Construction and Building Materials 268 (2021) 121145

Contents lists available at ScienceDirect

Construction and Building Materials

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

Evaluation of the repeatability and reproducibility of the double punch test

André Baltazar Nogueira [a,*], Luana de Carvalho Ribeiro Simão [a], Renata Monte [a], Renan P. Salvador [b], Antônio Domingues de Figueiredo [a]

[a] *Department of Civil Construction Engineering, Polytechnic School of University of São Paulo, Professor Almeida Prado Av, Trav 2, 83, 05424-970 São Paulo, Brazil*
[b] *Department of Civil Engineering, São Judas Tadeu University, 546 Taquari St., 03166-000 São Paulo, Brazil*

ARTICLE INFO

Article history:
Received 27 June 2020
Received in revised form 21 September 2020
Accepted 28 September 2020
Available online 21 October 2020

Keywords:
Fiber reinforced concrete
Double punch test
Repeatability
Reproducibility

ABSTRACT

The constant advances in the use of fiber reinforced concrete (FRC) bring the need of developing practical tests for quality control programs. The double punch test (DPT) is a promising alternative technique in relation to the flexural test of notched beams as proposed by the fib Model Code 2010. In that sense, this work aims to investigate the possibility of finding repeatability and reproducibility of the crack and residual loads determined by DPT. For the repeatability study, an experimental program was developed in a single laboratory, involving two steel fiber contents (30 kg/m^3 and 55 kg/m^3) in three equivalent concrete batches. For the reproducibility study, an interlaboratory program was developed, involving the same fiber and contents and the same concrete matrix, tested in six different laboratories. In both parts, analysis of variance and the ASTM E691 method were used to compare average and variance values. In addition, the minimum number of specimens necessary to provide adequate results of average and variance was determined. The results showed that the DPT is sensitive to variations in the mechanical properties after cracking of the FRC that are of special interest for structural application. In addition, it is possible to find repeatability and reproducibility of residual loads in terms of average values with a limited number of specimens. The crack load was repeatable in the same fiber content, but not reproducible. Despite this, DPT can be considered as efficient for the ordinary control of the quality of the FRC.

1. Introduction

The constant advances on the use of fiber reinforced concrete (FRC) for structural purposes bring the need of definitions of quality control programs. These programs should be based on tests that provide reliable results and, preferably, easy to apply [1,2]. The test method should also guarantee the precise evaluation of the parameters that give the FRC the possibility of being used as a structural material. An important reference in that sense is the *fib* Model Code 2010 [3], which guides the FRC parameterization for structural applications.

The fib Model Code 2010 selected the three-point bending test (3PBT) EN 14651 [4,5] as a reference method for FRC structural parameterization. This test is performed with prismatic notched beams for the determination of the flexural strength and the residual strength of the composite at different crack opening levels. Nevertheless, the execution of this test requires a sophisticated equipment, with closed loop control, which is not easily available in common laboratories in Brazil and other underdevelopment countries and. Therefore, the use of this type of equipment is considered a restriction to conduct quality control procedures.

Another drawback of this test is related to the volume and weight of the specimen used. A single beam for the 3PBT contains 12.4 L of concrete, weighs about 30 kg, and, consequently, it is not practical for a laboratory labor routine as it increases the risk to the workers' health. In addition, it is impracticable to extract a prismatic specimen from a concrete structure due to the large specimen size and difficulties associated to the sawing process. As a consequence, the flexural test is not applicable for existing structural evaluations

The double-punch test (DPT) was developed aiming its use as an alternative test method for FRC [1,6,7]. The test was standardized in Spain by the standard AENOR UNE 83515 [8]. The test performs a double punch in cylindrical specimens measuring 150 mm in diameter that weigh approximately 5.5 kg, which is less than one fifth of the weight of a prismatic specimen used in the 3PBT.

* Corresponding author.
E-mail addresses: andre.nog88@gmail.com (A.B. Nogueira), luana@unit.com.br (L. de Carvalho Ribeiro Simão), renata.monte@usp.br (R. Monte), antonio.figueiredo@usp.br (A.D. de Figueiredo).

https://doi.org/10.1016/j.conbuildmat.2020.121145
0950-0618/

The DPT may also be used to evaluate the parameters associated with the post-crack behavior of FRC.

The standard test makes use of a circumferential extensometer placed at specimen mid height, measuring the total circumferential opening displacement (TCOD). However, there is a proposed simplification of this method, using the acquisition of the axial displacement of the test machine instead of the TCOD [9]. This new configuration turns the test even simpler to perform and reduces the level of requirements for the test machine, turning the DPT even more attractive.

The *fib* Model Code 2010 allows the use of alternative tests to obtain the residual strength of FRC if appropriately previously correlated to the 3PBT [3]. Therefore, experimental researches have been performed to seek for correlations between flexural and double punch tests [10–13]. In that sense, the DPT reveals itself as a more promising alternative for a quality control test due to its simpler test procedure compared to closed loop tests, such as the 3PBT. A possible drawback of the DPT is the risk of negative influence of instabilities associated to the open loop testing machine, but recent studies demonstrate that it does not influence the determination of residual strength negatively [14].

All these technical features previously described present the DPT as a simpler solution for FRC systematic control. However, no previous research work was found addressing the repeatability and reproducibility conditions of the DPT. This subject is also important when aiming at the DPT implementation for regular quality control programs.

In this context, the objective of this study is to evaluate if the DPT is robust enough to repeat and reproduce average values and, also, variances. To do so, an experimental program was developed, conducted in one laboratory to evaluate repeatability and in six independent laboratories to evaluate the reproducibility.

2. Methodology

According to the ASTM E691, Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method, repeatability is the ability of repeating a single result in the same specimen, tested in the same location using the same equipment. From another perspective, reproducibility is the ability of reproducing a single result in the same specimen, tested in different locations using different equipment [15]. The problem is that both tests, DPT and 3PBT, are destructive, which restricts the use of the same specimen for the verification of repeatability and reproducibility. Even with the need to adapt this concept, several studies indicate that the ASTM E691 standard can be a tool to evaluate the confidence in tests results [16,17].

Another important aspect to be studied is the coefficient of variation (CV) obtained in each kind of test for the residual strength of FRC [18]. Aspects such as the specimen geometry, casting conditions and equipment precision become important parameters that influence the scatter of the test results and, consequently, the capability of repeat and reproduce results. With these aspects in mind, an experimental plan was developed with the objective of evaluating the repeatability and reproducibility conditions of the DPT and some of the aspects that can affect these conditions.

The experimental program was divided in two parts. The first part comprised a preliminary characterization of the repeatability of DPT. Three batches of FRC specimens were cast on different dates and tested at the age of 21 days in the same laboratory. The repeatability of the DPT results was evaluated by the comparison of average and variance values of crack and residual loads (F_{cr} and $F_{0.5}$; $F_{1.5}$; $F_{2.5}$; $F_{3.5}$) using analysis of variance (ANOVA) and the standard method ASTM E691 [15]. In addition, this evaluation allows the determination of the minimum number of test specimens required to properly plan part 2 of the study and ensure the reliability of the experiment.

The second part consisted in an interlaboratory comparative analysis of the reproducibility of DPT, incorporating conclusions from part 1. One batch of FRC for each fiber content was produced and tested with more than 180 days. Tests were performed in six different laboratories, with different testing machines, load capacity and data acquisition frequency. Specimens were distributed to laboratories to be tested according to a pre-established procedure. As in part 1, results of average and CV values of crack and residual loads obtained were compared by means of ANOVA and by the standard method ASTM E691 [15].

For both parts of the experimental program, all the specimens were produced under the same conditions at the Laboratory of Materials, Components and Construction Processes from the Polytechnic School of University of São Paulo. Two mixtures were designed using one concrete matrix and two fiber contents (30 kg/m^3 and 55 kg/m^3). These mixtures were identified as T30 and T55 for fiber contents of 30 kg/m^3 and 55 kg/m^3, respectively. Fig. 1 represents the experimental program schematically.

3. Part 1: Preliminary study (repeatability analysis)

3.1. Materials

The concrete matrix was composed by a Portland cement type CEM II/B-S 42.5 R, siliceous aggregates, tap water, a polycarboxylate type superplasticizer and Dramix 80/60 hooked-end steel fibers (length, diameter and aspect ratio equal to 60 mm, 0.75 mm and 80, respectively). The mix composition adopted is described in Table 1. Concrete presented a slump value equal to (100 ± 20) mm and an average compressive strength of 40 MPa at 28 days.

3.2. Mixing and casting procedures

A concrete mixer with a nominal capacity of 120 L (considered as a low energy rotation mixer) was used. The sequence of addition of the materials in the mixer was kept constant: first adding the coarse aggregate and 1/3 of total water, then quartz river sand, artificial sand and cement. After an initial homogenization of 2 min, the remaining water and the fibers were added. The mini-

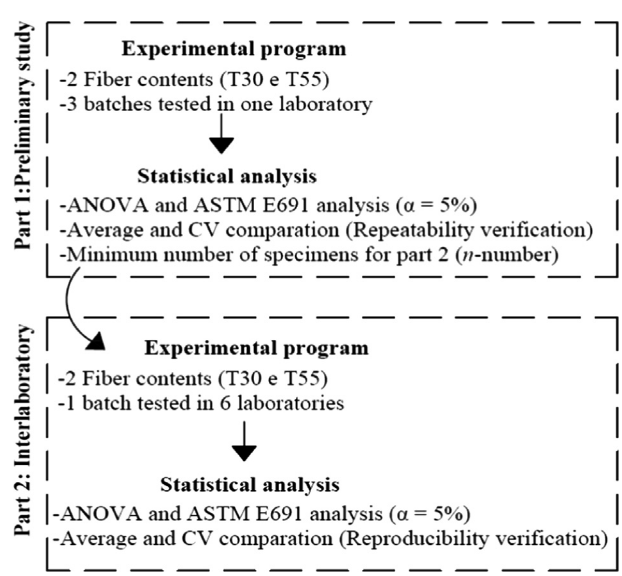

Fig. 1. Scheme of the experimental program conducted in this study.

Table 1
Concrete mix composition.

Material	Dosage (kg/m^3)
Cement	380
Quartz river sand	447
Artificial sand (crushed rock)	239
Coarse aggregate (d_{max} = 19 mm)	1061
Water	192
Superplasticizer	1.9
Steel fiber	30 / 55

mum mixing time was established as 10 min, after which slump was measured.

The first batch of concrete T55 did not present a good homogenization and fibers were not uniformly distributed in the matrix. Even with these problems, this batch was retained in the study in order to evaluate the effect of possible mixing problems in the process and if the DPT is sensitive to detect such a problem. The finishing quality of the specimen surface was considered adequate according to AENOR UNE 83515 recommendations [8].

Three batches of T30 and T55 were produced for the repeatability analysis. For each batch, three cylindrical specimens measuring 150 mm in diameter and 300 mm in height were cast for the DPT and two cylindrical specimens measuring 100 mm in diameter and 200 mm in height were cast for the determination of compressive strength. Specimens were consolidated using a vibrating table and cured in a humid chamber at the temperature of (23 ± 2) °C and humidity above 95%. Specimens destined to the DPT were sawn at mid-height at the age of 14 days, in order to obtain cylinders with height/diameter ratio equal to 1, according to the test standard [8]. After sawed, specimens returned to the humid chamber, to continue the curing process. For the first part of the program, all specimens (both for DPT and compressive strength) were tested at the age of 21 days.

3.3. Double punch test

Concrete cylindrical specimens measuring 150 mm in diameter and 150 mm in height were tested. Two load wedges measuring 37.5 mm in diameter and 30 mm in height, made of tempered steel hardness of 55 HRC, were placed at each face of the specimen (see Fig. 2a). As a simplification of the standard AENOR UNE 83515 [8], the axial displacement of the specimen (instead of the TCOD) was kept constant at 0.5 mm/min. Several other researches also use the DPT simplification to determine crack and residual loads for FRC [2,9,10,14]

This test provides a load–displacement curve, which is analyzed in five points, similarly to the procedure of the flexural test EN 14651 [4]. These points represent the cracking load resisted by the cementitious matrix (F_{cr}) and residual loads at the axial displacements of 0.5; 1.5; 2.5; 3.5 mm ($F_{0.5}$; $F_{1.5}$; $F_{2.5}$; $F_{3.5}$). These five points are the focus of the study, so the initial curve, before reaching F_{cr}, is discarded. Fig. 2b shows the final scheme of the curves, with the F_{cr} positioned at origin of the abscissa axis in softening behavior and in hardening behavior (2c). Notice that when hardening behavior occurs, the maximum load may or may not be the F_{cr}.

The test standard [8] does not fix a minimum test machine capacity. However, 200 kN is the minimum recommended because loads may easily exceed 100 kN during the test depending on the concrete compressive strength and fiber content [1,11]. The DPT may be performed in open loop machines, as is the case of all equipment used in this work. A data acquisition system capable of recording the axial displacements for each applied load must be used. The testing machine used for part 1 was a Shimadzu UH-2000 kN XR, operating in open loop, capacity of 2000 kN and data acquisition frequency of 20 Hz.

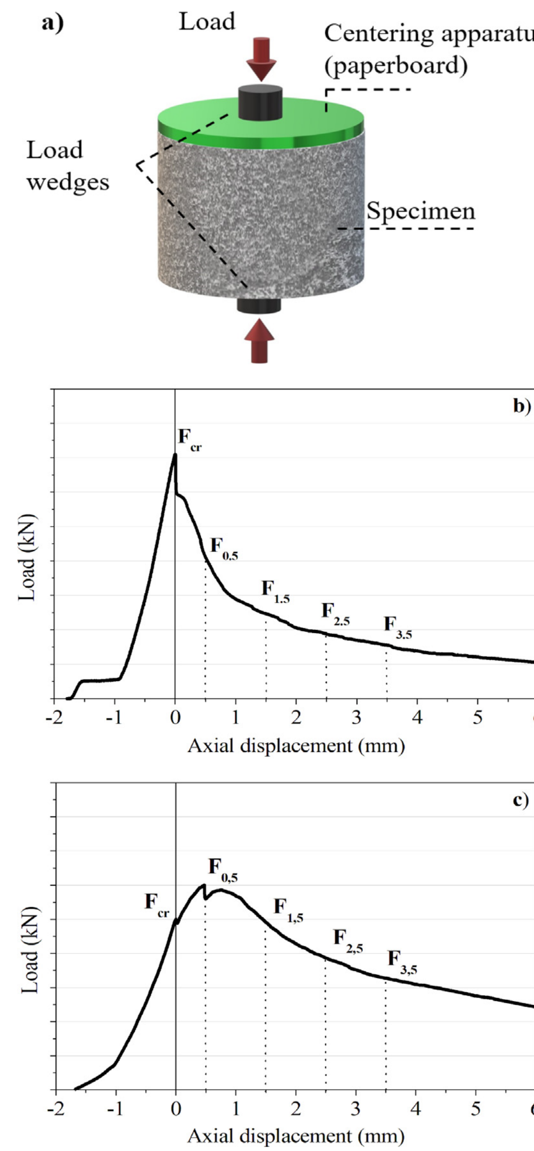

Fig. 2. Final set of DPT, with the centering apparatus (a); final load-axial displacement curve with F_{cr} dislocated to zero value in abscissa axis for softening (b) and hardening (c) behavior.

3.4. Methodology for statistical analysis

The methodology to analyze all results of crack load and residual loads followed a pre-determinate sequence: statistical Analysis of variance (ANOVA) tests, the methodology of ASTM E691 [15–17] and the determination for minimum sample size (number n) for planning part 2 of the study. It is important to emphasize that the standard AENOR UNE 83515 does not establish a minimum number of specimens, neither an acceptable standard deviation for the test [8]. Therefore, 6 specimens per mixture were adopted for part 1 of the study based on previous works [1,7].

The software Minitab® 17 was used to perform ANOVA tests, with a significance level (α) of 5%. The general linear model procedure was used and the hypothesis that the average value of several populations is equal was tested. The test requires a response or measurement taken from the units sampled and one or more factors. The aim was to check repeatability of average values of all five points of the load-axial displacement curve (F_{cr}; $F_{0.5}$; $F_{1.5}$; $F_{2.5}$; $F_{3.5}$). The *p*-values determined for a given variable shows whether the effect for that variable is significant or not. Thus, the influence of two variables in each of the five points was evaluated in terms of *p*: *Average* values among the batches and *Fiber Content* among the batches. In addition, the interaction of *Average* values and *Fiber Content* was evaluated.

If a specific point of the curve (F_{cr}; $F_{0.5}$; $F_{1.5}$; $F_{2.5}$; $F_{3.5}$) shows no repeatability of averages in ANOVA test, an in-depth analysis of this specific point was conducted. For this analysis, a Tukey test with a significance level (α) of 5% was performed in T30 and T55. In addition to the analysis of average values, the Levene test is employed to check hypotheses of equality of variances. This test establishes an initial hypothesis of equality of variances among batches 1, 2 and 3 (H_0). If the *p*-value for a test point is higher than 0.05, the variances may be considered equal among batches. If the *p*-value is lower than 0.05, H_0 should be rejected.

In order to complement the ANOVA results, the statistical analysis proposed by ASTM E691 [15] was conducted. Two dimensionless parameters were calculated: *h* and *k*. The parameter *h* is related to the average values of all results obtained, while *k* is related to the standard deviations results. After that, for each of the five points of the load-axial displacement, *h* and *k* are compared to tabulate critical parameters (h_{crt} and k_{crt}). To positively conclude about the repeatability of results, *h* and *k* must be lower than their respective critical values. The h_{crt} and k_{crt} values for the part 1 of the developed program are, respectively, 1.15 and 1.37. The equations to obtain the calculated and critical parameters are properly presented in the standard ASTM E691 [15].

The last statistical analysis was the determination of number *n* of samples. The objective is to plan the part 2 of the program, calculating the minimum sample size. Also, this methodology allows to check if six specimens were enough to ensure an adequate representation of the average and CV values in the results from part 1. To determine *n*, an average value of CV (CV_m) was calculated from de results, excluding the F_{cr} values, since this parameter is governed predominantly by the concrete matrix. Equation (1) was used to determine *n*, where σ is the standard deviation of the test, Z_y is the critical value found in the normal distribution table and equal to 1.64 (significance level α of 5%) and ε is the adopted error. This error is the value of CV_m that affects the average value of a residual test point ($F_{0.5}$; $F_{1.5}$; $F_{2.5}$; $F_{3.5}$). The equation assumes that the population follows a Gaussian distribution.

$$n = \frac{\sigma^2 \cdot Z_y^2}{\varepsilon^2} \tag{1}$$

3.5. Analysis of results

3.5.1. Characterization of the concrete matrix

Table 2 presents the results of concrete consistency (slump) and compressive strength, from the three FRC batches (part 1).

The average value of compressive strength obtained with part 1 was 44 MPa, considered satisfactory for reaching pre-established compressive strength of 40 MPa in less than 28 days. No significant difference in the compressive strength between the batches and fiber contents was noticed. The low variation of compressive strength values indicates that the influence of matrix characteristics on the FRC behavior was restricted, as expected [19,20]. In addition, it was possible to reach a slump of (100 ± 20) mm with a controlled level of polycarboxylate admixture (0.5%).

Table 2
Slump (mm) and compressive strength (MPa) of the three FRC batches analyzed.

Identification	Batch	Slump (mm)	f_{cm} (MPa)	CV (%)
T30	1	110	45.9	5.7
	2	95	42.1	3.4
	3	95	42.4	9.0
T55	1	115	45.6	3.3
	2	120	43.8	1.9
	3	100	41.8	3.2

3.5.2. Evaluation of DPT repeatability

Fig. 3 shows the load-axial displacement curves resulted from the batches T30 (Fig. 3a) and T55 (Fig. 3b). The thicker line indicates the average curve of batches 1, 2 and 3. Table 3 shows the average and CV results determined with cracking and residual loads. During the test, all specimens showed the expected cracking pattern, from two to four radial cracks in both punching faces, which usually reach full height of cylinder, accompanied by secondary cracks that do not open completely [7]. As an example of the cracking pattern achieved, in Fig. 4, it can be observed two specimens, one from T30 (Fig. 4a) and one from T55 (Fig. 4b), with similar cracking pattern.

Fig. 3a indicates that all specimens presented a softening behavior. In addition, the post-crack behavior of all batches was similar, and the average curves are very close to each other, which is a good indication of repeatability of averages values. Fig. 3b shows the load-axial displacement curves resulted from the three batches of T55 tested. 50% of the specimens showed a hardening behavior, which was already observed in previous studies [7,11].

Visually, the scatter of the results related to T55 is higher than T30 and the average curves are more separated (Fig. 3b). Clearly, at least one of the specimens presented a very distinguished pattern. This disparate result is associated to the batch that presented mix-

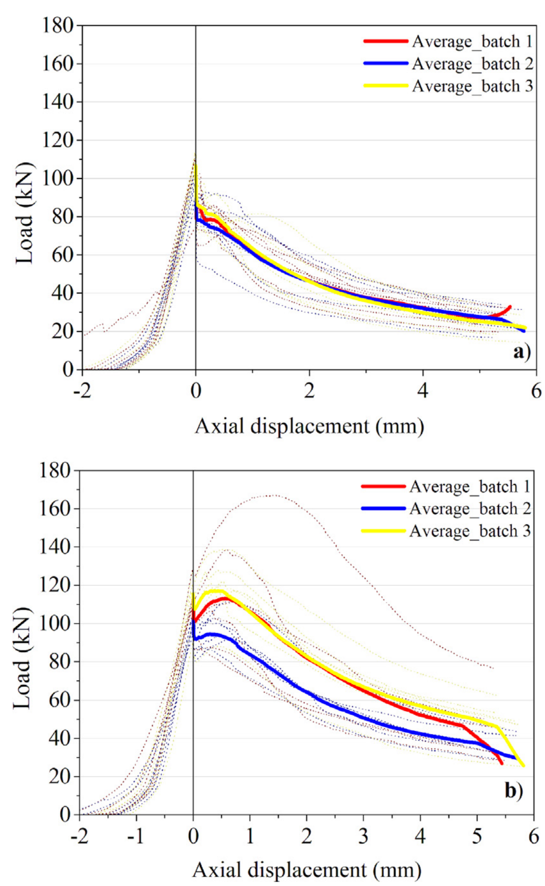

Fig. 3. Part 1: load-axial displacement curve from T30(a) and T55(b).

Table 3
Average results and CV from part 1.

Id.	Batch	F_{cr} Average (kN)	CV (%)	$F_{0.5}$ Average (kN)	CV (%)	$F_{1.5}$ Average (kN)	CV (%)	$F_{2.5}$ Average (kN)	CV (%)	$F_{3.5}$ Average (kN)	CV (%)
T30	1	104.4	14.3	72.6	11.0	53.0	21.0	41.7	20.5	35.1	20.5
	2	96.9	6.5	71.3	20.9	52.1	25.3	41.3	22.4	34.5	19.8
	3	106.7	8.0	76.1	10.1	53.7	27.0	40.2	27.0	32.6	27.4
T55	1	112.6	8.3	112.8	23.5	93.1	43.6	74.2	50.0	58.1	46.6
	2	102.9	5.6	93.7	8.1	72.3	9.6	55.4	10.9	45.6	14.1
	3	115.4	11.4	116.9	14.5	93.3	198	73.9	22.1	61.3	23.5

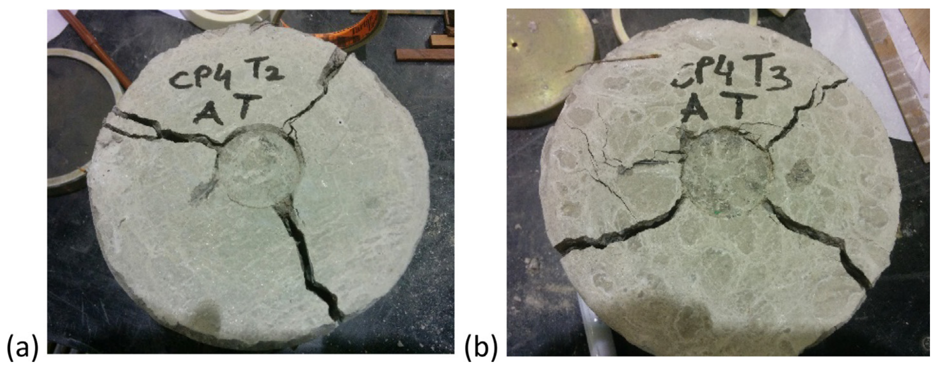

Fig. 4. Specimen from T30 (a) and specimen from T55 (b).

Table 4
Comparison in repeatability (equally of average values).

Variable	F_{cr}	$F_{0.5}$	$F_{1.5}$	$F_{2.5}$	$F_{3.5}$	
Average*Fiber Content	p-value	0.94	0.26	0.41	0.34	0.30
	Significant	No	No	No	No	No
Fiber Content	p-value	0.03	0.00	0.00	0.00	0.00
	Significant	Yes	Yes	Yes	Yes	Yes
Average values	p-value	0.03	0.08	0.33	0.37	0.40
	Significant	Yes	No	No	No	No

ing problems, which highlights the importance of mixing control to reduce the variability of the test. Nevertheless, none of the results was discarded in order to evaluate the influence of this kind of problem.

Results presented in Table 3 denote a wide region of uncertainty for the post-crack behavior of T55 concrete from batch 1, which reached a CV above 40% in $F_{1.5}$; $F_{2.5}$ and $F_{3.5}$. This is related to the poor fiber homogenization that was observed during concrete casting. This situation may have led to variations in the amount of incorporated fiber among specimens. This level of scatter showed that the casting process had an important influence on the total scatter [18].

As described in item 3.4, the first statistical analysis was ANOVA for equality of average values. The test is done by general linear mode, with α of 0.05. Table 4 summarizes the results.

Table 4 shows that the interaction between the two variables (Average among batches * Fiber Content) was not significant in the process (p greater than 0.05). Therefore, it is possible to analyze the variables separately. The variable *Fiber Content* presented p-values lower than 0.05 for all points. Therefore, *Fiber Content* is significant in the process and the test is totally sensitive to changes in fiber content, which was expected for DPT [1–2,7,9–13]. Finally, comparing *Average* values (from the five curve points) among batches, all residual load points presented p-values above 0.05, indicating that differences in *Averages* values among batches are not significant. Therefore, repeatability of averages was achieved.

An important fact to be highlighted is that the repeatability was achieved even with the batch considering mixing problems. On the one hand, this may demonstrate that the test method is robust to ensure repeatability of the parameters associated with the contribution of the fiber. On the other hand, this may indicate that the test method is not sensitive, in terms of average values results, to detect problems such as this occurred in the casting process.

The F_{cr} was the only one that did not show repeatability of averages, with p-value lower than 0.05. In order to have a better understanding of the F_{cr} behavior, an in-depth analysis was performed. Table 5 presents a paired Tukey test comparison for T30 and for T55.

From results of Table 5, is possible to state that, with a level of 95% confidence, all F_{cr} results from T30 can be grouped in B, while T55 can be grouped in A. Thus, even with the same concrete matrix, F_{cr} is only repeatable in the same fiber content. In other words, it is possible to control a group of F_{cr} results of the DPT if only one fiber content is employed.

Using the Levene test, the initial hypothesis (H_0) of equal variances was rejected for $F_{1.5}$ and $F_{2.5}$ (Table 6). Only F_{cr}, $F_{0.5}$ and $F_{3.5}$ showed repetition of variances, with p value greater than 0.05. Again, the high CV from batch 1 affected the analyses, compromising the repeatability of variances. In this sense, the DPT was sensi-

Table 5
Tukey comparison for F_{cr} in repeatability for part 1.

Identification	FRC batch	N° of specimens	F_{cr} average (kN)	Grouping*
T30	1	6	106.7	A-B
	2	6	104.4	B
	3	6	96.8	A-B
T55	1	6	115.4	A
	2	6	112.6	A-B
	3	6	103.0	A

*Average values that do not share a letter A or B are significantly different.

Table 6
Comparison in repeatability (equally of variances).

Point	p-value	Result
F_{cr}	0.96	Accept H_0
$F_{0.5}$	0.08	Accept H_0
$F_{1.5}$	0.047	Reject H_0
$F_{2.5}$	0.025	Reject H_0
$F_{3.5}$	0.10	Accept H_0

tive to detect mixing problems by increasing the dispersion of the results. Therefore, specifications that indicates maximum limits for variances will be useful for quality control process.

Fig. 5a shows the *h*-results of the ASTM E691 method [15]. The red lines indicate the critical value of 1.15 in modulus. Fig. 5b shows the *k*-results, method that evaluate the variations. The critical value for *k* is 1.37 and is also indicated with a red line.

Results presented in Fig. 5a shows that all h-values are lower than 1.15. Consequently, the repetition of average values between the batches was achieved, confirming the same conclusion obtained by the ANOVA method. For the k-values (Fig. 5b), the poor homogenization of the batch 1 in the T55 led to values above the critical (1.37) for F0.5; F1.5; F2.5 and F3.5. The results of this methodology also carry the same conclusions of ANOVA method for residual loads.

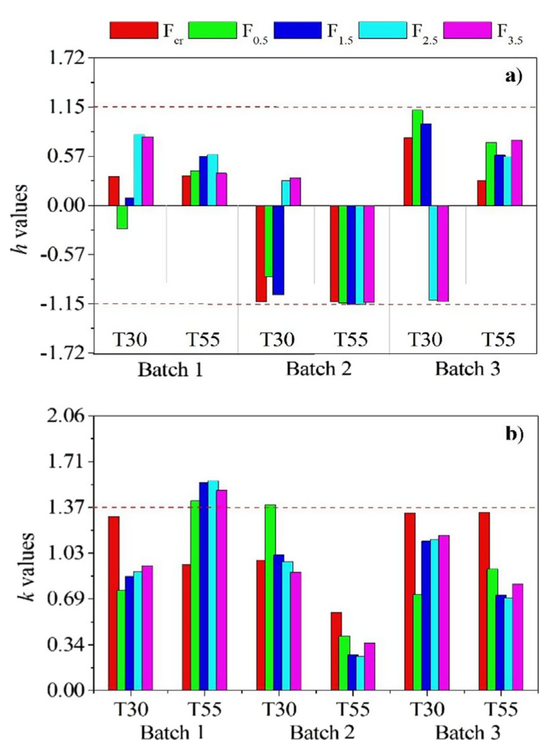

Fig. 5. Results of *h* (a); and *k* (b), for part 1.

Last statistical analysis aims to determinate number *n* by Eq (1). Results are detailed in Table 7.

Results from Table 7 show that only $F_{1.5}$, $F_{2.5}$ and $F_{3.5}$ of the batch 1 from T55 presented values above 6. This group is formed by the exact same specimens that had high scatter in results due to lack of homogeneity. All the other points in the three batches and fiber contents presented number *n* less than 4.

To summarize the conclusions of part 1, even when DPT tested is performed in a critical situation, it was able to repeat post-crack average values. F_{cr} is also repeatable when isolating fiber content in analysis. The poor homogenization of fibers is a possible scenario when mixing high steel fibers contents, and in this situation, affect directly the repetition of variance. For the post-crack values, the DPT was considered partially repeatable in part 1 (only averages values), but with high potential to fully repeat averages and variances if good homogenization is guaranteed.

4. Part 2: Interlaboratory study (reproducibility)

4.1. Materials

Part 2 of the program used the same FRC mixtures (T30 and T55), described in item 3.1.

4.2. Mixing and casting procedures

Mixing, casting, curing and sawing procedures were the same as the ones employed for part 1. A concrete mixer with a capacity equal to 400 L was used and one batch of T30 and T55 were produced. For each batch, 12 cylindrical specimens measuring 150 mm in diameter and 300 mm in height were cast for the DPT and 8 cylindrical specimens measuring 100 mm in diameter and 200 mm in height were cast for the determination of compressive strength. As in part 1, all specimens destined to the DPT were sawed at mid-height at the age of 14 days, before returning to the humid chamber. Each laboratory participating in the study received 4 specimens of each fiber content for the DPT (diameter and height equal to 150 mm)

The concrete casting process occurred in one single day and the specimens remained in curing chamber for more than 180 days before being distributed to all laboratories. This long age was derived to the fact that was difficult to arrange the same schedule for tests simultaneously with all six laboratories involved. Therefore, as the tests had to be performed within a week, the long age prevented the possibility of some significant variation of the characteristics of the material due to different maturity levels. Specimens were transported to the laboratories and stored inside the facilities, in dry environments, protected from moisture and sun. Specimens were kept stored in these situations for 50 days.

4.3. Participating laboratories

The DPT interlaboratory program was planned to involve 6 Brazilian laboratories located in São Paulo state. As the main objec-

Table 7
Minimum number of specimens (n) for CV_m of 22%

Identification	Batch	$F_{0.5}$	$F_{1.5}$	$F_{2.5}$	$F_{3.5}$
T30	1	0.6	2.2	2.2	2.3
	2	2.1	3.1	2.6	2.0
	3	0.6	3.8	3.6	3.5
T55	1	3.1	11.3	15.2	12.3
	2	0.3	0.3	0.6	0.7
	3	1.3	2.3	2.9	3.5

Table 8
Laboratories participating in the reproducibility experimental program.

Lab	Main activity	Machine brand	Machine capacity (kN)	Frequency of data acquisition (Hz)
1	Comercial	Kratos	300	0.4
2	Comercial	Emic	300	2
3	Comercial	Emic	1000	2
4	Comercial	Emic	200	2
5	Research	Instron	300	10
6 (USP)	Research	Shimadzu	2000	20

tive of the DPT is to suit as an FRC quality control test, the laboratories choices were guided aiming to select companies related with concrete systematic control in worksite. Also, two research laboratories where selected.

These laboratories are shown in Table 8 in conjunction with the main characteristics of their test machines. As DPT could be performed using an open loop test machine, the data acquisition frequency was not fixed [21]. As a consequence, there was a marked diversity of equipment involved in the study as a whole, which is a particularly interesting condition for assessing reproducibility.

As the participants (excluding USP laboratory) did not have any experience with the DPT, a technical instruction was created and previously distributed for each laboratory. The instruction contained a summary of the DPT method, establishing the same test procedure (described in item 3.3) for all laboratories. In addition, 6 centralization apparatus and 6 identical pairs of load wedges were produced and distributed to all laboratories, standardizing test conditions. In some laboratories, a meeting was necessary to solve all doubts about the DPT method.

4.4. Methodology for statistical analysis

The same approach used in part 1 was adopted to verify the reproducibility using the ANOVA model (equality of average values and variances), but with the different sample size, of four specimens per fiber content. For the ASTM E691 analysis, h_{crt} and k_{crt} correspond to 1.66 and 1.54 respectively, again due to the fact that a reduced number of specimens for each sample was used [15].

4.5. Analysis of results

4.5.1. Characterization of the concrete matrix

Results presented in Table 9 denote a higher slump value for the concrete produced to the interlaboratory study compared to the concrete from part 1. That fact could be associated to the better mixing process with the higher energy mixing machine. The average value of compressive strength obtained with part 2 was 47 MPa, considered satisfactory. Results from part 2 are higher than results from part 1 because of the age of testing, which is above 180 days.

Table 9
Slump (mm) and compressive strength (MPa), all tested in laboratory 6 (USP).

Identification	Slump (mm)	fc (MPa)	CV (%)
T30	130	44.3	4.6
		41.5	
T55	160	53.2	6.3
		48.6	

4.5.2. Evaluation of DPT reproducibility

Fig. 6 shows the load-axial displacement curves resulted from T30 (Fig. 6a) and T55 (Fig. 6b). Table 10 shows the average and CV results determined with cracking and residual loads. As occurred in the preliminary study, all specimens tested presented

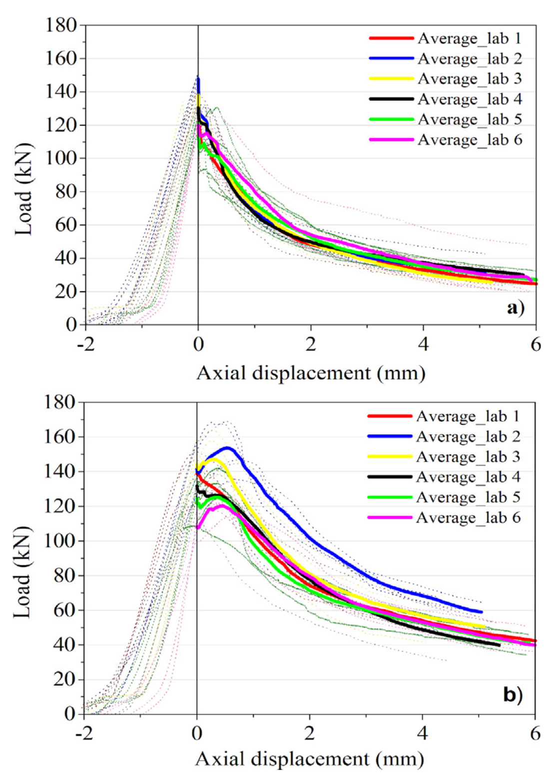

Fig. 6. Part 2: load-axial displacement curve from T30(a) and T55(b).

Table 10
Average results and CV from part 2.

Id.	Lab.	F_{cr} Average (kN)	CV (%)	$F_{0.5}$ Average (kN)	CV (%)	$F_{1.5}$ Average (kN)	CV (%)	$F_{2.5}$ Average (kN)	CV (%)	$F_{3.5}$ Average (kN)	CV (%)
T30	1	133.7	9.1	89.0	14.7	55.4	17.6	43.8	14.4	36.2	13.9
	2	147.8	3.0	92.0	7.6	57.4	9.5	44.8	15.7	37.7	21.2
	3	138.2	2.6	91.2	10.5	56.1	19.0	43.6	16.8	34.5	21.5
	4	130.5	3.9	88.1	12.7	56.0	9.3	45.4	6.3	42.2	8.9
	5	123.4	4.9	95.0	22.6	61.6	18.1	46.4	16.4	38.6	15.3
	6	119.4	4.1	91.0	9.7	56.9	3.6	42.6	8.4	33.5	10.2
T55	1	139.5	9.0	126.3	8.9	87.7	10.6	68.8	14.3	58.5	12.3
	2	141.9	12.0	153.5	11.3	117.8	6.3	90.7	4.1	72.9	5.3
	3	144.3	9.8	142.4	9.4	95.1	14.3	72.3	19.5	60.9	20.1
	4	131.8	4.4	125.0	18.3	91.7	26.8	68.0	25.3	53.6	24.9
	5	125.3	9.7	123.6	13.9	81.8	12.5	65.0	14.4	56.5	11.4
	6	108.3	6.4	119.8	4.6	91.2	14.1	68.9	9.5	57.2	11.9

Table 11
Comparison in reproducibility (equally of average values).

Variable		F_{cr}	$F_{0.5}$	$F_{1.5}$	$F_{2.5}$	$F_{3.5}$
Average*Fiber Content	*p-value*	0.43	0.12	0.06	0.09	0.13
	Significant	No	No	No	No	No
Fiber Content	*p-value*	0.88	0.00	0.00	0.00	0.00
	Significant	No	Yes	Yes	Yes	Yes
Average values	*p-value*	0.00	0.24	0.12	0.17	0.44
	Significant	Yes	No	No	No	No

Table 12
Tukey comparison for F_{cr} in reproducibility.

Identification	Laboratory	N° of specimens	F_{cr} average (kN)	Grouping*
T30	1	4	138.2	A-B
	2	4	147.8	A
	3	4	138.2	A-B
	4	4	130.5	A-B-C
	5	4	123.4	B-C
	6	4	120.1	B-C
T55	1	4	139.5	A-B
	2	4	141.9	A-B
	3	4	144.3	A-B
	4	4	131.8	A-B-C
	5	4	125.3	A-B-C
	6	4	108.3	C

*Average values that do not share a letter A, B or C are significantly different

the same expected cracking pattern (from two to four major cracks), observed in Fig. 4.

Fig. 6a shows the load-axial displacement curves resulted from the six laboratories for T30. The results showed a very similar pattern to the observed in part 1. The average curves are overlapping each other, which is a good indication of reproducibility of averages values. Fig. 6b shows the load-axial displacement curves resulted from the six laboratories for T55. 66% of the specimens showed a hardening behavior due to the high steel fiber content [7,11]. From Fig. 6b it is possible to see the average curves more separated than the T30 curves. Table 11 summarizes the results from general linear mode from ANOVA with α of 0.05.

Table 11 shows that the interaction between the two variables (Average value among laboratories*Fiber Content) was not significant in the process (p greater than 0.05), same conclusion from part 1, as expected for an FRC mechanical test. The variable *Fiber Content* resulted a p greater than 0.05 only for F_{cr}. This condition indicates that F_{cr} is not sensitive to fiber content changes when tested in different locations. In other words, no pattern was found, and it is possible to obtain wide variation of F_{cr} results using 30 kg/m^3 or 55 kg/m^3 of steel fiber. Comparing *Average values* among laboratories, all points related to the residual regimes presented values above 0.05, indicating that through different laboratories, differences in average values were not significant. This finding means that, by ANOVA, reproducibility of average values was achieved.

F_{cr} did not show reproducibility (p less than 0.05). A similar result was obtained in part 1. To have a better understanding of the F_{cr} behavior, Table 12 presents a paired Tukey test comparison with 95% reliability for T30 and for T55.

From results of Table 12, even isolating the two fiber contents, a pattern in results was not found, especially due to laboratory 2 that is isolated from T30 group and laboratory 6, isolated from T55 group. This outcome suggests that, with 95% confidence level, controlling F_{cr} with different testing machines can result in different values.

Table 13 show the Levene test for equality of variances. The test showed the similar conclusion from ANOVA and rejects the initial hypothesis of equal variances (H_0) only for F_{cr}, with p-value of

Table 13
Comparison in reproducibility (equally of variances).

Point	*p*-value	Result
F_{cr}	0.028	Reject H_0
$F_{0.5}$	0.49	Accept H_0
$F_{1.5}$	0.59	Accept H_0
$F_{2.5}$	0.6	Accept H_0
$F_{3.5}$	0.82	Accept H_0

0.028 (lower than 0.05). All post-crack points accepted H_0, showing equal variances, which indicates results accuracy [22]. This result contrasts with the first part demonstrating that DPT is able to control homogenization problems. That is, the variance not being maintained under conditions of uniformity can be considered as indicative of problems of homogenization.

Fig. 7a summarizes the results of *h*-values in the ASTM E691 method (*h* is related to the averages), with the red lines indicating 1.66 as the critical value in modulus. Fig. 7b shows the *k*-values (*k* is related to the variations), with the red lines indicating the critical value of 1.54.

According to the results from Fig. 7a, from 48 post-crack points analyzed for each laboratory and fiber content, 6% of them did not show reproducibility of average values ($F_{1.5}$; $F_{2.5}$; and $F_{3.5}$ from laboratory 2). Throughout the study, this was the highest percentage of error for average values in post-crack regime. The F_{cr} value of laboratory 6 was not reproducible, which was already observed in Tukey test. According to Fig. 7b, from *k*-results, only one post-crack result, out of 48 analyzed, did not show reproduction of variations ($F_{1.5}$ from laboratory 4). This is the lowest error occurred for equal variances in this study.

To summarize, the reproducibility experimental program was evaluated in a critical situation. It was tested in different locals, manipulated with different operators and tested in completely different machines, with limited number of specimens per fiber content. Even so, the DPT was able to reproduce averages and variances values in post-crack regime. A different conclusion can be made for the F_{cr}. This point did not seem to be able to reproduce averages and variances through different testing machines.

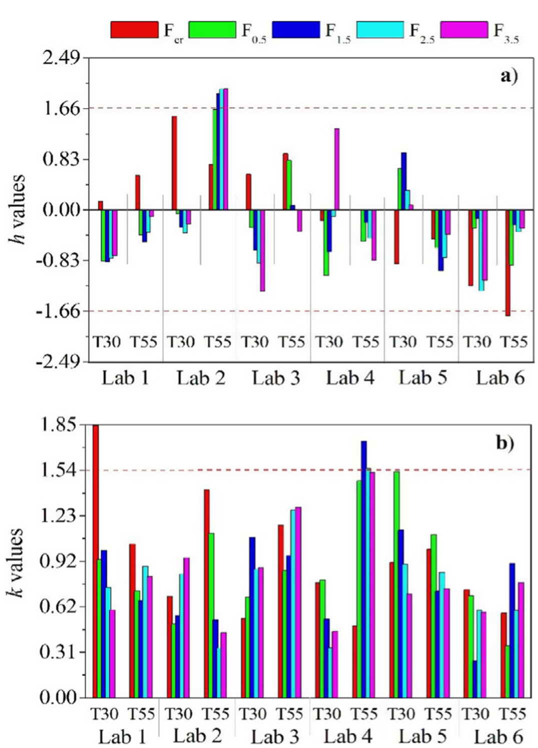

Fig. 7. Results of *h* (a); and *k* (b), for part 2.

5. Conclusions

A test would be easily considered reproducible and repeatable if it was not sufficiently sensitive to the variations in the materials properties. Therefore, one of the main conclusions achieved in this study is the fact that the DPT is sensitive to variations in the residual strength of FRC, which is of special interest for structural applications. Once the mixing process is a key factor to guarantee the proper homogenization of the composite, the DPT has proved to be capable to detect such inadequacy by means of increased dispersion of the results. However, once a good homogeneity of mixture is achieved, the equivalence of residual reference loads is obtained in terms of average and variance values using four specimens. In this sense, the DPT can serve as an instrument to control the homogeneity of the FRC if the condition of repeatability of the variance is verified. Therefore, for FRC structural applications, it is important to establish tolerances for the minimum average values and also for the variances of these results.

F_{cr} values is repeatable but it was not possible to fully reproduce F_{cr} by means of different testing machines. Two laboratories (2 and 6) had the equivalence of results rejected by the Tukey test. However, the fiber content has negligible influence in F_{cr}, which is related to the matrix characteristics. Therefore, the quality control of FRC for structural applications could be easily complemented through conventional compressive tests and the DPT could be used mainly for residual strength quality control.

No significant effect was observed due to differences in terms of testing machines, such as load capacity, and data acquisition frequency for post-crack mechanical properties determination. The experimental study focusing the reproducibility proves that DPT can fully reproduce residual loads in terms of average and variance with limited number of specimens. Therefore, the post-crack values of the DPT can be fully repeated and reproduced. Consequently, a systematic quality control programs of FRC is perfectly possible even in worksite. Therefore, the DPT test could be considered a robust test to be used as a reliable tool for FRC regular quality control in terms of post-crack behavior.

CRediT authorship contribution statement

André Baltazar Nogueira: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization, Project administration. **Luana de Carvalho Ribeiro Simão:** Investigation, Writing - review & editing. **Renata Monte:** Conceptualization, Methodology, Validation, Writing - review & editing. **Renan P. Salvador:** Conceptualization, Methodology, Validation, Validation, Visualization. **Antônio Domingues de Figueiredo:** Validation, Resources, Visualization, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Antonio D. de Figueiredo gratefully acknowledge the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq) for the support (Proc. N°: 305055/2019-4). Renata Monte would also like to acknowledge the financial support of the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq) Proc. N°: 437143/2018-0.

References

[1] C. Molins, A. Aguado, S. Saludes, Double Punch Test to control the energy dissipation in tension of FRC (Barcelona test), Mater Struct 42 (4) (2009) 415–425.
[2] R. Monte, A. de la Fuente, A.D. Figueiredo, A. Aguado, Barcelona Test as an Alternative Method to Control and Design Fiber-Reinforced Concrete Pipes, ACI. Struct J. 113 (2016) 6, https://doi.org/10.14359/51689018.
[3] Fédération Internationale DU Béton – FIB, Fib Model Code for Concrete Structures 2010, (2013).
[4] EN 14651, Test method for metallic fibre concrete – Measuring the flexural tensile strength (limit of proportionality (LOP), residual), (2007).
[5] M. Di Prisco, M. Colombo, D. Dozio, Fibre-reinforced concrete in fib Model Code, Principles, models and test validation, Struct. Concr. 14 (2013) (2010) 342–361, https://doi.org/10.1002/suco.201300021.
[6] W.F. Chen, Double Punch test for tensile strength of concrete, ACI Mater. J. 62 (1970) 993–995.
[7] S.C. Malatesta, A. Aguado, C. Molins, Generalization of the Barcelona test for toughness control FRC, Mater. Struct. 45 (2012) 1053–1069, https://doi.org/10.1617/s11527-011-9816-8.
[8] Asociación Española DE Normalización Y Certificación, AENOR 83515: Hormigones con fibras. Determinación de la resistencia a fisuración, tenacidad y resistencia residual a tracción. Método Barcelona, (2010).
[9] Pablo Pujadas Ana Blanco Sergio Cavalaro Albert de la Fuente Antonio Aguado NEW ANALYTICAL MODEL TO GENERALIZE THE BARCELONA TEST USING AXIAL DISPLACEMENT 19 2 2013 259 271 http://journals.vgtu.lt/index.php/JCEM/article/view/3835
[10] Isaac Galobardes Cesar L. Silva Antonio Figueiredo Sergio H.P. Cavalaro Chris I. Goodier Alternative quality control of steel fibre reinforced sprayed concrete (SFRSC) Construction and Building Materials 223 2019 1008 1015
[11] E. Galeote, A. Blanco, S.H.P. Cavalaro, A. de la Fuente, Correlation between the Barcelona test and the bending test in fibre reinforced concrete, Constr. Build. Mater. 152 (2017) 529–538.
[12] S. Carmona, C. Molins, A. Aguado, Correlation between bending test and Barcelona tests to determine FRC properties, Constr. Build. Mater. 181 (2018) 673–686.
[13] S. Carmona, C. Molins, Use of BCN test for controlling tension capacity of fiber reinforced shotcrete in mining works, Constr. Build. Mater. 198 (2019) 399–410.
[14] L. C.R. Simão, A.B. Nogueira, R. Monte, R.P. Salvador, A.D. de Figueiredo, Influence of the instability of the double punch test on the post-crack response of fiber-reinforced concrete, Constr. Build. Mater. 217 (2019) 185–192.
[15] AMERICAN STANDARD FOR TESTING MATERIALS, ASTM E691-16: Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method, (2016).
[16] J.M. Gudimettla, G.L. Crawford, J. Tanesi, A.A. Ardani, Interlaboratory Study and Precision Statement for AASHTO T 336 Test Method, Transp. Res. Rec. 2508 (1) (2015) 13–21.
[17] J. Mohammadi, W. South, Measuring Chromium in General Purpose (GP) Cement, J. Test. Eval. 45 (4) (2017) 20150457, https://doi.org/10.1520/JTE20150457.
[18] S.H.P. Cavalaro, A. Aguado, Intrinsic scatter of FRC: an alternative philosophy to estimate characteristic values, Mater Struct 48 (11) (2015) 3537–3555.
[19] A. Bentur, S. Mindess, Fibre reinforced cementitious composites, Lond. N.Y.: Mod. Concr. Technol. 2 (2007).
[20] D. R. Morgan, S. Mindess, L. Chen, Testing and Specifying Toughness for Fiber Reinforced Concrete and Shotcrete, In: Second. Univ.-Ind. Fiber Reinf. Concr. Other Adv. Mater. Tor. Can. Proc. (1995) 29-50.
[21] R. Gettu, B. Mobasher, S. Carmona, D.C. Jansen, Testing of concrete under closed-loop control, Adv. Cem. Based Mater. 3 (2) (1996) 54–71.
[22] A.H. Goaiz, T. Yu, M.N.S. Hadi, Quality Evaluation Tests for Tensile Strength of Reactive Powder Concrete, J. Mater. Civ. Eng. 30 (5) (2018) 04018070, https://doi.org/10.1061/(ASCE)MT.1943-5533.0002257.