Variability case study based on in-situ rebound hardness testing of concrete
Part 1. Statistical analysis of inherent variability parameters

ADORJÁN BOROSNÝÓI • Assoc. Prof., BME, Dept. of Construction Materials and Engineering Geology • borosnyoi.adorjan@epito.bme.hu

Received: 28. 08. 2014. • Érkezett: 2014. 08. 28. • http://dx.doi.org/10.14382/epitoanyag.jsbcm.2014.16

Abstract
Statistical analysis has been performed on rebound index data collected in-situ at the bottom surface of a concrete floor over 130 m² of testing area. Different measures of inherent variability were calculated. Correlations between inherent variability measures were studied. It has been shown that skewed distributions can be fit with the best goodness of fit for the vast majority of the local statistical parameters studied. Application of the Peirce’s criterion filtering was demonstrated. Inferior local normality of the distribution of rebound index was found for the in-situ rebound hardness test results.
Keywords: Structural concrete; Nondestructive testing (NDT); Hardness; Rebound index; Inherent variability

1. Introduction
Numerous non-destructive testing (NDT) methods are available to estimate the strength of structural concrete [1-6]. Surface hardness testing is a long established NDT method for concrete structures [7,8]. Nowadays, the rebound hammer is the surface hardness testing device for concrete of the most widespread use [9-11].

Structural concrete is a multiple-level heterogeneous composite material [12,13]. On a macro-level, its heterogeneity is derived from concrete construction practice, i.e. random differences of compaction and curing at different locations, possible non-uniform supply of material and structural effects (reinforcement, changes in dimension or shape, connecting structural elements). On a meso-level, its heterogeneity is derived from concrete composition, i.e. type and grading of coarse aggregate and the paste content of fresh concrete. On a micro-level, its heterogeneity is derived from hydration performance that is governed by the type, particle size, hydraulic activity and amount of cement as well as that of supplementary cementing materials, the w/c-ratio, the type and amount of admixtures and curing compounds and the method and effectiveness of curing. On a nano-level, its heterogeneity is derived from the petrographic/chemical characteristics of fine aggregate and that of non-hydraulic supplementary materials and the molecular structure of hydrated cement paste.

This multiple-level heterogeneity results both inherent (local) variability and spatial (regional) variability of the performance properties. Local variability can be analysed by conventional mathematical statistical methods. Analysis of regional variability needs spatial models – that are typically used in geostatistics.

2. Scope of the studies
The present series of papers introduces path finding, comparative statistical analyses for the inherent variability and the spatial variability parameters calculated for in-situ rebound hammer tests performed on structural concrete.

In the first part, inherent variability parameters are analysed corresponding to the independent test areas in terms of statistical location parameters (mean, median, mode), statistical variance parameters (standard deviation, coefficient of variation), statistical dispersion parameters (range, relative range, studentized range) and normality parameters (skewness, excess kurtosis, W-statistic by Shapiro and Wilk, extreme studentized deviate by Peirce’s criterion).

Aim of the present studies is to demonstrate the statistical performance of rebound hardness for in-situ testing of concrete and to add to the better understanding of sampling and modelling by rebound hardness testing.

3. Experimental
The bottom surface of the top concrete slab of a framed, monolithic, subsoil concrete tunnel was studied, with dimensions of 25.0 m × 7.5 m and a thickness of 0.48 m [14]. The measuring region on the bottom surface was 22.0 m × 6.0 m. A total number of 42 test areas were selected for Schmidt rebound hammer testing. N-type original Schmidt rebound hammer was used. Eleven individual rebound index readings were recorded at each test area. The measurements were performed by the same operator.

4. Robustness of data for test areas
Random occurrence of outliers is expected in the test area recordings of in-situ rebound hammer tests due to the uneven surface texture and local, near-surface heterogeneity of structural concrete. If the operator uses mechanical rebound hammer for the measurements (as was the case for the present studies), any common rule of thumb may be applied for the immediate rejection of extreme outlier values from the data recorded. The filtering may be automatic or can be carried out during data analysis for rebound hammers recording digitally. Technical literature does not confirm, but as well, does not reject the hypothesis of normality for the probability distribution of individual rebound index readings.
for a test area. Therefore, it may be reasonable to select and run an appropriate extreme studentized deviate (ESD) test for normality to identify any (single or multiple) outlier in the data set of test areas [15]. Peirce’s criterion is a rigorous method based on probability theory that can be used to identify outliers in a sample. Details of the method are given in [16]. By Peirce’s criterion the maximum deviation from the mean is calculated and compared with tabulated values. The process is relatively simple and can be converted easily into an algorithm. If the maximum deviation of interest is greater than the tabulated value, then the suspicious value is removed as outlier. If no observation exceeds the tabulated value, then it cannot be claimed that there is a statistical outlier. The method assumes normal probability distribution of the data.

The extreme studentized deviate test by Peirce’s criterion was run for the 42 independent test areas. It resulted that at 11 test areas a total number of 14 individual rebound index readings can be considered as statistical outlier. It means 3% of the total number of 462 individual rebound index readings. It can be concluded that the data set is robust enough and the recording index does not change by location within the same structure or not. The 11-bin frequency histograms corresponding to the 42 independent test areas are indicated in Fig. 2. On a simple visual inspection it can be postulated that the normality of the observations is expected to be weak. Majority of the observations suggest platykurtic, almost uniform distribution.

There are about 40 normality tests available in the technical literature [18], however, the most common normality test procedures of statistical analyses are the Shapiro-Wilk test, the Kolmogorov-Smirnov test, the Anderson-Darling test and the Lilliefors test. It is demonstrated in the technical literature [19] that the Shapiro-Wilk test is the most powerful normality test from the above four, therefore, the Shapiro-Wilk normality test was run for the present analysis. Details of the method can be found in [20,21].

According to the null hypothesis of the Shapiro-Wilk normality test, the population has normal probability distribution. The calculation of the so called W-statistic by the method results the probability (p-value) that corresponds to the null hypothesis and one can see at which significance level the assumption that the population has normal probability distribution can be accepted. If the p-value is smaller than a preliminarily chosen significance level, then the null hypothesis is rejected; otherwise the null hypothesis is accepted at the chosen significance level.

It was realized during the present study that the hypothesis of normality can be generally accepted at a very low significance level for the 42 independent test areas analysed separately. The mean for the p-value is 40.0%. The p-value is smaller than 60.5% for 75% of the test areas. The p-value is smaller than 76.2% for 95% of the test areas. The p-value is smaller than 85.0% for 99% of the test areas. The maximum p-value is found to be 90.2%. Frequency histograms for the p-value and the W-statistic are shown in Fig. 3.

5. Normality of data for test areas

The overall empirical frequency histogram of the individual rebound indices (no separation by location) is indicated in Fig. 1 together with the best goodness of fit Fisher-Tippett (Generalized extreme value, GEV) probability density function. Considering the rebound hammer test, it can be assumed that the rebound index reading sets of separate test areas are independent and identically distributed (i.i.d.) random variables since it can be supposed that the probability distribution of the rebound index does not change by location within the same structure and the separate test areas can be considered to be mutually independent [17]. Based on these assumptions, the central limit theorem applies for the rebound hammer test; i.e. the probability distribution of the resultant of the rebound index reading sets of separate test areas (each with finite mean and finite variance) approaches the normal distribution if sufficiently large number of the i.i.d. random variables is available. Fig. 1 also indicates fitted normal distribution for comparison. It can be realised by a simple visual inspection that the resultant of the 42 independent test areas apparently approaches rather well to the normal distribution. For the 426 individual rebound indices, the global mean value is 52.6, the global median value is 53, the global mode value is 53, the global skewness is 0.11 and the global excess kurtosis is –0.11. It should be noted, however, that the statistical goodness of fit for the normal distribution is fairly less than that of the Fisher-Tippett (Generalized extreme value, GEV) distribution. It was shown by Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared goodness of fit tests that at least twenty different types of probability distributions (e.g. Pearson V, Gamma, Beta, Lognormal, Fréchet and even uniform distributions) provide better goodness of fit than the normal distribution, apart from the overall best fit Fisher-Tippett distribution. Therefore, it can not be concluded that statistically the normal probability distribution applies for the present population of rebound index readings. From an engineering point of view, however, the normal probability seems to be suitable as a rough estimate.
Further statistical information is obtained from the calculation of the skewness and excess kurtosis. Frequency histograms are shown in Fig. 3. The mean for the skewness is 0.068 that would not indicate pronounced deviation from normality on its own, however, a definite bimodality is visible on the frequency histogram that needs further considerations. An apparent mode at about –0.4 and at about 0.5 is seen. The sum of the 42 individual skewness values may indicate, if any, positive or negative skewness tendencies. The sum is 2.87 that reveals a stronger positive skewness tendency. The mean for the excess kurtosis is –0.19, but the median is more meaningful for the presented positively skewed frequency histogram, being –0.38. The sum of the 42 individual excess kurtosis values is –8.28. The findings confirm that platykurtic distribution governs the observations.

It may be assumed that removal of the 14 statistical outlier individual rebound index values that were determined with the extreme studentized deviate test by Peirce’s criterion would help reducing the deviation from normality. Therefore, skewness and excess kurtosis values were calculated for the 42 independent test areas after the removal of the suspected outliers (indicated in red colour in Fig. 2). The removal resulted in improvement in skewness performance: the mean is shifted closer to zero, to the value of 0.006 and the sum of the skewness values is reduced to 0.272. Surprisingly, the excess kurtosis performance became worse: the mean is shifted farther from zero, to the value of –0.59 and the sum of the excess kurtosis values is increased to –25.01.

6. Statistical location parameters at test areas

The mean, the median and the mode values were calculated for the 42 independent test areas and were published earlier [14]. Rather close correlation of the mean, median and mode values can be supposed for the present observations. The influence of the removal of the 14 outlier individual rebound indices selected by Peirce’s criterion was also tested during the present analyses. It was realized that the application of Peirce’s criterion does not result considerable change in the frequency histograms of the mean, median or mode values.

7. Statistical variance parameters at test areas

The standard deviation and the coefficient of variation values were calculated for the 42 independent test areas. Frequency histograms are shown in Fig. 4 together with the best goodness of fit Fisher-Tippett (Generalized extreme value, GEV) probability
density functions. Positive skewness is observable for both parameters, in accordance with literature data [17]. The influence of the removal of the 14 outlier individual rebound indices selected by Peirce’s criterion was again tested. Outlier removal was resulted in considerable loss in the peakedness of both histograms, becoming more uniform, however, the mean of the two statistical variance parameters did not change considerably.

4. Statistical analysis of the statistical variance parameters

The range, the relative range and the studentized range values were calculated for the 42 independent test areas. Frequency histograms are shown in Fig. 5 together with the best goodness of fit probability density functions. Best goodness of fit of Dagum distribution was found for the range values and Wakeby distribution for the relative range and the studentized range values. The observed positive skewness in the distribution of the range values is in accordance with literature data [17]. The influence of the removal of the 14 outlier individual rebound indices selected by Peirce’s criterion was also tested. Outlier removal was resulted in a shift toward smaller values of the mean for the range, the relative range and the studentized range and the skewness and peakedness of the frequency histograms have been changed as well. It seems that the extreme studentized deviate test by Peirce’s criterion has the strongest influence on the statistical dispersion parameters, as it was expected.

Fig. 4. Statistical analysis of the statistical variance parameters

a) standard deviation ($s_{R,j}$) for test areas

b) coefficient of variation ($V_{R,j}$) for test areas

4. ábra  Variancia paraméterek statisztikai vizsgálata

a) szórás ($s_{R,j}$) hisztogramja

b) relatív szórás ($V_{R,j}$) hisztogramja

8. Statistical dispersion parameters at test areas

The range, the relative range and the studentized range values were calculated for the 42 independent test areas. Frequency histograms are shown in Fig. 5 together with the best goodness of fit probability density functions. Best goodness of fit of Dagum distribution was found for the range values and Wakeby distribution for the relative range and the studentized range values. The observed positive skewness in the distribution of the range values is in accordance with literature data [17]. The influence of the removal of the 14 outlier individual rebound indices selected by Peirce’s criterion was also tested. Outlier removal was resulted in a shift toward smaller values of the mean for the range, the relative range and the studentized range and the skewness and peakedness of the frequency histograms have been changed as well. It seems that the extreme studentized deviate test by Peirce’s criterion has the strongest influence on the statistical dispersion parameters, as it was expected.

Fig. 5. Statistical analysis of the statistical dispersion parameters

a) range ($r_{R,j}$) for test areas

b) relative range ($\rho_{R,j}$) for test areas

c) studentized range ($\theta_{R,j}$) for test areas

5. ábra  Terjedelem paraméterek statisztikai vizsgálata

a) terjedelem ($r_{R,j}$) hisztogramja

b) relatív terjedelem ($\rho_{R,j}$) hisztogramja

c) studentizált terjedelem ($\theta_{R,j}$) hisztogramja

Fig. 6. Correlations between the mean, the range and the normality parameters

6. ábra  Várhatóérték, terjedelem és normalitási paraméterek korrelációi
9. Discussion

The strength estimation of concrete by rebound hammer testing usually applies empirical relationships between statistical location parameters (mostly the mean and the median) and the compressive strength. It can be realized that the three statistical location parameters are interrelated and a rather strong correlation is found between the mean and median values for the concrete floor in the present study [14]. Frequency histograms for the mean and median values are rather similar and can be modelled by the same type of probability distribution. The mode values, however, indicate weaker correlation with the other two statistical location parameters and the peakedness of the corresponding frequency histogram is more pronounced [14].

Frequency histograms of statistical location parameters and statistical distribution parameters were found to be well fit by the Wakeby distribution for the concrete floor in the present study. Wakeby distribution is typically used for flood frequency analysis and for other extreme events in hydrology [22]. Wakeby distribution has five parameters, therefore, it can mimic the shapes of many commonly used skewed distributions. Wakeby distribution is very flexible; this is the reason why it can be well fit to observed rainfall data, or flood peaks and other extreme events in hydrology [23].

Control parameters for the data of test areas can be the range and the standard deviation according to ASTM C805 International Standard [24].

It can be realized for the present study that no correlation is found between the statistical location parameters and the statistical dispersion parameters (mean vs. range and mean vs. studentized range is indicated in Fig. 6). It can be also seen that no correlation is found between the statistical location parameters and normality parameters as well as no correlation is found between the statistical dispersion parameters and normality parameters (particular plots are indicated in Fig. 6). Normality parameters (p-value, skewness, excess kurtosis) show no correlation with each other either.

It can be realized by studying the correlation (Fig. 7) between the statistical dispersion parameters and the statistical variance parameters that the range and the standard deviation are interrelated and a rather strong correlation is found between them for the concrete floor in the present study. Frequency histograms also confirm the observation. Studentized range – by definition – is determined by the ratio of the range and the standard deviation, therefore, the isouquant range and isouquant standard deviation responses are clearly visible in the studentized range vs. standard deviation and the range vs. studentized range representations, respectively (Fig. 7). It can be also demonstrated for the concrete floor in the present study that the studentized range indicates a value that is connected to the peakedness of the frequency histogram of rebound indices at the test areas; a visible, rather good correlation is demonstrated between studentized range and excess kurtosis (Fig. 7).

The statistical variance parameters determine each other by their nature (Fig. 4) and are expected to show positive skewness in their probability distributions according to the technical literature [17] that is visible in the frequency histograms (Fig. 4). The mean values for the recorded rebound indices are found to be relatively high. A comparison with literature data for the recorded mean rebound index vs. standard deviation as well as mean rebound index vs. coefficient of variation responses corresponding to the test areas is shown in Fig. 8.
10. Conclusions

Inherent variability analyses were carried out on in-situ rebound hammer test results collected at the bottom surface of a concrete floor over 130 m² of area tested. The following observations can be highlighted:

1. Robust data set of rebound indices is possible to be recorded by the subjective rejection of the suspected outliers by the operator during testing. Application of the extreme studentized deviate test by Peirce’s criterion resulted only 3% suspected outliers in the complete data set of 462 rebound index readings recorded at 42 individual test areas for the present tests. It was found that statistical dispersion parameters are the most sensitive to the Peirce’s criterion filtering and statistical location parameters are the least sensitive to the Peirce’s criterion filtering.

2. From an engineering point of view, normal probability distribution may be suggested only as a rough estimate for the frequency histogram of rebound indices collected at the 42 individual test areas if no separation of data by location is applied. It should be noted, however, that the goodness of fit of normal distribution was found to be inferior to skewed (or even uniform) distributions; best goodness of fit was found for the Fisher-Tippett (Generalized extreme value, GEV) distribution.

3. The hypothesis of normality (tested by Shapiro-Wilk test) can be generally accepted at a very low significance level for the 42 independent test areas analysed separately. Platykurtic behaviour governs the observations. Application of the Peirce’s criterion filtering decreases the deviation from normality in skewness, but can not improve excess kurtosis performance.

4. Skewed distributions can be fit with the best goodness of fit for the vast majority of frequency histograms of the local statistical parameters studied.

11. Acknowledgement

Author gratefully acknowledges the support of the Hungarian Scientific Research Fund project “Durability and performance characteristics of concretes with novel type supplementary materials” (OTKA K 109233).

References

[1] ACI (2003) In-Place Methods to Estimate Concrete Strength - ACI 228.1R-03. American Concrete Institute, Farmington Hills, Michigan
http://dx.doi.org/10.1016/j.conbuildmat.2013.11.113
Appendix. Formulae

The following local statistical measures were calculated for the inherent variability during the present analyses (in the formulae, $R_i$ accounts for the individual rebound index readings and $i$ denotes the $i^{th}$ test area):

- **mean value** (=arithmetic mean): $R_{m,j} = \frac{1}{n} \sum_{i=1}^{n} R_i$
- **standard deviation**: $s_{R,j} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (R_i - R_{m,j})^2}$
- **coefficient of variation**: $V_{R,j} = \frac{s_{R,j}}{R_{m,j}} \times 100$
- **range**: $r_{R,j} = R_{i,\text{max}} - R_{i,\text{min}}$
- **relative range**: $\rho_{R,j} = \frac{r_{R,j}}{R_{m,j}} \times 100$
- **studentized range**: $\theta_{R,j} = \frac{r_{R,j}}{s_{R,j}}$
- **extreme studentized deviate**: $\zeta_{i,j} = \frac{|R_i - R_{m,j}|}{s_{R,j}}$
- **skewness**: $\gamma_1 = \frac{\sum_{i=1}^{n} (R_i - R_{m,j})^3}{(n-1) \cdot (n-2) s_{R,j}}$
- **excess kurtosis**: $\gamma_2 = \frac{\sum_{i=1}^{n} (R_i - R_{m,j})^4}{(n-1) \cdot (n-2) \cdot (n-3) s_{R,j}^2} - \frac{3(n-1)^2}{(n-2)(n-3)}$