Tri-level Framework for Realistic Estimation of Concrete Strength Using Bayesian Data Fusion of UPV and Guided Coring

Sharvil Alex Faroz1 · Siddhartha Ghosh2 · Alex Michael Faroz3

Received: 9 July 2022 / Accepted: 29 October 2022 / Published online: 8 November 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Identification of in situ concrete characteristic strength (CCS) of deteriorating RC structures directly affects their rehabilitation. Contemporary practices lack the ability to yield CCS or at most produce an unrealistic estimate, because: (i) core tests are few due to technical and economic constraints and their positioning is random, inducing bias in the strength estimate, (ii) non-destructive tests (NDT) are themselves not good estimators of project specific strength, (iii) tests are only local indicators and most practising engineers are unaware of how to process the spatial variability. This paper overcomes these limitations with a novel and implementable framework, for an asset manager to make scientific estimates of CCS. This is a tri-level framework: (1) NDT are performed to capture the spatial variability of the investigated concrete; (2) core extractions are guided by the prior NDT information, to avoid bias in test location selection; (3) NDT and coring results are fused via Bayesian NDT test calibration and updating of the probability distribution of the concrete strength. The proposed framework is successfully implemented for the assessment and rehabilitation of a critical hospital complex in India.

Keywords Characteristic strength · Homogeneous zone · Guided coring · Calibration · Bayesian NDT data fusion · Rehabilitation

1 Motivation
Reinforced concrete (RC) is a major construction material for civil structures. The characteristic strength of concrete $f_{ck}$ (or the specified compressive strength $f'_{c}$ in USA) is arguably the most important basic parameter used in the analysis and design of RC structures. For example, it is used in the calculation of the design strength of a compression member or that of a flexural member [1]. Moreover, the in situ concrete strength also plays an important part in reliability analyses [2,3], fragility analyses [4,5] and remaining life analyses [6,7] of deteriorating infrastructures.

Generally, the collapse of RC framed structures is strongly influenced by a deficiency in concrete strength, which greatly affects the capacity of columns to carry compressive forces [8,9]. Investigating the in situ mechanical property of concrete is a crucial step to evaluate the structural capacity of both old (affected by gradual and possibly shock deteriorations, like earthquakes or fire) and of new (to check non-conformity of strength) constructions, thereby assisting in rehabilitation measures (eg. [10]) when necessary. Under-estimation of the concrete strength results in excessive rehabilitation incurring unnecessary costs, and its over-estimation compromises the safety of the users in case the structure is continued to be used with little or no intervention. A realistic estimate of concrete characteristic strength on the other hand can maintain a balance between safety and economy.

For an existing structure, although the concrete is of a designated grade, variability of in-place strength is expected due to within-batch, batch-to-batch and systemic (i.e., within...
core samples is very small; engineers many a times limit to
crete strength estimation, particularly when the number of
Such random coring can induce significant bias in the con-
variation through VI and specify appropriate core locations.
It is difficult for any inspection engineer to actually detect the spatial
location solely based on its ease of access for testing. It is dif-
structural elements of that building. This work overcame the subjectivity of choosing the number of tests to be performed for a given project. Moreover, the outcomes of the NDTs were used to identify zones of concrete with similar properties indicating structural element-wise variability. The generally adopted random selection of core locations was transformed by Pfister et al. [18] to a systematic process. Their basic idea was that the core location defined according to the information previously provided by NDT data offer a better representation of the variability in strength of the concrete under inspection. Since cores are generally few in numbers, a conversion model between NDT and core crushing strength is useful to estimate local strength values by converting NDT test results performed at different test locations within the structure. A deterministic regression calibration of these models was adopted [19,20]. However the uncertainty in the calibration process could not be quantified in a deterministic regression. Ali-Benyahia et al. [21,22] proposed the minimum number of core samples required to sufficiently calibrate a conversion model. Further the NDT converted to strength values were used to obtain mean and standard deviation of the strength values [15]. However they did not provide a methodology to obtain a robust probability distribution of concrete strength for different zones of a building. A non-robust estimate can give a false over estimation of concrete strength [23]. Models were developed which provide the variability of concrete strength for individual NDT test values [24]. However their work did not provide an adoptable method for engineers/inspectors to use this information on how to combine results from multiple tests and from multiple locations in the repair-maintenance decision-making process. While these scattered efforts tried to address the uncertainties involved in the NDT results to relate NDT results with more reliable tests like core tests, none of these provided a wholistic and systematic approach towards decision-making for maintenance, repair or rehabilitation of a deteriorated RC structure, addressing (i) identification of zones having different concrete strengths, (ii) incorporation of uncertainty in the measuring instrument and avoiding point estimates in favour of full probabilistic characterization, (iii) interpretation of concrete strength from small number of test data, (iv) quantification and handling of uncertainty of NDT and onsite variability and (v) probabilistic NDT data fusion to obtain $f_{ck}$, in a wholistic manner.

In the proposed framework, at first, all the components of a structure—such as beams, columns and slabs—are identified and numbered. Among these, the components to which access is available are separately noted. In the next step, the conventional practise of conducting NDT and CT in parallel (or irrespective of the other’s occurrence) is intentionally avoided. Rather, an extensive scanning of the structure is performed using only a NDT method. The relatively inexpensive NDT is carried out on accessible components to capture the spatial variability of concrete and then to identify physical zones of statistically homogeneous concrete in the structure. Following this, a coring scheme is adopted—guided by the NDT results—at a location within each homogeneous zone. This guided coring reduces the probability of selecting cores from less frequent concrete strength ranges. Thus, it removes the outliers’ bias and helps in acquiring data from the whole structure with a minimum number of expensive core tests.

The guided coring is followed by a Bayesian NDT calibration for incorporating the test uncertainty as per the site conditions. This allows for a statistical characterization of crushing strength at every location of NDT test, without performing the expensive core tests at each of these locations. Moreover this uncertainty quantification captures the scatter and makes the analysis closer to reality. Then, a statistical characterisation of concrete strength distribution is performed for each homogeneous zone using Bayesian data fusion of the outcomes of CT and NDT. This combination resolves the issue of non-congruent outcomes on sound mathematical basis. The characteristic strength ($f_{ck}$), the specified compressive strength ($f'_c$) or any other strength parameter, can then be easily computed from this distribution. The steps of the proposed framework are illustrated in Fig. 1. This framework is successfully demonstrated for a hospital building case study.
Fig. 2  
(a) Stages of assessment and rehabilitation for the hospital building;  
(b) Distressed structural components in the hospital;  
(c) In-plan structural layout of the building

3 Project Description

The structure under consideration is a “Ground + 2” floor hospital complex having a C shape in plan (Fig. 2a), constructed circa 1995. The building is located at Vasai, near Mumbai in India. The building was found to be visibly distressed at many places (Fig. 2b) as a result of which the hospital was vacated despite huge inflow of patients. Weighing in the huge demand of this critical infrastructure facility, the owner decided that it would be quicker to undertake rehabilitation of the hospital rather than demolishing and completely rebuilding it. The structural assessment and rehabilitation are performed in three stages (Fig. 2a). This stage-wise breakup is opted for, to allow for moving delicate and expensive medical equipment from one part of the building to another while the first part is being tested and rehabilitated.

Considering the absence of original structural drawings, “as-is” architectural and virtual structural models for this building are created in the very first step of this rehabilitation project. This work involves detailed visual reconnaissance,
physical measurements of the whole premise and of the structural members. Non-invasive techniques are used to locate and acquire reinforcement details, and the residual diameter of corroded rebars are measured. In the same reconnaissance phase, structural components which are accessible for conducting tests are also identified. Post data assimilation, the structural plans of the building are prepared (Fig. 2c). Figure 2a shows the part of the hospital considered in Stage 1. The built-up area of each floor in this part is 168.25 m² and is composed of 32 columns, 52 beams and 20 slab panels per floor.

4 Level 1: Non-destructive Testing of the Building

Non-destructive testing forms the next stage after a detailed reconnaissance. Usually both RH and UPV tests are conducted for such investigations; however past research showed that in most cases, using just the UPV test yields more realistic results [18]. Further, the relevant Indian Standard recommends that RH should not be performed if the concrete is designated as ‘doubtful’ according to the UPV test [16], which is the case in the present project. It should be noted that RH is greatly affected by carbonation—overestimating the concrete strength by 50% [16] leading to erroneous estimates [17], whereas UPV results are negligibly affected by carbonation [25]. Considering these, UPV test is selected as the preferred and only NDT for this project.

4.1 NDT Design

As discussed previously, NDTs widely explore the structure and thus their number should be decided to adequately cover the expanse of a structure. The minimum number of NDTs to be conducted depends on the number of primary structural components and the total floor area. For an extended level of testing, Masi et al. [15] suggested to use at least 9 test points or 11% of the total primary components as NDT test points for a density of 80 structural components (column+beam) per floor area of 300 m². Using this density in the present structure (104 components per 168.25 m²), along with the minimum values prescribed for an extended level of testing, we arrive at either (i) seven UPVs or (ii) five UPVs. Adopting the higher value, seven tests are performed on all columns, beams and slabs per floor. Since it is usually difficult to perform tests on beams and slabs, only 25% of total tests on these components are performed with at least three test points [15]. It is decided to perform four tests on beams and slabs per floor. Thus 15 test locations per floor and in total 45 locations in the building are chosen for NDT in this project. These suggestions (i.e. [15]) should not be considered as rigid and the number of NDT points can be decided based on the condition of the project at hand. For example, when long span frames are involved, the density of structural components is low and the minimum number of NDT as per this suggestion comes out to be very small.

The UPV test locations are selected only in the accessible parts of the building. Since ‘guided coring’ (introduced later) is to follow in sub-domains of the UPV test regions, the NDT test points are chosen considering that future coring in these locations would not weaken the structure in any significant way. Additionally, using a rebar locator it is also seen that these corings would not take place, cutting through rebars underneath. For beams it is best to test the lower part on the vertical face at mid-span. The reason is that this part of the beam is in tension under sustained/gravity loads and core extraction does not affect its capacity much. For slabs it is decided that only UPV test shall be conducted since it is massively difficult to core on a slab. For the columns, recognizing that the static pressure due to consolidation after concrete casting causes its strength to vary along the member height [11], tests are carried out at column mid-height to average out such effects. Besides, this is generally the position with the least bending moment and with least shear reinforcements. Other recommended care, as suggested by relevant codes of practice, are incorporated in this process.

The nomenclature used for test location is ‘XYZ’, where X indicates the floor: G (ground), F (first), or S (second); Y is the structural component type: C (column), B (beam), or S (slab); and Z is the serial number of a particular component. The UPV test locations are shown in Fig. 3 for the ground, first and second floors, respectively. On each structural component shown in the figures, a single UPV test is conducted at the test location. The UPV values (V) for different locations around the hospital building are presented in Table 1.

4.2 Identification of Homogeneous Concrete Zones

Concrete can be considered homogeneous when the coefficient of variation (CoV) of concrete compressive strength varies within the range of 10 to 20% [26]. Non-homogeneity of concrete over a structure can happen due to multiple reasons, including replacement of the contractor or the concrete production company, change in the source of construction material supply, and changes in the exposure after casting, in addition to the casting sequence of different components and different floors, mentioned earlier. It is intuitively expected that components within each separate casting phases will have relatively similar characteristics. However, it may be possible to have homogeneous areas within the same floor and within multiple floors too.

The important aspect to consider in processing the achieved test results (Table 1) is therefore, whether the entire collected data is from a single population (implying that the structure is homogeneous in terms of concrete strength) or,
as often happens, due to the various construction stages of the building, there exists multiple homogeneous areas (i.e., more populations). The UPV test in this project is used to recognize areas which can be considered homogeneous with respect to the concrete quality based on the velocity values, since the mechanical property of concrete corresponds to the UPV test [8]. The homogeneous parts of the building are identified in three steps.

In the first step, the structure is divided into sub-groups, that can be treated as statistically homogeneous a priori: C0, C1 and C2 are the group of columns on ground, first and second floors, receptively; BS0, BS1 and BS2 are the group of beams + slabs on the ground, first and second floors, respectively. For statistical analysis, working with a Gaussian distribution is simple; however, as UPV values take positive real numbers it may not be properly represented by a Gaussian distribution. To take the advantage of the Gaussian simplicity still, we consider the logarithm of the UPV (lnV) values henceforth. The variability of the UPV test results for these groups is based on the evaluation of the CoV of the test results. In this project we have adopted a limiting CoV of 15% for homogeneity. Thus, the groups themselves homogeneous if CoV ≤ 15% for each group. Table 2 reports the sample size (n), sample mean (m), unbiased sample standard deviation (s), and CoV for the tests. From this table, we see that each group has CoV well below 15% and therefore each group is considered homogeneous within itself. These statistics can be visualized with a two
Table 1 UPV values for each test location

| Element ID | V  | ID | V  | ID | V  |
|------------|----|----|----|----|----|
|            | (km/s) | (km/s) | (km/s) | (km/s) | (km/s) |
| Column GC1 | 2.84 | FC1 | 2.81 | SC1 | 2.91 |
| Column GC2 | 2.89 | FC2 | 2.80 | SC2 | 2.77 |
| Column GC3 | 2.77 | FC3 | 2.87 | SC3 | 2.70 |
| Column GC4 | 2.83 | FC4 | 2.77 | SC4 | 2.81 |
| Column GC5 | 2.79 | FC5 | 2.89 | SC5 | 2.80 |
| Column GC6 | 2.87 | FC6 | 2.90 | SC6 | 2.72 |
| Column GC7 | 2.85 | FC7 | 2.93 | SC7 | 2.66 |
| Slab GS1   | 2.82 | FS1 | 2.96 | SS1 | 2.87 |
| Slab GS2   | 2.96 | FS2 | 2.87 | SS2 | 3.03 |
| Slab GS3   | 2.92 | FS3 | 2.72 | SS3 | 2.90 |
| Slab GS4   | 2.82 | FS4 | 2.84 | SS4 | 2.79 |
| Beam GB1   | 2.90 | FB1 | 2.82 | SB1 | 2.77 |
| Beam GB2   | 2.76 | FB2 | 2.86 | SB2 | 2.84 |
| Beam GB3   | 2.89 | FB3 | 2.97 | SB3 | 2.73 |
| Beam GB4   | 3.01 | FB4 | 2.80 | SB4 | 2.92 |

Table 2 Statistical values of ln V

| C0 | C1 | C2 | BS0 | BS1 | BS2 |
|----|----|----|-----|-----|-----|
| n  | 7  | 7  | 7   | 8   | 8   | 8   |
| m  | 1.042 | 1.048 | 1.017 | 1.060 | 1.049 | 1.049 |
| s  | 0.015 | 0.021 | 0.030 | 0.028 | 0.029 | 0.033 |
| CoV (%) | 1.438 | 1.998 | 2.940 | 2.678 | 2.744 | 3.181 |

The second step is to determine if the homogeneous subgroups belong to the same or a different population. A simple test adopted here consists of comparing the mean of ln V of two groups; if equal, they belong to the same population, and otherwise if found unequal. For any two groups P and Q, of size \( n_P, Q \) with sample mean \( \bar{m}_P, Q \) and sample standard deviation \( s_{P, Q} \), two alternate hypotheses can be evaluated:

1. \( H_0 : \bar{m}_P = \bar{m}_Q \) (‘null hypothesis’)
2. \( H_1 : \bar{m}_P \neq \bar{m}_Q \) (‘alternate hypothesis’)

These hypotheses are examined with the t-test, for the test statistic:

\[
t = \frac{\bar{m}_P - \bar{m}_Q}{\sqrt{\frac{\sum (m_P - \bar{m}_P)^2}{n_P} + \frac{\sum (m_Q - \bar{m}_Q)^2}{n_Q}}},
\]

where \( \Sigma^2 = \frac{(n_P - 1)s_P^2 + (n_Q - 1)s_Q^2}{n_P + n_Q - 2} \). The hypothesis \( H_0 \) is accepted if \( t < t_c \), where \( t_c \) is the critical value corresponding to a probability \( (1 - \frac{\alpha}{2}) \) for a Student’s t-distribution with a degree of freedom \( (n_P + n_Q - 2) \). There is a probability of error that a wrong hypothesis may be selected for such cases. In this regard there are two possible errors [27]:

1. Type I error → rejecting \( H_0 \) when in fact it is true
2. Type II error → accepting \( H_0 \) when in fact it is false

The selection of \( \alpha \) depends on the importance attributed to the two types of errors. In the present case, it is preferable to assume any two actually homogeneous zones to be inhomogeneous rather than having the opposite error; for this reason, a larger value of \( \alpha = 0.20 \) is used. The t-test is used to compare groups in pair at a time from the available groups. Table 3 provides the result of hypothesis testing. For \( t < t_c \), the groups in pair are considered similar.

Based on this analysis, it is possible to create “hyper groups” comprising of many original groups. Each original group is checked with all others to see if they are similar as per t-test and are combined into a hyper group. Fig. 4b, c, and d show three possible combinations or hyper groups. In
these figures the connecting line between each group implies that they are similar as per t-test from Table 3. As C2 has failed to be similar with any of the groups as per t-test, it is treated as a separate group all by itself. Groups C1, BS1, BS2 show similarity to C0 (Fig. 4b) and to BS0 (Fig. 4d) separately. However the pair C0, BS0 itself has failed the t-test (Table 3). As such Fig. 4c can be supposed to comprise of two categories and it must be tested if inclusion of C0 and BS0 in a single hyper group can cause it to fail the hypothesis of belonging to the same population.

In the third step, ANalysis Of VARIance (ANOVA) is used to test the hyper groups if they belong to the same population. For any groups P, Q, R, with mean \( \bar{m}_P, \bar{m}_Q, \bar{m}_R \), two alternate hypotheses can be evaluated:

1. \( H_0: \bar{m}_P = \bar{m}_Q = \bar{m}_R \) (null hypothesis)
2. \( H_1: \bar{m}_P \neq \bar{m}_Q, \bar{m}_P \neq \bar{m}_R, \bar{m}_Q \neq \bar{m}_R, \bar{m}_P \neq \bar{m}_Q \neq \bar{m}_R \) (alternate hypothesis)

When applying ANOVA a significance level of \( \alpha = 0.20 \) is used. The null hypothesis is accepted if the test statistic, \( F < F_c \) (critical value). Table 4 reports the findings of the ANOVA test, indicating that all hyper groups are accepted. We choose the hyper group C0-C1-BS0-BS1-BS2 as an all encompassing group for further analysis. Thus we have two major homogeneous zones (Z1, Z2) for the building viz:

1. Z1 : C0-C1-BS0-BS1-BS2
2. Z2 : C2

The CI for these two new hyper groups is shown in Fig. 4e. The CI for both groups is apart, thus confirming that they can be considered as belonging to different populations. A comparison of Fig. 4a and e shows that C2 was always separate from the other groups since its inception. Thus before conducting the t-test in step two, the CI can be used to eliminate groups whose CI do not overlap with any of the other groups, thus saving computation effort of the t-test.

### 5 Level 2: Guided Coring

First, to decide the number of cores to extract, an initial reference is made to the work of Masi et al. [15]. Accordingly, for an extended level of testing, at least 4 in number or 5% of the total number of primary components must be extracted per homogeneous zone. Through this, we arrive at (i) 4 minimum number of cores and (ii) with 2 and 1 cores as percent of total primary components in Z1 and Z2, respectively. We could choose 4 cores per zone, however we find it quite unfair to take equal cores in both zones, considering Z1 has 38 elements and Z2 has only 7 elements and at least 3 cores are suggested by IS 456 [1]. Seeking reference elsewhere, we find that Giannini et al. [8] suggested to use at least 5 cores, although they did not provide any basis for the same. On the other hand, Ali-Benyahia et al. [21] concluded that when using a single NDT method, a minimum of 9 cores should be used for effective assessment. This conclusion was based on a deterministic model fitting based on two indicators: root mean square error and coefficient of determination. Although such deterministic criteria are challenged by probabilistic methods [28], nevertheless, we have used total 9 cores in total for this project and distributed \( N_{c1} = 6 \) cores in Z1 and \( N_{c2} = 3 \) cores in Z2.

If these cores are extracted randomly, the outcome may introduce bias in the estimation of concrete strength, particularly when the number of cores is small (in this project, it is merely 2.88% of the total number of structural components in the building). Since the mechanical property of concrete corresponds to the UPV test [8], the ingenuity is to select the core locations using the “prior” information provided by the NDT exploration [18,26]. Considering that the core location is guided by the UPV results previously done, we call this method ‘guided coring’. Fundamentally, we choose the location of cores in such a way that it provides an unbiased sample of cores and reduces the probability that all cores belong to a limited range of strength. Guided coring thus only requires

### Table 3 Results of the Student’s t-test

| Pairs                | t    | \( t_c \) | Decision |
|----------------------|------|----------|----------|
| C0 - C1              | 0.6617 | 1.3562  | Accept   |
| C0 - C2              | 1.919 | 1.3562  | Reject   |
| C0 - BS0             | 1.4576 | 1.3502  | Reject   |
| C0 - BS1             | 0.5782 | 1.3502  | Accept   |
| C0 - BS2             | 0.5339 | 1.3502  | Accept   |
| C1 - C2              | 2.2248 | 1.3502  | Reject   |
| C1 - BS0             | 0.846 | 1.3502  | Accept   |
| C1 - BS1             | 0.0437 | 1.3502  | Accept   |
| C1 - BS2             | 0.0605 | 1.3502  | Accept   |
| C2 - BS0             | 2.7727 | 1.3502  | Reject   |
| C2 - BS1             | 2.0622 | 1.3502  | Reject   |
| C2 - BS2             | 1.9183 | 1.3502  | Reject   |
| BS0 - BS1            | 0.7324 | 1.345   | Accept   |
| BS0 - BS2            | 0.6557 | 1.345   | Accept   |
| BS1 - BS2            | 0.02  | 1.345   | Accept   |

### Table 4 Results of the ANOVA test

| Pairs                | F   | \( F_c \) | Decision |
|----------------------|-----|----------|----------|
| C0-C1-BS1-BS2        | 0.1293 | 1.6601   | Accept   |
| C0-C1-BS0-BS1-BS2    | 0.4248 | 1.5904   | Accept   |
| C1-BS0-BS1-BS2       | 0.2723 | 1.6558   | Accept   |

Henceforth, the NDT exploration [18,26]
that NDT test results are available before extracting cores and as such, do not incur any additional cost for achieving so.

The $\ln V$ values of each zone ($Z_l, l = 1, 2, \ldots$) are arranged from the lowest to the highest value. This data is then subdivided into $n_l$ subsets for each zone and the guided coring proceeds in either of the two ways:

1. $n_l < N_{C_l}$: The number of cores in each subset is selected proportional to the relative frequency of the UPV data in each subset. The core location corresponding to the existing UPV location, can be chosen to cover the range of that UPV subset.
2. $n_l = N_{C_l}$: One core is selected within the set or at the median value of the data in each UPV subset.

We adopt the first way to divide $Z_1$ into $n_1 = 3$ parts and $Z_2$ is divided into $n_2 = 2$ parts. The histogram of UPV is presented in Fig. 5a and b for $Z_1$ and $Z_2$, respectively. For zone $Z_1$, 2 cores are allocated in the lower subset, 3 in the middle and 1 in the upper subset. The distribution of cores is shown in Fig. 5c. Similarly for $Z_2$, 2 cores are allocated in the lower subset and 1 in the upper subset, for which the histogram can be seen in Fig. 5d. One can see that the shape of core distribution is similar to that of the histogram of the $\ln V$ values. These cores are positioned on locations corresponding to the locations of UPV from Fig. 3. The cores extracted following the guided coring scheme are shown in Fig. 5e.
Fig. 6 a UPV testing of extracted cores, b uncertainty quantification in measurement

6 Level 3: Bayesian Data Fusion

6.1 Bayesian UPV Test Calibration

NDTs are affected by many site conditions and cause considerable measurement uncertainty [6]. Inevitably, interpreting actual results from the uncertain measurements requires calibration of such tests [12,23,29]. In order to perform such a calibration, prior to testing the concrete cores in crushing, they are tested with UPV (Fig. 6a). Then the cores are crushed to generate the $f_c$ values. The results of this exercise are reported in Table 5. This dataset is denoted as, $\mathcal{D} = \{(V_1, f_{c1}), \ldots, (V_j, f_{cj}), \ldots, (V_n=9, f_{cn=9})\}$.

The measurement conversion model (MCM) relating concrete crushing strength ($f_c$) and UPV ($V$) is [8]  

$$\ln f_c = \beta_0 + \beta_1 \ln V + \varepsilon$$  \hspace{1cm} (1)$$

$\beta_0 \in \mathbb{R}$ and $\beta_1 \in \mathbb{R}^+$ are called the constant bias and scaling bias respectively, collectively known as the systematic biases of an NDT instrument [6]. Typically, they represent the epistemic uncertainty, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon)$ denotes the additive random error in measurements, which represents the aleatory uncertainty. Thus the MCM parameters are $\theta = [\beta_0 \ \beta_1 \ \sigma_\varepsilon]^T$, which are to be identified/updated with the training data $\mathcal{D}$. There are three basic approaches of parameter identification [23]: (i) the least squares method, (ii) the maximum likelihood estimation and (iii) the Bayesian method. The first and the second methods have the drawback that they give just point estimates of the model parameters. In contrast, the Bayesian method provides the distribution of the model parameters as well. A Bayesian framework is thus deemed more suitable and is adopted for calibration. As we do not have specific prior information of these values, assuming an improper prior for $\theta$, the posterior estimative distribution for $Y = \ln f_c$, given $X = \ln V$ is [30]:

$$f(Y|X, \mathcal{D}) = ST(Y|\hat{\beta}_0 + \hat{\beta}_1 \cdot X, \frac{n \cdot h(X)}{n-2}, n-2) \hspace{1cm} (2)$$

where $ST(\cdot)$ indicates the Student’s t-distribution. The probabilistic estimation of calibrated UPV is shown in Fig. 6b for the training data. The uncertainty in relating NDT with core is quantified in the process. It is evident that, over a deterministic regression, a Bayesian calibration can capture the scatter of the training data. The posterior estimative distribution of $f_c(= e^Y)$ is determined considering: $F_{f_c}(f) = \mathbb{P}(f_c \leq f) = \mathbb{P}(Y \leq \ln f) = F_Y(\ln f)$. Applying derivative to this equation, it follows:

$$f(f_c|V, \mathcal{D}) = \frac{1}{f_c} f(Y|X, \mathcal{D}) \hspace{1cm} (3)$$

Once the UPV test is thus calibrated, for every value of $V$ measured in situ, a PDF for $f_c$ is obtained. It thus becomes possible to get the strength of concrete without extracting an actual core form all the test points in the structure, plus...
we have also tuned the NDT with the specific structure. For future projects using similar equipment, the calibration results obtained in this work can be treated as prior which can be updated using data collected from that specific project.

6.2 Data Fusion

Recognizing the fact that core crushing values are the most reliable values but few in numbers while the UPV are numerous, practically scanning the entire structure under investigation, it is a natural conclusion to combine the two, in order to characterize the concrete strength and its variability. Combination of multiple tests has been suggested to improve the reliability for the interpretation of results [31]. The concept of combining information from multiple sources or tests in a Bayesian setting is called Bayesian data fusion [6,32].

In general, the normal and lognormal distribution models are considered to describe concrete strength [33]. Usually a normal distribution is suitable for pristine concrete where quality control is maintained. While assessing existing or degraded structures however, using a normal distribution may yield unrealistic results. The reason behind this is that due to a high variability of strength in such structures the normal distribution may well extend to the negative domain and can have considerable probability in that domain, even exceeding 5% ($f_{ck}$). Therefore in this project a lognormal model is adopted for concrete strength:

$$f(f_{c}|\mu, \sigma) = \frac{1}{f_{c}\sqrt{2\pi}\sigma^2} \exp \left[ -\frac{(\ln f_{c} - \mu)^2}{2\sigma^2} \right]$$

(4)

The objective is to identify the model parameters $[\mu \ \sigma]^T$ which describe the variability characteristic of concrete in each zone $Z_1$ or $Z_2$, through Bayesian data fusion.

6.2.1 Prior Information

The model parameters are: $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$. We define $\sigma$ in an alternate form: $\sigma = e^2$, thus the parameter set becomes $\xi = [\mu \ s_2]^T$. Prior distribution of the concrete strength parameters may be of three types: i) informative; ii) partially informative (based on information provided by the relevant standard); iii) non-informative [32]. Jalayer et al. [34] constructed a prior concrete strength distribution to represent typical values found in post World War II construction in Italy. Caspeele and Taerwe [35] proposed a set of informative priors for different concrete classes based on concrete production data from Germany. Although their prior may be adapted to different countries, its applicability to RC structures for which there is no information regarding the expected concrete class is difficult. In the present case study of the hospital building, original design details are unavailable. Thus constructing an appropriate informative prior for concrete strength is not feasible here. Moreover a prior for the original concrete strength may still be possible to adopt based on the minimum strength requirement as per the design code prevalent at site and the time of construction. However, that prior would be valid only for the pristine concrete and not the degraded present condition. As the specific information of concrete used for the building is not available, we assume a non-informative prior distribution; $f(\mu) = h_1, f(s_2) = h_2$. This uniform density is defined over the entire admissible range of these parameters, which is ($-\infty, +\infty$) [36]. The joint prior distribution is taken as the product of the densities of individual parameters, by assuming each parameter to be independent of the other: $f(\xi) = h_1 \times h_2$, where $h_1, h_2$ are very small values.

6.2.2 Likelihood Formulation

We have two types of data available as outcome of the investigation, (i) core crushing values, as point values, and (ii) local distribution of $f_{c}$ based on the UPV reading. Such information are available in both the zones $Z_1$ and $Z_2$.

The likelihood for a single core crushing value $f_{ck}$ is $f(f_{c}|\xi)$ from Eq. 4. For $N_{C}$ core values in a given zone $Z_i$ of this project, the likelihood considering each value to be independent of the other is

$$L_c(\xi) = \prod_{j=1}^{N_{C}} f(f_{c_j}|\xi)$$

(5)

In each zone ($Z_1, Z_2$) the location where core has not been extracted, Eq. 3 is used for estimation of $f_{c}$ based on UPV. $f_{c}$ thus obtained from the calibrated UPV test is not in the form of a point value (unlike core), rather it is a distribution (Eq. 3) for each UPV reading $V_i$. Thus, Eq. 4 cannot be directly used to form the likelihood function as in the case of core values. The likelihood function of $f_{c}$, for a single measurement of $V_i$ is thus formulated as [6,32]

$$L_i(\xi) = \int_{f_{c}} f(f_{c}|\xi) \cdot f(f_{c}|V_i, D)df_{c} \approx \frac{1}{S} \sum_{k=1}^{S} f(f_{c}^{(k)}|\xi)$$

(6)

where, $f_{c}^{(k)}$ are $k = 1, 2, \ldots, S$ samples from $f(f_{c}|V_i, D)$. Theoretically, a large number of $f_{c}^{(k)}$ values are required in Eq. 6 for better approximation. However more samples also increase computation. Here we evaluate the integral for $S = 30$, obtained using Latin Hypercube Sampling from $f(f_{c}|V_i, D)$. The number 30 is chosen analogous to the code’s recommendation of 30 samples for satisfactory statistical assessment [14]. Later it is also verified (Fig. 7) that
30 samples are indeed sufficient to approximate the likelihood function. An important aspect to note is that the UPV values at the location where core is extracted are not used to form the likelihood, since the core values themselves are available. Thus for a total \( N_{v} \) UPV values in each zone, only \( N_{v} - N_{c} \) UPV values are considered to form the likelihood:

\[
\mathcal{L}_{v}(\xi) = \prod_{i=1}^{N_{v} - N_{c}} L_{i}(\xi) \tag{7}
\]

The overall likelihood combining the point values and the PDF values of \( f_{c} \) is

\[
\mathcal{L}(\xi) = \mathcal{L}_{c}(\xi) \times \mathcal{L}_{\xi}(\xi) \tag{8}
\]

### 6.2.3 Posterior Estimation

The concrete distribution parameters are updated through Eq. 8 by Bayesian updating for a non-informative prior [6]:

\[
f(\xi) \propto \mathcal{L}(\xi), \quad N = 5,000 \text{ sample values of the parameters } \xi \text{ are generated after thinning using the Delayed Rejection Adaptive Metropolis (DRAM) algorithm. Once updated the parameters } \xi \text{ are available in the form of sample values } (\xi^{(k)}; \quad k = 1, 2, \ldots, N). \text{ A robust distribution of } f_{c}, \text{ which includes the variability due to the sample values of } \xi, \text{ is obtained as } [6]
\]

\[
f(f_{c}) = \int_{\xi} f(f_{c}|\xi) f(\xi)d\xi \approx \frac{1}{N} \sum_{k=1}^{N} f(f_{c}|\xi^{(k)}) \tag{9}
\]

Using Eq. 9, a distribution for concrete crushing strength \( f_{c} \) is obtained for all zones in the structure viz. \( Z_{1} \) and \( Z_{2} \).

In research, many times the sample mean, \( \bar{\xi} = \left[ \frac{1}{N} \sum_{k=1}^{N} \bar{\xi}^{(k)} \right]^{T}, \) or the sample median, is used as parameter values for inference, in contrary to Eq. 9 [6,23]. The approximate distribution in such cases will be:

\[
f(f_{c}) \approx f(f_{c}|\bar{\xi}). \text{ The purpose here is to compare and check if it can effectively approximate Eq. 9, (which requires more computation) and simplify the estimation.}
\]

### 7 Estimation of Characteristic Strength \( f_{ck} \)

The characteristic value \( f_{ck} \) which is the 5 percentile value of \( f_{c} \), is computed by solving

\[
0.05 = \int_{0}^{f_{ck}} f(f_{c})d f_{c} \tag{10}
\]

Similarly the specified strength \( f'_{c} \) can be obtained as the 1 percentile value of the crushing strength. To check the effect of the number of samples \( S \) (for approximating the likelihood function in Eq. 6) on the evaluation of \( f_{ck} \), the estimate (Eq. 10) is evaluated for increasing values of \( S \). These \( S \) values are samples from Eq. 3 using the Latin Hypercube Sampling. When only a single sample is drawn \( (S = 1) \), the expected value of the whole distribution is considered or taken to be the sample. From the result in Fig. 7a, it is seen that initially there is fluctuation in the estimation of \( f_{ck} \), but beyond \( S = 30 \) the estimate stabilizes. Therefore, for the present project \( S = 30 \) samples are adopted to give a reliable approximation.

The concrete strength distribution for both different zones is shown in Fig. 7b along with their \( f_{ck} \) value. For \( Z_{1} \), \( f_{ck} = 4.42 \) MPa and for \( Z_{2} \), \( f_{ck} = 3.64 \) MPa. For the approximate distribution, the characteristic value for \( Z_{1} \), \( f_{ck}^{approx} = 4.58 \) MPa and for \( Z_{2} \), \( f_{ck}^{approx} = 4.41 \) MPa. Clearly, this is an overestimate of 1.026 and 1.161 times respectively when compared to \( f_{ck} \) obtained from Eq. 9. Presently no clear empirical relation can be seen to relate these two values in general, irrespective of the zones considered. However, based on Fig. 4e, it can be observed that for the zone with a lesser spread the \( f_{ck}^{approx} \) value is closer to the \( f_{ck} \) value obtained from Eq. 9.

### 8 Rehabilitation Scheme

For the scope of this paper, a single column in zone \( Z_{1} \) on the ground floor (GC4 of Fig. 3) is chosen to demonstrate the rehabilitation process adopted. The process is repeatable for other structural members. A detailed structural analysis model of the hospital building is created in ETABS [37]. The dead load and the imposed load are assigned as per Parts I and II, respectively, of IS 875 [38]. Additionally a load of 0.5 kN/m² is assigned to the top slab to account for the solar panels installed on the roof of the hospital. Seismic loads and corresponding load combinations are adopted from IS 1893 (Part 1) [39], and a dynamic seismic analysis following the response spectrum method is performed. The load combination 0.9DL+1.5EQ is found to be the most critical, where DL is the dead load and EQ is the seismic load. The demand on the column is evaluated as: axial load \( P = 257.10 \) kN, moments \( M_{11} = 44.80 \) kNm, \( M_{22} = 130.56 \) kNm. The axes 11 and 22 are parallel to the shorter and longer side of the column respectively.

The original size of this column is measured to be 230 mm × 400 mm. Since much of the cover concrete (40 mm) has spalled, the contribution of cover concrete to the capacity computation is ignored. For the purpose of strengthening, the loose cover concrete is chipped off. Reinforcements in this column are found to have corroded. The rust on the reinforcements is cleaned using a wire brush and the residual diameter of the bars is measured using a digital Vernier caliper. The measurement is taken only on exposed bars. For main bars, the diameters measured are: [9.4, 10.3, 10.0, 11.3, 11.8, 10.8, 10.4, 10.1].
Fig. 7  a Strength estimates with increasing number of samples ($S$) to approximate the likelihood function, b Evaluation of $f_{ck}$

8.6, 10.6, 11.3, 11.0, 11.4] mm. Their mean value of 10.5 mm is used in the capacity computation. Using this information and adopting $f_{ck} = 4.42$ MPa evaluated earlier in the framework proposed here, the existing capacity of the column is evaluated using the standard guidebook SP 16 [40]. This capacity is deducted from the demand on the column and rehabilitation is proposed using concrete jacket following IS 15988 [41] for the residual demand. The advantage of the framework is (due to realistic $f_{ck}$), the existing capacity estimation of the column is realistic and the rehabilitation scheme is deemed neither under-designed nor excessive, thus balancing between safety and economy. The strengthened column along with its section is shown in Fig. 8.

9 Conclusions

Existing RC structures can be inherently deficient due to a lack of sufficient detailing and/or corrosion of reinforcement, missing members, overloading, damage, etc. In all cases, determining the actual or in situ concrete characteristic strength $f_{ck}$ (or the specified compressive strength, $f'_c$) is one of the most important step in the assessment and rehabilitation of existing structures. Due to a lack of available systematic assessment techniques, this essential parameter is not determined in a scientific way.

This paper provides a strategic framework for the inspection and assessment engineers to evaluate realistic in situ strength of concrete, particularly in projects dealing with deteriorated structures. This framework attempts to eliminate the subjectivity in assessment (elaborated in the opening of this paper), which is inherent in most current practices. It also gives a systematic methodology to the practising engineers to investigate and come to a decision on the necessary rehabilitation. This system is based on establishing and quantifying the spatial variability and identifying strength homogeneous zones (parts of structure with similar concrete) in the domain of investigation through extensive NDT, which is impossible through only visual inspections. The bias in strength estimate, due to limited core numbers and random selection of core locations, is overcome through guided coring. This proposed coring strategy uses the variability captured by the NDT to determine core locations so that the chance of selecting cores from only a limited strength range is reduced. Subsequently, the disconnect between any NDT and the site concrete is addressed through a Bayesian calibration of the NDT data. Through calibration, it is technically possible to establish the crushing strength at every NDT location without actually extracting a core at that location. This overcomes the barrier of 30 concrete samples required for a satisfac-
tory statistical assessment. Through Bayesian data fusion, the information collected on statistical variability through NDT and actual strength through cores are combined to determine the probability distribution of in situ crushing strength and the characteristic value in each homogeneous zone.

This framework is recommended for a realistic estimation of concrete strength and hence the capacity of structure prior to rehabilitation. In this paper the proposed framework is implemented for a deteriorated hospital complex in India which had to be quickly rehabilitated due to large inflow of patients. Being a critical health infrastructure, the risk of an unrealistic assessment of concrete strength could be catastrophic and this made a strong case for the application of the proposed framework. Future evolution of this framework will focus on the maximisation of information gain through the inspections and testing stages when performed under economic constraints.

**Acknowledgements** The authors would like to thank the anonymous reviewers whose critical comments and useful suggestions have been valuable in enhancing the quality of the paper. Authors also thank Nihar Gonsalves for data collection and Kewal Raval for retrofit design of this project.

**Author Contributions** Conceptualization: [SAF]; Methodology: [SAF, SG]; Formal analysis and investigation: [SAF]; Writing - original draft preparation: [SAF]; Writing - review and editing: [SG]; Resources: [AMF]; Supervision: [SG]; Project administration: [AMF]; Investigation:[AMF]

**Funding** Not applicable.

**Data Availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

**Declarations**

**Conflict of interest** The authors declare that there is no conflict of interest.
22. Ali-Benyahia, K., Sbartaï, Z.-M., Breysse, D., Ghrici, M., Kenai, N., Pereira, N., Romão, X.: Assessing concrete strength variability in structures. ACI Mater. J. 96(2), 261–270 (1999)

10. Siddika, A., Alyousef, R., Amran, Y.H.M., Al Mamun, A., et al.: Strengthening of concrete beams by using fiber-reinforced polymer composites: a review. J. Build. Eng. 25, 100798 (2019). https://doi.org/10.1016/j.jobe.2019.100798

19. Rashid, K., Waqas, R.: Compressive strength evaluation by non-destructive techniques: an automated approach in construction industry. J. Build. Eng. 12, 147–154 (2017). https://doi.org/10.1016/j.jobe.2017.05.010

27. Ang, A.H.-S., Tang, W.H.: Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering, 2nd edn. Wiley, New York (2007)

28. Yuen, K.-V.: Recent developments of Bayesian model class selection and applications in civil engineering. Struct. Saf. 32, 338–346 (2010). https://doi.org/10.1016/j.strusafe.2010.03.011

30. Bernardo, J.M.: Bayesian methodology in statistics. In: Brown, S.D., Tauler, R., Walczak, B. (eds.) Comprehensive Chemometrics, pp. 213–245. Elsevier, Oxford (2009). https://doi.org/10.1016/B978-044452701-1.00095-8

18. Pfister, V., Tundo, A., Luprano, V.A.M.: Evaluation of concrete strength by means of ultrasonic waves: a method for the selection of core position. Construct. Build. Mater. 61, 278–284 (2014). https://doi.org/10.1016/j.conbuildmat.2014.03.017

36. Sankararaman, S., Mahadevan, S.: Separating the contributions of variability and parameter uncertainty in probability distributions. Reliab. Eng. Syst. Saf. 112, 187–199 (2013). https://doi.org/10.1016/j.ress.2012.11.024

12. Lad, D.B., Faroz, S.A., Ghosh, S.: Fusion of Rebound Number and Ultrasonic Pulse Velocity Data for Evaluating the concrete Strength Using Bayesian Updating. In: Varde, P.V., Prakash, R.V., Vinod, G. (eds.) Reliability, Safety and Hazard Assessment for Risk-Based Technologies. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9008-1_35

23. Ali-Benyahia, K., Sbartaï, Z.-M., Breysse, D., Ghrici, M., Kenai, N.: Analysis of the single and combined non-destructive test approaches for on-site concrete strength assessment: general statements based on a real case-study. Case Stud. Construct. Mater. 6, 109–119 (2017). https://doi.org/10.1016/j.cscsm.2017.01.004

22. Ali-Benyahia, K., Sbartaï, Z.-M., Breysse, D., Ghrici, M., Kenai, N.: Improvement of non-destructive assessment of on-site concrete strength: Influence of the selection process of cores location on the assessment quality for single and combined ndt techniques. Construct. Build. Mater. 195, 613–622 (2019). https://doi.org/10.1016/j.conbuildmat.2018.10.032

26. Krivok, G.J., Scanlon, A.: Estimation of compressive strength of concrete in existing structures based on in situ test data. ACI Mater. J. 84(3), 235–245 (1987)