estimated values is always a questionable issue because of the uncertainties associated with the strength predictions based upon NDT measurements. To improve the reliability, the uncertainties must be reduced by specifying and controlling their influencing factors. In this study, the focus was given mainly to the controlled factors (i.e. factors that can be modified or changed) in order to improve the quality of assessment. Therefore, the objective of this thesis was to study the current assessment methodology in order to provide practical recommendations that can improve the reliability of assessing the in-situ strength in existing concrete structures by nondestructive tests and cores.

The first axis of this thesis was devoted to analyze the current assessment methodology. This analysis was subdivided into three main stages. The first one focused on studying the effects of several of the methodology controlled factors. The considered factors were: the number of test locations for cores (NC) which are used for identifying the conversion model, the quality of NDT measurements (within-test variability), the true in-situ variability of the concrete strength, the model type, and the combination of NDT techniques. For analyzing these effects, three sources of datasets (in-situ data, laboratory studies data, and generated synthetic data) were considered. $RMSE$ and r^2 errors were calculated for the estimated local strengths in order to be as indicators about the quality of fitted model and its ability to estimate strength. The errors were calculated in two ways: considering only the test locations with cores that used to identify the model (fitting error), and considering all test locations with NDT measurements but without cores (prediction error).

From the analysis at this stage, we showed that the errors in the estimated strengths are affected by the number of test locations for cores, NC . For small NC , the fitting error is small while the prediction error is large, but as NC increases the prediction error is progressively reduced, however it is always larger than the fitting error. This important issue warns engineers to avoid using the fitting error (usually r^2) as an indication about the quality of assessment and they should use the prediction error for this purpose. Furthermore, it was shown that the true in-situ concrete strength variability, $s(f_{ist})$, has significant effect on the quality of assessment. As $s(f_{ist})$ increases, the predictive $RMSE$ increases while the ratio $RMSE/s(f_{ist})$ decreases. Consequently, one must pay attention to how the quality of the assessment is quantified: looking at the absolute error, $RMSE$, or relative error, $RMSE/s(f_{ist})$,

would lead to different conclusions. Regarding the quality of measurements, we found that it has significant effect on the prediction error so it deserves a lot of attention on site. As compared to the previous three effects, it was found that the effect of model type (linear or power) is less significant since it is generally limited to the case of small NC with smaller error provided by the linear model. The last effect that was analyzed in this stage was the efficiency of combining two NDT techniques (pulse velocity and rebound hammer) with the destructive tests (cores). It was shown that the combination is not always efficient and its efficiency depends on NC and the quality of measurements. This explains the reasons why in the literature the efficiency of combining NDT is controversial (combination is said to bring some added-value in some cases but not in others).

In common practice, one of the challenges for engineers is the selection of an assessment strategy among many possible strategies having the same investigation budget. Therefore, in the second stage of the first axis, several strategies defined by experts in an international benchmark devoted to assess the concrete mean strength and strength variability were analyzed. The efficiency of these strategies was studied by simulating their application to a series of 1000 synthetic buildings having the same statistical properties and by defining the indicators of the quality of assessment.

It was shown that the scatters in the estimated mean strength and estimated strength variability values confirm that it is impossible to evaluate a strategy from a unique result “one shot study” as provided by the contributors to this benchmark. These scatters also emphasize that the variabilities in the estimated mean strength and estimated strength variability are significant and cannot be neglected whatever the assessment strategy. The other finding of this analysis was that the assessment of strength variability is a much more difficult challenge than the assessment of the mean strength. The strategies analyzed here have shown a handicap for assessing the concrete strength variability. This analysis has also confirmed the significant effect of the quality of measurements on the quality of assessment. After studying the effect of the way of selecting the core locations, a method was proposed for selecting these locations depending on NDT measurements “conditional selection”. The main finding of this analysis was that this method can improve the quality of assessment without any additional cost, therefore it is strongly recommended for real practice.

The third stage of the first axis was devoted to analyze the existing model identification approaches (regression approach and calibration approaches (Δ -method and k -method)). These approaches were applied to eight real datasets (4 with rebound measurements and 4 with pulse velocity measurements) in order to assess the mean strength and strength variability. From comparing the quality of assessment of these approaches, it was found that all approaches (regression, calibration) can efficiently estimate the mean concrete strength. Furthermore, increasing NC can significantly improve the estimation efficiency. On contrary, it was shown that the regression approach remains limited in capturing the concrete strength variability, it always underestimates the concrete variability and increasing NC has no significant effect on the quality of estimation. Regarding the calibration approaches, it was found that they produce unstable results (quality of assessment varies as the prior model is changed) so their efficiency depends on chance. Moreover, increasing NC could not improve the results of calibration approaches especially for Δ -method which was proved its independence from the effect of NC . These findings were obtained from both type of NDT techniques (rebound hammer or pulse velocity) which indicate the generality of them.

In the first axis, the handicap of the existing approaches for assessing strength variability was confirmed. Therefore, the second axis of this thesis was directed to propose a new model identification approach “bi-objective” in order to capture concrete strength variability in addition to mean strength. The derivation of the bi-objective approach was provided for both

linear and power models. The validity of this approach was proved by applying it on several datasets (some datasets with rebound hammer and other with pulse velocity) from in-situ and laboratory studies in order to predict the mean strength and concrete variability. Moreover, the prediction ability of the bi-objective approach was compared with that of the existing approaches.

It was found that the bi-objective approach shows high efficiency (on contrary to other approaches) in capturing the concrete strength variability in addition to its efficiency in capturing the mean strength with better efficiency as NC increases. It was shown that the bi-objective approach is valid with both the ultrasonic pulse velocity and rebound hammer measurements. Furthermore, from mathematical viewpoint, it can be used with any NDT technique combined with DT technique (cores) in order to identify a conversion model. Regarding the model type, it was found that the bi-objective approach works with power model as well as with the linear model but the uncertainty arising from using power model is larger than that of linear model when NC is small.

In the third axis of this thesis, the quality of assessment was analyzed and it was correlated graphically with its affecting factors: number of test location for cores (NC) used to identify the model, quality of measurements (within-test variability), the way of selecting the core locations, the true in-situ concrete variability, the type of NDT technique, using single or combination of NDT techniques, the model identification approach, and the quantity to be assessed (mean strength or strength variability). A large campaign of synthetic datasets (36000- (V, R, f_c) test results) having different characteristics were generated for this analysis. Furthermore, in order to confirm the robustness of the analysis findings with the synthetic datasets, the analysis of the factors listed above were also done using real datasets but with rebound hammer only (2500- (R, f_c) test results).

From these analyses, the significant effects of these factors on the quality of assessment were confirmed with the most important role played by the quality of assessment. Furthermore, it was shown that, for prescribed relative error or uncertainty level around the true value ($\pm U\%$), assessing the mean strength in less variable concrete is more reliable. On contrary, the assessment of concrete strength variability in less variable concrete is less reliable. From the graphical correlation between the quality of assessment and the affecting factors, we derived the minimum values for NC . It was found a wide range of NC values for each estimated quantity. As a result, the minimum NC is not a simple unique value that can be applied in all situations, as recommended by the standards (ACI 228.1R, 2003; EN 13791, 2007). Consequently, we think that this issue needs to be seriously revised in the current standards in order to provide minimum NC values (in tables or formulas) with detailed information about their application limitations.

Since the proposed bi-objective approach gives encouraging results with single NDT techniques for assessing efficiently the mean strength and strength variability, our first perspective is to develop this approach to be applicable for the case of combination of NDT techniques. We think that this development provides more improvement in the quality of assessment.

The second perspective is studying the effect of the number of test locations for NDT measurements (NT) (i.e. sample size) on the quality of the assessment. From statistical aspects the scatter of sampling distribution of mean strength (for example) depends on the ratio σ/\sqrt{NT} , where σ is the standard deviation of concrete strength population. Therefore increasing the sample size leads to reduce one of the uncertainty sources (sampling uncertainty) and, as a result, reduces the scatter in the estimated values. This factor was not

considered in the present study, however we think that it deserves to be considered in future works.

The quality of assessment obviously improves when both *NC* and quality of measurements increase. However, when the total amount of resources is fixed (fixed investigation budget), the optimal balance between these factors does not follow simple rules and it deserves to be studied through a more comprehensive analysis.

INTRODUCTION GÉNÉRALE

Problématique

Le béton est un matériau composite couramment utilisé dans le domaine de la construction en raison de sa résistance et de sa durabilité, de sa disponibilité, du faible prix des matières premières, et du fait de sa capacité à se couler dans des formes architecturales complexes. Par conséquent, de nos jours, il existe un nombre important de structures en béton et, jour après jour, ce nombre est en constante évolution. Parmi les propriétés mécaniques et physiques du béton, sa résistance à la compression est la propriété la plus importante car elle est essentielle pour concevoir un élément structurel ou calculer sa capacité portante. La résistance à la compression n'est pas une valeur fixe pendant la durée de vie du béton en raison de plusieurs effets comme le durcissement au jeune âge et des fissures internes développées (en raison des effets environnementaux ou de chargement) à l'âge avancé. Par conséquent, l'évaluation de la résistance à la compression in situ du béton est requise dans de nombreuses situations telles que:

- le contrôle de la qualité du béton pour assurer la conformité avec les spécifications pour le cas de nouvelles constructions,
- les litiges après la non-conformité du matériau testé,
- le doute sur la qualité du béton dans une nouvelle construction,
- la modification de l'utilisation (ou de la fonction) d'une structure afin de calculer avec précision la capacité structurelle,
- les dommages résultant d'un incendie, d'une fatigue, d'une surcharge ou d'une dégradation par l'environnement,
- la question de la réhabilitation sismique qui se pose aujourd'hui dans plusieurs pays (comme l'Italie et la Turquie) et qui nécessite également une évaluation in situ précise de la résistance du béton dans les structures existantes (Biondi & Candigliota, 2008; Gunes, 2015).

Il est clair que les trois premières situations concernent les nouvelles structures alors que les trois dernières font l'objet d'une évaluation in situ dans les structures existantes. De nombreuses méthodes d'essai sont disponibles pour estimer la résistance in situ. Le test sur des carottes est la méthode la plus directe utilisée. Cependant, elle présente de nombreux inconvénients : coûteuses, parfois limitée en raison de l'accès difficile de la machine à carottage, seulement représentatif d'un petit volume du béton et il a un effet localement destructeur sur la structure (Meynink & Samarin, 1979).

D'autres méthodes d'essai utilisées pour évaluer la résistance in situ du béton sont les méthodes non destructives, telles que : le rebond, la vitesse d'impulsion ultrasonore, le test de pénétration, le pullout, etc. Les principales caractéristiques de ces méthodes (par rapport aux tests destructifs) sont leur simplicité, leur rapidité et leur faible coût. Cependant, leur principal inconvénient est que ce sont des méthodes indirectes puisqu'elles ne peuvent pas mesurer directement la résistance in situ. Ils mesurent d'autres paramètres physiques (ex : vitesse ultrasonore, indice rebond, force d'arrachement, etc.) et la force est dérivée d'un "modèle de conversion" de corrélation entre la résistance et les propriétés mesurées. En raison du fait que le béton est un matériau largement hétérogène, il n'existe pas un modèle unique pour tous les bétons (Breyse, 2012).

Par conséquent, la méthodologie actuelle repose sur l'établissement d'un modèle de conversion pour chaque cas particulier à l'aide de tests destructifs et non destructifs. Ensuite,

le modèle dérivé est utilisé pour estimer la résistance du béton aux emplacements de tests non destructifs. Cette méthodologie présente de nombreux degrés de liberté :

- Combien d'emplacements de tests sont nécessaires pour les prélèvements destructifs ?
- Quelle (s) méthode (s) non destructive (s) utiliser ?
- Combien d'emplacements de tests sont nécessaires pour les tests non destructifs ?
- Quelle est la manière de choisir les emplacements de tests ?
- Quelle est la forme mathématique du modèle de conversion ?
- Quelle approche d'identification du modèle ?
- Utiliser une ou combiner plusieurs méthodes non destructives ?
- Quelle est la qualité de l'évaluation de la résistance prédite ?
- Comment la qualité de l'évaluation peut-elle varier lorsqu'une réponse aux questions ci-dessus change ?

Les normes (ACI 228.1R, 2003; EN 13791, 2007) tentent de fournir des réponses à certaines de ces questions, mais elles ne fournissent pas une méthodologie générale qui permette de répondre à toutes ces questions.

La portée de ce travail

Le problème général présenté dans la section ci-dessus ne peut pas être résolu avec une seule thèse de doctorat, par conséquent cette thèse a eu pour objectifs de traiter une partie de cette problématique générale. Premièrement, cette recherche porte sur l'évaluation de la résistance in situ dans les structures existantes, car elle est plus essentielle aujourd'hui avec les besoins de la réhabilitation sismique. Par exemple, en Turquie, plusieurs millions de bâtiments ont été classés comme des bâtiments à risque et, par conséquent, des informations fiables sur leur résistance sont requises pour toute décision concernant le réaménagement ou le renouvellement (Gunes, 2015). Afin de faire la distinction entre les structures existantes et les structures nouvelles, on adopte la définition suivante pour les structures existantes, en accord avec les directives de la RILEM TC 249 ISC : les structures en béton armé existantes sont définies comme des structures où les effets d'âge (craquage, corrosion, etc.) et les barres d'acier de renforcement ont une grande influence sur les résultats des mesures non destructives et la résistance prédite. Une autre caractéristique des structures existantes est le manque dans les informations détaillées sur le béton utilisé.

La résistance mécanique du béton d'une structure réelle à n'importe quel emplacement de test est toujours inconnue. Par conséquent dans le présent travail, la valeur de référence pour cette résistance inconnue sera la résistance mesurée sur des prélèvements (carottes) à ce même emplacement de test.

Comme il a été indiqué dans la section précédente, il existe plusieurs techniques non destructives qui peuvent être utilisées pour évaluer la résistance à la compression du béton. Cependant, dans ce travail de doctorat, seules les méthodes du rebond et celle de la vitesse ultrasonore sont considérées car elles sont les méthodes les plus largement utilisées sur les structures en béton armé.

Dans la pratique, la résistance à la compression moyenne et la résistance à la compression caractéristique sont les valeurs les plus couramment évaluées, ces dernières étant des entrées importantes pour le calcul structurel. L'évaluation de la résistance caractéristique dépend de la résistance moyenne et de l'écart-type des résistances à la compression (variabilité de la

résistance du béton). La variabilité du béton est donc également une valeur requise. Le coefficient de variation, $CV(f_c)$, dû à la variation in-situ de la résistance du béton au sein d'une structure (c.-à-d. La variabilité du béton / résistance moyenne) est de 13% (ACI 214.4R, 2003). Cependant (Masi & Chiauzzi, 2013) ont trouvé une valeur $CV(f_c)$ de 21% dans un élément d'une structure. (Masi & Vona, 2009) ont étudié la variabilité du béton dans de nombreux bâtiments en Italie et ont observé que les valeurs probables de $CV(f_c)$ se situent entre 15 et 35%. (Pucinotti, 2013) a également montré que dans de nombreux cas le $CV(f_c)$ atteint 35%. C'est pourquoi l'évaluation de la variabilité du béton est nécessaire en plus de la valeur moyenne de la résistance. Le présent travail de doctorat porte donc sur l'estimation des valeurs de résistances locales, de la résistance moyenne et de la variabilité de la résistance du béton. La résistance caractéristique peut être déduite une fois que ces propriétés sont connues, mais elle n'est pas traitée dans le cas de ces travaux de thèse.

Lorsque la variabilité de la résistance du béton est élevée ou lorsque les informations disponibles indiquent que le béton dans la structure entière (ou les parties considérées) ayant une composition ou une qualité différente, les normes (ACI 228.1R, 2003; EN 13791, 2007) recommandent de diviser la structure étudiée aux plusieurs régions de tests. À cette fin, des chercheurs (Giannini, et al., 2014; Masi, et al., 2016) ont proposé des critères pour diviser la structure étudiée en zones homogènes (régions de tests). Dans le présent travail, chaque jeu de données qui sera utilisé dans l'analyse est considéré comme appartenant à une région de test, c'est-à-dire que la question de la subdivision est hors de notre portée.

Dans la méthodologie actuelle, un modèle de conversion entre la résistance du béton et les mesures non destructives est identifié. Dans la forme courante de ce modèle, la résistance du béton représente la variable dépendante (réponse) tandis que les propriétés mesurées non destructives représentent les variables indépendantes (entrées). Cependant, une possibilité proposée par certains chercheurs, voir par exemple (Kheder, 1999; Atici, 2011; Szilágyi, et al., 2011; Martínez-Molina, et al., 2014) consiste à utiliser (en plus des mesures non destructives) plusieurs caractéristiques du béton (rapport eau / ciment, rapport agrégat / ciment, teneur en mélange, densité du béton, âge, etc.) comme variables indépendantes dans le modèle. Cependant, le principal inconvénient de ce type de modèles est la nécessité de connaître les caractéristiques du béton pour alimenter le modèle alors qu'ils restent généralement inconnus dans les structures anciennes. Par conséquent, dans cette étude, les modèles avec des mesures non destructives seulement comme des entrées seront considérés.

Objectif et méthodologie générale de recherche

Cette thèse vise à étudier la méthodologie d'évaluation actuelle afin de fournir des recommandations pratiques qui peuvent améliorer la fiabilité de l'évaluation de la résistance des bétons dans les structures existantes par des essais non destructifs et des prélèvements destructifs (carottes).

Pour atteindre cet objectif, la présente thèse suivra la méthodologie de recherche présentée sur la Figure 1. Après avoir synthétisé le problème et les facteurs d'influence, nous définirons les éléments de la stratégie d'évaluation. Ensuite, les outils nécessaires à l'analyse de cette thèse seront fournis. Le premier outil est la construction d'un simulateur utilisant le code VBA. Ce simulateur a deux objectifs: la génération de données synthétiques, et la simulation de la stratégie d'évaluation, qui peut être répétée plusieurs fois. Le deuxième outil concerne les ensembles de données de mesures réelles qui seront utilisés dans l'analyse. De nombreux ensembles de données seront mis à disposition. Ils ont été recueillis à partir de la collaboration avec d'autres recherches et de la littérature scientifique. Les détails seront présentés dans le Chapitre 3.

L'étude de cette thèse peut être classée en trois axes principaux. Le premier axe (Chapitre 4) sera l'analyse de la méthodologie d'évaluation actuelle. Cette analyse se subdivise en trois parties principales : l'analyse des facteurs les plus influents de la méthodologie d'évaluation, l'analyse et la comparaison de plusieurs stratégies d'évaluation à partir d'un benchmark international effectué par la RILEM TC 249 ISC et l'analyse des différentes approches existantes de l'identification du modèle.

Notre travail couvrira également les contributions originales dans les deuxième et troisième axes. Nous allons d'abord développer une nouvelle approche d'identification de modèle pour évaluer à la fois la résistance moyenne et la variabilité de la résistance du béton dans le deuxième axe (Chapitre 5). Ensuite, dans le troisième axe, la qualité de l'évaluation sera analysée en relation avec les principaux facteurs de la méthodologie d'évaluation (Chapitre 6). Enfin, nous proposons quelques recommandations pour améliorer la pratique des ingénieurs (Chapitre 6).

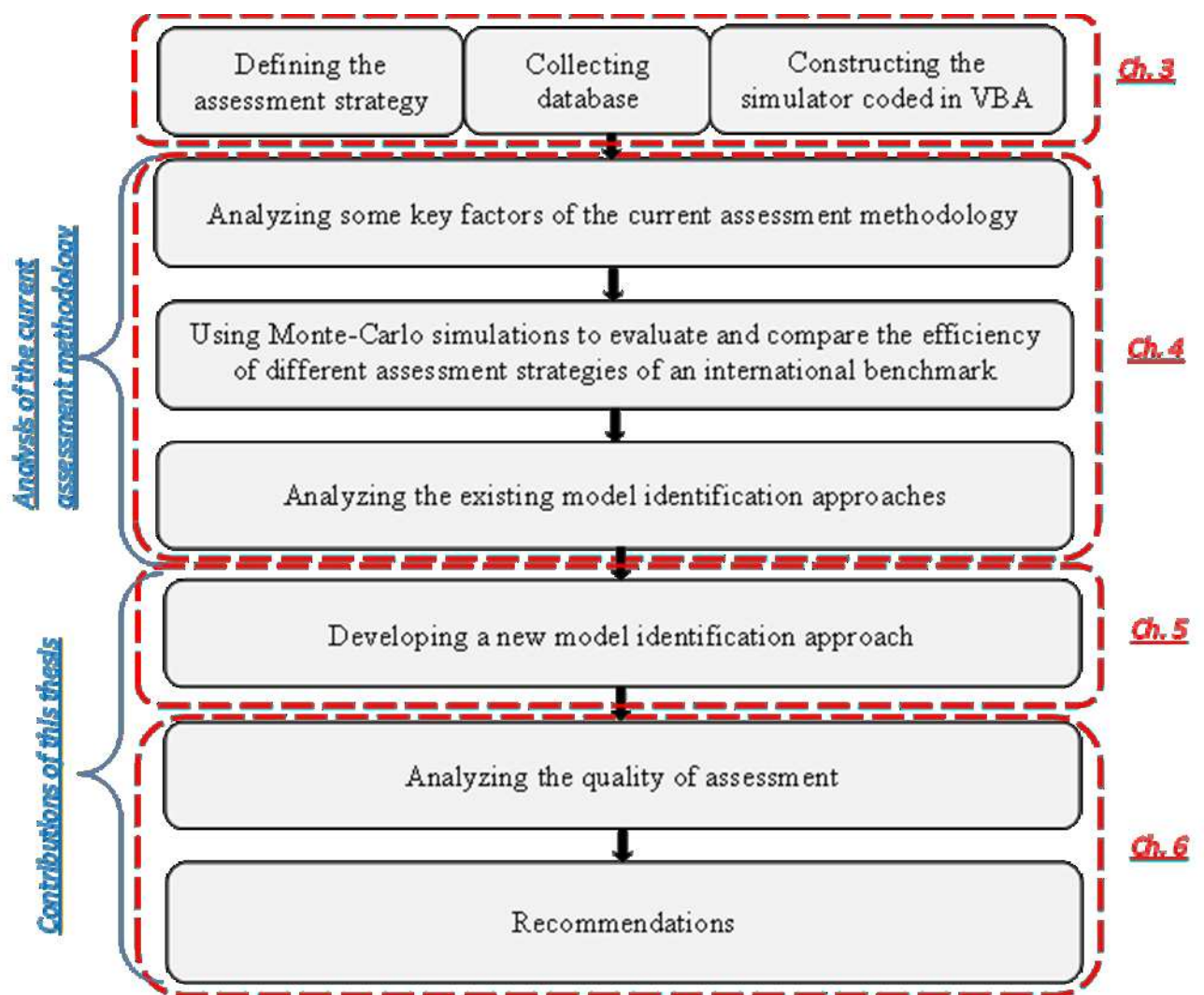


Figure 1. Méthodologie générale de recherche

CONCLUSION GÉNÉRALE

L'évaluation de la résistance à la compression des bétons dans les structures existantes est toujours un défi pour les ingénieurs. La méthodologie actuelle consiste à effectuer des mesures non destructives (ND) comme le rebond ou / et la vitesse d'impulsion ultrasonore sur plusieurs emplacements de tests dans la structure. Ensuite, des prélèvements destructifs (carottes) sont extraits de certains de ces emplacements pour des essais de compression afin de mettre en œuvre un "modèle de conversion" entre la résistance à la compression et les mesures ND. Le modèle de conversion est utilisé pour estimer la valeur de résistance locale à chaque emplacement de test en utilisant la valeur ND correspondante. Ensuite, on calcule les valeurs de résistance moyenne estimée et / ou de l'écart type de la résistance estimée (variabilité de la résistance du béton). Cependant, la fiabilité de ces valeurs estimées est toujours une question discutable en raison des incertitudes associées aux prédictions de résistance basées sur les mesures ND. Pour améliorer la fiabilité, les incertitudes doivent être réduites en spécifiant et en contrôlant leurs facteurs d'influence. Dans cette étude, l'accent a été mis principalement sur les facteurs contrôlés (c'est-à-dire les facteurs qui peuvent être modifiés) afin d'améliorer la qualité de l'évaluation. L'objectif de cette thèse était donc d'étudier la méthodologie d'évaluation actuelle afin de fournir des recommandations pratiques susceptibles d'améliorer la fiabilité de l'évaluation de la résistance in-situ des structures en béton par des essais non destructifs et des prélèvements.

Le premier axe de recherche de cette thèse a été consacré à l'analyse de la méthodologie d'évaluation actuelle. Cette analyse a été subdivisée en trois étapes principales. La première a porté sur l'étude des effets de plusieurs facteurs contrôlés par la méthodologie. Les facteurs considérés sont : le nombre d'emplacements de tests pour les carottes utilisées pour identifier le modèle (NC), la qualité des mesures ND (répétabilité des mesures), la variabilité in-situ de la résistance du béton, la forme du modèle et la combinaison de techniques de CND. Pour l'analyse de ces effets, trois sources de données (données in-situ, données d'études de laboratoire et données synthétiques générées) ont été examinées. Les erreurs $RMSE$ et r^2 ont été calculées pour les résistances locales estimées afin d'être des indicateurs de la qualité du modèle ajusté et de sa capacité à estimer la résistance. Les erreurs ont été calculées de deux manières: en considérant uniquement les emplacements de tests avec des carottes servant à identifier le modèle (erreur d'ajustement) et en considérant toutes les emplacements de tests avec des mesures ND mais sans carottes (erreur de prédiction).

À partir de l'analyse de cette étude, nous avons montré que les erreurs dans les résistances estimées sont affectées par le nombre d'emplacements où l'on prélève des carottes, NC . Pour un petit NC , l'erreur d'ajustement est faible alors que l'erreur de prédiction est importante, mais à mesure que NC augmente, l'erreur de prédiction est progressivement réduite mais elle est toujours plus grande que l'erreur d'ajustement. Cette importante question avertit les ingénieurs d'éviter d'utiliser l'erreur d'ajustement (habituellement r^2) comme indication de la qualité de l'évaluation et ils devraient utiliser l'erreur de prédiction à cette fin. De plus, il a été démontré que la variabilité in-situ de la résistance du béton, $s(f_{ist})$, a également un effet significatif sur la qualité de l'évaluation. Comme $s(f_{ist})$ augmente, la $RMSE$ prédictif augmente tandis que le rapport $RMSE/s(f_{ist})$ diminue. Par conséquent, il faut prêter attention à la façon dont la qualité de l'évaluation est quantifiée: en regardant l'erreur absolue, $RMSE$, ou l'erreur relative, $RMSE/s(f_{ist})$, conduirait à des conclusions différentes. En ce qui concerne la qualité des mesures, il a été constaté que ceci a un effet significatif sur l'erreur de prédiction de sorte qu'il mérite une attention particulière. Par rapport aux trois effets précédents, on a constaté que l'effet de la forme du modèle (linéaire ou de puissance) est moins significatif puisqu'il est généralement limité au cas de NC petite avec une erreur plus

petite fournie par le modèle linéaire. Le dernier effet qui a été analysé à cette étape était l'efficacité de la combinaison de deux techniques de mesures ND (vitesse d'impulsion et indice rebond) avec les tests destructifs (carottes). Il a été démontré que la combinaison n'est pas toujours efficace et que son efficacité dépend du *NC* et de la qualité des mesures. Cela explique les raisons pour lesquelles, dans la littérature, l'efficacité de la combinaison de *CND* est controversée (la combinaison apporte une valeur ajoutée dans certains cas, mais pas dans d'autres).

Dans la pratique courante, l'un des défis pour les ingénieurs est la sélection d'une stratégie d'évaluation parmi les stratégies possibles avec le même budget d'auscultation. Par conséquent, dans la deuxième étape du premier axe, plusieurs stratégies définies par des experts dans un benchmark international consacré à l'évaluation de la résistance moyenne et de la variabilité de la résistance du béton ont été analysées. L'efficacité de ces stratégies a été étudiée en simulant leur application à une série de 1000 bâtiments synthétiques ayant les mêmes propriétés statistiques et en définissant les indicateurs de la qualité de l'évaluation.

Il a été démontré que les dispersions dans les valeurs estimées de la résistance moyenne et de la variabilité de la résistance confirment qu'il est impossible d'évaluer une stratégie à partir d'un seul résultat tel que fourni par les contributeurs à ce benchmark international. Ces dispersions soulignent également que la variabilité de la résistance moyenne estimée et la variabilité de la variabilité de la résistance estimée sont importantes et ne peuvent être négligées quelle que soit la stratégie d'évaluation. L'autre conclusion de cette analyse était que l'évaluation de la variabilité de la résistance est un défi beaucoup plus difficile que l'évaluation de la résistance moyenne. Les stratégies analysées ici ont montré un handicap pour évaluer la variabilité de la résistance du béton. Cette analyse a également confirmé l'effet significatif de la qualité des mesures sur la qualité de l'évaluation. Après avoir étudié l'effet du mode de sélection des emplacements pour les carottes, une méthode a été proposée pour sélectionner ces emplacements en fonction des mesures *CND* "sélection conditionnelle" qui améliore la qualité de l'évaluation sans coût supplémentaire, il est donc fortement recommandé.

La troisième étape du premier axe a été consacrée à l'analyse des approches existantes d'identification de modèles (approche de régression et approches de calibration (méthode- Δ et méthode- k)). Ces approches ont été appliquées à huit ensembles de données réelles (4 avec des mesures de rebond et 4 avec des mesures de vitesse ultrasonore) afin d'évaluer la résistance moyenne et la variabilité de résistance. En comparant la qualité de l'évaluation de ces approches on a constaté que toutes les approches (régression et calibration) peuvent estimer efficacement la résistance moyenne du béton. En outre, l'augmentation de *NC* peut améliorer significativement l'efficacité d'estimation. Au contraire, il a été démontré que l'approche de régression reste limitée pour évaluer la variabilité de la résistance du béton et elle sous-estime toujours la variabilité du béton et l'augmentation de *NC* n'a pas d'effet significatif sur la qualité de l'estimation. En ce qui concerne les approches de calibration, on a constaté qu'elles produisaient des résultats instables (la qualité de l'évaluation varie en fonction du modèle a priori), de sorte que leur efficacité dépend du hasard. De plus, l'augmentation de *NC* ne pouvait pas améliorer les résultats des approches de calibration en particulier pour la méthode Δ qui a prouvé son indépendance par rapport à l'effet de *NC*. Ces résultats ont été obtenus à partir de deux types de techniques de mesures ND (rebond ou vitesse ultrasonore).

Dans le premier axe, le handicap des approches existantes pour évaluer la variabilité de la résistance a été confirmé. Par conséquent, le deuxième axe de cette thèse visait à proposer une nouvelle approche d'identification de modèle "approche bi-objectif" afin d'évaluer la variabilité de la résistance du béton en plus de la résistance moyenne. La dérivation de

l'approche bi-objectif a été fournie pour les modèles linéaires et puissances. La validité de cette approche a été prouvée en l'appliquant sur plusieurs ensembles de données (certains jeux de données avec le rebond et d'autres avec la vitesse ultrasonore) à partir d'études in situ et en laboratoire afin de prédire la résistance moyenne et la variabilité du béton. De plus, la capacité de prédiction de l'approche bi-objectif a été comparée à celle des approches existantes.

Il a été constaté que l'approche bi-objectif montre un résultat plus fiable (contrairement à d'autres approches) pour évaluer la variabilité in-situ du béton en plus de son efficacité à évaluer la résistance moyenne avec une meilleure efficacité quand NC augmente. Il a été démontré également que l'approche bi-objectif est valide avec les mesures de la vitesse d'impulsion ultrasonore et du rebond. De plus, du point de vue mathématique, elle peut être utilisée avec n'importe quelle technique de CND combinée avec la technique destructive (carottage) afin d'identifier un modèle de conversion. En ce qui concerne la forme du modèle, on a constaté que l'approche bi-objectif fonctionne avec le modèle de puissance ainsi qu'avec le modèle linéaire, mais l'incertitude découlant de l'utilisation du modèle de puissance est plus grande que celle du modèle linéaire lorsque NC est faible.

Dans le troisième axe, la qualité de l'évaluation a été analysée et elle a été corrélée graphiquement avec ses facteurs affectants: nombre d'emplacements de tests pour les carottes utilisées pour identifier le modèle NC , qualité des mesures (répétabilité des mesures), mode de sélection des emplacements de tests pour les carottes, la technique CND , l'utilisation d'une ou de plusieurs techniques de CND , l'approche d'identification du modèle et la quantité à évaluer (résistance moyenne ou variabilité). Une grande campagne d'ensembles de données synthétiques ($36000-(V, R, f_c)$) ayant des caractéristiques différentes a été générée pour cette analyse. Cependant, afin de confirmer la robustesse des résultats de l'analyse avec les ensembles de données synthétiques, l'analyse des facteurs énumérés ci-dessus a été effectuée en utilisant des ensembles de données réelles, mais avec la technique de rebond seulement ($2500-(R, f_c)$).

A partir de ces analyses, les effets significatifs de ces facteurs sur la qualité de l'évaluation ont été confirmés avec le rôle le plus important joué par la qualité des mesures. En outre, il a été démontré que, pour une erreur relative prescrite ou un niveau d'incertitude autour de la valeur réelle ($\pm U\%$), l'évaluation de la résistance moyenne en béton moins variable est plus fiable. Au contraire, l'évaluation de la variabilité de la résistance du béton moins variable est moins fiable. À partir de la corrélation graphique entre la qualité de l'évaluation et les facteurs affectants, nous avons calculé les valeurs minimales de NC . On a trouvé un large éventail de valeurs NC pour chaque quantité estimée. Par conséquent, le NC minimum n'est pas une valeur unique qui peut être appliquée dans toutes les situations comme le prévoient par les normes (ACI 228.1R, 2003; EN 13791, 2007). Par conséquent, nous pensons que cette question doit être sérieusement révisée dans les normes actuelles afin de fournir des valeurs minimales NC (dans les tableaux ou les formules) avec des informations détaillées sur leurs limites d'applications.

Étant donné que l'approche bi-objectif proposée donne des résultats encourageants avec des techniques de mesures ND pour évaluer efficacement la résistance moyenne et la variabilité de résistance, notre première perspective est de développer cette approche pour être applicable pour le cas de combinaison de techniques de mesures ND . Nous pensons que cette évolution permet d'améliorer davantage la qualité de l'évaluation.

La deuxième perspective consiste à étudier l'effet du nombre d'emplacements de tests pour les mesures ND (c'est-à-dire la taille de l'échantillon, NT) sur la qualité de l'évaluation. Pour les aspects statistiques, la répartition de la distribution de la résistance moyenne (par exemple)

dépend du rapport σ/\sqrt{NT} , où σ est l'écart-type de la population de la résistance du béton. Par conséquent, l'augmentation de la taille de l'échantillon conduit à réduire l'une des sources d'incertitude (incertitude d'échantillonnage) et réduit donc la dispersion des valeurs estimées. Ce facteur n'a pas été pris en compte dans la présente étude. Cependant nous pensons qu'il mérite d'être pris en compte dans les travaux futurs.

La qualité de l'évaluation évidemment s'améliore lorsque NC et la qualité des mesures augmentent. Cependant, lorsque le montant total des ressources est fixé (budget d'investigation fixe), l'équilibre optimal entre ces facteurs ne suit pas des règles simples et mérite d'être étudié par une analyse plus complète.

REFERENCES

[A]

- ACI 214.4R, 2003. *Guide for Obtaining Cores and Interpreting Compressive Strength Results*, Detroit: American Concrete Institute.
- ACI 228.1R, 2003. *In-Place Methods to Estimate Concrete Strength*, Detroit: American Concrete Institute.
- Al-Ameeri, A. S., Al-Hussain , K. & Essa, M., 2013. Constructing a mathematical models to predict compressive strength of concrete from non-destructive testing. *International Journal of Civil Engineering and Technology (IJCIET)*, 4(4), pp. 1-20.
- Ali-Benyahia, K., Ghrici, . M. & Kenai , S., 2010. *Corrélations entre essais non destructifs et essais destructifs du béton à faible résistance mécanique*. La Bourboule, Rencontres Universitaires de Génie Civil.
- Ali-Benyahia, K. et al., 2017. Analysis of the single and combined non-destructive test approaches for on-site concrete strength assessment: General statements based on a real case-study. *Case Studies in Construction Materials*, Volume 6, pp. 109-119.
- ASTM C39, 2014. *Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens*, Philadelphia: ASTM International.
- ASTM C42, 2012. *Standard Test Method for Obtaining and Testing Drilled Cores and Sawed Beams of Concrete*, Philadelphia: ASTM International.
- ASTM C597, 2002. *Standard Test Method for Pulse Velocity Through Concrete*, Philadelphia: ASTM International.
- ASTM C805, 1997. *Standard Test Method for Rebound Number of Hardened Concrete*, Philadelphia: ASTM International.
- Atici, U., 2011. Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network. *Expert Systems with Applications*, 38(8), pp. 9609-9618.

[B]

- Balayssac, J. P. et al., 2012. Description of the general outlines of the French project SENSO–Quality assessment and limits of different NDT methods. *Construction and Building Materials*, Volume 35, pp. 131-138.
- Bartlett, F. M. & MacGregor, J. G., 1994. *Assessment of concrete strength in existing structures*, Alberta: Structural Engineering Report No. 198.
- Bell, S., 2001. *A beginner's guide to uncertainty of measurement*, Teddington, UK: National Physical Laboratory.
- Biondi, S. & Candigliota, E., 2008. *In situ tests for seismic assessment of RC structures*. Beijing, 14th World Conference on Earthquake Engineering.
- Bogas, J. A., Gomes, M. G. & Gomes, A., 2013. Compressive strength evaluation of structural lightweight concrete by non-destructive ultrasonic pulse velocity method. *Ultrasonics*, 53(5), pp. 962-972.
- Brencich, A., Cassini, G., Pera, D. & Riotto, G., 2013. Calibration and Reliability of the Rebound (Schmidt) Hammer Test. *Civil Engineering and Architecture*, 1(3), pp. 66-78.

Breysse, D., 2012. Nondestructive evaluation of concrete strength: an historical review and a new perspective by combining NDT methods. *Construction and Building Materials*, 33(21), pp. 139-163.

Breysse, D. et al., 2017. Non destructive assessment of in situ concrete strength: comparison of approaches through an international benchmark. *Materials and Structures*, 50(133).

Breysse, D. & Martínez-Fernández, J. L., 2014. Assessing concrete strength with rebound hammer: review of key issues and ideas for more reliable conclusions. *Materials and Structures*, 47(9), p. 1589–1604.

Brozovsky, J., 2009. *Evaluation of calculation correlation efficiency as mentioned in EN13791 in order to determination concrete compression strength by nondestructive testing*. Ljubljana, Slovenian, 10th International conference of the Slovenian society for NDT, pp. 221-231.

Brozovsky, J., Benes, D. & Zach, J., 2013. NDT of LWC with Expanded Clay. In: O. Büyüköztürk & et al., eds. *Nondestructive Testing of Materials and Structures*. RILEM Bookseries 6 ed. Istanbul: Springer, pp. 335-340.

BS 1881: Part 203, 1986. *Testing Concrete Part 203: Recommendations for measurement of velocity of ultrasonic pulses in concrete*, London: British Standards Institution.

BS 1881: Part 120, 1983. *Testing concrete Part 120: Method for determination of the compressive strength of concrete cores*, London: British Standards Institution.

Bungey, J., 1980. The validity of ultrasonic pulse velocity testing of in-place concrete for strength. *NDT International*, 13(6), pp. 296-300.

Bungey, J. H., Millard, S. G. & Grantham, M. G., 2006. *Testing of Concrete in Structures*. 4th ed. London: Taylor and Francis.

[C]

Carino, N., 1994. Nondestructive testing of concrete: history and challenges. In: P. Mehta, ed. *Concrete technology – past, present and future*. Detroit: American Concrete Institute, SP 144-30, pp. 623-678.

Carino, N., 2008. Nondestructive test methods. In: E. Nawy, ed. *Concrete construction engineering handbook*. 2 ed. Boca Raton: CRC Press, pp. 21/1-74.

Celik, A. O., Kiliac, K., Tuncan, M. & Tuncan, A., 2012. Distributions of compressive strength obtained from various diameter cores. *ACI Materials Journal*, 109(6), pp. 597-606.

Cianfrone, F. & Facaoaru, I., 1979. Study on the introduction into Italy on the combined non-destructive method, for the determination of in situ concrete strength. *Matériaux et Constructions*, 12(71), pp. 413-424.

Concrete Society Technical Report No.11, 1987. *Concrete core testing for strength*, London: Concrete Society.

[D]

D'Ambrisi, A., Cristofaro, M. T. & De Stefano, M., 2008. *Predictive models for evaluating concrete compressive strength in existing buildings*. Beijing, China, 14th world conf on earthquake engineering.

DeCoursey, W. J., 2003. *Statistics and Probability for Engineering Applications*. 1 ed. Saskatoon: Elsevier Science.

del Rio, L. M. et al., 2004. Characterization and hardening of concrete with ultrasonic testing. *Ultrasonics*, Volume 42, pp. 527-530.

Demirboğa, R., Türkmen, İ. & Karakoc, M. B., 2004. Relationship between ultrasonic velocity and compressive strength for high-volume mineral-admixed concrete. *Cement and Concrete Research*, 34(12), pp. 2329-2336.

Der Kiureghian, A. & Ditlevsen, O., 2009. Aleatory or epistemic? Does it matter?. *Structural Safety*, 31(2), pp. 105-112.

Devore, J. L. & Berk, K. N., 2007. *Modern Mathematical Statistics with Applications*. 1 ed. California: ThomsonBrooks/Cole.

Di Leo, A., Pascale, G. & Viola, E., 1984. Core sampling size in non destructive testing of concrete structures. *ACI SP "Special publication"*, Volume 82, pp. 459-478.

[E]

El Mir, A. & Nehme, S. G., 2016. A comparative study on ultrasonic pulse velocity for normally vibrated and self-compacting concretes. *Concrete Structures*, Volume 17, pp. 8-12.

EN 12390-1, 2000. *Testing hardened concrete - Part 1: Shape, dimensions and other requirements for specimens and moulds*, Brussels: CEN.

EN 12390-3, 2009. *Testing hardened concrete - Part 3: Compressive strength of test specimens*, Brussels: CEN.

EN 12504-1, 2000. *Testing concrete in structures - Part 1: Cored specimens- Taking, examining and testing in compression*, Brussels: CEN.

EN 12504-2, 2012. *Testing concrete in structures - Part 2: Non-destructive testing - Determination of rebound number*, Brussels: CEN.

EN 12504-4, 2004. *Testing concrete - Part 4: Determination of ultrasonic pulse velocity*, Brussels: CEN.

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

Ergün, A. & Kürklü, G., 2012. assessing the relationship between the compressive strength of concrete cores and molded specimens. *Gazi University Journal of Science*, 25(3), pp. 737-750.

EUROLAB Technical Report 1/2006, 2006. *Guide to the evaluation of measurement uncertainty for quantitative test results*, Paris: European Federation of National Associations of Measurement, Testing and Analytical Laboratories.

[F]

FHWA, 1997. *Guide to non destructive testing of concrete*, Washington DC: US Department of Transportation, Federal Highway Administration.

Fiore, A., Porco, F., Uva, G. & Mezzina, M., 2013. On the dispersion of data collected by in situ diagnostic of the existing concrete. *Construction and Building Materials*, Volume 47, pp. 208-217.

Fischli, M. & Moczko, A., 2012. Rebound hammer. In: D. Breyse, ed. *Non-Destructive Assessment of Concrete Structures: Reliability and Limits of Single and Combined Techniques, State-of-the-Art Report of the RILEM Technical Committee TC 207-INR*. Dordrecht: Springer, pp. 101-110.

[G]

Galan, A., 1984. Détermination des caractéristiques de résistance du béton d'après les mesures de propagation d'ultra-sons combinées avec les données sclérométriques. *Matériaux et Construction*, 17(3), pp. 201-206.

Garnier, V., 2012. Ultrasound through transmission. In: D. Breyse, ed. *Non-Destructive Assessment of Concrete Structures: Reliability and Limits of Single and Combined Techniques, State-of-the-Art Report of the RILEM Technical Committee TC 207-INR*. Dordrecht: Springer, pp. 17-27.

Giannini, R., Sguerri, L., Paolacci, F. & Alessandri, S., 2014. Assessment of concrete strength combining direct and NDT measures via Bayesian inference. *Engineering Structures*, Volume 64, pp. 68-77.

Grantham, M., 2012. *Using combined NDT methods and the Sonreb method in concrete dispute resolution*. Edinburgh, Structural Faults and repairs.

Gunes, O., 2015. Turkey's grand challenge: Disaster-proof building inventory within 20 years. *Case Studies in Construction Materials*, Volume 2, pp. 18-34.

[H]

Hajjeh, H. R., 2012. Correlation between destructive and non-destructive strengths of concrete cubes using regression analysis. *Contemporary engineering sciences*, 5(10), pp. 493-509.

Hannachi, S. & Guetteche, M. N., 2012. Application of the combined method for evaluating the compressive strength of concrete on site. *Open Journal of civil Engineering*, Volume 2, pp. 16-21.

Hobbs, B. & Kebir, M. T., 2007. Non-destructive testing techniques for the forensic engineering investigation of reinforced concrete buildings. *Forensic science international*, 167(2), pp. 167-172.

Huang, Q., Gardoni, P. & Hurlebaus, S., 2011. Predicting Concrete Compressive Strength Using Ultrasonic Pulse Velocity and Rebound Number. *ACI Materials Journal*, 108(4), pp. 403-412.

Hyndman, R. J. & Koehler, A. B., 2006. Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4), pp. 679-688.

[I]

IAEA, 2002. *Guidebook on non-destructive testing of concrete structures*, Vienna: International Atomic Energy Agency..

Idrissou, M. M., 2006. *Reliability in interpreting non-destructive testing (NDT) results of concrete structures*, PhD thesis, University Technology Malaysia.

[J]

Jain, A. et al., 2013. Combined use of non-destructive tests for assessment of strength of concrete in structure. *Procedia Engineering*, Volume 54, pp. 241-251.

JCGM 100, 2008. *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*, Geneva, Switzerland: JCGM 2008.

Jones, R. & Façoaru , I., 1969. Recommendations for testing concrete by the ultrasonic pulse method. *Materiaux et Constructions*, 2(4), pp. 275-284.

[K]

Kaplan, M. F., 1959. The effects of age and water/cement ratio upon the relation between ultrasonic pulse velocity and compressive strength of concrete. *Magazine of Concrete Research*, 11(32), pp. 85-92.

Kheder, G., 1999. A two stage procedure for assessment of in situ concrete strength using combined non-destructive testing. *Materials and Structures*, Volume 32, pp. 410-417.

Knaze, P. & Beno, P., 1984. The use of combined non-destructive testing methods to determine the compressive strength of concrete. *Matériaux et Constructions*, 17(3), pp. 207-210.

Komlos, K. et al., 1996. Ultrasonic Pulse Velocity Test of Concrete Properties as Specified in Various Standards. *Cement and Concrete Composites*, Volume 18, pp. 357-364.

[L]

Lencis , U., Udris , A. & Korjakins , A., 2013. Moisture effect on the ultrasonic pulse velocity in concrete cured under normal conditions and at elevated temperature. *Construction Science Journal*, Volume 14, pp. 71-78.

Leshchinsky, A. M., 1991. Combined methods of determining control measures of concrete quality. *Materials and structures*, 24(3), pp. 177-184.

Lin, Y., Kuo, S. F., Hsiao, C. & Lai, C. P., 2007. Investigation of pulse velocity-strength relationship of hardened concrete. *ACI materials journal*, 104(4), pp. 344-350.

Liu, J. C., Sue, M. L. & Kou, C. H., 2009. Estimating the Strength of Concrete Using Surface Rebound Value and Design Parameters of Concrete Material. *Tamkang Journal of Science and Engineering*, 12(1), pp. 1-7.

Luprano, V. et al., 2008. *Nondestructive test methods for evaluation of concrete: the case study of Punta Perotti (Italy)*. Istanbul, 11 DBMC International Conference on Durability of Building Materials and Components.

[M]

Machado , M. D., Shehata , L. C. & Shehata, I. A., 2009. Correlation curves to characterize concretes used in Rio de Janeiro by means of non destructive tests. *Revista IBRACON de Estruturas e Materiais*, 2(2), pp. 100-123.

Mahmoudipour, M., 2009. *Statistical case study on schmidt hammer, Ultrasonic and Core Compression strength tests Results Performed on Cores obtained from Behbahan Cement Factory in Iran*. Prague, Czech Republic, 5th International Workshop of NDT Experts.

Mahure, N. V. et al., 2011. Correlation between pulse velocity and compressive strength of concrete. *International Journal of Earth Sciences and Engineering*, 4(6), pp. 871-874.

Malhotra, V., 2004. Surface hardness methods. In: V. Malhotra & . N. Carino, eds. *Handbook on non-destructive testing of concrete*. 2 ed. Boca Raton: CRC Press, pp. 1/1-15.

Martínez-Molina , W. et al., 2014. Predicting Concrete Compressive Strength and Modulus of Rupture Using Different NDT Techniques. *Advances in Materials Science and Engineering*, Volume 2014, pp. 1-15.

- Masi, A. & Chiauzzi, L., 2013. An experimental study on the within-member variability of in situ concrete strength in RC building structures. *Construction and Building Materials*, Volume 47, pp. 951-961.
- Masi, A., Chiauzzi, L. & Manfredi, V., 2016. Criteria for identifying concrete homogeneous areas for the estimation of in-situ strength in RC buildings. *Construction and Building Materials*, Volume 121, pp. 576-587.
- Masi, A., Digrisolo, A. & Santarsiero, G., 2013. Experimental Evaluation of Drilling Damage on the Strength of cores Extracted from RC Buildings. *World Academy of Science, Engineering and Technology*, Volume 79, pp. 697-703.
- Masi, A. & Vona, M., 2008. *La stima della resistenza del calcestruzzo in-situ: impostazione delle indagini ed elaborazione dei risultati*. Potenza, Progettazione sismica, No. 1/2009, IUSS Press.
- Masi, A. & Vona, M., 2009. *Estimation of the in-situ concrete strength: provisions of the European and Italian seismic codes and possible improvements*. Naples, Proceedings of the Eurocode 8 Perspectives from the Italian Standpoint Workshop, pp. 67-77.
- Meynink, P. & Samarin, A., 1979. *Assessment of compressive strength of concrete by cylinders, cores, and non destructive tests*. Stockholm, Swedish Concrete Research Institute.
- Ministry of railways-India, 2009. *Guidelines on non-destructive testing of bridges BS – 103*, Lucknow: B&S Directorate.
- Mohammed, B., Azmi, N. & Abdullahi, M., 2011. Evaluation of rubbercrete based on ultrasonic pulse velocity and rebound hammer tests. *Construction and Building Materials*, 25(3), pp. 1388-1397.
- Monteiro, A. & Gonçalves, A., 2009. *Assessment of characteristic compressive strength in structures by the rebound hammer test according to EN 13791: 2007*. Nantes, France, NDTCE 09.
- Montgomery, D. & Runger, G., 2003. *Applied statistics and probability for engineers*. USA: John Wiley & Sons Inc..
- Musmar, M. A. & Abedalhadi, N., 2008. Relationship between ultrasonic pulse velocity and standard concrete cube crushing strength. *Journal of Engineering Sciences, Assiut University*, 36(1), pp. 51-59.

[N]

- Naik, T., Malhotra, V. & Popovics, J., 2004. The Ultrasonic Pulse Velocity Method. In: V. Malhotra & N. Carino, eds. *Handbook on non-destructive testing of concrete*. Boca Raton: CRC Press, pp. 8/1-18.
- Nash't, I., A'bour Saeed, H. & Sadoon, A., 2005. *Finding an unified relationship between crushing strength of concrete and non-destructive tests*. Manama, Middle East Nondestructive Testing Conference & Exhibition.
- Neville, A. M. & Brooks, J. J., 2010. *Properties of Concrete*. 2nd éd. London: Prentice Hall.
- Nikhil, M. et al., 2015. The use of combined non destructive testing in the concrete strength assessment from laboratory specimens and existing buildings. *International Journal of Current Engineering and Scientific Research*, 2(5), pp. 55-59.
- Nobile, L., 2015. Prediction of concrete compressive strength by combined non-destructive methods. *Meccanica*, 50(2), pp. 411-417.

Nobile, L. & Bonagura, M., 2013. *Accuracy of non-destructive evaluation of concrete compression strength*. Portoroz, Slovenian, 12th International Conference of the Slovenian Society for Non-Destructive Testing.

[O]

Oixian, L. & Bungey, J. H., 1996. Using compression wave ultrasonic transducers to measure the velocity of surface waves and hence determine dynamic modulus of elasticity for concrete. *Construction and Building Materials*, 10(4), pp. 237-242.

Oktar, O. N., Moral, H. & Taşdemir, M. A., 1996. Sensitivity of concrete properties to the pore structure of hardened cement paste. *Cement and concrete research*, 26(11), pp. 1619-1627.

[P]

Pereira, N. & Romao, X., 2016. Assessment of the concrete strength in existing buildings using finite population approach. *Construction and Building Materials*, Volume 110, pp. 106-116.

Pfister, V., Tundo, A. & Luprano, V. A., 2014. Evaluation of concrete strength by means of ultrasonic waves: A method for the selection of coring position. *Construction and Building Materials*, Volume 61, pp. 278-284.

Popovics, S., 2001. Analysis of the concrete strength versus ultrasonic pulse velocity relationship. *Materials Evaluation*, 59(2), pp. 123-130.

Popovics, S., Rose, J. L. & Popovics, J. S., 1990. The behavior of ultrasonic pulses in concrete. *Cement and Concrete Research*, 20(2), pp. 259-270.

Postacioglu, B., 1985. Nouvelles significations de l'indice sclérométrique Schmidt et de la vitesse de propagation des ultra-sons. *Materials and Structures*, 18(6), pp. 447-451.

Proceq SA, 2016. *Operating Instructions - Concrete Test Hammer N/L NR/NL Manual*, Schwerzenbach: Proceq.

Proverbio, E. & Venturi, V., 2005. *Reliability of nondestructive tests for on site concrete strength*. Lyon, s.n.

Pucinotti, R., 2013. Assessment of in situ characteristic concrete strength. *Construction and Building Materials*, Volume 44, pp. 63-73.

Pucinotti, R., 2015. Reinforced concrete structure: nondestructive in situ strength assessment of concrete. *Construction and Building Materials*, Volume 75, p. 331-341.

[Q]

Qasrawi, H. Y., 2000. Concrete strength by combined nondestructive methods simply and reliably predicted. *Cement and Concrete Research*, Volume 30, pp. 739-746.

[R]

RILEM TC 43-CND, 1983. Recommendations for testing concrete by hardness methods. *Matériaux et Constructions*, 16(95), pp. 368-376.

RILEM TC 43-CND, 1993. Draft recommendation for in situ concrete strength determination by combined non-destructive methods. *Materials and Structures*, 26(155), pp. 43-49.

RILEM TC 7- NDT, 1977. Recommendations for testing concrete by hardness methods. *Matériaux et Constructions*, 10(59), pp. 313-316.

- Ris, R. C., Holthuijsen, L. H. & Booij, N., 1999. A third-generation wave model for coastal regions: 2. Verification. *Journal of Geophysical Research: Oceans*, 104(C4), pp. 7667-7681.
- Rojas-Henao, L., Fernández-Gómez, J. & López-Agüí, J. C., 2012. Rebound Hammer, Pulse Velocity, and Core Tests in Self-Consolidating Concrete. *ACI Materials Journal*, 109(2), pp. 235-243.
- Ross, S. M., 2009. *Introduction to probability and statistics for engineers and scientists*. 4 ed. California: Elsevier.

[S]

- Samarin, A., 2004. Combined Methods. In: V. M. Malhotra & N. J. Carino, eds. *Handbook on non-destructive testing of concrete*. Boca Raton: CRC Press, pp. 9/1-12.
- Sbartaï , Z.-M., Breysse , D., Larget , M. & Balayssac, J.-P., 2012. Combining NDT techniques for improved evaluation of concrete properties. *Cement & Concrete Composites*, Volume 34, p. 725–733.
- Sbartaï, Z., Nguyen, N. T., Breysse, D. & Bos, F., 2012. *méthodologie de combinaison des technique de CND pour meilleure évaluation des béton*. Chambéry, Rencontres Universitaires de Génie Civil.
- Schickert, G., 1984. Critical reflections on non-destructive testing of concrete. *Matériaux et Construction*, 17(3), pp. 217-223.
- Shariati , M. et al., 2011. Assessing the strength of reinforced concrete structures through Ultrasonic Pulse Velocity and Schmidt Rebound Hammer tests. *Scientific Research and Essays*, 6(1), pp. 213-220.
- Shcherbakov , M. V. et al., 2013. A survey of forecast error measures. *World Applied Sciences Journal*, Volume 24, pp. 171-176.
- Soshiroda, T., Voraputhaporn, K. & Nozaki, Y., 2006. Early-stage inspection of concrete quality in structures by combined nondestructive method. *Materials and Structures*, 39(2), pp. 149-160.
- Soutsos, M. N. et al., 2012. Estimation of on-site compressive strength of concrete. In: D. Breysse, ed. *Non-Destructive Assessment of Concrete Structures: Reliability and Limits of Single and Combined Techniques, State-of-the-Art Report of the RILEM Technical Committee TC 207-INR*. Dordrecht: Springer, pp. 119-186.
- Sravindrajah, R., Loo, Y. H. & Tam, C. T., 1988. Strength evaluation of recycled-aggregate concrete by in-situ tests. *Materials and Structures*, 21(4), pp. 289-295.
- Szilágyi , K., 2013. *Rebound surface hardness and related properties of concrete*., PhD thesis, Budapest University of Technology and Economics.
- Szilágyi , K. & Borosnyói , A., 2009. 50 years of experience with the Schmidt rebound hammer. *Concrete Structures*, Volume 10, pp. 46-56.
- Szilágyi , K., Borosnyói , A. & Zsigovics , I., 2014. Extensive statistical analysis of the variability of concrete rebound hardness based on a large database of 60 years experience. *Construction and Building Materials*, Volume 53, pp. 333-347.
- Szilágyi, K., Borosnyói, A. & Zsigovics, I., 2011. Rebound surface hardness of concrete: Introduction of an empirical constitutive model. *Construction and Building Materials*, 25(5), pp. 2480-2487.

[T]

True, G., 2003. Core sampling and testing. Dans: J. Newman & B. S. Choo, éd. *Advanced Concrete Technology: Testing and Quality*. 1st éd. London: Elsevier.

[U]

Uva, G., Porco, F., Fiore, A. & Mezzina, M., 2013. Proposal of a methodology for assessing the reliability of in-situ concrete tests and improving the estimate of the compressive strength. *Construction and Building Materials*, Volume 38, pp. 72-83.

[W]

Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K., 2012. *Probability & Statistics for Engineers & Scientists*. 9 ed. Boston: Prentice Hall.

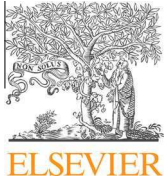
When, R., 1974. Non-destructive testing of concrete. *Building Science*, 9(3), pp. 157-166.

APPENDIX (A)

This appendix includes our published paper:

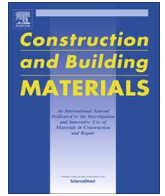
Alwash, M., Breysse, D., Sbartai, ZM., 2015. Non-destructive strength evaluation of concrete: Analysis of some key factors using synthetic simulations. *Construction and Building Materials*, 99, pp. 235-245.

[<http://dx.doi.org/10.1016/j.conbuildmat.2015.09.023>]



Contents lists available at ScienceDirect

Construction and Building Materials

journal homepage: www.elsevier.com/locate/conbuildmat

Non-destructive strength evaluation of concrete: Analysis of some key factors using synthetic simulations

Maitham Alwash^{a,b,*}, Denys Breysse^a, Zoubir Mehdi Sbartaï^a^a I2M-GCE, University of Bordeaux, UMR 5295, 33400 Talence, France^b Civil Eng. Dept., College of Engineering, University of Babylon, Babel, Iraq

HIGHLIGHTS

- Fitting error cannot be a real indication about the quality of concrete strength assessment by NDT.
- A misleading model is produced if we fit using cores number equal to the number of model parameters.
- To assess a fitted model, looking at: r^2 , $RMSE$ or $RMSE/s(f_c)$ leads to different conclusions.
- To fit a SonReb model there is a minimal number of cores above which the combination is efficient.

ARTICLE INFO

Article history:

Received 8 October 2014

Received in revised form 4 September 2015

Accepted 19 September 2015

Keywords:

Concrete strength
 Ultrasonic pulse velocity
 Rebound hammer
 Combination
 NDT techniques
 Synthetic simulation

ABSTRACT

Non-destructive techniques (NDT) like rebound hammer (RH) and ultrasonic pulse velocity (UPV) are widely used in conjunction with destructive techniques (core tests) for assessing the concrete strength in existing buildings. The methodology consists in fitting regression models between NDT techniques and destructive tests on a limited number of cores. The quality of the model is affected by many influencing factors such as: the number of cores, the quality of NDT measurements, the variability in concrete strength, the existence and magnitude of possible uncontrolled factors (like saturation rate) and the combination of techniques. In this paper, the effects of these factors are studied using a synthetic simulation approach in order to well understand them and consequently to develop a methodology for improving the quality of strength assessment. In order to assess the quality of fitted model and its ability to estimate strength, $RMSE$ and r^2 errors are calculated and it is found that the calculation of r^2 alone may give misleading indication since r^2 is very sensitive to the scattering of the explanatory variable. Another important result of the present study is that there is a critical minimal number of cores which makes the combination efficient while for a lower number the use of single technique is preferable. This number depends on the qualities of the two techniques to be combined.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

In the real practice, the structural engineer always needs to carry out tests in existing structures in order to make the right decision about the condition of the structure. The testing of existing structures is usually related to an assessment of structural integrity. When the assessment is based only on destructive testing (DT) by extracting cores for compression testing, the cost of coring and testing may allow only a relatively small number of tests to be carried out on a large structure, which may be misleading [1]. Thus non-destructive techniques (NDT) are used for the assessment of

concrete strength in existing building in conjunction with destructive tests. Many guidelines and specifications are available [1–4], which indicate the increasing use of NDT in real practice. The strength estimation requires a model which can be identified by establishing a statistical correlation between DT and NDT results. Many works have been published in which each author has identified his/her specific model [5–11]. State of the art papers [12,13] have recently identified a huge number of models that have been proposed by different researchers.

The ultrasonic pulse velocity and rebound hammer methods are frequently combined for a better estimation of concrete strength. This is very convenient since these two techniques are sensitive to the variations in some concrete properties in opposite directions. For instance, the increase in moisture content of concrete raises the pulse velocity but lowers the rebound number [14].

* Corresponding author at: I2M-GCE, University of Bordeaux, UMR 5295, 33400 Talence, France.

E-mail address: maitham-fadhil-abbas.alwash@u-bordeaux.fr (M. Alwash).

The combination of NDT techniques was proposed firstly by RILEM (Technical committees 7 and 43) based on seminal work from Façoaaru [15]. There is a huge number of works that deal with the combination of NDT methods and provide mathematical models [16–20]. These works lead to controversial conclusions about the efficiency of combination. Some works found that “the combined use of ultrasonic pulse velocity and rebound hammer greatly improves the accuracy of the process of assessment of concrete strength especially if information about concrete mix proportions and density are available” [21], while others did not find a significant improve in the concrete assessment by combining methods [22,23].

The effectiveness of combination is studied in the present paper in order to explain the controversial results obtained by different researchers.

It is commonly agreed that none of the produced models is able to predict the concrete strength with enough accuracy for using the assessed value for further structural computations [24]. The reason is, on one hand, the influence of both measurements uncertainties and uncontrolled factors (like concrete humidity, carbonation...), and on the other hand, the lack of certainty upon the “best way” to establish the conservation model (number of cores, calibration method, the use of a single NDT method or of a combination...).

A synthetic simulation approach has been proposed [12] in order to deepen the analysis of this issue since the ability of real data sets (experimental work) for deepening the analysis remains limited. In this paper, the effects of some influencing factors such as: the number of cores which are used for the regression/calibration of the model, the quality of NDT measurements, the variability in the concrete strength, the variability in the concrete degree of saturation and the combination of NDT techniques are studied. The synthetic simulation approach is used in order to well understand these effects and consequently to improve the methodology of model development and, as a result, to improve the quality of concrete strength assessment.

2. Synthetic simulation approach

The basic idea of this approach is to simulate statistically the problem of concrete strength evaluation using NDT techniques within the computer by creating a synthetic world which mimics as closely as possible the real world, in order to make possible an in-depth analysis and a parametric study of influencing factors. While it is of course not possible to reproduce *in silico* the real world, the simulation must point what influencing factors (input data of the synthetic model) are considered, and how they influence (sign, magnitude, possible coupling effects) the physical properties measured with NDT methods (outputs of the synthetic model). In the version used in this paper, the synthetic model has been developed in order to correctly reproduce the relationships between:

- strength and moisture content considered as input data on one hand,
- ultrasonic pulse velocity and rebound value considered as output data on the other hand.

Other possible influencing factors (carbonation, cracking, aggregate type and size...) are not considered. While they may be considered in a next version of the synthetic model (a revised version, work under progress, including the influence of carbonation is presently developed by the authors for an international benchmark prepared for RILEM committee TC 249-ISC), it must be pointed that the present model does not pretend to be “the truth”.

It is however the conviction of the authors that most conclusions drawn in this study have very probably a high level of generalization, that will remain to be confirmed in the future.

The detailed principles of the synthetic approach have already been published [12,24,25], to which the reader is invited to refer. In order to avoid duplication, only main patterns of the numerical process are described here.

The first step is the generation of concrete properties: true concrete strength f_c is generated by assuming a Gaussian distribution $N(f_{cm}, s(f_c))$ while a truncated Gaussian distribution, $N(S_r, s(S_r))$ with $S_r \leq 100\%$, is used to generate the values for the degree of saturation S_r which appears as an uncontrolled factor, see Appendix B for more numerical details.

True values (in the synthetic world) for the velocity V (ultrasonic pulse velocity technique) and the rebound number R (rebound hammer technique) which represent the NDT measurements are produced using relationships established after an in-depth literature review of available experimental results, Eqs. (1) and (2) as proposed by Breyse [24]:

$$V = V_{ref} (f_c / f_{cref})^{1/bf} (S_r / S_{ref})^{1/bs} \quad (1)$$

$$R = R_{ref} (f_c / f_{cref})^{1/cf} (S_r / S_{ref})^{1/cs} \quad (2)$$

where the reference values (*ref*) are arbitrary values introduced in order to normalize the equations, and have no influence on the general behavior. The exponents quantify the relative sensitivity of V and R to the variations in strength and humidity. The reference values are $R_{ref} = 40$, $V_{ref} = 4000$ m/s, $S_{ref} = 85\%$ and $f_{cref} = 40$ MPa. The exponent values have been carefully chosen, in order to accurately describe what is observed in practice. The strength sensitivity exponents bf and cf have respectively been taken equal to 4.90 and 2.10. The humidity sensitivity exponents bs and cs have respectively been taken equal to 7.14 and -3.33 [24]. The bs and cs values respectively correspond to an increase of 6% in V and a decrease of 12% in R as the concrete humidity changes from dry-air specimen (assumed at $S_r = 65\%$) to fully saturated condition ($S_r = 100\%$). These values are in agreement with what we have found in literature. For example for V , [26] states that the pulse velocity in saturated concrete may be up to 5% higher than in dry concrete. [27] finds an increase of 19% in V between totally dry and maximum saturated conditions i.e. 6.65% between $S_r = 65\%$ and saturated conditions. Experimental results from [21] show an increase in V of (200–400) m/s between air dry and wet conditions. Also for R , the results from [21] show a decrease of (3–4) points in R values i.e. less than 10%. [28] states that well-cured, air-dried specimens, when soaked in water and tested in the saturated surface-dried condition, show rebound readings 5 points lower than when tested dry. Czech Standard CSN 731373 after [29] indicates a 20% decrease in R .

The relationships (1) and (2) can be combined together in order to obtain an equation in the form ($f_c = f(V, R)$). One thus gets:

$$f_c = f_{cref} (1/V_{ref})^{-bsk/cs} (1/R_{ref})^k V^{-bsk/cs} R^k \quad (3)$$

$$\text{where : } k = cf * bf * cs / (cs * bf - cf * bs) \quad (4)$$

Eq. (3) provides the true (synthetic) value of strength that could be identified by: (a) measuring rebound and velocity on at least two measurement points without any measurement error, (b) choosing a double-power law conversion model. Errors on strength estimates will result from both measurement errors (the main part) and possible choice of a model having a different mathematical shape.

Eq. (3) is used to produce the iso-strength curves shown in Fig. 1a whose shape can be compared with the curves used in practice by engineers like those provided by RILEM TC-43 [15], see

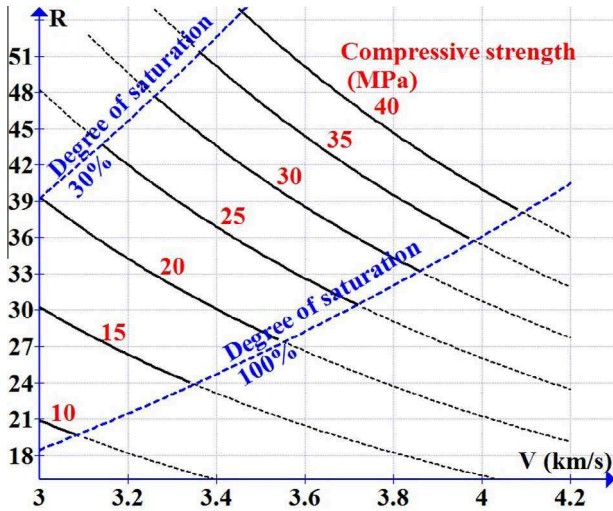


Fig. 1a. Iso-strength curves produced by synthetic simulation, Eq. (3).

Fig. 1b. The fact that synthetic curves do not exactly fit empirical curves is not a problem since several other empirical curves can be used as an alternative [1,5,30,31].

The chart built with the synthetic model, Fig. 1a, exhibits features (sign and magnitude of variations) similar to those exhibited in various empirical charts available in the literature. For instance, the induced changes in R and V values when the saturation degree varies within the limits of its practical range have the same magnitude as those exhibited in practical studies. We must emphasize that this chart does not intend to be the “real” one, but it only shows the ability of synthetic process in producing data that have similar features as the real data.

As it is the case in the real world, measurement errors (ϵ_V and ϵ_R) are added to the generated true values of V and R . The magnitude of these errors are obtained by assuming a Gaussian distribution $N(0, s(V))$ or $(0, s(R))$ with zero average and standard deviation, $s(V)$ or $s(R)$, represents the local variability of the measurements. The possible range of measurement uncertainties for R and V measurements has been widely documented in literature [12,32].

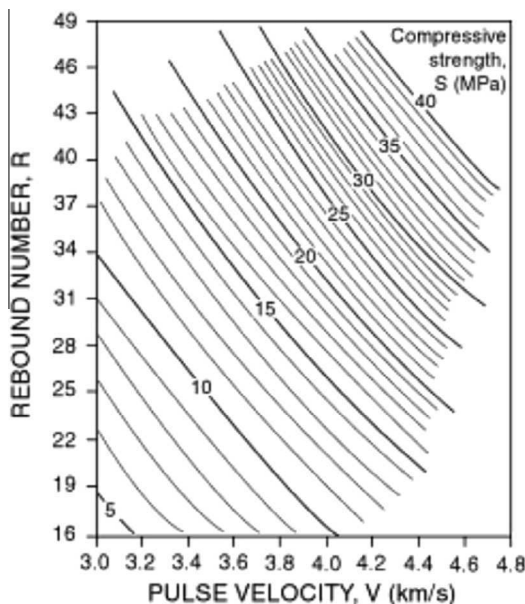


Fig. 1b. Iso-strength curves produced by RILEM TC-43 [15].

Table 1

The values of standard deviation $s(V)$ and $s(R)$ for different quality levels of measurements [12].

Measurements quality level	$s(V)$ m/s	$s(R)$
High	50	1
Average	100	2
Low	200	4

Table 1 gives the values for these standard deviations for different quality levels. The magnitude of the measurement error decreases as the measurement quality increases. Since the “measurement quality” results from a variety of causes (among which the fulfillment to standards or guidelines, the expertise of operators...), the equivalence between a high “measurement quality” and low “measurement variability” can be seen as a shortcut. We will use it and consider that these two concepts have the same meaning for avoiding too long considerations. The main interest of “measurement variability” is that it provides a quantified value, which can easily be quantified in practice by repeating the same measurement in a close neighborhood of a given point.

From the data set of cores strengths (which are selected randomly from the generated strengths values) and the corresponding NDT measurements, a regression analysis can be carried out. A linear model is thus identified by using the least squares method, then it is used to calculate the estimated strengths from NDT values. The quality of assessment is determined by calculating the errors, $RMSE$ (root mean squared error) and r^2 (coefficient of determination), see Appendix A, between true and estimated strengths.

Because of the random character of the random process (random generation of true properties and of errors, and random choice of cores), these results (identified set of model parameters, $RMSE$ and r^2) vary from one simulation to another. To have representative outputs, it is required to repeat the simulation NR times for each data set, and to analyze the distribution of results. Average values and standard deviations of all results (values of model parameters, $RMSE$ and r^2) are calculated. All results discussed in this study have been obtained with $NR = 200$, for which the convergence and stability has been observed.

In order to create the synthetic world and perform the simulation process a computer program has been developed using VBA software. The flowchart in Fig. 2 illustrates how the process works.

The objective is to analyze the effects of several factors which have a significant influence on the quality of strength assessment in the real world practice. Simulations are carried out by varying some parameters and analyzing how $RMSE$ and r^2 on assessed strengths varies. Five varying parameters are considered in this study: (a) the number of cores (NC) for model calibration, (b) the quality of NDT measurements, (c) the range of concrete strength, (d) the variability of concrete saturation rate, and (e) the possibility of combining two NDT methods. Some parameters remain fixed in all simulations:

- The models identified are linear models (respectively $f_c = a_1 + b_1 V$, $f_c = a_2 + b_2 R$, $f_c = a + b V + c R$, depending on the fact that strength is estimated either from one or two NDT measurements),
- The average concrete saturation rate $S_{m} = 80\%$ which corresponds to a rather humid concrete, when unspecified, the standard deviation is $s(S_r) = 5\%$,
- Maximum number of NDT tested points (NT) = 100,
- The simulations are repeated ($NR = 200$) times in order to obtain relevant statistical information. Convergence analysis has empirically proved that this number is much larger than the one necessary to get stable results.

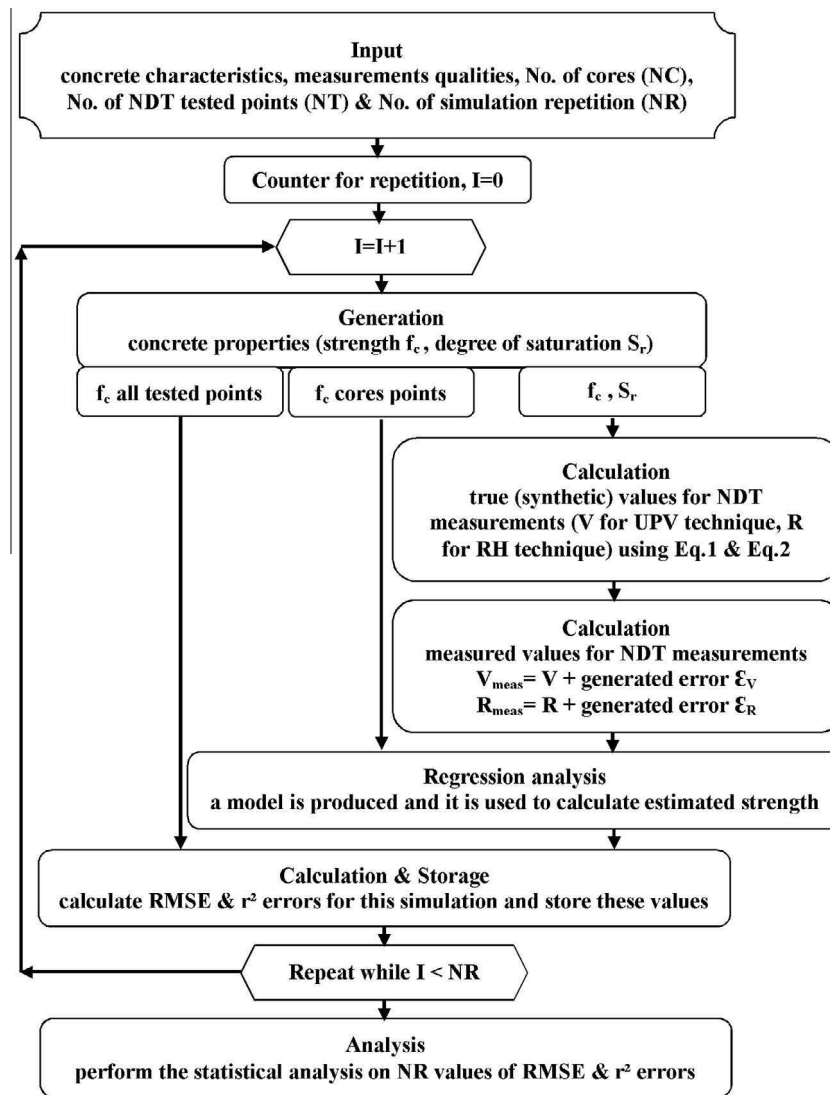


Fig. 2. Flowchart of the synthetic simulation approach.

When unspecified, the concrete with $f_{cm} = 45$ MPa and $s(f_c) = 5$ MPa is supposed. Otherwise, the effect of strength variability is studied by varying the value of $s(f_c)$ from 1 to 10 MPa. The effect of variability in concrete saturation rate is studied by varying the value of $s(S_r)$ from 1% to 10%. One must underline here that the selection of linear model does not mean that this form of model is the best one, although it's widely used in practice. We select this form here because it is a simple way in the synthetic simulations to account for model error, since this model differs from that used for the generation of "true" values, Eqs. (1) and (2).

3. Results and discussion

3.1. The effect of number of cores

In Figs. 3–5, the effect of core number on the *RMSE* values for NDT measurements with different quality levels (high, average and low) is studied. For each quality level, the core number is varied from 3 to 20 and for each core number two values of *RMSE* are calculated: one for the set of *NC* cores on which the model has been fitted, and the other for (*NT-NC*) points at which only NDT is available i.e. where the model is used in a predictive way. Two curves

are drawn, one (on cores) quantifies the fitting error *FEC*, while the other *PEC* (*FEC* and *PEC* are defined in Appendix A) quantifies the effective capacity of the process to estimate strength from NDT only (prediction error), which is the real issue of concern. The difference between the two (predictive and fitting) is the extrapolation error. All *RMSE* plotted on the figures correspond to average values calculated from the series of *NR = 200* repetitions for each data set configuration. From Figs. 3–5 the following notes can be drawn:

- The general features are identical for all strategies (*V* and *R* measurements alone or in combination),
- One must point here that when NDT strength assessment is studied in the scientific literature, authors often limit their analysis to the first (fitting) stage, considering the r^2 and *RMSE* from fitting as a representative of assessment quality. However it is clear that the values of *PEC* are always larger than the corresponding values of *FEC* thus using the fitting error cannot represent the reality.
- *RMSE* values for fitting error increase with the core number because the number of points to be fitted using a model having a fixed number of parameters is increased. However, the prediction error exhibits an adverse pattern, since it decreases while

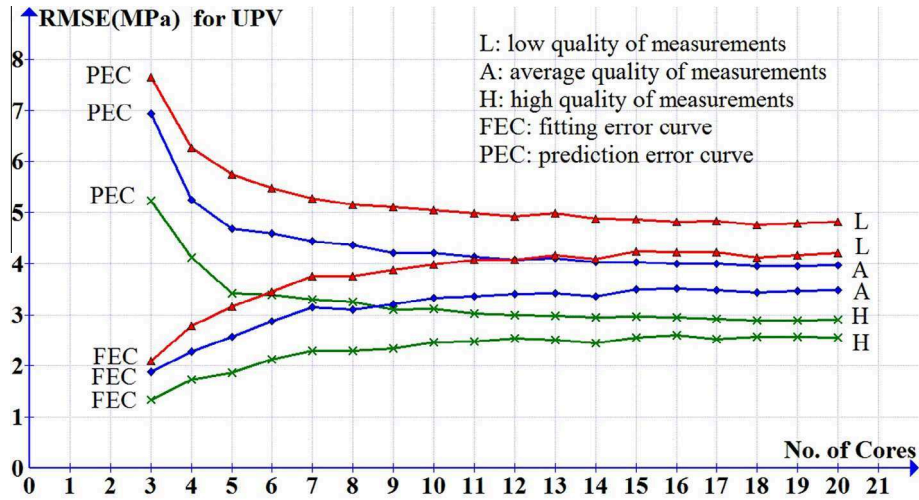


Fig. 3. RMSE versus core number for different qualities of UPV measurements.

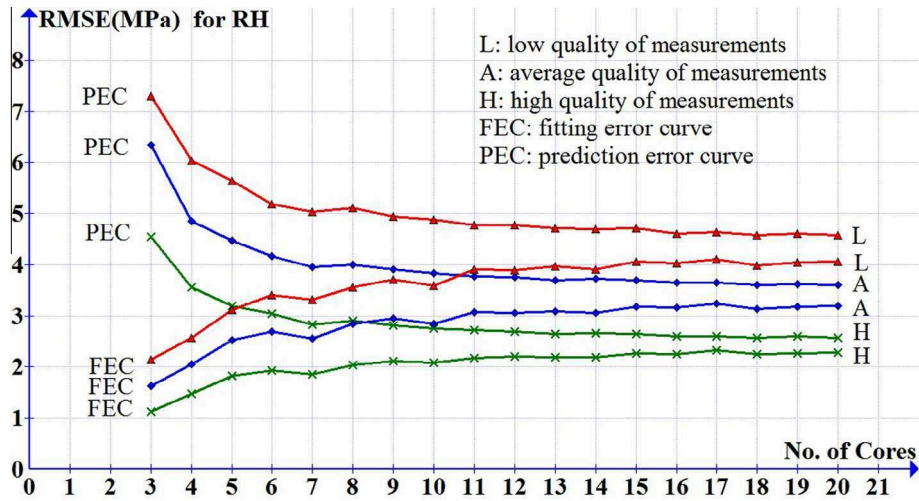


Fig. 4. RMSE versus core number for different qualities of RH measurements.

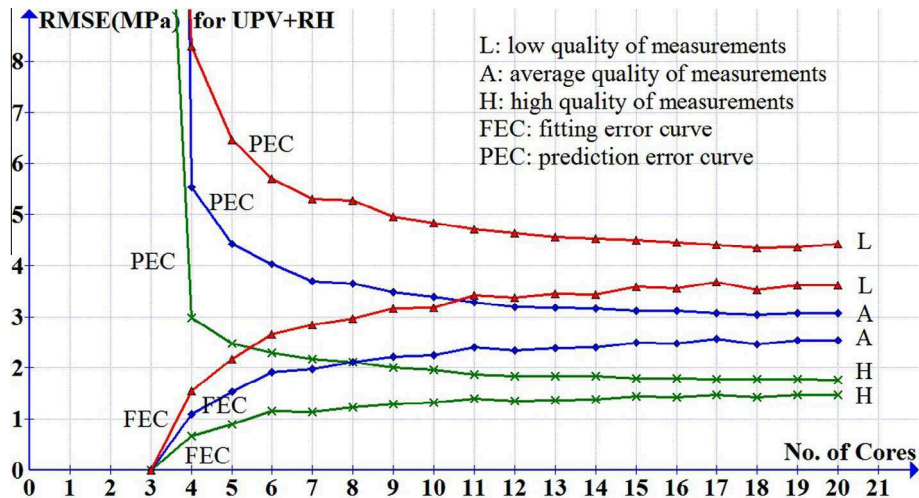


Fig. 5. RMSE versus core number for different qualities of combined (UPV + RH) measurements.

the number of cores increases: by increasing the core number, despite of increasing fitting error, the model provides a better picture of true strengths. By considering only FEC, one would obtain a totally wrong picture. It must be noted that, in practical studies, the assessor balances between contradictory objectives, since he must (a) increase the number of cores for model calibration, in order to reinforce the stability of the conversion model, (b) increase the number of additional cores for checking the model and evaluating PEC, while (c) limiting the total number of cores to an acceptable economical level. One interest of synthetic simulations is that the effect of changing both NC and NT can be analyzed efficiently without any cost.

- For any quality level, the difference between the two curves (extrapolating error) reduces as the number of cores increases. Also, the value of this error depends on the quality of measurement, and decreases when the quality increases.
- Measurement of a lower quality (with more measurement noise) results in a larger predictive $RMSE$.
- On Fig. 5 one can see that, with 3 cores, the $RMSE$ value is zero, which means that one has enough data to identify the three parameters (a, b, c) of the model, without any fitting error. However, the model has a very poor predictive ability and the predictive error is large. This discrepancy must be pointed, since a very low number of cores is the usual practice in current engineering. Scientific literature offers many cases in which the identified model in such a situation is said to be good, this statement being just based on r^2 fitting value, when the predictive ability is in fact very poor.

In order to illustrate and emphasize the practical consequences of this statement, this situation was reproduced on data drawn from a real case-study (original data are taken from measurements collected by Nobile and Bonagura [33]). They consist in a set of 9 combinations of in-situ core strength, velocity and rebound number (Table 2). The standard deviation of strengths is 4.4 MPa. A linear regression model identified on the full data set comes to: $f_{cest} = 0.00988 V_r + 1.0211 R_r - 47.592$, with $r^2 = 0.62$ and $RMSE = 2.5$ MPa. Then it was decided to randomly select 3 subsets of 3 cores, and to identify a model for each subset (this mimics the case where only 3 cores are taken on the structure), while calculating the PEC value on the remaining 6 cores (not used for fitting the model). The data and models summarized in Table 2 respectively lead to PEC values of 9.8, 5.9 and 4.7 MPa, which are varying a lot and are even larger than the original standard deviation. Of course, selecting other subsets would have led to different values, but this suffices to point the risk of relying on a too small number of cores, and on fitting error estimates only (in this case, the models perfectly fit the available data, $r^2 = 1$, but this is meaningless).

3.2. The effect of strength variability

Effect of strength variability on the values of predictive $RMSE$ and r^2 is respectively illustrated on Figs. 6 and 7. On both figures, the error values from single and combined average quality techniques and with $NC = 10$ cores are plotted as the standard deviation of strength is varied from 1 to 10 MPa. These figures show that the increase in the strength variability leads to a simultaneous increase in predictive $RMSE$ and r^2 . From the determination coefficient evolution [34], also see Appendix A, it appears that r^2 is better for a concrete having large strength variability (see Appendix B for the numerical illustration of this observation), while the $RMSE$ increases at the same time. This apparent contradiction can be further analyzed: on one hand, a larger scatter in concrete strength has a “stabilizing effect” on the model regression, thus increasing r^2 , while on the other hand, this does not imply that the strength is accurately assessed, since $RMSE$ still increases with $s(f_c)$.

Table 2

The analysis of real data set provided by [33] to show the risk of using three cores only to fit a SonReb regression model.

Real in-situ measurements			Regression model estimated from a subset of three cores
f_{cr} MPa	V_r m/s	R_r	
17.8	3250	31.9	$f_{cest} = -0.0274V_r + 0.3094R_r + 96.912 \quad r^2 = 1$
27.8	2965	39	
23.25	3140	39.78	
18.7	2930	38.83	$f_{cest} = 0.0026V_r + 0.1731R_r + 4.302 \quad r^2 = 1$
18.9	3285	34.61	
18.5	2960	37.22	
25.6	3500	36.84	$f_{cest} = 0.0193V_r + 0.9301R_r - 76.132 \quad r^2 = 1$
29.3	3470	41.44	
20.6	3120	39.34	

From Fig. 7, r^2 values are negative for small values of variability (less than 3 MPa) because for small values of strength variability the probability for obtaining a wrong model (even with negative slope) is large. Such a wrong model leads to produce SSE value (error sum of squares) larger than SST value (total sum of squares) and, as a result, a negative value of r^2 (see Appendix A for the definition of r^2) [34]. In this case, one could say that (even with number of cores $NC = 10$) we have a risk of obtaining a wrong model. One can also realize that the $RMSE/s(f_c)$ ratio decreases, meaning that the absolute error increases while the relative error decreases. This confirms that one must pay attention to how the quality of the model is quantified: looking at r^2 , absolute error $RMSE$ or relative error $RMSE/s(f_c)$ would lead to different conclusions.

3.3. The effect of variability in concrete saturation rate

Fig. 8 shows the effect of variability in concrete saturation rate on the values of predictive $RMSE$. In this figure, the error values from single and combined techniques with average qualities are plotted as the standard deviation for degree of saturation is varied from 1% to 10%. From these curves the following notes can be drawn:

- Larger values of $RMSE$ values are obtained for single techniques when the saturation variability increases because saturation rate is an uncontrolled factor which induces more noise when it is not explicitly accounted for in the model.
- When the strength variability changes from 1 to 10 MPa the error increases by 400% (Fig. 6), while the increase in the error as the saturation rate variability varies from 1% to 10% is only 14% (Fig. 8), thus the variability in saturation rate has a much smaller effect on the values of $RMSE$ than variability in strength.
- The curve of combined techniques is nearly horizontal which means that the noise induced by this uncontrolled factor (variability of saturation rate) can be eliminated by using a proper combination of NDT techniques.

3.4. The efficiency of combination of techniques

Figs. 9–11 illustrate how using techniques in combination affects the quality of estimation as the number of cores varies from 3 to 20, when UPV is used alone or if rebound (RH) is added in combination. The figures illustrate how the quality of techniques influences the results, and the efficiency of combination. Similar results would have been obtained with RH as the first single technique.

Fig. 9 shows that, considering the first (UPV) technique has a high quality, when a second technique is used in combination, it always leads to a more accurate assessment if the number of cores (NC) is equal or greater than 9. The better the quality of this second technique, the more efficient the combination. However, when NC

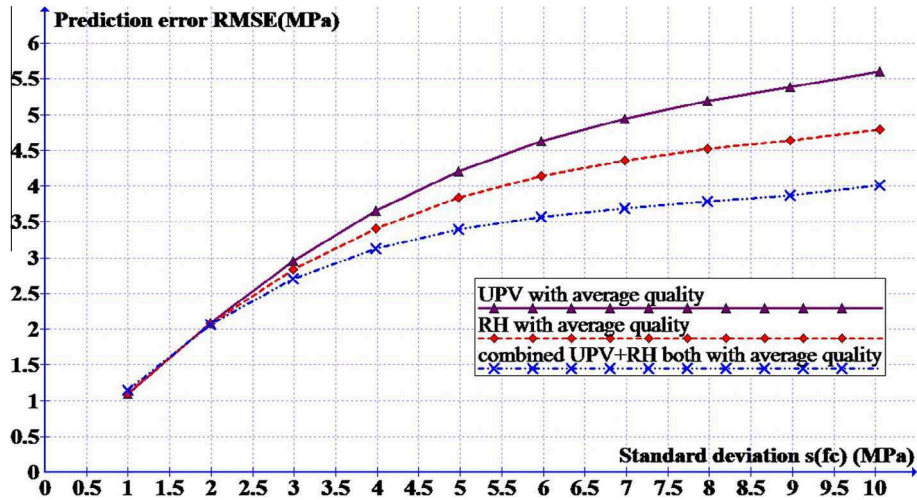


Fig. 6. The effect of strength variability on the values of predictive RMSE for single and combined techniques with average qualities (No. of cores = 10).

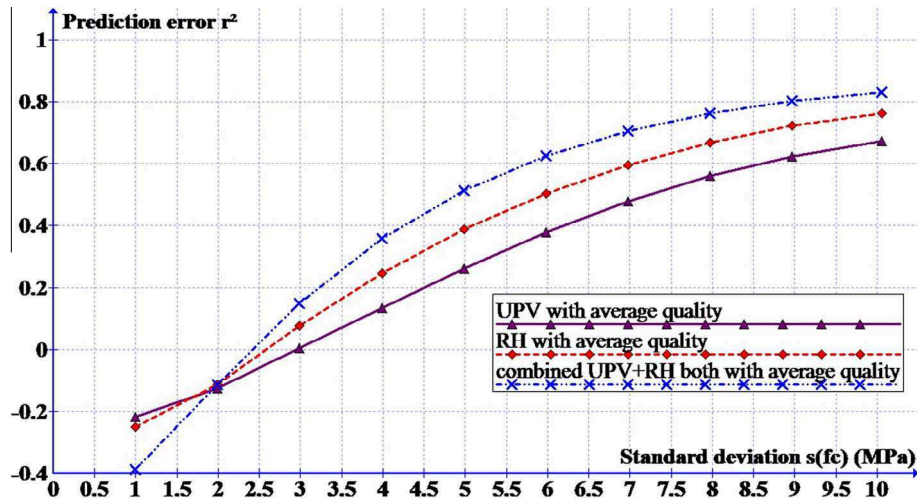


Fig. 7. The effect of strength variability on the values of predictive r^2 for single and combined techniques with average qualities (No. of cores = 10).

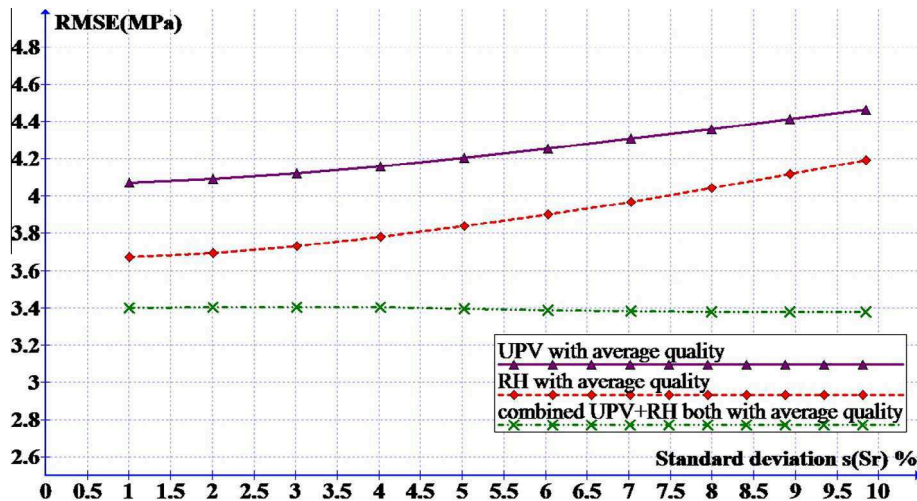


Fig. 8. The effect of variability in concrete saturation rate on the values of RMSE for single and combined techniques with average qualities (No. of cores = 10).

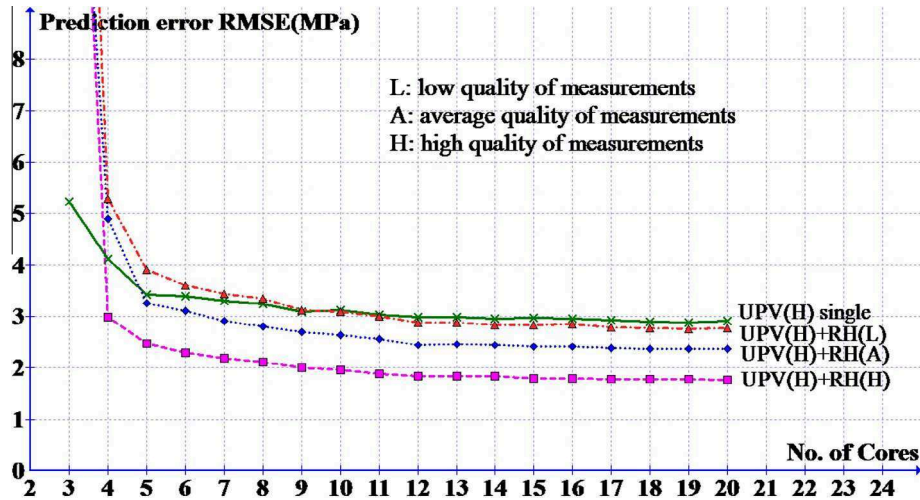


Fig. 9. Efficiency of NDT, alone or in combination, for UPV with high quality and different quality levels of RH measurements, NC = 10 cores.

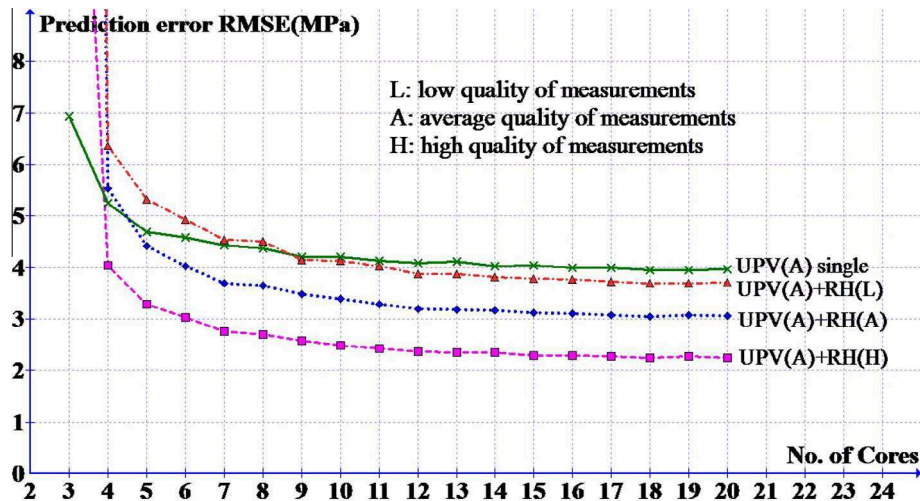


Fig. 10. Efficiency of NDT, alone or in combination, for UPV with average quality and different quality levels of RH measurements, NC = 10 cores.

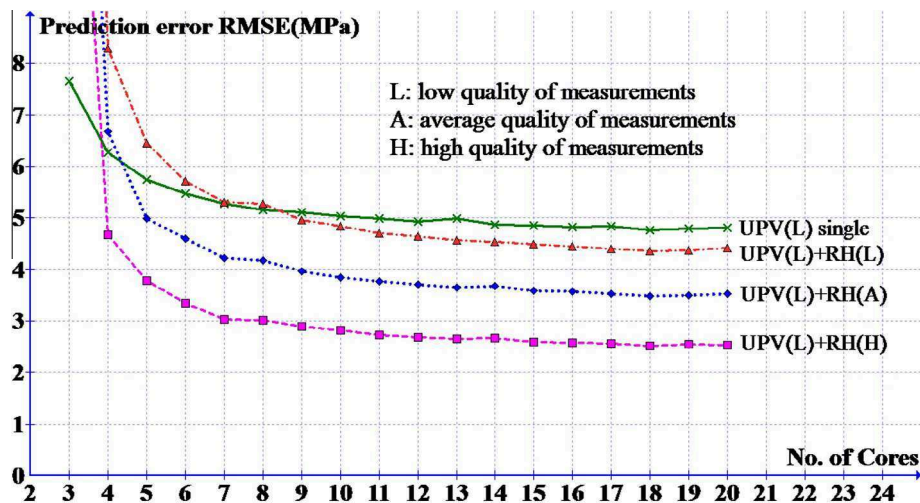


Fig. 11. Efficiency of NDT, alone or in combination, for UPV with low quality and different quality levels of RH measurements, NC = 10 cores.

is between 5 and 9, the combination is efficient only if the quality of a second technique is high or average. For such NC values, combining a (second) low quality technique to the (first) high quality technique would lead to larger $RMSE$ and would be prejudicial, like it was found in some cases in the literature. If one has no more than 4 cores, the only efficient combination is with a second high quality technique. This simple graph explains the reasons why in the literature the efficiency of combining NDT is controversial (combination is said to bring some added-value in some cases but not in others): the reason lies in the relative quality of the two techniques (and the result also depends on the number of cores).

Figs. 10 and 11 show the results when the quality of the first technique (UPV) is respectively average and low. They exhibit a very similar pattern to that shown on Fig. 9.

It can be noted (for $NC \geq 9$) that the minimum value of $RMSE$ is obtained from the combination of two techniques with high qualities while this error reaches its maximum value when both qualities are low. Thus the qualities of measurement widely control the efficiency of combination.

From above, it seems that there is a critical minimal number of cores above which the combination is efficient while for a small number of cores the use of single technique will be better. This number is affected by the qualities of the two techniques to be combined. We think that this important observation is a general property but it needs further investigations to be approved.

4. Conclusions

In order to understand the effects of several factors which affect the assessment of concrete strength in existing structures by NDT measurements a parametric study has been performed using the synthetic simulation approach. The synthetic simulation is a powerful approach making it possible to explain the (sometimes apparently controversial) results that NDT researchers obtain on site or in laboratory when they make the NDT measurements, while providing a more general view on the problem treated. The following conclusions can be drawn:

- Errors in the estimated strength are affected by the number of cores. For small number of cores, the fitting error is small while the prediction error is large, but as the number of cores increases the prediction error is progressively reduced. This error depends on the quality of measurement as, for the same number of cores, it decreases with better quality measurements. For the set of examples treated in this study, it appears that the added-value of increasing the number of cores beyond 7 or 8 is only limited.
- Using three cores to fit a model having three parameters will produce a fitting error equal to zero, which means that one has enough data to identify the three parameters of the model. However, the model has a very poor predictive ability (the prediction error is large). Real practice offers many cases in which

the identified model in such a situation is considered to be good, this statement being just based on r^2 fitting value, when the predictive ability is in fact very poor.

- The quality of the model produced after calibration is affected by the concrete variability. When the concrete variability is small, a wrong model is obtained and the fitted model has a very low r^2 value even with a large number of cores ($NC = 10$ in our case).
- In order to assess the quality of fitted model, one must pay attention to how this quality is quantified: looking at r^2 , absolute error $RMSE$ or relative error $RMSE/s(f_c)$ would lead to different conclusions. The calculation of r^2 alone may give misleading indication since this parameter is very sensitive to the scattering of the explanatory variable. $RMSE$ provides a more straightforward estimation of the differences between f_{creal} and f_{cest} and appears to be the most appropriate estimation.
- In practice, the relevant $RMSE$ is the prediction error, which can be much larger than fitting error, and can be estimated only with additional cores.
- The concrete saturation rate is an uncontrolled factor and increasing its variability adds noise to the system. Its effect on $RMSE$ appears (for the set of examples treated in this study), less significant than that of variability in strength.
- The effect of uncontrolled factors, like moisture has been considered in this study, can be efficiently removed by using the combination of NDT techniques.
- For the present case study, it has been shown that, a first technique being given, there is a critical minimal number of cores for making the NDT combination efficient. This number depends on the quality of the second technique and decreases when this quality increases. Synthetic simulations have explained why the combination can worsen the assessment if the number of cores is too small or if the quality of measurements is too poor. This points the main reasons why various experimental studies on this issue have led to controversial results about the interest of combining two NDT.

The synthetic approach had not the ambition of covering all the complexity of real concrete structures. Regarding the conclusions drawn in this paper, our opinion is that all conclusions are correct from a qualitative viewpoint, while they must be considered with some care regarding quantitative statements (critical number of cores, $RMSE$ values, quantitative estimates of sensitivity...), whose estimates probably depends on how the synthetic models Eqs. (1) and (2) are close from real structure.

Many other parameters (not considered in the present paper) are known to influence the investigation strategy and its efficiency. Among them, one can cite for instance the carbonation effect, the existence of damaged zones, or the multi-scale variability (within member, between components, between floors in a multi-story building...). Some work is in progress in order to develop synthetic simulations based on information captured in real case studies. It is

Table A1

Definitions of several terms appeared in this study.

Name	Symbol	Definition
Degree of saturation	S_r	A ratio described the amount of humidity in the hardening concrete and it is changed from 100% for fully saturated condition to 0% for totally dried condition
Root mean squared error	$RMSE$	Its formula from statistic: $RMSE = \sqrt{\sum_{i=1}^n (f_{ci} - f_{cesti})^2 / n}$ where f_{ci} , f_{cesti} respectively the true generated strength at point i and the corresponding estimated strength resulted from the regression model
Coefficient of determination	r^2	Its formula from statistic: $r^2 = (SST - SSE) / SST$ where $SST = \sum_{i=1}^n (f_{ci} - \bar{f}_c)^2$, $\bar{f}_c = (\sum_{i=1}^n f_{ci}) / n$, $SSE = \sum_{i=1}^n (f_{ci} - f_{cesti})^2$
Fitting error curve	FEC	Error calculated for cores points only, using the above formulas with $n = NC$
Prediction error curve	PEC	Error calculated for all points (except cores points), using the above formulas with $n = (NT - NC)$

Table B1

A numerical example presents a sample of generated strengths and NDT values for one simulation.

f_c MPa	S_r	V m/s	R	V_{meas} m/s	R_{meas}
47.7	0.804	4114	44.2	4134	43.1
42.4	0.838	4040	41.3	3820	42.7
49.5	0.828	4162	44.6	3993	44.2
50.5	0.840	4188	44.8	4155	48.4
50.7	0.721	4102	47.0	4266	46.3
45.3	0.836	4094	42.7	3933	43.1
44.6	0.774	4036	43.3	4068	44.1
41.8	0.771	3982	42.1	4077	44.0
46.1	0.910	4158	41.9	4292	40.4
47.6	0.903	4179	42.7	4109	42.8

thus possible to reproduce specific properties like the NDT values distribution or the estimated measurement uncertainty. The general (and ambitious) objective is to contribute to guidelines for optimal investigation strategies for on-site concrete structures.

Appendix A. Definitions

The definitions for several terms appeared in the manuscript are given in Table A1.

Appendix B. Illustrated numerical examples

B.1. An example illustrates the process of generating strengths and NDT values

As it is stated in Section 2, the true concrete strength f_c is assumed to be Gaussian distributed with mean value f_{cm} and standard deviation $s(f_c)$, while a truncated Gaussian distribution (mean value S_{rm} and standard deviation $s(S_r)$) with $S_r \leq 100\%$, is assumed for the degree of saturation S_r . Using the cumulated normal distribution function (in VBA) we generate values for f_c and S_r corresponding to random probabilities (Random function in VBA). Applying these generated values to Eqs. (1) and (2), the true values for V and R are produced. The measurements errors (ε_V and ε_R) are assumed to be distributed in the form of Gaussian distribution with zero mean value and standard deviation (from Table 1): $s(V)$ for velocity test and $s(R)$ for rebound test. Again, using the cumulated normal distribution function (in VBA) we generate values for (ε_V and ε_R) corresponding to random probabilities (Random function in VBA). The errors values are added to the corresponding true values of measurements to produce the measured values of measurements V_{meas} and R_{meas} . A numerical example is presented in Table B1 below for the generated values from one simulation with the following input data:

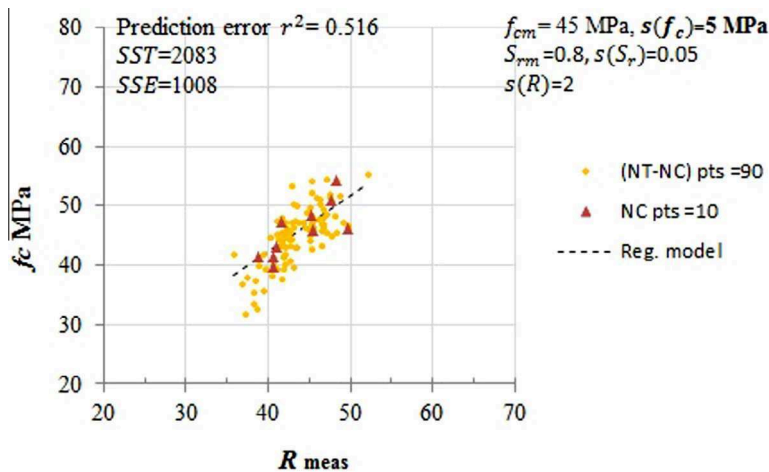


Fig. B1. Studying the effect of concrete strength variability on r^2 , concrete with $s(f_c) = 5$ MPa.

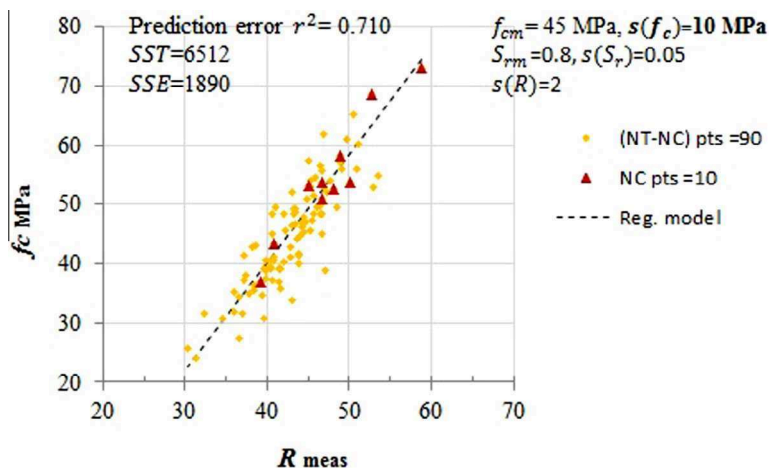


Fig. B2. Studying the effect of concrete strength variability on r^2 , concrete with $s(f_c) = 10$ MPa.

$f_{cm} = 45$ MPa, $s(f_c) = 5$ MPa, $S_m = 80\%$, $s(S_r) = 5\%$,
 $s(V) = 100$ m/s, $s(R) = 2$ and $NT = 10$

B.2. A numerical example explaining why the large scatter in concrete strength has a stabilizing effect on r^2

Two concretes are considered with strength variability of 5 MPa for the first one and 10 MPa for the other. The clouds of points (R_{meas}, f_c) for the two concretes are shown in Figs. B1 and B2. For each case a regression model is produced using 10 cores, then this model is used to calculate the estimated strengths and the values of SST, SSE (see definition of r^2 in Appendix A) and corresponding prediction errors r^2 (all these values are presented in Figs. B1 and B2). It can be noted that increasing $s(f_c)$ from 5 to 10 increases SST value more than three times while SSE is less than twice. Thus when the concrete variability $s(f_c)$ increases, while SST and SSE both increase, the increase rate is larger for SST, and as a result r^2 value increases.

References

- [1] IAEA International Atomic Energy Agency, Guidebook on non-destructive testing of concrete structures, Testing Training Course Series No. 17, 2002.
- [2] CECS 02:2005, Technical specification for detecting strength of concrete by ultrasonic-rebound combined methods, Chinese Standards, 2005.
- [3] EN 13791, Assessment of in situ compressive strength in structures and precast concrete, CEN, Brussels, 2007.
- [4] S.P. Pessiki (chair), In-place methods to estimate concrete strengths, ACI 228.1R-03 report, 2003.
- [5] H. Qasrawi, Concrete strength by combined nondestructive methods simply and reliably predicted, *Cem. Concr. Res.* 30 (2000) 739–746.
- [6] K. Szilágyi, A. Borosnyói, I. Zsigovics, Rebound surface hardness of concrete: introduction of an empirical constitutive model, *Constr. Build. Mater.* 25 (2011) 2480–2487.
- [7] D. Mikulic, Z. Pause, V. Ukraincik, Determination of concrete quality in a structure by combination of destructive and non-destructive methods, *Mater. Struct.* 25 (1992) 65–69.
- [8] R. Pucinotti, Reinforced concrete structure: nondestructive in situ strength assessment of concrete, *Constr. Build. Mater.* 75 (2015) 331–341.
- [9] N. Fawzi, A. Said, A. Jassim, Prediction of compressive strength of reinforced concrete structural elements by using combined non-destructive tests, *Univ. Baghdad J. Eng.* 19 (10) (2013) 1189–1211.
- [10] J. Brozovsky, D. Prochaska, D. Benes, Determination of high performance concrete strength by means of impact hammer, in: 10th Int. Conf. of the Slovenian Society For NDT, Slovenian, Ljubljana, 1–3 September, 2009, pp. 233–241.
- [11] H. Rong, C. Qian, J. Han, C. Wang, The testing strength curves of lightweight aggregate concrete by rebound method and ultrasonic-rebound combined method, *J. Wuhan Univ. Technol. Mater. Sci.* 26 (5) (2011) 1010–1017.
- [12] D. Breyse, Nondestructive evaluation of concrete strength: an historical review and new perspective by combining NDT methods, *Constr. Build. Mater.* 33 (21) (2012) 139–163.
- [13] K. Szilágyi, A. Borosnyói, 50 years of experience with the Schmidt rebound hammer, *Concr. Struct.* 10 (2009) 46–56.
- [14] Ö. Arioz, A. Tuncan, M. Tuncan, T. Kavas, K. Ramyar, K. Kilinc, B. Karasu, Use of combined non-destructive methods to assess the strength of concrete in structures, Afyon Kocatepe University, J. Sci. (2009) 147–154.
- [15] I. Facaoaru, Draft recommendation for in situ concrete strength determination by combined non-destructive methods RILEM Draft Recommendation, *Materials and Structures* 26 (1993) 43–49. RILEM TC 43-CND.
- [16] P. Meynink, A. Samarin, Assessment of compressive strength of concrete by cylinders, cores and nondestructive tests, in: RILEM Symp proc on quality control of concrete structures, Swedish Concr Res Inst, Stockholm, 1979, pp. 127–134.
- [17] Z. Sbartai, D. Breyse, M. Larget, J. Balayssac, Combining NDT techniques for improved evaluation of concrete properties, *Cement Concr. Compos.* 34 (6) (2012) 725–733.
- [18] S. Biondi, E. Candigliota, In situ tests for seismic assessment of RC structures, in: International 14th World Conference on Earthquake Engineering, Beijing, 2008, pp.12–17.
- [19] Q. Huang, P. Gardoni, S. Hurlbeaus, Predicting concrete compressive strength using ultrasonic pulse velocity and rebound number, *ACI Mater. J.* 108 (4) (2011) 403–412.
- [20] A. Al-Ameeri, K. Al-Hussain, M. Essa, Constructing a mathematical models to predict compressive strength of concrete from non-destructive testing, *Int. J. Civ. Eng. Technol.* 4 (2013) 1–20.
- [21] G.F. Kheder, A two stage procedure for assessment of in situ concrete strength using combined non-destructive testing, *Mater. Struct.* 32 (1999) 410–417.
- [22] V. Luprano, F. Caretto, N. Labia, G. Ciniglio, A. Tati, A. Tundo, V. Pfister. Nondestructive test methods for evaluation of concrete: the case study of Punta Perotti (Italy). International Conference on Durability of Building Materials and Components, Istanbul, Turkey 11–14 May, 2008.
- [23] K. Komloš, S. Popovics, T. Nürnbergerová, B. Babál, J.S. Popovics, Ultrasonic pulse velocity test of concrete properties as specified in various standards, *Cement Concr. Compos.* 18 (1996) 357–364.
- [24] D. Breyse, J. Martinez-Fernandez, Assessing concrete strength with rebound hammer: review of key issues and ideas for more reliable conclusions, *Mater. Struct.* 47 (2014) 1589–1604.
- [25] M. Alwash, D. Breyse, Z.M. Sbartai, A parametric study using synthetic simulation for improving the evaluation of concrete strength in existing buildings by NDT techniques, 32th AUGC meeting, Orléans 4–6 (2014).
- [26] ASTM C597–02, Standard test method for pulse velocity through concrete, ASTM Standards, West Conshohocken, 2002.
- [27] U. Lencis, A. Udris, A. Korjakins, Moisture effect on the ultrasonic pulse velocity in concrete cured under normal conditions and at elevated temperature, *Constr. Sci.* 14 (2013) 71–78.
- [28] V. Malhotra, N. Carino, Handbook on nondestructive testing of concrete, second ed., CRC Press, 2004.
- [29] Brozovsky J. Evaluation of calculation correlation efficiency as mentioned in EN13791 in order to determination concrete compression strength by nondestructive testing, in: 10th International conference of the Slovenian society for NDT, Ljubljana, 2009, 221–231.
- [30] A. Galan, Détermination des caractéristiques de résistance du béton d'après les mesures d'ultrasons combinés avec les données sclérométriques, *Matériaux et Constr.* 17 (99) (1984) 201–206.
- [31] P. Knaze, P. Beno, The use of combined non-destructive testing methods to determine the compressive strength of concrete, *Matériaux et Constr.* 17 (99) (1984) 207–210.
- [32] K. Szilágyi, A. Borosnyói, I. Zsigovics, Extensive statistical analysis of the variability of concrete rebound hardness based on a large database of 60 years experience, *Constr. Build. Mater.* 53 (2014) 333–347.
- [33] L. Nobile, M. Bonagura, Accuracy of non-destructive evaluation of concrete compression strength, in: 12th Int. Conf. of the Slovenian Society for NDT, Slovenian, Portoroz, 4–6 2013, pp. 57–64.
- [34] D. Montgomery, G. Runger, Applied statistics and probability for engineers, John Wiley & Sons Inc., USA, 2003.

APPENDIX (B)

This appendix includes our published paper:

Alwash, M., Breyse, D., Sbartāi, ZM., 2017. Using Monte-Carlo simulations to evaluate the efficiency of different strategies for nondestructive assessment of concrete strength. *Materials and Structure*, 50, 14p.

[[DOI 10.1617/s11527-016-0962-x](https://doi.org/10.1617/s11527-016-0962-x)]

Using Monte-Carlo simulations to evaluate the efficiency of different strategies for nondestructive assessment of concrete strength

Maitham Alwash · Denys Breyse · Zoubir Mehdi Sbartai

Received: 24 June 2016 / Accepted: 18 October 2016
© RILEM 2016

Abstract An international benchmark comparing a large variety of concrete strength assessment strategies was proposed by NDT experts of the RILEM committee TC-249. It was based on synthetic data and devoted to the estimation of average strength and concrete strength variability of a series of concrete columns of a single floor building using fixed budget. Based on information gathered after this benchmark, the main contribution of this paper is to simulate, using Monte Carlo simulation, the experts' assessment strategies (i.e. NDT investigation program and data analysis process). These strategies are repeated a certain number of times in order to establish a representative assessment of each strategy. The quality of estimation by each strategy is evaluated using several indicators. The results show that this quality can be improved by using/changing several controlling factors: improving the quality of measurements, selecting core locations depending on the NDT test results (conditional cores), combining several NDT techniques and increasing the amount of resources.

Keywords Benchmark · Concrete strength · Engineering practice · Non-destructive techniques · Monte Carlo simulation · Assessment strategy

1 Introduction

A RILEM Technical Committee TC-249 ISC (non-destructive in situ strength assessment of concrete) was created with the aim of establishing guidelines for an efficient use of non-destructive techniques (NDT) for concrete strength assessment. TC members have decided to carry out a benchmark whose objectives were to identify and compare the expert practices regarding NDT assessment of concrete strength. In this benchmark, depending on the amount of available resources, three knowledge levels (KL1, KL2 and KL3) were considered [1]. Eighteen contributors have proposed different strategies in order to assess the concrete strength of a synthetic building. The main elements of this benchmark have been detailed in [2]. Each strategy has provided two values for each knowledge level:

- the average value of estimated strengths
- the standard deviation of estimated strengths (concrete strength variability).

These two values were given in [2] for all strategies. It is important to highlight the fact that if each strategy is applied to another building having the same

M. Alwash (✉) · D. Breyse · Z. M. Sbartai
I2 M-GCE, University of Bordeaux, CNRS UMR 5295,
33405 Talence, France
e-mail: maitham-fadhil-abbas.alwash@u-bordeaux.fr

M. Alwash
Civil Engineering Department, College of Engineering,
University of Babylon, Babel, Iraq

properties (statistically speaking) then different estimated values of average strength and standard deviation might be derived. Hence, a general conclusion about the efficiency of each strategy cannot simply be drawn from this single result because the effect of chance. The random component of the result cannot be neglected and can be estimated only by repeating the simulation a certain number of times. This would be time-consuming and unpractical without an automatic process.

In this study, using Monte-Carlo simulations [3–6], we will study the efficiency of several strategies proposed in the benchmark after repeating each one 1000 times. For each strategy, the results will be a distribution of 1000-pairs of estimated average strength and standard deviation. These results will be analyzed in order to reach general statements about the efficiency of each strategy and the role of most influencing factors.

Five strategies have been selected among those proposed by the contributors to the benchmark. The selection criterion was the exhaustiveness of information about the process (i.e. about how the measured values are exactly processed and how estimations are derived) and the fact that the process would be fully automatic without any expert interaction. These five strategies (namely B, E, F, J and O5 in [2]) have the advantage of covering the three main possibilities regarding the conversion model identification:

- using a prior model without any calibration (contributor F),
- calibrating a prior model by using NDT test results and core strengths (contributors E, J and O5),
- fitting a specific model between NDT test results and core strengths using the regression analysis (contributor B).

Three of these strategies use a single NDT (rebound R for contributors J and O5, ultrasonic velocity V for the contributor B), while contributors F and E use combined NDT.

2 Monte-Carlo simulation

The simulation process adopted in the present study is exactly the same as that presented and detailed in [2]. The case study is also identical i.e. the same synthetic building (dimensions, properties and applied rules).

The only difference is that the process is repeated herein 1000 times, thus generating 1000 synthetic buildings instead of a single one.

A software coded in VBA has been built by the authors in order to perform the simulation process. Each of the five selected strategies is reproduced within the computer so that it works like proposed by the contributor (i.e. the same type, number, location and quality of measurements and the same method for strength estimation). The flowchart in Fig. 1 illustrates the main steps of the simulation produced in the software.

As it is shown, after selecting a contributor's strategy and the corresponding characteristics, the simulation process is repeated 1000 times. For each simulation, synthetic true strengths are generated at all test locations in the building (i.e. 620 test locations) then the average and standard deviation pair of strengths is derived ($\bar{f}_c, s(f_c)$).

On the same simulation, strength can be estimated from the processing of NDT measurements according to the selected strategy. As a result a pair of average and standard deviation values for the estimated strengths ($\bar{f}_{cest}, s(f_{cest})$) is also derived. After 1000 simulations, for any given strategy, the result is a dataset of 1000 pairs for true strengths ($\bar{f}_c, s(f_c)$) and a dataset of 1000 pairs for estimated strengths ($\bar{f}_{cest}, s(f_{cest})$).

3 Simulation of strategies and the obtained results

The details of the five strategies (B, E, F, J, O5) as proposed by the contributors are presented in this section. Furthermore, the results obtained from simulating and repeating 1000 times these strategies are shown in Figs. 2, 3, 4, 5 and 6. In each figure, two clouds of points are plotted: the cloud of true values 1000 pairs of ($\bar{f}_c, s(f_c)$) and the cloud of estimated values [1000 pairs of ($\bar{f}_{cest}, s(f_{cest})$)]. The results provided by the contributor during the benchmark are added as a single point within the estimated cloud.

3.1 Strategy of contributor B

The contributor identified a specific power model ($f_c = aV^b$) between ultrasonic velocity V measurements and core compressive strengths using regression analysis. He selected 17 test locations for carrying out 17 V measurements having low quality (the options



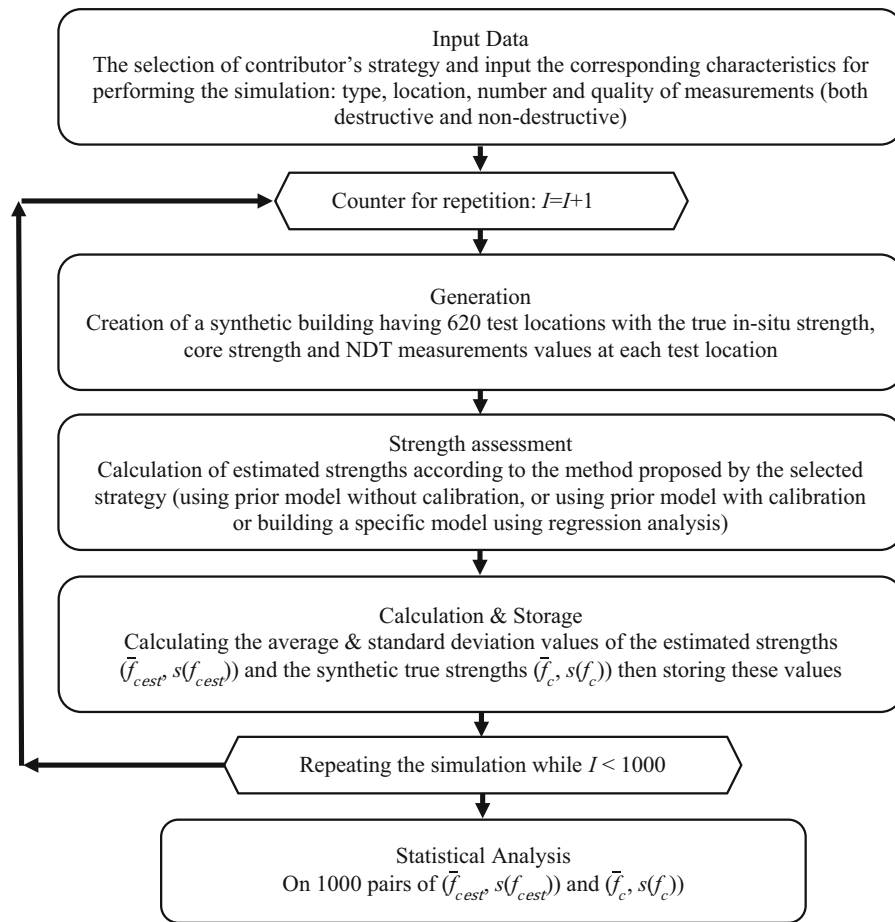


Fig. 1 Flowchart illustrating the main steps of the simulation process

regarding measurement quality were presented in [2]). These V measurements were used to select five low quality cores, whose location was defined based on the value of V measurements, with a specific attention to extreme values: two cores were taken at locations of lowest V , two were taken at locations of highest V and one was taken at a location of median V . Therefore the contributor selected cores in order to cover the whole range of NDT measurements. Cores selected depending on NDT test results are called “conditional cores” in this paper.

When the model was built, the strengths corresponding to 12 V measurements were calculated (test locations without cores). The weighted average

strength \bar{f}_{cest} and weighted strength standard deviation $s(f_{cest})$ values were calculated from 17 strength values, with weight = 1 for core strengths (five strength values) and weight = 1/3 for calculated estimated strengths (12 strength values). The reason of weighting lies on the confidence attributed to the values, which is higher for direct measurements (on cores) than for derived values (from NDT and model).

This strategy is exactly reproduced in the simulator (Fig. 1). The resulting 1000 pairs of $(\bar{f}_{cest}, s(f_{cest}))$ and the corresponding true values $(\bar{f}_c, s(f_c))$ are plotted on Fig. 2. The point $\bar{f}_{cest} = 27.6$ MPa, $s(f_{cest}) = 3.7$ MPa that was provided by the contributor is also shown in this figure.

3.2 Strategy of contributor E

The strategy used average quality measurements (six test locations for V and 12 for R). Three cores were selected by the contributor at predefined test locations (along one diagonal of the building plan: two at extreme corners and the third core at the column closest to the center). The contributor's approach was based on calibrating the prior bivariate power model given in Eq. 1 [5]:

$$f_{\text{cest}} = kV^{2.6}R^{1.3} \quad (1)$$

where k was calculated according to the core strengths and corresponding NDT test results (for more details about the calculation of k see [4]). The model in Eq. 1 was used to assess the estimated strengths at all test locations from NDT test results. Since V and R had been measured together at only six test locations, at additional test locations where only R test result was available, the average of all V values was substituted in Eq. 1 for estimating the strength.

The simulation of this strategy produces the estimated cloud of values shown in Fig. 3. The point $\bar{f}_{\text{cest}} = 25.6$ MPa, $s(f_{\text{cest}}) = 2.3$ MPa provided by the contributor is also plotted on this figure.

3.3 Strategy of contributor F

The approach was based on the use of prior mono-variate models for ultrasonic velocity V and rebound hammer R measurements without calibration, Eqs. 2 and 3:

$$f_{\text{cest}} = 0.0228V - 65.1 \quad (2)$$

$$f_{\text{cest}} = 1.76R - 35.4 \quad (3)$$

The contributor used average quality measurements with 9 V values (seven on site and two on cores) and 12 R values (ten on site and two on cores) in addition to three cores having predefined locations (one on an inner column and the others on side-columns: one column on the long side and the second on the short side of the building). The final average strength and standard deviation of concrete strengths (concrete variability) were calculated from 24 strength values resulting from:

Eq. 2 (9 estimates), Eq. 3 (12 estimates) and cores strengths (3 direct strengths on cores).

Figure 4 shows the results from 1000 simulations, where the point provided earlier by the contributor ($\bar{f}_{\text{cest}} = 23.1$ MPa, $s(f_{\text{cest}}) = 3.7$ MPa) is also highlighted.

3.4 Strategy of contributor J

The approach was based on a prior model which had been identified from a large set of previous data, Eq. 4, [7]:

$$f_{\text{cest}} = k(0.013R^{2.2271}) \quad (4)$$

The contributor selected 29 test locations that cover all columns in the building in order to carry out rebound average quality measurements. Three cores were selected at predefined test locations (one at a corner column and the others at two different interior columns in the building). Like for strategy E, the parameter k was calculated according to the core strengths and corresponding NDT measurements. Each f_{cest} value was multiplied by 0.81 in order to adjust it to the case of cylindrical sample (Eq. 4 gives the strength of cube samples).

Both clouds (true and estimated) and the point provided by the contributor ($\bar{f}_{\text{cest}} = 22.7$ MPa, $s(f_{\text{cest}}) = 4.4$ MPa) are plotted on Fig. 5.

3.5 Strategy of contributor O5

The contributor selected 40 test locations. The first 20 test locations were located at the mid-height of all columns. The other 20 locations were selected to study the variation along the height of several columns. At each test location, average quality rebound measurement was carried out. Two cores were taken at test locations selected according to R test results (conditional cores). However the way of core selections differs from that proposed by contributor B. The cores were chosen at locations respectively corresponding to the median and minimum values of the first 20 R test results. The contributor adopted the prior model provided by EN 13791 standard [8] with calibrating it by a shifting factor Δ , Eq. 5:

$$f_{cest} = 1.73R - 34.5 + \Delta \tag{5}$$

Δ was calculated according to the core strengths and the corresponding R test results (for more details about the calculation of Δ see [4]).

The calibrated model was used to assess the estimated strengths at the 40 test locations. Then the

average strength and strength standard deviation values were calculated.

The estimated cloud resulting from 1000 simulations is shown in Fig. 6 together with the true cloud and the point provided by the contributor ($\bar{f}_{cest} = 28.7$ MPa, $s(f_{cest}) = 4.1$ MPa).

Fig. 2 Comparison of estimated (average strength and strength standard deviation) to the true values; results from 1000-simulations of contributor's strategy B at KL3

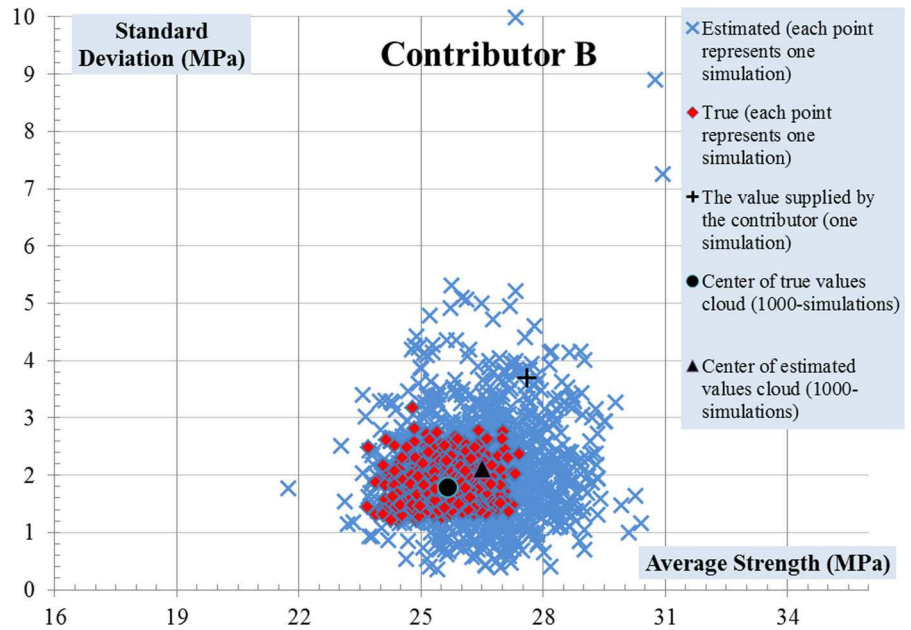


Fig. 3 Comparison of estimated (average strength and strength standard deviation) to the true values; results from 1000-simulations of contributor's strategy E at KL3

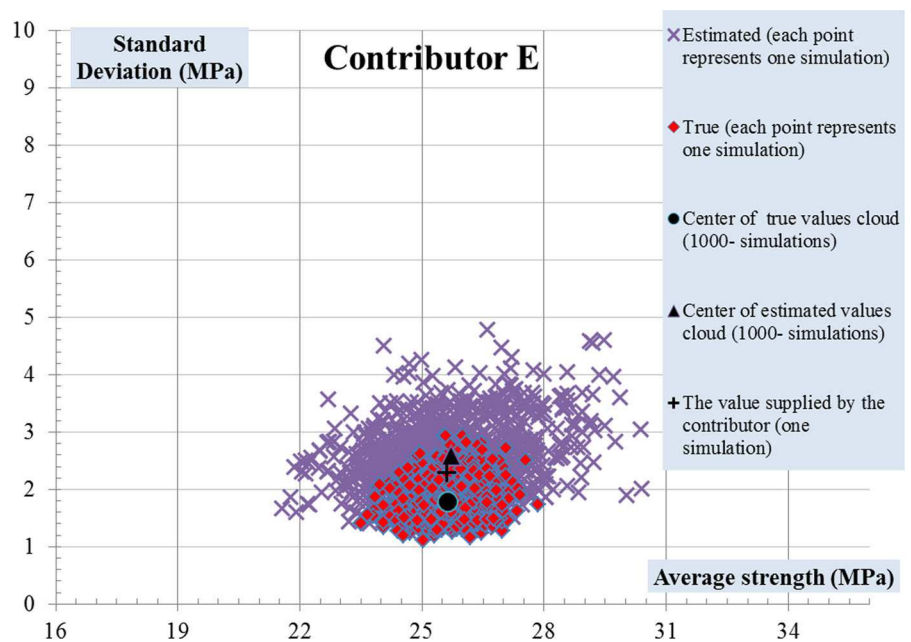


Fig. 4 Comparison of estimated (average strength and strength standard deviation) to the true values; results from 1000-simulations of contributor's strategy F at KL3

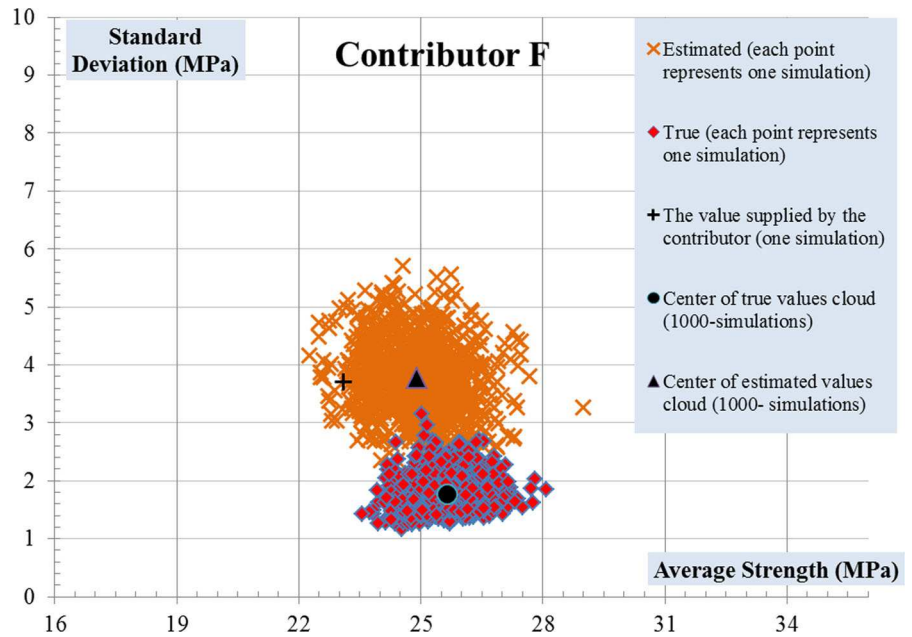


Fig. 5 Comparison of estimated (average strength and strength standard deviation) to the true values; results from 1000-simulations of contributor's strategy J at KL3

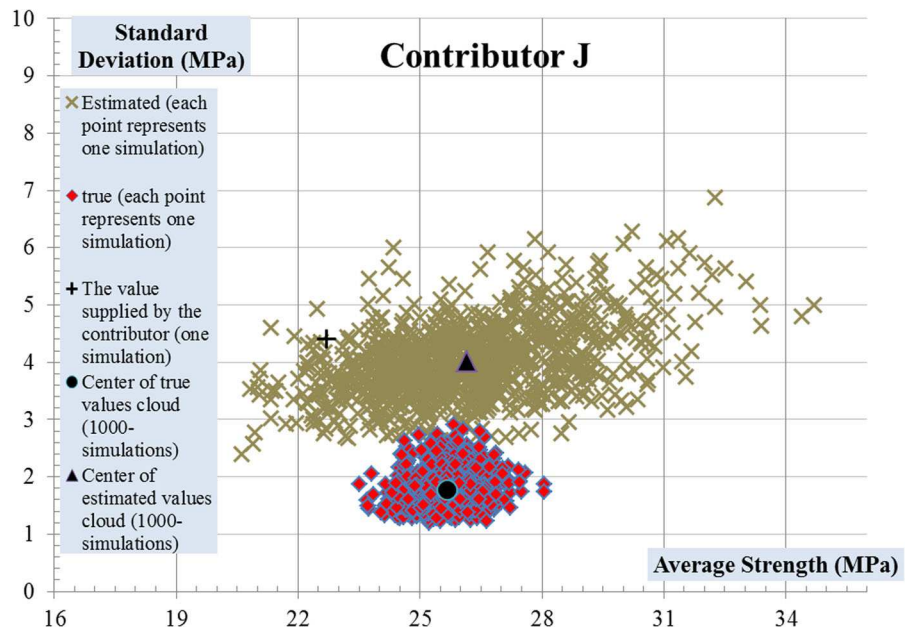
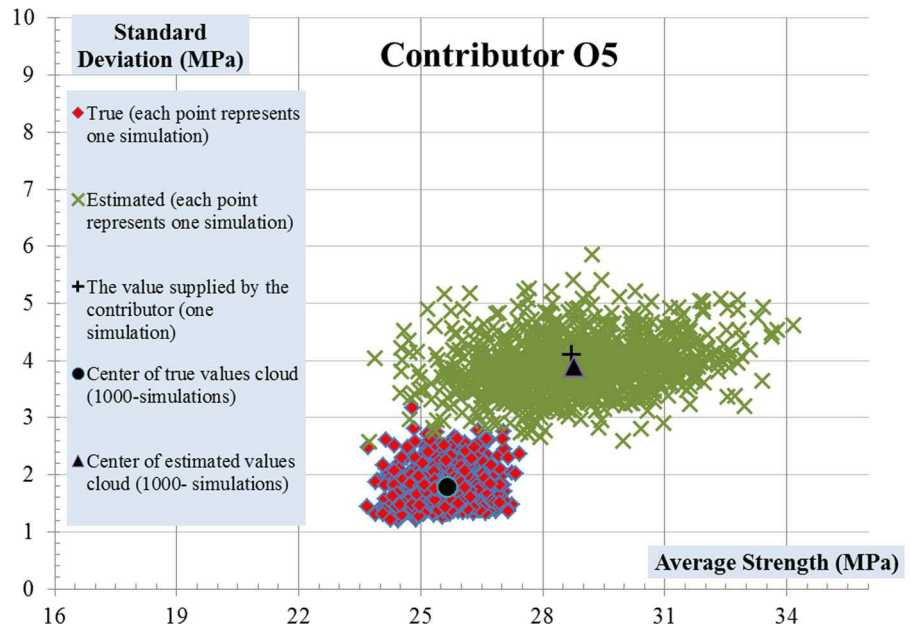


Fig. 6 Comparison of estimated (average strength and strength standard deviation) to the true values; results from 1000-simulations of contributor's strategy O5 at KL3



4 Analysis of simulation results

From Figs. 2, 3, 4, 5 and 6, one can note that the location of the unique point provided by each contributor during the benchmark is not representative of the strategy efficiency. Because of chance, some of these points are close to the center of the estimated cloud, while others lay on its border. This emphasizes that the application of an assessment strategy only once (one point) cannot provide a real picture about the efficiency of this strategy. This justifies the need of simulation for getting a more representative result.

A more refined look on “true clouds” in Figs. 2, 3, 4, 5 and 6 reveals tiny variations between these clouds. This is due to the fact that simulations for the five strategies have been carried on independently, thus resulting in different sets of synthetic buildings. However the statistical properties of these five “true clouds” are similar, with the same average values (\bar{f}_c and $s(\bar{f}_c)$ are respectively equal to 25.6 and 1.8 MPa), the same shape and extension.

The comparison between these figures also highlights that the location of the estimated clouds differ from one strategy to another. Furthermore, the shapes change from one strategy to another but in all cases estimated clouds are more scattered than true clouds. So to evaluate the efficiency of any strategy it is

necessary to quantify the location and shape of its estimated cloud using indicators of estimation quality.

4.1 Indicators of estimation quality

In order to evaluate the quality of the estimation (accuracy and precision) of each strategy, two types of indicators are considered. They respectively correspond to (a) the difference between true values and estimated values (b) to the scatter of estimated values.

The first indicator is the root mean square error (RMSE) which is devoted to quantify the accuracy of forecasting [9, 10]. For each estimated cloud, two values of RMSE are calculated regarding average strength and standard deviation, respectively:

$$\text{RMSE}_{\text{av}} = \sqrt{\sum_{i=1}^{\text{NI}} (\bar{f}_{\text{cest}i} - \bar{f}_{\text{ci}})^2 / \text{NI}} \quad (6)$$

$$\text{RMSE}_{\text{sd}} = \sqrt{\sum_{i=1}^{\text{NI}} [s(\bar{f}_{\text{cest}i}) - s(\bar{f}_{\text{ci}})]^2 / \text{NI}} \quad (7)$$

where NI is the number of simulation repetitions (NI = 1000). These indicators are normalized [11]:

$$\text{NRMSE}_{\text{av}} = \text{RMSE}_{\text{av}} / \bar{f}_c \quad (8)$$

$$\text{NRMSE}_{\text{sd}} = \text{RMSE}_{\text{sd}} / \overline{s(f_c)} \tag{9}$$

where $\overline{f_c}$ and $\overline{s(f_c)}$ are the respective averages of 1000 values of true average strength and strength standard deviation.

The second indicator quantifies the horizontal and vertical scatters of the estimated cloud i.e. the standard deviation of the average strengths $s(\overline{f_{\text{cest}}})$, and the standard deviation of the strength standard deviations $s(s(f_{\text{cest}}))$. The reference values corresponding to the true cloud are $s(\overline{f_c}) = 0.66\text{MPa}$ and $s(s(f_c)) = 0.30\text{MPa}$. The scatters of the true cloud correspond to the randomness of the simulation process, for a given set of concrete properties.

Table 1 summarizes the main features of the selected five strategies and the values of the four indicators. The efficiency of the strategy increases as normalized indicators decrease and as the two other indicators are closer to reference values.

The efficiency of each strategy can be analyzed through the values of the four indicators given in Table 1 and through the location and shape of the estimated clouds in Figs. 2, 3, 4, 5 and 6.

4.2 Comparison of the strategies efficiency

In order to directly compare the five strategies, in addition to Table 1, the cumulative distributed functions (CDF) of estimated average strengths are drawn

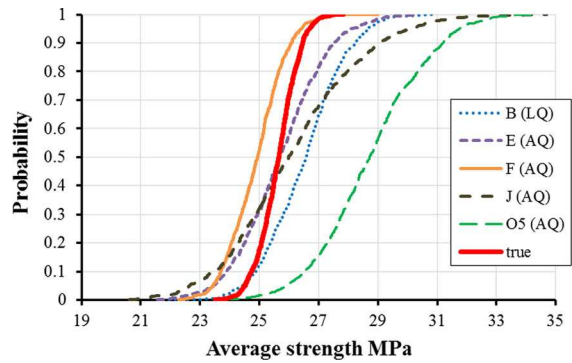


Fig. 7 CDF curves of estimated average strength of the five strategies as compared with the true curve; results from 1000-simulations of contributors' strategies at KL3

on Fig. 7 for all strategies. Figure 8 presents the CDF curves of the estimated concrete variability.

Regarding the estimation of average concrete strength, all strategies can be considered efficient since the normalized error NRMSE_{av} is less than 10 % (strategies B, E, F, J) and 15 % for the strategy O5. However, some strategies are not robust, with a significant probability of having a large over or under estimation, directly linked to the $s(\overline{f_{\text{cest}}})$ value. This problem is particularly relevant with Strategy J and, to a lower extent, with Strategy O5. Strategy F appears to be the most efficient because it has the minimum values of RMSE_{av} and $s(\overline{f_{\text{cest}}})$ but this strategy

Table 1 Characteristics of the five strategies at KL3 and the values of the quality indicators

Str.	Q	No. of measurements			Way of core select	Model identification approach	Quality indicators			
		Core	V	R			NRMSE_{av} %	NRMSE_{sd} %	$s(\overline{f_{\text{cest}}})$ MPa	$s(s(f_{\text{cest}}))$ MPa
B	LQ	5	17	0	CC	Regression analysis	6	54	1.30	0.92
E	AQ	3	6	12	PC	Calibration (k-method)	5	55	1.45	0.57
F	AQ	3	9	12	PC	No calibration	4	116	0.93	0.61
J	AQ	3	0	29	PC	Calibration (k-method)	8	129	2.22	0.64
O5	AQ	2	0	40	CC	Calibration (Δ -method)	14	121	1.78	0.48

Q: quality of measurements (LQ: low quality, AQ: average quality, HQ: high quality)

PC: core locations predefined before the starting of the investigation program

CC: conditional core locations i.e. defined after the first series of NDT measurements



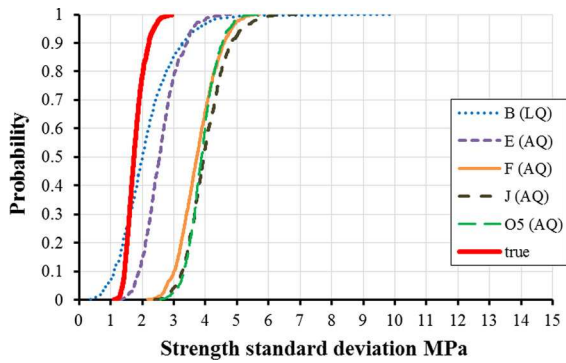


Fig. 8 CDF curves of estimated concrete variability of the five strategies as compared with the true curve; results from 1000-simulations of contributors’ strategies KL3

deserves further comments. It must be pointed that it uses an uncalibrated model. With such a method, the uncalibrated model can in some situations (by the effect of chance) more or less fit with the true properties, but it can also be very far from these true properties. The issue of calibration has been widely documented [8, 12, 13] and uncalibrated models are risky. In other words, in this study the contributor was lucky in his arbitrary of the prior model, which was (by chance) convenient for the concrete under consideration. To analyze this issue more into details, the simulation of strategy F was done again using various alternative models selected from literature. The quality indicators resulting from these simulations are given in Table 2.

Table 2 shows some very bad results (model 3) and some very good results (model 4). Good results are obtained when, because of chance, the prior model fits

well with the specific context. Since there is no guarantee for such a good fit, using an uncalibrated model is by no way a robust method and must be avoided. Therefore the most efficient strategies regarding average strength estimation are strategies B and E.

Regarding the variability of strength, there is a general tendency to overestimate it, with particularly very bad indicators for strategies F, J and O5, for which the overestimation exceeds 100 % (notwithstanding with a large scatter). Strategies E and B get better results, with a lower overestimation. This is probably the beneficial consequence of options privileged in these approaches: (a) the combination of two NDT for strategy E, (b) the conditional cores, using five cores and the weighted average of strength estimates for strategy B. However, the scatter $s(s(f_{cest}))$ is larger for strategy B due to the uncertainty arising from using low quality measurements. Nevertheless, with $NRMSE_{sd}$ values always larger than 50 % these strategies are not fully satisfactory regarding the estimation of concrete variability.

5 How the efficiency of assessment strategies can be improved?

The simple answer to this question is to reduce as much as possible the uncertainty. Many sources of uncertainty exist and affect the global estimation process and the final quality of the diagnostic: statistical uncertainties of sampling [14],

Table 2 Simulation results for alternative models applied with strategy F

No.	Alternative models	References	Quality indicators			
			$NRMSE_{av}$ %	$NRMSE_{sd}$ %	$s(\bar{f}_{cest})$ MPa	$s(s(f_{cest}))$ MPa
1	$f_{cest} = 0.03672V - 129.077$ $f_{cest} = 1.353R - 17.393$	[17]	23	218	0.79	0.56
2	$f_{cest} = 0.05418V - 206.27$ $f_{cest} = 1.47R - 16.85$	[18]	25	448	1.00	0.84
3	$f_{cest} = 0.07692V - 310$ $f_{cest} = 2.5R - 46$	[19]	32	844	1.48	1.23
4	$f_{cest} = 0.0189V - 42.04$ $f_{cest} = 0.9294R - 1.1219$	[20]	2	36	0.51	0.32

A conversion factor (0.8) was applied to transform the cube strength into the equivalent core strength



Table 3 Characteristics of the selected strategies for the parametric study and the values of the quality indicators (in bold: characteristics differ from those in Table 1)

Str.	Q	No. of measurements			Way of core select	Model identification approach	Indicators of quality				
		Core	V	R			NRMSE _{av} %	NRMSE _{sd} %	$s(\bar{f}_{cest})$ MPa	$s(s(f_{cest}))$ MPa	
Studying the effect of amount of resources (KL1)											
B	LQ	2	4	0	CC	Regression analysis	8	114	2.01	2.03	
E	AQ	1	2	6	PC	Calibration (k-method)	10	59	2.71	0.86	
Studying the effect of quality of measurements (KL3)											
B	AQ	3	11	0	CC	Regression analysis	7	82	1.60	1.47	
E	HQ	2	4	8	PC	Calibration (k-method)	4	28	1.24	0.54	
J	HQ	2	0	20	PC	Calibration (k-method)	7	54	1.59	0.48	
Studying the effect of way of core selection (KL3)											
E	AQ	3	6	12	CCA	Calibration (k-method)	5	55	1.40	0.57	
J	AQ	3	0	29	CCA	Calibration (k-method)	7	134	1.48	0.62	
O5	AQ	2	0	40	CCA	Calibration (Δ -method)	6	121	1.68	0.48	

measurement uncertainties [15], influence of uncontrolled factors [16] and model uncertainties. The analysis of the five strategies presented in this paper highlighted the effects of several influencing factors on the quality of estimates. We will analyze in more details these effects, by considering them in a more systematic way. The focus will be given on three items: the quality of measurements, the way of selecting the location of cores and the amount of resources allocated to the investigation program. The general idea is to keep all main features of the strategies while simply varying one of these influencing factors and analyzing how the quality of the assessment is affected. Table 3 summarizes all results, which can be compared to those obtained with original strategies (Table 1).

5.1 Effect of amount of resources

Three cost levels (KL1, KL2 and KL3) have been considered in the benchmark [2] corresponding to a progressively increasing amount of resources. All results discussed above had been obtained at KL3. It seems interesting to see how the quality of assessment

varies when KL changes, Therefore two strategies (B, E) have been selected to be analyzed at level KL1 (amount of resources at KL1 = 1/3 KL3). The numbers of measurements are globally divided by a factor 3 and are provided in Table 3, which can be compared to the numbers for the same strategies at KL3 in Table 1. All other characteristics (quality of measurements, the way of core selection and model identification approach) of these two strategies remain unchanged.

The selected strategies are simulated within the computer and repeated 1000 times. The final results of estimated average strengths and estimated concrete variabilities are respectively plotted on Figs. 9 and 10 besides with the CDF curves provided at KL3. The curves show that reducing the amount of resources from KL3 to KL1 increases significantly the metrics (NRMSE_{av}, NRMSE_{sd}, $s(\bar{f}_{cest})$ and $s(s(f_{cest}))$), see Table 3. This behavior is expected since the reduction in the number of measurements (DT and NDT) causes the uncertainties to grow up. For example, strategy B herein is affected by a dramatical increase in the model uncertainty due to the fact that the two parameters of the regression model are derived from information



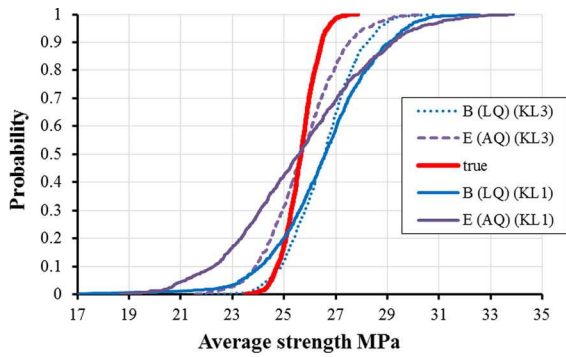


Fig. 9 Effect of KL level on the CDF curves of estimated average strength of the B, E strategies as compared with the true curve (1000 simulations)

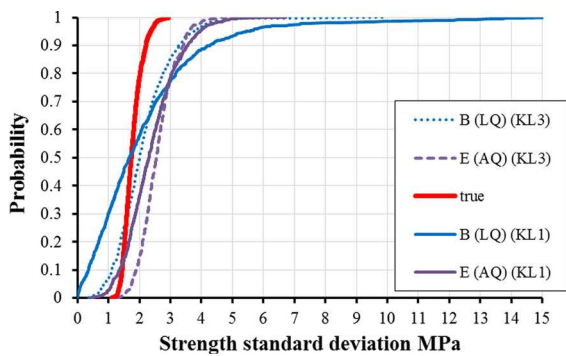


Fig. 10 Effect of KL level on the CDF curves of estimated concrete variability of the B, E strategies as compared with the true curve (1000 simulations)

gathered on only two cores instead of three (Strategy B) or one core instead of two (Strategy E). The result of the identification is thus highly affected by uncertainties.

5.2 Effect of quality of measurements

In the benchmark three quality levels of measurements are considered according to the local variability of repeating measurements within small area. These levels are: high quality (HQ), average quality (AQ) and low quality (LQ). In order to further study this effect, three strategies (B, E and J) have been selected among the five strategies considered in this paper. According to the rules of the benchmark the cost of measurements increases as the quality level improves. The available amount of resources being fixed, the number of measurements reduces when the quality level improves. For the selected three strategies, the

quality level has been increased (from LQ to AQ for strategy B and from AQ to HQ for strategies E and J). Consequently the numbers of measurements reduced accordingly. These numbers are provided in Table 3 and can be compared with the initial values given in Table 1.

After 1000 simulations, CDF curves for the three strategies are plotted in Figs. 11 and 12 for average strength and concrete variability respectively. As a comparison, the original CDF curves for these three strategies, from Figs. 7 and 8, are also plotted herein. The metrics of estimation quality ($NRMSE_{av}$, $NRMSE_{sd}$, $s(\bar{f}_{cest})$ and $s(s(f_{cest}))$) are given in Table 3. From Figs. 11 and 12, Tables 1 and 3, it is clear that improving the quality of measurements has a slightly positive effect on the indicators of the average strength, $NRMSE_{av}$ and $s(\bar{f}_{cest})$, for strategies E and J, while this effect is negative for strategy B. This

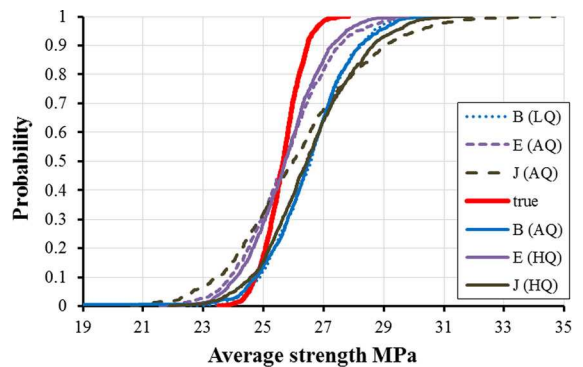


Fig. 11 Effect of improving the quality of measurements on the CDF curves of estimated average strength of the B, E, J strategies

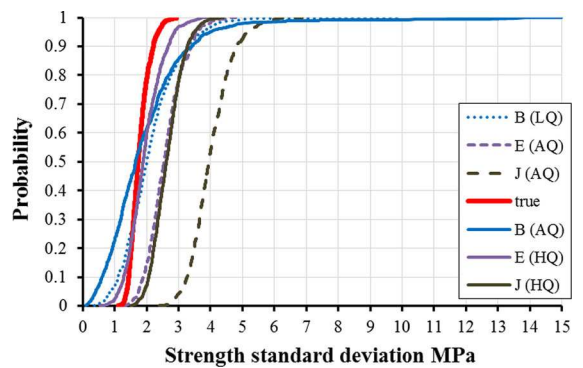


Fig. 12 Effect of improving the quality of measurements on the CDF curves of estimated concrete variability of the B, E, J strategies

opposite effect is more remarkable for the concrete variability: for strategies E and J, NRMSE_{sd} reduces to one half their values in Table 1 while the reduction in $s(s(f_{\text{cest}}))$ is small. On the contrary, for strategy B, NRMSE_{sd} and $s(s(f_{\text{cest}}))$ significantly increase in comparison with the corresponding values given in Table 1.

These apparently inconsistent results can be explained by the controversial effects of improving the quality of measurement on one hand while the number of tests is reduced on the other hand. The result may be either positive for strategies (E, J) or negative for strategy B. The measurement uncertainty has a predominant effect for strategies E and J in which quality increased from average to high.

However, for strategy B improving the quality of measurement (from low to average) reduced the number of cores from 5 to 3 and this second factor had a predominant effect (it can be noted that this strategy is based on a regression model which requires to identify two model parameters).

5.3 Effect of the way of core selection

Two ways of selecting the core location have been considered by the benchmark contributors: predefined cores “PC” and conditional cores “CC”. The idea of the conditional cores is to select cores that cover, as much as possible, the whole range of the NDT measurements distribution. Thus we propose the following rule for selecting cores: (a) rank all NDT test results from the lowest to the highest, (b) subdivide the set into NC subsets (NC = number of cores), (c) take a core at the location which has the NDT value closest to the median value of the subset. This rule will be denominated “CCA”. To study the effect of the way of core selection on the quality of estimation, three strategies (E, J and O5) have been selected. The contributors of the strategies (E, J) used the PC option, while contributor O5 used CC but following a different rule (see §3.5). The CCA rule will be applied to these three strategies. The other characteristics of the strategies are unchanged including the number of measurements, since CCA implies no additional cost.

After 1000 simulations the results obtained with the conditional cores CCA rule for average strength and concrete variability are respectively plotted in Figs. 13 and 14. The values of the indicators (NRMSE_{av} ,

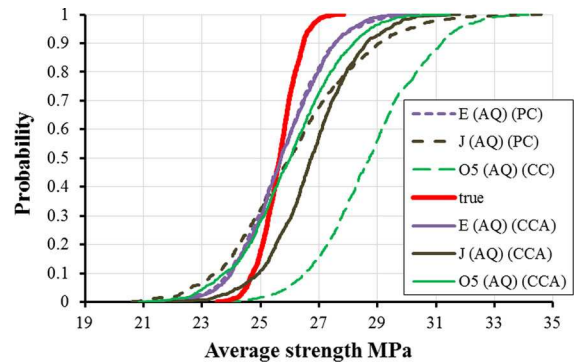


Fig. 13 Effect of conditional cores on the CDF curves of estimated average strength of the E, J, O5 strategies

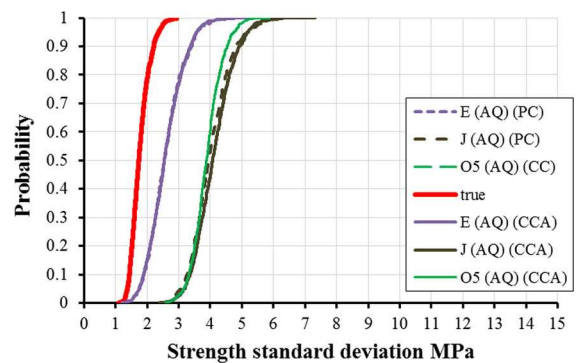


Fig. 14 Effect of conditional coring on the CDF curves of estimated concrete variability of the E, J, O5 strategies

NRMSE_{sd} , $s(\bar{f}_{\text{cest}})$ and $s(s(f_{\text{cest}}))$ are given in Table 3 and can be compared with those of Table 1. It appears that conditional cores can improve both the accuracy and precision (reducing the scatter) of the average strength estimation while it generally has no effect on the accuracy and precision of the estimated concrete variability. Moreover, for the strategy E that already obtained good results, it has effect neither on average strengths nor on concrete variability estimates. The main effect is the decrease of the horizontal scatter of the cloud of the estimated strength, which is particularly visible for strategy J. The explanation is that with a small number of cores (NC = 3 here), the model uncertainties resulting from a too narrow range of variation of the NDT test results may be very large. Conditional cores, by ensuring a better coverage of the strength range reduce these uncertainties. Since conditional cores induce no additional cost and since it can have only positive effect, this process must be strongly recommended.

6 Conclusions and perspectives

In this work, several strategies defined by experts in an international benchmark devoted to assess the concrete strength and its variability have been selected. The efficiency of these strategies has been studied by simulating their application to a series of 1000 synthetic buildings having the same statistical properties and by defining quality indicators of the assessment, in terms of accuracy and precision.

The comparison of the simulated results with those provided by the contributors to the benchmark, confirmed that it is impossible to evaluate a strategy from a unique result (“one shot study”). This statement is valid for synthetic values but can, for the same reasons, be extended to real in situ studies, since the effect of chance (or lack of chance) may prevent any valid statement.

Monte-Carlo simulations are a good tool to study the efficiency of strategies and to get more representative results. The influence on the quality and precision of the assessment of the main parameters which define what is an “assessment strategy” has been analyzed. This work has confirmed that:

- it is very dangerous to consider only the estimated strengths (average values, standard deviation) without considering the variability of these estimates. This will justify to define and quantify the confidence level attached to any uncertainty interval around a true value. With average or HQ test results, few cores with a properly defined strategy may be enough to get a reasonably accurate estimate of average strength. The magnitude of uncertainty related to a given number of cores will be systematically quantified in future studies.
 - the quality of estimates obviously increases when both quantity and quality of test results increase. However when the total amount of resources is fixed, the optimal balance between quantity and quality of test results does not follow simple rules and deserves to be studied through a more comprehensive analysis.
 - properly assessing the concrete variability is a much more difficult challenge than assessing the average concrete strength. The strategies proposed here have shown a general tendency to overestimate the standard deviation, which is conservative but not economically safe.
- the quality of test results, or test result uncertainty, that can be easily assessed on site by simply repeating the test several times is a key factor regarding the efficiency of assessment.
 - conditional cores can improve the quality of assessment without any additional cost and must be strongly recommended,
 - similarly, the weighted average, combining strength measured directly on cores with estimated strength obtained from NDT test results and a calibrated conversion models, has a positive effect, since it reduces uncertainty.

These results will contribute to improve the engineering practice. The RILEM TC ISC-249 is preparing guidelines that will take profit of these conclusions. The quantitative results are probably too case-specific to be considered as general, but a more systematic analysis for a variety of concretes is currently under progress and will lead to quantify the phenomena identified in this paper. The expected result will be a series of quantified relationships between the test result uncertainty and the number of test results on one hand and the final uncertainty on local strength values, average strength and strength variability on the other hand. Once this information will be validated, it will be possible to define an investigation program compatible with the quality of assessment which is looked for.

Acknowledgments The authors gratefully acknowledge all members of RILEM Technical Committee TC-249 ISC for their efforts in preparing this benchmark. Special thanks are due to S. Biondi, M. Fontan, V. Luprano S. and K. Szilágyi for sharing the information about their strategies in the benchmark.

Compliance with ethical standards

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

References

1. Biondi S (2008) The knowledge level in existing buildings assessment. In: 14th world conference on earthquake engineering, Beijing, China, October 12–17, 2008, CAEE Chinese Association. Earthquake Engineering, IAEE International Association. Earthquake engineering, digital paper ID 05-01-0447, Mira Digital Publishing
2. Breyse D, Balayssac JP, Biondi S, Borosnyói A, Cangioliota E, Chiauzzi L, Garnier V, Grantham M, Gunes O, Luprano V, Masi A, Pfister V, Sbartai ZM, Szilágyi K,

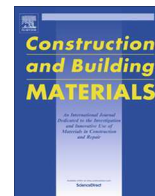
- Fontan M (2015) Comparing investigation approaches and NDT methodologies for concrete strength estimation: an international benchmark. In: NDT-CE conference, Berlin, 15–17 Sept 2015
3. Alwash M, Breyse D, Sbartai ZM (2015) Non-destructive strength evaluation of concrete: analysis of some key factors using synthetic simulations. *Constr Build Mater* 99:235–245
 4. Alwash M, Sbartai ZM, Breyse D (2016) Non-destructive assessment of both mean strength and variability of concrete: a new bi-objective approach. *Constr Build Mater* 113: 880–889
 5. Breyse D (2012) Nondestructive evaluation of concrete strength: an historical review and new perspective by combining NDT methods. *Const Build Mater* 33:139–163
 6. Breyse D, Martinez-Fernandez J (2014) Assessing concrete strength with rebound hammer: review of key issues and ideas for more reliable conclusions. *Mater Struct* 47:1589–1604
 7. Szilagyi K (2013) Rebound surface hardness and related properties of concrete. Ph. D. thesis, University of Technology and Economics, Budapest
 8. EN 13791 (2007) Assessment of in situ compressive strength in structures and precast concrete. CEN, Brussels
 9. Chai T, Draxler R (2014) Root mean square error (RMSE) or mean absolute error (MAE)? —Arguments against avoiding RMSE in the literature. *Geosci Model Dev J* 7:1247–1250
 10. Hyndman R, Koehler A (2006) Another look at measures of forecast accuracy. *Int J Forecast* 22:679–688
 11. Ris R, Holthuijsen L, Booij N (1999) A third-generation wave model for coastal regions. *J Geophys Res* 104:7667–7681
 12. IAEA International Atomic Energy Agency (2002) Guidebook on non-destructive testing of concrete structures. Testing training course series n° 17
 13. Masi A, Vona M (2009) Estimation of the in situ concrete strength: provisions of the European and Italian seismic codes and possible improvements. In: Cosenza E (ed) Proceedings of the eurocode 8 perspectives from the Italian standpoint workshop, Naples, April 1–3, pp 67–77
 14. Pereira N, Romao X (2016) Assessment of the concrete strength in existing buildings using finite population approach. *Constr Build Mater* 110:106–116
 15. American Concrete Institute (2003) In-place methods to estimate concrete strengths, ACI 228.1R-03 report; 2003
 16. 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
 17. Qasrawi H (2000) Concrete strength by combined nondestructive methods simply and reliably predicted. *Cem Concr Res* 30:739–746
 18. Soshiroda T, Voraputhaporn K (1999) Recommended method for earlier inspection of concrete quality by non-destructive testing. In: Proceedings symposium concrete durability and repair technology, Dundee, pp 27–36
 19. Mikulic D, Pause Z, Ukraincik V (1992) Determination of concrete quality in a structure by combination of destructive and non-destructive methods. *Mater Struct* 25:65–69
 20. Mahmoudipour M (2009) Statistical case study on Schmidt hammer ultrasonic and core compression strength tests' results performed on cores obtained from Behbahan cement factory in Iran. NDT in progress 2009. In: 5th international workshop of NDT experts, Prague, 12–14 Oct, pp 189–196

APPENDIX (C)

This appendix includes our published paper:

Alwash, M., Sbartai, ZM., Breysse D., 2016. Non-destructive assessment of both mean strength and variability of concrete: a new bi-objective approach. *Construction and Building Materials*, 113, pp. 880-889.

[<http://dx.doi.org/10.1016/j.conbuildmat.2016.03.120>]



Non-destructive assessment of both mean strength and variability of concrete: A new bi-objective approach



Maitham Alwash^{a,b,*}, Zoubir Mehdi Sbartaï^a, Denys Breyse^a

^aI2M-GCE, University of Bordeaux, UMR 5295, 33405 Talence, France

^bCivil Eng. Dept., College of Engineering, University of Babylon, Babel, Iraq

HIGHLIGHTS

- Bi-objective approach shows high efficiency in capturing the concrete strength variability.
- Regression approach has limited ability in capturing the concrete strength variability.
- Calibration approaches cannot be used to estimate the concrete strength variability.
- Bi-objective, regression and calibration approaches can efficiently assess the mean strength.

ARTICLE INFO

Article history:

Received 15 November 2015

Received in revised form 8 March 2016

Accepted 21 March 2016

Keywords:

Concrete strength
Concrete variability
Characteristic strength
In situ assessment
Rebound hammer
NDT techniques

ABSTRACT

Using non-destructive techniques (NDT) like rebound hammer in combination with destructive techniques (DT) like core test is a common practice. Two approaches are widely used to produce a model for assessing the concrete strength. The first approach consists in fitting a specific model between NDT measurements and cores using the regression analysis. The second approach uses a prior model which is calibrated according to measured core strengths. The EN 13791 and ACI standards require a large number of cores to estimate mean concrete strength and concrete strength variability and consequently to calculate the characteristic strength value which depends on these two inputs. In this work, we propose a new approach for identifying the models based on NDT and DT tests in order to capture both mean strength and concrete strength variability. This approach is first illustrated by synthetic simulations which are a good way to study a problem having many degrees of freedom. The proposed approach is then tested on a real data set. In both cases, it is confirmed that the common approaches are able to estimate the mean strength but they fail, even with a large number of cores, to accurately estimate the concrete variability and hence the characteristic strength. Reversely, the new approach shows its high efficiency in capturing the concrete variability (in addition to the mean strength) with a number of cores lower than that prescribed by the standards.

© 2016 Elsevier Ltd. All rights reserved.

Abbreviations: Superscript⁻, mean value of the variable under consideration; $s(\cdot)$, standard deviation of the variable under consideration; *Test location*, limited area selected for measurements used to provide one test result; $f_{c,core}$, core compressive strength, corresponding to one test location; $f_{c,est}$, estimated individual strength of concrete, corresponding to one test location; $f_{c,uncal}$, estimated individual strength of concrete, corresponding to one test location, produced from using an uncalibrated prior model; R , Rebound number, test result, it is the mean of rebound hammer readings corresponding to one test location; NC , Number of cores; NI , Number of repetitions; NR , Number of test locations for rebound hammer measurements; $RMSE$, Root Mean Squared Error.

* Corresponding author at: I2M-GCE, University of Bordeaux, UMR 5295, 33405 Talence, France.

E-mail address: maitham-fadhil-abbas.alwash@u-bordeaux.fr (M. Alwash).

1. Introduction

Evaluating the concrete compressive strength in existing structures is a common requirement. For example, the change in the use of a structure may require the determination of the concrete strength to accurately assess the structural capacity. There also may be a need to evaluate concrete strength after a structural failure like fire damage or environmental degradation [1]. The seismic retrofitting issue arises nowadays in several countries (like Italy and Turkey) that also emphasizes the need for an accurate in situ assessment of concrete strength in existing structures [2,3].

Destructive technique DT (core test) has many drawbacks: it is expensive, time consuming, sometimes difficult access of coring

machine, only representative of small volume of concrete and has some locally destructive effect on the structure [4]. To overcome these drawbacks, non-destructive techniques NDT can be combined with cores in order to provide more economical evaluation of the concrete compressive strength in the structure. The main challenge is to identify a relationship “conversion model” between the NDT test results and the concrete strength. The existing model identification approaches can be classified into two main categories: regression approaches by identifying a specific model using a limited data set of core strengths and NDT results, and calibration approaches in which a prior model is modified for best agreement with an experimental data.

In the real practice, the mean compressive strength and the characteristic compressive strength are the most common assessed values. The assessment of characteristic strength depends on the mean strength and on the standard deviation of the compressive strengths (concrete variability), thus the concrete variability is also a required value. Furthermore, the ACI 214.4R-03 [5] reported that the coefficient of variation (CV) due to in situ concrete strength variation within a structure (i.e. concrete variability/mean strength) is 13%. However Masi and Chiauzzi [6] found a CV value of 21% within one member of a structure. Masi and Vona [2] studied the concrete variability in many buildings in Italy and they observed that the probable values of CV range between 15% and 35%. Pucinotti [7] also stated that in many cases the CV reaches 35%. That is why the assessment of concrete variability within some homogenous zones (one floor for example) or the whole structure is needed in addition to mean strength value.

Using NDT methods, European Standard EN 13791 [8] allows two approaches (Alternative 1 and 2) for assessing the individual compressive strength values then the mean strength and concrete variability and as a result the characteristic strength. According to the requirements of this standard, the minimum number of cores (NC) is respectively 18 for Alternative 1 (regression analysis approach) and 9 for Alternative 2 (calibration approach). ACI 228.1R-03 standard [1] also requires at least 12 cores (six test locations with two cores at each location) to develop an adequate strength relationship.

In this paper, we present a new model identification approach “bi-objective” that is devoted to capture two material characteristics: the mean and standard deviation of the concrete strength values. Then the prediction capability of bi-objective approach is compared with that of the existing approaches.

The synthetic simulation [9–14] is adopted here to generate a data set (NDT test result and strength values) representative of a synthetic building. These data are used through the present study for testing and validating the proposed approach. The new approach is also applied on a real data set obtained from the scientific literature. In this paper the linear shape of conversion model is considered for all model identification approaches (calibration, regression and bi-objective).

2. Existing approaches for assessing the compressive strength by NDT techniques

The assessment of concrete strength always needs a conversion model establishing the relationship between the compressive strength of concrete and the test results drawn from NDT measurements. There is a consensus to say that there is no universal conversion model that could be used whatever the concrete. In practice, two groups of model identification approaches are widely used to produce a conversion model for assessing the compressive strength.

2.1. Regression approaches

These approaches consist in fitting a specific model between NDT measurements and compressive strength of cores using ordinary least squares method [15–17] or its modified form developed by Mandel [1,18].

2.2. Calibration approaches

They use a prior model (many models exist in literature or standards [8,19–24]) which is calibrated according to the measured core strengths. Two possibilities for calibration are often used in real practice: (1) multiplying factor method, and (2) shifting factor method. In this section, a brief description of each method is presented for the case of rebound hammer technique as an example. However, the same principles are also valid for any other NDT technique like for instance ultrasonic wave pulse velocity.

2.2.1. Multiplying factor method (*k*-method)

The principle comes to update an uncalibrated prior model by a coefficient k to produce a calibrated model,

$$f_{cest}(R) = kf_{c\ uncal}. \quad (1)$$

The coefficient k is calculated as in the following steps:

- Calculate the mean value of core strengths \bar{f}_{ccore} ,
- Use the uncalibrated prior model to calculate the estimated strengths at core locations then take the mean of these values, $\bar{f}_{c\ uncal}$.
- Calculate the calibration factor $k = \bar{f}_{ccore}/\bar{f}_{c\ uncal}$.

2.2.2. Shifting factor method (Δ -method)

The concept here is to shift the uncalibrated prior model by a coefficient Δ ,

$$f_{cest}(R) = f_{c\ uncal}(R) + \Delta \quad (2)$$

The coefficient Δ is calculated as in the following steps:

- Use the uncalibrated prior model to calculate the estimated strength at each core location $f_{c\ uncal, i}$ then,
- Calculate the shifting factor $\Delta = \sum_{i=1}^{NC} (f_{ccore\ i} - f_{c\ uncal, i})/NC$

where $f_{ccore\ i}$ is the compressive strength of core i .

3. The principles of bi-objective approach

From the basics of the existing approaches, it is obvious that none of these approaches has the objective to capture the concrete variability although the standards recommend some of these approaches to estimate the concrete variability because it is an essential parameter in the calculation of the characteristic strength of concrete. Thus we propose here a new “bi-objective” approach which is devoted to capture the variability of concrete strengths in addition to their mean value.

The basic idea is that any investigation program with NDT technique (rebound hammer for example) provides a data set of NC-pairs of (R, f_{ccore}) , where the rebound measurements and core strengths are measured at the same test locations. This data set is used to identify a relationship (conversion model) between concrete strength and the rebound number test results. Usual mathematical shapes of such models have two parameters [9,25]. It is the case for the most common ones: (a) linear models $f_{cest} = aR + b$, (b) exponential models $f_{cest} = a \exp(bR)$, (c) power-law models $f_{cest} = aR^b$. Analytically, two conditions are required in order to

derive the values of the parameters a and b . For bi-objective approach, we consider our two objectives as the conditions for obtaining the unknown parameters, i.e. by ensuring that both mean strength and standard deviation are identical for real values and estimated ones:

$$\bar{f}_{cest} = \bar{f}_{ccore} \quad (3)$$

$$s(f_{cest}) = s(f_{ccore}) \quad (4)$$

where \bar{f}_{cest} , \bar{f}_{ccore} are the estimated and measured (on cores) mean strength values respectively, while $s(f_{cest})$, $s(f_{ccore})$ are the estimated strength variability and the variability calculated from the core strengths respectively.

Applying the first condition on a linear model $f_{cest} = aR + b$, we get:

$$\bar{f}_{cest} = a\bar{R} + b = \bar{f}_{ccore} \quad (3')$$

while the second condition can be rewritten in the form:

$$\begin{aligned} s(f_{cest}) &= s(f_{ccore}) \rightarrow s^2(f_{cest}) = s^2(f_{ccore}) \rightarrow var(f_{cest}) = var(f_{ccore}) \\ var(aR + b) &= var(f_{ccore}) \\ a^2 var(R) &= var(f_{ccore}) \end{aligned} \quad (4')$$

Consequently, the values of the unknown parameters are:

$$a = s(f_{ccore})/s(R) \quad (5)$$

$$b = \bar{f}_{ccore} - a\bar{R} \quad (6)$$

where \bar{R} , $s(R)$ are respectively the mean and standard deviation values of R test results corresponding to the core locations.

The bi-objective approach is presented here with the rebound hammer as NDT technique. However the same principles remain valid for other NDT techniques such as the ultrasonic pulse velocity. This is because in the above development of the parameters (a and b) there are no constraints related to the type of NDT approach.

4. Prediction capability of bi-objective approach as compared with the existing approaches

4.1. Source of data

In order to compare the prediction capability of the above model identification approaches, a data set is required. In this study, synthetic simulation is used to provide a synthetic data by creating a synthetic building in which all information is known: relationships between NDT test results and strength, strength values and NDT test results. To adequately mimic what happens in real life, synthetic data consider both variability multi-scale patterns like they are documented in real case studies and measurement uncertainty. Synthetic data have already been used by the authors in series of applications and have shown their high potential for improving the analysis of data [10,14]. Several publications have detailed how the simulation works, which will not be described here [9,11–13].

The synthetic building considered in this study is a one-story concrete structure in which only columns are considered, as shown in Fig. 1. Each column has 31 test locations distributed along its elevation so the size of data set generated for this building corresponds to $(20 \times 31 = 620)$ test locations. This means that we have 620-pairs of (R, f_{ccore}) and the corresponding true (in the synthetic simulation) values, see Fig. 2 as an example of generated synthetic data. The details of this synthetic building and the generation process were well illustrated in [11].

The main interests of synthetic simulation are: (a) its ability to consider many influencing factors and its flexibility in changing the values of these factors, (b) its ability to generate a huge number of (NDT, DT) pairs within a synthetic building while it is impossible

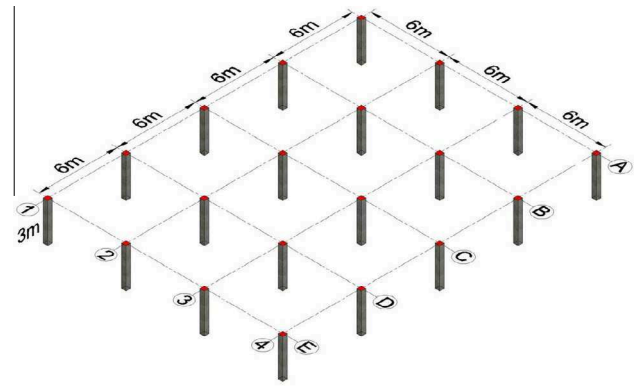


Fig. 1. Synthetic building.

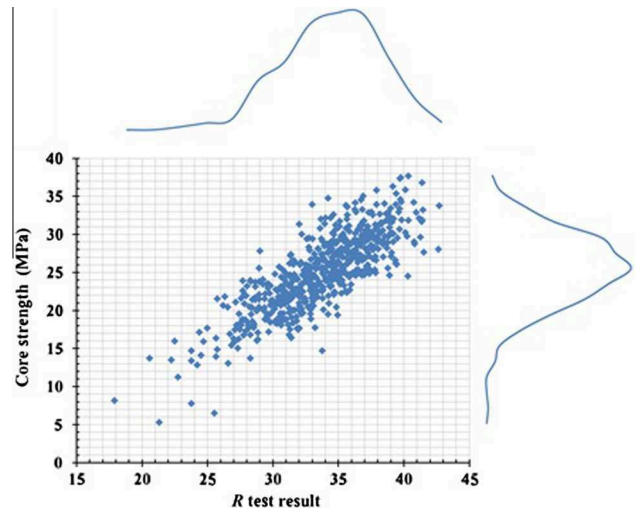


Fig. 2. Synthetic data for a synthetic building having: true mean strength 25 MPa, true concrete variability 4.8 MPa and true mean of concrete saturation rate 65% with a coefficient of variation 4%.

with real building due to the cost and time constraints, (c) the ability of repeating the simulations many times in order to reduce the influence of chance, and (d) the easiness to quantify the efficiency of the approach since “true strength value” is known for each test location in the synthetic building and consequently the difference between “true” and estimate values can be calculated.

4.2. Characteristic of the synthetic data case study

The considered case study has the following characteristics:

- synthetic building having: true concrete mean strength of 25 MPa, true concrete standard deviation of 4.8 MPa. Many influencing factors can affect the NDT evaluation of concrete strength in the concrete structures such as concrete humidity and carbonation. However in this synthetic building the effect of carbonation is considered to be non-significant and only the effect of humidity is taken into account assuming a mean value of concrete saturation rate of 65% (air dried condition) with a coefficient of variation of 4%. The generated synthetic data for this case study is shown in Fig. 2,
- Cores are considered to be extracted horizontally from the concrete zones between the steel reinforcement and they have a standard length to diameter ratio. In order to maintain their internal humidity, cores are assumed to be sealed in a tight envelope,

- single NDT technique: rebound hammer (hammer applied horizontally), $NR = 160$, i.e. 160 test locations selected in each simulation from the total test locations (620),
- test locations are selected randomly.

In this study, all approaches (calibration, regression and bi-objective) deal with linear models. Therefore for calibration approaches several uncalibrated prior linear models are selected from scientific literature [26–29]. Table 1 shows the testing conditions of these models. Moreover, they are illustrated graphically in Fig. 3 with the range of R for which they are derived (the solid part of each line). These four models are selected because they have different slopes that cover a wide range of the existing models for the rebound - strength relationship [9,25].

4.3. Algorithm of comparison between the approaches

In order to illustrate the process of comparison between the approaches, an algorithm has been developed which is presented in Fig. 4. As it is shown, after the generation of the synthetic data the test locations of NDT measurements and cores are selected randomly. Thus NC -pairs of (R, f_{ccore}) are specified. These pairs are used by each approach to produce a conversion model for this approach and to calculate the individual strength values and consequently the values of mean strength, concrete variability and errors $RMSE$ in the individual strengths.

Because this process includes the random selection of the test locations, any conclusion requires its repetition a large number of times (from the selection of test locations to the calculation of errors). Hence this process is repeated NI times. At each repetition, new test locations are selected from the whole synthetic data and as a result new NC -pairs of (R, f_{ccore}) , which leads to a new conversion model for each approach and consequently to new values of mean strength, concrete variability and $RMSE$. The number of repetitions NI depends on the required width of the confidence interval, see Appendix A for the selection criterion of NI . In the present case study $NI = 2500$ is selected to ensure that the one-half width of confidence interval will be less than 0.1 MPa, see Figs. A1 and A2. After the NI repetitions, for each approach, we have NI values of \bar{f}_{cest} , $s(\bar{f}_{cest})$ and $RMSE$. Finally, these results are post-processed in order to derive:

- \bar{f}_{cest} and $s(\bar{f}_{cest})$, which are respectively the mean value and standard deviation of estimated mean concrete strengths, the $s(\bar{f}_{cest})$ is so-called “standard error of mean” by the statisticians,

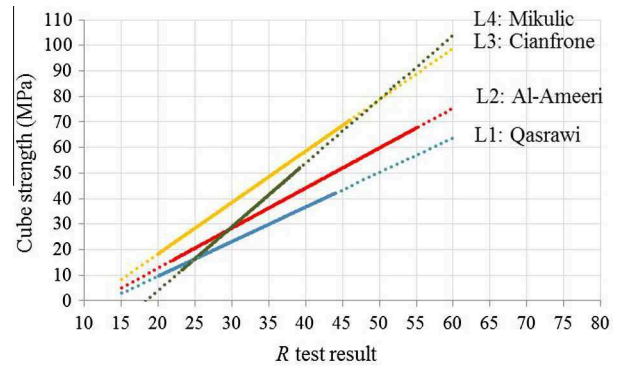


Fig. 3. Uncalibrated prior models selected from literatures [26–29].

- $s(\bar{f}_{cest})$ and $s(s(\bar{f}_{cest}))$ which are respectively the mean value and standard deviation of estimated standard deviation values of concrete strength, the $s(s(\bar{f}_{cest}))$ is so-called “standard error of standard deviation” by the statisticians, and
- \overline{RMSE} , $s(RMSE)$ which are respectively the mean value and standard deviation of the root mean square error calculated on each repetition for the estimated individual strength values.

These six values are calculated for each approach and for several values of NC (from 1 to 20). The results are summarized in Figs. 5–7 and Table 2.

In order to create the synthetic building and perform the process illustrated in this algorithm we developed a computer program using VBA.

4.4. Results and discussion

In this study, we aim to compare the prediction capability of the approaches in estimating three elements: Concrete strength variability, Mean concrete strength and Errors “ $RMSE$ ” in the estimated individual strengths.

4.4.1. Concrete strength variability

The concrete variability estimations resulting from applying different approaches are compared in Fig. 5. It shows four cases L1, L2, L3 and L4 that are identical except regarding the prior model applied with the calibration approaches. Therefore each case corresponds to one of prior models that were selected from scientific literature [26–29], see Fig.3 and Table 1 for more details about these models.

Table 1 Testing conditions of the uncalibrated prior models selected from literatures.

Symb.	Authors	Rebound number	Cube strength (MPa)	Specimen	Concrete age (days)	Concrete humidity	Hammer direction
L1	Qasrawi, 2000 [26]	20–44	6–42.5	cubes 15 cm	Concrete cubes of unknown history made under site conditions were also brought from various sites for testing.	All cubes were immersed under water for a min period of 24 h before testing, just before testing, the cubes were rubbed with a clean dry cloth 24 h after the casting the specimens were demolded and put in water for curing at temp. $(24 \pm 3)^\circ\text{C}$ until the testing day	Horiz.
L2	Al-Ameeri et al., 2013 [27]	22–55	16–72	Cubes 15 cm and 10 cm	28	24 h after the casting the specimens were demolded and put in water for curing at temp. $(24 \pm 3)^\circ\text{C}$ until the testing day	Horiz.
L3	Cianfrone et al., 1979 [28]	20–46	16–76	Cubes 15 cm and 16 cm	1–30	Testing program that was carried out on samples prepared and kept in standard conditions	Horiz.
L4	Mikulic et al., 1992 [29]	23.2–39.2	13–50	Cube 15 cm	28	Not specified	Not specified

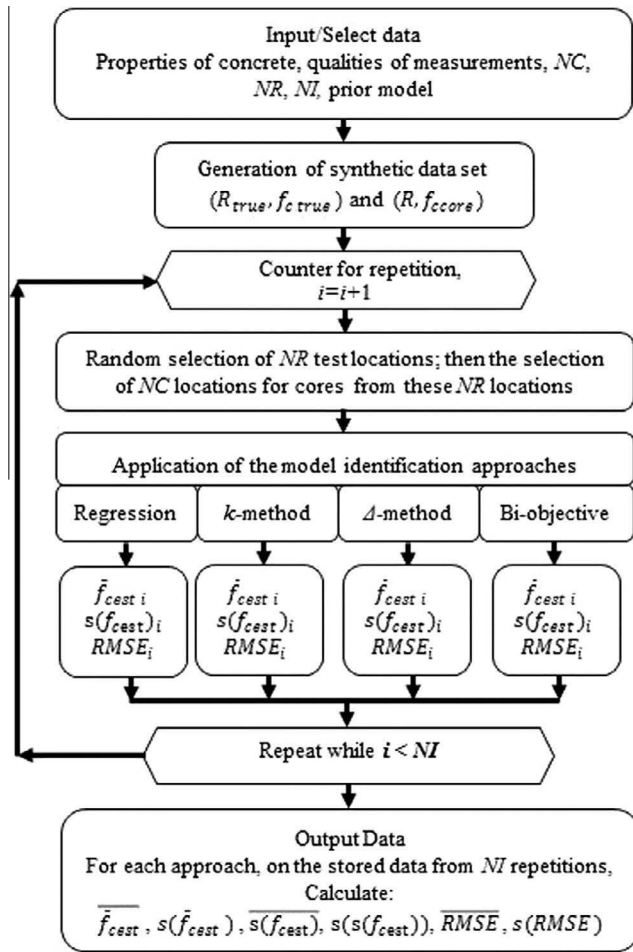


Fig. 4. Algorithm of comparison between the approaches.

The interest of using four prior models is to compare the estimations provided by the approaches that are independent from the prior model (regression and bi-objective approaches) with different cases of the calibration approaches. As a consequence we can get a clear picture about the behavior of the calibration approaches. Each point in Fig. 5 is the mean of NI values of concrete variability, $\overline{s(f_{cest})}$, while the standard deviation of these values, $s(s(f_{cest}))$, is given in Table 2, for case L1 only, in order to avoid the congestion that will be produced if we plot them on Fig. 5.

From Fig. 5, it is clear that the bi-objective approach is the only one which can capture the true concrete variability even with a number of cores lower than the minimum number required by the standards (18 for regression approach and 9 for calibration approach according to EN 13791, and 12 for regression approach according to ACI 228.1R-03).

Furthermore, as NC increases, Fig. 5 shows that the bi-objective curve stabilizes near the value of 5 MPa i.e. it could not reduce the difference from the red line whatever the increase in NC , (the red line represents the true strength variability in the synthetic building considered in the present study, 4.8 MPa). In fact this is not a shortcoming in the bi-objective approach because it well captures its input data (variability of core strengths 5 MPa) which already have uncertainty due to the effects of coring process and other influencing factors. This is interesting because it indicates that the bi-objective approach can accurately capture their inputs.

For the regression approach, it underestimates the concrete variability by about 20%, and increasing NC has no significant effect on the prediction capability. It can also be noted that the curves of

regression approach are nearly similar in the four cases (the same observation can be noted for the curves of bi-objective approach) because they are independent of the prior models, however the slight differences being only due to the random effect of repetitions.

Regarding the calibration approaches (k -method and Δ -method), the behavior is similar in the four cases, nearly horizontal lines but with different values i.e. in some cases they are overestimating the concrete variability while in the other they are underestimating or perfectly fit the true value as it is shown in case L2. These differences are due to the prior model ability to represent the concrete under consideration. Since this ability cannot be evaluated in the real building, the efficiency of estimation by the calibration approaches mainly depends on chance.

As it is mentioned above, each point in Fig. 5 is the mean of NI values of estimated concrete variability, $\overline{s(f_{cest})}$, however it is interesting to know the estimation interval within which the estimated concrete variability value will fall if one has only one iteration (as it is the case in practice). This interval depends on the values of $s(s(f_{cest}))$, in addition to the values of $\overline{s(f_{cest})}$, which are presented in Table 2.

The $s(s(f_{cest}))$ values given in Table 2 decrease as the NC increases for all approaches except the Δ -method in which the values are constant as the NC increases. To understand this result of Δ -method, let's come back to Eq. (2) and take the variance of the estimated strength: $var(f_{cest}) = var(f_{c uncal} + \Delta) = var(f_{c uncal})$, this means that the variation of the estimated strengths is independent of Δ and consequently it is independent of cores (because in Δ -method the effect of cores is taken into account by Δ).

4.4.2. Mean concrete strength

A comparison between the mean strengths predicted by all approaches for the four cases L1, L2, L3 and L4 is illustrated in Fig. 6. Like Fig. 5, each point in Fig. 6 is the mean of NI values of mean strength, $\overline{f_{cest}}$, while the standard deviation of these values, $s(f_{cest})$, is given in Table 2 for case L1 only. It is obvious that all approaches are able to predict the mean strength with acceptable precision as soon as $NC \geq 3$ (cases L1, L2, L3) or $NC \geq 6$ (case L4). However the decision about the suitable value of NC to predict the mean strength also depends on $s(\overline{f_{cest}})$ which controls the estimation interval that corresponds to a single iteration. From Table 2, for all approaches, the values of $s(f_{cest})$ decrease as NC increases with more or less identical decreasing rate.

The bi-objective and regression approaches always have the same value for $NC = 2$ (bold numbers in Table 2). This is logical due to the mathematical fact that for identifying a linear model with two unknown parameters and only two pairs of $(R, f_{c core})$, there is a unique solution whatever the approach, bi-objective or regression. Moreover, the behavior of the bi-objective and regression approaches is quite similar and the two curves are very close especially as NC increases.

For the calibration approaches, the behavior essentially depends on the prior model and its ability to fit the true values of the concrete under consideration. However, increasing NC has a significant effect on the prediction capability of the calibration approaches.

4.4.3. Errors "RMSE" in the estimated individual strengths

In fact there are two types of errors which can be calculated using the individual strength values (individual strength is the strength value corresponding to one test location): fitting error (error calculated for strengths estimated at the test locations where cores are available and have been used for fitting a model) and prediction error (error calculated for strengths estimated at

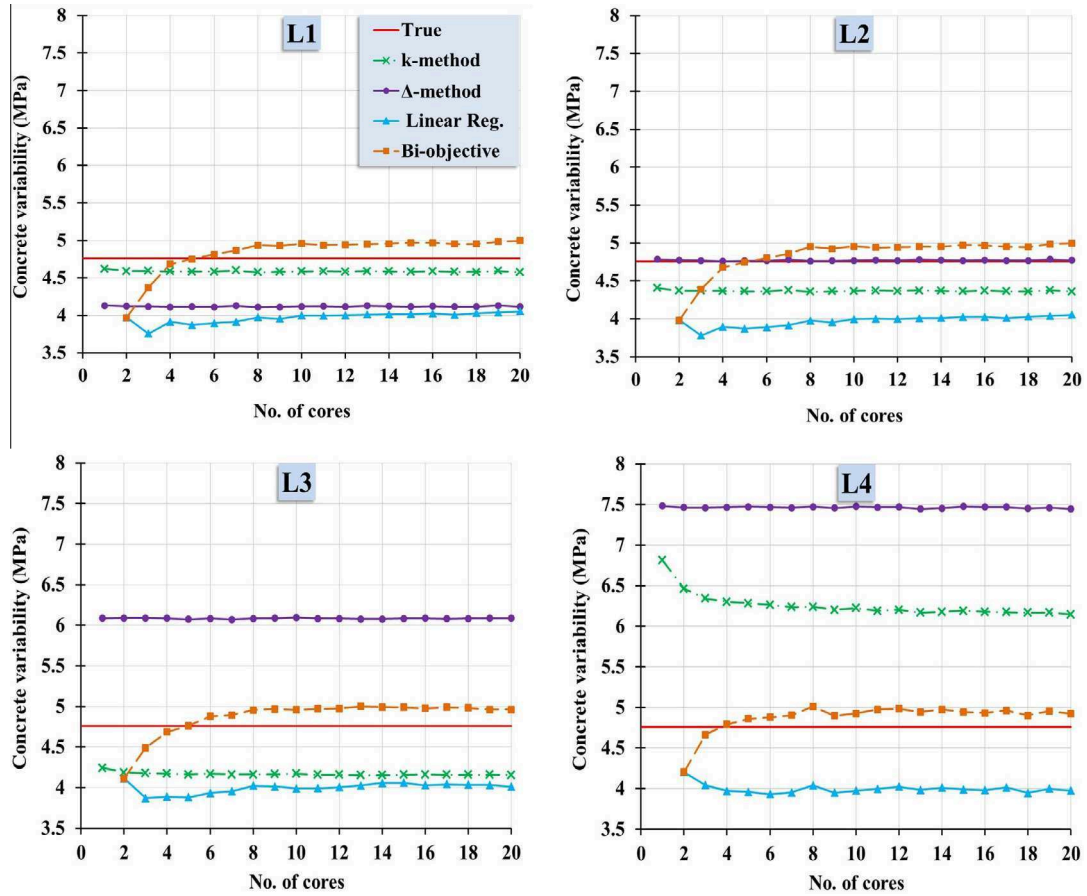


Fig. 5. Concrete variability estimated by different approaches (synthetic data case study).

test locations where only NDT measurements are available) [12]. The second type is more representative of the prediction capability because it measures the errors in values other than those used to derive the model. Thus the prediction error is adopted here to calculate the *RMSE* in the estimated individual strengths. Fig. 7 synthesizes the results (mean *RMSE* values for *NI* repetitions, \overline{RMSE}) for all approaches and the four linear prior models while Table 2 gives the values of $s(RMSE)$ for case L1 only.

In all cases, in Fig. 7, the prediction error reduces as the number of cores *NC* increases. However, no general rule can be derived about the more efficient approach i.e. there is no approach which always produces the minimum error in the four cases. The bi-objective and regression approaches give curves which are close to each other with a better estimation efficiency of strength values by the regression approach as *NC* increases. The behavior of the bi-objective approach (the same observation is noted for regression approach) is similar in the four cases, since they are independent of the prior model and the only differences are due to the randomness of repetitions. Regarding the calibration approaches, in two cases (L1 and L2) the curves of the two methods are nearly identical while in the two other cases (L3 and L4) the *k*-method is the better. It seems that the *k*-method is more efficient than the Δ -method but we cannot confirm the generality of this statement. Moreover, for case L4, the error values resulting from applying the calibration approaches are larger than those given by the other cases. This can be explained by comparing the *R*-values range of the synthetic data (22–42) shown in Fig. 2 with the applicable ranges for the models L1 to L4 shown in Fig. 3 and given in Table 1. It is clear that the synthetic data range is within the applicable ranges for models L1 to L3, while this is not the case for model

L4. Consequently it is recommended to use the models within their applicable ranges.

4.4.4. Failure of the model identified by each approach

A last criterion that deserves to be analyzed when comparing the relative merits of the approaches is related to the fact that, in some situations, the *NC*-pairs of (R, f_{core}) which are used for model identification may produce a very bad or even physically impossible conversion model. It is the case for instance in the following situations:

- when the estimated concrete strength at a test location obtained from using the conversion model has zero or negative value,
- when the conversion model has a negative slope, which would mean that the strength decreases while the NDT value increases, which generally has no physical meaning. This situation can be encountered only with regression approach because the other approaches always guarantee a positive slope.

The risk of having the above situations is larger if the number of cores is small, and if the measurement uncertainties are relatively large as compared to the range of variation of concrete strength in the domain of investigation.

In this study, the developed software includes two conditions (if $f_{cest} \leq 0$ or if slope < 0). When one or both of them is satisfied, the software will reject this model and return to identify another one. Counters are inserted in the software to count the number of failed repetitions (repetitions identify wrong models) for each approach.

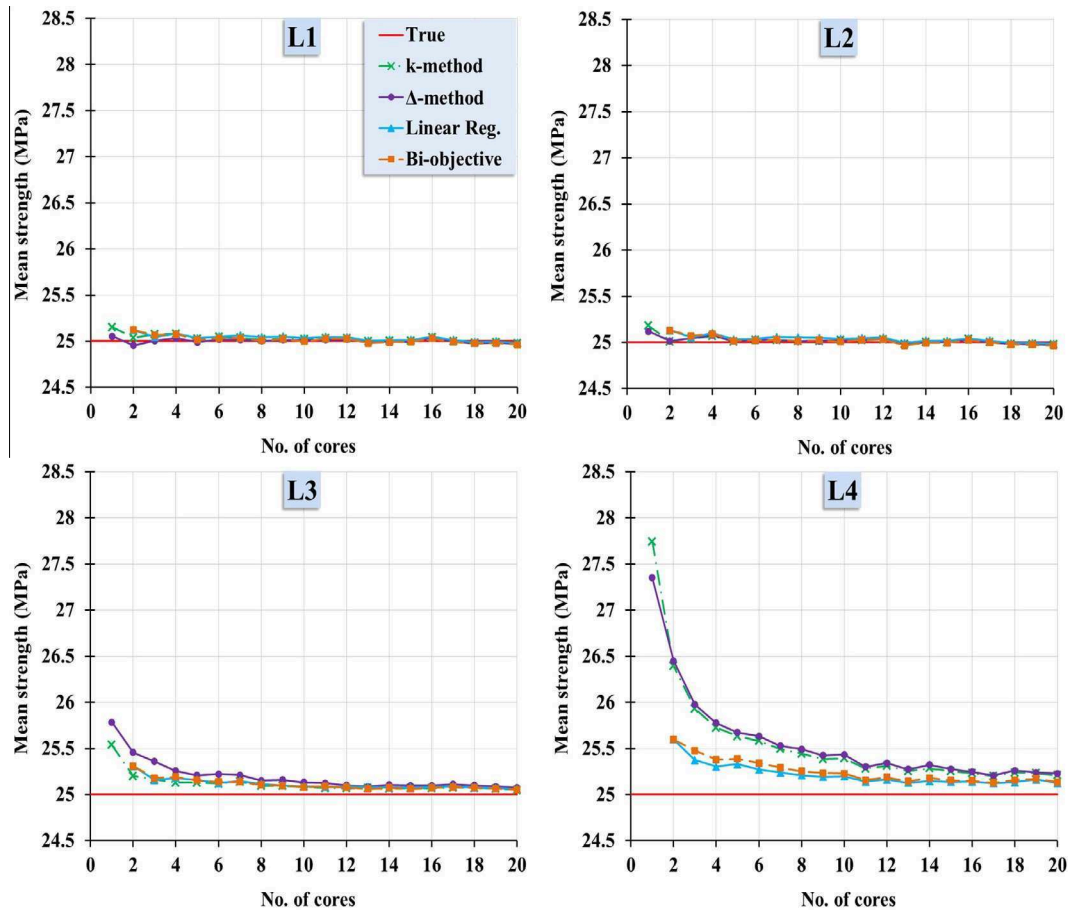


Fig. 6. Mean strength estimated by different approaches (synthetic data case study).

The percentage of failed repetitions with respect to the total number of repetitions for each approach is given in Table 3 for case L4.

For regression and bi-objective approaches, this table shows large failure percentages when NC is small, however these values reduce with increasing NC . The values corresponding to calibration approaches are high due to fact that the prior model L4 has a high slope and as a result the risk of producing a wrong model is high. On the contrary, these values become zero for the case L1 which emphasize the dominative role of the prior model.

When the regression approach and bi-objective approach are compared, the first one appears to be more robust. This criterion provides additional information to be taken into account when choosing the number of cores and the model identification approach. For instance, if the target is the mean strength, the regression approach will be sufficient. On the other hand, if the concrete standard deviation is a target, the bi-objective approach is the only one adapted, but it implies to increase a little bit the number of cores to reduce the risk of having a wrong model.

5. Application of bi-objective approach on a real data set

In order to test the prediction capability of the approaches on real data, a case study is selected from scientific literature. This real data set was provided by [28] as a result of an extensive laboratory testing program that was carried out on samples prepared and kept in standard conditions. It consists in values of concrete strength obtained by carrying out the compressive test on 80 cube samples

(tested at ages from 1 to 30 days), and the corresponding R test results (each test result is the mean value of 18 readings of rebound hammer test applied horizontally on each cube). In the synthetic case, the true values of mean strength and concrete variability were known but they are unknown in this real case. Consequently, they are replaced by the mean and standard deviation values calculated from all cube strengths which are respectively 51.5 MPa and 11.9 MPa.

The algorithm illustrated in Fig. 4 is applied here but with: real data set instead of synthetic one, $NR = 80$, number of samples (cubes) used to identify a model varying from 1 to 10 and for case L1 only, the results are shown in Figs. 8 and 9. Each point in Fig. 8 is the mean of NI values of concrete variability, $\overline{s(f_{cest})}$, while the standard deviation of these values $s(s(f_{cest}))$, is given in Table 4. As the results of synthetic simulation, the bi-objective approach shows its capability to capture the real concrete variability (11.9 MPa in this case) with a number of samples, used to identify a model, much lower than the minimum number required by standards. The regression approach tends to underestimate the concrete variability while the good or bad results of the calibration approaches are only due to the chance (the prior model effect).

Fig. 9 shows the estimated mean strength by all approaches as a function of the number of samples. Each point represents the mean of NI values of mean strength, $\overline{f_{cest}}$, and the standard deviation of these values, $s(\overline{f_{cest}})$, is presented in Table 4. It confirms that the approaches are able to estimate the real mean strength (51.5 MPa in this case) with an acceptable precision when $NC \geq 3$.

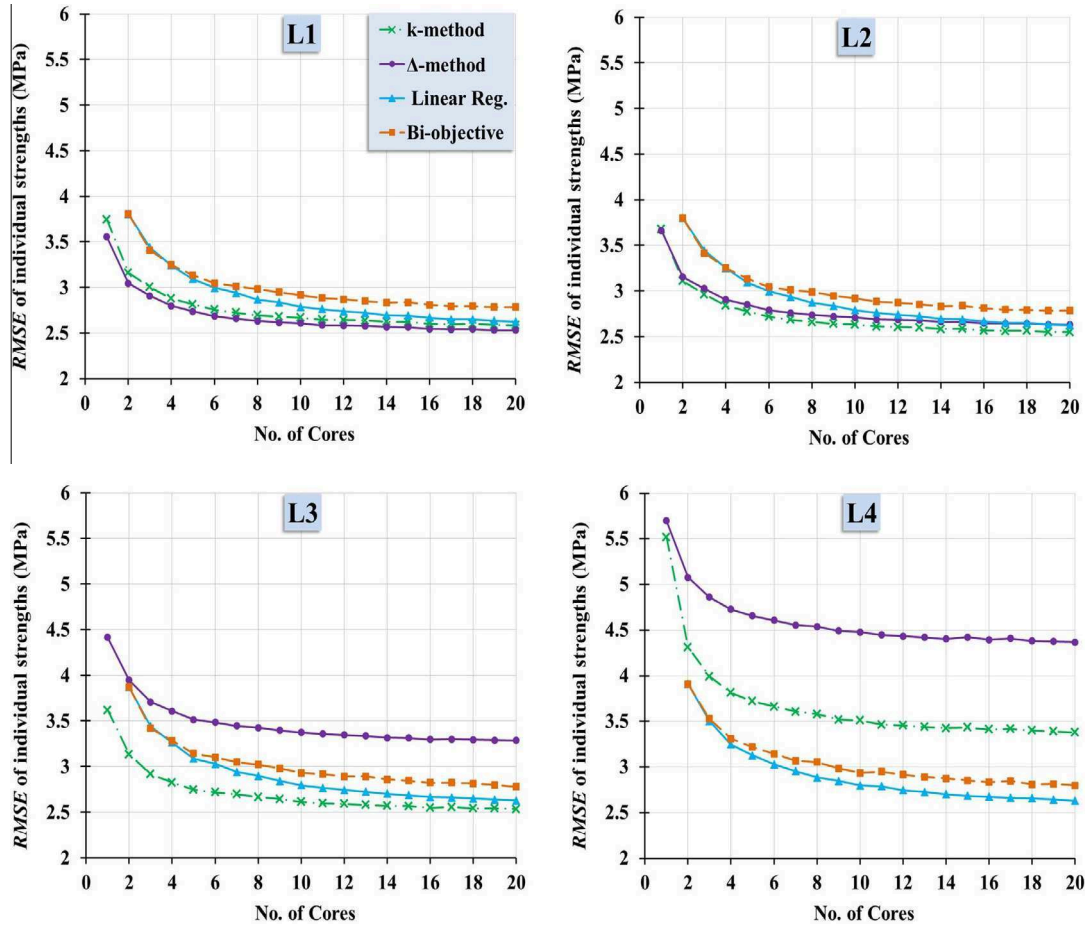


Fig. 7. RMSE in the individual strengths estimated by different approaches (synthetic data case study).

Table 2

Standard deviation values of NI-values of concrete variability, mean strength and RMSE estimated by each approach (L1 prior model for *k*-app. and Δ -app.) – all results in MPa (synthetic data case study).

NC	$s(s(f_{cest}))$				$s(\bar{f}_{cest})$				$s(RMSE)$			
	<i>k</i> -app.	Δ -app.	Reg. app.	Bi-obj.	<i>k</i> -app.	Δ -app.	Reg. app.	Bi-obj.	<i>k</i> -app.	Δ -app.	Reg. app.	Bi-obj.
1	0.6	0.2	–	–	3.2	2.9	–	–	1.7	1.3	–	–
2	0.5	0.2	2.1	2.1	2.1	2.0	2.5	2.5	0.9	0.8	1.4	1.4
3	0.4	0.2	1.7	1.7	1.8	1.7	1.9	1.9	0.6	0.6	0.9	1.0
4	0.4	0.2	1.5	1.5	1.5	1.4	1.6	1.6	0.5	0.5	0.8	0.8
5	0.3	0.2	1.4	1.3	1.3	1.3	1.4	1.4	0.4	0.4	0.7	0.7
6	0.3	0.2	1.3	1.2	1.2	1.2	1.3	1.3	0.4	0.3	0.6	0.6
7	0.3	0.2	1.2	1.2	1.1	1.1	1.2	1.2	0.3	0.3	0.5	0.6
8	0.3	0.2	1.1	1.1	1.0	1.0	1.1	1.1	0.3	0.3	0.4	0.6
9	0.3	0.2	1.1	1.0	1.0	1.0	1.0	1.1	0.3	0.3	0.4	0.5
10	0.3	0.2	1.0	1.0	1.0	0.9	1.0	1.0	0.3	0.2	0.4	0.5
11	0.3	0.2	1.0	0.9	0.9	0.9	1.0	1.0	0.2	0.2	0.4	0.5
12	0.3	0.2	0.9	0.9	0.9	0.9	0.9	0.9	0.2	0.2	0.3	0.4
13	0.3	0.2	0.9	0.9	0.9	0.9	0.9	0.9	0.2	0.2	0.3	0.4
14	0.3	0.2	0.8	0.8	0.8	0.8	0.8	0.8	0.2	0.2	0.3	0.4
15	0.3	0.2	0.8	0.8	0.8	0.8	0.8	0.8	0.2	0.2	0.3	0.4
16	0.3	0.2	0.8	0.8	0.8	0.7	0.8	0.8	0.2	0.2	0.3	0.4
17	0.3	0.2	0.8	0.7	0.7	0.7	0.7	0.8	0.2	0.2	0.3	0.4
18	0.3	0.2	0.7	0.7	0.7	0.7	0.7	0.8	0.2	0.2	0.2	0.4
19	0.3	0.2	0.7	0.7	0.7	0.7	0.7	0.7	0.2	0.2	0.2	0.4
20	0.3	0.2	0.7	0.7	0.7	0.7	0.7	0.7	0.2	0.2	0.2	0.3

Like the case of synthetic data (Table 2), the values in Table 4 decrease as NC increases and for the Δ -method the values of $s(s(f_{cest}))$ exhibit an independent behavior from the changes in NC. However, the values in Table 4 are larger than the correspond-

ing values in Table 2. The last observation can be understood from the following statistical relations: $s(\bar{f}_{cest}) = \sigma/\sqrt{NR}$ and $s(s(f_{cest})) \approx \sigma/\sqrt{2(NR-1)}$ where σ is the standard deviation of the estimated strength population [15,30].

Table 3
Model failure percentage for case L4.

NC	k-app.	Δ -app.	Reg. app.	Bi-obj.
1	26.3	49.8	–	–
2	26.0	49.3	41.2	29.3
3	26.8	48.7	20.7	20.3
4	26.0	48.6	12.4	17.2
5	25.8	48.3	7.4	12.9
6	25.6	47.9	4.4	10.0
7	26.1	47.9	3.0	9.4
8	26.9	47.9	2.2	8.1
9	25.4	46.7	1.5	6.7
10	25.3	46.6	1.3	6.2
11	26.6	47.5	0.7	4.9
12	26.0	46.9	0.7	4.4
13	25.6	45.8	0.5	4.1
14	24.4	45.6	0.6	3.5
15	25.9	46.3	0.3	3.4
16	24.9	46.7	0.2	2.9
17	25.9	46.3	0.1	2.4
18	25.1	47.0	0.1	2.6
19	24.5	45.9	0.1	2.4
20	26.2	47.8	0.2	2.5

Table 4
Standard deviation values of NI -values of concrete variability and mean strength estimated by each approach (L1 prior model for k-app. and Δ -app.) – all results in MPa (real data case study).

NC	$s(s(\hat{f}_{cest}))$				$s(\hat{f}_{cest})$			
	k-app.	Δ -app.	Reg. app.	Bi-obj.	k-app.	Δ -app.	Reg. app.	Bi-obj.
1	1.1	0.0	–	–	4.8	5.7	–	–
2	0.8	0.0	4.2	4.2	3.5	4.1	4.3	4.3
3	0.6	0.0	3.5	3.3	2.7	3.3	3.2	3.2
4	0.5	0.0	3.0	2.7	2.3	2.7	2.5	2.5
5	0.5	0.0	2.7	2.4	2.1	2.5	2.3	2.3
6	0.4	0.0	2.4	2.1	1.8	2.2	2.0	2.0
7	0.4	0.0	2.2	1.9	1.7	2.0	1.8	1.8
8	0.4	0.0	1.9	1.7	1.5	1.9	1.6	1.7
9	0.3	0.0	1.8	1.6	1.4	1.8	1.5	1.6
10	0.3	0.0	1.6	1.4	1.4	1.7	1.5	1.5

6. Conclusions

In this work, the efficiency of calibration and regression approaches is compared with the bi-objective approach (proposed in the present study) for predicting the concrete mean strength and concrete variability (standard deviation of strength). Synthetic data and real data case studies are considered. From both cases, the following conclusions can be drawn:

- All approaches (regression, calibration and bi-objective) can efficiently estimate the mean concrete strength even with NC lower than the minimum number required by the standards. Furthermore, increasing NC can significantly improve the estimation efficiency.
- When trying to quantify concrete strength variability, the bi-objective approach is the only one that shows high efficiency in capturing the true concrete variability even with a number of cores (NC) lower than the minimum number required by the standards. On the contrary, the regression approach remains limited in capturing the concrete variability, it always underestimates the concrete variability and increasing NC has no significant effect on the estimation efficiency. Regarding the calibration approaches, they produce unstable results (estimation efficiency varies as the prior model is changed) so their efficiency depends on chance. Moreover, increasing NC could not improve the results of calibration approaches especially for Δ -method which is proved its independence from the effect of NC .
- For all approaches, the prediction error in the individual strengths reduces as NC increases. However, the comparison between the prediction error curves from various approaches could not lead to a general statement about the best approach for estimating the individual strengths. However, it seems that the regression approach has better estimates of individual strengths than the bi-objective approach.

In order to generalize these conclusions, the bi-objective approach needs a further analysis by testing variety of real data sets that cover a wide range of concrete strengths.

Appendix A. Criterion for selecting the number of repetition NI

The process illustrated in Section 4.3 and Fig. 4 is repeated NI times in order to have more representative results from each approach because the application of an approach one time cannot give real picture about its efficiency. From single application of an approach, one value of \hat{f}_{cest} is produced. Let

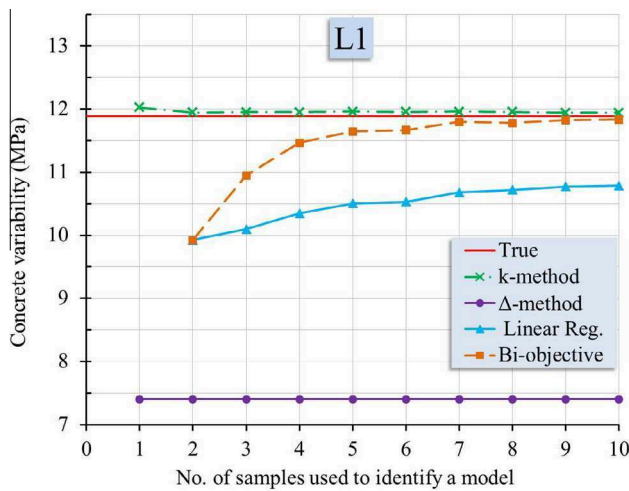


Fig. 8. Concrete variability estimated by different approach (real data case study).

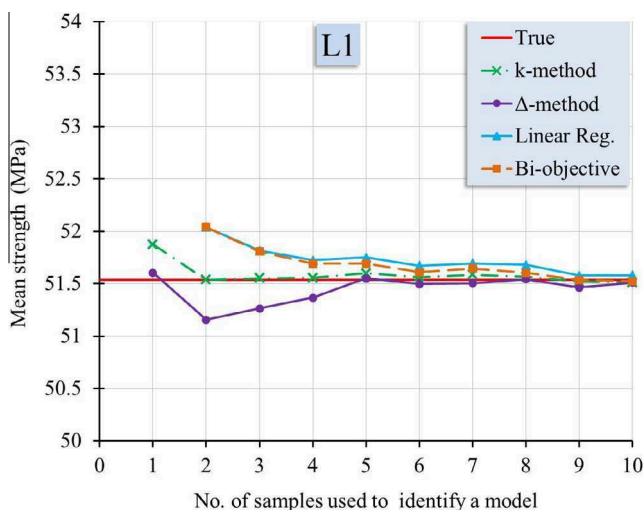


Fig. 9. Mean strength estimated by different approach (real data case study).

$\bar{f}_{cest1}, \bar{f}_{cest2}, \bar{f}_{cest3}, \dots, \bar{f}_{cestNI}$ be a random sample from a population with unknown mean μ and variance σ^2 . Now if the sample size NI is large, the central limit theorem implies that \bar{f}_{cest} has approximately a normal distribution with mean μ and variance σ^2/NI . Therefore $Z = (\bar{f}_{cest} - \mu)/(\sigma/\sqrt{NI})$ has approximately a standard normal distribution. This ratio could be used as a pivotal quantity to produce an approximate confidence interval for μ . However, the standard deviation σ is unknown. It turns out that when NI is large, replacing σ by the sample standard deviation $s(\bar{f}_{cest})$ has little effect on the distribution of Z . This leads to the following confidence interval [15–16],

$$(\bar{f}_{cest} - z_{\alpha/2}s(\bar{f}_{cest})/\sqrt{NI}) \leq \mu \leq (\bar{f}_{cest} + z_{\alpha/2}s(\bar{f}_{cest})/\sqrt{NI}) \quad (A1)$$

so for 95% the one-half width of this interval is $1.96 s(\bar{f}_{cest})/\sqrt{NI}$. In the same way we can find the one-half width of the confidence interval corresponding to concrete variability to be: $1.96 s(s(f_{cest}))/\sqrt{NI}$. For the case study characteristics described in Section 4, using bi-objective approach these two quantities are drawn as a function of NI for (2–10) cores as shown in Figs. A1 and A2. From these figures, for the desired half width of the confidence interval the corresponding NI value can be determined.

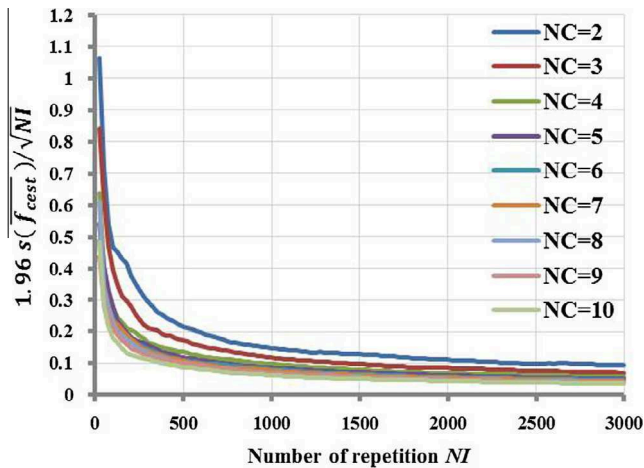


Fig. A1. One-half of confidence interval width for the estimated mean strength as a function of NI .

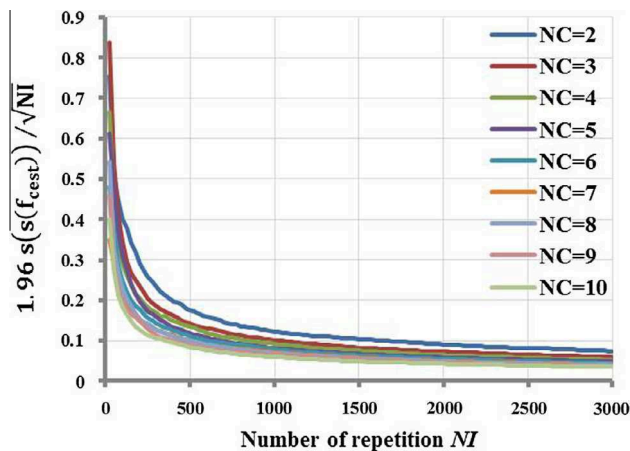


Fig. A2. One-half of confidence interval width for the estimated concrete variability as a function of NI .

References

- [1] S.P. Pessiki (chair). In-place methods to estimate concrete strengths. ACI 228.1R-03 report; 2003.
- [2] A. Masi, M. Vona, Estimation of the in situ concrete strength: provisions of the European and Italian seismic codes and possible improvements, in: E Cosenza (Ed.), Proc. of the Eurocode 8 Perspectives From the Italian Standpoint Workshop, Naples; April 1–3, 2009, 2009, pp. 67–77.
- [3] O. Gunes, Turkey's grand challenge: disaster-proof building inventory within 20 years, Case Stud. Constr. Mater. 2 (2015) 18–34.
- [4] P. Meynink, A. Samarin, Assessment of compressive strength of concrete by cylinders, cores and nondestructive tests, in: RILEM Symp Proc on Quality Control of Concrete Structures, Swedish Concr Res Inst, Stockholm, 1979, pp. 127–134.
- [5] J. Cook (chair). Guide for obtaining cores and interpreting compressive strength results. ACI 214.4R-03, 2003.
- [6] A. Masi, L. Chiauzzi, An experimental study on the within-member variability of in situ concrete strength in RC building structures, Constr. Build. Mater. 47 (2013) 951–961.
- [7] R. Pucinotti, Assessment of in situ characteristic concrete strength, Constr. Build. Mater. 44 (2013) 63–73.
- [8] EN 13791, Assessment of In Situ Compressive Strength in Structures and Precast Concrete, CEN, Brussels, 2007.
- [9] D. Breyse, Nondestructive evaluation of concrete strength: an historical review and new perspective by combining NDT methods, Constr. Build. Mater. 33 (21) (2012) 139–163.
- [10] D. Breyse, J. Martinez-Fernandez, Assessing concrete strength with rebound hammer: review of key issues and ideas for more reliable conclusions, Mater. Struct. 47 (2014) 1589–1604.
- [11] D. Breyse et al., Comparing investigation approaches and NDT methodologies for concrete strength estimation: an international benchmark, in: NDT-CE Conf., Berlin, 15–17 September 2015, 2015.
- [12] M. Alwash, D. Breyse, Z.M. Sbartaï, Non-destructive strength evaluation of concrete: analysis of some key factors using synthetic simulations, Constr. Build. Mater. 99 (2015) 235–245.
- [13] M. Alwash, D. Breyse, Z.M. Sbartaï, Amélioration de la méthodologie d'évaluation non destructive de la résistance mécanique du béton dans les structures existantes, in: 33th AUGC Meeting, Anglet, 27–29 May 2015, 2015, pp. 27–29.
- [14] M. Alwash, D. Breyse, Z.M. Sbartaï, Analysis of key influent factors of investigation for NDT concrete strength estimation in relation with an international benchmark, in: NDT-CE Conf., Berlin, 15–17 September 2015, 2015.
- [15] D. Montgomery, G. Runger, Applied Statistics and Probability for Engineers, John Wiley & Sons Inc., USA, 2003.
- [16] M.S. Ross, Introduction to Probability and Statistics for Engineers and Scientists, fourth ed., Elsevier Inc., USA, 2009.
- [17] W.J. DeCoursey, Statistics and Probability for Engineering Applications, Elsevier Inc., Saskatoon, 2003.
- [18] J. Mandel, Fitting straight lines when both variables are subject to error, J. Qual. Technol. 16 (1) (1984) 1–14.
- [19] G.F. Kheder, A two stage procedure for assessment of in situ concrete strength using combined non-destructive testing, Mater. Struct. 32 (1999) 410–417.
- [20] A. D'Ambrisi, M. Cristofaro, M. De Stefano, Predictive models for evaluating concrete compressive strength in existing buildings, in: International 14th World Conference on Earthquake Engineering, Beijing, Beijing, 12–17 October 2008, 2008.
- [21] R. Pucinotti, Reinforced concrete structure: nondestructive in situ strength assessment of concrete, Constr. Build. Mater. 75 (2015) 331–341.
- [22] L. Rojas-Henao, J. Fernández-Gómez, J. López-Agüi, Rebound hammer, pulse velocity and core tests in self-consolidating concrete, ACI Mater. J. 109 (2) (2012) 235–243.
- [23] E. Proverbio, V. Venturi, Reliability of nondestructive tests for on site concrete strength, in: 10th International Conf. on Durability of Building Materials and Components, Lyon, Lyon, 17–20 April 2005, 2005.
- [24] P. Knaze, P. Beno, The use of combined non-destructive testing methods to determine the compressive strength of concrete, Matér. Constr. 17 (99) (1984) 207–210.
- [25] K. Szilagyi, A. Borosnyoi, 50 years of experience with the Schmidt rebound hammer, Concr. Struct. 10 (2009) 46–56.
- [26] H. Qasrawi, Concrete strength by combined nondestructive methods simply and reliably predicted, Cem. Concr. Res. 30 (2000) 739–746.
- [27] A. Al-Ameeri, K. Al-Hussain, M. Essa, Constructing a mathematical models to predict compressive strength of concrete from non-destructive testing, Int J Civil Eng Technol 4 (2013) 1–20.
- [28] F. Cianfrone, I. Façaoaru, Study on the introduction into Italy on the combined non-destructive method, for the determination of in situ concrete strength, Matér. Constr. 12 (71) (1979) 413–424.
- [29] D. Mikulic, Z. Pause, V. Ukraincik, Determination of concrete quality in a structure by combination of destructive and non-destructive methods, Mater. Struct. 25 (1992) 65–69.
- [30] B. Harding, C. Tremblay, D. Cousineau, Standard errors: a review and evaluation of standard error estimators using Monte Carlo simulations, Quant. Methods Psychol. 10 (2) (2014) 107–123.

APPENDIX (D)

This appendix includes our published paper:

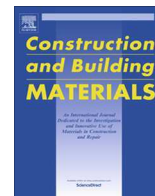
Alwash, M., Breysse, D., Sbartai, ZM., Szilágyi, K., Borosnyói, A., 2017. Factors affecting the reliability of assessing the concrete strength by rebound hammer and cores. *Construction and Building Materials*, 140, pp. 354-363.

[[DOI: 10.1016/j.conbuildmat.2017.02.129](https://doi.org/10.1016/j.conbuildmat.2017.02.129)]



Contents lists available at ScienceDirect

Construction and Building Materials

journal homepage: www.elsevier.com/locate/conbuildmat

Factors affecting the reliability of assessing the concrete strength by rebound hammer and cores



Maitham Alwash^{a,b,*}, Denys Breyse^a, Zoubir Mehdi Sbartai^a, Katalin Szilágyi^c, Adorján Borosnyói^c

^a I2M-GCE, University of Bordeaux, UMR 5295, 33405 Talence, France

^b Civil Eng. Dept., College of Engineering, University of Babylon, Babel, Iraq

^c Department of Constructing Materials and Engineering Geology, Budapest University of Technology and Economics (BME), Muegyetem rkp. 3, 1111 Budapest, Hungary

HIGHLIGHTS

- Selecting core locations based on rebound measurements improves the assessment reliability.
- Reducing repeatability of rebound measurements improves the assessment reliability.
- Increasing the number of cores improves the assessment reliability.
- Concrete intrinsic variability plays an important role on the assessment reliability.

ARTICLE INFO

Article history:

Received 27 October 2016

Received in revised form 2 February 2017

Accepted 22 February 2017

Keywords:

Concrete
Assessment reliability
Mean strength
Strength variability
Rebound hammer
NDT techniques
Cores

ABSTRACT

To assess concrete strength in a structure, nondestructive technique (NDT) like rebound hammer is combined with destructive technique (coring tests) in order to implement a relationship “conversion model” between the compressive strength and, NDT measurements. The conversion model is used to estimate the local strength value at each test location using the corresponding NDT value. Then the estimated mean strength and/or estimated strength standard deviation (concrete strength variability) values are calculated. However, the reliability of these estimated values is always a questionable issue because of the uncertainties associated with the strength predictions based upon NDT measurements. To improve the reliability, the uncertainties must be reduced by specifying and controlling their influencing factors. The objective of this paper is to study the reliability of assessment by analyzing the effects of several influencing factors: number of test locations used to identify a conversion model between strength and rebound measurement NC (number of cores), true value of concrete strength variability, within-test variability of rebound measurements, accepted uncertainty level, quantity to be assessed (mean strength, strength variability), model identification approach (like regression) and the way of selection core locations (random or conditional i.e. selection based on NDT measurements from preliminary investigation). To this end, a large campaign of laboratory studies datasets (1700 test results) was considered for the analysis in the present study.

Results show that NC , within-test variability of rebound measurements and true concrete strength variability have significant effects on the assessment reliability. Conditional selection of cores has also an important effect on improving the reliability so it is strongly recommended.

© 2017 Elsevier Ltd. All rights reserved.

Abbreviations: Superscript $\bar{\cdot}$, mean value of the variable under consideration; $s(\cdot)$, standard deviation of the variable under consideration; $CV(\cdot)$, coefficient of variation of the variable under consideration; Test location, limited area selected for measurements used to provide one test result; f_c , core (or cube) compressive strength corresponding to one test location; f_{cest} , estimated strength of concrete corresponding to one test location; R , Rebound number, test result, it is the mean of rebound hammer readings corresponding to one test location; CV_R , within-test variability of rebound measurements (in terms of the coefficient of variation); NC , represents the number of cores (or cube) used to identify the model, in present study, one core is considered to be provided by each test location, consequently NC also refers to number of test locations for cores; NI , number of repetitions; NR , number of test locations for rebound hammer measurements; NT , total number of test locations of the reduced dataset; U , accepted level of uncertainty, relative value %; T , true reference value; $U \times T$, accepted level of uncertainty, absolute value.

* Corresponding author at: I2M-GCE, University of Bordeaux, UMR 5295, 33405 Talence, France.

E-mail address: maitham-fadhil-abbas.alwash@u-bordeaux.fr (M. Alwash).

1. Introduction

Assessment of in-situ concrete strength in structures is always a challenge for engineers. In the current methodology, nondestructive techniques (NDT) are combined with destructive techniques (coring tests) in order to implement a relationship "conversion model" between the compressive strength and NDT measurements. Regression approach is the most popular approach that is used to identify the conversion model [1–5]. However, in real practice, engineers also use "calibration approach" as a model identification approach [5–7]. This approach is based on calibrating a prior model or basic curve (from literature or standards) according to the measured core strengths. Alwash et al. [7] have recently proposed a new model identification approach so-called "bi-objective approach" which is devoted to capture the concrete strength variability in addition to mean strength, Appendix A provides the principle of this approach.

Thereafter, the conversion model, whatever the model identification approach, is used to estimate the local strength value at each test location using the corresponding NDT value. Then the estimated mean strength and the estimated strength standard deviation (concrete strength variability) values are calculated. However, the reliability of these estimated values is always a questionable issue because repeating an investigation program (i.e. same number of measurements, same techniques for the same building) several times will produce different estimated values.

For decades, studying the reliability of assessing the concrete strength by rebound hammer measurements has been the objective of many scientific researches. However, this issue is quite controversial [8–9]. Some researchers [10–12] are pessimistic considering that rebound hammer is unable to give a reliable estimate of the concrete strength. However, the combination of rebound hammer with the ultrasonic pulse velocity may improve the reliability of assessment [13–14]. On contrast, other researchers like [15] consider that the accuracy of estimation of compressive strength of test specimens cast, cured, and tested under laboratory conditions by a properly calibrated hammer lies between ± 15 and $\pm 20\%$. Furthermore, the probable accuracy of estimation of concrete strength in a structure is $\pm 25\%$ [15–16]. Szilágyi and Borosnyói [17] indicate that the expected error of the strength estimation by the Schmidt rebound hammer under general service circumstances is about $\pm 30\%$. FHWA [18] states that the accuracy of in-situ strength assessment with rebound hammer is between ± 30 and 40% .

Many sources of uncertainty exist and affect the global concrete strength prediction process and the final reliability of the assessment: measurement uncertainties [2,19], true strength variability [20], model uncertainties [20], statistical uncertainties of sampling [21], and influence of uncontrolled factors such as concrete degree of saturation and carbonation [1,22–24].

Moreover, it is necessary to indicate here that the effects of the sources of uncertainty in old structure can differ from that in newly-built structure due to the age effects (i.e. cracks, local damage, steel reinforcement corrosion, etc.). Because of the age effects, more uncertainty is expected in the case of old structure and consequently less reliable assessment.

To improve the reliability, the uncertainties must be reduced by controlling their influencing factors. The objective of this paper is to study the reliability of assessment by analyzing the effects of several influencing factors: number of test locations used to identify a conversion model between strength and rebound measurement NC (number of cores), true value of concrete strength variability, within-test variability of rebound measurements, accepted uncertainty level, quantity to be assessed (mean strength, strength variability), model identification approach (regression,

bi-objective) and the way of selection core locations (random or conditional i.e. NDT based selection). To this end, a large campaign of laboratory studies datasets was considered for the analysis in the present study.

2. Datasets

In order to study the assessment reliability, datasets are required to perform the analysis. Seventeen datasets that belong to different laboratory studies presented in [25] were considered in this paper. Each dataset resulted from one specific laboratory study and specific testing conditions, but the specific study can include one or several mixes with variety of concrete characteristics (mix properties, age, curing, and admixture). The size of datasets varies from 100 to 216 test result pairs (rebound number R , cube strength f_c), i.e. in total more than 2500 test result pairs. For comparison purposes and avoiding statistical biases due to the effect of sampling uncertainty [21], we reduce the size of each dataset by selecting only 100 test result pairs from its original results. Consequently we have a fixed size for all datasets ($NT = 100$). For each dataset, the selection was carried out by ranking R values from minimum to maximum then subdivided them into 100 groups and the median value of each group was selected to be in the reduced dataset. This process of selection ensures that the reduced datasets (each have $NT = 100$) well represented their original datasets.

Table 1 gives the necessary information about these datasets. They cover a wide range of concrete mean strength \bar{f}_c (36–77 MPa) and concrete variability (in terms of strength standard deviation $s(f_c)$ from 6.4 to 17.4 MPa or in terms of strength coefficient of variation $CV(f_c)$ from 11 to 33%). Regarding the rebound number values R (test results), the range of R corresponding to each dataset is also shown in Table 1. Moreover, each R test result value represents the average value of 10 replicates (10 rebound hammer readings on the same surface of a concrete specimen during the laboratory tests). Therefore, within-test variability (or repeatability) of rebound measurements (in terms of the coefficient of variation, CV_R) is known for each test result and the average values of CV_R for each dataset is given in Table 1. Through this study each dataset is represented by the letter D followed by the mean strength value then the value of strength coefficient of variation. The mean strength and concrete variability values given in Table 1 will be called "true or reference" values and used as a reference: estimated strengths will be compared to these true strengths.

3. Research methodology

The methodology adopted in this research was subdivided into three main steps:

- assessing mean strength and strength standard deviation and plotting the cumulative distribution function (CDF) curves,
- Assessing the quality of these estimates by developing risk curves
- Studying the effect of the way of selection the NC test locations on the reliability of assessment.

3.1. Assessing the values of \bar{f}_{cest} and $s(f_{cest})$ for all datasets and plotting the CDF curves

In real practice, to assess the concrete strength in a structure, the engineer establishes an investigation program: he carries out NDT measurements (rebound hammer in this study) at a number of test locations (NR) and from some of these test locations (NC)

Table 1
Datasets characteristics ($NT = 100$ for each dataset).

Dataset No.	Symbol	\bar{f}_c (MPa)	$s(f_c)$, MPa	$CV(f_c)$ %	R-Range	CV_R %
1	D67-11	67.19	7.65	11.38	40–55	3.24
2	D77-15	77.40	11.24	14.52	42–53	3.12
3	D69-17	69.11	12.05	17.43	34–53	2.93
4	D36-18	35.78	6.42	17.95	32–44	5.77
5	D60-19	59.96	11.34	18.92	34–51	3.79
6	D63-19	63.24	12.02	19.00	42–57	3.39
7	D72-20	71.62	14.11	19.71	32–55	3.28
8	D65-20	64.92	12.88	19.84	34–59	5.12
9	D46-21	45.84	9.74	21.26	25–46	5.00
10	D42-22	41.82	9.29	22.21	27–38	8.84
11	D71-23	70.60	16.46	23.32	35–56	3.11
12	D43-24	42.50	10.16	23.90	29–45	4.82
13	D55-25	55.11	13.79	25.03	25–48	5.25
14	D44-25	44.36	11.18	25.21	24–45	5.33
15	D55-27	54.91	14.9	27.13	26–50	4.06
16	D63-27	63.45	17.35	27.34	23–49	4.86
17	D46-33	45.55	14.87	32.66	21–49	4.84

cores are extracted. Then he identifies a model between the concrete strength and NDT measurements using one of the model identification approaches (regression, calibration or bi-objective). In order to mimic the reality this process was simulated according to the flowchart shown in Fig. 1. This flowchart illustrates how the process was applied to each dataset given in Table 1. For each dataset ($NT = 100$) two successive random selections were carried: (a) the choice of NR test locations ($NR < NT$, here $NR = 60$), (b) the choice of the NC core locations, NC being iteratively increased from

3 to 15. Since the specimens in this study are cubes, then NC represents the number of cubes used in the model identification process. For each NC , and each simulation, two linear models ($f_{cest} = aR + b$) were identified using regression and bi-objective approaches. Then, each model was used to calculate the estimated strength values, test results, at NR test locations and as a result to calculate the estimated mean strength \bar{f}_{cest} and estimated concrete variability $s(f_{cest})$. This process was repeated NI times ($NI = 1000$) so the final outputs were, for each NC value and each approach, a

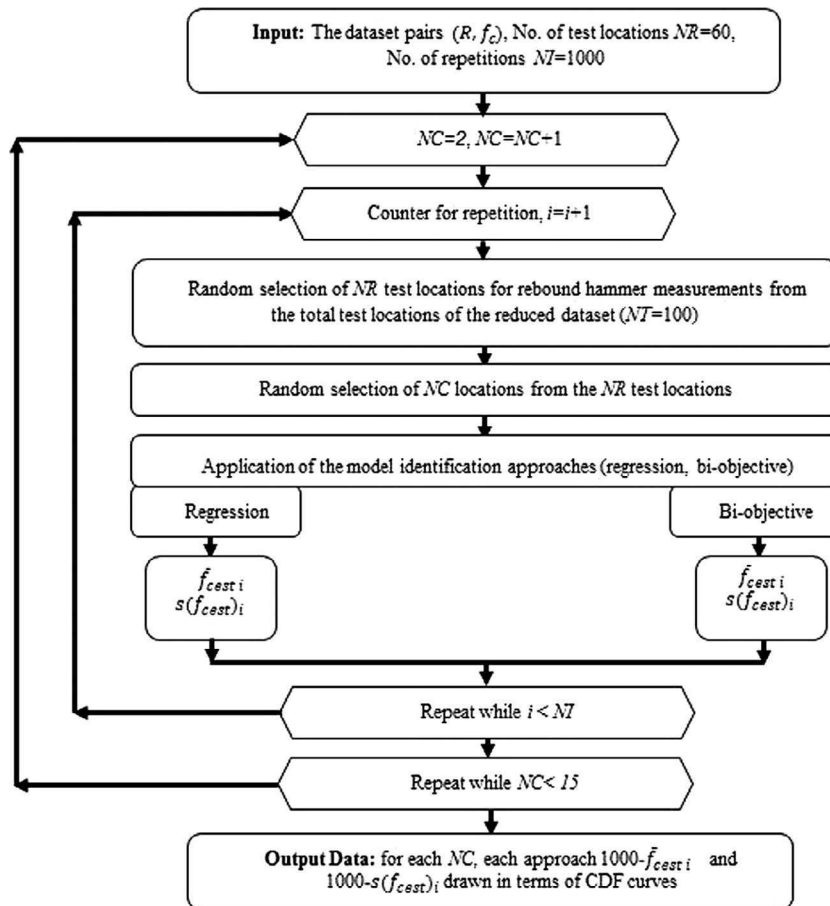


Fig. 1. Algorithm assessing the values of \bar{f}_{cest} and $s(f_{cest})$.

series of 1000 values of \bar{f}_{cest} and a series of 1000 values of $s(f_{cest})$ were produced. The adopted way to show these results is to draw each series in terms of cumulative distribution function CDF.

To achieve this process a simulator was built. It was coded using the programming language VBA (Visual Basic for Applications) that integrated with Excel, for details information about VBA see [26].

3.2. Assessing the quality of estimation by developing “Risk Curves”

As explained in the introduction, a common indication about the reliability of assessment using NDT is providing the estimated value with a $\pm U\%$ relative error (uncertainty). In fact, due to many sources of uncertainties, the estimated values are scattered and their CDF can be plotted. The reliability of estimation is improved as the scatter reduces. Therefore factors affecting the scatter can also control the reliability. To correlate these factors with the reliability of assessment it is preferable to proceed in two steps:

- (a) Using the CDF curve, fix an interval $\pm(U \times T)$ around the true reference/target value, T ,
- (b) Quantify the probability (or risk) that the estimated value lies outside the fixed interval.

The concept to derive the risk values (i.e. probability of a wrong prediction, outside the fixed interval) from CDF curve is illustrated in Fig. 2. Risk values can be finally plotted as a function of NC to establish new curves so-called “Risk curves”.

3.3. Studying the effect of the way of selection the NC test locations on the reliability of assessment

The way of selection the NC test locations may also play a role on the reliability of assessment. Methodology presented in Fig. 1 is based on the random selection of the NC test locations. However these NC locations can be defined in accordance with certain conditions applied on the NDT results “conditional selection”. The idea of the conditional selection is to select the NC test locations that cover, as much as possible, the whole range of the NDT measurements distribution. The flowchart of Fig. 1 was thus applied with a simple change (conditional selection instead of random selection

of NC locations). The following conditional selection rule was applied: (a) rank all NDT test results from the lowest to the highest value, (b) subdivide the set into NC subsets, (c) select a test location where the NDT value is closest to the median value of each subset. The same rule for the conditional core selection was earlier applied on a synthetic dataset (generated dataset using Monte-Carlo simulation) presented in a benchmark carried out by the members of RILEM TC-249 committee [27–28]. It shows an improvement in the reliability of the assessment so it is desired herein to apply it on real laboratory datasets. However, this is not the only possible rule for the conditional selection of cores depending on the NDT values. For example, [29] proposes a method that is based on identifying the location of a set of NC coring points which can provide a statistical distribution of NDT measurements similar to that provided by all NDT measurements.

4. Results and discussion

4.1. Cumulative distribution function CDF of \bar{f}_{cest} and $s(f_{cest})$

As illustrated in Section 3.1, for each dataset each NC value and each approach, a series of 1000 values of \bar{f}_{cest} and a series of 1000 values of $s(f_{cest})$ were produced and presented in terms of cumulative distribution function CDF. For each dataset, the simulator provided CDF curves in four figures corresponding to the two estimated quantities (\bar{f}_{cest} and $s(f_{cest})$) with two approaches (regression and bi-objective) for each estimated quantity. Since it is impractical to display herein the figures for all datasets, only the figures corresponding to dataset D42-22 are shown in Fig. 3.

Fig. 3 illustrates how the CDF curves vary with NC . In fact the analysis was performed for NC values (3–15) however to avoid the curves congestion, the figures show only the curves that corresponding to NC values (3, 6, 9, 12, 15). Fig. 3a and b refer respectively to the mean strength estimation by regression and bi-objective approaches while Fig. 3c and d refer to strength variability estimation. On the four figures the reference line (red vertical line) represents the true reference value of the dataset which is the target of assessment, given in Table 1, respectively 41.82 MPa for \bar{f}_c and 9.29 MPa for $s(f_c)$. Therefore the reliability of assessment is better as the CDF curve is closer to the reference line (i.e. less

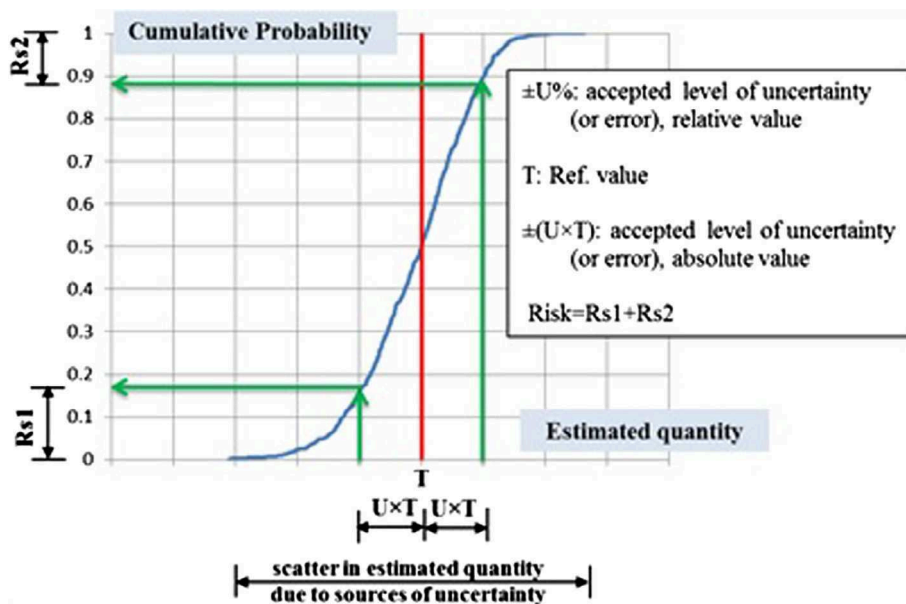


Fig. 2. Proposed concept to derive risk value that corresponding to certain uncertainty using CDF curve.

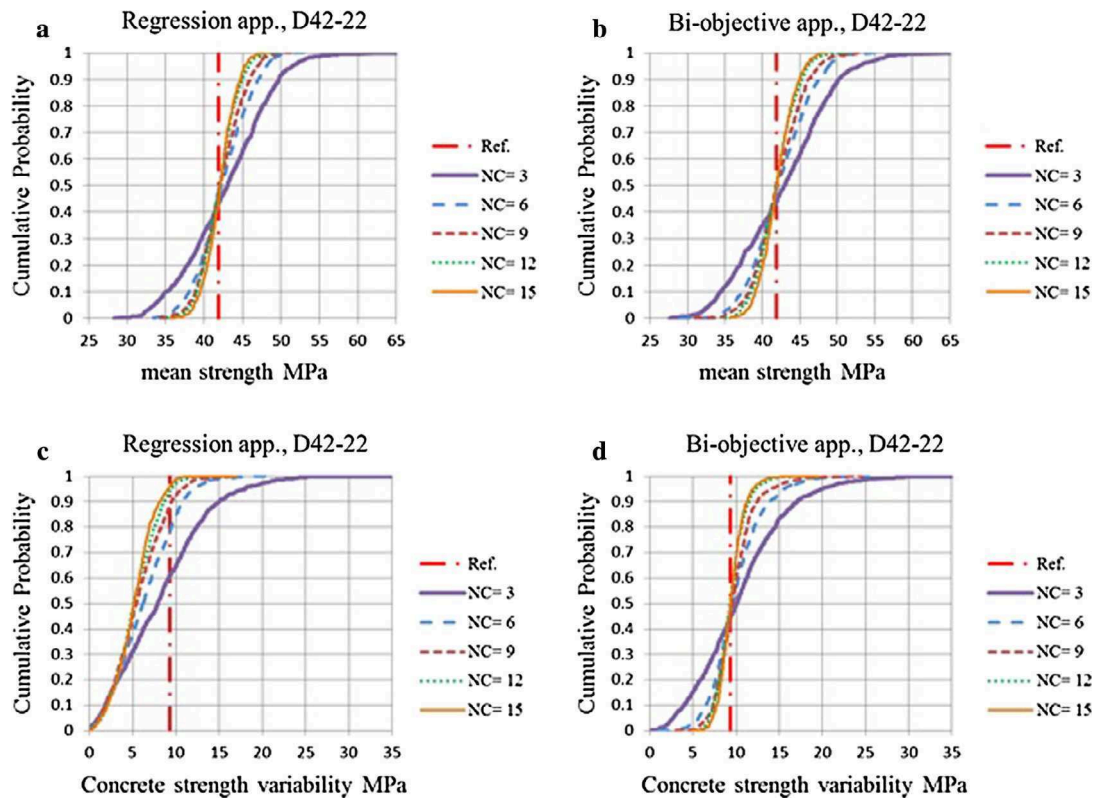


Fig. 3. CDF curves of concrete strength mean and variability estimation for concrete D42-22.

scatter). In an ideal case (no scatter), the CDF curve would coincide with the reference line, but the sources of uncertainty (model uncertainty, statistical uncertainty of sampling, measurement uncertainty, uncontrolled factors) always exist and as a result some scatter is unavoidable. However, as shown in Fig. 3a and b, the scatter decreases as NC increases. Moreover, the reduction in the scatter (i.e. improvement in reliability) appears remarkable for low values of NC while it becomes almost negligible (i.e. expensive without being effective) for higher NC values. The comparison of CDF curves of regression approach with those of bi-objective approach indicates that these two approaches are close in their reliability for assessing the mean strength.

However, regarding strength variability assessment presented in Fig. 3c and d, the result is different. For the regression approach, Fig. 3c, increasing NC reduces the scatter but does not guarantee the convergence towards the reference. The CDF curves show, even with large NC, a left bias (meaning general underestimation of strength variability). In contrast, the CDF curves corresponding to bi-objective approach, Fig. 3d, show a significant reduction in the scatter on both sides and a tendency to converge towards the true reference value as NC increases. This finding confirms that the common regression approach cannot be used to estimate the strength variability and that the bi-objective approach answers this need as previously shown by [7].

4.2. Analysis of the control factors using the risk curves

From previous section, it is clear that there is a relationship between the reliability of assessment and the sources of uncertainty so to quantify this relation and deepen the analysis the CDF curves are used to establish risk curves as illustrated in Section 3.2. The idea is for a certain interval of error (or accepted level of uncertainty) around the true reference value, what is the risk

(probability) to have an estimated value outside this target interval.

4.2.1. Analysis of risk curves corresponding to mean strength estimation

To construct the risk curves of mean strength estimation, the accepted uncertainty level was fixed ($U = \pm 10\%$) for the 17 datasets considered in this study. Fig. 4a illustrates the risk curves that belong to regression approach while curves corresponding to bi-objective approach are shown in Fig. 4b. Looking at these graphs, it is clear that the two approaches provide very similar results with an improvement in reliability (reduction in risk) as NC increases for all datasets. In accordance with the statistical aspects, the true concrete strength variability affects the scatter in estimated values [30–31] and as a result it plays as a controlling factor in the assessment reliability. Therefore, in the legend of Fig. 4, the 17 datasets were arranged from 1 to 17 according to the true strength variability (in terms of $CV(f_c)$). Hence, we expected that the dataset with the smallest variability (D67-11) produces the lowest risk values because the small strength variability leads to CDF curve having small scatter and as a result small risk (see Fig. 2). The risk curves for dataset D67-11 in Fig. 4a and b are in agreement with our expectation. However, the dataset with the largest variability (D46-33) does not provide the highest risk values, which is produced by the dataset D42-22. This indicates the presence of other control factors that have a significant role. From Table 1, it can be seen that D42-22 the within-test variability $CV_R = 8.84\%$ which is higher than the value corresponding to D46-33 ($CV_R = 4.84\%$). Hence beside the true strength variability, the within-test variability (or repeatability) also plays as a control factor.

The effects of these controlled factors (concrete variability, NDT within-test variability), in addition to uncontrolled factors, are implicitly shown in Fig. 4 where the contribution of each factor

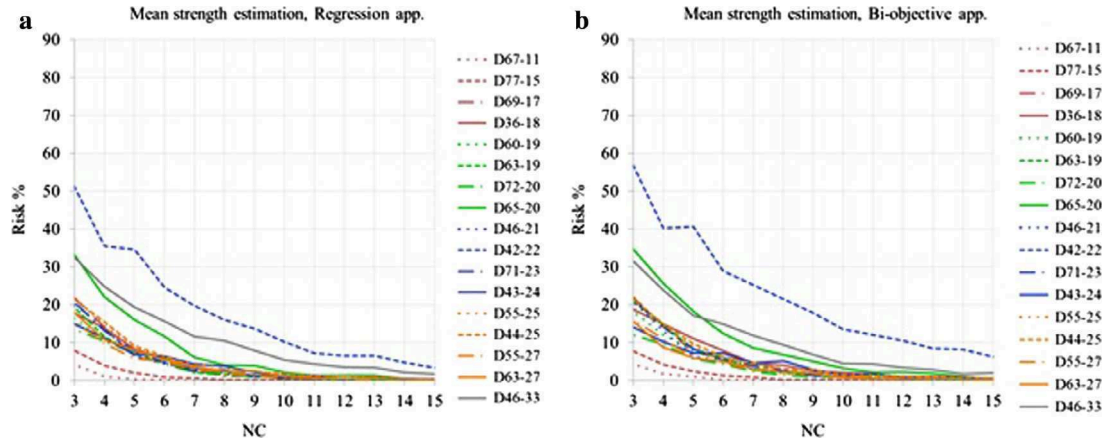


Fig. 4. Risk curves of all datasets: for mean strength estimation with $U = \pm 10\%$.

is difficult to analyze. Additional figures were plotted to get a clearer view upon this issue, by varying only one factor while the other is fixed as much as possible. Fig. 5a and b show four risk curves (corresponding to four datasets) selected from Fig. 4. The two datasets in each figure (D67-11 and D72-20 in Fig. 5a and D77-15 and D71-23 in Fig. 5b) have nearly the same within-test variability (i.e. CV_R). As a consequence, in each figure, the difference between the two risk curves illustrates the effect of concrete variability on the assessment reliability. It is obvious that, for the same value of NC, for the more variable dataset, the assessment of mean

strength is less reliable. However, the absolute difference between the risk values decreases as NC increases and becomes negligible when $NC > 9$.

Fig. 5c and d study the effect of within-test variability, CV_R , on the assessment reliability by comparing risk curves corresponding to four datasets respectively D72-20 and D65-20 in Fig. 5c and D60-19 and D63-19 in Fig. 5d. In each figure the two curves have nearly the same true concrete variability, $CV(f_c)$, and a close value of mean strength. From Fig. 5c one can deduce that, for the same NC, using bad measurement (i.e. larger CV_R) leads to less reliable

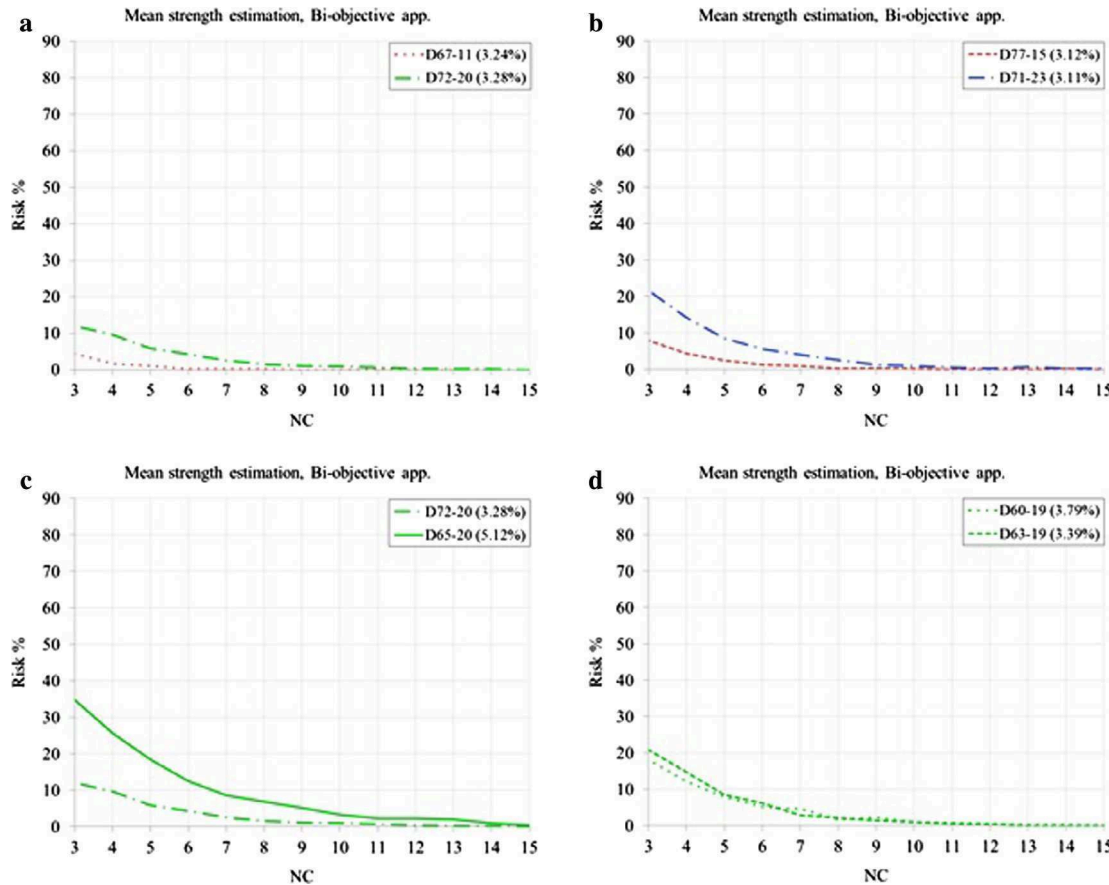


Fig. 5. Risk curves showing the effects of controlled factors on the assessment reliability of mean strength when $U = \pm 10\%$, for each dataset CV_R is indicated between parentheses.

assessment of mean strength. Fig. 5d shows that the curves for two cases with a similar concrete variability and similar within-test variability are very close.

4.2.2. Analysis of risk curves corresponding to strength variability estimation

Risk curves of strength variability estimation for an accepted uncertainty level ($U = \pm 25\%$) for the 17 datasets were shown in Fig. 6a and b. Fig. 6a illustrates the handicap of regression approach in estimating the strength variability since increasing NC offers no guarantee regarding the reduction of risk values. On the contrary, NC has significant effect on the risk curves with bi-objective approach, as presented in Fig. 6b. For the risk curves of strength

variability estimation, the implicit effects of the factors deduced in the case of mean strength (concrete variability, within-test variability) can also be noted.

To analyze these effects, several datasets selected from Fig. 6b were shown in Fig. 7. These datasets are the same as those analyzed in Fig. 5a–d. From Fig. 7a and b for same NC and comparable within-test variability CV_R , the assessment reliability is better for the concrete having more strength variability. It can be noted that this effect is adverse than that presented in Fig. 5a and b for the case of mean strength estimation; decreasing the concrete variability makes the mean strength assessment more reliable, but the concrete variability assessment less reliable. For the case of mean strength estimation, the behavior is rational because capturing

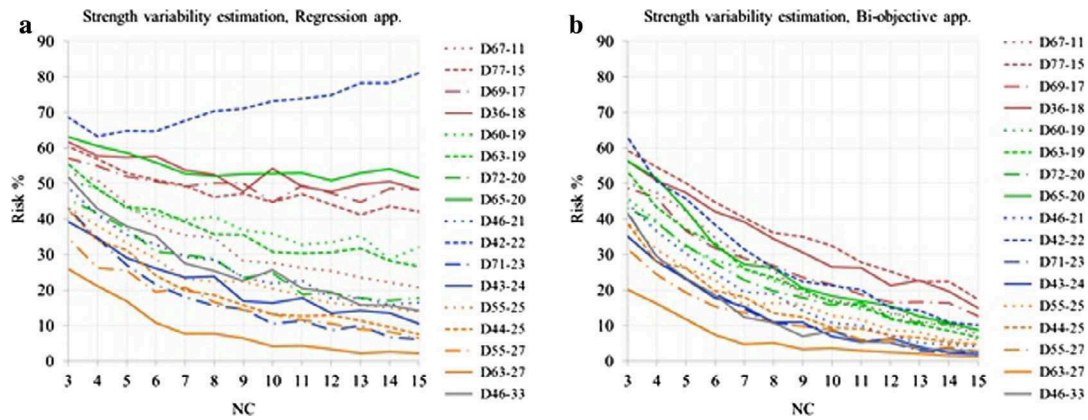


Fig. 6. Risk curves of all datasets: for strength variability estimation with $U = \pm 25\%$.

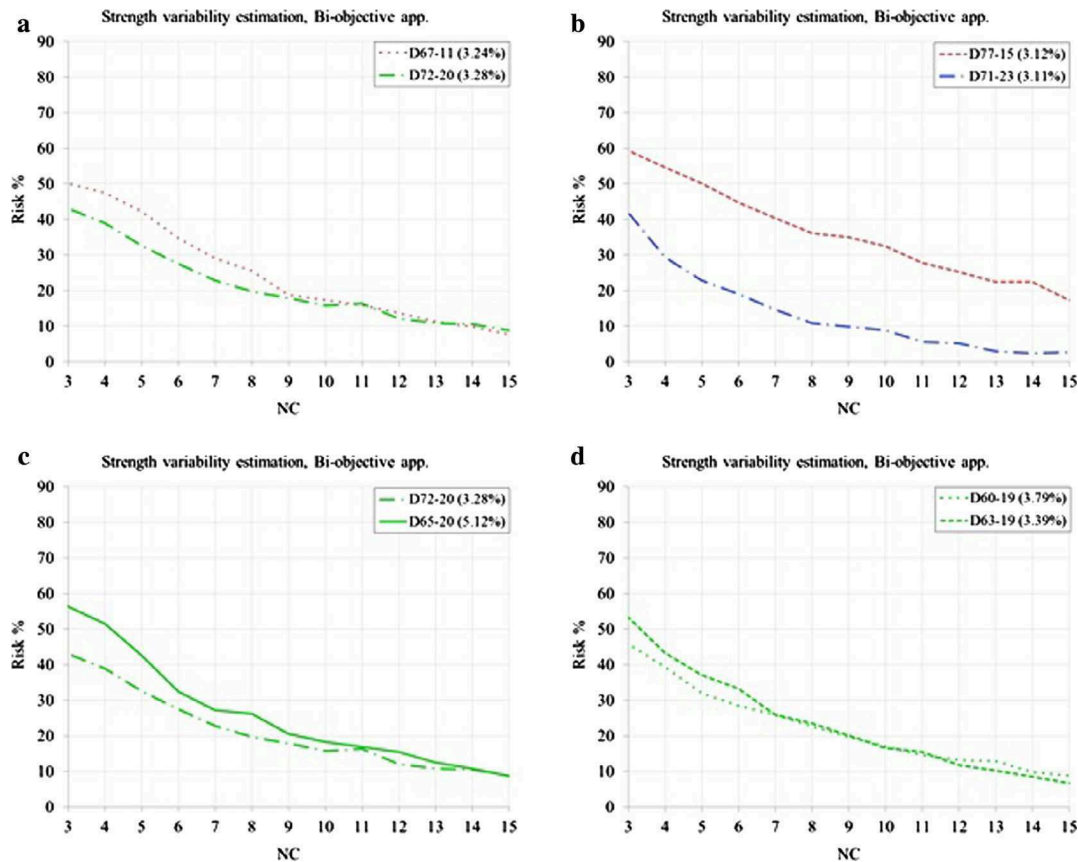


Fig. 7. Risk curves showing the effects of controlled factors on the assessment reliability of strength variability when $U = \pm 25\%$, for each dataset CV_R is indicated between parentheses.

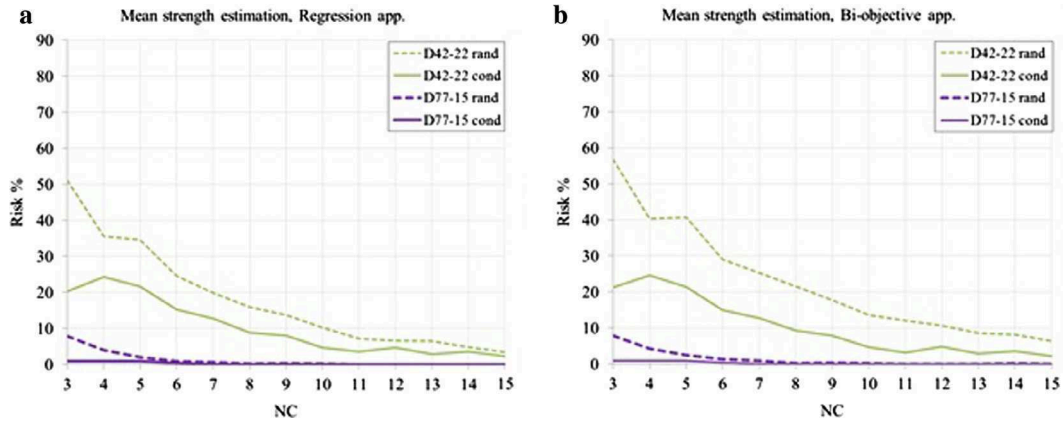


Fig. 8. Risk curves showing the effect of the way of selection the NC test locations on the assessment reliability of mean strength when $U = \pm 10\%$.

mean strength is easier in a less variable concrete. Whilst, for the case of strength variability estimation, the behavior (i.e. assessing the strength variability is easier in a more variable concrete) seems irrational. To understand the second behavior, it is necessary to bear in mind the fact that the risk curves in this study were derived from fixing a relative value for the accepted uncertainty level (i.e. fixing $U\%$ consequently the absolute value ($U \times T$) varies depending on T). However fixing an absolute value for the uncertainty level (i.e. using a fixed value instead of ($U \times T$) in Fig. 2) leads to reverse behavior i.e. the strength variability assessment becomes easier in a less variable concrete. Therefore the statement about the effect of true concrete variability on the reliability of assessing the concrete strength variability depends on the way in which the accepted uncertainty level is represented.

The effect of within-test variability on the assessment of strength variability is illustrated in Fig. 7c and d. This effect is qualitatively identical to that was shown for mean strength assessment, since a better (smaller) within-test variability decreases the risk values (Fig. 7c) while the risk curves for two similar concretes are very close to each other, Fig. 7d.

The comparison between risk curves plotted on Figs. 4 and 6 also shows that the convergence is slower for variability assessment than for mean strength assessment. Thus for the same values of uncertainty $U\%$ and prescribed risk, the NC value required for the assessment of strength variability will be larger than that required for the mean strength assessment.

4.3. Effect of the way of selection the NC test locations on the reliability of assessment

All above results were obtained from the random selection of the NC test locations. However, in this section the conditional selection process detailed in Section 3.3 was applied in order to see its effect on the assessment reliability. All steps of the methodology illustrated in Fig. 1 were followed with the only change regarding the conditional selection of the NC test locations. Therefore, after reanalyzing the 17 datasets considered in this study, series of 1000 values of \bar{f}_{cest} and a series of 1000 values of $s(f_{cest})$ were obtained for each dataset, each NC value and each approach. The risk values were calculated and used to build risk curves as previously mentioned.

To study the effect of conditional selection of NC test locations on the mean strength estimation, the risk curves corresponding to two datasets (respectively D42-22 and D77-15) that were the extreme cases on Fig. 4 were selected, and the two approaches (regression or bi-objective method) were used. The curves with random selection were redrawn on Fig. 8 beside the corresponding

risk curves built with the conditional selection. The comparison between random and conditional cases shows a significant reduction in the risk values (improvement in reliability) when using the conditional selection. The behavior is very similar for regression and bi-objective approaches. The magnitude of risk reduction is high for the case of highly initial risk values. However for all cases the amount of reduction decreases as NC increases.

Regarding the strength variability estimation, the risk curves resulted from the two options of NC selection (random and conditional) are plotted on Fig. 9 for two datasets that represented the extreme curves in Fig. 6b. Regression approach is not considered since it fails to address the strength variability. Fig. 9 shows a reduction in risk values (improvement in reliability) when NC locations are selected conditionally as compared with the values corresponding to the case of random selection. The magnitude of the reduction is larger for the dataset with high initial risk values.

Consequently, since conditional selection induces no additional cost and since it can improve the reliability of assessment for both estimated quantities (mean strength and strength variability), this process must be strongly recommended. It must be noted that conditional selection is fully compatible with in-situ concrete assessment: it only requires that the preliminary NDT investigation is carried out before the core location is effectively decided.

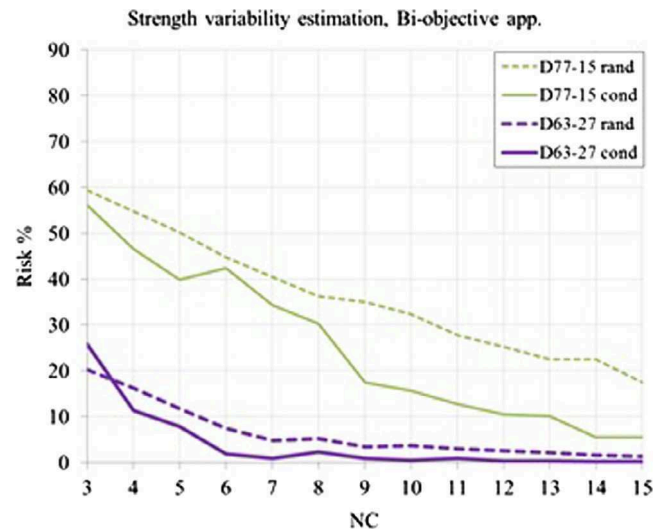


Fig. 9. Risk curves showing the effect of the way of selection the NC test locations on the assessment reliability of strength variability when $U = \pm 25\%$.

Table 2
Summary of the effects of factors considered in this study on the reliability of assessment.

Factor	Mean strength estimation		Strength variability estimation	
	Regression	Bi-objective	Regression	Bi-objective
Reducing the number of test locations used to identify the model between strength and rebound measurement NC	Worsening	Worsening	Approach provides unreliable assessment	Worsening
Reducing within-test variability of rebound measurements CV_R	Improvement	Improvement		Improvement
Reducing the true concrete strength variability $CV(f_c)$	Improvement	Improvement		Worsening
Using conditional selection for NC test locations instead of random selection	Improvement	Improvement		Improvement

Other factors exist that may improve the reliability of assessment. One of these factors is the number of test locations for NDT measurements (NR) or in statistician wording “sample size”. From statistical aspects the scatter of sampling distribution [30–31] of mean strength depends on the ratio σ/\sqrt{NR} , where σ is the standard deviation of concrete strength population. Therefore increasing the sample size leads to reduce one of the uncertainty sources (sampling uncertainty [21]) and, as a result, reduces the scatter of CDF curves. In this study all the results have been obtained with $NR = 60$ and larger NR may have some reduction effect on the risk values.

Finally, it is necessary to note that all risk values presented in this paper were corresponding to fixed values of accepted uncertainty level, ($U = \pm 10\%$) for mean strength estimation and ($U = \pm 25\%$) for strength variability estimation. Changing these values would of course lead to different risk values as illustrated in the Fig. 2 since widening the accepted uncertainty interval reduces the risk values and vice versa. However all effects illustrated in this study would keep unchanged.

5. Conclusions

Reliability of assessing the concrete strength using NDT measurements (rebound hammer in this study) and destructive tests was studied in this paper. Seventeen datasets that covered a wide range of true mean strength \bar{f}_c (36–77 MPa) and true strength variability (in terms of strength standard deviation $s(f_c)$ (6.4–17.4 MPa) or in terms of strength coefficient of variation $CV(f_c)$ (11–33%) were considered. Two estimated quantities (mean strength and strength variability) had been the target of the assessment using two model identification approaches: regression and bi-objective. The effects of several factors on the reliability of assessment were analyzed. These are: within-test variability of rebound measurements, true concrete strength variability, number of test locations used to identify the model between strength and rebound measurement NC (number of cores in the case of real structure), way of selection the NC test locations (random or conditional), and the model identification approach (regression or bi-objective). The effects of these factors on the reliability of assessment have been confirmed and are summarized in Table 2.

As shown in Table 2, the assessment of concrete strength variability using the regression approach provides unreliable assessment. Regarding the effect of NC , reducing NC leads to less reliable estimations of the mean strength and concrete strength variability whatever the model identification approach (regression or bi-objective). Whilst reducing the within-test variability of the rebound measurements leads to more reliable assessment. Regarding the true strength variability of the concrete under investigation, less variable concrete gives more reliable estimation of concrete mean strength. However the estimation of the concrete strength variability will be less reliable (this statement is based on considering relative uncertainty or error around the true value). Concerning the way of selection core locations, since conditional selection requires no additional cost and since it can improve the reliability of assessment for both estimated quantities (mean

strength and strength variability), this process must be strongly recommended.

The NC plays a major role since the increase of this factor always ensures a better reliability. To obtain assessed values of mean strength and strength variability having the same reliability, the NC value required for the assessment of strength variability should be larger.

Application of the bi-objective approach on datasets that covered a wide range of concrete mean strength and strength variability has confirmed the robustness of this approach and its ability in assessing concrete strength variability in addition to mean strength.

While playing an important role, the concrete intrinsic variability is usually not a degree of freedom during the investigation: it can be assessed, but not changed. On the contrary, the interest of minimizing the within-test variability has been confirmed. This issue will be addressed further by comparing the reliability of assessment with NDT methods having different values of within-test variability (typically rebound hammer versus ultrasonic velocity).

This analysis also paves the way towards practical rules for quantifying the relation between the number of cores NC and the reliability of the assessment, or for defining the practical number of cores required to reach a prescribed reliability.

Appendix A. The principle of bi-objective approach

Any investigation program with NDT technique (rebound hammer for example) provides a data of NC -pairs of (R, f_c) i.e. NC -test locations with the rebound measurement, R , and core strength, f_c , values corresponding to each test location. This data is used to identify a conversion model between concrete strength and the rebound number test results. Usual mathematical shapes of such models have two parameters [6,17]. For the case of linear model, the model formula is $f_{cest} = aR + b$. Analytically, two conditions are required in order to derive the values of the parameters a and b . For bi-objective approach, we consider our two objectives as the conditions for obtaining the unknown parameters, i.e. by ensuring that both mean strength and standard deviation are identical for real values and estimated ones:

$$\bar{f}_{cest} = \bar{f}_c \quad (A1)$$

$$s(f_{cest}) = s(f_c) \quad (A2)$$

where $\bar{f}_{cest}, \bar{f}_c$ are the estimated and measured (on cores) mean strength values respectively, while $s(f_{cest}), s(f_c)$ are the estimated strength variability and the variability calculated from the core strengths respectively.

Applying the first condition on a model $f_{cest} = aR + b$, we get:

$$\bar{f}_{cest} = a\bar{R} + b = \bar{f}_c \quad (A1')$$

while the second condition can be rewritten in the form:

$$s(f_{cest}) = s(f_c) \rightarrow s^2(f_{cest}) = s^2(f_c) \rightarrow var(f_{cest}) = var(f_c)$$

$$\text{var}(aR + b) = \text{var}(f_c) \quad (\text{A2}')$$

$$a^2 \text{var}(R) = \text{var}(f_c)$$

Consequently, the values of the unknown parameters are:

$$a = s(f_c)/s(R) \quad (\text{A3})$$

$$b = \bar{f}_c - a\bar{R} \quad (\text{A4})$$

where \bar{R} , $s(R)$ are respectively the mean and standard deviation values of R test results corresponding to the core locations (i.e. NC locations).

References

- [1] RILEM NDT-7, Recommendations for testing concrete by hardness methods, Tentative recommendation, 7-NDT committee – nondestructive testing, Mater. Struct. 10 (5) (1977) 313–316.
- [2] Pessiki SP (chair). In-place methods to estimate concrete strengths. ACI 228.1R-03 report; 2003.
- [3] IAEA International Atomic Energy Agency. Guidebook on non-destructive testing of concrete structures. Testing Training Course Series No. 17; 2002.
- [4] A. Masi, L. Chiauzzi, V. Manfredi, Criteria for identifying concrete homogeneous areas for the estimation of in-situ strength in RC buildings, Constr. Build. Mater. 121 (2016) 576–587.
- [5] EN 13791, Assessment of in situ compressive strength in structures and precast concrete, CEN, Brussels, 2007.
- [6] D. Breyse, Nondestructive evaluation of concrete strength: an historical review and new perspective by combining NDT methods, Constr. Build. Mater. 33 (21) (2012) 139–163.
- [7] M. Alwash, Z.M. Sbartai, D. Breyse, Non-destructive assessment of both mean strength and variability of concrete: a new bi-objective approach, Constr. Build. Mater. 113 (2016) 880–889.
- [8] Proverbio E, Venturi V. Reliability of nondestructive tests for on site concrete strength. 10th International Conf. on Durability of Building Materials and Components, Lyon, 17–20 April 2005.
- [9] M. Fischli, A. Moczko, Rebound hammer, in: D. Breyse (Ed.), Non-Destructive Assessment of Concrete Structures: Reliability and Limits of Single and Combined Techniques, State-of-the-Art Report of the RILEM Technical Committee TC 207-INR, Springer, 2012, pp. 101–110.
- [10] N.J. Carino, in: E.G. Nawy (Ed.), Concrete Construction Engineering Handbook, CRC press, Boca Raton, FL, Nawy, 1997. chapter 19, 19/1–68.
- [11] A. Brencich, G. Cassini, D. Pera, G. Riotta, Calibration and reliability of the rebound (Schmidt) hammer test, Civil Eng. Archit. 1 (3) (2013) 66–78.
- [12] R. Pucinotti, Reinforced concrete structure: nondestructive in situ strength assessment of concrete, Constr. Build. Mater. 75 (2015) 331–341.
- [13] G.F. Kheder, A two stage procedure for assessment of in situ concrete strength using combined non-destructive testing, Mater. Struct. 32 (1999) 410–417.
- [14] H. Qasrawi, Concrete strength by combined nondestructive methods simply and reliably predicted, Cem. Concr. Res. 30 (2000) 739–746.
- [15] V.M. Malhotra, N.J. Carino, Handbook on Non Destructive Testing of Concrete, CRC Press, 2004.
- [16] Ministry of railways (India). Guidelines on non-destructive testing of bridges BS – 103 August, 2009.
- [17] K. Szilágyi, A. Borosnyói, 50 years of experience with the Schmidt rebound hammer, Concrete Structures 10 (2009) 46–56.
- [18] FHWA. Guide to non-destructive testing of concrete, FHWA-SA-97-105, USDOT, Washington DC, 1997.
- [19] K. Szilágyi, A. Borosnyói, I. Zsigovics, Extensive statistical analysis of the variability of concrete rebound hardness based on a large database of 60 years experience, Constr. Build. Mater. 53 (2014) 333–347.
- [20] N.J. Carino, Nondestructive testing of concrete: history and challenges, SP 144–30, in: P.K. Mehta (Ed.), Concrete Technology – Past, Present and Future, ACI, Detroit, MI, 1994, pp. 623–678.
- [21] N. Pereira, X. Romao, Assessment of the concrete strength in existing buildings using finite population approach, Constr. Build. Mater. 110 (2016) 106–116.
- [22] ASTM C805–97. Standard test method for rebound number of hardened concrete. ASTM Standards, West Conshohocken; 1997.
- [23] J.H. Bungey, S.G. Millard, Testing of Concrete in Structures, 3rd edition., Chapman and Hall, 1996.
- [24] M. Alwash, D. Breyse, Z.M. Sbartai, Non-destructive strength evaluation of concrete: analysis of some key factors using synthetic simulations, Constr. Build. Mater. 99 (2015) 235–245.
- [25] K. Szilágyi, Rebound surface Hardness and Related Properties of Concrete (P.h.D. thesis), Budapest University of Technology and Economics, 2013.
- [26] J. Walkenbach, Excel VBA Programming for Dummies, John Wiley & Sons Inc., USA, 2013.
- [27] Breyse D et al. Comparing investigation approaches and NDT methodologies for concrete strength estimation: an international benchmark. NDT-CE conf. Berlin, 15–17 Sept. 2015.
- [28] M. Alwash, D. Breyse, Z. Sbartai, Using Monte-Carlo simulations to evaluate the efficiency of different strategies for nondestructive assessment of concrete strength, Mater. Struct. 50 (2017) 14p, <http://dx.doi.org/10.1617/s11527-016-0962-x>.
- [29] V. Pfister, A. Tundo, V. Luprano, Evaluation of concrete strength by means of ultrasonic waves: a method for the selection of coring position, Constr. Build. Mater. 61 (2014) 278–284.
- [30] D. Montgomery, G. Runger, Applied Statistics and Probability for Engineers, John Wiley & Sons Inc., USA, 2003.
- [31] B. Harding, C. Tremblay, D. Cousineau, Standard errors: a review and evaluation of standard error estimators using Monte Carlo simulations, Quant. Methods Psychol. 10 (2) (2014) 107–123.

Résumé

Pour évaluer la résistance mécanique du béton dans un ouvrage existant, la méthodologie courante combine des mesures non destructives (CND) comme le rebond ou/et la vitesse des ondes ultrasoniques avec la technique destructive (carottes) afin de produire une relation “modèle de conversion” entre la résistance mécanique et les mesures CND. Le modèle de conversion est utilisé pour estimer la valeur locale de résistance mécanique à chaque emplacement de test en utilisant la valeur CND correspondante. Ensuite, on calcule les estimations de la résistance moyenne et/ou de l'écart-type de la résistance (variabilité de la résistance du béton). Cependant, la fiabilité d'estimation est toujours discutable en raison des incertitudes associées à l'évaluation de la résistance basée sur les mesures CND. Pour améliorer la fiabilité, les incertitudes doivent être réduites en spécifiant et en contrôlant leurs facteurs d'influence. Par conséquent, l'objectif de cette thèse est d'analyser la méthodologie d'évaluation courante afin de fournir des recommandations pratiques qui peuvent améliorer la fiabilité de l'évaluation de la résistance in-situ du béton dans les ouvrages existantes par des tests non destructifs et des carottes. Pour ce but, un simulateur a été construit afin d'analyser les effets des facteurs les plus influents en utilisant une vaste campagne de données provenant de sources différentes (études in situ ou en laboratoire et données synthétiques générées). La première contribution de ce travail est le développement d'une nouvelle approche d'identification du modèle “bi-objectif” qui peut efficacement capturer la variabilité de la résistance mécanique en plus de la moyenne. Après avoir étudié l'effet du mode de sélection des emplacements pour les carottes, une méthode a été proposée pour sélectionner ces emplacements en fonction des mesures CND “sélection conditionnelle” qui améliore la qualité de l'évaluation sans coût supplémentaire. Une dernière innovation est l'établissement de courbes de risque qui quantifient la relation entre le nombre de carottes et la précision de l'estimation. Enfin, des recommandations ont été formulées afin de fournir des estimations plus fiables.

Mots-clés: résistance in-situ du béton, résistance mécanique moyenne, variabilité du béton, vitesse des ondes ultrasoniques, rebond, CND, combinaison, simulation, incertitude, stratégie d'évaluation, fiabilité d'évaluation

Abstract

To assess concrete strength in an existing structure, the current methodology combines nondestructive measurements (NDT) like rebound hammer or/and pulse velocity with destructive technique (cores) in order to implement a relationship “conversion model” between the compressive strength and NDT measurements. The conversion model is used to estimate the local strength value at each test location using the corresponding NDT value. Then the estimated mean strength and/or estimated strength standard deviation (concrete strength variability) values are calculated. However, the reliability of these estimated values is always a questionable issue because of the uncertainties associated with the strength assessment based upon NDT measurements. To improve the reliability, the uncertainties must be reduced by specifying and controlling their influencing factors. Therefore, the objective of this thesis is to analyze the current assessment methodology in order to provide practical recommendations that can improve the reliability of assessing the in-situ strength in existing concrete structures by nondestructive tests and cores. To this end, a simulator was built in order to analyze the effects of the most influencing factors using a large campaign of datasets from different sources (in-situ or laboratory studies, and generated synthetic data). The first contribution of this work is the development of a new model identification approach “bi-objective” that can efficiently capture the strength variability in addition to the mean strength. After studying the effect of the way of selection the core locations, a method was proposed to select these locations depending on the NDT measurements “conditional selection” that improves the quality of assessment without additional cost. A third innovation was the development of a procedure to identify the relation between the number of cores and the accuracy of the estimation. Finally recommendations were derived in order to provide more reliable estimated values.

Keywords : in-situ concrete strength, mean strength, concrete variability, ultrasonic pulse velocity, rebound hammer, NDT, combination, simulation, uncertainty, assessment strategy, assessment reliability