What We Can and Cannot Learn from a Single Shear Test of a Very Large RC Beam

Houlin Xu[1]; A. Abdullah Dönmez[2]; Hoang T. Nguyen[3]; and Zdeněk P. Bažant, Hon.M.ASCE[4]

Abstract: In the existing database on shear load capacity, tests of very large beams are scarce. Valuable additions to the database have recently been made in 2021 at the University of California, Berkeley (UCB), and in 2015 at the University of Toronto. These two tests were the largest ever among the standard three-point-bend type tests conducted so far. They verified the effects of beam size and of steel stirrups on the ultimate load, V_u, provided that the same concrete and steel are used. The present analysis, which deals in detail only with the UCB test, shows that the subsequent public blind competitions to predict the V_u measured in both tests were meritless and potentially misleading. The reason is that, similar to design codes, the only information provided to the competitors (besides the E modulus) was the required concrete compression strength, f'_c, whereas the mean compressive and tensile strengths, fracture energy, initial creep data, and so on, were not provided. The fault of a competition of this kind is evidenced by (1) finite-element fracture simulations, (2) analysis of the huge statistical scatter of a database of 784 tests and a previous database in which f'_c was also the only concrete property used, like in the design code, and (3) estimation of the statistical error due to anchoring code provisions to the classical shear strength approximation $2\sqrt{f'_c}$ (psi), which was set at about 65% below the mean of the data cloud in the database. The winning prediction of the UCB competition had an error of only 2.7% of the measured failure load, even though the probability of success is here shown to have been between 0.14% and 8.46%, with 0.90% being the best estimate. Hence, competitions of this type are, in essence, a lottery. Furthermore, the fact that the winning predictions in both competitions happened to be obtained by cross-section strain analysis based on beam mechanics, and no fracture mechanics, is potentially misleading. This, of course, does not detract from the value of the UCB and Toronto experiments as important and unique additions to the database and as verifications of the load capacity for the particular concrete used. **DOI: 10.1061/JSENDH.STENG-12242.**

Author keywords: Beam shear failures; Concrete design code; Experimental scatter; Ill-informed prediction; Large scale tests; Prediction probability; Size effect; Strength databases; Structure strength.

Introduction

An important test of shear failure of a reinforced concrete beam was conducted in 2021 at the University of California, Berkeley (UCB), under the direction of Professor Jack Moehle, with an advisory committee consisting of Evan Bentz, Michael Collins, David Fields, Neil Hawkins, and Dominic Kelly (Zhai et al. 2022). The test was unique and important because, compared with a laboratory test, the beam was very large (total depth 3.56 m and span 21.34 m) and because cases without and with stirrups were compared. Only one similar beam of this kind (4.0 m deep), with and without stirrups, has been tested so far, at the University of Toronto (Collins et al. 2015). Both these beams represented thin slices or a very thick and wide one-way slab. Another very large beam, subjected to a uniform load, was tested in Tokyo (Iguro et al. 1984; Shioya and Akiyama 1994). Beams of depths up to 2.0 m failing in shear were tested by Bernat and García (2020) in Barcelona, Spain, and 1.89 m by Yoshida et al. (2000) in Toronto.

To determine the shear load capacity without and with stirrups from one and the same large beam, vertical stirrups were installed only in the left half-span (Fig. 1). On the first loading (Phase 1), the shear failure occurred on the right side having no stirrups. The loading was stopped right after the maximum load, and the damaged side still retained some integrity and suffered only partial damage. The damaged side with no stirrups was then retrofitted with straps consisting of high strength steel bars on the beam sides, so as to make the failed half-span much stronger than the other one. Furthermore, extra flexural reinforcement bars were added by means of grouting. This achieved a second loading (Phase 2) that caused failure in the left half-span with stirrups under a much higher load. What has been learned by the contractor is the load capacities of its very thick foundation slabs to be built using the same concrete and reinforcement.

A blind public competition for predicting the load capacities with and without stirrups was organized, with 59 participants. The winning prediction, whose errors in load capacity was +2.7% without stirrups and −9.8% with stirrups, was obtained by Lawrence Burkett, who used spreadsheet calculations to determine the shear force and moment distributions along the beam and then the computer program Response 2000 developed by E. Bentz at University of Toronto. Only the UCB competition is here analyzed in detail, although the conclusions will also apply to a competition of the same kind organized by E. Bentz after the Toronto test.

[1]Graduate Research Assistant, Dept. of Civil and Environmental Engineering, Northwestern Univ., Evanston, IL 60208. ORCID: https://orcid.org/0000-0001-8262-0294

[2]Visiting Assistant Professor, Dept. of Civil and Environmental Engineering, Northwestern Univ., Evanston, IL 60208. ORCID: https://orcid.org/0000-0002-2448-7090

[3]Postdoctoral Research Assistant, Dept. of Civil and Environmental Engineering, Northwestern Univ., Evanston, IL 60208.

[4]McCormick Institute Professor and W. P. Murphy Professor of Civil and Mechanical Engineering and Materials Science, Northwestern Univ., Evanston, IL 60208 (corresponding author). Email: z-bazant@northwestern.edu

Note. This manuscript was submitted on November 11, 2022; approved on March 27, 2023; published online on June 17, 2023. Discussion period open until November 17, 2023; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Structural Engineering***, © ASCE, ISSN 0733-9445.

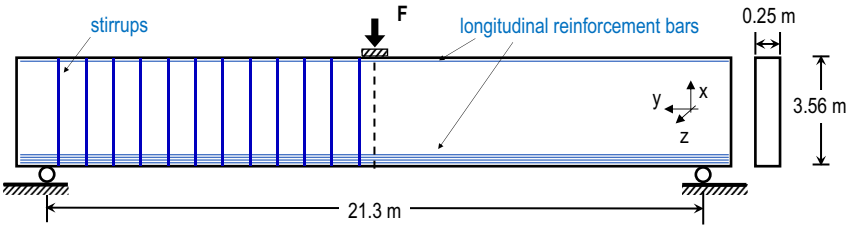

Fig. 1. The beam tested, with the dimensions.

The main objective of this study is to determine whether or not the method of the winning prediction is of any use for design code improvement.

Test Description

The specified compressive strength of concrete f_c' was 30MPa, from which the Young's modulus was estimated as $E = 25$ GPa. Coarse aggregate ASTM C33 #67 of maximum size 25 mm and fine aggregate ASTM C33 were used (ASTM 2003). The longitudinal reinforcement ratio is $\rho_w = A_s/b_w d$, where d is the effective depth, which is the distance from the top face to the centroid of the flexural reinforcement of area A_s. This longitudinal reinforcement ratio was 0.85% in the left half-span and 0.46% in the right half-span. The longitudinal reinforcing bars consisted of steel ASTM A1035 Grade 120 at the bottom and A1035 Grade 100 at the top (ASTM 2015). The stirrups of shear reinforcement in the left half-span consisted of steel No. 5 A615 Grade 60. In all simulations, the steel bars were assumed to be attached to the mesh nodes, with no slip (although slip between the mesh nodes was unhindered).

Missing Input Data and Questionable Aspects of the Competition

Similar to the previous Toronto competition (Collins et al. 2015), the following basic material parameters important for FE crack band analysis were not made available to the competitors:

$$q_1 = f_t, \qquad q_2 = G_f, \qquad q_3 = E_{\text{eff}}, \qquad q_4 = E_{\text{eq}}, \qquad q_5 = k_f \tag{1}$$

where f_t = mean tensile strength of concrete; G_f = material fracture energy; $E_{\text{eff}} = E/(1+\varphi)$ is the effective modulus, where φ is the short-time creep coefficient, which is not negligible for the 2-day duration of each loading and for an even longer duration of self-weight loading ($\varphi = 0.1$–0.2 as estimated here); E_{eq} = equivalent elastic modulus of the damaged half-span accounting for the strengthening effect of steel straps (its effect on the load capacity is negligible but not on the maximum deflection); and k_f = friction coefficient at the sliding support. All the parameters q_1, q_2, q_3, q_4, and q_5 can be measured by inexpensive standardized procedures (G_f according to the 1989 RILEM Recommendation TC89-FMT based on the size-effect tests feasible even in old testing machines without servocontrol) (Bažant et al. 2021; Bažant and Kazemi 1990; Bažant and Planas 2019; RILEM 1991).

Doubtless the reason these data in Eq. (1) were not measured was that the design code ACI 318 (ACI 2019) does not, and cannot, require them. The code must be universal, applicable to all concretes. However, the competition is not a design code. It necessitates these data because the prediction is to be made for one particular concrete, not for concretes in general.

The safety margin of the design code is much larger than one might infer from the load factors and the understrength factors. The total safety factor, defined as the failure load divided by the design load, was estimated to range from 1.7 to 8 for the RC beam shear [it was about 1.3 to 8 before the size effect was introduced into the code equation (ACI 2019)]. The reason is that the code equation was not set as the optimum mean fit of the cloud of worldwide database of the test results for different concretes (Reineck et al. 2014). Rather it was set at the lower margin of the data cloud, which was about 65% below the optimum mean fit [Fig. 1(a) in Bažant and Yu (2006)]. This amounts to a covert safety margin.

This margin is further complicated by the indirect size effect hidden in excessive dead load factors [discussed in detail by Bažant and Frangopol (2002)]. These factors are 1.4 or 1.2 in ACI (2019) depending on whether earth or water pressure, etc., is simultaneous with self-weight and 1.4 in the European Model Code (*fib* 2013), although in practice, the error in the weight of concrete cannot exceed 3% (except as a blunder). The excessive factor penalizes large beams compared with small beams because the self-weight of the latter is negligible. This amounts to a covert size-effect factor, which, however, does not have a realistic form.

Given that the design code equation taking into account only concrete strength gives the load capacities in the database with such an enormous error, how can the winning prediction, having the same limited information (concrete strength f_c' only) predict the UCB test result with an error of only 2.7%? Only by luck.

Can the Crack Band Analysis Fit the UCB Test Data?

Lacking relevant input data, we are forced to answer here an inverse question: can one find plausible and realistic values of input parameters q_1, q_2, q_3, q_4, and q_5 for which the finite-element (FE) fracture analysis would give close agreement with the test results? If one can, it will, of course, not suffice to validate the FE fracture analysis. But it will suffice to show that a poor FE prediction in this competition does not invalidate the FE fracture analysis, and that a negative conclusion about the FE would be unfounded.

Here, we answer this question by using the FE crack band model (Bažant and Oh 1983) with the microplane model M7 as the material constitutive damage law (Caner and Bažant 2013; Nguyen et al. 2021). This model came out superior by far when, in an extensive recent study (Bažant et al. 2022), eight computational models for fracture and damage, including the phase field, peridynamic and crack band (M7) models, were compared with a set of 11 distinctive fracture and failure experiments relevant to practice, all but one on concrete. In view of the success of the FE crack band model in these extensive comparisons, it would be quite disturbing if the FE crack band model were found incapable of fitting the UBC test.

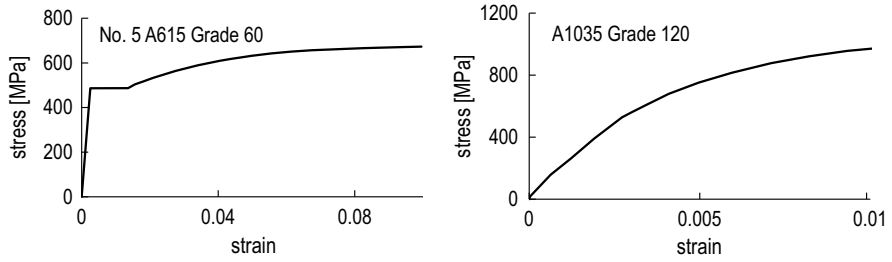

Fig. 2. Constitutive laws of steel materials used in simulations.

FE Simulations with M7 Crack Band Model

The stress–strain curves of the longitudinal bars and of stirrups are shown in Fig. 2. These curves, digitized, were directly used as steel properties in the FE code. The friction coefficient of the sliding of the beam over the end support was not reported, although it has a nonnegligible effect on the load capacity. It was considered in the computations as 0.03.

In the crack band model, the finite-element size, h, is not arbitrary. In the damage zone, h must be treated as a material property, the material characteristic length. It affects the results and represents a material fracture characteristic that defines the G_f. The simulation used $h = 30$ mm, which is reasonable for the maximum aggregate size of about 25 mm.

It is important to note that the loading in the test was quite slow, proceeding at a rate too low to ignore the effect of short-time creep of concrete. It took about 1 day to reach the maximum load in the right half-span (Phase 1), and about 2 days in the left half-span (Phase 2). The interval between these two phases was about 3 weeks, during which the creep under self-weight was increasing the deflection. This is to be compared with the typical initial (short-time) creep curve of concrete under sustained constant stress shown in Fig. 3(a) (the magnitude of this creep is not widely appreciated). The curve in Fig. 3(a) is based on extensive creep test data filtered for errors (Bažant and Jirásek 2018; Bažant et al. 2015; Rasoolinejad et al. 2018).

In view of the lack of detailed short-time creep data and precise load history, the only way to take creep at least approximately into account is by elastic analysis based on the effective modulus, $E_{\text{eff}}(t) = E_0/[1 + \varphi(t)] = 1/J(t)$, where E is the Young's elastic modulus of concrete, and $\varphi(t)$ is the creep coefficient representing the relative increase of strain after sudden load application under constant stress ($<f'_c$) sustained up to current time t. It is clear that the $E_{\text{eff}}(t)$ for loading durations of 1 or 2 days (representing the inverse of compliance J) is significantly smaller than the standard elastic modulus E_0, which corresponds roughly to a 10-min load duration. Because E_0 was not measured, it was estimated from f'_c using the ACI formula (ACI 2019), which gave $E_0 = 25.7$ GPa. Upon consideration of these facts, the effective elastic modulus for slow loading was taken as $E = 22.5$ GPa.

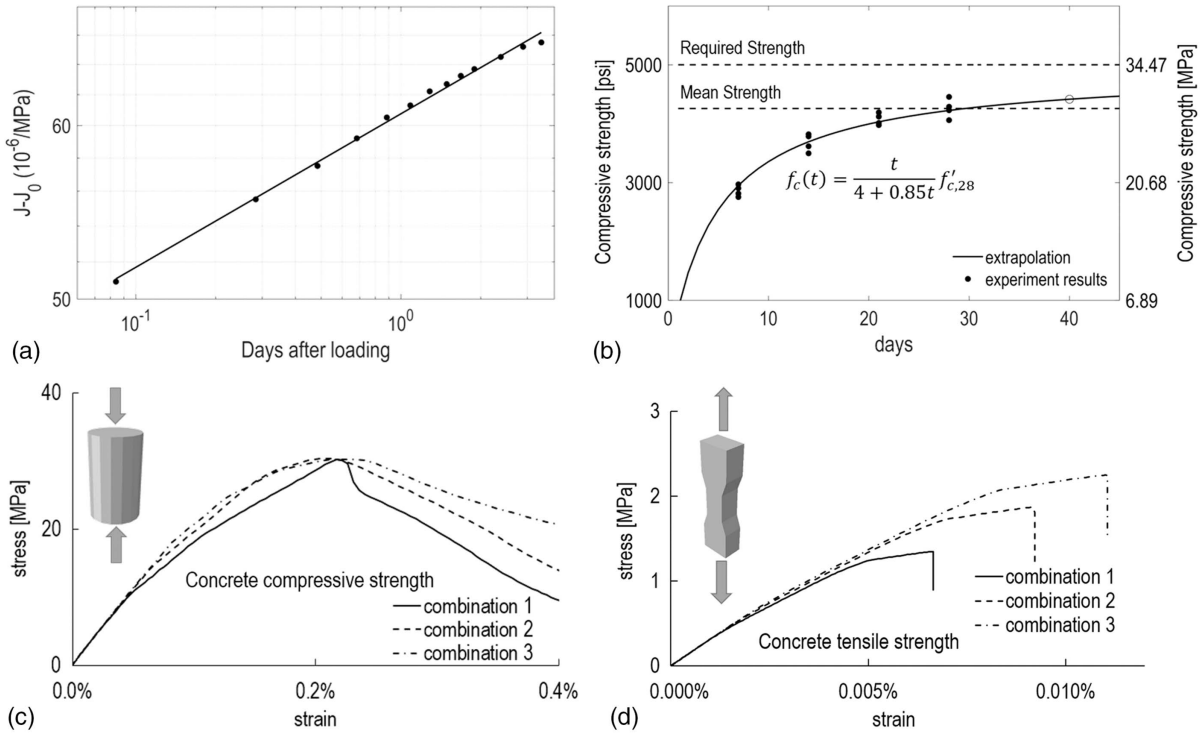

Fig. 3. (a) Compliance function for short-time creep curve (after error filtering) based on the test data of Müller et al. (1999); (b) evolution of compressive strength of concrete used in the test (1 psi = 6,894.76 Pa); and (c and d) simulation of concrete compressive and tensile tests based on three different calibrations.

Phase 2 presents even more difficult questions. The loading jack was unloaded, although the self-weight remained. The strengthening of the damaged right half-span with high-strength steel straps took, as it seems, up to a week. Then the beam was reloaded, which took again about 2 days. More creep occurred due to self-weight, and also due to reloading in Phase 2 after the pause in experiment. Even more uncertain is the stiffness of the retrofitted right half-span, in which the damage reduced the stiffness of the concrete per se, whereas the straps increased the stiffness significantly. As a rough guess, the effective elastic modulus of the right half-span, treated as an elastic body with no strength limit, was taken as $E = 25$ GPa, which made it stiffer overall than the left half-span. For predicting the maximum load in Phase 2, this guess is not important but plays some role, albeit minor, in predicting the maximum deflection in Phase 2, which was measured and should be matched. The left half-span underwent in Phase 2 further creep, and so its effective elastic modulus for Phase 2 analysis was reduced to 21 GPa. As expected, the Phase 2 loading produced shear failure in the left half-span with stirrups.

In the microplane model M7 (Caner and Bažant 2013; Nguyen et al. 2021), which was adopted as the material damage constitutive model for concrete, the free input parameters k_1, k_2, k_3, and k_4 control the compressive and tensile strength of the concrete. For a given mean compression strength f_c, the tensile strength f_t can vary considerably, and varying the input free parameters of M7 reflects that. For the same f_c, the values of fracture energy G_f can differ even more. The M7 coding and a sample FE crack band coding of the present analysis can be freely downloaded (Bažant 2023).

In the load-deflection diagrams of the beam shown in Figs. 3(c and d), the postpeak softening must be ignored because of the dynamic postpeak softening instability. A loading system stiff enough to prevent this instability would be difficult and costly to install for such a large beam, and it would even be unnecessary because only the load capacity is of interest.

Comparisons of FE Crack Band Predictions with Experiments

The FE crack band simulations of the tests used the ABAQUS Explicit 2022 finite-element code with microplane model M7 introduced through the user's subroutine VUMAT. The inevitable fictitious kinetic energy in the explicit simulations based on dynamic relaxation was kept negligible compared with the strain energy to ensure quasi-static response.

Due to the hydration of the concrete (Sakthivel et al. 2019), we extrapolated the compressive strength from given experimental results in Fig. 3(b). Among the three combinations of the M7 input parameters giving the compressive and tensile uniaxial stress–strain curves shown in Figs. 3(c and d), Combination 1 yielded good fits of the UCB test both in Phase 1 and Phase 2. So this is the combination that was used. Its parameters were $k_1 = 55 \times 10^{-6}$, $k_2 = 36$, $k_3 = 80$, and $k_4 = 90$, for which

$$E = 25 \text{ GPa}, \quad f_t = 1.34 \text{ MPa}, \quad G_f = 25.6 \text{ J/m}^2 \text{ (Combination 1)} \tag{2}$$

It led to close fits of the measured load capacities without and with stirrups. More than that, it also gave good fits of the measured maximum deflections [Figs. 4(a and b)]. The self-weight effect was automatically included in the FE analysis. Also, the simulations of Phases 1 and 2 should not follow the loading-unloading loop because in the experiment this loop must have eliminated the seating deformations at supports which surely distorted the initial loading [Figs. 4(a and b)].

The midspan deflections at the peak load in Phases I and II were predicted as 24 and 85 mm, respectively, which agreed closely with the observed deflection peaks. Fig. 5(a) further shows the predicted fracture patterns at maximum load. They all look realistic.

It appeared, however, that the same sample test data could also be fitted with two other, equally realistic, combinations of M7 material parameters. One was $k_1 = 76.5 \times 10^{-6}$, $k_2 = 70$, $k_3 = 30$, and $k_4 = 65$, for which

$$E = 25 \text{ GPa}, \quad f_t = 1.87 \text{ MPa}, \quad G_f = 44.9 \text{ J/m}^2 \text{ (Combination 2)} \tag{3}$$

and the other was $k_1 = 92 \times 10^{-6}$, $k_2 = 160$, $k_3 = 20$, and $k_4 = 65$, for which

$$E = 25 \text{ GPa}, \quad f_t = 2.25 \text{ MPa}, \quad G_f = 64.7 \text{ J/m}^2 \text{ (Combination 3)} \tag{4}$$

while the value of φ was kept the same. Importantly, the measured f_c value [Figs. 3(c and d)] was also matched for all three combinations. For the three parameter combinations, the FE crack band model with M7 delivered, respectively, the load capacities given in Table 1. From Fig. 5(b), it is also evident that the crack profiles are similar for all three combinations, even though the error of Combination 3 is large. Hence, it is insufficient to judge models only by the crack profiles.

Fig. 6 shows another interesting feature, namely, the calculated stresses in the stirrups are very nonuniform (although this

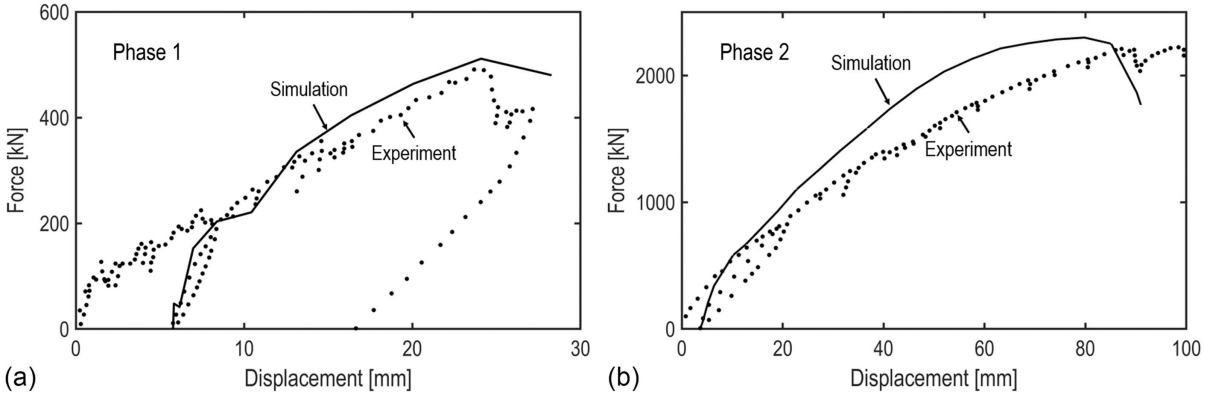

Fig. 4. Comparison of the load-displacement curve between the simulation (with M7) and the test result without and with stirrups.

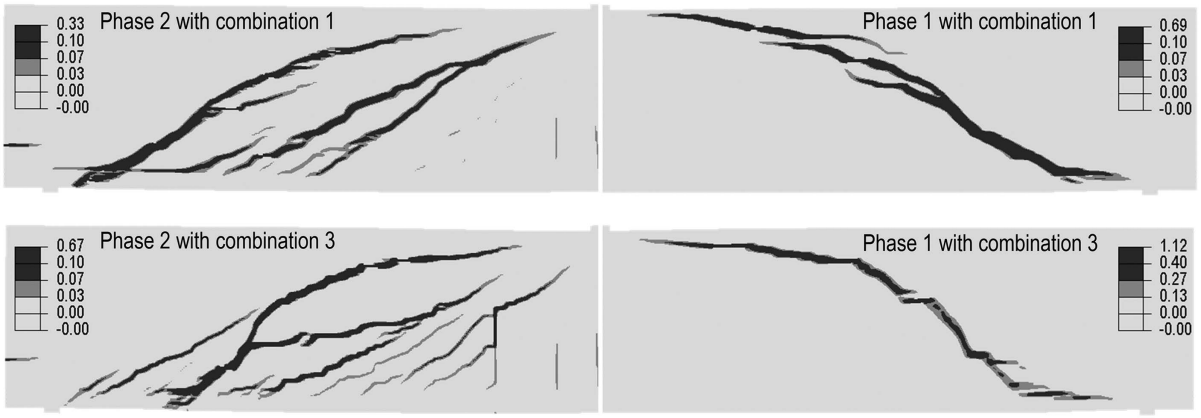

Fig. 5. Crack profile (logarithmic strain patterns) at the peak load without and with stirrups, resulting from Combinations 1 and 3.

Table 1. Peak load for different combinations in simulation

Simulations	No stirrups (kN)	Error (%)	With stirrups (kN)	Error (%)
Combination 1	511	3.4	2,298	3.1
Combination 2	673	36.2	2,455	10.1
Combination 3	884	78.9	2,529	13.5

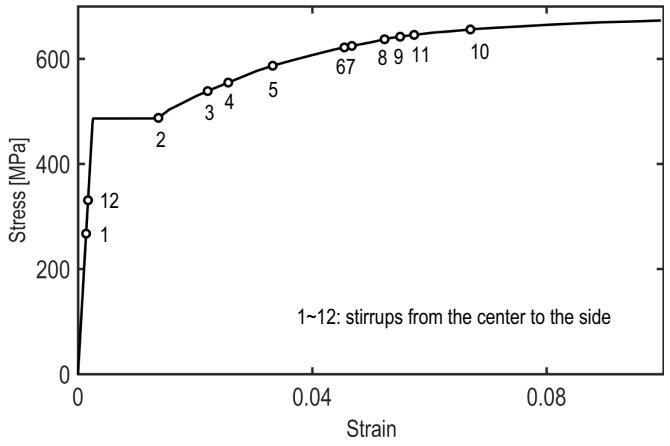

Fig. 6. Maximum stress values along each stirrup at the peak load.

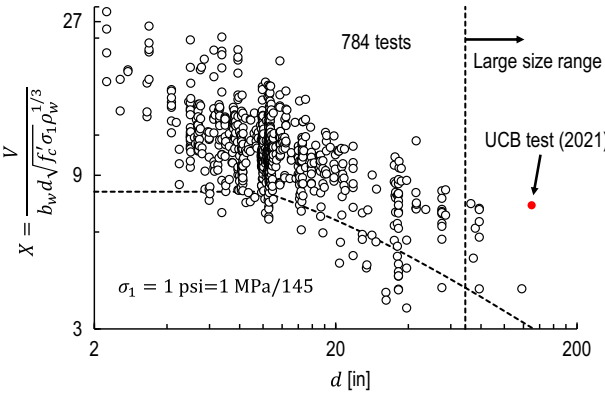

Fig. 7. Scatter of the shear strength exhibited from the ACI 445D database of 784 weighted points (1 in. = 25.4 mm).

nonuniformity cannot be verified because no strain gauges were bonded to the stirrups). Stirrups 1 and 12, which never yielded, were outside the range of the main diagonal crack. All the others were in the range of plastic hardening of stirrups. The limit analysis, underlying the design codes, assumes all the stirrups to have at maximum load the same stress, f_y, but this is not true here.

The foregoing results based on fracture mechanics are one-way to document the problem of misleading nature of ill-informed predictions based on incomplete material data, missing f_t, G_f, E_{eff} or φ, and E_{eq}. Another way to document this problem is by statistical analysis of the existing huge databases on beam shear failure, such as the database of ACI 445d/DAfStb (Reineck et al. 2014), which represents a modest expansion of the previous ACI 445F database with 784 data (Reineck et al. 2003; Fig. 7), which is, in turn, a major expansion of the original Northwestern University database of 296 data (Bažant and Kim 1984). The 2015 test at University of Toronto was not included in the ACI 445d/DAfStb database because this database was created earlier.

Statistical Estimation of Uncertainty of V When Only f_c Is Known

Estimation Based on Shear Stress Effect in ACI-ASCE Database without Size Effect

Fig. 8(a) shows in dimensionless form the classical experimental database of ACI-ASCE Committee 426 (MacGregor et al. 1997; ACI 1977), which describes how the cross-section average shear stress at maximum load, characterized by the dimensionless ordinate $V/(\sqrt{f'_c}\sigma_1 b_w d)$ in which $\sigma_1 = 1$ psi $= 1$ MPa/145, depends on the longitudinal reinforcement ratio $A_s/b_w d$. As usual, $\sqrt{f'_c}$ introduces the generally accepted approximate effect of f'_c on shear strength. This classical database involves no correction for the size effect of d, which means that the persistent narrow Europe–Japan–US disagreements about the proper form of the size-effect equation for design codes do not affect the data statistics.

The shaded narrow vertical strip in the figure shows a group of data points chosen so that the number of points was significant (25 in this case) and that the test parameters were relatively close those of the UCB test. The coefficient of variations of the ordinates of the points in this group is 22.4%. Because the probability

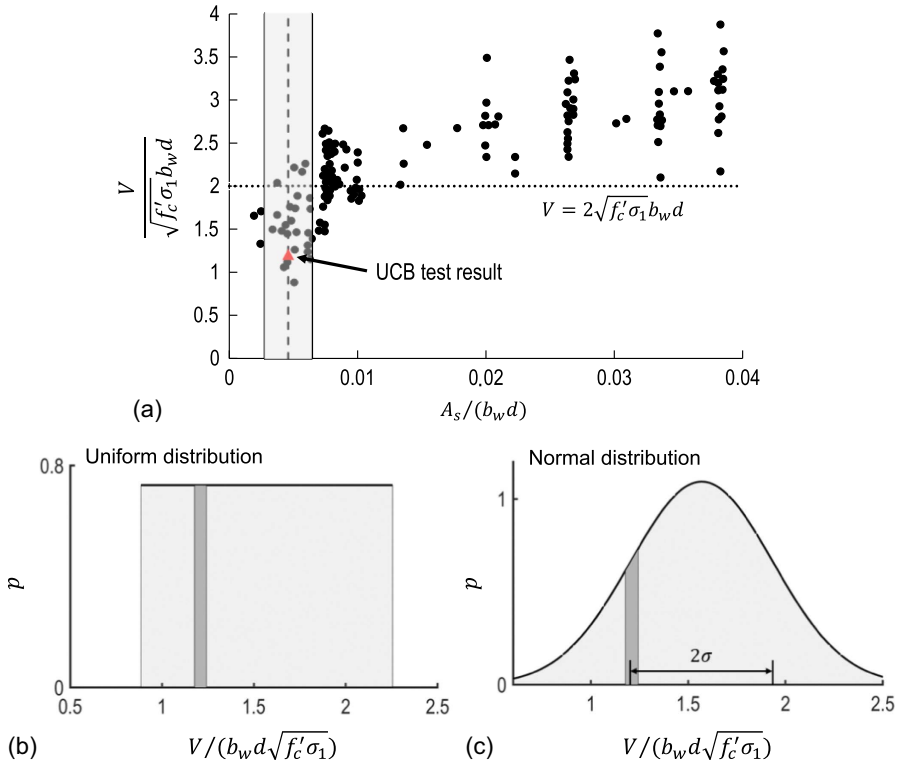

Fig. 8. (a) Relationship between the relative shear load capacity and the reinforcement ratio; and (b and c) probability of a good prediction when considering no size effect using a uniform distribution and normal distribution.

distribution is not known, we considered two extreme choices: (1) a uniform (rectangular) distribution between the margin points, and (2) the normal (or Gaussian) distribution with coefficient of variation 22.4%, as shown in Figs. 8(b and c). From the probability content of the strips of width ±2.7% in Figs. 8(b and c) that are centered at the ordinate corresponding to the UCB test, we infered that the probability of the wining prediction was as follows:

$$P = 4.77\% \quad \text{for uniform distribution}$$
$$P = 4.40\% \quad \text{for normal distribution} \tag{5}$$

So, our uncertainty about the type of probability distribution has almost no effect.

Estimation Based on d > 1.75 m from ACI 445F Database without Size-Effect Shift

Fig. 9(a) shows in double logarithmic scale the ACI 445F worldwide database of 784 tests [ACI Committee 318 (ACI 2019)] of beams of various sizes d, without stirrups, plotted as relative shear strength $X = V/V_0$ with $V_0 = 14.5 b_w d\sqrt{f_c'\sigma_1}\rho_w^{1/3}$, versus beam depth d. Here, V_0 has been determined by optimal fitting (or nonlinear regression) of the whole database with the energetic size-effect law embedded in the shear provisions of ACI 445/2019.

Because there is no statistical data basis for beams comparable to the UCB test, we must widen the scope by considering from the database the group of nine tests of depth >1.75 m marked in Fig. 9(a) by a vertical dashed line. We assumed that the coefficient of variation of this group, found to be $\omega_D = 22.0\%$, was approximately the same (Luo et al. 2021) as could be expected if the UCB test was repeated on different concretes. The probability analysis was, and must be, carried out in the linear scale even though the data in the figures are shown in the logarithmic scale.

First, let us ignore the size effect on the statistics. The group of nine data points does not suffice to determine the type of probability distribution. So we again assumed two extremes: (1) the rectangular distribution spanning the extreme points, and (2) the normal distribution, both shown in Figs. 9(b and c). For each of them, we located the shaded strips of width ±2.7% [Figs. 9(b and c)] that are centered at coordinate X_c representing the centroid of the group of nine data points, The centroid was found to be $X_c = 5.87$.

Now we want to find the probability of the prediction being within ±2.7% of the load capacity observed in the UCB test. The data do not suffice to decide the probability distribution function (PDF) of shear load capacity [which was analyzed in detail by Luo et al. (2021)]. The width of this interval (or strip) is $\Delta X = 2 \times 2.7\% V/(b_w d\sqrt{f_c'\sigma_1}\rho_w^{1/3})$ and its center lies in Fig. 7 at coordinate $X_c = V/(b_w d\sqrt{f_c'\sigma_1}\rho_w^{1/3})$. The probability of the winning prediction is thus estimated to be

$$P = 8.46\% \quad \text{for uniform distribution}$$
$$P = 6.78\% \quad \text{for normal distribution} \quad \text{(but size effect ignored)} \tag{6}$$

Estimation Based on d > 1.75 m from ACI 445F Database with Size Effect Considered

The size effect is an important factor. According to the energetic size-effect law (SEL), $V/V_0 = (1 + d/d_0)^{-1/2}$ [incorporated with a cutoff into ACI 318-2019 (ACI 2019)], we denote $V_0 = c \times b_w d (f_c'\sigma_1)^{1/2}\rho_w^{1/3}$, where $c = 14.5$ is a constant to be determined. A 3.56-m-deep beam that corresponds to the dimension of UCB test (2021) is marked in Fig. 9(a). For the data of large beams,

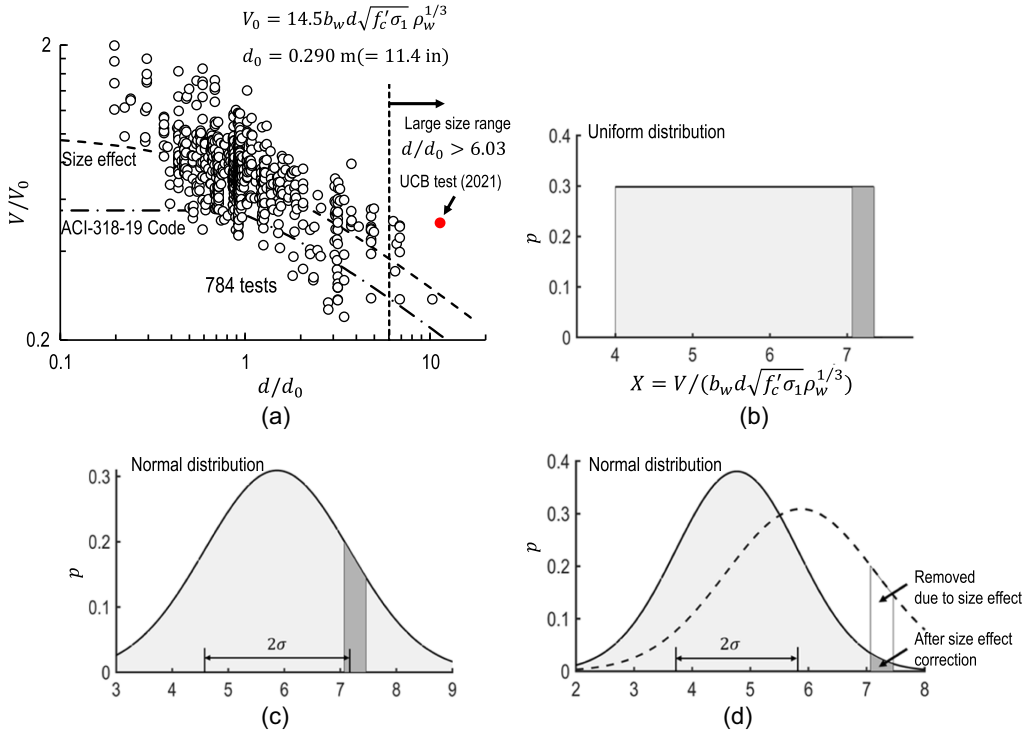

Fig. 9. (a) Size-effect law fitting of the ACI 445D database; (b and c) calculation for a good prediction using different probability distributions (uniform and normal); and (d) corrected calculation of probability of a good prediction due to size-effect law.

the theoretical V_{SE}/V_0 value, corresponding to the group of nine data of average size $\overline{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ ($n = 9, i = 1, 2, \ldots, 9$ numbers of data points), was obtained by means of the size-effect law

$$\xi_1 = V_{SE}/V_0 = (1 + \overline{d}/d_0)^{-1/2} \tag{7}$$

where the subscript SE labels the V calculated from the size-effect law. The mean of the group of nine data is

$$\xi_2 = \frac{1}{n}\sum_{i=1}^{n}(V_i/V_{0i}) \tag{8}$$

By comparing these two numbers, we can get the correction factor $\eta = \xi_2/\xi_1 = 1.15$ from the size-effect curve fitted to the original data. This factor is further applied as the size-effect correction for the UCB test of large beam. Thus, the mean is obtained

$$\overline{V/V_0} = \eta(1 + d/d_0)^{-1/2} = 0.33, \quad \text{where } d = 3.3 \text{ m } (=130 \text{ in.}) \tag{9}$$

The mean of the PDF will change accordingly in the normal distribution. In this *most realistic* estimation, the value for the probability of the winning estimate of load capacity [Fig. 9(d)] is thus found to be

$$P = 0.90\% \quad \text{(normal distribution corrected by SEL)} \tag{10}$$

One reason for this low value is the factor $\rho^{1/3}$ in the design code equation. The ρ value in the UCB test was unusually low for beams, which causes that this factor shifts the data point in Fig. 9(a) higher than it would lie for beams with a normal ρ-value. Doubtless, the reason for the low ρ value is that the UCB test was intended to represent a vertical slice of foundation footings, which normally have a much lower ρ than structural RC beams.

Among all the present estimates, $P = 0.90\%$ (Eq. 10), looks most realistic. However, because statistics are always debatable and often face many hidden obstacles [as recently elucidated by Nobel Prize winner Kahneman et al. (2021)], we present here all our estimates of the probability of winning prediction, which ranged from 0.14% to 8.46%. Whichever one the reader may prefer, it is clear that the winning prediction was obtained by some divine inspiration.

Nevertheless, this is also a point that makes the UCB competition (as well as the Toronto competition) valuable. It clarifies the big uncertainty in shear strength predictions based on insufficient data. But in no way can one draw any conclusions for the choice of size-effect equation for the design code ACI 318/2019 (Moehle 2019).

Role of Self-Weight in Cross-Section Models

The strong variation of shear force along the beam caused by big self-weight loading raises a further question about all predictions that did not use FE fracture analysis and evaluated V at some critical cross section (Bentz 2000) whose distance, x_V, from the support was either estimated or implied by some simplified argument. The critical x_V does not occur at the traditionally assumed distance d from the support.

The estimation of x_V played a big role in a previous similar competition to predict the shear capacity of a beam 4 m deep, without and with stirrups, tested at the University of Toronto in 2015. For a design equation using the size-effect law, it was observed that due to big self-weight, a small shift of x_V would change a poor prediction to a perfect one.

Conclusions

Based on the analyses conducted and results obtained, the following conclusions can be made:

1. A close FE prediction of failure load in a RC beam shear is impossible without knowing, in addition to f'_c, five relevant material parameters that vary widely from one concrete to another. They include the tensile and compressive strengths, material fracture energy, effective elastic modulus as reduced by short-time creep, and the mean (rather than the required) strength of concrete test cylinders.
2. The present FE analysis showed that there exist realistic values of the aforementioned five parameters that lead to a close fit of the data measured in the UCB test. This, of course, does not prove the validity of the FE fracture approach. But it shows that the UCB test does not disprove it either.
3. For the predictions lacking the relevant input data, the prediction competition is a matter of chance, i.e., the prediction competition is essentially a lottery. Nevertheless, the two prediction competitions have been useful in demonstrating (1) the enormous degree of uncertainty of load capacity predictions when only the required compression strength of concrete, f'_c, is specified, and (2) the importance of testing additional material parameters enabling realistic predictions.
4. The design code equation, as well as a simple strain-based cross-section analysis, can yield a close prediction of one individual test only by luck. The design equation has been set to give estimates that lie at the lower margin of the data cloud in a database covering all possible concretes (by contrast, the trends indicated by the form of this equation, e.g., the form of the equation describing the effect of the beam size or of the shear span, are what should closely agree with a series of individual tests of different sizes on one and the same concrete).
5. The huge scatter of the database also explains why the understrength factors stated in the code and further hidden understrength factors built into the code equation have to be so large.
6. Considering the ranges of several large databases that are relevant to the UCB test, and assuming that the test data within these ranges follow either the uniform distribution or the normal distribution, the probability of winning prediction according to various degrees of statistical simplification was found to range from 0.14% to 8.46% based on the ACI 445F database without size-effect shifts, 0.90% for the same but with size-effect shifts, and <5% for the classical ACI-ASCE shear database with no size effect. The most realistic estimate was 0.90%. Interestingly, the consideration of size effect reduces the probability of the winning prediction.
7. It further follows that the competitions held do not validate the method of analysis used for the winning predictions, which was the same for both UCB and Toronto tests.
8. Because the design code must apply to all concretes, and because thousands of designs have to be conducted on the basis of the required strength only, the precise type of concrete often is not, and cannot, be specified in advance. Consequently, large statistical variability of structural strength must, unfortunately, be accepted, which makes very large overall safety factors inevitable. This uncertainty can be reduced by a design based on realistic finite-element fracture analysis considering all the material parameters. However, ignoring the design code would deprive the designer from the legal protection that it provides.
9. In a beam as heavy as in the UCB contest, the variation of shear force along the beam due to self-weight is so strong that, in absence of FE fracture analysis, the choice of location of the critical cross section for sectional analysis is a big unknown. Knowing the correct location can greatly improve the prediction. This is another factor that makes the predictions based on simple cross-section analysis a matter of chance.
10. As a byproduct of the M7 FE crack band analysis, it was found that the stresses in stirrups at maximum load of the UCB beam are not uniform and not equal to yield limit f_y. Rather, they vary strongly from one stirrup to another. This is not what is assumed in design codes and represents an additional reason why the safety factors have to be so large. In view of this observation, strain gages should, in future tests, be mounted on the stirrups.
11. From a single test (or even a few tests), nothing can be inferred for the correctness of the design code. The code predicts, for given f'_c, the ultimate load value that is at the margin of a wide data band and, in the case of beam shear, equals roughly 35% of the data mean. Applying the understrength and load factors to that load value, one gets a design strength estimate that is, by experience, safe. But the load of failure probability of one in a million, which statistical experts consider safe, cannot be calculated from a failure load that corresponds to the data cloud margin, even if the type of probability distribution is known, because this distribution must be anchored in the mean (or in the regression curve) of the database.
12. The valuable contributions of the UCB test, as well as the Toronto test, were that they significantly improved the database and made blatantly clear the pitfalls of competitions based on the required concrete strength only. Further large-scale beam shear tests should be performed to enlarge the database.
13. To make future competitions for large-scale test prediction meaningful, it will be necessary to carry out and report to the competitors the tensile and compressive strengths, material fracture energy, and effective elastic modulus as reduced by short-time creep. All these tests are standardized (for fracture energy in RILEM TC89-FMT), and their cost is negligible compared with the large-beam test).

In summary, succinctly, the two big beam tests were valuable, but the competitions were misleading.

Broader Implications for CO_2 Emissions

The safety factors, over or covert, have consequences for CO_2 emissions. The worldwide CO_2 emissions from cement production are about to exceed those from all the cars and trucks in the world and are rising sharply. This problem is aggravated by both excessive and inadequate safety margins. To explain it, imagine hypothetically that all the structures are arranged, as in Fig. 10, in a

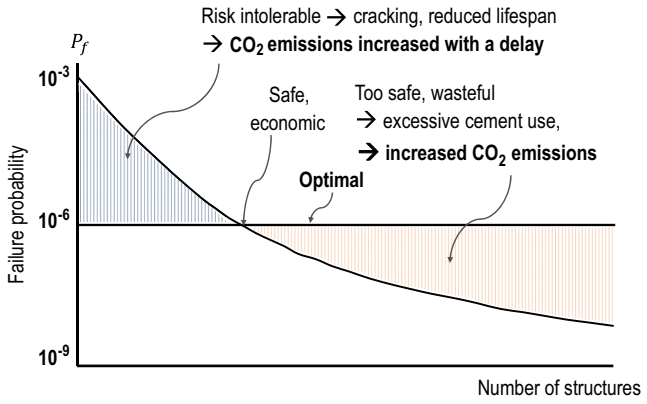

Fig. 10. Estimated failure probability versus the number of structures.

decreasing sequence of their failure probability (as if known). The excessively high safety factors highlighted by this study, giving unnecessarily low p_f, lead to an excessive use of concrete, and thus of cement. This increases CO_2 emissions.

However, excessively low safety factors also increase the emissions—in the long run. They lead to more cracking, which allows ingress of water and various corrosive ions. This impairs the durability and requires early replacement of the concrete structure.

If the lifetimes of concrete structures could be doubled, the environmental benefit would be enormous. For that goal, the safety factors and structural analysis methods matter.

Appendix. Alternative Method to Calculate the Probability of Prediction within ±2.7% of Test Result

Fig. 11(a) shows in double logarithmic scale the relative shear load capacities V/V_0 without stirrups as revealed by 784 tests assembled from the worldwide ACI 445F database (V is the shear force and V_0 is the normalizing constant) (Yu et al. 2016). The diameter of each data circle is proportional to the weight of each data point, which was obtained by dividing the range of log d into vertical strips of equal width $\Delta \log d$ and assigning the same combined weight to each strip. This weighting minimizes the statistical bias due to data crowding for small sizes and scarcity at large sizes.

To quantify scatter in the range corresponding to the UCB test, we mark in the figure a narrow vertical strip corresponding to the parameters of this test (we cannot consider a vertical line, i.e, a strip of zero width, because the chance of a test point lying on such a line would be negligible). The coefficient of variation of the ordinates of the weighted data points on this strip, corresponding to the case of no stirrups, has been evaluated as $\omega_D = 32.5\%$.

But ω_D is not a correct characteristic of scatter because the equation for the normalizing factors V_0 and d_0 of the ordinates and coordinates of data points is also uncertain. This equation is characterized by a certain coefficient of variation ω_0. It is obvious that random scatter of V_0 and d_0 must increase the coefficient of variation of the data on $y = V/V_0$ when plotted in these uncertain coordinates. So the ordinates y_i $(1, 2, \ldots, n_i)$ of the points in the vertical strip, characterized by a certain mean $\bar{y}$ and coefficient of variation ω_D, have been spread apart due to the uncertainty in V_0 and d_0. This uncertainty is characterized by further coefficients of variation ω_{V0} and ω_{d0}, and also by the difference $\Delta y = y_n - y_1$ between the highest and lowest points. Taking antilog of each coordinate, we find $y_n = 0.60$, $y_1 = 0.20$, and $\Delta y = 0.40$. For lack of information and for the sake of simplicity, we will assume that the distribution of the y-data in this strip are rectangular. This means that the probability density of y in this strip is uniform, equal to $1/\Delta y$, and the mean is $\bar{y} = (y_1 + y_n)/2 = 0.40$.

Aside from V_0 and d_0, one must also consider the random error of the form itself of the size-effect law, which was used by Bažant and Yu (2005) to construct the data plot in the figure. This error is characterized by a certain coefficient of variation ω_{SE}. Its value can be inferred from size-effect tests on one and the same concrete, and according to the numerous plots of beam shear tests made by Bažant and Yu (2005), we have, roughly, $\omega_{SE} = 5\%$. Furthermore, the uncertainty in function $\sqrt{f'_c}$ converting f_c to f_t should also be considered, being roughly characterized by a certain coefficient of variation $\omega_f = 5\%$.

As known from statistics (Benjamin and Cornell 2014; Bulmer 1979; Haldar and Mahadevan 2000), if the correlations of random variables are negligible, we have the approximation $\omega_D^2 = \omega_V^2 + \omega_{V_0}^2 + \omega_{d_0}^2 + \omega_{SE}^2 + \omega_f^2$. From this it follows that if the UCB test could be repeated for many different concretes with different f_c, the true coefficient of variation, ω_V, of the UCB shear load capacity in the UCB test would be approximately

$$\omega_V = (\omega_D^2 - \omega_{V_0}^2 - \omega_{d_0}^2 - \omega_{SE}^2 - \omega_f^2)^{1/2} \tag{11}$$

The coefficient of variation ω_{V0} characterizes the uncertainty in the form of the dependencies of V_0 on the shear span a/d (especially in the exponent 0.7) and on f'_c [including uncertainty of exponent $1/2$, which is considered as $1/3$ in *fib* Model Code (*fib* 2013)]. Approximately $\omega_{V0} = 0.10$. The ω_{d0} characterizes the uncertainty in the form of the dependence of transitional size d_0 on f'_c (especially in exponent -0.7). Approximately again, $\omega_{d0} = 0.10$ (the values of d, d_0, f'_c, and a/d are different for different data points on the vertical strip). Based on these considerations, Eq. (7) yields, for the case of no stirrups, the following estimate:

$$\omega_V = (0.325^2 - 0.10^2 - 0.10^2 - 0.05^2 - 0.05^2)^{1/2} = 28.4\% \tag{12}$$

This characterizes the error of shear force when only f_c is known, and after the bias in the coordinates due to randomness of parameters V_0 and d_0 has been eliminated. This decrease in

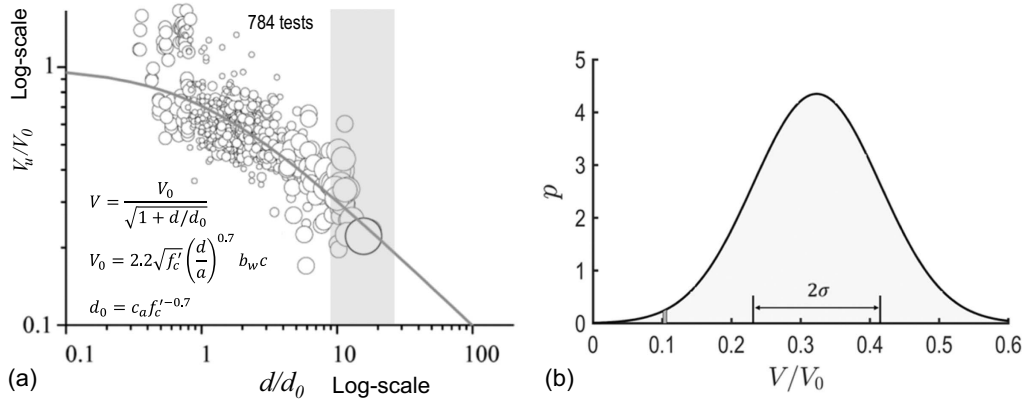

Fig. 11. (a) Scatter of shear strength in ACI 445D database of 784 tests' weighted points, weighted to counter sparsity and normalized for known trends (reproduced, curves are unrelated); and (b) calculation for a good prediction using normal probability distributions.

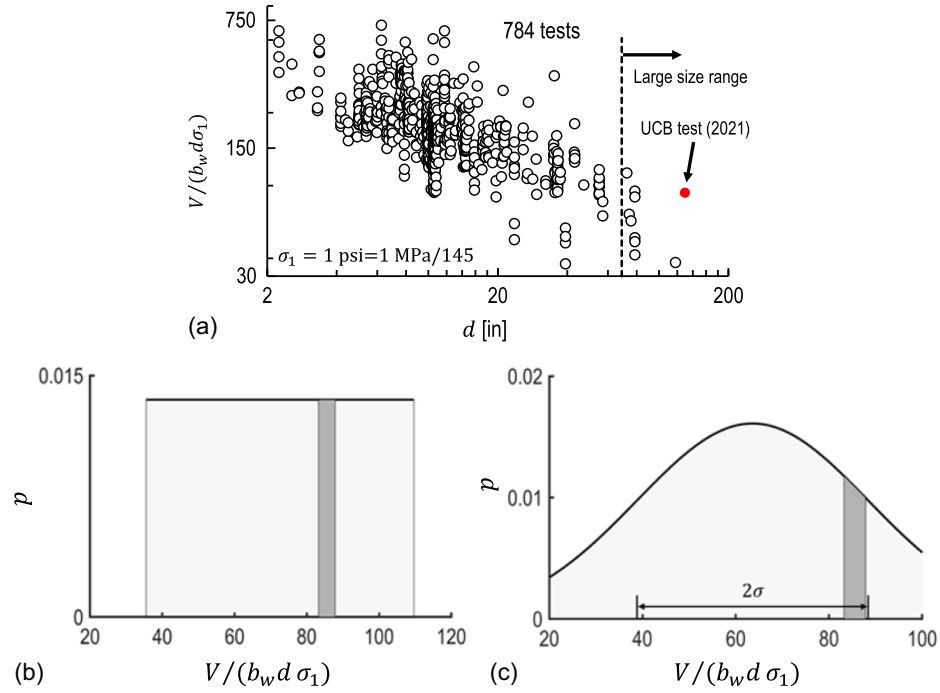

Fig. 12. (a) Scatter of shear strength in ACI 445D database of 784 tests' weighted points, normalized in a different way (1 in. = 25.4 mm and 1 psi = 1/145 MPa); and (b and c) calculation for a good prediction using uniform and normal probability distributions.

coefficient of variation means that the spread of the rectangular distribution must be reduced from Δy to $\Delta y' = (0.284/0.325)\,\Delta y = 0.874\Delta y$.

If we say a good prediction is within $\pm 2.7\%$ of the measured maximum (or ultimate) load V_u, we may again consider a uniform probability distribution of V_u/V_0 (normalized value) within the interval $(-0.027, 0.027)$ V_u/V_0 centered at the measured V_u/V_0. The probability content of this interval is equal to the fraction of this interval within the spread $\Delta y'$ of the unbiased rectangular probability distribution of V_0/V_u when only f_c known, which is $(0.054V_u/V_0)/\Delta y'$. This is the probability of a good prediction. But after calculation, we find that the UCB test is out of the range of the existing database ($P = 0$), which means that the uniform distribution is not appllicable in this situation. Furthermore, if we consider the distribution of V_u/V_0 as a normal distribution, then the probability of a good prediction will be changed to 0.14%, as already shown in Fig. 11(b).

If we choose to use the relative shear load capacities $V/b_w d$ without stirrups, as revealed in double logarithmic scale by 784 tests assembled from the worldwide ACI 445F database and shown in Fig. 12(a), then the probability will change accordingly, as shown in Figs. 11(b and c). Then, if the deterministic part of the size effect is ignored

$$\begin{aligned} P &= 6.23\% \quad \text{for uniform distribution} \\ P &= 5.02\% \quad \text{for normal distribution} \end{aligned} \tag{13}$$

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

Partial financial support under NSF Grant No. CMMI-202964 and ARO Grant No. W911NF-19-1-003, both to Northwestern University, is gratefully acknowledged. Dan Frangopol, Jia-Liang Le, and Wen Luo are thanked for valuable comments.

References

ACI (American Concrete Institute). 1977. *Suggested revisions to shear provisions for building codes*. ACI-ASCE Committee 426. Farmington Hills, MI: ACI.

ACI (American Concrete Institute). 2019. *Building code requirements for structural concrete: (ACI 318-19); and commentary.* ACI 318R-19. Farmington Hills, MI: ACI.

ASTM. 2003. *Specification for concrete aggregates*. ASTM C33. West Conshohocken, PA: ASTM. https://doi.org/10.1520/C0033-03.

ASTM. 2015. *Standard specification for deformed and plain, low-carbon, chromium, steel bars for concrete reinforcement*. ASTM A1035/A1035M-15. West Conshohocken, PA: ASTM.

Bažant, Z. P. 2023. "Free coding of explicit microplane constitutive model M7 for concrete (with extensions to fiber concrete and fatigue loading)." Accessed May 30, 2023. http://www.civil.northwestern.edu/people/bazant/.

Bažant, Z. P., A. Dönmez, E. Masoero, and S. R. Aghdam. 2015. "Interaction of concrete creep, shrinkage and swelling with water, hydration, and damage: Nano-macro-chemo." In *Proc.*, *CONCREEP 10*, 1–12. Reston, VA: ASCE. https://doi.org/doi:10.1061/9780784479346.001.

Bažant, Z. P., and D. M. Frangopol. 2002. "Size effect hidden in excessive dead load factor." *J. Struct. Eng.* 128 (1): 80–86. https://doi.org/10.1061/(ASCE)0733-9445(2002)128:1(80).

Bažant, Z. P., and M. Jirásek. 2018. *Creep and hygrothermal effects in concrete structures*. Dordrecht, Netherlands: Springer.

Bažant, Z. P., and M. T. Kazemi. 1990. "Determination of fracture energy, process zone length and brittleness number from size effect, with application to rock and concrete." *Int. J. Fract.* 44 (2): 111–131.

Bažant, Z. P., and J. K. Kim. 1984. "Size effect in shear failure of longitudinally reinforced beams." *J. Am. Concr. Inst.* 81 (5): 456–468.
Bažant, Z. P., J.-L. Le, and M. Salviato. 2021. *Quasibrittle fracture mechanics and size effect: A first course*. Oxford, UK: Oxford University Press.
Bažant, Z. P., H. T. Nguyen, and A. Abdullah Dönmez. 2022. "Critical comparison of phase-field, peridynamics, and crack band model M7 in light of gap test and classical fracture tests." *J. Appl. Mech.* 89 (6): 061008. https://doi.org/10.1115/1.4054221.
Bažant, Z. P., and B. H. Oh. 1983. "Crack band theory for fracture of concrete." *Matér. Constr.* 16 (3): 155–177. https://doi.org/10.1007/BF02486267.
Bažant, Z. P., and J. Planas. 2019. *Fracture and size effect in concrete and other quasibrittle materials*. Boca Raton, FL: Routledge.
Bažant, Z. P., and Q. Yu. 2005. "Designing against size effect on shear strength of reinforced concrete beams without stirrups: II. Verification and calibration." *J. Struct. Eng.* 131 (12): 1886–1897. https://doi.org/10.1061/(ASCE)0733-9445(2005)131:12(1886).
Bažant, Z. P., and Q. Yu. 2006. "Reliability, brittleness, covert understrength factors, and fringe formulas in concrete design codes." *J. Struct. Eng.* 132 (1): 3–12. https://doi.org/10.1061/(ASCE)0733-9445(2006)132:1(3).
Benjamin, J. R., and C. A. Cornell. 2014. *Probability, statistics, and decision for civil engineers*. New York: Courier Corporation.
Bentz, E. C. 2000. *Sectional analysis of reinforced concrete members*. Toronto: Univ. of Toronto.
Bernat, R. M., and J. M. B. García. 2020. *Project HORVITAL (No. BIA201564672-C4-1-R)*. Barcelona, Spain: Univ. Poliecnica Catalunya.
Bulmer, M. G. 1979. *Principles of statistics*. North Chelmsford, MA: Courier Corporation.
Caner, F. C., and Z. P. Bažant. 2013. "Microplane model M7 for plain concrete. I: Formulation." *J. Eng. Mech.* 139 (12): 1714–1723. https://doi.org/10.1061/(asce)em.1943-7889.0000570.
Collins, M. P., E. C. Bentz, P. T. Quach, and G. T. Proestos. 2015. "The challenge of predicting the shear strength of very thick slabs." *Concr. Int.* 37 (11): 29–37.
fib (Fédération Internationale du Béton). 2013. *Fib model code for concrete structures 2010*. Berlin: Wilhelm Ernst & Sohn Verlag fur Architektur und Technische.
Haldar, A., and S. Mahadevan. 2000. *Probability, reliability, and statistical methods in engineering design*. New York: Wiley.
Iguro, M., T. Shioya, Y. Nojiri, and H. Akiyama. 1984. "Experimental studies on shear strength of large reinforced concrete beams under uniformly distributed load." *Doboku Gakkai Ronbunshu* 1984 (348): 175–184. https://doi.org/10.2208/jscej.1984.348_175.
Kahneman, D., O. Sibony, and C. R. Sunstein. 2021. *Noise: A flaw in human judgment*. New York: Little, Brown.
Luo, W., J.-L. Le, M. Rasoolinejad, and Z. P. Bažant. 2021. "Coefficient of variation of shear strength of RC beams and size effect." *J. Eng. Mech.* 39 (2): 1–10. https://doi.org/10.3901/JME.2003.02.001.
MacGregor, J. G., J. K. Wight, S. Teng, and P. Irawan. 1997. *Reinforced concrete: Mechanics and design*. Hoboken, NJ: Prentice Hall.
Moehle, J. P. 2019. "Key changes in the 2019 edition of the ACI building code (ACI 318-19)." *Concr. Int.* 41 (8): 21–27.
Müller, H. S., Z. P. Bažant, and C. H. Kuttner. 2013. *Data base on creep and shrinkage tests*. RILEM Subcommittee 5 report RILEM TC 107-CSP. Champs-sur-Marne, France: RILEM.
Nguyen, H. T., F. C. Caner, and Z. P. Bažant. 2021. "Conversion of explicit microplane model with boundaries to a constitutive subroutine for implicit finite element programs." *Int. J. Numer. Methods Eng.* 122 (6): 1563–1577. https://doi.org/10.1002/nme.6590.
Rasoolinejad, M., S. Rahimi-Aghdam, and Z. P. Bažant. 2018. "Statistical filtering of useful concrete creep data from imperfect laboratory tests." *Mater. Struct. Constr.* 51 (6): 1–14. https://doi.org/10.1617/s11527-018-1278-9.
Reineck, K. H., E. Bentz, B. Fitik, D. A. Kuchma, and O. Bayrak. 2014. "ACI-DAfStb databases for shear tests on slender reinforced concrete beams with stirrups." *ACI Struct. J.* 111 (5): 1147–1156. https://doi.org/10.14359/51686819.
Reineck, K.-H., D. A. Kuchma, K. S. Kim, and S. Marx. 2003. "Shear database for reinforced concrete members without shear reinforcement." *Struct. J.* 100 (2): 240–249.
RILEM. 1991. "TC89-FMT: Fracture mechanics of concrete test methods." *Mater. Struct.* 23: 457–460.
Sakthivel, T., R. Gettu, and R. G. Pillai. 2019. "Compressive strength and elastic modulus of concretes with fly ash and slag." *J. Inst. Eng. Ser. A* 100 (4): 575–584. https://doi.org/10.1007/s40030-019-00376-w.
Shioya, Y., and H. Akiyama. 1994. "Application to design of size effect in reinforced concrete structures." In *Size effect in concrete structures*, edited by H. Mihashi, H. Okamura, and Z. P. Bažant, 409–416. London: E & FN Spon.
Yoshida, Y., E. Bentz, and M. P. Collins. 2000. *Results of large beam tests*. Toronto: Univ. of Toronto.
Yu, Q., J.-L. Le, M. H. Hubler, R. Wendner, G. Cusatis, and Z. P. Bažant. 2016. "Comparison of main models for size effect on shear strength of reinforced and prestressed concrete beams." *Struct. Concr.* 17 (5): 778–789. https://doi.org/10.1002/suco.201500126.
Zhai, J., C. I. Chen, B. L. Worsfold, and J. P. Moehle. 2022. "Blind prediction competition of a full-scale shear-critical reinforced concrete mat foundation slice." In *Proc., 12th National Conf. Earthquake Engineering*. Oakland, CA: Earthquake Engineering Research Institute.