WARSAW UNIVERSITY OF TECHNOLOGY Index 351733
FACULTY OF CIVIL ENGINEERING
COMMITTEE FOR CIVIL AND WATER ENGINEERING
POLISH ACADEMY OF SCIENCES ISSN 1230-2945

DOI: 10.24425/ace.2022.140154

ARCHIVES OF CIVIL ENGINEERING

Vol. LXVIII ISSUE 1 2022

pp. 27–44

Research paper

Stochastic number of concrete families and the likelihood of such a value

Józef Jasiczak[1], Marcin Kanoniczak[2], Łukasz Smaga[3]

Abstract: Modern construction standards, both from the ACI, EN, ISO, as well as EC group, introduced numerous statistical procedures for the interpretation of concrete compressive strength results obtained on an ongoing basis (in the course of structure implementation), the values of which are subject to various impacts, e.g., arising from climatic conditions, manufacturing variability and component property variability, which are also described by specific random variables. Such an approach is a consequence of introducing the method of limit states in the calculations of building structures, which takes into account a set of various factors influencing structural safety. The term "concrete family" was also introduced, however, the principle of distributing the result or, even more so, the statistically significant size of results within a family was not specified. Deficiencies in the procedures were partially supplemented by the authors of the article, who published papers in the field of distributing results of strength test time series using the Pearson, t-Student, and Mann–Whitney U tests. However, the publications of the authors define neither the size of obtained subset and their distribution nor the probability of their occurrence. This study fills this gap by showing the size of a statistically determined concrete family, with a defined distribution of the probability of its isolation.

Keywords: compressive strength of concrete, continuous concreting process control, concrete family, statistical method

[1]Prof., DSC., PhD., Eng., Poznan University of Technology, Faculty of Civil and Transport Engineering, Piotrowo 5, 60-965 Poznań, Poland, e-mail: jozef.jasiczak@put.poznan.pl, ORCID: 0000-0003-3643-9819
[2]Assoc. Prof., PhD., Eng., Poznan University of Technology, Faculty of Civil and Transport Engineering, Piotrowo 5, 60-965 Poznań, Poland, e-mail: marcin.kanoniczak@put.poznan.pl, ORCID: 0000-0001-8017-645X
[3]Assoc. Prof., DSC.,PhD., Adam Mickiewicz University, Faculty of Mathematics and Computer Science, 61-614 Poznań, Poland, e-mail: ls@amu.edu.pl, ORCID: 0000-0002-2442-8816

1. Introduction

The subject of this article is included in the broadly understood statistical control of concrete quality and is another study by the authors on this subject. The previous articles [8–10] discussed: the description of the method, the selection of tests (ACME 2015), the example of extracting concrete families (C&C, 2014), the use of the proprietary method to assess seasonality in the production of concrete mix (JCE & IT, 2017), and in the next (ACE, 2021) the multiplicity of the set making up a concrete family. All articles were inspired by situations from practice in which, after the introduction of the PN-EN 206-1 standard, there were problems with the definition of a concrete family defined in the standard as "(...) a group of concretes with a defined and documented relationship between the relevant properties (...)".

An additional difficulty reported by concrete contractors was assigning the obtained concrete properties to a specific period. An example is the construction of a motorway pavement made of cement concrete, of a fixed class (C30/37, C35/45), divided into separate fields with a side of 6 m. The fields are numbered, concreted on the following days, and the contractor wants to know if the concrete has the required class. During the construction of the concrete highway Swiecko – Nowy Tomysl, in the case of material defects, not the entire section was replaced, but a specific field. A similar assessment was used in the construction of a concrete terminal for customs clearance in Swiecko.

When carrying out opinions on the relationship between the strength and the dates of laying concrete on the construction site, the authors initially recommended the use of the Shewhart individuals control chart with the lower and upper tolerance limits. On the card, the concrete strength was marked on the vertical axis, and the concreting dates on the horizontal axis. Activities on the card were entered manually, but it turned out that with a large number of results, the processes should be automated by creating special numerical software. For this reason, the cyclical production process was treated as a time series being the implementation of a stochastic process, the domain of which is precisely time. Concrete strength is a random process. To divide the strength time series into groups with comparable properties (according to the definition of the concrete family), the authors initially used Pearson's x^2 statistics as a criterion for separation (but also for interval compliance) in situations where tolerance limits were defined in advance.

Often, however, the obtained results were outside the limits of tolerance (the problem is described in the author's publication "The concept of 'over-strength of concrete' in the tender procedure for concrete objects of communication infrastructure", BTA, 1/2017 [12]) and for this reason, other tests were adopted (in the work *t*-Student and U-Mann–Whitney) and the corresponding numerical procedure according to R Core Team: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2015 [16].

Thanks to such a procedure, in this article, the boundaries of the separation and the test results were determined using two methods, the number of concrete families was calculated and the probability of such a number was established.

The authors of the study are aware of other possibilities of solving the problem, the more so that in the new edition of PN-EN 206-1 [23], the possibility of using CUSUM control cards was introduced (propagated in the works [4]). Some authors even suggest use in control processes Fuzzy models or sets – mathematical means of representing vagueness and imprecise information, which include randomly shaped concrete strength [3,5,14,17]. Also in these methods, a formal decision-making method in the form of, for example, a truncated V-mask is established to determine whether a process is out of control.

In the control procedure authors of this article, a similar regulatory function is performed by including the test results which differ from the rest of the group to the new concrete family.

2. Scientific grounds for the division into families of concrete

2.1. General assumptions

The European Standard PN-EN 206-1 [22] introduced the concept of a concrete family, which is defined as a group of concrete, with a specified and documented relationship between relevant properties, however without stating the stabilization of features over any periods. Assigning concrete to a family is strictly related to the relationship between strength and process conditions. The concept of a concrete family has been discussed in source literature [2,7,13,15,19,20].

Defining separate families of concrete is a division of a sequence of concrete compressive strength results into groups of statistically stabilized strength parameters, over specified periods of execution time. When continuously producing large quantities of concrete mix, correctly estimating a concrete family is justified from the perspective of reliability of buildings operated in the future [17,18,20,21].

Therefore, the analysis concerns a specified number of concrete compressive strength test results obtained by studying the strength of concrete test coupons sampled during on-site concreting structural elements.

The results of concrete compressive strength tests are subject to variability control in created subsets characterizing a specified concrete mix-manufacturing date and variability test of continuous subsets making up a certain closed number of analysed results.

To determine the specific, maximum, and mean compressive strength, that is, the parameters representing an entire set of the aforementioned results, it is necessary to develop a population histogram for such a set. Population, mean strength, an arithmetic mean of a distribution series were determined for each of the distinguished classes – strength ranges. The standard deviation for the entire result set was calculated as per the formula:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} n_i} \tag{2.1}$$

where:

n – means the sum of the population in all classes,

n_i – the size of the i-th class calculated in the range: $\langle x_i; x_{in} \rangle$,

$\overline{x}_i$ – mean compressive strength value of the i-th class and

$\overline{x}$ – arithmetic mean of the distribution series.

To compare the strength parameters of concrete in individual subsets, two results of random samples are compared, assuming that two independent random samples X and Y come from a normal distribution of unknown parameters, respectively: $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$.

A sequence of n numbered working plots, which characterize a specific number of produced concrete mix is tested. Each plot has an assigned sequence $x_i = \left(x_{i,1}, x_{i,2}, \ldots, x_{i,n_i}\right)$, $i = 1, 2, \ldots, n$ of concrete compressive strength test results.

The following hypotheses on expected values were assumed for verification:

$$\begin{cases} H_0: & \mu_{\{x_m, x_{m+1}, \ldots, x_{m+r}\}} = \mu_{\{x_{m+r+1}\}} \\ H_1: & \mu_{\{x_m, x_{m+1}, \ldots, x_{m+r}\}} \neq \mu_{\{x_{m+r+1}\}} \end{cases} \tag{2.2}$$

where:

$\mu_{\{x_m, x_{m+1}, \ldots, x_{m+r}\}}$ – an assumed, existing expected value of probability distribution of the combined sample $\{x_m, x_{m+1}, \ldots, x_{m+r}\}$,

m – numerical parameter determining the first subset of a sequence of compared subsets, containing several individual concrete compressive strength test results,

r – numerical parameter determining the number of subsets outside of the first subset within the analysed sequence.

The hypotheses (2.2) above is subject to verification, which involves analysing the strength parameters of individual subsets in the following order [1, 11, 17]:

1. Checking whether it belongs to a concrete family of subsets 1 and 2 (i.e. X_1 and X_2) by verifying the hypotheses (2.2) for $m = 1$ and $r = 0$:

$$\begin{cases} H_0: & \mu_{\{x_1\}} = \mu_{\{x_2\}} \\ H_1: & \mu_{\{x_1\}} \neq \mu_{\{x_2\}} \end{cases} \tag{2.3}$$

If the null hypothesis is not rejected, i.e., subset 1 is a concrete family with subset 2, move to point 2. If the null hypothesis is rejected, i.e., subset 1 is not a concrete family with subset 2, move to point 3.

2. Checking whether it belongs to a concrete family of subsets 1–2 and 3 (i.e. $X_{1,2}$ and X_3) by verifying the hypotheses (2.2) for $m = 1$ 1 and $r = 1$:

$$\begin{cases} H_0: & \mu_{\{x_1, x_2\}} = \mu_{\{x_3\}} \\ H_1: & \mu_{\{x_1, x_2\}} \neq \mu_{\{x_3\}} \end{cases} \tag{2.4}$$

If the null hypothesis is not rejected, i.e., subset 1–2 is a concrete family with subset 3, continue the verification of the hypotheses (2.2) as per point 2 for $m = 1$ and $r = 2$. If the null hypothesis is rejected, i.e., subset 1–2 is not a concrete family with subset 3, move to point 3 for $m = 3$ and $r = 0$.

3. Checking whether it belongs to a concrete family of subsets 2 and 3 (i.e. X_2 and X_3) by verifying the hypotheses (2.2) for $m = 2$ and $r = 0$:

$$\begin{cases} H_0 : \mu_{\{x_2\}} = \mu_{\{x_3\}} \\ H_1 : \mu_{\{x_2\}} \neq \mu_{\{x_3\}} \end{cases} \tag{2.5}$$

If the null hypothesis is not rejected, i.e., subset 2 is a concrete family with subset 3, move to point 2 and continue the calculation $m = 2$ and $r = 1$. If the null hypothesis is rejected, i.e., subset 2 is not a concrete family with subset 3, continue the calculations as per point 3 for $m = 3$ and $r = 0$.

Rejecting the null hypothesis H_0 will consistently mean adopting the assumption that samples come from various families of concrete. Whereas not rejecting the null hypothesis will consistently mean that the samples come from the same concrete family.

Two different statistical tests to compare the results of two random samples were used to verify the aforementioned hypotheses for expected values (2.3) and its non-parametric version. The analysis covered two independent random samples X and Y (representing accordingly two sets of results, each with a specified number of concrete compressive strength test results), coming from a population with continuous distributions. In the first calculation step, sample X was the first subset of the group of all analysed subsets, and sample Y was the second subset of the group of all analysed subsets. In the second calculation step, and then, analogically in subsequent ones, sample X was a set of results representing subsets, which were used in the previous step to make up a concrete family, and if such a family was not created – sample X was a set of results representing sample Y in the previous calculation step. Sample Y was also another previously unused subset of results.

The mean and variance of samples X and Y were calculated, and the following hypotheses – null hypothesis: $H_0: \mu_X = \mu_Y$ and alternative: $H_1: \mu_X \neq \mu_Y$, were assumed.

2.2. Verification whether a concrete compressive strength test results set belongs to a within a concrete family using the *t*-Student test

Studying concrete family classification using the t-Student test for two independent samples (concrete strength analyses for each of the plot are conducted independently) were conducted following several steps, using strength parameters determined based on the aforementioned subsets and from an entire set of concrete compressive strength test results.

The first step was to check whether the analysed data (results) come from a normal distribution. This involved conducting calculations using the Shapiro–Wilk test [18]. Next, samples X and Y were used to calculate the mean values and variances, respectively. Another operation was to verify the hypothesis on the equality of variances in two populations using the significance test for two variances, i.e., the F test [18]. The last essential element of the conducted analysis was to compare the results of two samples and check, whether they came from populations with the same expected values, using the significance test for two expected values, i.e., the t-Student test [18]. Two cases were considered. The first case, where samples X and Y are independent and come from a normal distribution, respectively: $N(\mu_X;\ \sigma_X^2)$; $N(\mu_Y;\ \sigma_Y^2)$ and there are equal variances for both samples ($\sigma_X^2 = \sigma_Y^2$) and a second case where samples X and Y are independent and come from a normal distribution, respectively: $N(\mu_X;\ \sigma_X^2)$; $N(\mu_Y;\ \sigma_Y^2)$ and the variances of both samples are not equal ($\sigma_X^2 \neq \sigma_Y^2$).

Not rejecting the null hypothesis $H_0: \mu_X = \mu_Y$ with an alternative hypothesis $H_1: \mu_X \neq \mu_Y$, i.e. when: $T(X,Y) < t(1-\alpha/2,\ n{+}m{-}2)$ for the first case or $T(X,Y) < t(1-\alpha/2,\ \beta)$ for the second case, means adopting the hypothesis on the classification of a group of results representing samples X and Y within one concrete family.

Whereas, if the null hypothesis is rejected, i.e., if for the first case $T(X,Y) > t(1-\alpha/2,\ n{+}m{-}2)$ or if: $T(X,Y) > t(1-\alpha/2,\ \beta)$ for the second case, the hypothesis on the classification of a group of results representing samples X and Y within one concrete family is rejected.

2.3. Verification of the hypothesis on the classification of a set of concrete compressive strength test results within a concrete family using the Mann–Whitney U test

Studying the classification within a concrete family involves verifying hypotheses on the equality of distribution cumulative functions for two samples:

$$\begin{cases} H_0: & F_{\{x_m, x_{m+1}, \ldots, x_{m+r}\}} = F_{\{x_{m+r+1}\}} \\ H_1: & F_{\{x_m, x_{m+1}, \ldots, x_{m+r}\}} \neq F_{\{x_{m+r+1}\}} \end{cases} \tag{2.6}$$

where: $F_{\{x_m, x_{m+1}, \ldots, x_{m+r}\}}$ – mean a probability distribution function of the combined sample $\{x_m,\ x_{m+1},\ \ldots,\ x_{m+r}\}$.

Hypotheses (2.6) are subject to the same verification as the hypotheses (2.2), however, with the use of the Mann–Whitney U test [18]. The null hypothesis H_0 assumes that samples X and Y were collected from the same distribution, whereas the alternative hypothesis H_1 assumes that samples X and Y were not collected from the same distribution. Studying the classification within a concrete family using the Mann–Whitney U test for two independent samples involves the consideration of two sets, from which independent random samples n and m are collected. All observations are subject to ordering in ascending order. In the event of the same observations in samples X and Y, one should apply a correction, which involves supplementing the value of the statistic U with half of the number of pairs (x, y) such that $x = y$.

Not rejecting the null hypothesis H_0: $F_X = F_Y$ with an alternative hypothesis H_1: $F_X \neq F_Y$, i.e. $U(X,Y)$ does not belong to the range $C = [0,\ u(n,m,\alpha/2)] \cup [u(n,m,1-\alpha/2),\infty]$ means adopting the hypothesis on the classification of a group of results representing samples X and Y within one concrete family. Whereas, if the null hypothesis is rejected, i.e. when: $U(X,Y)$ belongs to the range $C = [0, u(n,m,\alpha/2)] \cup [u(n,m,1-\alpha/2),\infty]$ means rejecting the hypothesis on the classification of a group of results representing samples X and Y within one concrete family.

3. Calculation examples

3.1. Analysing an annual set of concrete compressive strength test results

A hydrotechnical structure consisting of several reinforced concrete tanks made of C35/45 concrete was completed within the Poznań area in recent years. The contract obliges the manufacturer of the concrete mix to maintain a constant concrete recipe throughout the year. For this reason, the manufacturer adopted variable concrete recipes at different times of the year to maintain constant concrete parameters (Table 1). The observations involved tanks made of concrete over one year, from January to December. The control covered the stabilization of the characteristics of both the mixture (described in this article) and the technology of making the tanks themselves. Test samples 15 × 15 × 15 cm were collected on each day of concreting and the concrete compressive strength was evaluated after 28 days of curing.

Table 1. Summer and winter mix recipes for maintaining constant strength of concrete throughout the year

Parameters	Summer season	Winter season
Class of concrete	C35/45	C35/45
Kind of cement	CEM I 42.5 N	CEM I 42.5 N
W/c ratio	0.44	0.44
Sand 0–2 mm	548	546
Gravel 2–8 mm	531	530
Gravel 8–16 mm	581	579
Fly ash	45	45
Plasticizer	CER – 0.25%	–
Superplasticizer	O132 – 0.85%	O146 – 0.85%
Air-entraining admixture	AIR A10 – 0.15%	AIR A10 – 0.15%
Retarding admixture	TARD – 0.30%	–
Accelerating admixture	–	X384 – 0.90%
28 day strength, MPa	52.1	52.7

Strength time series (a total of 134 ranges representing concreting dates within the year) were used to determine the fundamental strength parameters for the year-long set of concrete compressive strength test results. The concrete families were also determined. A collective population histogram (Fig. 1) was also developed to confirm that concrete achieved the designed class. This enabled achieving parameters representing the entire set of data (719 individual results of concrete compressive strength tests), without relation to the time of concrete mix production.

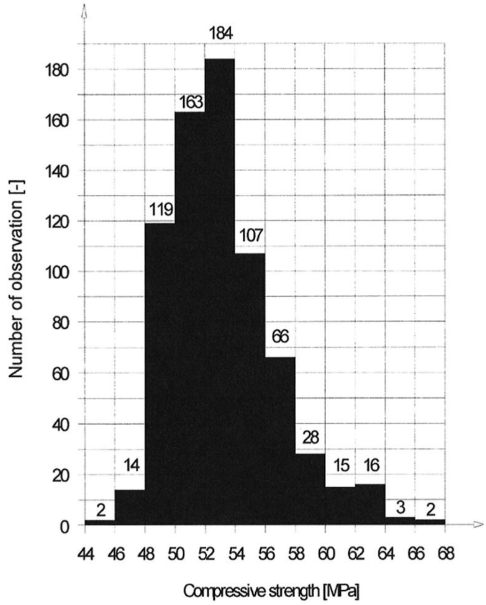

Fig. 1. Population histogram for an annual set of concrete compressive strength test results

The entire set of results exhibits an average compressive strength at a level of 53.1 MPa, a standard deviation of 3.6 MPa, and specific strength f_{ck} = 47.2 MPa, higher than the initially assumed f_{ck} = 45.0 MPa, but belonging to the same C35/45 concrete design class. The obtained large data set allowed for the identification of concrete families and the determination of their number. By analyzing the strength test results appearing successively during the year, it is possible to confirm their high variability but only selected tests can show the distinctiveness resulting from belonging to different concrete families.

3.2. Division of concrete compressive strength test result sequence into concrete families

The authors verified the assumed hypotheses (2.2) on the classification of concrete compressive strength test result set within a concrete family using the *t*-Student and Mann–Whitney U tests, as per the calculation procedure discussed in point 2.

The outcome of the first test was a division of the entire result set into 68 concrete families, whereas the second test divided the entire result set into 71 concrete families.

Publication [4] summarizes the basic data on the test result sequence, their corresponding standard deviations, and ranges with statistically stabilized parameters determined through the verification of statistical hypotheses using the *t*-Student and Mann–Whitney U test.

The results of verification of the aforementioned hypotheses for the beginning of a test result sequence are given in Table 2. The outcome of the verification calculations is the

Table 2. Summarized data on the general evaluation of a set and division into concrete families

Entire set of concrete compressive strength test results	Size of set	**719**														
	Mean value, standard deviation, f_{ck}, MPa	**53.1; 3.6; 47.2;**														
Division into sequences into concrete families		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Division using *t*-Student test	Concrete family size	5	20			7	7	7	15		26				5	7
	Mean compressive strength, MPa	55.4	50.3			52.9	56.1	62.0	52.5		58.8				50.0	60.3
Division using Mann–Whitney U test	Concrete family size	5	20			7	14		15		26				5	7
	Mean compressive strength, MPa	55.4	50.3			52.9	59.0		52.5		58.8				50.0	60.3
Division into concrete families – continued		16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Division using *t*-Student test	Concrete family size	16			12		10		17			10		5	5	5
	Mean compressive Strength, MPa	51.5			56.8		49.6		54.3			61.2		54.0	57.2	53.4
Division using Mann–Whitney U test	Concrete family size	10		6	12		10		17			10		5	5	5
	Mean compressive strength, MPa	51.8		51.0	56.8		49.6		54.3			61.2		54.0	57.2	53.4
Division into concrete families – continued		31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Division using *t*-Student test	Concrete family size	10		5	15			11		10		5	10		12	
	Mean compressive strength, MPa	56.2		54.0	49.6			52.5		50.4		57.6	54.6		59.2	
Division using Mann–Whitney U test	Concrete family size	10		5	15			5	6	10		5	5	5	12	
	Mean compressive strength, MPa	56.2		54.0	49.6			51.6	53.7	50.4		57.6	53.8	55.2	59.2	

histograms for 1, 2, 3, . . . , n-element subsets, corresponding to the established concrete families is shown in Fig. 2.

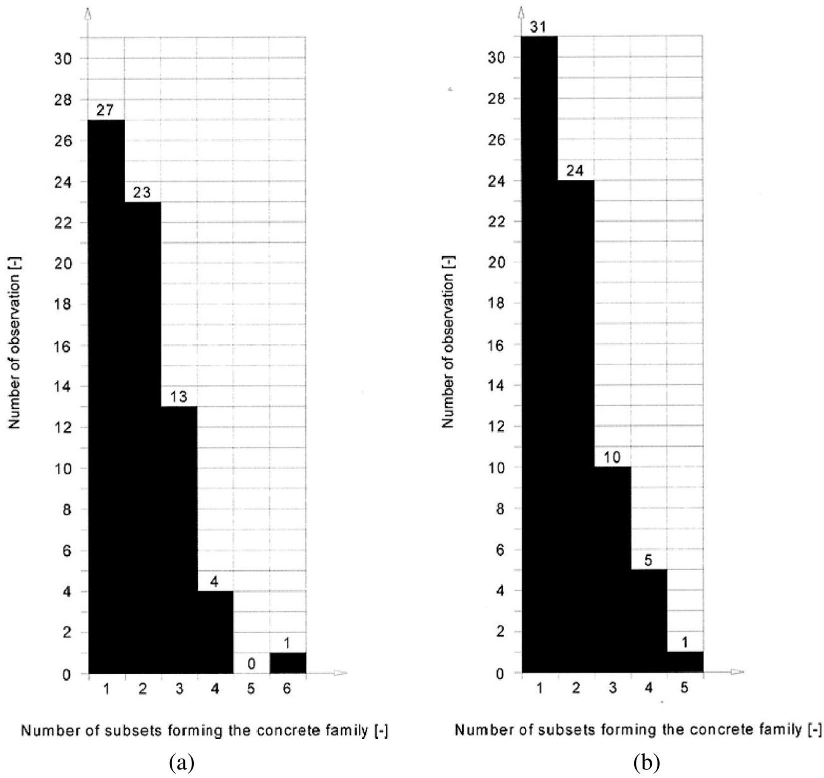

Fig. 2. Numbers of subsets forming concrete families according to: a) t-Student test, b) Mann–Whitney U test

3.3. Number of subsets making up a concrete family

This chapter discusses the distribution of the number of subsets making up a concrete family for the data considered in Section 2. The descriptive statistics methods [14], which present an empirical distribution of this number were used and two theoretical distributions modelling this distribution were suggested for this purpose.

The empirical distribution was described by descriptive statistics (mean, median, and standard deviation), empirical probabilities (percent) of the occurrence of individual numbers of subsets making up a concrete family, and bar charts. The number of subsets making up a concrete family was theoretically described with the use of two discrete distributions since this number can be a natural number, i.e. take values within the set $\{1, 2, \ldots\}$. Two distributions, Poisson distribution, and binomial distribution were matched out of the known ones. However, in both these distributions, the lowest value was zero, which does not appear as a value of the number of subsets making up a concrete family. This is why the modifications of these two distributions were considered. Namely, they were shifted by one, and then the lowest value of the new distributions will be equal to one. A detailed description of these statistical models is presented below.

Let X_1 and X_2 be random variables with Poisson $\pi(\lambda)$ and binomial $b(m, p)$ distributions, respectively, where $\lambda > 0$ and $p \in (0,\ 1)$ are unknown parameters and $m \in \{0,\ 1, \ldots\}$ is fixed ([14], pp. 149–152). In the binomial distribution, we adopt the number $m = 133$ as the number of experiments, since our data set considered 134 observation days. We take into account two random variables $L_1 = X_1 + 1$ and $L_2 = X_2 + 1$, which will model the number of subsets making up a concrete family. The first one adopts values that are natural numbers $(1,\ 2, \ldots)$, and the second one adopts $1, 2, \ldots, m + 1$. This is reasonable since the number of subsets making up a concrete family cannot be higher than $m + 1 = 134$, and moreover, its high value is unlikely. Theoretical distribution models for the number of subsets making up a concrete family based on the distributions of random variables L_1 and L_2 will be called Models 1 and 2, respectively.

Distributions of random variables L_1 and L_2 are described by the following probability functions

$$(3.1)\qquad \begin{aligned} P(L_1 = l) &= \frac{\lambda^{l-1} e^{-\lambda}}{(l-1)!}, && \text{for } l = 1, 2, \ldots \\ P(L_2 = l) &= \binom{m}{l-1} p^{l-1} (1-p)^{m-l+1}, && \text{for } l = 1, 2, \ldots, m+1 \end{aligned}$$

Expected values, medians, and standard deviations of random variables L_1 and L_2 are as follows:

$$(3.2)\qquad \begin{aligned} &E(L_1) = \lambda + 1, && M_e(L_1) \approx \left\lfloor \lambda + \frac{1}{3} - \frac{0.02}{\lambda} \right\rfloor + 1, && S_d(L_1) = \sqrt{\lambda} \\ &E(L_2) = mp + 1, && M_e(L_2) = \lfloor mp \rfloor + 1 \vee \lceil mp \rceil + 1, && S_d(L_2) = \sqrt{mp(1-p)} \end{aligned}$$

The values depend on unknown parameters λ and p, which are estimated using the maximum likelihood method (Górecki, 2011, p. 195). Let $L = (L_1, L_2, \ldots, L_n)^T$ be a simple sample from the population for a random variable L, which is the number of subsets making up a concrete family. Estimators of parameters λ and p are as follows:

$$\hat{\lambda} = \overline{L} - 1, \qquad \hat{p} = \frac{\overline{L} - 1}{m}$$

where

$$\overline{L} = n^{-1} \sum_{i=1}^{n} L_i$$

is the sample mean. Therefore, in the maximum likelihood method, the estimators for expected values, medians, and standard deviations in Models 1 and 2 are as follows:

$$E(\hat{L}_1) = \overline{L}, \quad M_e(\hat{L}_1) = \left\lfloor \overline{L} - \frac{2}{3} - \frac{0.02}{\overline{L} - 1} \right\rfloor + 1, \quad S_d(\hat{L}_1) = \sqrt{\overline{L} - 1}$$

(3.3)

$$E(\hat{L}_2) = \overline{L}, \quad M_e(\hat{L}_2) = \lfloor \overline{L} \rfloor \vee \lceil \overline{L} \rceil, \qquad S_d(\hat{L}_2) = \sqrt{\left(\overline{L} - 1\right)\left(1 - \frac{\overline{L}-1}{m}\right)}$$

respectively.

The distribution of the number of subsets making up a concrete family was estimated using the data discussed in chapter 2, and an empirical distribution and the theoretical distribution of Models 1 and 2 in a tabular and graphic manner.

Table 3. Summary of the number of subsets obtained based on the t-Student test, using an empirical distribution and two theoretical distributions (Models 1 and 2).
The meaning of the rows is as follows: Estimator – values of estimators for the parameters of theoretical distributions; P-value – p-values of the Pearson's chi-square test for the conformity of the theoretical distribution; Mean, Median and Standard deviation – values of estimators for the expected values, medians, and standard deviations; 1, 2, 3, 4, 6 – estimated probabilities for the occurrence of subset number equal to 1, 2, 3, 4, 6

Description	Empirical	Model 1	Model 2
Estimator	N/A	0.9705882	0.0072977
P-value	N/A	0.9584091	0.9562437
Mean	1.9705882	1.9705882	1.9705882
Median	2.0000000	2.0000000	2.0000000
Standard deviation	1.0362186	0.9851844	0.9815830
1	0.3970588	0.3788601	0.3775142
2	0.3382353	0.3677172	0.3691044
3	0.1911765	0.1784510	0.1790843
4	0.0588235	0.0577341	0.0574873
6	0.0147059	0.0027194	0.0026050

Namely, Tables 3 and 4 state the values of estimators for the expected value, median and standard deviation of the number of subsets making up a concrete family, and the estimations of probabilities for the occurrence of the numbers of subsets obtained within the collected data.

These tables also state the values of estimators for the parameters of the theoretical distributions of Models 1 and 2 and the p-value of the Pearson's chi-square test [11, 14] with a given theoretical distribution, which, in all cases, are significantly higher than the significance level $\alpha = 0.05$. Therefore, these values indicate a very good match of both

Table 4. Summary of the number of subsets obtained based on the Mann–Whitney U test, using an empirical distribution and two theoretical distributions (Models 1 and 2).
The meaning of the rows is as follows: Estimator – values of estimators for the parameters of theoretical distributions; P-value – p-values of the Pearson's chi-square test for conformity of the theoretical distribution; Mean, Median and Standard deviation – values of estimators for the expected values, medians, and standard deviations; 1, 2, 3, 4, 5 – estimated probabilities for the occurrence of subset number equal to 1, 2, 3, 4, 5

Description	Empirical	Model 1	Model 2
Estimator	N/A	0.8873239	0.0066716
P-value	N/A	0.7514698	0.7368091
Mean	1.8873239	1.8873239	1.8873239
Median	2.0000000	2.0000000	2.0000000
Standard deviation	0.9935405	0.9419787	0.9388312
1	0.4366197	0.4117562	0.4105338
2	0.3380282	0.3653611	0.3667231
3	0.1408451	0.1620968	0.1625623
4	0.0704225	0.0479441	0.0476769
5	0.0140845	0.0106355	0.0104071

theoretical models with the data. This is also confirmed by the bar charts (Fig. 3 and 4) for the number of subsets making up a concrete family.

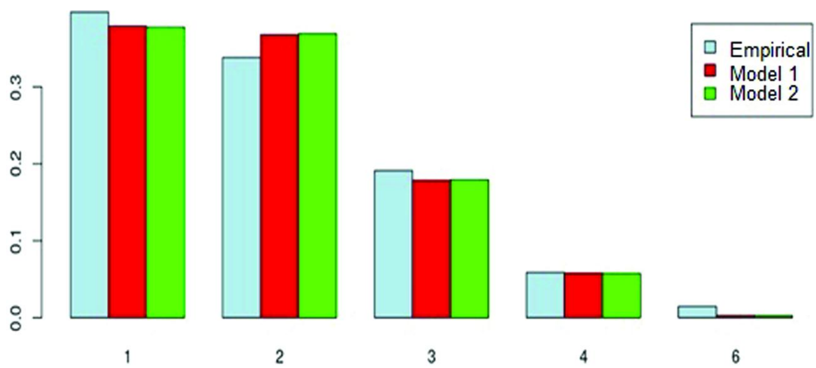

Fig. 3. Bar charts covering the number of subsets obtained based on the t-Student test, using an empirical distribution and two theoretical distributions (Models 1 and 2)

The obtained numerical values, as well as the bar charts, indicate that a low number of subsets making up a concrete family, i.e. 1, 2, and 3 is the most probable, and a comment to these findings has been added in the conclusions.

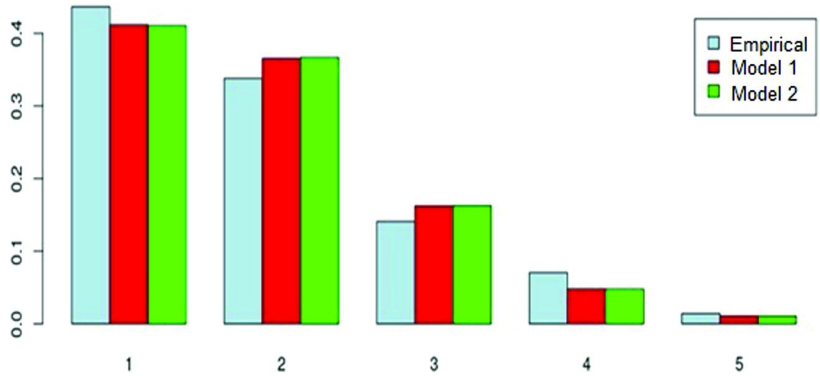

Fig. 4. Bar charts covering the number of subsets obtained based on the Mann–Whitney U test, using an empirical distribution and two theoretical distributions (Models 1 and 2)

4. Final analysis and conclusions

4.1. Final analysis

The outcome of verifying the assumed hypothesis (2.2) on the classification of a concrete compressive strength test result set within a concrete family using the *t*-Student and Mann–Whitney U tests, as per the calculation procedure presented in point 2, is a division of the entire result set into 68 concrete families in the case of the first test and 71 families in the case of the second test. This comment will concern two issues, i.e. shaping of the concrete class value for individual families and the frequency of incidence of range numbers within these families from $1, \ldots, n$, and the probable number of sample sets within a single-family.

Regarding the distribution of 719 test results and 134 individual ranges resulting in 68 concrete families (71 in the case of the second test), it can be noticed that as many as 64% of the families were classified in the higher classes, C40/50 or even C45/55, and only 36% in the assumed C35/45 class. In the case of the second test, the numbers are 66.7% and 33.3%, respectively. The aforementioned results indicate the existence of strength instability within the entire result set. There are local overestimations within the time intervals, significantly exceeding the design concrete class. This is a safe side error but it comes with increased expenditure associated with producing concrete within a higher strength range, unnecessary from the point of view of structural safety. The commonly applied passive control can ensure structural safety but in an economically inefficient manner.

By treating the obtained yearly sequence of concrete compressive strength test results as a stochastic process, one can obtain a locally unstable system, and the outcome is a frequent (every 1, 2, or 3 production days) change in the assignment to specific families, hence, a too low number of result ranges assigned to a single-family (for the *t*-Student test division results: almost 40% of single-range families, 34% of two-range families, 19% of three-range families, and multi-range families are only 7%; whereas for the Mann–Whitney

U test: single range families make up almost 44%, two-range families 34%, three-range families 14% and multi-range families only 8%). Since one range represents 5-8 individual strength test results, two ranges 10-16 and three ranges 18-21, estimating the characteristics of a family must follow two formulas:

$$\text{size} < 15: f_{cm} \geq f_{ck} + 4 \quad \text{and} \quad f_{ci} \geq f_{ck} - 4, \ \text{MPa}$$
$$\text{size} \geq 15: f_{cm} \geq f_{ck} + 1.48 s_R \quad \text{and} \quad f_{ci} \geq f_{ck} - 4 \ \text{MPa}$$

where:
f_{cm} – average from n strength test results for n series of samples,
f_{ck} – specific compressive strength (concrete class),
f_{ci} – single strength test result out of n series of samples,
s_R – standard deviation for concrete family.

The above confirms the thesis on the need to apply the suggested distribution methods in practice, even though approx. 40% of the estimation are based on a so-called small sample ($n < 15$, which in turn can lead to underestimating value f_{ck} [11].

The considerations regarding a small number of result ranges assigned to a single concrete family, summed up in conclusion 4.4 have been described with and confirmed by two statistical models. Models 1 and 2 were based on Poisson and binomial distributions shifted by one, respectively, since the number of ranges can be at least one. The parameters of these models were estimated using the maximum likelihood method. Furthermore, the adequacy of Models 1 and 2 was justified not only using descriptive statistics models but also their statistical significance was demonstrated through the Pearson's chi-square conformity test. Models 1 and 2 theoretically confirm that, for the discussed data regarding concrete compressive strength and their division using the t-Student and Mann–Whitney U tests, the number of result ranges assigned to one concrete family is minor and is mainly 1, 2, and 3, in proportions given previously.

5. Conclusions

The entire concrete mix production process requires continuous monitoring. The compressive strength of concrete is a fundamental parameter subject to inspection. This strength is tested based on test coupons systematically sampled, specified quantities of the mix, at a specified time of the process. The consequential result set is then evaluated. A yearly set of concrete compressive strength test results can be evaluated globally, in terms of the strength parameters representing all results from a given set. This exhibits an average compressive strength at a level of 53.1 MPa, the standard deviation of 3.6 MPa, and specific strength $f_{ck} = 47.2$ MPa, higher than the initially assumed $f_{ck} = 45.0$ MPa, but belonging to the same C35/45 concrete design class. However, the analysis of all results in a set does not enable detecting a lowered or overvalued concrete strength, which can occur at certain production process time intervals.

The efficiency of a strength evaluation system can be significantly improved by thoroughly considering concrete families, and implementing active control assessment based on the verification of the statistical hypotheses described in the paper. Analysing the strength parameters of concrete manufactured at a given time of the production process, represented by a group of results characterizing the statistical invariability of the parameters, namely, belonging to a single concrete family, enables detecting inadequate strength of a produced element (group of elements). This is fully justified from the perspective of structural reliability and the economic optimization of concrete mix production.

Acknowledgements

This work was supported by the Poznan University of Technology – Polish Academy of Science under Grant No. 0412/SBAD/0022.

References

[1] A. Sarja, "Durability design of cocnrete structures – Committee report 130-CSL", *Materials and Structures*, 2017, vol. 33, pp. 14–20, DOI: 10.1007/BF02481691.

[2] *Concrete according to standard PN EN 206-1 – commentary – collective work supervised by prof. Lech Czarnecki*. Kraków: Polski Cement, 2004.

[3] I. Skrzypczak, W. Kokoszka, J. Zięba, A. Leśniak, D. Bajno, Ł. Bednarz, "A Proposal of a Method for Ready-Mixed Concrete Quality Assessment Based on Statistical-Fuzzy Approach", *Materials*, 2020, vol. 13, no. 24, DOI: 10.3390/ma13245674.

[4] I. Skrzypczak, L. Buda-Ożóg, J. Zięba, "Dual CUSUM chart for the quality control of concrete family", *Cement Wapno Beton, CWB*, 2019, vol. 24, no. 4, pp. 276–285, DOI: 10.32047/CWB.2019.24.4.3.

[5] I. Skrzypczak, L. Buda-Ożóg, T. Pytlowany, "Fuzzy method of conformity control for compressive strength of concrete on the basis of computational numerical analysis", *Meccanica*, 2016, vol. 51, pp. 383–389, DOI: 10.1007/s11012-015-0291-0.

[6] J. Jasiczak, "Probabilistic Criteria for the Control of Compressive Strength Stabiilization in Concrete", *Foundations of Civil and Environmental Engineering*, 2011, no. 14, pp. 47–61.

[7] J. Jasiczak, M. Kanoniczak, Ł. Smaga, "Standardized concept of a concrete family on the example of continuous Spiroll board production", *Budownictwo i Architektura*, 2014, vol. 13, no. 2, pp. 99–108.

[8] J. Jasiczak, M. Kanoniczak, Ł. Smaga, "Statistical division of compressive strength results on the aspect of concrete family concept", *Computers and Concrete*, 2014, vol. 14, no. 2, pp. 145–161.

[9] J. Jasiczak, M. Kanoniczak, L. Smaga, "Stochastic identity of test result series of the compressive strength of concrete in industrial production conditions", *Archives of Civil and Mechanical Engineering*, 2015, vol. 15, pp. 584–592.

[10] J. Jasiczak, M. Kanoniczak, Ł. Smaga, "Division of Series of Concrete Compressive Strength Results into Concrete Families in Terms of Seasons within Annual Work Period", *Journal of Computer Engineering&Information Technology*, 2017, vol. 6, no. 3, pp. 1–9, DOI: 10.4172/2324-9307.1000198.

[11] J. Jasiczak, M. Kanoniczak, "Justified adoption of normative values f_{ci} and f_{cm} in the estimation of concrete classification for small samples", *Journal of Civil Engineering, Environment and Architecture, JCEEA,* 2017, vol. XXXIV, no. 64 (3/I/17), pp. 203–212, DOI: 10.7862/rb.2017.115.

[12] J. Jasiczak, "The concept of 'over-strength of concrete' in the tender procedure for concrete objects of communication infrastructure", *BTA,* 2017, no. 1, pp. 64–68 (in Polish).

[13] L. Taerwe, "Basic aspect of quality control of concrete", in *"Utilizing Redy Mix Concrete and Mortar", Proceedings of the International Conference*. UK, Scotland, 1999, pp. 221–235.

[14] N.K. Nagwani, "Estimating the concrete compressive strength using hard clustering and fuzzy clustering based regression techniques", *The Scientific World Journal*, 2014, vol. 2014, DOI: 10.1155/2014/381549.

[15] R. Caspeele, L. Taerwe, "Conformity control of concrete based on the 'concrete family' concept", in *Proceedings of the 5th International Probabilistic Control, 28–29 Nov.2007*. Ghent, 2007, pp. 241-252.

[16] *R Core Team: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. [Online]. Available: http://www.R-project.org/.

[17] S. Woliński, "Evaluating the quality of concrete using standardized methods and according to fuzzy logic", in *"Dni Betonu" Conference*, Kraków: Polski Cement, 2006, pp. 1121–1131 (in Polish).

[18] T. Górecki, *Basics of statistics with examples in R.* Legionowo: BTC, 2011.

[19] Z. Kohutek, "Concrete family – concept genesis, terminology, criteria and general creation principles", *Przegląd Budowlany*, 2010, no. 10, pp. 26–31 (in Polish).

[20] EN 1992:2008 Eurocode 2: Design of concrete structures.

[21] ISO 2394:2000 General principles on reliability for structures.

[22] PN–EN 206–1: 2003 Concrete. Part 1: Requirements, properties, production and conformity.

[23] PN-EN 206+A1:2016-12. Concrete. English version.

Stochastyczna liczba wyników badań tworzących rodzinę betonu i prawdopodobieństwo wystąpienia takich wartości

Słowa kluczowe: wytrzymałość betonu na ściskanie, ciągła kontrola procesu betonowania, rodzina betonu, metody statystyczne

Streszczenie:

Współczesne normy budowlane zarówno z grupy EN, ISO jak i EC wprowadziły wiele procedur statystycznych do interpretacji uzyskiwanych na bieżąco (w trakcie realizacji obiektu) wyników badań wytrzymałości betonu na ściskanie, której wartości podlegają różnym przypadkowym wpływom, na przykład wynikającym z warunków klimatycznych, zmienności produkcji zmienności właściwości składników, które również opisują określone zmienne losowe.

Podejście takie jest konsekwencją wprowadzenia do obliczeń konstrukcji budowlanych metody stanów granicznych uwzględniającej zbiór różnych czynników wpływających na bezpieczeństwo konstrukcji. Z tego powodu wdrożono w ostatnich latach wiele procedur kontrolujących i regulujących dotrzymanie przez producenta betonu granicznych parametrów mieszanki, poczynając od statystycznej, globalnej oceny wytrzymałości, poprzez procedury przedziałowe (karty kontrolne Shewarta), po skomplikowane analizy stochastyczne, zawierające drobnoprzedziałowe oceny ciągów wyników badań o ujednoliconej, statystycznie istotnej, wartości parametrów podstawowych wytrzymałości. Szczególnie dużo uwagi poświęca się, zarówno w praktyce budowlanej jak i w rozważaniach teoretycznych, zagwarantowaniu przez producenta mieszanki wytrzymałości betonu z 95% prawdopodobieństwem jej wystąpienia.

W normie europejskiej PN-EN 206-1 wprowadzono dodatkowo termin rodzina betonów (ang. Family of concrete concept), którą określono jako "(...) *grupę betonów o ustalonej i udokumentowanej zależności pomiędzy odpowiednimi właściwościami"*, bez podania jednak oznaczeń ilościowych odnośnie wielkości tej grupy i stabilizacji cech (na przykład wytrzymałości betonu na ściskanie) w jakichkolwiek przedziałach czasowych. Przy wytwarzaniu w sposób ciągły dużych ilości mieszanki betonowej, poprawne oszacowanie rodziny betonów jest zasadne z punktu widzenia niezawodności eksploatowanych później konstrukcji budowlanych, o czym świadczy bogata literatura zacytowana w artykule.

Przyporządkowanie betonu do rodziny jest ściśle związane z relacją pomiędzy wytrzymałością a uwarunkowaniami technologicznymi. Wyznaczenie oddzielnych zbiorów (rodzin betonów)

jest podziałem ciągu wyników badań wytrzymałości betonu na ściskanie na grupy o statystycznie ustabilizowanych parametrach wytrzymałościowych w określonych przedziałach czasowych ich wykonania.

Przedmiotem analiz zamieszczonych w niniejszej pracy jest więc określona, szczególnie duża liczba wyników badań wytrzymałości betonu na ściskanie zebranych w ciągu jednego roku podczas betonowania kilku obiektów hydrotechnicznych o takiej samej klasie wytrzymałościowej C35/45 i stałej recepturze z drobnymi modyfikacjami sezonowymi (lato, zima).

W części teoretycznej pracy podano podstawy weryfikacji hipotez o wyodrębnieniu szeregów czasowych wytrzymałości o statystycznej zwartości, tworzących tzw. rodziny betonu. Rozdziału wyników badań szeregu czasowego wytrzymałości dokonano stosując testy Pearsona, t-Studenta i Manna–Whitneya. Na wybranym przykładzie określono liczby uzyskanych podzbiorów oraz prawdopodobieństwa wystąpienia takich liczności. Jest to istotne wzbogacenie teorii jakości betonu o liczność statystycznie wydzielonej rodziny betonu z określeniem rozkładu prawdopodobieństwa jej wyodrębnienia.

Received: 10.02.2021, Revised: 24.06.2021