

Article

Quantifying the Laboratory Effect in 3D Concrete Printing: A Meta-Analysis, Cross-Study Reliability Assessment, and Minimum Reporting Set

Salih Özdemir *  and Sema Alaçam 

Faculty of Architecture, Istanbul Technical University, Taşkışla Campus, Harbiye Mahallesi, Taşkışla Cad. No. 2, Şişli, 34367 Istanbul, Türkiye; alacams@itu.edu.tr

* Correspondence: salih.ozdemir@itu.edu.tr

Abstract

Process dependence is a defining feature of 3D concrete printing (3DCP), and any pooled inference across laboratories must quantify rather than ignore the laboratory effect. We assembled a harmonized corpus of 80 primary 3DCP studies (465 experiments and 1083 hardened mechanical observations) under PRISMA-style screening, applied controlled vocabularies and unit canonicalization, and estimated the share of variance attributable to the study cluster. The intraclass correlation coefficient is 0.77 (95% CI [0.59, 0.86]; 345 records and 40 studies), so more than three quarters of the dispersion in 28-day compressive strength lies between laboratories rather than between mixtures within the same laboratory. A 7-to-28-day calibration evaluated under leave-one-study-out (LOSO) cross-validation produces $R^2 = 0.625$, $MAE = 8.67$ MPa and $RMSE = 11.38$ MPa, with 95% prediction intervals covering 96.9% of held-out observations; the band is wide enough (≈ 85 MPa at the median) that the calibration is suitable for early-screening of mixture and process candidates rather than for code-compatible structural verification. Most of this between-study scatter traces back to processing variables, curing regimes and loading directions that are inconsistently reported. We propose a three-tier Minimum Reporting Set (MSRS) and demonstrate, through a retrospective audit of the included corpus, that adoption is feasible at the entry tier (63% already comply) while higher tiers identify the metadata that must be added before pooled inference can support structural-design decisions for printed elements.

Keywords: 3D concrete printing; compressive strength; digital construction; structural reliability; inter-laboratory variability; data harmonization



Academic Editor: Suvash Chandra Paul

Received: 14 May 2026

Revised: 2 June 2026

Accepted: 3 June 2026

Published: 10 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Three-dimensional concrete printing (3DCP) places construction-grade cementitious materials onto a build surface in continuous filaments without formwork. Layer-wise deposition, time-dependent rheological evolution (i.e., structural build-up and rebuild kinetics under continuous shear), and sensitivity to environmental conditions make the printed product more variable than the same mixture cast into a mould. Civil engineering codes (e.g., EN 1992-1-1 [1] and ACI 318-19 [2]) calibrate safety factors against decades of cast-concrete data; transferring these factors directly to printed elements is not defensible without a quantitative account of the new sources of dispersion. The standardization process for 3DCP structural elements is still developing. Achieving code-compatible design for printed structures will likely require greater data standardization and cross-laboratory consistency.

Process dependence is itself the central reason that cross-study inference is needed. Because individual laboratories use different mixers, pumps, nozzles, control loops, and curing arrangements, a strength model fitted on one laboratory's data rarely transfers to another. The question we address is not whether such variability exists, which is uncontroversial, but how large it is once mixture composition is held fixed in expectation, and what fraction of it can be attributed to undocumented processing decisions. Quantifying this share is a precondition for treating any printed mixture as a candidate for code-compatible design.

Fresh-state rheology, principally static and dynamic yield stress (τ_0) and plastic viscosity, governs both pumpability and post-deposition buildability [3]. Printable windows reflect a tension between viscosity low enough for continuous extrusion and structural build-up rapid enough to resist filament collapse. Mineral additions such as sepiolite or hydrated lime raise early yield stress in fly-ash systems [4] and capillary suction and moisture migration in fine-grained mixtures contribute additional effective confining stress [5]. At the same time, sustainability imperatives are accelerating the move from Ordinary Portland Cement (OPC) toward blended and alternative binders: recycled construction-and-demolition waste (C&DW) powders lower embodied carbon [6], lightweight expanded clay aggregates (LECA) improve thermal performance [7], and ground granulated blast-furnace slag has been shown to raise green strength substantially while fly ash extends open time at the cost of early build-up [8].

To support optimization across this expanding material space, various machine learning approaches have been developed. For instance, LightGBM, XGBoost, and deep learning models have been applied to predict compressive and flexural strengths [9,10]. Certain algorithms target anisotropic compressive strength and slump flow [11], whereas others predict mixture proportions for conventional and alkali-activated printed concretes [12]. Additionally, artificial neural networks have been utilized to model interlayer bond strength [13]. These approaches are valuable when their inferential scope is matched to the underlying data structure. A still-open question is how the openly available datasets support the cross-laboratory generalization that any pooled-data model implicitly assumes. Currently, these datasets show variations in feature selection and reporting depth. For example, some studies use augmented datasets to address missing values and assist in mixture prediction [14], while others combine limited laboratory samples with synthetic data generation for rheological optimization [15]. In addition, specific datasets focusing on low-carbon binders [16] are not always directly compatible with traditional Portland cement databases. Existing databases, review papers, and machine-learning models in the 3DCP literature often treat pooled records as independent and identically distributed (i.e., exchangeable) observations, which can overlook laboratory-specific correlation structures and understate prediction uncertainty. Recent extensive inter-laboratory studies by the RILEM TC 304-ADC committee spanning 30 laboratories, 34 datasets, and approximately 5000 test specimens, have demonstrated that the current lack of standardized testing protocols directly drives inconsistent reporting and order-of-magnitude variations in compressive strength [17]. Building on this empirical foundation, this study provides a statistical meta-analysis to decompose and quantify these variance components across a corpus of 80 published studies. While RILEM's physical round-robin experiment identifies testing discrepancies, our work evaluates the laboratory effect under literature-pooling conditions, treating it as a clustering factor rather than random independent noise.

When critical metadata, including explicit curing temperature and humidity, loading direction relative to the print path, and controller feedback traces [18,19], are not reported, models trained on pooled records have to rely on confounded proxies; the practical consequence is that residual variance is inflated by undocumented moderators rather than

by measurement error, and the conventional assumption that records are independent and identically distributed across laboratories no longer holds. Printed concretes typically exhibit weaker interlayer bonding and pronounced anisotropy compared with cast equivalents, yet matched orthogonal test data remain sparse [20–23], which makes this confounding harder to detect from in-sample diagnostics alone.

This study makes three contributions to the structural-design assessment of printed concrete. First, we assemble a study-aware machine-readable corpus of 80 3DCP studies and quantify the between-laboratory share of variance directly. Second, rather than presenting a pooled regression model as a standalone predictive end-product, we use it as a diagnostic tool. We demonstrate that because process metadata are so poorly documented across the literature, even optimal pooling models yield impractically wide prediction intervals (spanning ≈ 85 MPa). This quantitative bound establishes exactly why adopting our proposed three-tier Minimum Reporting Set (MSRS) is an immediate necessity for the field rather than a distant goal. Third, we assess, through a retrospective audit, what level of MSRS adoption the existing literature already supports. The analytical workflow that delivers these three outputs is summarized in Figure 1.

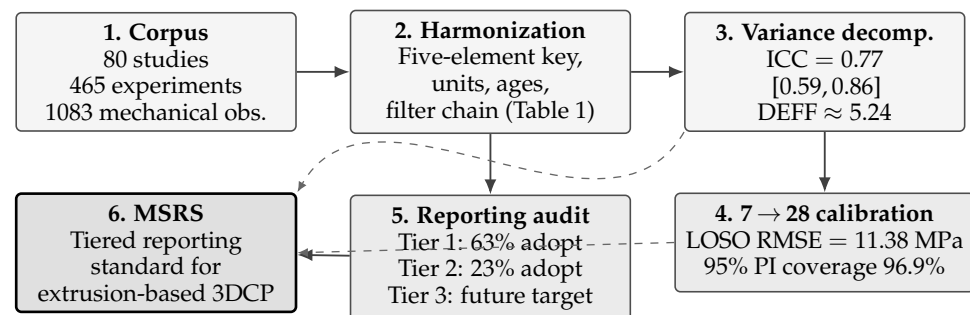


Figure 1. Analytical workflow. Stages 1–2 build the harmonized corpus; stages 3–5 produce the quantitative findings; and stage 6 is the proposed reporting standard.

2. Materials and Methods

2.1. Corpus Assembly and Screening

The corpus was assembled from the peer-reviewed primary literature published in 2017–2025 indexed in Scopus and Web of Science (Core Collection). To our knowledge, it is the most extensive PRISMA-screened literature corpus harmonized to a controlled vocabulary for cross-laboratory variance analysis in 3DCP at the time of writing; we use the more uniformly controlled experimental benchmark of Mechtcherine et al. [17] as a complementary point of reference rather than a competitor. Searches targeted TITLE–ABS–KEY in Scopus and TS in WoS for extrusion-based 3D concrete printing combined with mechanical-performance and rheology/printability terms; queries appear in full in the Supplementary Materials (Section S6).

To ensure a rigorous and transparent selection process, the following inclusion and exclusion criteria were applied:

- **Inclusion Criteria:** (i) Peer-reviewed journal articles and conference proceedings published between 2017 and 2025; (ii) studies focusing on extrusion-based 3DCP; (iii) studies reporting hardened mechanical-property observations (compressive, flexural, tensile, or interlayer bond strength) with explicit curing age and loading directions; and (iv) studies containing primary experimental data.
- **Exclusion Criteria:** (i) Review articles, meta-analyses, and compilations (to avoid double-counting); (ii) studies containing only fresh-state rheological properties or numerical simulations without experimental mechanical tests; and (iii) studies with

severely incomplete or unconvertible mix and process metadata (e.g., missing water-to-binder ratio (W/B) or age information).

After deduplication, screening, and full-text appraisal, the screening workflow progressed as follows:

1. **Identification:** Search queries yielded 73 records from Scopus (69 imported successfully) and 65 from Web of Science, totaling 138 initial records.
2. **Deduplication:** Automated title and DOI matching removed 53 duplicate records, leaving 85 records for screening.
3. **Screening:** Screening titles and abstracts excluded records that did not meet the basic 3DCP domain scope, leaving 80 records for eligibility assessment.
4. **Eligibility:** Full-text appraisal of these 80 studies excluded reviews and studies without extractable record-level mechanical data.
5. **Inclusion:** This yielded a final corpus of 62 studies (61 of which contribute qualified mechanical observations) comprising 465 distinct experiments and 1083 mechanical observations.

Studies that contained only contextual or aggregated findings without verifiable hardened-state metadata were retained for context but excluded from quantitative analysis [4,6,15,16,18,19,24–35]. To avoid double counting, review articles and compilations were excluded from the quantitative analysis. This exclusion encompasses broad overviews on sustainability and material properties [36], as well as systematic reviews of mixture design characteristics [37]. Similarly, reviews focusing specifically on fresh-state rheology [38], fiber reinforcement strategies [39], and advanced materials such as engineered polymer composites [40] were retained to provide background context but removed from the mechanical dataset.

Publication and Selection Bias

Any literature-based meta-analysis is subject to potential biases. First, *publication bias* can occur because mixtures achieving target strengths or successful printability are more likely to be reported, whereas less successful trials (such as those suffering nozzle clogging, collapse, or early cracking) may remain unpublished. The assembled corpus may, therefore, represent a somewhat optimistic view of 3DCP performance. Second, *selection bias* can be introduced by documentation requirements; studies omitting fundamental parameters (such as the W/B ratio or curing age) were excluded. While necessary to maintain model consistency, this criterion might systematically exclude research from settings with less comprehensive instrumentation.

2.2. Harmonization and Analysis Subsets

Each test record received a five-element key (study identifier, mixture label, specimen geometry, loading direction, and test age). Source priority for conflicting duplicates was table > figure > text. Strengths were converted to MPa, dimensions to millimetres, and ages to canonical buckets (1, 3, 7, 14, 28, 56, and 90 days). Mixtures with $W/B \geq 0.60$ that depended on advanced processing (viscosity-modifying admixtures and near-nozzle activation) were retained with a regime flag rather than censored ($n = 5$).

Data Harmonization Examples

Inconsistent raw data were standardly mapped into canonical formats. For specimen geometry, dimensions like “40 × 40 × 160 mm prisms” or “50 mm cubes” were recorded as height H and diameter/width D to compute the aspect ratio H/D . For loading direction, raw descriptions (e.g., “perpendicular to layer interface” and “parallel to print direction”) were mapped to standard orthogonal vectors relative to the print path: direction X (lon-

gitudinal, along the print path), Y (lateral, perpendicular to the path in the horizontal plane), and Z (perpendicular to the layers in the vertical plane). For binders, alternative systems were classified into categorical groups (OPC, OPC+SCM, and AAS/Geopolymer) based on raw mass percentages. Curing environments were harmonized by mapping water curing, fog rooms, and controlled chambers ($RH \geq 95\%$) to a “Moist” flag, while standard air-cured specimens were labeled “Ambient”.

The W/B is computed from total mixing water against total binder mass (cement plus all supplementary cementitious materials), with two protocol-driven adjustments. For alkali-activated and geopolymer systems, we use the activator-water content together with any free water as the numerator and the total reactive solid (precursor plus activator solids) as the denominator, so that the reported value is comparable with the OPC convention rather than with the activator-by-mass convention occasionally used in primary sources. For mixtures containing high-absorption aggregates such as lightweight expanded clay aggregate (LECA) or recycled construction-and-demolition powders, we use the effective W/B reported by the original authors when stated, and otherwise correct the as-mixed water by the aggregate’s saturated-surface-dry water demand using the value reported by the same source. Records for which neither convention can be applied unambiguously are flagged and excluded from regression analyses.

The mechanical analyses use nested subsets, summarized in Table 1. We make these explicit because reader confusion about “how many records?” is otherwise inevitable when several different counts appear in adjacent tables.

Table 1. Nested analysis subsets and their use in this paper. Filter rules are detailed in Supplementary Materials Section S1.

Subset	<i>n</i>	Used in
Total mechanical observations (compressive, flexural, tensile, bond, <i>E</i> , shear, other)	1083	Coverage tables, descriptive figures
Compressive tests (raw, all ages)	562	Coverage table, age plot
Compressive tests after 28 d power-law normalization	550	Dataset summary, W/B–strength panels
Records with ≥ 2 measurements per study within 25–31 d	345	ICC and DEFF (40 studies)
Within-window LOSO evaluation points	257	LOSO metrics, PI coverage
Matched 7-to-28 d pairs	71	Calibration slope estimation
Studies with ≥ 3 ages (growth exponent fit)	33	Power-law exponent histogram

2.3. Statistical Models

2.3.1. Age Normalization

All compressive measurements were referred to a 28-day equivalent through a power-law age model

$$\sigma(t) = \sigma_{28} \left(\frac{t}{28} \right)^b, \quad b \approx 0.47, \quad (1)$$

where *b* is estimated from within-study fits across studies that report at least three test ages (median *b* = 0.47, IQR 0.25–0.58; 33 studies). Records measured within 25–31 days were treated as 28-day equivalents directly and not back-projected. The pooled median *b* is a working approximation: alternative chemical binders (geopolymers, alkali-activated

materials, and high-volume SCM systems) follow age-strength kinetics that need not match Portland-cement-only systems. We use a single b because binder-class-stratified estimation is not yet supported by the available cluster sizes (Section 2.4); the IQR width reported above already absorbs much of this between-binder heterogeneity, and the limitation is revisited in Section 4.1.

2.3.2. Joint W/B–H/D Regression

Mix and process geometry are confounded across laboratories: studies with thicker layers tend to adopt lower W/B to maintain shape stability (Spearman correlation coefficient $\rho = -0.41$, p -value $p < 0.001$). The joint behavior of mixture proportions and extrusion parameters has been examined in various contexts, including investigations of recycled glass printability [41], water-to-solid ratios in foamed concretes [42], and composites for marine applications [43]. Other studies address printability by optimizing particle-size distributions in strain-hardening cementitious composites (SHCC) [44] or by examining the effects of viscosity-modifying admixtures and aggregate volumes [3]. These parameters have also been evaluated for mixtures incorporating waste tire rubber [45], recycled fibers [23], or in studies validating specific laboratory printer configurations [46]. However, the cross-study confounding with W/B is rarely treated explicitly. To separate these effects, we fitted a mixed-effects regression on age-normalized strength,

$$\sigma_{28} \sim W/B + H/D + W/B:H/D + (1 \mid \text{study}), \quad (2)$$

with mean-centred fixed effects and a study-level random intercept. Random slopes for W/B and H/D were considered but did not converge under restricted maximum likelihood: several studies report fewer than five mixtures, so the slope variance components are not identified. We, therefore, retain a random-intercept-only specification and treat this as a methodological limitation (Section 4.1). Cluster-robust (HC3, study-clustered) standard errors were used for fixed-effect inference.

2.3.3. Probability of Attainment

For target strengths $X \in \{30, 40, 50\}$ MPa, attainment probability is fitted as a logistic surface,

$$\Pr(\sigma \geq X) = \text{logit}^{-1}(\gamma_0 + \gamma_1 W/B + \gamma_2 H/D + \gamma_3 W/B:H/D), \quad (3)$$

restricted to the 5–95% observed envelope of W/B and H/D so that contour estimates are not extrapolated.

2.3.4. Cross-Laboratory Error Budget

Prediction error was estimated by leave-one-study-out (LOSO) cross-validation: each fold withholds one entire study, refits the calibration on the remainder, and predicts the held-out points. We deliberately chose LOSO over standard K-fold cross-validation to prevent data leakage. Given the high ICC (heavy clustering by laboratory), random K-fold splits would distribute records from the same laboratory into both training and validation sets, artificially inflating apparent accuracy. LOSO simulates the real-world challenge of predicting outcomes for an unseen laboratory. We report the coefficient of determination (R^2), mean absolute error, root-mean-squared error, and the empirical coverage of nominal 95% prediction intervals computed at the level of single observations. The slightly elevated RMSE/MAE ratio compared with normal-distribution expectations provides a quantitative indicator of heavy-tailed error distributions, consistent with undocumented catastrophic failures like cold-joint delamination or sudden collapse in 3DCP.

2.4. Sample-Size Adequacy

The reviewer-facing question is whether 80 studies and 1083 mechanical observations are sufficient for the inferences made. We treat each inference separately.

ICC and design effect. With 345 records nested in 40 studies and an average cluster size of $\bar{k} \approx 8.6$, the two-level approximation of Snijders and Bosker [47] gives $\geq 80\%$ power to reject $ICC = 0$ when the true value exceeds 0.10 at $\alpha = 0.05$. The estimated value is 0.77, so the test is not power-limited and the 95% confidence interval [0.59, 0.86] is the binding statement of precision.

Cross-laboratory calibration. The 7-to-28-day calibration uses 71 matched pairs from 21 contributing studies. Treating the pair set as the inferential unit, the standard error of the OLS slope is approximately 0.14 at the observed dispersion, giving $\geq 80\%$ power to detect a slope departure of $|b - 1| \geq 0.30$ at $\alpha = 0.05$. This is sufficient for the screening claim made (the 7-day strength is informative for the 28-day value to within roughly ± 11 MPa) but explicitly insufficient for sub-strata such as binder class, where we report only descriptive medians.

Anisotropy and curing strata. Sub-strata that fall below ten observations (notably the ambient-cured stratum and the matched orthogonal-direction stratum) are reported as descriptive statistics with explicit caveats and are not used to draw inferential conclusions. A power-aware reading of the corpus is that compressive screening models can be fitted reliably; bond, anisotropy, and elastic-modulus models cannot.

2.5. Software and Reproducibility

Analyses were performed in Python 3.11 using the pandas, statsmodels, and scikit-learn libraries; mixed-effects models were fitted with the linear mixed-models routine in statsmodels. Aggregated counts, filtering rules, the complete study inventory, and an evidence-backed version of the MSRS table are reproduced in the Supplementary Materials.

3. Results

3.1. Dataset Characteristics and Reporting Coverage

The harmonized corpus spans the contemporary design space of extrusion-based 3DCP. Table 2 reports central tendencies and dispersion for the principal numerical fields after 28-day normalization (compressive strength) or in their original units (process geometry). The substantial range observed both in the W/B (0.20–0.90) and in the layer-height-to-nozzle-diameter ratio (H/D; not directly reported in this table but derived later) sets the boundaries within which the regression models in Section 3.2 can be considered descriptive.

Table 2. Descriptive statistics for the harmonized analysis subset.

Variable	<i>n</i>	Mean	SD	Min	Max
Compressive strength (MPa, 28 d eq.)	550	46.8	23.3	1.2	122.9
Test age (days, raw)	496	21.8	10.4	0	90
W/B ratio	414	0.335	0.070	0.200	0.900
Layer height (mm)	272	11.4	4.8	5.0	30.0
Nozzle diameter (mm)	210	21.0	9.0	3.0	60.0
Print speed (mm/s)	255	55.5	27.4	8.3	200.0

Reporting coverage across the full corpus (Table 3) shows that mixture-level descriptors are documented at usable rates while critical extrusion descriptors are not. Path width is documented in roughly 9% of physical tests, extrusion pressure in 1.8%, and the Herschel–Bulkley parameters K and n in fewer than a quarter of the tests. Loading direction is tagged in 62 of 62 included studies but matched orthogonal pairs at a common age and property are absent, so a cross-study anisotropy ratio cannot be estimated from the present corpus.

Table 3. Reporting coverage of mixture, process, mechanical, and rheological parameters across the harmonized corpus. Counts are at the experiment level for mixture and process rows ($N = 465$); at the test/record level for mechanical rows (which can exceed the experiment count when a single experiment yields multiple specimens) and for rheology data-points; and at the experiment level for rheology experiment rows.

Category	Parameter	<i>n</i>	Type	Strength	Remark
Overall	Studies in corpus	80	–	–	PRISMA-included primary studies
Overall	Studies screened	85	–	–	Post-deduplication
Overall	Total experiments	465	–	–	Baseline corpus size
Mix	W/B ratio	170	numerical	medium	Tails sparse
Mix	Binder class	465	categorical	medium	Heuristic labelling
Process	Nozzle diameter	136	numerical	low	Under-reported
Process	Layer height	154	numerical	medium	Usable subset
Process	Print speed	148	numerical	medium	Canonicalized to mm/s
Process	Path width	98	numerical	low	Sparse
Process	Extrusion rate	68	numerical	low	Rare quantitative
Process	System type (1K vs. advanced)	465	categorical	medium	Heuristic; lower bound
Mechanical	Compressive (tests)	562	numerical	high	Abundant
Mechanical	Flexural (tests)	296	numerical	medium	Moderate
Mechanical	Splitting tensile (tests)	8	numerical	low	Very limited
Mechanical	Direct tensile (tests)	34	numerical	medium	Protocol variation
Mechanical	Bond/interlayer (tests)	76	numerical	low	Under-reported
Mechanical	Elastic modulus (tests)	24	numerical	low	Sparse
Mechanical	Shear (hardened tests)	0	numerical	very low	Nearly absent
Mechanical	Other/unmapped (tests)	83	numerical	–	Heterogeneous
Mechanical	Loading direction (tests)	374	categorical	low	Few orthogonal pairs
Rheology	H–B model (K, n) experiments	42	numerical	low	Heterogeneous fits
Rheology	Yield stress (experiments)	218	numerical	low	Few time series
Rheology	Yield stress (datapoints)	355	numerical	low	Multiple per experiment

A graphical view of overall reporting coverage (Figure 2) makes the imbalance between mixture-level and process-level documentation visible at corpus scale; the temporal distribution of included studies and measurement counts is provided in Supplementary Materials Section S4.

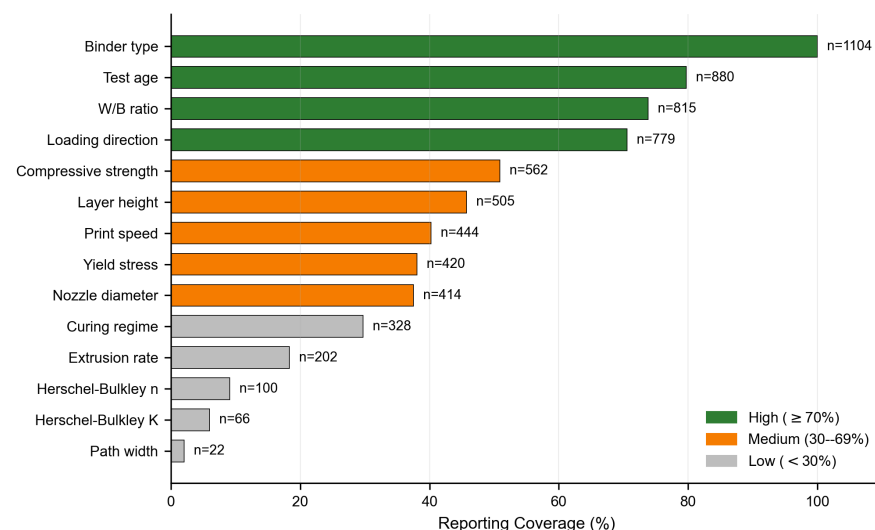


Figure 2. Reporting coverage for key mix, process, and hardened-property fields across the harmonized corpus, expressed at the record level ($N = 1104$). Shorter bars indicate sparser reporting.

3.2. The Laboratory Effect: ICC and Design Effect

Compressive strength is plotted against W/B in Figure 3. The expected negative relationship dominates the pooled regression line in panel (c), but the within-study fits also

show large vertical offsets between laboratories at any fixed W/B; that is, the effect that the laboratory cluster contributes to total scatter is comparable to, or larger than, the effect of W/B itself.

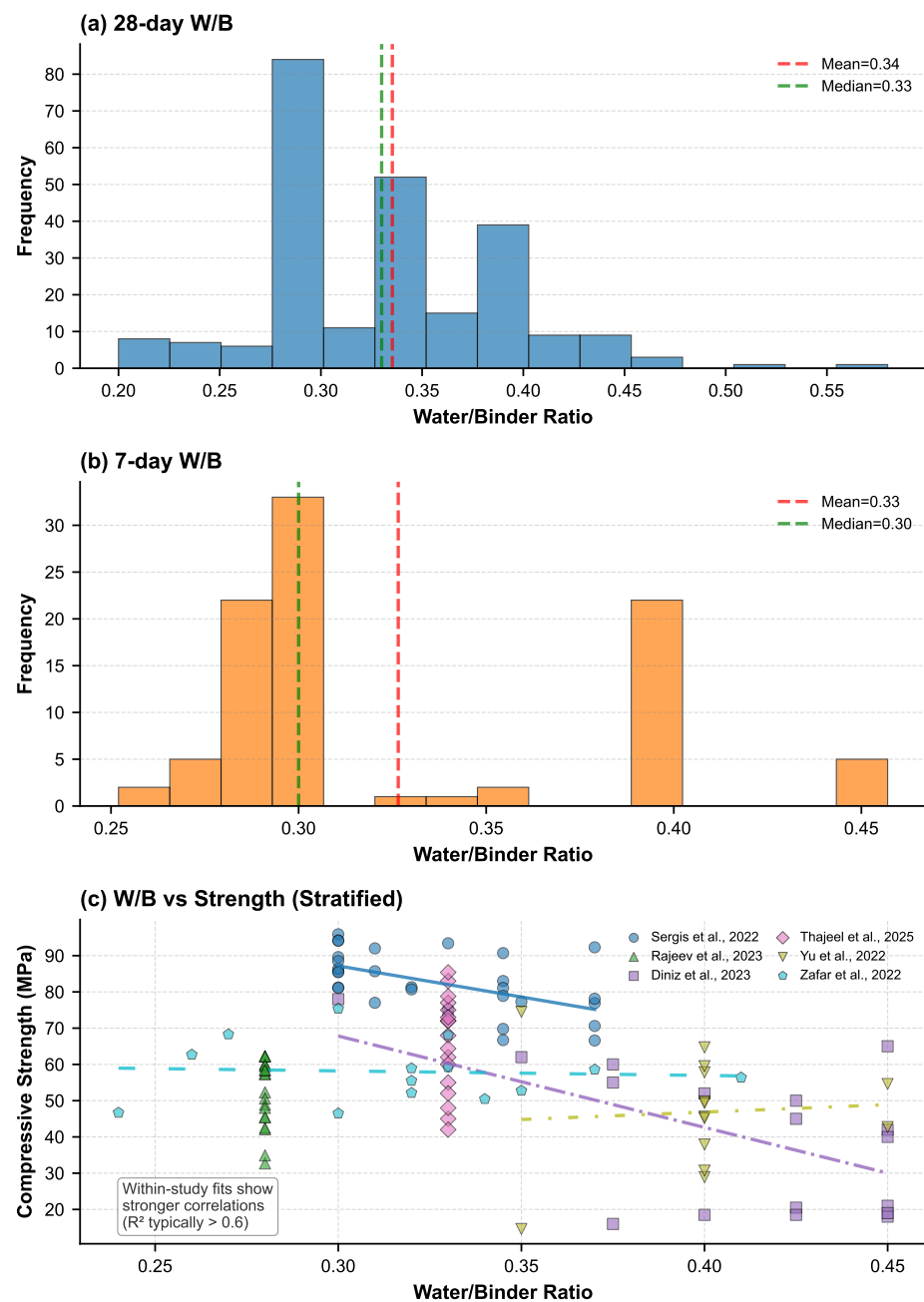


Figure 3. W/B distributions and the W/B–strength relationship. (a) 28-day W/B histogram; (b) 7-day W/B histogram; and (c) compressive strength against W/B with study-stratified (grey) and pooled (heavy) OLS fits [23,33,48–51].

To quantify the laboratory share directly we computed the intraclass correlation coefficient on the records in the 25–31 day window with two or more measurements per study (345 records in 40 studies; mean cluster size $\bar{k} \approx 8.6$, and median cluster size $\tilde{k} = 6.5$):

$$\text{ICC} = 0.77 \quad [95\% \text{ CI } 0.59, 0.86], \quad (4)$$

estimated by random-intercept REML fitting and bootstrapped over 1000 study-level resamples (Figure 4). In engineering terms, two specimens printed in the same laboratory are sta-

tistically much more similar to each other than to two specimens printed in different laboratories, and we are separating the share of variance that follows from this clustering from the share that follows from the mixture itself. More than three quarters of the variance in 28-day compressive strength lies between studies. Treating the corpus as if observations were exchangeable across laboratories collapses the design effect and substantially understates prediction uncertainty. Because the cluster-size distribution is right-skewed, we report the design effect at the median cluster size as $DEFF \approx 1 + (\bar{k} - 1) ICC = 1 + 5.5 \times 0.77 \approx 5.24$, which means the effective independent sample size for cluster-aware inference is roughly one fifth of the nominal record count. The mixed-effects model in Equation (2) was fitted on the subset of $N = 93$ records from 11 studies where both W/B and H/D are jointly documented. The estimated fixed-effect coefficients and their 95% confidence intervals (CIs) are as follows:

- Intercept (at the dataset centroid): 61.36 MPa (95% CI [37.11, 85.60], p -value $p < 0.001$);
- W/B (mean-centered): -210.00 MPa per unit (95% CI $[-416.69, -3.28]$, p -value $p = 0.046$);
- H/D (mean-centered): -19.59 MPa per unit (95% CI $[-108.61, 69.43]$, p -value $p = 0.666$);
- Interaction $W/B \times H/D$: 162.50 (95% CI $[-497.31, 822.30]$, p -value $p = 0.629$).

The W/B term exhibits a statistically significant negative gradient, confirming that lower water-to-binder ratios correlate with higher strengths across pooled studies, which is consistent with classic concrete chemistry. However, neither the main effect of H/D nor the interaction term is statistically significant, suggesting that geometric variation does not follow a simple linear trend once clustered by study.

The model resolves a between-study variance of $\hat{\sigma}_u^2 = 574.03 \text{ MPa}^2$ for the study intercept and a within-study residual variance of $\hat{\sigma}_e^2 = 361.71 \text{ MPa}^2$. A likelihood-ratio test rejects the null of zero between-study variance ($\chi^2(1) = 55.52$, p -value $p < 0.001$). The conditional R^2 (variance explained by fixed effects plus the random intercept) is approximately 0.697, while the marginal R^2 (fixed effects only) is approximately 0.215: most of the explained variance comes from knowing which laboratory produced the observation, not from the W/B and H/D values themselves.

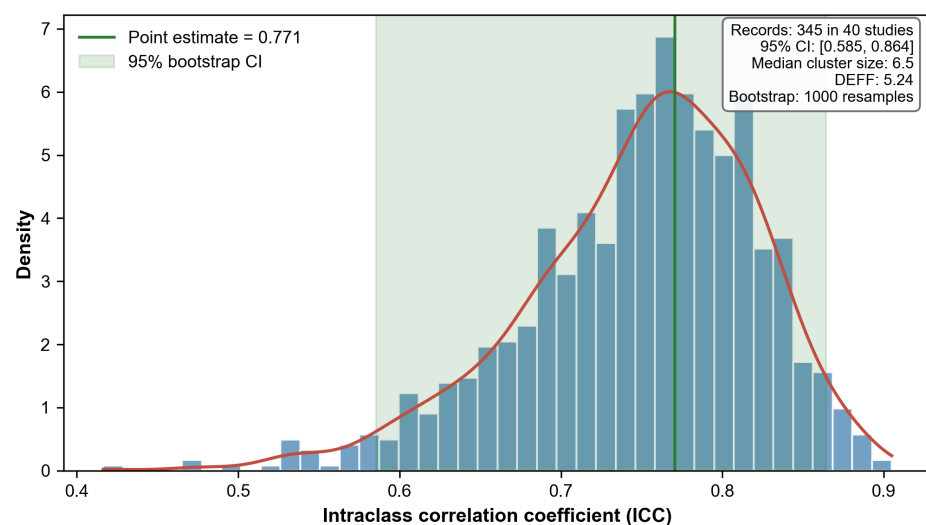


Figure 4. Bootstrap density of the intraclass correlation coefficient on 345 records nested in 40 studies in the 25–31 d window with ≥ 2 measurements per study. The point estimate is 0.771, with a study-level cluster bootstrap 95% CI of [0.585, 0.864] over 1000 resamples; and the median cluster size is 6.5, giving a design effect $DEFF \approx 5.24$.

3.3. Strength Development and a 7-to-28-Day Quality-Control Trajectory

Within-study age–strength growth curves are markedly more consistent than the cross-study spread (Figure 5): laboratories that report multiple ages produce well-behaved power-law trajectories even when their absolute levels differ. Power-law exponents fitted within each study cluster around a median $b = 0.47$ (IQR 0.25–0.58; 33 studies contributing ≥ 3 ages, Figure 6), which supports the use of Equation (1) to refer measurements made at neighbouring ages onto a 28-day-equivalent basis.

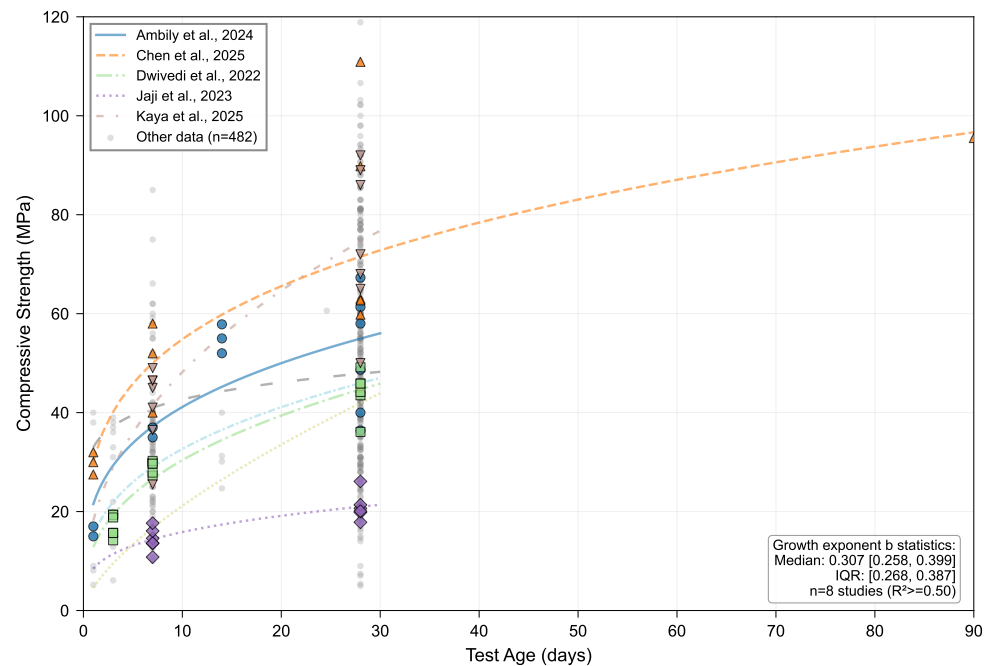


Figure 5. Compressive strength against test age (raw compressive subset, $n = 562$). Coloured trend lines are within-study power-law fits for five exemplar studies and grey points are the remaining records. The annotation box reports a stricter subset, restricted to fits with $R^2 \geq 0.50$ ($n = 8$ studies and median $b = 0.31$) and the looser filter used in Equation (1) retains 33 studies (Figure 6) [8,46,52–54].

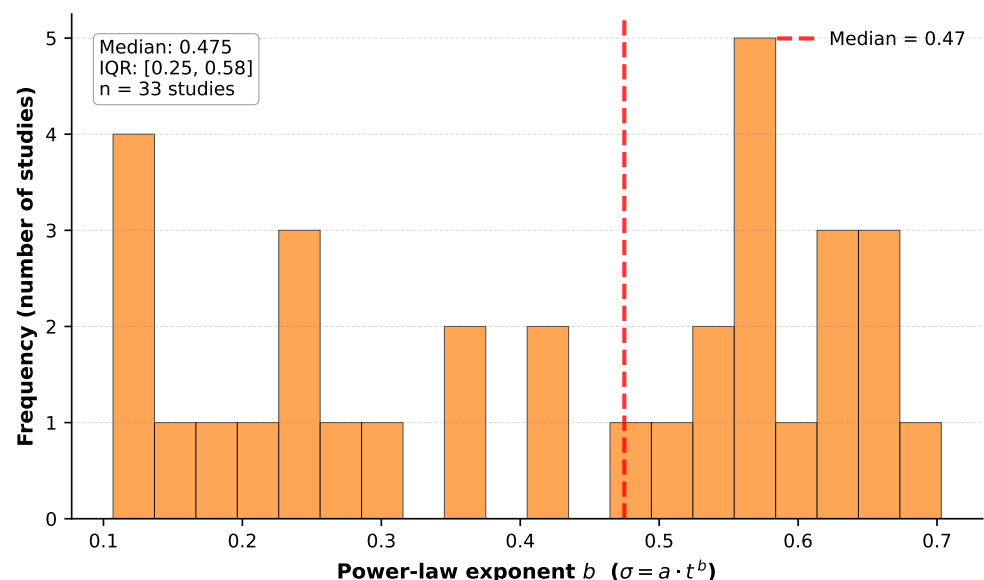


Figure 6. Distribution of within-study power-law growth exponents b from $\sigma = a t^b$ fits (33 studies that report at least three test ages) and median $b = 0.47$, IQR 0.25–0.58.

The early-age window most relevant to on-site quality control is the 7-to-28-day pair. We anchor the calibration here for three reasons. First, on-site practice routinely uses a 7-day measurement to forecast the 28-day acceptance value, so a calibration in this window has direct quality-control utility. Second, the 1-to-3-day range, although well represented in the corpus, is dominated by binder- and admixture-specific early-hydration kinetics: a single power-law exponent does not transfer across mixtures that differ in accelerator chemistry, fly-ash content, or near-nozzle activation, so a 1-to-28 calibration would convolve normalization error with cross-laboratory variability. Third, restricting the corpus to 28-day-only records would shrink the available evidence base by more than half because the bulk of contemporary 3DCP studies report 7- or 14-day strengths alongside, or in place of, the 28-day value. The 7-to-28 window, therefore, preserves coverage while keeping the age-normalization step on a part of the hydration trajectory where the C–S–H network is established and a single power-law model is defensible.

We use this window to address a practical question: given a 7-day compressive measurement on a printed mixture from an unseen laboratory, how tightly can a generic, uncalibrated model bracket the eventual 28-day strength? The calibration was fitted on 71 matched within-study pairs (median linear slope 1.06, MAE 7.90 MPa, and RMSE 10.41 MPa). Generalizability was then evaluated under leave-one-study-out cross-validation across 49 study folds (257 held-out evaluation points), which gives a realistic gauge of out-of-laboratory performance (Table 4).

Table 4. Cross-validation summary for the 7-to-28-day compressive calibration. Metrics are aggregated across leave-one-study-out folds at the record level.

Metric	Value
R^2 (pooled across folds)	0.625
Mean absolute error (MPa)	8.67
Root-mean-squared error (MPa)	11.38
Empirical 95% prediction-interval coverage	96.9%
Evaluation points (record-level n)	257
Study folds	49

The empirical coverage of the 95% prediction band is close to nominal across the observed strength range (Figures 7 and 8), and remains conservative under fold rotation. The width of the band (≈ 85 MPa at the median, including the residual variance from omitted moderators) is wide enough that the calibration must be read as a screening tool rather than as a substitute for a laboratory-specific model. We return to this point in Section 4.4.

A curing-stratified version of this calibration is provided in Supplementary Materials Section S1; we keep it separate from the main result because the ambient stratum ($n = 7$) is below the inferential threshold defined in Section 2.4 and is descriptive only.

The age-only attainment envelopes for 30/40/50 MPa targets, fitted on records with valid age and strength tags, are shown in Figure 9. They provide a single-variable lower bound on the probability of meeting a structural-grade target and are used in Section 4.4 to illustrate the practical penalty associated with omitted process metadata.

A descriptive curing-stratified breakdown of these 7-to-28-d-gain ratios is reported in Supplementary Materials Section S1; we have kept it out of the main results because the ambient stratum ($n = 7$, single study) falls below the inferential threshold defined in Section 2.4 and is descriptive only.

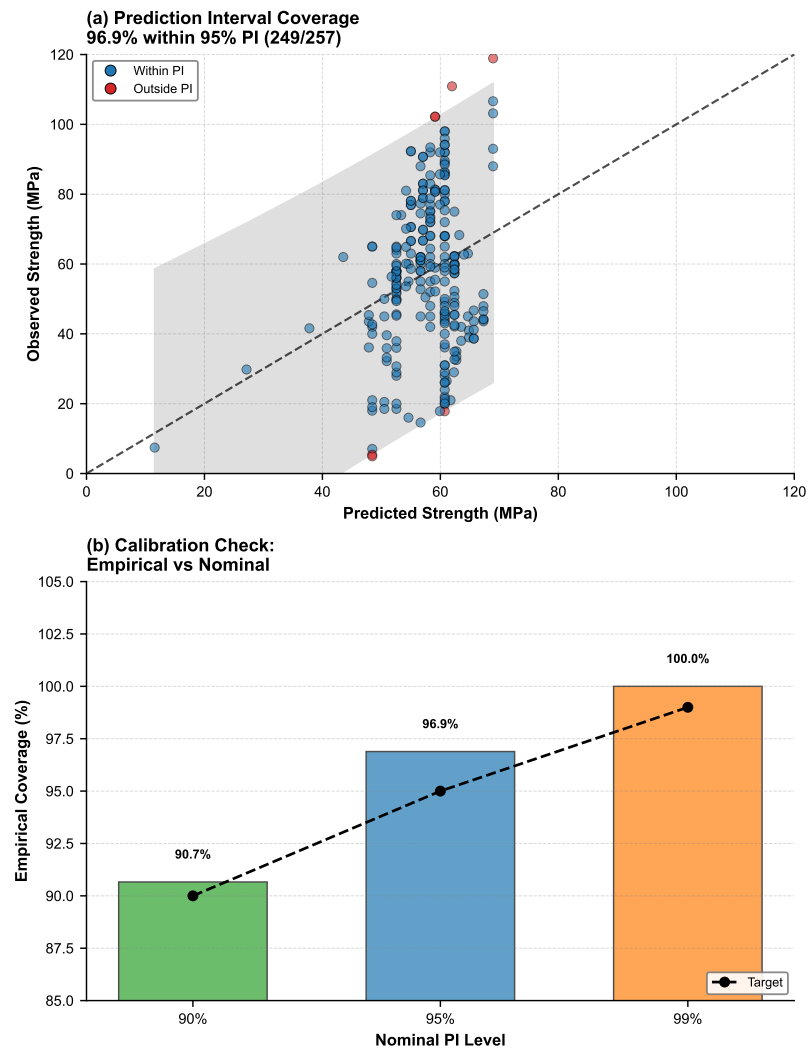


Figure 7. Prediction interval diagnostics under LOSO cross-validation. (a) Observed against predicted 28-day compressive strength with 95% prediction-interval shading and (b) empirical against nominal coverage at 90, 95, and 99% levels. The dashed line is the line of perfect coverage.

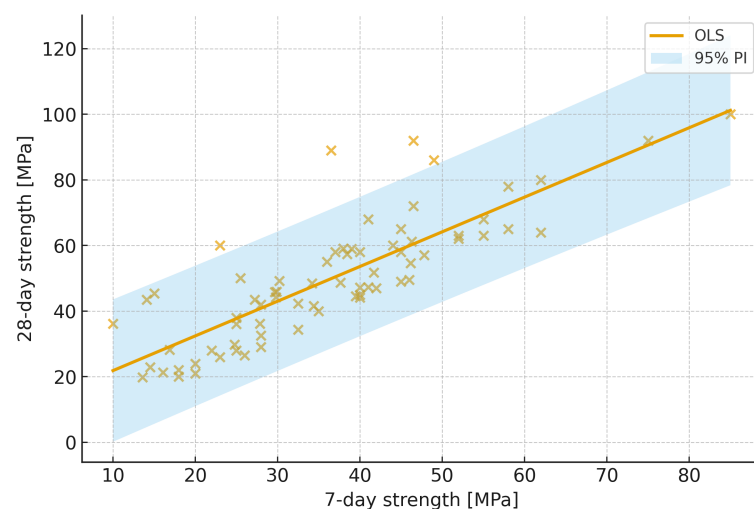


Figure 8. Attainment of 28-day-equivalent compressive strength from 7-day measurements (71 matched pairs). The solid line is the OLS fit and the shaded band is the calibrated 95% single-sample prediction interval.

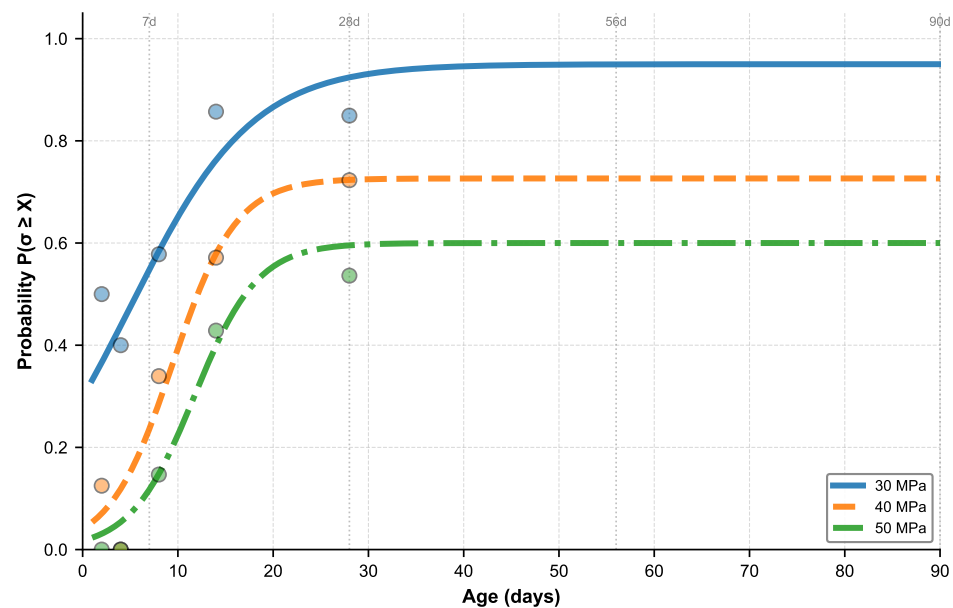


Figure 9. Age-only attainment probability for compressive targets $X \in \{30, 40, 50\}$ MPa, computed on records with valid age and strength tags. Curves are logistic fits; markers are empirical bin frequencies at canonical ages.

3.4. Mechanical Interactions and Composition Trends

Pairs of mechanical tests measured on the same mixture within a ± 1 day window allow direct estimation of cross-property ratios. The flexural-to-compressive ratio has a median of 0.19 (IQR 0.14–0.23) and the tensile-to-compressive ratio is approximately 0.07. The interlayer-bond-to-compressive ratio averages only 0.024, that is, the layer interface fails at a stress level roughly two orders of magnitude smaller than the bulk crushing strength of the same mixture. This ratio is the most direct quantitative signature, in the present corpus, of cold-joint formation between successively deposited layers, and it is sensitive to the inter-layer time gap, ambient temperature and humidity at the moment of deposition, and the pumping pressure profile maintained near the nozzle—all variables that the corpus reports inconsistently. The under-reporting of these moderators is a substantial driver of the laboratory effect identified in Section 3.2, and the case for elevating their reporting to MSRS Tier 2 and Tier 3 (Section 4.3) rests on this connection.

When the corpus is stratified by primary binder class, the W/B–strength relationship differs in terms of slope as well as level. Standard OPC-based mixtures ($n = 198$) show a median 28-day compressive strength of 38.2 MPa with a steep penalty gradient of -47.3 MPa per unit of W/B (p -value $p < 0.001$). Alternative chemical binders, namely geopolymers and alkali-activated materials in our corpus ($n = 42$), show a higher median (44.7 MPa) with a flatter, statistically non-significant gradient (-22.1 MPa per unit of W/B). The interpretation we draw is descriptive: in the alternative-binder subset, activator chemistry and water- and chemical-bonding budgets compete with the bulk water mass for control of terminal strength, so a single-W/B target carries less of the predictive weight [12,53]. Inferential subgroup tests are deferred until binder-class metadata are reported with sufficient density to satisfy the power threshold of Section 2.4.

4. Discussion

4.1. Physical Origins of the Laboratory Effect

The variance decomposition in Section 3.2 establishes that two laboratories using nominally identical W/B and H/D values can produce substantially different compressive

strengths. The W/B–strength relationship itself is reproducible and behaves as established cement-chemistry expectations dictate. The remaining dispersion has to come from somewhere, and the candidates are not exotic: they are the processing variables that the literature does not consistently report. At the micro-structural level, this variance traces to unrecorded differences in hydration kinetics driven by supplementary cementitious materials (SCMs), rheometer geometry differences (e.g., vane spindles, parallel plates, or coaxial cylinders) complicating yield-stress targets, and micro-cracking induced by early plastic shrinkage. Process-level differences in the structural build-up rate A_{thix} between mixtures, in the time-dependent yield-stress trajectory imposed by near-nozzle activation strategies, in the inter-layer time gap and the print speed that together determine interfacial conditioning at the moment a new filament is deposited, and in the curing history (especially the relative humidity and temperature profile during the first 24 h) all act as hidden moderators. Where they go undocumented, they are absorbed into the residual variance of pooled regressions, and the apparent error of the model overstates the underlying material variability. Viscosity-modifying admixtures further illustrate the trade-off: they can raise yield stress and improve shape stability while reducing the eventual compressive strength, so a W/B-only response surface gives a misleading picture of the workability–strength frontier in their presence [3,25,55].

The estimated ICC of 0.77 represents a combined study-level effect rather than a pure reflection of laboratory equipment performance. This value reflects several confounding differences across publications: (i) chemical binder configurations, such as OPC vs. geopolymers activator ratios or variable SCM reactivity; (ii) testing standards, testing machine stiffnesses, and loading rates; (iii) variations in curing regimes and early-age environmental conditions; and (iv) specimen geometry standardizations. In literature-pooling conditions, these factors can act in concert to shift the group-level intercept, contributing to the apparent cluster effect.

The literature examples illustrate how these unstandardized variables can influence the laboratory effect. Curing history is a significant factor, as drying in ambient air instead of moist curing can restrict hydration and induce micro-cracking, which tends to reduce long-term compressive strength. Regarding process parameters, changes in the H/D ratio alter the contact area and pressure between layers, potentially introducing macro-voids that lower compressive capacity. Additionally, printing speed can affect dynamic compaction and higher path velocities under constant volumetric flow can reduce compaction pressure and bulk density, leading to lower strength. When these parameters are not reported, their physical consequences are absorbed into the residual variance, contributing to the study-level clustering.

Mixed-effects estimation is not a stylistic preference but a methodological necessity here. Pooling treats every record as an independent draw from a common conditional distribution, and the data do not support that assumption. The random-intercept formulation we use in Equation (2) accepts this and absorbs the unexplained between-laboratory shift into a single parameter. A richer specification with random slopes for W/B or H/D would, in principle, separate the laboratory shift from a laboratory-specific sensitivity, but it requires either more mixtures per study than several of our included sources report, or stronger priors than we are willing to set in a non-Bayesian frequentist fit. The slope variance components do not converge under restricted maximum likelihood (REML) with the current data density. Forcing a parallel-slope assumption unavoidably introduces some materials-science bias as different extrusion systems and chemical admixtures react uniquely to W/B variations. We record this as an open methodological question: overcoming this REML convergence failure will require either exploring a Bayesian hierarchical

framework with weakly informative priors on a subset of highly-reported laboratories, or a shift in the field towards higher per-laboratory mixture coverage.

4.2. Confounding Between Mixture and Process Geometry

A second methodological consequence concerns the univariate effects of layer height or nozzle diameter, which cannot be read off the corpus directly. Univariate regressions return null effects ($p > 0.67$ for both), but the interpretation is not that geometry does not matter; it is that, across our sources, laboratories that use thicker layers tend to use lower W/B to maintain shape stability ($\rho = -0.41$, $p < 0.001$). The two predictors are, therefore, confounded at the design stage. The joint W/B $\times$ H/D specification in Equation (2) partly disentangles them but it cannot fully recover the causal structure; for that, factorial designs that vary geometry independently of W/B within a single laboratory are needed. The physical mechanism behind the H/D effect is well-established in principle: taller layers relative to nozzle diameter accumulate more interlayer interfaces per unit height. However, nozzle geometry is often analyzed in conjunction with the volumetric extrusion rate and printing speed as these parameters jointly influence the dynamic compaction pressure that affects bond-line strength. A causally interpretable estimate of this effect is, however, beyond what the present meta-corpus can provide.

The third related limitation is the severe under-reporting of orthogonal-direction pairs. Only 62 of the 62 included studies tag loading direction at all, preventing a cross-study estimate of the anisotropy ratio from our corpus. Yet, experimental evidence confirms that directionality and cold joints dictate failure. The RILEM TC 304-ADC benchmark reported compressive strength reductions of 20.3% in the perpendicular (V) direction and 18.2% in the longitudinal (W) direction compared with cast specimens, with cold-joint interfaces driving an additional $\sim 10\%$ penalty in the loading directions U and V while remaining negligible in W [17]. Crucially, this anisotropy destabilizes variance itself: the same benchmark reported extreme cases (e.g., Lab 05) where the intra-laboratory coefficient of variation reached 43% in the V direction while remaining at 7% in U and 4% in W [17]. From a structural-design perspective, these directional shifts mean that code-compatible safety factors must cover the weakest orientation, which cannot be quantified from the literature without matched directional reporting.

4.3. The Minimum Reporting Set (MSRS)

The combination of an ICC of 0.77 and the documentation gaps in Table 3 establishes a clear practical implication. Cross-laboratory predictive models for 3DCP will not become safe to use in design without a uniform reporting baseline that turns currently hidden moderators into explicit, machine-readable covariates. We propose a three-tier MSRS; Table 5 summarizes the parameter list and the tiered structure. The full evidence-backed table, with method-tag specifications and a per-parameter rationale, is reproduced in Section S5 of the Supplementary Materials so that practitioners can apply it directly without consulting the authors.

The correlation structure observed in the corpus (Figure 10) supports the design directly. The cells with the highest pairwise sample sizes show interpretable, physically reasonable associations: compressive–flexural strength, $\rho = 0.74$; flexural–curing relative humidity, $\rho = 0.81$; bond–H/D ratio, $\rho = 0.82$; and yield-stress–print-speed, $\rho = -0.75$. The W/B–compressive correlation is small and negative once binder class and age are pooled together, which is consistent with the binder-class slope contrast reported in Section 3.4 and is itself a piece of evidence for the need to disaggregate. Many cells, especially those involving the Herschel–Bulkley parameters or the bond strength, return

“–” because the pairwise sample size falls below five; these are precisely the cells the MSRS is designed to populate, tier by tier.

Table 5. Proposed Three-Tier Minimum Standardized Reporting Set (MSRS) for 3DCP. Priority levels, formats, target user groups, and mechanistic benefits are defined to standardize literature reporting.

Tier	Parameters	Reporting Format	Priority Level	Target User Group	Mechanistic Rationale/Benefit
Tier 1	- Water-to-binder (W/B)	- Mass ratio (report basis)	Essential for Publication (Minimum for basic paper review and comparability)	All authors, reviewers, and journals	Enables basic normalization, orientation classification, and first-order dimensional check. Compliance is low-effort.
	- Loading direction (X/Y/Z)	- Angle relative to path			
	- Exact specimen age	- Days (decimal for <24 h)			
	- Layer height (H) & Nozzle diameter (D)	- Millimeters (mm)			
	- Print speed	- mm/s			
Tier 2	- Extrusion mass flow rate	- g/s	Recommended for Pooled Modeling (Allows cross-laboratory predictive calibration)	Data-driven modelers and material developers	Captures hydration kinetics, shrinkage/cracking risks, filament compaction, and cold-joint formation. Narrows model prediction intervals.
	- Curing environment	- Temp (°C), RH (%), duration			
	- Stand-off distance	- mm			
	- Interlayer time interval	- Seconds (s)			
	- Binder class	- Mass percentages			
Tier 3	- Controller pressure/torque	- kPa or N·m (time-series)	Required for Structural Design (Necessary for code-compliant safety factor calibration)	Structural-design engineers and safety certifiers	Provides real-time process feedback, early-age shape stability, and interfacial bond capacity. Essential for serviceability (deflection) and structural reliability.
	- Fresh yield stress series $\tau_0(t)$	- kPa (vane geometry noted)			
	- Thixotropic build-up (A_{thix})	- Pa/s			
	- Elastic modulus (E)	- GPa			
	- Interlayer bond strength	- MPa (direct tension)			

The proposed MSRS primarily focuses on fresh-state rheology and early-to-standard strength parameters. Under a broader structural safety framework, similar to advancements in other engineering systems (e.g., steel–concrete composite walls), characterizing long-term creep and shrinkage deformations [56] and fire resistance under combined loading [57] remains a key consideration for service-life performance. As 3DCP structural-design codes mature, incorporating long-term stability and fire-safety metrics into reporting protocols could support more comprehensive design integration.

Bond strength is the clearest case for action. It is severely under-reported (less than 7% of mechanical observations) despite being the dominant failure mechanism for layered geometry, and elastic-modulus measurements total only 24 across the entire corpus. This limited availability of multi-directional and interface data is noteworthy because relying on compressive strength alone may not fully capture general structural performance. Printed concrete typically exhibits anisotropy and contains discrete interfaces. The boundaries between layers represent potential shear and tensile planes, where the bond strength is often a small fraction (averaging 2.4% in this corpus) of the bulk compressive strength. Consequently, failure modes in printed components can be governed by interface de-

lamination, interlayer sliding, or buckling, rather than material crushing. In addition, conventional design codes (e.g., ACI 318-19 and EN 1992-1-1) rely on empirical formulations that estimate tensile capacity, shear resistance, and elastic modulus directly from compressive strength. For 3DCP, these conventional scaling relationships may not directly apply, as recent benchmarks suggest that printed elements can require higher compressive strengths to achieve the same elastic modulus as cast equivalents [17]. The gap is particularly consequential for emerging printed concretes that depart from the conventional dense-matrix archetype, such as foamed and variable-density mixtures developed for thermally insulating envelopes [58]: their stiffness distributions are essentially undocumented at the cross-study level. Without reliable stiffness data, neither EN 1992-style deflection verifications nor ACI 318-style serviceability assumptions can be transferred to printed elements with code-compatible confidence.

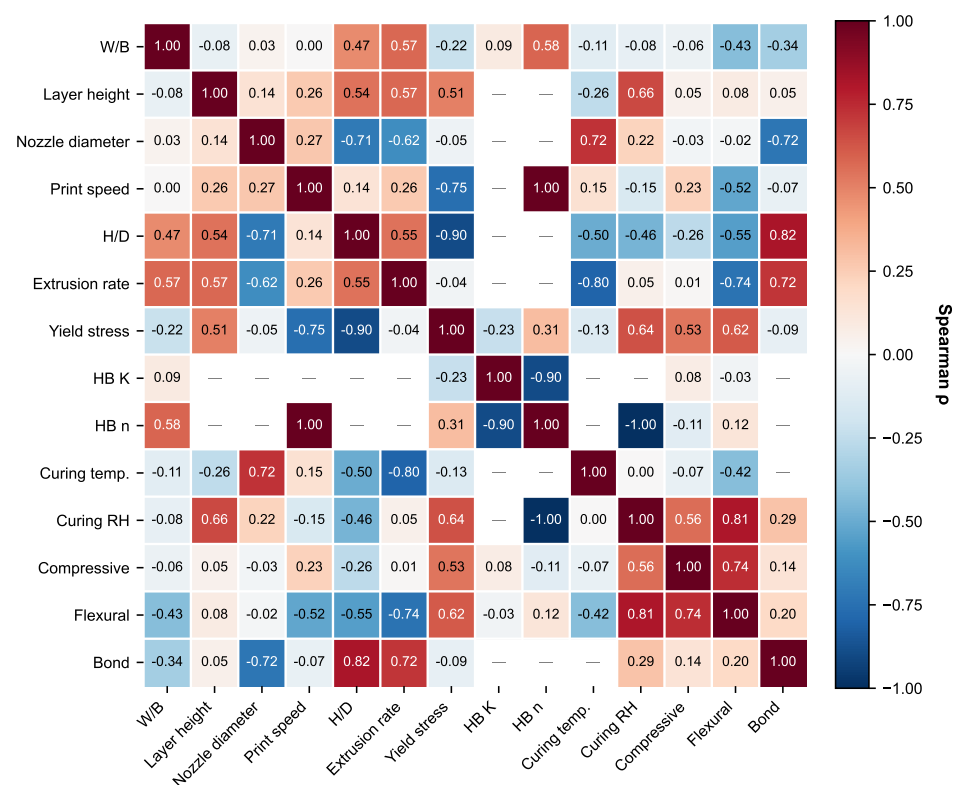


Figure 10. Spearman rank correlations between mixture, process, rheological, curing, and mechanical descriptors at the experiment level for records in the 25–31 d window. Cells display ρ and “--” marks pairs with fewer than five overlapping experiments. The accompanying pairwise- N matrix is provided as a CSV alongside the Supplementary Figures S1–S3.

A retrospective audit of the included corpus indicates that 63% of studies already satisfy Tier 1 in full, 23% satisfy Tier 2, and none satisfy Tier 3 in its entirety. The interpretation is not that the field is failing to report; it is that the bar for cross-laboratory predictive modeling has moved, and a coordinated step-change in reporting at Tiers 2–3 is what now blocks code-compatible inference.

4.4. Implications for Screening and Design

The cross-validated calibration produces a prediction interval whose mean width is approximately 85 MPa under LOSO. The 95% empirical coverage is close to nominal, so the band is honest about the variability the corpus contains; but it is too wide to be

substituted for a laboratory-specific verification programme. The screening role we propose is the following.

Consider a practitioner targeting $f'_{c,28} \geq 40$ MPa with $W/B = 0.35$ and $H/D = 0.50$ in an unseen laboratory. The point prediction from Equation (2) is $\hat{\sigma}_{28} \approx 44.2$ MPa, and the calibrated 95% prediction interval spans [21.5, 66.9] MPa. The age-only attainment surface gives $\Pr(\sigma \geq 40) \approx 0.58$ at this design point. To reach $\Pr \geq 0.80$ without on-site data, the practitioner has to move to $W/B \approx 0.28$ or $H/D \approx 0.35$. Alternatively, even a small set of laboratory-specific reference measurements would, under the random-intercept model, cut the prediction interval by an estimated 35–50% by replacing the pooled intercept with a study-specific estimate. The implication for structural design is direct: an uncalibrated pooled model is useful for screening alternatives at the planning stage, but code-compatible verification has to remain laboratory-specific until process-level reporting matures to the point where mechanistic moderators can be entered as covariates.

4.5. Limitations

Three limitations bound the inferences in this paper and are worth stating explicitly.

(i) *Single age-normalization exponent.* The 28-day equivalent in Equation (1) uses a pooled median power-law exponent $b \approx 0.47$. Different binder classes Portland-cement-only systems, alkali-activated and geopolymer systems, and high-volume SCM blends are not expected to share an exponent, and the IQR observed in the corpus (0.25–0.58) is wide precisely to reflect this chemical diversity. Forcing a uniform $b = 0.47$ introduces notable distortion when estimating 28-day strengths for non-OPC systems (where SCM hydration continues strongly into later ages, or where geopolymer polycondensation deviates entirely from C–S–H kinetics). We retain a single b because binder-class-stratified estimation is not yet supported by the per-stratum cluster sizes (Section 2.4); the LOSO prediction interval already absorbs the resulting normalization error, and binder-class-specific exponents should become identifiable as the literature accumulates. Until then, ages outside the 25–31-day window for alternative-binder mixtures should be referred to 28 d with caution.

(ii) *Effective W/B for non-Portland systems.* The W/B ratio is computed under the OPC convention with corrections for activator-bearing solids in alkali-activated systems and for high-absorption aggregates in LECA and recycled-powder mixtures (Section 2.2). Records that cannot be placed unambiguously into either convention are flagged and excluded from the regressions. When primary studies do not state which convention they use, we read the highest-information value (typically the as-mixed water against total binder mass) and inherit any ambiguity. Future round-robin programmes should report both the as-mixed and the effective W/B explicitly.

(iii) *Random-slopes identifiability.* Equation (2) uses a random intercept only. Random slopes for W/B or H/D would, in principle, separate the laboratory shift from a laboratory-specific sensitivity, but the slope variance components do not converge under restricted maximum likelihood (REML) with the present cluster sizes. Forcing a parallel-slope assumption unavoidably introduces some materials-science bias, as different extrusion systems and chemical admixtures react uniquely to W/B variations. We treat this as an open methodological question; overcoming this REML convergence failure will likely require exploring Bayesian hierarchical models with weakly informative priors on a subset of highly-reported laboratories, or waiting until per-study mixture coverage naturally expands.

5. Conclusions

By treating laboratory identity as a first-class statistical object rather than as nuisance noise, we have measured the share of variance in 3DCP compressive strength that

originates between laboratories rather than between mixtures within them. The estimate, $ICC = 0.77 [0.59, 0.86]$ over 345 records in 40 studies, makes three quantitative claims. First, the corresponding design effect ($DEFF \approx 5.24$ at the median cluster size) reduces the effective independent sample size of the corpus to roughly one fifth of its nominal record count, and any pooled inference that ignores this clustering will systematically understate uncertainty. Second, a 7-to-28-day calibration evaluated under leave-one-study-out cross-validation produces $RMSE = 11.38$ MPa and 95% prediction-interval coverage of 96.9%; the band is wide (≈ 85 MPa at the median) but honest, and it is suitable for screening rather than for code-compatible verification. This wide interval suggests caution when applying uncalibrated, literature-pooled models to direct structural design. Without laboratory-specific calibration or standardized process control, generic predictions are associated with notable uncertainty, indicating that site-specific physical testing remains necessary to verify structural reliability. Third, a retrospective audit shows that 63% of the included studies already satisfy Tier 1 of the proposed MSRS, 23% satisfy Tier 2, and none satisfy Tier 3, which identifies process-level documentation (controller traces and early-age yield-stress time series) as the binding constraint on causally-informed predictive modeling.

The practical message is, therefore, that cross-laboratory predictive models for 3DCP do not need to be more flexible; they need to be fed richer process metadata. Forward validation of the MSRS through a coordinated round-robin experimental programme is the natural next step. The mixed-effects architecture used here can be extended directly with additional fixed covariates as round-robin reporting accumulates, narrowing the gap between screening-level estimates and the site-specific reliability targets that structural-design codes require.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/buildings16142737/s1>, Section S1: Calibration and prediction interval diagnostics; Section S2: Parameter-space coverage and observed envelope; Section S3: Study inventory (PRISMA-included); Section S4: Reporting coverage and co-reporting distributions; Section S5: Evidence-backed MSRS (full table); Section S6: PRISMA queries and flow; and Section S7: ICC and uncertainty.

Author Contributions: Conceptualization, S.Ö. and S.A.; methodology, S.Ö.; software, S.Ö.; formal analysis, S.Ö.; investigation, S.Ö. and S.A.; data curation, S.Ö.; writing, original draft preparation, S.Ö.; writing, review and editing, S.A.; visualization, S.Ö.; supervision, S.A.; and project administration, S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Scientific Research Projects Coordination Unit of Istanbul Technical University. Project Number: ILAP2-2024-46209. The Article Processing Charge (APC) was also funded by the Scientific Research Projects Coordination Unit of Istanbul Technical University under Project Number: ILAP2-2024-46209.

Data Availability Statement: The original contributions presented in this study are included in the article and the Supplementary Materials. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors thank the contributing primary investigators of the included studies for the openness of their published reporting, which made cross-laboratory analysis possible. During the preparation of this manuscript, the authors used a large language model (ChatGPT-4.5, OpenAI) for language and style editing on the authors' draft text. The authors reviewed and edited the output and take full responsibility for the content of the publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

3DCP	3D concrete printing
ACI	American Concrete Institute
C&DW	Construction and demolition waste
DEFF	Design effect
EN	European Norm
HC3	Heteroskedasticity-consistent variance estimator (type 3)
H/D	Layer-height to nozzle-diameter ratio
ICC	Intraclass correlation coefficient
IQR	Inter-quartile range
LECA	Lightweight expanded clay aggregate
LOSO	Leave-one-study-out (cross-validation)
MAE	Mean absolute error
MSRS	Minimum (Standardized) Reporting Set
OLS	Ordinary least squares
OPC	Ordinary Portland Cement
PI	Prediction interval
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RMSE	Root-mean-squared error
SCM	Supplementary cementitious material
W/B	Water-to-binder ratio

References

1. EN 1992-1-1:2004; Eurocode 2: Design of Concrete Structures—Part 1-1: General Rules and Rules for Buildings. European Committee for Standardization (CEN): Brussels, Belgium, 2004.
2. ACI 318-19; Building Code Requirements for Structural Concrete (ACI 318-19) and Commentary (ACI 318R-19). American Concrete Institute: Farmington Hills, MI, USA, 2019.
3. Kaya, E.; Ciza, B.; Yalçinkaya, Ç.; Felekoğlu, B.; Yazıcı, H.; Çopuroğlu, O. Effect of Hydroxypropyl Methylcellulose and Aggregate Volume on Fresh and Hardened Properties of 3D Printable Concrete. *Constr. Build. Mater.* **2024**, *456*, 139253. [\[CrossRef\]](#)
4. Benam, S.S.; Sandalci, I.; Kara, B.; Bebek, O.; Bundur, Z.B. Improving the 3D Printability of High-Volume Fly Ash Mixtures Through Addition of Mineral Admixtures. In *Proceedings of the Rilem Spring Convention and Conference 2024, RSCC 2024*; RILEM Bookseries; Ferrara, L., Muciaccia, G., DiSumma, D., Eds.; Springer: Cham, Switzerland, 2025; Volume 56, pp. 265–273. [\[CrossRef\]](#)
5. Lee, J.H.; Kim, J.H. Matric Suction of Fine Sand and Its Effect on the Shape Stability of 3D Printed Cement Mortar. *Constr. Build. Mater.* **2022**, *341*, 127618. [\[CrossRef\]](#)
6. Cavalcante, T.C.; Filho, R.D.T.; Reales, O.A.M. Rheological and Environmental Implications of Recycled Concrete Powder as Filler in Concrete 3D Printing. *Buildings* **2025**, *15*, 1280. [\[CrossRef\]](#)
7. Araújo, R.A.; Martinelli, A.E.; Cabral, K.C.; Dantas, A.F.; Silva, I.F.; Xavier, A.A.; Santos, A.L. Thermal Performance of Cement-Leca Composites for 3D Printing. *Constr. Build. Mater.* **2022**, *349*, 128771. [\[CrossRef\]](#)
8. Kaya, E.; Ciza, B.; Yalçinkaya, Ç.; Felekoğlu, B.; Yazıcı, H.; Çopuroğlu, O. A Comparative Study on the Effectiveness of Fly Ash and Blast Furnace Slag as Partial Cement Substitution in 3D Printable Concrete. *J. Build. Eng.* **2025**, *108*, 112841. [\[CrossRef\]](#)
9. Uddin, M.N.; Ye, J.H.; Deng, B.Y.; Li, L.Z.; Yu, K.Q. Interpretable Machine Learning for Predicting the Strength of 3D Printed Fiber-Reinforced Concrete (3DP-FRC). *J. Build. Eng.* **2023**, *72*, 106648. [\[CrossRef\]](#)
10. Lai, X.; Gong, C.; He, E.; Li, Y.; Zhou, Y.; Zhang, F. Deep Learning for Predicting the Strength of 3D-Printable Engineered Cementitious Composites. In *Proceedings of the IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, Changchun, China, 27–29 February 2024; pp. 52–56. [\[CrossRef\]](#)
11. Malik, U.J.; Riaz, R.D.; Rehman, S.U.; Usman, M.; Riaz, R.E.; Hamza, R. Advancing Mix Design Prediction in 3D Printed Concrete: Predicting Anisotropic Compressive Strength and Slump Flow. *Case Stud. Constr. Mater.* **2024**, *21*, e03510. [\[CrossRef\]](#)
12. Chippagiri, R.; Lanjewar, B.; Kamath, M.; Ralegaonkar, R. Prediction of Construction Materials for Conventional and Alkali-Activated 3D Concrete Printing. In *Green Infrastructure and Construction for Sustainable Future*; Sustainable Civil Infrastructures; Springer: Cham, Switzerland, 2025; pp. 87–93. [\[CrossRef\]](#)
13. Mousavi, M.; Bengar, H.A.; Mousavi, F.; Mahdavinia, P.; Bengar, M.A. Interlayer Bond Strength Prediction of 3D Printable Concrete Using Artificial Neural Network: Experimental and Modeling Study. *Structures* **2025**, *71*, 108147. [\[CrossRef\]](#)

14. Rehman, S.U.; Riaz, R.D.; Usman, M.; Kim, I.H. Augmented Data-Driven Approach Towards 3D Printed Concrete Mix Prediction. *Appl. Sci.* **2024**, *14*, 7231. [\[CrossRef\]](#)
15. Zafar, M.S.; Javadnejad, F.; Hojati, M. Optimizing Rheological Properties of 3D Printed Cementitious Materials via Ensemble Machine Learning. *Addit. Manuf.* **2025**, *109*, 104889. [\[CrossRef\]](#)
16. Iqbal, M.; Nazar, S.; Yang, J.; Mahmoud, H.A. Integrated Machine Learning and Response Surface Methodology for Comprehensive Rheological Characterization of Low-Carbon Binders. *Case Stud. Constr. Mater.* **2025**, *22*, e04475. [\[CrossRef\]](#)
17. Mechtcherine, V.; Muthukrishnan, S.; Robens-Radermacher, A.; Wolfs, R.; Versteeg, J.; Menna, C.; Ozturk, O.; Ozyurt, N.; Roupec, J.; Richter, C.; et al. Mechanical Properties of 3D Printed Concrete: A RILEM 304-ADC Interlaboratory Study—Compressive Strength and Modulus of Elasticity. *Mater. Struct.* **2025**, *58*, 181. [\[CrossRef\]](#)
18. Ahi, O.; Ertuğ, Ö.; Bundur, Z.B.; Bebek, Ö. Automated Flow Rate Control of Extrusion for 3D Concrete Printing Incorporating Rheological Parameters. *Autom. Constr.* **2024**, *160*, 105319. [\[CrossRef\]](#)
19. Rehman, A.U.; Kim, I.G.; Kim, J.H. Towards Full Automation in 3D Concrete Printing Construction: Development of an Automated and Inline Sensor-Printer Integrated Instrument for in-Situ Assessment of Structural Build-up and Quality of Concrete. *Dev. Built Environ.* **2024**, *17*, 100344. [\[CrossRef\]](#)
20. Aslani, F.; Dale, R.; Hamidi, F.; Valizadeh, A. Mechanical and Shrinkage Performance of 3D-printed Rubberised Engineered Cementitious Composites. *Constr. Build. Mater.* **2022**, *339*, 127665. [\[CrossRef\]](#)
21. Zeng, J.J.; Hu, X.; Sun, H.Q.; Liu, Y.; Chen, W.J.; Zhuge, Y. Triaxial Compressive Behavior of 3D Printed PE Fiber-Reinforced Ultra-High Performance Concrete. *Cem. Concr. Compos.* **2025**, *155*, 105816. [\[CrossRef\]](#)
22. Ahadi, B.; Lopez, M.V. Zigzag Reinforcement Method for 3D Concrete Printing. *Buildings* **2025**, *15*, 1721. [\[CrossRef\]](#)
23. Rajeev, P.; Ramesh, A.; Navaratnam, S.; Sanjayan, J. Using Fibre Recovered from Face Mask Waste to Improve Printability in 3D Concrete Printing. *Cem. Concr. Compos.* **2023**, *139*, 105047. [\[CrossRef\]](#)
24. Araújo, R.A.; Martinelli, A.E.; Cabral, K.C.; Nunes, U.S.; Silva, I.F.D.; Xavier, A.A.C. Effect of Lightweight Expanded Clay Aggregate (LECA) on the Printability of Cementitious Compositions for 3D Printing. *Constr. Build. Mater.* **2025**, *470*, 140581. [\[CrossRef\]](#)
25. Arunothayan, A.R.; Nematollahi, B.; Khayat, K.H.; Ramesh, A.; Sanjayan, J.G. Rheological Characterization of Ultra-High Performance Concrete for 3D Printing. *Cem. Concr. Compos.* **2023**, *136*, 104854. [\[CrossRef\]](#)
26. Dahlenburg, M.; Hechtl, C.; Matthaus, C.; Fottner, J. 3D-Concrete Printing—Graded Concrete Extrusion. In Proceedings of the International Conference on Computation, Automation and Knowledge Management (ICCAKM), Dubai, United Arab Emirates, 15–17 November 2022; pp. 1–5. [\[CrossRef\]](#)
27. Dey, D.; Srinivas, D.; Panda, B.; Sitharam, T. Processing of Cementitious Materials for 3D Concrete Printing. In *Industry 4.0 and Advanced Manufacturing*; Lecture Notes in Mechanical Engineering; Springer: Singapore, 2023; pp. 283–291. [\[CrossRef\]](#)
28. Du, S.; Teng, F.; Zhuang, Z.; Zhang, D.; Li, M.; Li, H.; Weng, Y. A BIM-enabled Robot Control System for Automated Integration Between Rebar Reinforcement and 3D Concrete Printing. *Virtual Phys. Prototyp.* **2024**, *19*, e2332423. [\[CrossRef\]](#)
29. Gao, H.; Jin, L.; Chen, Y.; Chen, Q.; Liu, X.; Yu, Q. Rheological Behavior of 3D Printed Concrete: Influential Factors and Printability Prediction Scheme. *J. Build. Eng.* **2024**, *91*, 109626. [\[CrossRef\]](#)
30. Lim, S.G.; Tan, M.J. A Rheological Model for Concrete Additive Manufacturing. *Constr. Build. Mater.* **2024**, *451*, 138771. [\[CrossRef\]](#)
31. Maury-Ramírez, A.; Rinke, M.; Blom, J. Low-Carbon Embodied, Self-Cleaning, and Air-Purifying Building Envelope Components Using TiO₂ Photocatalysis, 3D Printing, and Recycling. *Coatings* **2024**, *14*, 1228. [\[CrossRef\]](#)
32. Pham, L.; Nguyen, T.; Trinh, T.; Nguyen, A.; Do, Q.; Bui, B.; Xiao, J. Development of 3D Printers for Concrete Structures: Mix Proportion Design Approach and Laboratory Testing. *Smart Sustain. Built Environ.* **2023**, *12*, 1056–1073. [\[CrossRef\]](#)
33. Sergis, V.; Ouellet-Plamondon, C.M. Fractional Factorial Design to Study Admixtures Used for 3D Concrete Printing Applications. *Mater. Lett.* **2022**, *324*, 132697. [\[CrossRef\]](#)
34. Sun, Z.; Zhao, Y.; Hou, D.; Li, Z.; Chen, B. Rheology Control of Cement Paste by In-Situ Polymerization for 3D Printing Applications. *Cem. Concr. Res.* **2025**, *187*, 107731. [\[CrossRef\]](#)
35. Xu, N.; Qian, Y. Effects of Fiber Volume Fraction, Fiber Length, Water-Binder Ratio, and Nanoclay Addition on the 3D Printability of Strain-Hardening Cementitious Composites (SHCC). *Cem. Concr. Compos.* **2023**, *139*, 105066. [\[CrossRef\]](#)
36. Mohamed, O.; Mishra, A.; Isam, F. An Overview of 3D Printed Concrete for Building Structures: Material Properties, Sustainability, Future Opportunities, and Challenges. *Structures* **2025**, *78*, 109284. [\[CrossRef\]](#)
37. Stout, I.; Godfrey, G.; Dayley, J.; Rodriguez, D.; Guthrie, W.S.; Sorensen, T.J. Concrete Mixture Properties and Designs for Additive Manufacturing: A Review of 3D Concrete Printing. In Proceedings of the 2024 Intermountain Engineering, Technology and Computing (IETC), Logan, UT, USA, 13–14 May 2024; pp. 285–290. [\[CrossRef\]](#)
38. Boddepalli, U.; Panda, B.; Ranjani Gandhi, I. Rheology and Printability of Portland Cement Based Materials: A Review. *J. Sustain. Cem.-Based Mater.* **2023**, *12*, 789–807. [\[CrossRef\]](#)
39. Warsi, S.B.F.; Panda, B.; Biswas, P. Exploring Fibre Addition Methods and Mechanical Properties of Fibre-Reinforced 3D Printed Concrete: A Review. *Dev. Built Environ.* **2023**, *16*, 100295. [\[CrossRef\]](#)

40. Zhong, H.; Zhang, M. Engineered Geopolymer Composites: A State-of-the-Art Review. *Cem. Concr. Compos.* **2023**, *135*, 104850. [[CrossRef](#)]
41. Andrew Ting, G.H.; Noel Quah, T.K.; Lim, J.H.; Daniel Tay, Y.W.; Tan, M.J. Extrudable Region Parametrical Study of 3D Printable Concrete Using Recycled Glass Concrete. *J. Build. Eng.* **2022**, *50*, 104091. [[CrossRef](#)]
42. Boddepalli, U.; Panda, B.; Gandhi, I.S.R. Impact of Water to Solids Ratio on Stability, Printability, and Mechanical Characteristics of 3D Printable Foam Concrete. In *Proceedings of the Construction 3D Printing, 4-Ic3dcp Conference 2023*; Springer Tracts in Additive Manufacturing; Tan, M.J., Li, M., Tay, Y.W.D., Wong, T.N., Bartolo, P., Eds.; Springer: Cham, Switzerland, 2024; pp. 165–173. [[CrossRef](#)]
43. Srinivas, D.; Panda, B.; Suraneni, P.; Sitharam, T.G. Mix Design Optimization of 3D-Printed Cementitious Composites for Marine Applications: Impact of Binder Composition, Accelerated Carbonation, and PVA Fibers on Strength and Durability. *Constr. Build. Mater.* **2025**, *489*, 142389. [[CrossRef](#)]
44. van Overmeir, A.L.; Figueiredo, S.C.; Savija, B.; Bos, F.P.; Schlangen, E. Design and Analyses of Printable Strain Hardening Cementitious Composites with Optimized Particle-Size Distribution. *Constr. Build. Mater.* **2022**, *324*, 126411. [[CrossRef](#)]
45. Valente, M.; Sambucci, M.; Chougan, M.; Ghaffar, S. Composite Alkali-Activated Materials with Waste Tire Rubber Designed for Additive Manufacturing: An Eco-Sustainable and Energy Saving Approach. *J. Mater. Res. Technol.* **2023**, *24*, 3098–3117. [[CrossRef](#)]
46. Dwivedi, A.; Pal, A.; Patel, S.S.; Chourasia, A.; Jain, A. Evaluation of Model 3D Printer and Design Mix for 3D Concrete Printing. In *Advances in Construction Materials and Sustainable Environment*; Lecture Notes in Civil Engineering; Springer: Singapore, 2022; Volume 196, pp. 837–847. [[CrossRef](#)]
47. Snijders, T.A.B.; Bosker, R.J. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, 1st ed.; SAGE Publications: London, UK, 1999.
48. Diniz, H.A.A.; Martinelli, A.E.; Cabral, K.C.; Ferreira, R.L.d.S.; da Silva, I.F.D. Synergistic Effects of the Use of Metakaolin, Sand and Water on the Properties of Cementitious Composites for 3D Printing. *Constr. Build. Mater.* **2023**, *366*, 130277. [[CrossRef](#)]
49. Thajeel, M.M.; Kopecskó, K.; Balázs, G.L. Enhancing Printability of 3D Printed Concrete by Using Metakaolin and Silica Fume. *Struct. Concr.* **2026**, *27*, 1068–1084. [[CrossRef](#)]
50. Yu, S.; Sanjayan, J.; Du, H. Effects of Cement Mortar Characteristics on Aggregate-Bed 3D Concrete Printing. *Addit. Manuf.* **2022**, *58*, 103024. [[CrossRef](#)]
51. Zafar, M.; Bakshi, A.; Hojati, M. Toward 3D-Printable Engineered Cementitious Composites: Mix Design Proportioning, Flowability, and Mechanical Performance. In *Proceedings of Tran-SET 2022*, San Antonio, TX, USA, 31 August–2 September 2022; pp. 102–114. [[CrossRef](#)]
52. Ambily, P.S.; Rajendran, N.; Kaliyavaradhan, S.K. Mix Design, Optimisation and Performance Evaluation of Three-Dimensional Printable Concrete. *Proc. Inst. Civ. Eng. - Constr. Mater.* **2024**, *177*, 302–320. [[CrossRef](#)]
53. Chen, Y.; El Cheikh, K.; Rahier, H. Methodology for the Design and Optimization of Potassium Silicate-Activated Slag Used as the Binder of 3D Printable Materials. *Constr. Build. Mater.* **2025**, *490*, 142536. [[CrossRef](#)]
54. Jaji, M.B.; van Zijl, G.P.A.G.; Babafemi, A.J. Slag-Modified Metakaolin-Based Geopolymer for 3D Concrete Printing Application: Evaluating Fresh and Hardened Properties. *Clean. Eng. Technol.* **2023**, *15*, 100665. [[CrossRef](#)]
55. Shoaie, P.; Gallantree-Smith, H.; Martínez Pacheco, V.; Pamies, R.; Kjøniksen, A.L.; Pilehvar, S. Comparative Analysis of 3D Printing of Portland Cement Mortars with Hydroxypropyl Methylcellulose and Microfibrillated Cellulose as Viscosity Modifying Agents. *Mater. Des.* **2024**, *244*, 113124. [[CrossRef](#)]
56. Wang, Y.; Khan, M.; Uy, B.; Katwal, U.; Tao, Z.; Thai, H.T.; Ngo, T. Long-term performance of steel-concrete composite wall panels under axial compression. *Structures* **2024**, *64*, 106606. [[CrossRef](#)]
57. Wang, Y.; Khan, M.; Uy, B.; Lim, L.; Thai, H.T.; Ngo, T. Behaviour, design and performance of steel-concrete composite walls in fire. *J. Constr. Steel Res.* **2025**, *227*, 109408. [[CrossRef](#)]
58. Parmigiani, S.; Falliano, D.; Moro, S.; Ferro, G.A.; Restuccia, L. Preliminary Study on Multi-Functional Building Components Utilizing Variable Density Foamed Concrete via 3D Printing. In *Proceedings of the Fourth Rilem International Conference on Concrete and Digital Fabrication, DC 2024*; RILEM Bookseries; Lowke, D., Freund, N., Bohler, D., Herding, F., Eds.; Springer: Cham, Switzerland, 2024; Volume 53, pp. 65–73. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.