# CONFIDENCE IN EXPERIMENTAL DATA – SUMMARY FOR THE FIELD OF CONCRETE TESTING

*Alena Horská* *

Katedra betonových a zděných konstrukcí, Fakulta stavební,
České vysoké učení technické v Praze, Thákurova 7/2077, 166 29 Praha 6, Česká republika.
`alena.horska@fsv.cvut.cz`

## ABSTRAKT

Článek poskytuje celkové shrnutí požadavků na zkoušení betonu a zpracování experimentálně získaných dat z hlediska reprezentativnosti výsledků. Je zde řešena oblast zahrnující proces výběru vzorků, vlastního experimentu a získávání materiálových parametrů betonu, využití těchto parametrů jakožto vstupů do následné analýzy chování betonových prvků a získání odezvy těchto prvků na určité zatěžovací podmínky. Ve všech těchto uvedených procesech dochází k vnášení chyb a nejistot do výsledků analýzy chování prvku. Článek popisuje možné zdroje těchto chyb a nejistot. Důležitost množství testovaných vzorků je demonstrována na jednoduchém ilustrativním příkladu.

## KLÍČOVÁ SLOVA

Experimentální zkoušení • Počet vzorků • Zkoušení betonu

## ABSTRACT

The article provides an overall summary of requirements on experimental concrete testing and processing the obtained data, in terms of representativeness of the results. The examined process includes sampling, experimental testing and obtaining material parameters, their following use as an input into an analysis, and simulating the structural response to a certain loading conditions. All of the stages of this process bring errors and uncertainties into the results of the analysis. The origin of these errors and uncertainties is described in the article. The importance of the sample size is demonstrated on a simple illustrative example.

## KEYWORDS

Experimental testing • Sample size • Concrete testing

## 1. INTRODUCTION

In engineering, experimental testing plays a crucial part in research. Very often experimentally obtained data are used as an input for further analysis. Very common example of that in the area of structural engineering is experimental testing of material properties that are then used as an input into FEM simulation in order to observe structural behaviour under certain loading conditions, as shown in Figure 1.

Satisfactory representativeness and factual correctness of these input data are essential, for no matter how precise the follow-up analysis is, it cannot provide correct results if the input data are incorrect or not representative enough.

This article provides a summary of requirements on experimentally obtained data, their correct statistical evaluation and interpretation, and points out possible errors that can occur in evaluation of experimental data in the field of concrete testing.

## 2. STANDARD APPROACH TO SAFE DESIGN

The Eurocode approaches the issue of data representativeness in a very unprecise but conventionally safe way. The respective part of the code can be found in EN 1990.

The Eurocode allows small sample sizes when determining material parameters experimentally. However, it realizes the small sample sizes lead to large standard deviations and, therefore, a large range of possible results. This results in a high uncertainty in obtained material parameters.

The Eurocode also provides methods of structural analysis that uses these uncertain material parameters as an input. The methods of analysis very often idealize the real physical nature of the structural behaviour and simplify the matter. These simplifications lead to a distortion of the results.

The combination of the uncertain material parameters and the result distortion due to simplifications in the follow-up structural analysis cause that the structural response obtained by calculations according to the Eurocode can be far from the real behaviour. For that, the Eurocode provides compensation in a form of safety factors. The material parameters are reduced by safety factors so that despite the error caused by the uncertainty and simplifications is large, the unprecise value of the material parameter is reduced enough to be proclaimed safe anyway.

Such approach is useful for practical applications when the financial resources for experimental testing of a structural material are limited, the design of the structure is limited by time, and the structures have a common use and do not require any special design procedures. In this case, the use of the Eurocode is completely sufficient.

---

* Školitel: prof. Ing. Alena Kohoutková, CSc., FEng.

However, in research, the goal is to be able to describe the behaviour of a material/structure with the utmost precision. For that, the uncertain, simplified and safety factor-distorted values given by the Eurocode are not sufficient enough, and more precise values of material parameters are required.

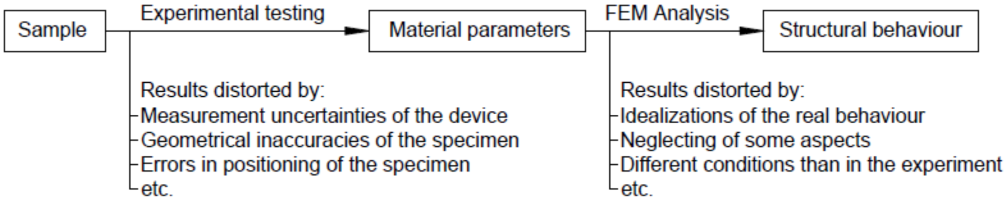

Figure 1: *Scheme of experimentally obtaining data and processing.*

## 3. UNCERTAINTIES IN THE ANALYSIS BASED ON EXPERIMENTAL DATA

When scientifically analysing structural behaviour of a concrete element, the scheme plotted in Figure 1 is followed. A sample is chosen, and is subjected to experimental testing in order to obtain its material properties. Then, these properties are used as an input into FEM analysis that simulates loading conditions and provides structural response. Each of these steps has its own specifics. This chapter describes these specifics, and points out the pitfalls of each step, focusing on the issue of uncertainties of the input and output data.

### 3.1. Sample

In experimental testing, it is desired for the sample (i.e. set of finite number of specimens; subset of population) not to refer only to itself, but to represent the whole population as accurately as possible (the word population in this context is used for the infinite number of specimens, i.e. for all of them). By that, it is meant that the mean value of the tested material characteristic (e.g. compressive strength of concrete) should be, ideally, identical to the mean value of the characteristic of the population. However, it is never possible to reach this ideal case as the identity of the sample and the population could be proven only if the sample size was equal to the population. Therefore, the identity of the sample and population never can be verified with 100% reliability. For that, the identity of the sample and population can be declared with a certain confidence level expressed as a percentage, less than 100 %. This is one of the major uncertainties that need to be noticed, and its effect needs to be incorporated when using such data in further analysis. The probability that the actual structural response is identical to the results of FEM modelling is proportional to the probability that the sample material parameters are identical to the parameters of the population.

Notice that at this point, because of what was described in the paragraph above, we start using probabilistic approach in this field. All of the values obtained anywhere in the process of structural analysis have their confidence level, which is a probability of their occurrence. When testing material properties experimentally, it is never possible to reach 100% confidence level in any of the values obtained.

The confidence level is strongly influenced by the sample size (i.e. number of specimens in the sample). It is obvious that increasing sample size has a positive effect on the precision, and, therefore, the confidence level increases with increasing number of specimens. However, this qualitative description of this dependence between sample size and result accuracy is not sufficient. It is necessary to quantify it in order to know how reliable the results of the analysis are. The quantification of the influence of sample size on reliability of results is demonstrated on a simple numerical experiment in the fourth section of this paper.

Furthermore, it is important to realize that even when the concrete sample is correctly picked and its confidence level is determined, the results of any follow-up analysis can be proclaimed valid for the whole population only if the manufacturing process of the concrete, or any other structural material, is identical to the sample manufacturing.

### 3.2. Experimental Testing of Material Parameters

The experimental testing itself is a rich source of errors, that influence obtained values of material parameters. The origin of them can be either instrumental or human induced.

Instrumental errors are caused by measurement inaccuracies of the device. From that, it is obvious, that the errors are strongly dependent on the measurement method and technique. The instrumental errors are unique for every measurement device, and should be declared in the manual of the device.

The human induced errors are caused by imperfections in human senses. On an example of the compressive test on concrete cubes, these errors might have a form of eccentric placement of the specimen into the device, inaccurate reading of the dimensions when measuring the specimen, or defects caused to the specimen during manipulation with it.

The errors mentioned above cannot be entirely avoided. Therefore, the errors need to be considered when evaluating the experimental outcomes. As the possible errors in the experimental testing are multiple, their cumulating should be calculated. Direct calculation of error cumulating can be done in case the experimental outputs are analytically expressed as a

---

* Školitel: prof. Ing. Alena Kohoutková, CSc., FEng.

function based on the inputs. The resulting error of multiple cumulated errors is a sum of partial derivations of that function by the inputs.

This method, however, has two major issues – it cannot be used in cases when the inputs are co-dependent, and when the relationship between the inputs and outputs does not have an analytical expression as a function. In one or the other case, the stochastic methods are the only option to obtain the dependence between the outputs and inputs that incorporates the errors of the measurement. Unfortunately, using such stochastic approach disables us to distinguish the precise solution and the errors, as it provides the outcomes including the errors without knowing their contribution to the results.

### 3.3. FEM Analysis and Structural Response

An analysis based on material parameters obtained experimentally can have several forms, depending on the simulated experiment and demands on the results. FEM analysis provides very accurate simulation of the experimentally tested problem when used correctly. However, it can also be a source of errors and uncertainties of the final outcome.

In order to save computational time, many variables are often neglected, either by the analytical method itself, or by the person conducting the analysis. Before anyone decides to neglect some inputs into the analysis, the sensitivity check should be run. Sensitivity check provides information about how much influence does a specific input variable have on the output. Without sensitivity check, no input neglecting should be done.

The analytical methods that connect the geometrical and material inputs with the structural response are more or less idealized. The idealization is not a choice, as the actual physical laws cannot be simulated with total precision for their complexity. The idealization distorts the results and for that it should be always on the safe side, providing the worse results than the actual, not the better ones. Quantification of the error caused by idealization is a major issue that cannot be solved in any simple way; it can be only projected to the results as a decrease of their confidence level.

## 4. EXAMPLES DEMONSTRATING SAMPLE SIZE SENSITIVITY

This chapter provides a simple example of how sample size influences the results obtained experimentally.

Let us have a population of concrete elements. To make the example easy to quantify and interpret, the population in this case does not have a count of infinity but only 400 pieces. With this finite number of pieces, we are able to calculate the exact statistic of this whole population, so that we can compare it to a statistic obtained from a sample picked up from this population. This allows us to see the difference between the results obtained on the sample and population.

In this example, we examine the values of compressive strength of the concrete elements. Since this is only an hypothetical example, the compressive strenght values of all 400 concrete elements in the population were not really measured on real concrete cubes. The strength values were randomly generated with normal distribution, having set the initial mean value at 40 MPa and standard deviation at 5 MPa. After generation of the values, the actual mean and standard deviation differ slightly from the inital generative values because the number of elements is finite. The values describing the population distribution are shown in Table 1.

Table 1: *Population of concrete elements tested for compressive strength*

| population of 400 pcs | | |
|---|---|---|
| | **values for generation** | **generated values** |
| **$f_{c,mean,\ pop}$ [MPa]** | 40.0 | 40.6 |
| **$\sigma_{pop}$ [MPa]** | 5.0 | 4.8 |

This example compares the accuracy of results for two different sample sizes: a) sample size $n$ = 3 specimens; b) sample size $n$ = 10 specimens. The examined feature of the specimens is compressive strength $f_c$.

In the case a), 10 samples were randomly picked from the population, each containing 3 specimens – see Table 2. Each specimen has its value of compressive strength $f_{c,specimen}$. For each sample, the mean compressive strength $f_{c,mean,spl}$ was calculated, obtaining values between 34.0 MPa and 45.0 MPa. Then, the difference between the mean values of the sample and the actual mean of the whole population $\Delta f_{c,mean}$ was calculated - see Equation (1).

$$\Delta f_{c,mean} = \left| f_{c,mean,spl} - f_{c,mean,pop} \right| \quad (1)$$

This difference in means $\Delta f_{c,mean}$ is a simple expression of the error caused by having the sample of smaller size than the population. The average of this error $\varnothing\Delta f_{c,mean}$ for ten samples was calculated, obtaining the value of 2.1 MPa.

Table 2: *Strength values of 10 samples; each sample containing 3 specimens*

| 3 specimens in 1 sample (strength in [MPa]) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **sample no.** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **$f_{c,specimen}$** | 43 | 34 | 38 | 38 | 44 | 49 | 49 | 47 | 35 | 35 |
| | 35 | 35 | 35 | 44 | 42 | 46 | 31 | 35 | 44 | 41 |
| | 41 | 33 | 45 | 43 | 39 | 40 | 39 | 44 | 43 | 35 |
| **$f_{c,mean,spl}$** | 39.7 | 34.0 | 39.3 | 41.7 | 41.7 | 45.0 | 39.7 | 42.0 | 40.7 | 37.0 |
| **$\Delta f_{c,mean}$** | 0.9 | 6.6 | 1.2 | 1.1 | 1.1 | 4.4 | 0.9 | 1.4 | 0.1 | 3.6 |
| **$\varnothing\Delta f_{c,mean}$** | 2.1 | | | | | | | | | |

---

* Školitel: prof. Ing. Alena Kohoutková, CSc., FEng.

The same procedure was conducted for the case b). Again, there were 10 samples picked randomly from the 400-piece population of compressive strength values. However, in this case, each sample contains 20 specimens - see Table 3. Processing the strength values of the specimen, the average error $ø\Delta f_{c,mean}$ expressed as a difference between the sample and population mean has the value of 0.6 MPa.

Table 3: *Strength values of 10 samples; each sample containing 20 specimens*

| 20 specimens in 1 sample (strength in [MPa]) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **sample no.** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **$f_{c,specimen}$** | 41 | 43 | 39 | 45 | 40 | 39 | 32 | 44 | 43 | 34 |
| | 41 | 49 | 35 | 33 | 32 | 37 | 44 | 31 | 39 | 45 |
| | 44 | 35 | 44 | 44 | 33 | 42 | 32 | 40 | 49 | 43 |
| | 40 | 35 | 40 | 32 | 44 | 33 | 44 | 34 | 49 | 40 |
| | 39 | 36 | 44 | 42 | 42 | 42 | 40 | 37 | 44 | 49 |
| | 48 | 41 | 31 | 50 | 39 | 52 | 36 | 44 | 43 | 38 |
| | 39 | 42 | 54 | 47 | 35 | 54 | 36 | 38 | 40 | 37 |
| | 39 | 39 | 43 | 38 | 44 | 43 | 49 | 34 | 39 | 49 |
| | 41 | 40 | 36 | 33 | 39 | 45 | 43 | 38 | 40 | 42 |
| | 42 | 36 | 33 | 42 | 44 | 35 | 44 | 43 | 38 | 45 |
| | 41 | 45 | 45 | 41 | 43 | 49 | 45 | 33 | 36 | 40 |
| | 41 | 42 | 45 | 46 | 45 | 36 | 41 | 41 | 31 | 36 |
| | 39 | 49 | 45 | 37 | 40 | 49 | 40 | 48 | 41 | 40 |
| | 41 | 52 | 32 | 32 | 49 | 41 | 44 | 34 | 48 | 40 |
| | 30 | 36 | 36 | 34 | 33 | 38 | 39 | 35 | 46 | 38 |
| | 41 | 37 | 43 | 36 | 43 | 42 | 41 | 42 | 36 | 40 |
| | 41 | 40 | 52 | 32 | 41 | 40 | 39 | 35 | 45 | 43 |
| | 39 | 39 | 43 | 52 | 34 | 49 | 42 | 52 | 32 | 40 |
| | 44 | 41 | 45 | 41 | 32 | 38 | 42 | 38 | 36 | 44 |
| | 36 | 39 | 42 | 48 | 36 | 49 | 36 | 52 | 38 | 39 |
| **$f_{c,mean,spl}$** | 40.4 | 40.8 | 41.4 | 40.3 | 39.4 | 42.7 | 40.5 | 39.7 | 40.7 | 41.1 |
| **$\Delta f_{c,mean}$** | 0.2 | 0.2 | 0.8 | 0.3 | 1.2 | 2.1 | 0.1 | 0.9 | 0.1 | 0.5 |
| **$ø\Delta f_{c,mean}$** | 0.6 | | | | | | | | | |

This simple demonstration provides a comparison of possible errors rising up from number of specimens. In this particular case, the error for the three-piece sample was 2.1 MPa and the error for the twenty-piece sample was 0.6 MPa, which makes the twenty-piece sample 3.5 times more accurate than the three-piece sample.

Of course, these particular values cannot be taken as fixed errors in such cases. The randomly picked values for the samples could have been picked different, which would change the mean values within the samples, which would affect the final error value, leading to a different result. But generally, the error is in vast majority of cases smaller for the larger sample sizes. There is always a chance that the smaller sample would provide more precise mean values than the larger sample, but the probability of that happening is very low, and gets lower with increasing difference between the sample sizes.

## 5. CONCLUSIONS

Experimental testing of material parameters of concrete and use of those parameters for structural analysis is burdened with many errors and uncertainties. These errors and uncertainties are brought into the data processing in each step, from sampling to simulating structural response.

The ways to deal with the issue of uncertainties and errors differ depending on whether the purpose of the experimental testing is practical or scientifical. For practical design of structures, standardization provides methods based on safety factors that compensate the errors and uncertainties in a conservative way, distorting the values of material parameters like strength to have them safely low.

In sciestific research, it is desired to have the data as precise as possible, with no distortions introduced by the safety factors. As the 100% precision can never be reached and, therefore, it is impossible to take one value as the only possible result, it is necessary to determine the range of possible result values, and, ideally, to assign a probability of occurrence to each of these values. For that, the knowledge of all idividual errors and uncertainties is crucial as well as the ability to evaluate them.

## ACKNOWLEDGEMENTS

The financial support of student grant project of CTU no. SGS19/034/OHK1/1T/11 is gratefully acknowledged.

# References

ČSN EN 1990 ed. 2. *Eurokód: Zásady navrhování konstrukcí.* Úřad pro technickou normalizaci, metrologii a státní zkušebnictví, 2011. 100 p.

Kelliher, D.; Sutton-Swaby, K. Stochastic representation of blast load damage in a reinforced concrete building. *Structural Safety* **2012**, *34*, 407–417.

Portilla, J. Probability and Statistics for Business and Data Science. Udemy. https://www.udemy.com/probability-and-statistics-for-business-and-data-science/learn/v4/content.