

Article

A Surrogate-Based Digital Durability Atlas for Textile-Reinforced Mortar Under Acidic, Alkaline, and Saline Aging

Nima Azimi ^{1,*}, Omid Hassanshahi ¹, Mohammad Bakhshi ², Mohammed H. Alzard ¹ and Diāna Bajāre ¹

¹ Institute of Sustainable Building Materials and Engineering Systems, Faculty of Civil and Mechanical Engineering, Riga Technical University, Kipsalas 6A, LV-1048 Riga, Latvia; omid.hassanshahi@rtu.lv (O.H.); mohammed-h-h.alzard@rtu.lv (M.H.A.); diana.bajare@rtu.lv (D.B.)

² Institute for Sustainability and Innovation in Structural Engineering (ISISE), Advanced Production and Intelligent Systems Associated Laboratory (ARISE), Department of Civil Engineering, University of Minho, 4800-058 Guimarães, Portugal; mohammad.bakhshi.ngd@gmail.com

* Correspondence: nima32@gmail.com or nima.azimi-resketi@rtu.lv

Abstract

Textile-reinforced mortar (TRM) systems are used to strengthen masonry, but their durability across aggressive environments is rarely evaluated within one consistent framework. This study integrates coordinated acidic, alkaline, and saline aging experiments (0–5000 h) into an audited database containing 577 specimen records, 96 groups, and six mechanical families. Parallel full and master-retained datasets were analyzed to quantify sensitivity to specimen-retention decisions. Of 402 genuine reference-factor comparisons, 44.3% had bootstrap 95% confidence intervals including unity, and 25.9% of 201 paired comparisons showed high sensitivity to dataset version. Baseline-retention trajectories were represented using surrogate models selected by leave-one-time-level-out validation. Of 60 trajectories, 50 supported modeling: 27 selected linear models, 20 quadratic models, and 3 conditionally validated Gaussian processes; 10 pull-out-basalt trajectories remained unresolved because their baseline reference group could not be verified. Variability was highest for pull-out basalt. Although the ranking of acidic, alkaline, and saline conditions remained consistent across the three reference definitions within each mechanical family, it differed between families. The resulting durability atlas provides an uncertainty-aware comparison of TRM responses within the observed domain without implying mechanistic degradation laws, interpolation between environments, or prediction beyond 5000 h.

Keywords: textile-reinforced mortar; glass textile; basalt textile; durability; bond behavior; environmental aging; uncertainty quantification; surrogate model; data provenance; cross-environment comparison



Received: 26 August 2026

Revised: 14 September 2026

Accepted: 21 September 2026

Published: 1 October 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

Textile-reinforced mortar (TRM), also termed fabric-reinforced cementitious matrix (FRCM), is established for strengthening masonry and concrete because an inorganic matrix offers compatibility, fire resistance, and vapor permeability while technical textiles provide tensile capacity [1–10]. System response is governed by textile architecture, matrix properties, and stress transfer at textile–mortar and composite–substrate interfaces [11–19]. Consequently, durability must be assessed at multiple scales rather than inferred from mortar strength alone.

Environmental durability remains a central barrier to reliable long-term design. Reviews identify matrix alkalinity, fiber chemistry, exposure medium, moisture history, and

test configuration as interacting variables [20–26]. Experimental studies report the alkali-related degradation of glass and basalt reinforcement [27–31], salt crystallization and moisture effects on FRCM–masonry joints [10,32,33], and exposure-dependent changes in composite tensile and bond behavior [34–36]. Direct comparison is difficult because materials, conditioning protocols, and response definitions differ across studies.

The present work integrates three coordinated acidic, alkaline, and saline programs [37–39]. These source studies share materials, reference conditions, exposure times, and specimen-group identifiers while retaining medium-specific characterization. Related publications by the authors establish the tensile-test configuration and layer effects [11,12,40], report single-lap bond behavior under alkaline and acidic exposure [37,38,41], and connect acidic mortar degradation to microstructural change through chemo-mechanical analysis [42,43]. The current contribution is therefore not a new experimental campaign; it is an audited cross-environment synthesis of the underlying specimen records.

Two methodological needs motivate the analysis. First, aggressive-medium effects should be evaluated against both the initial state and time-matched dry or water-conditioned references. Second, uncertainty and data curation must remain visible: formal outlier procedures exist [44], and data provenance is recognized as important for reliable durability modeling [45], yet 55 specimen omissions in the source workbooks have no recorded justification. Bootstrap intervals [46] and parallel full versus master-retained analyses quantify how sampling variability and retention decisions influence interpretation. For sparse trajectories, simple regression and Gaussian-process candidates [47] are selected by validation rather than flexibility alone, with the known limitations of aggregated leave-one-out assessment acknowledged [48].

Accordingly, this study constructs a categorical, uncertainty-aware durability atlas for six matrix- and textile-bond response families. It integrates 577 specimen records, computes three reference-factor formulations under two dataset versions, quantifies uncertainty and curation sensitivity, and validates low-order time trajectories within 0–5000 h. A digital durability atlas is defined here as an uncertainty-aware, categorical synthesis of experimentally observed durability trajectories indexed by mechanical response family, textile type, exposure environment and exposure time, in which every trajectory is reported together with its validation error and with an explicit statement of how far it depends on specimen-retention decisions. The term is introduced descriptively in this study and is not an established TRM/FRCM designation. The atlas is a retrospective synthesis of discrete tested conditions and is not a continuous chemical-severity surface, a mechanistic degradation law, a continuously updated model, or an operational digital twin [49].

Three questions follow from this framing and organize the analysis. First, how far do reference-factor estimates and their uncertainty intervals depend on specimen-retention decisions that carry no recorded justification, and are those decisions distributed evenly across the dataset or concentrated in particular parts of it? Second, what functional form can be supported for a durability trajectory that is observed at only four exposure times, and how stable is that choice against the tolerances used to make it? Third, what is learned by placing acidic, alkaline and saline ageing in one harmonized framework that is not visible in the three source studies taken separately? The contribution is therefore integrative and methodological rather than experimental or mechanistic: the materials, specimens and exposure protocols are those of the source publications, and what is new is the audited specimen-level integration, the explicit quantification of data-curation sensitivity, and the cross-environment comparison built on them.

2. Materials and Methods

Materials, specimen geometries, mortar and textile compositions, and exposure protocols are reported in full in the three source publications and are summarized here at the level needed to interpret the present integrated analysis. A parameter-by-parameter summary of the three campaigns, separating what was identical across them from what differed, is given in Table S1 of the Supplementary Materials.

All three source studies used the same commercially available hydraulic lime-based mortar, Planitop HDM Restauro (MAPEI S.p.A., Milan, Italy). The mortar was reinforced with either Mapegrid G 220 alkali-resistant glass-fiber textile or Mapegrid B 250 basalt-fiber textile (both manufactured by MAPEI S.p.A., Milan, Italy) to produce the TRM composite specimens. Two reference aging conditions were investigated alongside each aggressive medium: a dry/Witness condition and distilled-water immersion. The aggressive conditions comprised an acidic solution, an alkaline solution, and a 3.5% NaCl saline solution. Specimens were tested at nominal exposure durations of 1000, 3000, and 5000 h, in addition to a shared 90-day (pre-exposure) baseline condition. Bond behavior between textile and mortar was assessed using pull-out and single-lap shear configurations, following the source publications' own experimental procedures. Where genuinely comparable, single-lap TRM-to-substrate bond testing can be contextualized against the Recommendation of RILEM Technical Committee 250-CSM for TRM-to-substrate bond characterization [50]; that recommendation addresses TRM-to-substrate bond testing specifically and is not cited here as the specifying authority for the internal textile-to-mortar pull-out configuration used in the source publications, which follows those publications' own verified procedures. For the purposes of this integration, the three campaigns are directly comparable in matrix, textiles, brick substrate, specimen coding, pre-exposure curing, nominal exposure durations, nominal replicate number and measured response (peak load, in kN), and differ only in the aggressive medium itself and in the medium-specific chemical and microstructural characterization that each source study reports. The dry/Witness and distilled-water reference groups are not three repeated experiments, but one set of measurements reused by all three publications, which is stated explicitly in the saline source study and is confirmed by the shared specimen-group identifiers in the archived records.

The three source publications report a broader experimental and characterization campaign than is analyzed quantitatively in this integration study. In addition to the six mechanical families analyzed here, the source publications include the elastic-modulus testing of the mortar, tensile testing of the bare (unembedded) dry textiles, and mineralogical/chemical characterization of the matrix and interface, discussed further in the Discussion and reported in full in the respective source publications. The six analyzed mechanical families, associated textile types, measured responses, exposure environments, and nominal exposure times are summarized in Table 1.

The present integrated quantitative analysis is restricted to the six core mechanical families for which specimen-level, group-resolved data were reconciled into the audited database: compressive, flexural, pull-out-glass, pull-out-basalt, single-lap-glass, and single-lap-basalt testing, evaluating bulk mortar mechanical properties and TRM bond behavior for both reinforcing textile types. Elastic-modulus testing, bare/dry-textile tensile testing, and mineralogical/chemical characterization datasets are outside the scope of the specimen-level retention-factor and surrogate-trajectory analysis presented in this study; they are described here only to characterize the source publications accurately, not as part of the integrated dataset. Full experimental details (specimen dimensions, curing conditions, solution concentrations, textile properties, test rig configurations, and loading rates) are reported in [37] (acidic), [38] (alkaline), and [39] (saline). The parameters that bear on

comparability between the three campaigns, and the values that are quoted identically in all three, are collected in Table S1.

Table 1. Mechanical families and exposure conditions included in the analysis.

Family	Textile (s)	Response Type	Environments	Nominal Exposure Times (h)
Compressive	n/a (mortar)	Bulk compressive strength	Witness, Water, Acidic, Alkaline, Saline	0 ¹ , 1000, 3000, 5000
Flexural	n/a (mortar)	Bulk flexural strength	Witness, Water, Acidic, Alkaline, Saline	0 ¹ , 1000, 3000, 5000
Pull-out Glass	Glass	Textile–mortar bond (pull-out)	Witness, Water, Acidic, Alkaline, Saline	0 ¹ , 1000, 3000, 5000
Pull-out Basalt	Basalt	Textile–mortar bond (pull-out)	Witness, Water, Acidic, Alkaline, Saline	0 ² , 1000, 3000, 5000
Single-lap Glass	Glass	TRM bond (single-lap shear)	Witness, Water, Acidic, Alkaline, Saline	0 ¹ , 1000, 3000, 5000
Single-lap Basalt	Basalt	TRM bond (single-lap shear)	Witness, Water, Acidic, Alkaline, Saline	0 ¹ , 1000, 3000, 5000

¹ The 0 h (90-day pre-exposure) baseline condition is shared across environments for the same test type and textile. ² The pull-out basalt 0 h baseline reference group is unresolved and excluded from baseline-retention calculations for this family only; time-matched ratios are unaffected.

2.1. Data Harmonization and Quality Assurance

The three studies' underlying specimen-level records were reconciled into a single audited specimen database (577 records across 96 experimental groups) through a multi-stage process: (i) a cell-level audit of the three master summary workbooks against their own internal formulas and stated statistics; (ii) a specimen-by-specimen reconciliation of the individually analyzed per-group files against the corresponding master-workbook entries; (iii) an evidence review classifying every specimen present in the analyzed files but absent from the master workbook ("omission") by the strength of documentary evidence available for its absence (raw-file annotations, failure-mode records, or explicit removal instructions), rather than assuming an omission was scientifically justified by default.

This review identified 59 genuine master-workbook omissions (distinct from a single specimen that was never measured in either source and is therefore not counted as an omission). Of these 59, three have a documented reason (a recorded test failure or equipment problem), one carries an explicit removal instruction in a raw-file name with no stated reason, and 55 have no reason or instruction that could be recovered from the archived project files available for the present audit. Formal statistical procedures exist for evaluating suspected outlying observations in replicate materials-testing data [44,51], but no record of such a procedure having been applied to any of these 55 omissions could be recovered. An unexplained omission is, on its own, neither evidence that the omitted specimen was erroneous nor evidence that it was valid; absent a documented basis for judging it either way, this study adopts a conservative provenance and sensitivity-analysis policy rather than presuming the omission correct. Consequently, every downstream quantity in this study is computed under two parallel dataset versions: a full dataset (every specimen value found in the analyzed files, including the 59 master-omitted specimens) and a master-retained subset (only the specimens also retained in the original master workbooks). The two versions are never merged, and no version is treated as inherently more correct than the other; the extent to which conclusions depend on this choice is itself reported as a primary result. Three points of wording should be made explicit here because the distinction matters for how these omissions are treated. The statement is about what remains traceable in the archived record, not about what the original investigators knew

when the analyses were performed: these omissions were made several years before the present retrospective audit, and the absence of a recoverable specimen-level note does not imply that no reason existed at the time. The omissions are also separated by the kind of evidence available rather than merged into a single category, distinguishing a specimen that was never measured, a specimen omitted with a documented test failure or equipment problem, a specimen carrying an explicit removal instruction but no stated reason, and a specimen for which no documentary trace of either kind survives. How the 55 records in the last category are distributed across mechanical families, textiles, environments and exposure times is reported in Section 3.1 and Table S2.

One additional group—the baseline (90-day, pre-exposure) Witness condition for the pull-out basalt configuration, internally labeled GD42—could not be resolved to a confident specimen-to-record linkage after this review, in the sense that no traceable specimen-level documentary record linking it to the analyzed files could be recovered from the archived material available for this audit, and it remains excluded from every calculation that requires it as a baseline reference with no substitute value used in its place.

2.2. Retention/Reference-Factor Definitions

The group mean $M(G)$ of an experimental group G was first computed as

$$M(G) = (1/n) \sum_{(i=1)}^{(n)} x_i \quad (1)$$

where n is the number of specimens in the group and x_i is the value of specimen i .

Three reference-factor formulations were then computed, each expressing a group mean as a ratio to a reference group mean of the same test type and textile. Baseline retention R_B is defined as

$$R_B(t) = M(G_t) / M(G_0) \quad (2)$$

where G_t is the group of specimens of a given test type, textile, and environment at exposure time t , and G_0 is the corresponding 90-day (pre-exposure) baseline group. This factor expresses total change from the initial condition and requires a resolved baseline group (unavailable for pull-out basalt).

The time-matched Witness ratio R_W is defined as

$$R_W(t) = M(G_{(t,env)}) / M(G_{(t,Witness)}) \quad (3)$$

comparing an aggressive-medium group to the dry/Witness group at the same nominal exposure duration t . This controls for specimen age and general ageing/testing variability common to both conditions at the same duration but does not isolate a specific chemical mechanism.

The time-matched Water ratio R_D is defined as

$$R_D(t) = M(G_{(t,env)}) / M(G_{(t,Water)}) \quad (4)$$

comparing an aggressive-medium group to the distilled-water-immersed group at the same nominal exposure duration t . This additionally controls for moisture-history effects common to both conditions but likewise does not isolate a unique chemical mechanism.

In all three definitions, a factor value greater than 1 indicates a higher measured group mean relative to its reference; it does not, by itself, indicate improved durability, since the underlying mean can be influenced by continued hydration/curing, moisture history, general ageing, or specimen-level testing variability, individually or in combination. The

mathematical definitions, reference conditions, and intended interpretation of the three retention and comparison factors are summarized in Table 2.

Table 2. Retention and reference-factor definitions.

Factor Type	Definition	Reference Condition	Controls for
Baseline retention	$R_B(t) = \frac{M(G_{t,env})}{M(G_0)}$	90-day (pre-exposure) baseline group	Nothing beyond the baseline itself; reflects total change from the initial condition
Time-matched Witness ratio	$R_W(t) = \frac{M(G_{t,env})}{M(G_{t,Witness})}$	Dry/Witness group at the same exposure time	Specimen age and the general ageing and testing variability common to both conditions; it does not control for immersion history
Time-matched Water ratio	$R_D(t) = \frac{M(G_{t,env})}{M(G_{t,Water})}$	Distilled-water-immersed group at the same exposure time	Specimen age and moisture-history effects common to both conditions

Neither time-matched ratio isolates a specific chemical mechanism; both reduce, but do not eliminate, age- and/or moisture-related confounding relative to baseline retention.

Ten of the 412 valid factors have an identical numerator and denominator group (a group's own baseline expressed as a ratio to itself). These equal exactly 1 by mathematical definition, not by empirical estimation, and are treated analytically rather than by independent bootstrap resampling; reporting them as independently estimated intervals would manufacture artificial uncertainty around a quantity that has none.

2.3. Bootstrap Uncertainty Analysis

For every valid factor comparing two genuinely distinct experimental groups, a 95% confidence interval was estimated by non-parametric bootstrap resampling at the specimen level following the percentile-bootstrap approach established by Efron and Tibshirani [46]. Numerator and denominator specimens were resampled independently, with replacement, each preserving its group's observed sample size, for 20,000 iterations per factor (fixed random seed, disclosed for reproducibility). The ratio of resampled means was computed per iteration; iterations in which the denominator resample mean equaled zero were recorded as invalid and excluded from the percentile and probability calculations (not replaced or imputed); a confidence interval was flagged unstable if more than 1% of iterations were invalid (0 factors met this threshold in this dataset). The 2.5th, 50th, and 97.5th percentiles of the valid resampled-ratio distribution were taken as the 95% confidence interval and bootstrap median; the proportion of valid resampled ratios above and below 1 were recorded as resampling probabilities—descriptions of the bootstrap distribution itself, not Bayesian posterior probabilities or frequentist significance statements. This study's typical per-group sample size ($n = 3\text{--}6$ specimens) is small relative to sample-size guidance sometimes cited for bootstrap coverage accuracy; this is disclosed as a limitation rather than resolved by the cited methodology.

For bootstrap replicate b , the factor was recalculated from independently resampled numerator and denominator specimens as

$$R^{*(b)} = \frac{\bar{x}_{\text{num}}^{*(b)}}{\bar{x}_{\text{den}}^{*(b)}}, \quad b = 1, \dots, B. \quad (5)$$

and the percentile interval was obtained directly from the empirical bootstrap distribution:

$$\text{CI}_{95\%}(R) = \left[Q_{0.025} \left(\left\{ R^{*(b)} \right\}_{b=1}^B \right), Q_{0.975} \left(\left\{ R^{*(b)} \right\}_{b=1}^B \right) \right]. \quad (6)$$

Group-mean uncertainty (independent of any ratio) was separately computed for every experimental group and dataset version as arithmetic mean, population and sample standard deviation, standard error of the mean, and a 95% confidence interval via the Student-t distribution (degrees of freedom = $n - 1$). Literal zero specimen values were retained and flagged in full-dataset calculations (as genuine observed, non-missing measurements); blank, missing, or unresolved records were excluded from calculation and counted separately.

Structural identity factors were excluded from independent bootstrap resampling and instead assigned, analytically: point estimate = 1, bootstrap median = 1, 95% confidence interval = [1, 1], interval width = 0, and no resampling probability. They retain the same reported bootstrap-iteration count and random seed as every other factor for schema consistency and traceability, but no independent resampling result underlies their reported values.

Because the percentile interval is the component of this procedure most exposed to the reviewer-noted small-sample concern, and because the typical group size here is $n = 3-6$, the percentile result was checked against two alternatives for every genuine comparison rather than being assumed adequate. The first is a bias-corrected and accelerated (BCa) bootstrap computed from the same 20,000 resamples, with the bias-correction constant obtained from the proportion of resampled ratios below the observed ratio and the acceleration constant from a jackknife over the pooled numerator and denominator specimens; BCa adjusts precisely for the median bias and skewness that make a percentile interval unreliable in small samples. The second is a resampling-free interval obtained by the delta method applied to the logarithm of the ratio, with Welch–Satterthwaite degrees of freedom, which removes the resampling assumption altogether at the cost of a normal-theory one. The percentile interval is retained as the primary reported uncertainty measure because it underlies every previously reported figure and because the three methods agree closely; the comparison is reported in Section 3.2 and in Table S9.

For every factor available under both dataset versions (206 comparisons: 5 structural identities + 201 genuine comparisons), the full-dataset and master-retained-subset estimates and confidence intervals were compared and classified into four ordered categories: critical (either version's bootstrap distribution is unstable, or the point-estimate direction relative to 1 flips between versions while both confidence intervals confidently exclude 1); high (the point-estimate direction relative to 1 changes between versions, or the conclusion of whether the confidence interval includes 1 changes, without both versions confidently conflicting); low (absolute difference ≤ 0.02 and relative difference $\leq 5\%$); and moderate (all remaining cases). Structural identity comparisons are reported separately and excluded from the substantive sensitivity rates. The 0.02 absolute and 5% relative values used to delimit the low-sensitivity category are pragmatic screening tolerances rather than physically derived constants; their influence on the resulting classification is examined in Section 3.3 and Table S5.

Before categorical classification, absolute and relative discrepancies between the full (F) and master-retained (M) estimates were calculated as

$$\Delta_{\text{abs}} = |R^{(F)} - R^{(M)}|, \quad \Delta_{\text{rel}} = \frac{|R^{(F)} - R^{(M)}|}{\max(|R^{(F)}|, \epsilon)}. \quad (7)$$

2.4. Surrogate-Model Construction and Validation

For each of the six mechanical test-type/textile families, five exposure environments, and two dataset versions (60 combinations), a baseline-retention trajectory was defined at up to four nominal exposure-time levels (0, 1000, 3000, 5000 h), using the shared 90-day baseline

as the common 0 h origin. Fifty of 60 trajectories have a complete four-point series and are eligible for surrogate-model fitting; the remaining ten (all pull-out Basalt) are blocked by the unresolved baseline group and carry no surrogate model. Exposure environments were treated strictly as categorical, mutually exclusive conditions; no continuous chemical-severity axis was constructed, and no interpolation was performed between environments. Four nominal exposure levels is a materially small basis for any trajectory model, and the consequences are stated here rather than left to the Discussion. A quadratic fit to four points retains one residual degree of freedom; a Gaussian process fit to four points has very little information with which to constrain its covariance hyperparameters; and any change in mechanical response that occurs between 0 and 1000 h, or between 1000 and 3000 h, is invisible to this analysis. The fitted forms are therefore treated throughout as descriptive, interpolative representations of the observed series within the tested domain. A selected functional form is a statement about which of the tested representations best reproduces a withheld observation, not evidence about the physical form of the underlying degradation process, and no forecast outside the observed domain is made from any of them.

Exposure time was normalized before fitting as

$$\tau = \frac{t}{5000}, \quad 0 \leq \tau \leq 1. \quad (8)$$

Four candidate model families were evaluated on each ready trajectory. The linear model is

$$\hat{R}(\tau) = a + b\tau. \quad (9)$$

The quadratic model is

$$\hat{R}(\tau) = a + b\tau + c\tau^2. \quad (10)$$

Shape-preserving piecewise cubic Hermite interpolation (PCHIP), an exact interpolant through the four calibration points, was offered only as a descriptive/visual aid, never as a validated predictive model; and a Gaussian-process (Kriging) regression used a covariance kernel of the form for two normalized exposure times τ and u ,

$$k(\tau, u) = \sigma_f^2 \exp \left[-\frac{(\tau - u)^2}{2\ell^2} \right] + \sigma_n^2 \delta_{\tau u}. \quad (11)$$

fit only where hyperparameter optimization converged without its length-scale (ℓ) or noise-level (σ_n) parameters reaching their imposed bounds [47]. No high-order polynomial, neural network, random forest, physics-informed neural network, or automated model-search procedures were used given the small number of calibration points (three to four) available per trajectory.

The Gaussian-process implementation is specified in full so that the classification of a fit as stable or disqualified can be reproduced. The covariance function is a constant kernel multiplied by a squared-exponential kernel with an additive white-noise term, with bounds of $[10^{-3}, 10^3]$ on the signal variance, $[10^{-2}, 10^2]$ on the length-scale (in normalized time) and $[10^{-5}, 10^5]$ on the noise level, all three initialized at unity. The response was standardized before fitting; hyperparameters were obtained by maximizing the log marginal likelihood using the L-BFGS-B optimizer implemented in SciPy version 1.18.1 through scikit-learn version 1.9.1 in Python version 3.15 from a single initialization. A deterministic optimization procedure was employed to ensure reproducible model fitting. Two diagnostics were then applied to every fitted curve. A fit was recorded as boundary-limited when the optimized length-scale terminated within one decade of its lower bound—that is, below 0.1 in normalized time—which is the regime in which the squared-exponential term carries no usable correlation across the observed exposure times, and the posterior

mean reverts to the prior mean with the residual variance absorbed into the noise term; a fit that was boundary-limited at four-point calibration was disqualified from selection. Oscillation was quantified rather than judged visually. The fitted curve was evaluated at 101 equally spaced points across the normalized domain: first, differences were taken, and sign changes were counted while ignoring differences smaller than 10^{-9} . A curve showing two or more sign changes contains at least two interior turning points, which three or four calibration points cannot support, and was flagged on that basis. Applying the identical criterion to the other candidate model families flagged none of the linear or quadratic fits and eight of the fifty PCHIP calibration fits, which confirms that the criterion discriminates between curvature that the data support and curvature that they do not, rather than flagging curvature indiscriminately. The fitted hyperparameters, their bound status and the oscillation count for every fit are reported in Table S7.

Model performance was assessed by leave-one-time-level-out cross-validation: each of the four time levels was withheld in turn, the candidate model refit on the remaining three, and the withheld point predicted. Withholding an interior level (1000 or 3000 h) is reported as interpolation validation; withholding a boundary level (0 or 5000 h) is reported as endpoint validation, and the two are always reported separately, never combined into one headline error metric, since prediction toward a domain boundary is a materially different task from prediction within it. No specimen-level random train/test splitting was used. This study's leave-one-time-level-out validation used only four time levels, of which only two (1000 and 3000 h) were interior holdouts, and was not checked against any independent external dataset; aggregated leave-one-out-style performance estimates of this kind can in general be affected by a distributional bias arising from dependence between a training fold's label distribution and its held-out label, and the resulting error figures reported here should be interpreted with this general statistical caution in mind, in addition to the small-sample limitation stated independently above.

For a validation subset V (interior or endpoint holdouts), the reported root-mean-square prediction error was

$$\text{RMSE}_{\text{LOTO}}(V) = \sqrt{\frac{1}{|V|} \sum_{k \in V} [R(t_k) - \hat{R}_{(-k)}(t_k)]^2} \quad (12)$$

A single model was recommended per ready trajectory using a fixed priority order: (1) the numerical validity and stability of the fit (a Gaussian process with boundary-hitting hyperparameters at calibration was disqualified); (2) best leave-one-time-level-out interpolation performance among valid, stable candidates; (3) minimal unsupported oscillation in the fitted curve; (4) the physical plausibility of predicted values; (5) simplicity/transparency, applied as a tiebreaker—a linear model was preferred over a better-performing candidate when its interpolation RMSE was within 10% of the best candidate's; and (6) compatibility with the established uncertainty quantification. The model with the lowest in-sample calibration error was never automatically selected, since a low-order model fit to only four points can reproduce those points closely without necessarily generalizing to a withheld point. The 10% figure in the fifth criterion is a parsimony tolerance fixed before the results were examined; it is a modelling convention expressing how much predictive performance is worth trading for a simpler functional form, not a physically derived threshold. Section 3.4 reports how the selected families change when the percentage varied between 0% and 20%.

The primary selection step can therefore be summarized as the constrained rule

$$m^* = \underset{m \in \mathcal{M}_{\text{stable}}}{\text{argmin}} \text{RMSE}_{\text{LOTO}}^{\text{interior}}(m). \quad (13)$$

where $\mathcal{M}_{\text{stable}}$ contains only numerically valid, physically plausible candidates without disqualifying hyperparameter-bound or oscillation flags. If the linear model's interior RMSE was within 10% of the minimum, linearity was selected by the predefined parsimony rule. Endpoint RMSE was retained as a separate diagnostic and did not overwrite the interpolation-based selection. The complete analytical workflow is summarized in Figure 1. The three coordinated source studies were first harmonized into an audited specimen-level database. Full and master-retained datasets were then maintained as parallel analytical versions throughout reference-factor calculation, bootstrap uncertainty analysis, and surrogate-model validation. The resulting validated trajectories form the categorical durability atlas presented in this study.

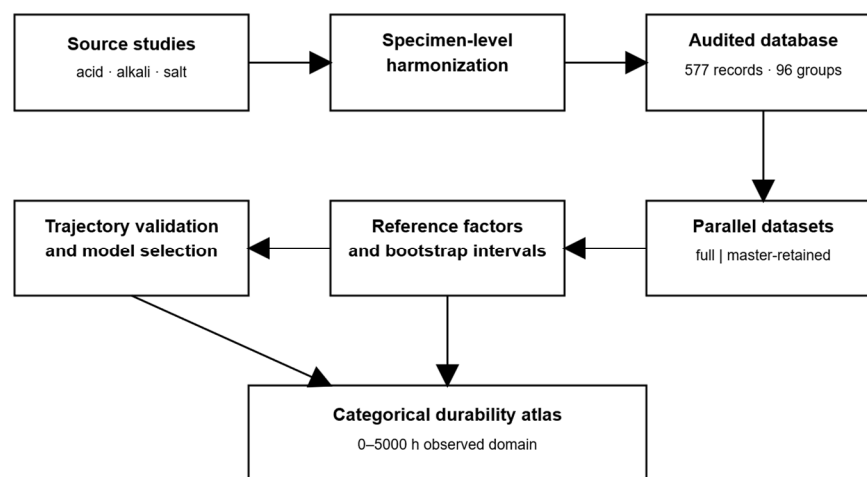


Figure 1. Workflow for the integrated durability analysis.

3. Results

3.1. Database Composition and Omission Accounting

The audited specimen database comprises 577 specimen-level records across 96 experimental groups spanning six mechanical test-type families (Table 1), two textile types (glass, basalt) where applicable, five exposure environments, and up to four nominal exposure durations (Figure 2). Of 60 records flagged as excluded from the master-workbook analysis, 59 are genuine master-workbook omissions and one was never measured in either source (Table 3). Of the 59 genuine omissions, three carry a documented reason, one carries an explicit removal instruction with no stated reason, and 55 have no recorded reason of any kind (Table 3, Figure 3). The resulting specimen-level database structure and the accounting of measured, omitted, and unresolved records are presented in Table 3. The 55 omissions with no recoverable documentation are not evenly spread through the dataset. Pull-out basalt alone accounts for 19 of them—34.5% of all such omissions and 20.9% of that family's full-dataset specimens—against 4 (4.2%) for pull-out glass and 4 (4.2%) for flexural; single-lap basalt accounts for 12 (12.5%), single-lap glass for 10 (10.4%) and compressive for 6 (6.3%). Aggregated by textile, basalt-reinforced specimens carry 31 of the 55 (16.6% of basalt specimens) against 14 for glass (7.3%) and 10 for the unreinforced mortar families (5.2%). A contingency test on omitted against retained specimens rejects independence for mechanical family ($\chi^2 = 21.95$, 5 df, Monte-Carlo $p = 0.001$) and for textile ($\chi^2 = 15.79$, 2 df, $p < 0.001$), and marginally for exposure environment ($\chi^2 = 9.80$, 4 df, $p = 0.043$, with alkaline at 15.7% and saline at 13.0% against 5.6% for distilled water), but not for exposure time ($\chi^2 = 2.34$, 3 df, $p = 0.51$; 9.4%, 11.7% and 8.9% at 1000, 3000 and 5000 h). The concentration is therefore in the specimen families that already show the largest specimen-to-specimen scatter, and not in any particular exposure duration. Counts, rates and the number of

affected groups for every stratification are given in Table S2. Because pull-out basalt is also the family carrying both the widest specimen-to-specimen variability and the unresolved baseline group, comparisons involving basalt bond behavior are the ones for which the full-versus-retained contrast should be read most carefully; Section 3.3 quantifies how often that contrast changes a stated conclusion.

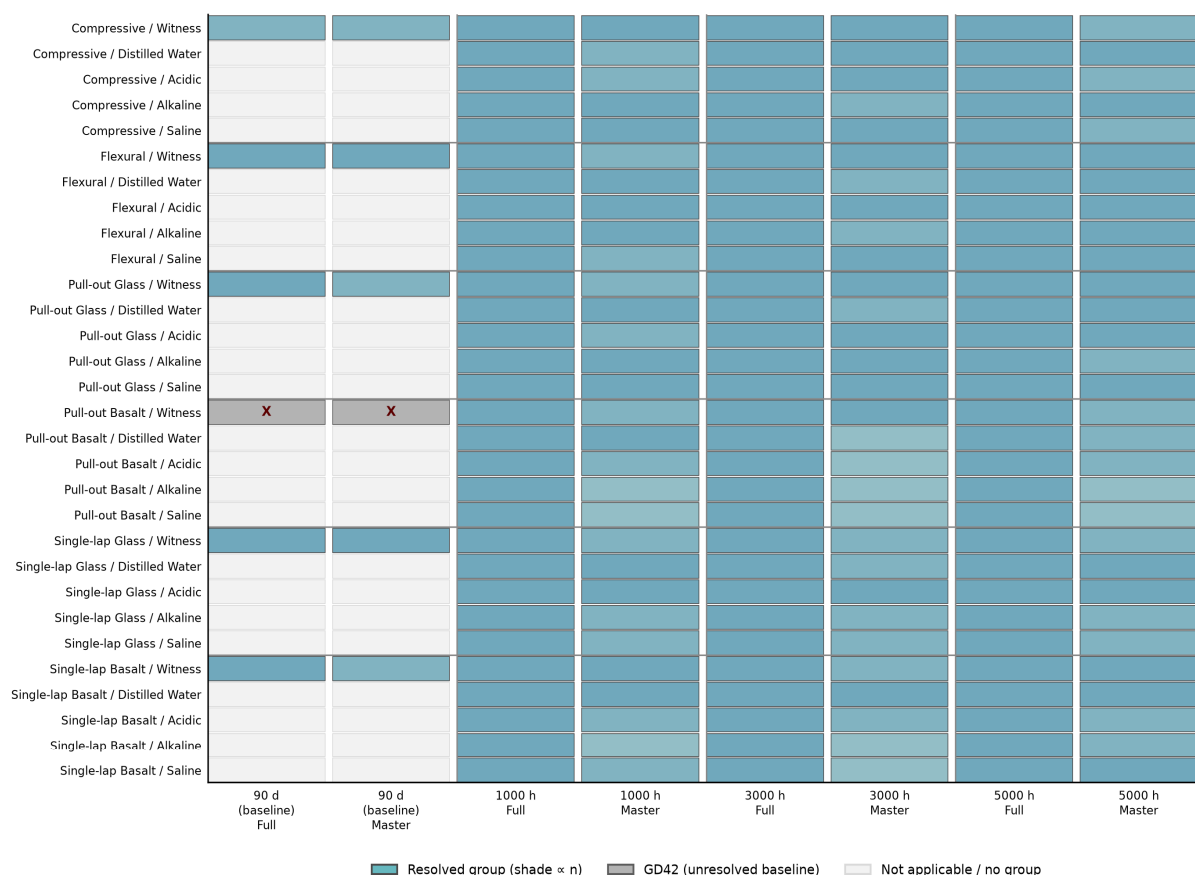


Figure 2. Dataset coverage by family, environment, time, and dataset version.

Table 3. Specimen and omission accounting.

Metric	Value
Total specimen-level records	577
Distinct experimental groups	96
Genuine master-workbook omissions	59
with a documented reason	3
with a removal instruction, no stated reason	1
unexplained (no reason or instruction found)	55
Unresolved reference group (pull-out basalt baseline)	1

The coverage of the two dataset versions across mechanical families, exposure environments, and nominal exposure times is shown in Figure 2. The unresolved pull-out basalt baseline group is identified separately and is not treated as a valid trajectory anchor.

Figure 3 summarizes the group-level coefficients of variation and the available documentary evidence for the 59 genuine specimen omissions. The basalt-textile bond-test families exhibit the widest variability, while most omissions lack a recorded justification.

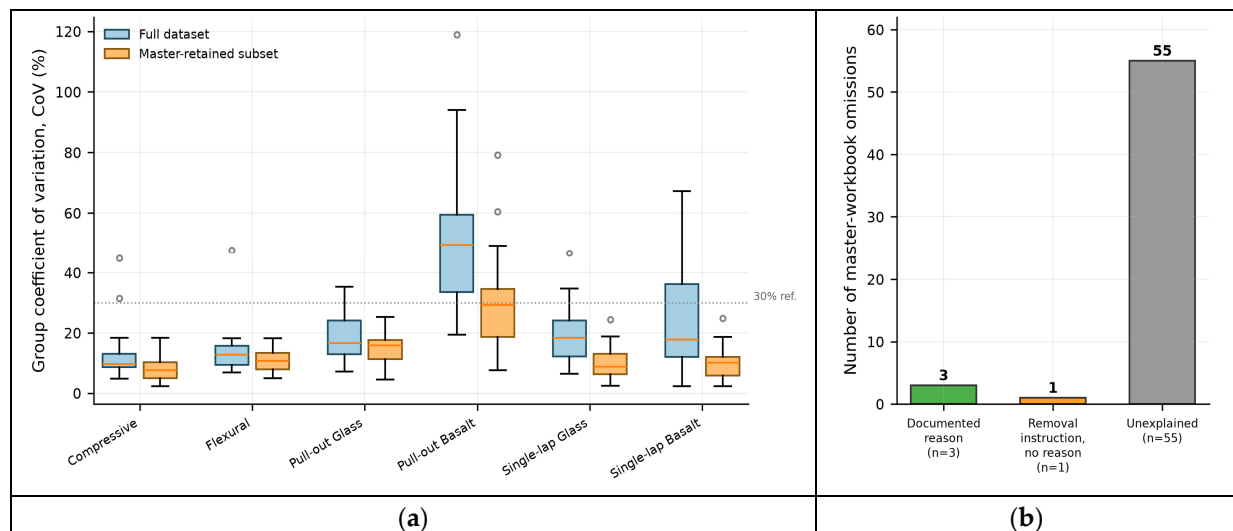


Figure 3. Group variability and specimen-omission evidence: (a) variability by mechanical family, with circles indicating values beyond 1.5 times the interquartile range; (b) evidence categories for the 59 omissions.

3.2. Uncertainty and Complementary Reference Views

Across 190 of the 192 group/dataset-version combinations with computable variability statistics (the remaining 2 are the unresolved pull-out basalt baseline rows), the coefficient of variation ranges from 2.2% to 118.8%, with the highest values concentrated in the two basalt-textile bond-test families (Figure 3a). Of 412 valid retention/comparison factors, 10 are structural identity factors (exactly $R = 1$ by definition) and 402 are genuine, independently bootstrapped comparisons. Among these 402, 44.3% (178/402) have a 95% bootstrap confidence interval that includes 1 (Table 4)—meaning that, given specimen-level variability at the observed sample sizes, the data do not clearly distinguish these comparisons from “no difference.” Representative baseline-retention, time-matched Witness-ratio, and time-matched Water-ratio trajectories, each with 95% bootstrap confidence intervals shown for both dataset versions, are presented for every family and environment in Figures 4–6. The extreme coefficients of variation quoted above occur in the two basalt bond families, but the medians locate the effect more precisely. Summarized by mechanical family, the median coefficient of variation is 9.1% for compressive (interquartile range 5.2–11.6%), 11.1% for flexural (7.9–14.8%), 16.5% for pull-out glass (11.9–18.8%), 12.6% for single-lap glass (8.1–18.8%), 12.0% for single-lap basalt (8.3–18.6%) and 34.5% for pull-out basalt (20.9–54.5%), each based on 30 to 32 group and dataset-version combinations. The elevated basalt variability is therefore concentrated in the pull-out configuration rather than being a property of basalt reinforcement across test scales: single-lap basalt and single-lap glass are indistinguishable in median terms. Pooled across configurations a difference between the two textiles remains (median 19.5% against 15.0%; Mann–Whitney $p = 0.002$), as does the difference between bond tests and matrix-dominated tests (20.0% for pull-out and 12.2% for single-lap against 9.9% for the matrix families; $p = 2.7 \times 10^{-7}$ and $p = 0.0006$ respectively). Full summaries are given in Table S6.

Replacing the percentile interval with a bias-corrected and accelerated interval changes neither the headline figure nor, in almost every case, the individual conclusion. Of the 402 genuine comparisons, 177 (44.0%) have a BCa interval that includes 1, against 178 (44.3%) for the percentile interval, and the two methods agree on whether the interval includes 1 in 397 of 402 comparisons (98.8%); the five disagreements are all cases in which one bound sits within 0.01 of unity. The median bias-correction constant is 0.0002, and the median acceleration constant is -0.0020 , so at these sample sizes, both corrections are

small, and the median interval width changes only from 0.370 to 0.379. The resampling-free delta-method interval is systematically wider (median width 0.495) and correspondingly more conservative, with 222 of 402 intervals (55.2%) including 1 and agreement with the percentile result in 358 of 402 comparisons (89.1%). The conclusion that a substantial share of the comparisons cannot be distinguished from unity at the available sample sizes therefore does not depend on the choice of the percentile bootstrap and is, if anything, stronger under the parametric alternative. This agreement between estimators is not evidence that any of them achieves nominal coverage at $n = 3\text{--}6$; the small-sample limitation stated in Section 2.3 is retained, and all three intervals are reported for every comparison in Table S9 so that a reader can apply whichever is preferred.

Table 4. Bootstrap uncertainty and dataset-version sensitivity.

Metric	Value
Valid retention/comparison factors	412
Structural identities	10
Genuine comparisons	402
Genuine factors with 95% CI including 1	178/402 (44.3%)
Dataset-version comparisons (genuine)	201
Low sensitivity	65
Moderate sensitivity	84
High sensitivity	52
Critical sensitivity	0
High-sensitivity rate	52/201 (25.9%)

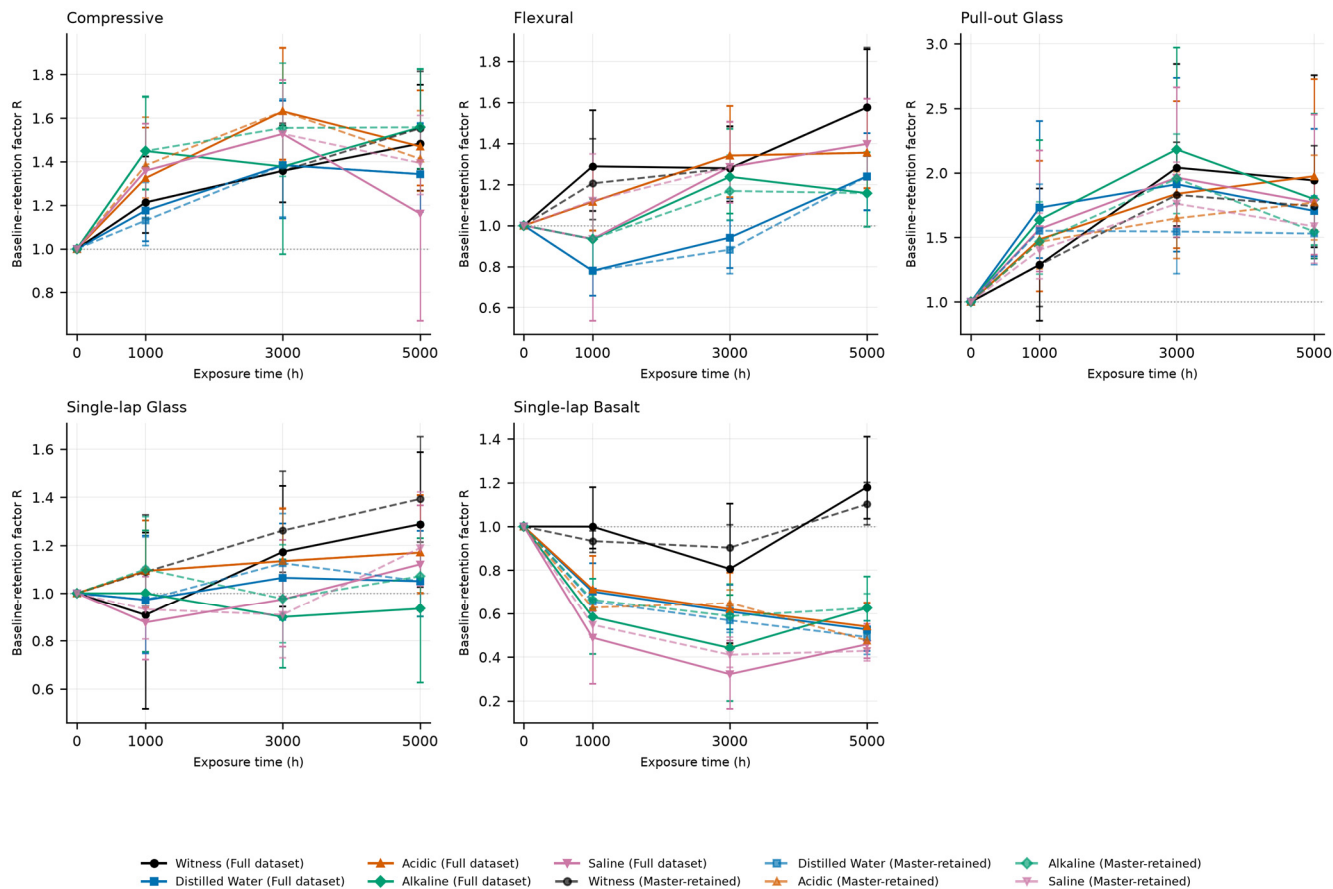


Figure 4. Baseline-retention trajectories for the five families with resolved baselines.

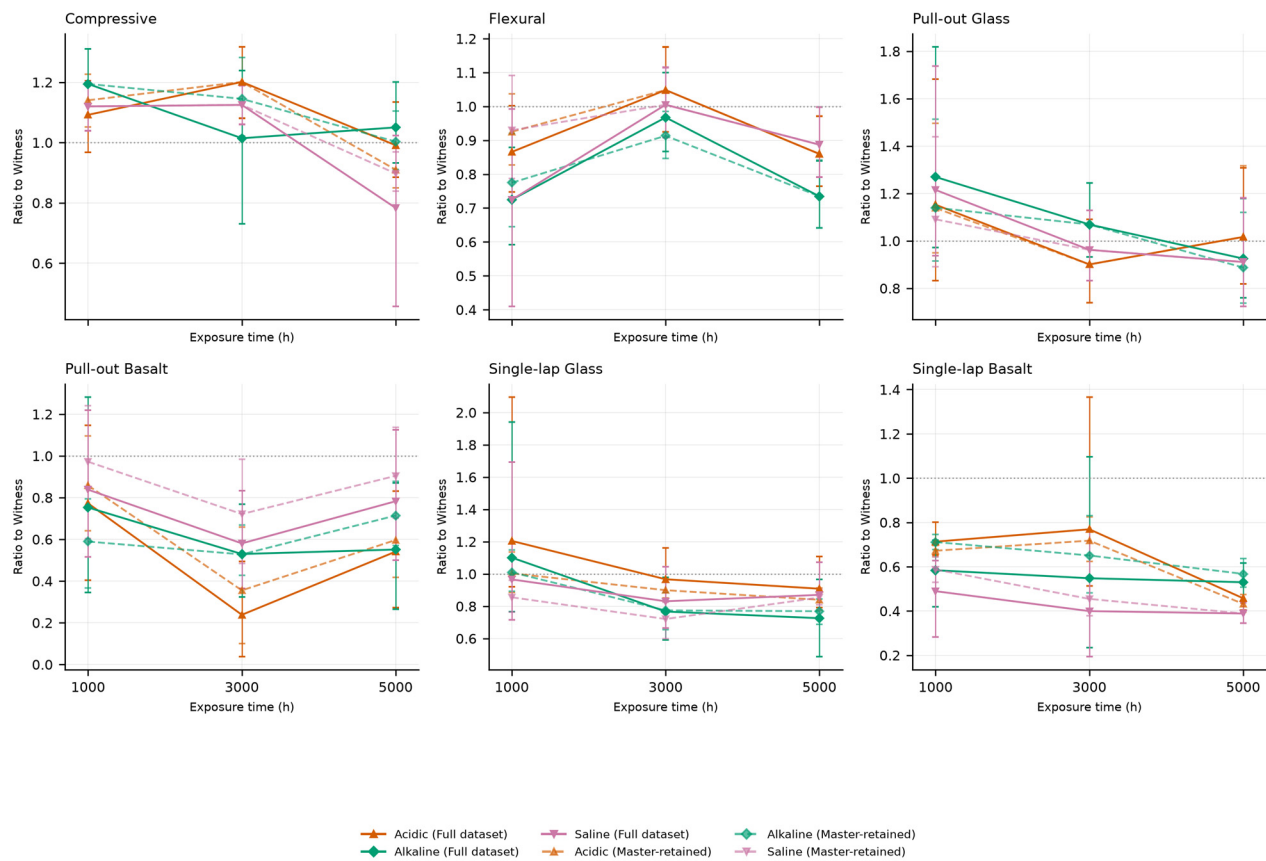


Figure 5. Aggressive-medium-to-Witness ratios.

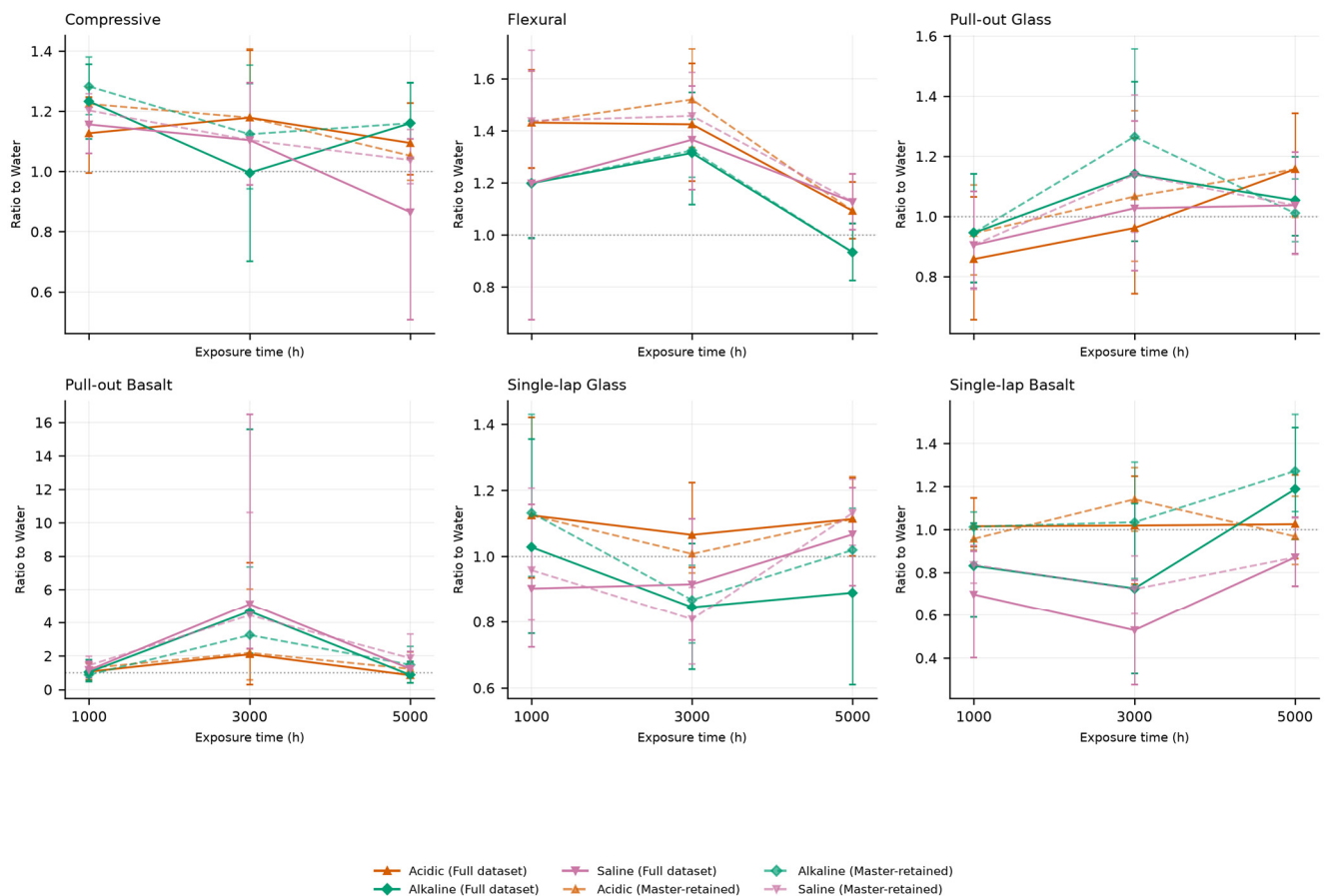


Figure 6. Aggressive-medium-to-Water ratios.

Figures 4–6 are complementary rather than interchangeable. Baseline retention measures total change from the shared pre-exposure state, whereas the time-matched Witness and Water ratios progressively introduce same-age reference conditions. Reading the three views together distinguishes an apparent change from baseline from a contrast that remains after age or immersion history is represented in the denominator.

3.2.1. Change Relative to the Pre-Exposure Baseline

Figure 4 shows the total response accumulated between the 90-day reference and each exposure time. Pull-out glass exhibited consistently elevated endpoint ratios, while single-lap basalt generally remained below its baseline in the aggressive media. Compressive and flexural responses often rise with time, including under Witness conditions, demonstrating why an upward baseline trajectory cannot be attributed uniquely to environmental resistance. Pull-out basalt is omitted from this figure because GD42 prevents a defensible 0 h denominator; no substitute baseline was introduced.

3.2.2. Aggressive Medium Relative to Time-Matched Witness

The Witness-normalized view changes the interpretation of several rising baseline trajectories. Ratios approaching 1 indicate that the aggressive-medium group and the same-age dry reference changed by similar proportions, even when both increased relative to 0 h. Confidence intervals crossing 1 remain common, so visual departures from the reference line should not be treated as clear evidence of an environmental effect without their bootstrap intervals. This comparison controls for nominal specimen age but not for immersion history.

3.2.3. Aggressive Medium Relative to Time-Matched Water

Water normalization provides the closest available experimental contrast for immersion-related effects. Differences between Figures 5 and 6 therefore indicate that the chosen reference history materially affects the calculated factor. The Pull-out Basalt alkaline example at 3000 h is especially sensitive in magnitude (4.663 for the full dataset and 3.242 for the master-retained subset), although both estimates remain above 1, and both intervals exclude 1. The Water ratio reduces age and moisture-history confounding but does not isolate a unique chemical mechanism.

3.3. Dataset-Version Sensitivity

Of 206 factor comparisons available under both dataset versions, five are structural identity comparisons and 201 are genuine comparisons. Among the 201 genuine comparisons: 65 (32.3%) are low sensitivity, 84 (41.8%) moderate, 52 (25.9%) high, and zero critical (Table 4, Figure 7). The point-estimate threshold used for an initial sensitivity screen (absolute difference >0.02 or relative difference $>5\%$) flags 132 of 206 total comparisons (64.1%); restricted to the 201 genuine comparisons, this is 132/201 ($\approx 65.7\%$)—essentially unchanged by excluding the structural identities, since those always have zero difference by definition. A representative example (Pull-out basalt, alkaline environment, 3000 h, time-matched Water ratio): the full-dataset point estimate is 4.663 versus 3.242 for the master-retained subset (Figures 6 and 7), classified as moderate sensitivity (direction relative to 1 unchanged in both versions; both confidence intervals exclude 1). Varying the screening thresholds does not change this result. Applying (0.01, 2.5%), (0.02, 5%), (0.03, 7.5%) and (0.05, 10%) in turn leaves the high-sensitivity count at 52 of 201 (25.9%) and the critical count at zero in every case because those two categories are defined by a change in the direction relative to 1, or in whether the interval includes 1, and not by any numerical tolerance. Only the boundary between the low and moderate categories moves, from 63 low and 86 moderate at the tightest thresholds to 83 and 66 at the loosest, reclassifying at

most 18 of the 201 comparisons relative to the published values (Table S5). The headline conclusion about dataset-version dependence is therefore a property of the comparisons themselves rather than of the thresholds chosen to summarize them.

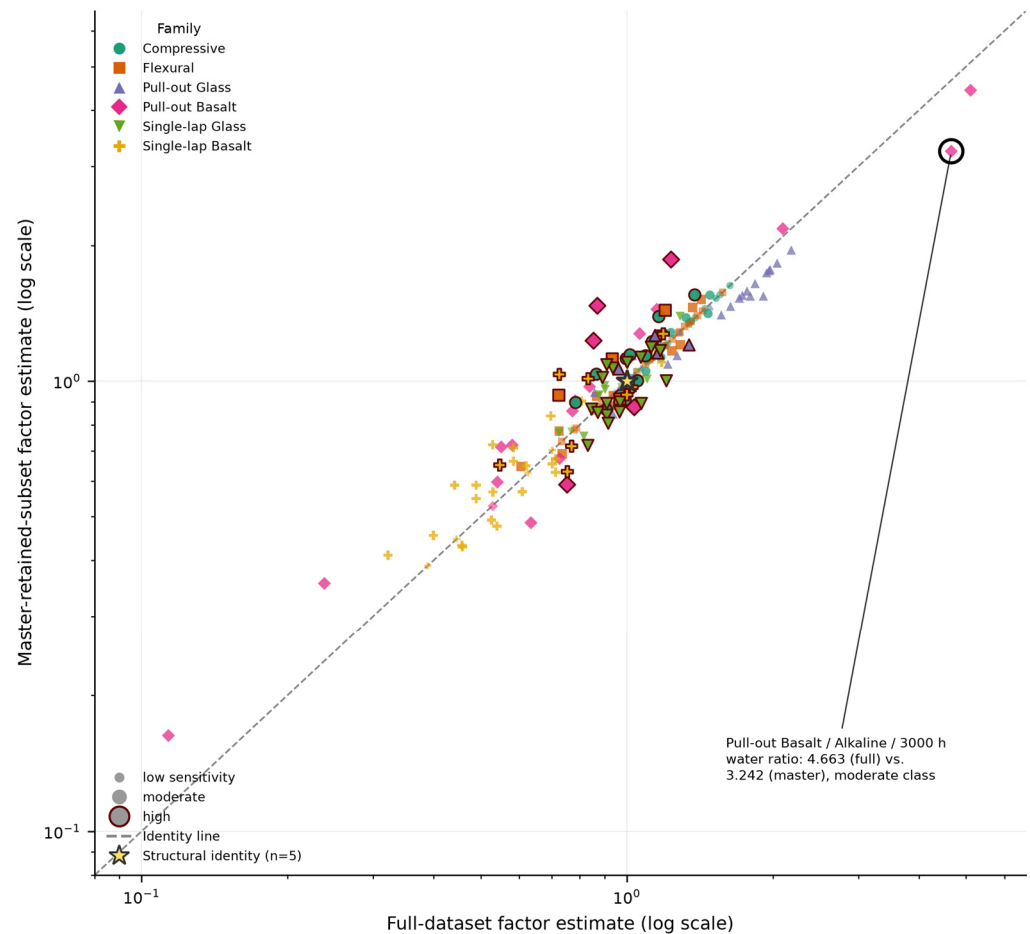


Figure 7. Factor sensitivity to dataset-version choice.

3.4. Trajectory Readiness and Surrogate-Model Selection

Of 60 defined baseline-retention trajectories, 50 have a complete four-point observed series and are eligible for surrogate-model fitting; the remaining 10 (all Pull-out Basalt, both dataset versions, all five environments) are blocked by the unresolved baseline reference group and carry no fitted model (Table 5, Figures 2 and 4). Among the 50 ready trajectories, the selected model family is linear in 27 cases, quadratic in 20, and a conditionally validated Gaussian process in three (Table 5, Figure 8a). No trajectory required a fallback to a purely descriptive or fully discrete representation. At full four-point calibration, 19 of 50 Gaussian-process fits (38%) showed boundary-hitting hyperparameters, disqualifying them from selection; among leave-one-time-level-out refits (three points), 97 of 200 Gaussian-process fits (48.5%) showed a hyperparameter warning and 117 of 200 (58.5%) showed detectable oscillation under the criterion defined in Section 2.4. Repeating the selection with the parsimony tolerance set to 0%, 5%, 10%, 15% and 20% gives 25, 26, 27, 27 and 28 linear selections, 21, 20, 20, 20 and 19 quadratic selections, and 4, 4, 3, 3 and 3 conditional Gaussian-process selections, respectively. Forty-seven of the 50 trajectories keep the same family at every tolerance, and only two of the 50 selections made at the adopted 10% value are decided by the parsimony rule at all; in the remaining 48, the linear model is either the best-performing eligible candidate on interior error or is not selected. The tolerance is therefore not load-bearing, and 10% sits in the middle of a plateau across which the outcome does not change (Table S4). Interior and endpoint cross-validation

errors for all four candidate families on every eligible trajectory, with the best candidate, the second-best candidate and the difference between them, are reported in Table S3; the median gap between the best and second-best eligible candidate is 0.042 in interior RMSE, so in a number of trajectories, the ranking rests on a small difference, and the selected form should be read as the best-supported representation among those tested rather than as an identified model. PCHIP attains the lowest interior error in 20 of the 50 trajectories but is never selected because it reproduces its calibration points exactly by construction and is retained only as a descriptive aid. Gaussian-process hyperparameters, their bound status and the oscillation count for every fit are listed in Table S7.

Table 5. Trajectory readiness and selected surrogate models.

Metric	Value
Baseline-retention trajectories defined	60
Ready (complete four-point series)	50
Blocked (unresolved reference group)	10
Selected model: linear	27
Selected model: quadratic	20
Selected model: conditional Gaussian process	3
Selected model: unresolved	10

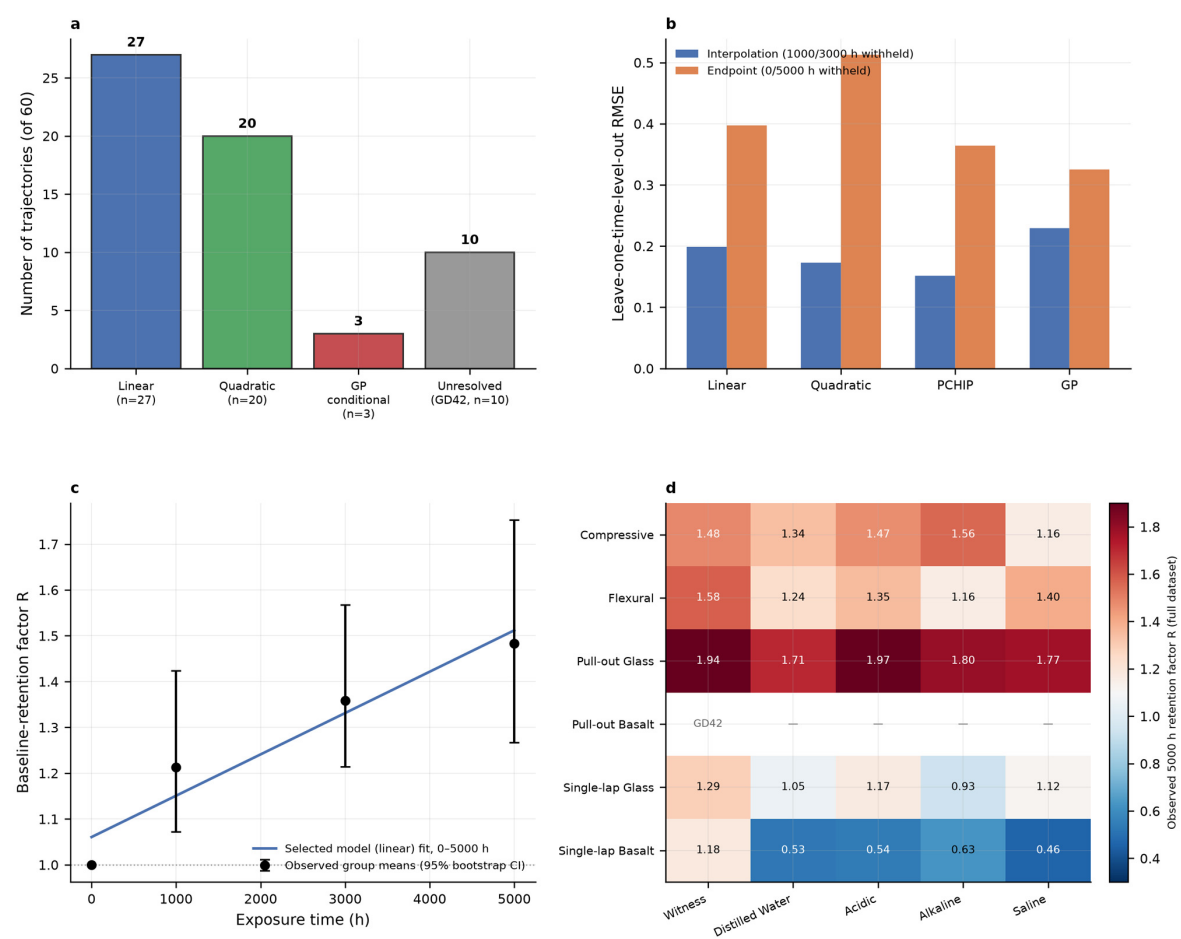


Figure 8. Model validation and categorical durability atlas: (a) selected model families and unresolved trajectories; (b) leave-one-time-level-out RMSE for interior and endpoint holdouts; (c) representative Compressive/Witness linear fit with observed means and 95% bootstrap confidence intervals; and (d) observed 5000 h baseline-retention atlas for the full dataset.

Figure 8c illustrates a representative Compressive/Witness linear trajectory together with the observed group means and their 95% bootstrap confidence intervals. The surrogate layer contributes more than graphical interpolation: it provides a reproducible rule for choosing trajectory complexity, quantifies withheld-time prediction error, and exposes whether specimen-retention decisions alter the selected functional form. For example, Compressive/Witness selects a linear model under both versions (interior RMSE 0.0734 full; 0.0671 retained), whereas Flexural/Saline changes from linear to quadratic (0.1276 full; 0.0127 retained). Pull-out glass/acidic changes from quadratic to linear, and single-lap basalt/Witness from linear to quadratic. These contrasts are summarized in Table 6 and show why the atlas preserves both versions rather than presenting one apparently definitive curve.

Table 6. Representative surrogate selections under the two dataset versions.

Trajectory	Full	Retained	Interior RMSE (F)	Interior RMSE (M)	Interpretive Role
Compressive/Witness	Linear	Linear	0.0734	0.0671	Stable form
Flexural/Saline	Linear	Quadratic	0.1276	0.0127	Form changes
Pull-out Glass/Acidic	Quadratic	Linear	0.1543	0.1962	Form changes
Single-lap Basalt/Witness	Linear	Quadratic	0.2098	0.0233	Form changes

RMSE values are leave-one-time-level-out interior errors; lower values denote better withheld-interior prediction within the observed domain.

3.5. Representative Atlas Patterns

Endpoint (5000 h) baseline-retention values vary substantially by family and environment (Table 5, Figure 8d): Pull-out glass shows the largest measured increases relative to baseline across all environments (full-dataset R ranging approximately 1.71–1.97), while single-lap basalt shows measured decreases relative to baseline in every environment except Witness (full-dataset R approximately 0.46–0.63 in the three aggressive media). These are reported strictly as observed, categorical, condition-specific, textile-dependent patterns; their interpretive limits are discussed below. These endpoint values are also the basis of the cross-environment comparison developed in Section 4.1, which examines whether the three aggressive media rank consistently within a family and whether that ranking is the same across families.

4. Discussion

The three reference views answer different scientific questions and should be interpreted as a sequence rather than as competing definitions. Baseline retention describes total change from the initial state; the Witness ratio asks whether the aggressive group differs from a same-age dry group; and the Water ratio asks whether it differs from a same-age immersed group. Thus, a trajectory may rise in Figure 4 but approach 1 in Figure 5 because the Witness condition also changed with age. Figure 6 then shows whether the contrast persists after immersion history is represented. These controls reduce, but do not eliminate, confounding and cannot identify a unique chemical mechanism. For the acidic program, the companion chemo-mechanical study attributes matrix-level trends to portlandite depletion and increased porosity [42]; that mechanism is not re-derived from the present ratios.

The 25.9% high-sensitivity rate among genuine dataset-version comparisons shows that a meaningful minority of conclusions about direction-relative-to-1, or about whether a comparison is distinguishable from “no difference,” depend on which of the 59 master-workbook omissions are treated as valid data. Because 55 of those 59 omissions carry no documented justification, this dependency cannot currently be resolved by appeal to a known experimental reason; it is reported as an open, quantified source of interpretive uncertainty (Table 4) rather than resolved in either direction. To the authors’ knowledge,

within the literature reviewed for this study, no directly comparable quantified dataset-curation-sensitivity analysis has previously been reported for a TRM/FRCM durability dataset; this absence is based on a targeted, not systematic, search, and is reported as such.

The six mechanical families show distinct, textile-dependent behavior. The largest specimen-to-specimen variability is confined to the pull-out basalt configuration (median coefficient of variation 34.5%) rather than being a general property of basalt reinforcement: single-lap basalt (12.0%) is indistinguishable from single-lap glass (12.6%), while the matrix-dominated tests remain the most stable (9.1% and 11.1%) (Figures 3 and 4, Table S6). The distinction matters because it locates the instability in a specific bond-test configuration rather than in the fiber. This agrees with extensive evidence that stress transfer depends on test configuration, embedded length, textile architecture, matrix layers, and interface fracture behavior [13,52–56]. Environmental studies further show that alkaline, saline, freeze–thaw, and natural aging can affect textiles, matrices, and interfaces differently [16,21,27,28,32,34,35,57–64]. These comparisons support the interpretation that bond response is especially sensitive, but they do not establish a unique degradation mechanism for the present specimens.

With only three to four observed time points per trajectory, a Gaussian process has very little information to constrain its covariance hyperparameters [47]; this manifested directly as boundary-hitting length-scale or noise parameters in 38% of full-calibration fits and in roughly half of the three-point leave-one-out refits. Linear and quadratic regressions, by contrast, always produce a well-defined, low-variance fit at this sample size and were disqualified only when a genuinely better-validated Gaussian-process fit happened to remain stable, which occurred in just three of 50 trajectories. This supports treating Gaussian-process/Kriging models as conditionally usable at best, not as a general-purpose model for this atlas; this instability rate is, to the authors' knowledge, an empirical result specific to this small-sample regime, consistent with general Gaussian-process theory but not previously quantified for a TRM/FRCM durability application within the literature reviewed. The reason is visible in the fitted hyperparameters themselves. With three or four points spread over a normalized domain of unit length, the marginal likelihood is nearly flat in the length-scale, and the optimizer frequently terminates with a length scale one to two orders of magnitude below the spacing of the observations, the signal variance driven towards its lower bound and effectively all of the variance assigned to the noise term. That is a degenerate fit which reverts to the sample mean, not a merely inaccurate one, which is why it is disqualified rather than penalized. A second symptom points the same way: re-running the fully specified implementation reproduces the archived held-out predictions exactly for every linear, quadratic and PCHIP fit (600 of 600), but for 166 of the 200 Gaussian-process refits, the residual difference arises from the non-convex marginal-likelihood surface rather than from any change of method. A model whose fitted form depends on where the optimiser happens to terminate is not a sound basis for a durability atlas at this sample size, and this reproducibility contrast is itself the clearest available evidence for the instability reported above.

The atlas constructed here is a categorical collection of validated, discrete-condition, textile-dependent trajectories: for each mechanical family, textile, and environment, a fitted curve within the observed 0–5000-h domain, accompanied by its leave-one-out validation error and its dataset-version sensitivity classification. It does not mean a continuous chemical-severity surface—no interpolation between acid, alkali, and salt exposure was performed or is supported by the categorical experimental design; it does not mean an operational digital twin capable of ingesting new field conditions—the constructed-facilities digital-twin literature defines that concept specifically by continuous, sensor-fed, real-time updating and operational/predictive integration [49], none of which this static, retro-

spective, fixed-dataset atlas provides; and it does not mean a universal mechanistic law relating chemical exposure to mechanical response—the surrogate models used are deliberately simple curve fits validated only within the tested domain, not physically derived degradation models. The descriptive label “durability atlas” used here for this categorical synthesis is not adopted from an established, field-recognized use of that specific term for a TRM/FRCM context in the literature reviewed for this study. Its practical value is correspondingly specific. For a researcher it provides, for each tested condition, a single reconciled estimate with an uncertainty interval, a stated functional form with the error incurred when a time level is withheld, and an explicit flag for whether that estimate depends on which specimens were retained. It identifies the conditions that are least well constrained and therefore most in need of further testing—the pull-out basalt family being the clearest example—and it allows textile systems and test configurations to be compared on a common basis, which the source studies individually do not permit. For an engineer it indicates which response families are comparatively robust and which are sensitive to the choice of reference, and therefore which published durability figures should be treated with caution. What it does not provide is a service-life estimate, a value for any condition that was not tested, or a basis for choosing between environments that were never compared experimentally.

The largest single source of unresolved uncertainty in this atlas, the blocked pull-out basalt baseline trajectories, stems from a single unresolved specimen-linkage gap, illustrating how one missing baseline record can propagate into ten blocked trajectories downstream. Future durability campaigns intending to support the surrogate-model or atlas-style synthesis of textile-dependent response would benefit from more than four time points per condition, to make quadratic or Gaussian-process fitting genuinely well-constrained rather than only nominally possible, and from explicit, contemporaneous documentation of any specimen exclusion decision at the time it is made, rather than relying on later forensic reconstruction. A related follow-up study examining acidity level, rather than only exposure duration, as an additional variable in TRM tensile performance under acidic aging [41] further illustrates that this study’s own fixed-concentration protocol leaves acidity level as an unexplored dimension, a natural direction for extending the present atlas, not a gap in the present study’s internal consistency.

The leave-one-time-level-out validation reported here evaluates whether a model generalizes across the observed 0–5000 h domain using this dataset alone; it does not constitute external validation against an independent dataset or exposure campaign, and it uses only two interior holdouts (1000 and 3000 h) and two endpoint holdouts (0 and 5000 h). Withholding a boundary time level is an endpoint holdout prediction, not extrapolation beyond the observed domain, the withheld value itself lies inside the tested 0–5000 h range, and no model in this study is ever evaluated outside it. Nonetheless, boundary-point prediction errors are consistently larger than interior (interpolation) prediction errors across every model family, underscoring that even within-domain boundary prediction is less reliable than interior prediction (Figure 8b). Extending this atlas with additional time points, beyond 5000 h or as additional interior points, would both permit the assessment of a wider set of candidate model families and provide a genuine external check on the trajectories reported here. Until then, this atlas should be read as a synthesis of the existing observed data, not as a validated forecasting tool.

Throughout, mechanical trends reported in this study are not, by themselves, interpreted as proof of a specific chemical mechanism, and no textile degradation mechanism is inferred solely from the composite mechanical behavior reported here; where mechanistic interpretation is offered, it is attributed explicitly to the relevant source publication or companion study. The three source publications report medium-specific characterization

evidence rather than an identical, shared method set: the acidic-exposure study [37] reports optical microscopy, scanning electron microscopy, X-ray diffraction, and thermogravimetric analysis; the alkaline-exposure study [38] reports X-ray diffraction and thermogravimetric analysis, but does not report scanning electron microscopy or optical microscopy as part of its published characterization; and the saline-exposure study [39] reports X-ray diffraction and optical microscopy. None of this characterization evidence is re-analyzed in the present integration study, and the method attribution above is stated only where directly verified from the corresponding source publication.

4.1. Cross-Environment Comparison of Ageing Responses

Integrating the three campaigns permits a comparison that none of the source studies can make on their own. Within a single mechanical family, measured on the same response variable and against the same reference conditions, how do acidic, alkaline and saline ageing rank against one another? Table 7 reports the 5000-h values for the six families under all three reference definitions for the full dataset, and two features of that table are the principal integrative result of this study.

Table 7. Cross-environment comparison at 5000 h: response of each mechanical family to the three aggressive media under the three reference definitions (full dataset).

Reference View	Family	Acidic	Alkaline	Saline	Ordering at 5000 h (Lowest First)
Baseline retention R_B	Compressive	1.470	1.558	1.162	Saline, acidic, alkaline
	Flexural	1.355	1.158	1.398	398Alkaline, acidic, saline
	Pull-out Glass	1.974	1.797	1.796	Saline, alkaline, acidic
	Pull-out Basalt	-	-	-	Unresolved baseline (GD42)
	Single-lap Glass	1.170	0.935	1.121	Alkaline, saline, acidic
	Single-lap Basalt	0.540	0.625	0.459	Saline, acidic, alkaline
Time-matched Witness Ratio R_W	Compressive	0.991	1.050	0.783	Saline, acidic, alkaline
	Flexural	0.860	0.735	0.887	Alkaline, acidic, saline
	Pull-out Glass	1.017	0.925	0.911	Saline, alkaline, acidic
	Pull-out Basalt	0.541	0.551	0.783	Acidic, alkaline, saline
	Single-lap Glass	0.908	0.726	0.870	Alkaline, saline, acidic
	Single-lap Basalt	0.457	0.530	0.389	Saline, acidic, alkaline
Time-matched Water Ratio R_D	Compressive	1.095	1.160	0.865	Saline, acidic, alkaline
	Flexural	1.092	0.933	1.127	Alkaline, acidic, saline
	Pull-out Glass	1.157	1.054	1.037	Saline, alkaline, acidic
	Pull-out Basalt	0.853	0.868	1.233	Acidic, alkaline, saline
	Single-lap Glass	1.114	0.890	1.067	Alkaline, saline, acidic
	Single-lap Basalt	1.025	1.188	0.871	Saline, acidic, alkaline

Full-dataset reference factors at 5000 h; 95% bootstrap confidence intervals for every entry are in Table S8. Ordering runs from the lowest to the highest value. Pull-out basalt baseline retention is unresolved (GD42).

The first is that the ordering of the three media within a family is identical under all three reference definitions. For compressive strength, the order at 5000 h is saline, then acidic, then alkaline under baseline retention, under the Witness ratio and under the Water ratio alike; the same internal consistency holds for flexural (alkaline, acidic, saline), pull-out glass (saline, alkaline, acidic), single-lap glass (alkaline, saline, acidic) and single-lap basalt (saline, acidic, alkaline), and for pull-out basalt under the two reference views available

to it (acidic, alkaline, saline). The absolute level of a ratio moves substantially when the reference changes, to the point that the same condition can sit above 1 in one view and below 1 in another, which happens in 25 of the 45 family, environment and time cells for which all three views exist. The relative ordering of the media nevertheless does not move. Comparative statements between media are therefore more robust in this dataset than statements about any individual ratio, which is the opposite of what the reference sensitivity of the individual factors alone would suggest, and it is an argument for reporting cross-environment contrasts rather than isolated retention values.

The second is that the ordering itself is not the same in different families. No medium is uniformly the most aggressive: saline gives the lowest 5000-h response for compressive, single-lap basalt and pull-out glass, alkaline for flexural and single-lap glass, and acidic for pull-out basalt. A durability ranking of exposure media obtained from one test configuration therefore does not transfer to another, and a single-medium study cannot establish which environment is the more severe for a TRM system in general. This is visible only once the campaigns are placed in one framework, and it is the clearest answer this study can give to the question of what integration adds.

Two further patterns carry through the comparison. Single-lap basalt under saline exposure is the most consistently separable response in the atlas: 16 of its 18 comparisons (88.9%) have intervals excluding 1, with a median ratio of 0.489, and it is the combination that falls furthest below its reference in every view. Acidic exposure shows the opposite structure. Twenty-three of its 30 baseline-retention comparisons (76.7%) exclude 1, but only 8 of 36 (22.2%) of its water-normalized comparisons do, so most of what appears as an acidic effect relative to the pre-exposure baseline is shared with plain water immersion at the same age; alkaline and saline exposure retain more of their separability under water normalization (30.6% and 44.4%). That distinction cannot be drawn from the acidic study alone because it requires the same water reference to be applied consistently across all three campaigns.

These are direct quantitative comparisons supported by the harmonized dataset, and they compare measured mechanical response rather than degradation mechanism. The characterization evidence available for the three media is not symmetric, as set out in the preceding paragraph, so an interpretation available for one medium may have no counterpart for another; mechanism is attributed only to the source study that measured it, and no microstructural inference is transferred between media. The comparison is also bounded by the experimental design in a further respect: each medium was applied at a single fixed concentration, so what is compared is three specific exposure conditions and not a general ordering of acid, alkali and salt severity. Reliability frameworks that proceed from a transport and corrosion model of the deterioration process, as has been done for the chloride-induced deterioration of reinforced concrete [65], differ from the present work in exactly this respect: they derive a probabilistic assessment from an assumed mechanism, whereas the atlas presented here remains empirical and categorical. The two approaches are complementary, and the gap between them indicates what a mechanistically grounded extension of this atlas would require.

4.2. Limitations

The atlas is bounded by the available evidence: each trajectory has only three to four time levels, no independent validation dataset exists, and no inference is made beyond 5000 h or between categorical exposure media. The unresolved GD42 baseline blocks 10 pull-out basalt trajectories, while 55 undocumented specimen omissions require the full-versus-retained sensitivity analysis. Time-matched ratios reduce but do not eliminate age and moisture confounding, and composite mechanical trends do not by themselves identify

textile-degradation mechanisms. These constraints make the atlas a transparent synthesis of observed behavior rather than a long-term forecasting model. Four exposure levels also mean that the analysis is blind to any change occurring between them, that a quadratic fit retains a single residual degree of freedom, and that a selected functional form states which of the tested representations best reproduces a withheld observation rather than establishing that the degradation process has that form. The cross-environment ordering reported in Section 4.1 compares measured responses in five categorical environments, each at one fixed concentration, and does not support interpolation to any untested medium, concentration or exposure duration. The uncertainty intervals reported here agree closely across three estimators, but agreement between estimators is not coverage: at $n = 3-6$, none of them can be assumed to attain their nominal level, and the intervals should be read as descriptions of the resampling or normal-theory spread rather than as calibrated coverage statements.

5. Conclusions

This study developed a surrogate-based digital durability atlas by integrating three coordinated experimental programs on glass- and basalt-textile-reinforced mortar exposed to acidic, alkaline, and saline environments. The analysis combined specimen-level data auditing, parallel dataset reconstruction, bootstrap uncertainty quantification, and cross-validated surrogate modeling. This approach enabled environmental and textile-dependent responses to be compared within a consistent framework while explicitly accounting for undocumented specimen omissions and limited temporal sampling. The principal conclusions are as follows:

- An audited specimen-level database comprising 577 records and 96 groups was constructed from three previously separate acidic, alkaline, and saline TRM durability studies. The full and master-retained dataset versions were preserved throughout, enabling the explicit quantification of the effects of specimen-retention decisions.
- Of the 402 genuine, non-identity retention and comparison factors, 44.3% had bootstrap 95% confidence intervals including 1. Thus, a substantial proportion of the measured differences could not be clearly distinguished from no difference at the available sample sizes. The same conclusion follows from a bias-corrected and accelerated bootstrap (177 of 402, 44.0%), which agreed with the percentile result in 98.8% of comparisons, and from a resampling-free delta-method interval, which was more conservative still.
- Dataset-version choice changed the qualitative interpretation of 25.9% of the genuine comparisons, although none reached the mutually contradictory critical classification. This finding underscores the importance of transparent specimen-retention reporting and warrants caution when reconstructed results are compared with published summary values.
- Fifty of the 60 defined baseline-retention trajectories supported validated low-order surrogate models: 27 linear models, 20 quadratic models, and three conditionally accepted Gaussian-process models. The remaining 10 pull-out basalt trajectories were blocked by one unresolved baseline reference group, demonstrating both the feasibility and the present limits of the atlas. These selections are stable against the tolerance used to make them: 47 of the 50 trajectories retain the same family when the parsimony tolerance is varied between 0% and 20%. The selected form identifies the best-supported low-order representation among the tested candidates within the observed data and is not evidence that the underlying degradation process takes that form.

- Gaussian-process/Kriging models frequently exhibited boundary-hitting hyperparameters with only three to four calibration points. Such instability occurred in approximately 38–48% of the evaluated fits, indicating that Gaussian processes should not be adopted as a general-purpose model family for similarly sparse durability datasets. The fitted hyperparameters show why: with three or four calibration points, the optimizer frequently terminates at a length-scale one to two orders of magnitude below the spacing of the observations, assigning effectively all variance to the noise term.
- Specimen-to-specimen variability was quantified rather than described: the median coefficient of variation was 34.5% for pull-out basalt against 16.5% for pull-out glass, 12.0% and 12.6% for single-lap basalt and single-lap glass, and 9.1% and 11.1% for the compressive and flexural families. The elevated basalt variability is concentrated in the pull-out configuration rather than being a general property of basalt reinforcement, so durability conclusions for TRM systems depend on test configuration at least as much as on textile type and should not be generalized across either.
- Placing the three campaigns in one framework showed that the ranking of acidic, alkaline and saline ageing within a mechanical family is identical under all three reference definitions but differs between families, so that no medium is uniformly the most aggressive. Single-lap basalt under saline exposure was the most consistently separable response in the atlas, whereas most of the apparent acidic effect relative to the pre-exposure baseline is shared with plain water immersion at the same age. Neither observation is available from the source studies individually.
- The resulting durability atlas is a categorical, uncertainty-aware, and validation-bounded synthesis of textile-dependent trajectories within the observed 0–5000-h domain. It should not be interpreted as a continuous chemical-severity surface, an operational digital twin, or a validated long-term forecasting tool.
- To the authors' knowledge, no directly comparable study was identified within the targeted literature that combines cross-environment specimen-level TRM integration, explicit dataset-curation sensitivity, bootstrap uncertainty, and validated low-order surrogate trajectories for glass- and basalt-textile systems. The contribution is therefore integrative and methodological rather than experimental or mechanistic.
- Future work should resolve the pull-out basalt baseline-data gap, introduce additional interior and longer-term exposure points, examine exposure-severity variables such as acidity level, and document specimen-exclusion decisions contemporaneously. These developments would support stronger external validation and more robust surrogate model selection.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/constrmater6050076/s1>. Tables S1–S9 provide detailed information on experimental comparability, specimen omissions, model validation, sensitivity analyses, variability, cross-environment comparisons, and uncertainty intervals.

Author Contributions: Conceptualization, N.A., O.H., M.B., M.H.A. and D.B.; methodology, N.A., O.H., M.B., M.H.A. and D.B.; software, N.A.; validation, N.A., O.H. and M.B.; formal analysis, N.A.; investigation, N.A., O.H., M.B. and D.B.; resources, D.B.; data curation, N.A.; writing—original draft preparation, N.A.; writing—review and editing, O.H., M.B., M.H.A. and D.B.; visualization, N.A.; supervision, M.B. and D.B.; project administration, N.A. and D.B.; funding acquisition, N.A., O.H. and D.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the 1.1.1.9 Activity “Post-doctoral Research”, research application No. 1.1.1.9/LZP/1/24/068, “Eco-efficient and circular building materials using local Baltic industrial and municipal by-products and wastes”. The second author acknowledges support from the 1.1.1.9 Activity “Post-doctoral Research”, research application No. 1.1.1.9/LZP/2/25/290,

“Use of Algal Biomass in the Production of Carbon Cured Cement, LC3, and Alkali Activated Geopolymer Composites”.

Data Availability Statement: The harmonized specimen-level database and analysis scripts supporting the findings of this study are available from the corresponding author upon reasonable request. The data require accompanying provenance and data-curation documentation for appropriate interpretation.

Acknowledgments: The authors acknowledge Riga Technical University for institutional support. The authors also acknowledge the previous experimental work and datasets that provided the basis for the validation case study used in this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

TRM	Textile-reinforced mortar
FRCM	Fabric-reinforced cementitious matrix
CoV	Coefficient of variation
CI	Confidence interval
LOTO	Leave-one-time-level-out (cross-validation)
GP	Gaussian process
RMSE	Root-mean-square error
MAE	Mean absolute error
PCHIP	Piecewise cubic Hermite interpolating polynomial
BCa	Bias-corrected and accelerated (bootstrap)

References

1. Koutas, L.N.; Tetta, Z.; Bournas, D.A.; Triantafillou, T.C. Strengthening of Concrete Structures with Textile Reinforced Mortars: State-of-the-Art Review. *J. Compos. Constr.* **2019**, *23*, 03118001. [\[CrossRef\]](#)
2. Boem, I. Masonry Elements Strengthened with TRM: A Review of Experimental, Design and Numerical Methods. *Buildings* **2022**, *12*, 1307. [\[CrossRef\]](#)
3. Papanicolaou, C.G.; Triantafillou, T.C.; Karlos, K.; Papathanasiou, M. Textile-Reinforced Mortar (TRM) versus FRP as Strengthening Material of URM Walls: In-Plane Cyclic Loading. *Mater. Struct./Mater. Constr.* **2007**, *40*, 1081–1097. [\[CrossRef\]](#)
4. de Felice, G.; De Santis, S.; Garmendia, L.; Ghiassi, B.; Larrinaga, P.; Lourenço, P.B.; Oliveira, D.V.; Paolacci, F.; Papanicolaou, C.G. Mortar-Based Systems for Externally Bonded Strengthening of Masonry. *Mater. Struct./Mater. Constr.* **2014**, *47*, 2021–2037. [\[CrossRef\]](#)
5. Lignola, G.P.; Caggegi, C.; Ceroni, F.; De Santis, S.; Krajewski, P.; Lourenço, P.B.; Morganti, M.; Papanicolaou, C.; Pellegrino, C.; Prota, A.; et al. Performance Assessment of Basalt FRCM for Retrofit Applications on Masonry. *Compos. B Eng.* **2017**, *128*, 1–18. [\[CrossRef\]](#)
6. Azimi, N.; Oliveira, D.V.; Ghiassi, B.; Dalalbashi, A.; Lourenço, P.B. Evaluation of Creep Effects on Textile-Reinforced Mortar via Single-Lap Shear Testing. *Constr. Build. Mater.* **2025**, *497*, 143897. [\[CrossRef\]](#)
7. Carloni, C.; Verre, S.; Sneed, L.H.; Ombres, L. Open Issues on the Investigation of PBO FRCM-Concrete Debonding. *Compos. Struct.* **2022**, *299*, 116062. [\[CrossRef\]](#)
8. Ombres, L.; Iorfida, A.; Mazzuca, S.; Verre, S. Bond Analysis of Thermally Conditioned FRCM-Masonry Joints. *Measurement* **2018**, *125*, 509–515. [\[CrossRef\]](#)
9. Anania, L. The Effectiveness of FRCM-U-Wrap Joist: An Alternative Solution in Flexural Strengthening of Reinforced Concrete Beams. *Mech. Adv. Mater. Struct.* **2021**, *28*, 237–251. [\[CrossRef\]](#)
10. Franzoni, E.; Gentilini, C.; Santandrea, M.; Zanutto, S.; Carloni, C. Durability of Steel FRCM-Masonry Joints: Effect of Water and Salt Crystallization. *Mater. Struct./Mater. Constr.* **2017**, *50*, 201. [\[CrossRef\]](#)
11. Azimi, N.; Dalalbashi, A.; Oliveira, D.V.; Ghiassi, B.; Lourenço, P.B. Tensile Behavior of Textile-Reinforced Mortar: Influence of Test Setup and Layer Arrangement. *Constr. Build. Mater.* **2023**, *394*, 132185. [\[CrossRef\]](#)
12. Azimi, N.; Dalalbashi, A.; Oliveira, D.V.; Ghiassi, B.; Lourenço, P.B. Tensile Behavior of Textile-Reinforced Mortar: Influence of the Number of Layers and Their Arrangement. *Key Eng. Mater.* **2022**, *916*, 91–97. [\[CrossRef\]](#)

13. De Santis, S.; Carozzi, F.G.; de Felice, G.; Poggi, C. Test Methods for Textile Reinforced Mortar Systems. *Compos. B Eng.* **2017**, *127*, 121–132. [\[CrossRef\]](#)
14. D’Antino, T.; Papanicolaou, C. Mechanical Characterization of Textile Reinforced Inorganic-Matrix Composites. *Compos. B Eng.* **2017**, *127*, 78–91. [\[CrossRef\]](#)
15. D’Antino, T.; Papanicolaou, C. (Corina) Comparison between Different Tensile Test Set-Ups for the Mechanical Characterization of Inorganic-Matrix Composites. *Constr. Build. Mater.* **2018**, *171*, 140–151. [\[CrossRef\]](#)
16. Butler, M.; Mechtcherine, V.; Hempel, S. Experimental Investigations on the Durability of Fibre-Matrix Interfaces in Textile-Reinforced Concrete. *Cem. Concr. Compos.* **2009**, *31*, 221–231. [\[CrossRef\]](#)
17. Kohan, L.; Azmi, N.; de Aguiar Souza, I.; Fioroni, C.; Azevedo, A.G.S.; Freitas, T.O.G.; Peixoto, J.J.; Ferreira, D.S.P.; Oliveira, D.V.; Barúque-Ramos, J. Coated Sisal/PET Hybrid Braided Textiles Interface Behavior into Cementitious Matrix. *Compos. Interfaces* **2025**, *32*, 663–687. [\[CrossRef\]](#)
18. Naik, D.L.; Sharma, A.; Chada, R.R.; Kiran, R.; Sirotiak, T. Modified Pullout Test for Indirect Characterization of Natural Fiber and Cementitious Matrix Interface Properties. *Constr. Build. Mater.* **2019**, *208*, 381–393. [\[CrossRef\]](#)
19. Agossou, O.G.; Homoro, O.; Amziane, S. Optimising the Textile-Matrix Interface for Enhanced Mechanical Performance in Vegetal FRCM Composites. *J. Build. Eng.* **2025**, *114*, 114334. [\[CrossRef\]](#)
20. Al-Lami, K.; D’Antino, T.; Colombi, P. Durability of Fabric-Reinforced Cementitious Matrix (FRCM) Composites: A Review. *Appl. Sci.* **2020**, *10*, 1714. [\[CrossRef\]](#)
21. Alma’aitah, M.; Ghiassi, B.; Dalalbashi, A. Durability of Textile Reinforced Concrete: Existing Knowledge and Current Gaps. *Appl. Sci.* **2021**, *11*, 2771. [\[CrossRef\]](#)
22. Mahboob, A.; Hasanshahi, O.; Safi, M.; Majid, T.A. Experimental Investigation of Eco-Friendly Fiber-Reinforced Concrete Using Recycled and Natural Fibers, Integrated with Recycled Aggregates. *Adv. Compos. Mater.* **2024**, *33*, 1101–1130. [\[CrossRef\]](#)
23. Hasanshahi, O.; Kameli, S.; Mohammadpour, R.; Bakhshi, M.; Mahboob, A. Experimental Investigation of Sustainable Roller-Compacted Concrete Pavement (RCCP) Reinforced with Natural Fibers: Mechanical Performance, Durability, and Environmental Impacts. *Int. J. Pavement Res. Technol.* **2025**. [\[CrossRef\]](#)
24. Ramezansefat, H.; Rezazadeh, M.; Barros, J.; Valente, I.; Bakshi, M. Modelling the High Strain Rate Tensile Behavior of Steel Fiber Reinforced Concrete Using Artificial Neural Network Approach. In *10th International Conference on FRP Composites in Civil Engineering: Proceedings of CICE 2020/2021*; Ilki, A., Ispir, M., Inci, P., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 1099–1109. [\[CrossRef\]](#)
25. Cascardi, A.; Verre, S.; Micelli, F.; Aiello, M.A. Durability-Aimed Performance of Glass FRCM-Confined Concrete Cylinders: Experimental Insights into Alkali Environmental Effects. *Mater. Struct.* **2025**, *58*, 329. [\[CrossRef\]](#)
26. Arboleda, D.; Babaeidarabad, S.; DiLaurenzio Hays, C.; Nanni, A. Durability of Fabric-Reinforced Cementitious Matrix (FRCM) Composites. In *Proceedings of the 7th International Conference on FRP Composites in Civil Engineering (CICE 2014)*, Vancouver, BC, Canada, 20–22 August 2014; pp. 384–390.
27. Wang, Q.; Ding, Y.; Randl, N. Investigation on the Alkali Resistance of Basalt Fiber and Its Textile in Different Alkaline Environments. *Constr. Build. Mater.* **2021**, *272*, 121670. [\[CrossRef\]](#)
28. Orłowsky, J.; Raupach, M. Durability Model for AR-Glass Fibres in Textile Reinforced Concrete. *Mater. Struct./Mater. Constr.* **2008**, *41*, 1225–1233. [\[CrossRef\]](#)
29. Donnini, J. Effects of Different Environments on Mechanical Properties. *Compos. Part B Eng.* **2019**, *174*, 107047. [\[CrossRef\]](#)
30. Nobili, A. Durability Assessment of Impregnated Glass Fabric Reinforced Cementitious Matrix (GFRCM) Composites in the Alkaline and Saline Environments. *Constr. Build. Mater.* **2016**, *105*, 465–471. [\[CrossRef\]](#)
31. Signorini, C.; Nobili, A. Comparing Durability of Steel-Reinforced Grout (SRG) and Textile-Reinforced Mortar (TRM) for Structural Retrofitting. *Mater. Struct.* **2021**, *54*, 131. [\[CrossRef\]](#)
32. Alatawna, A.; Nahum, L.; Sripada, R.; Birenboim, M.; Regev, O.; Peled, A. Textile-Reinforced Mortar: Durability in Salty Environment. *Cem. Concr. Compos.* **2022**, *130*, 104534. [\[CrossRef\]](#)
33. Franzoni, E.; Santandrea, M.; Gentilini, C.; Fregni, A.; Carloni, C. The Role of Mortar Matrix in the Bond Behavior and Salt Crystallization Resistance of FRCM Applied to Masonry. *Constr. Build. Mater.* **2019**, *209*, 592–605. [\[CrossRef\]](#)
34. Karakasis, I.C.; Papanicolaou, C.G.; Triantafyllou, T.C. Residual Textile Reinforced Mortar-to-Concrete Bond after Exposure to Harsh Environments. *Struct. Concr.* **2025**, *26*, 3597–3610. [\[CrossRef\]](#)
35. De Munck, M.; El Kadi, M.; Tsangouri, E.; Vervloet, J.; Verbruggen, S.; Wastiels, J.; Tysmans, T.; Remy, O. Influence of Environmental Loading on the Tensile and Cracking Behaviour of Textile Reinforced Cementitious Composites. *Constr. Build. Mater.* **2018**, *181*, 325–334. [\[CrossRef\]](#)
36. Ceroni, F.; Bonati, A.; Galimberti, V.; Occhiuzzi, A. Effects of Environmental Conditioning on the Bond Behavior of FRP and FRCM Systems Applied to Concrete Elements. *J. Eng. Mech.* **2018**, *144*, 04017144. [\[CrossRef\]](#)
37. Azimi, N.; Schollbach, K.; Oliveira, D.V.; Lourenço, P.B. Effect of Acidic Environment Exposure on Mechanical Properties of Trm Composites. *Constr. Build. Mater.* **2024**, *438*, 137142. [\[CrossRef\]](#)

38. Azimi, N.; Schollbach, K.; Oliveira, D.V.; Lourenço, P.B. Effect of Exposure to Alkaline Environment on the Mechanical Properties of TRM Composites. *J. Build. Eng.* **2025**, *105*, 112468. [\[CrossRef\]](#)
39. Azimi, N.; Oliveira, D.V.; Schollbach, K.; Lourenço, P.B. Durability and Bond Behavior of Textile-Reinforced Mortar Composites under Prolonged Saline Exposure. *J. Build. Eng.* **2025**, *112*, 113887. [\[CrossRef\]](#)
40. Taheri, A.; Azmi, N.; Oliveira, D.V.; Tinoco, J.; Lourenço, P.B. Integrating Experimental Analysis and Gradient Boosting for the Durability Assessment of Lime-Based Mortar in Acidic Environment. *Buildings* **2025**, *15*, 408. [\[CrossRef\]](#)
41. Azimi, N.; Oliveira, D.V.; D'Antino, T.; Schollbach, K.; Lourenço, P.B. Durability of Textile-Reinforced Mortar in Acidic Environments: Effects of Acidity Level and Exposure Time on Tensile Performance. *Mater. Struct.* **2026**, *59*, 151. [\[CrossRef\]](#)
42. Azimi, N.; Schollbach, K.; Oliveira, D.V.; D'Antino, T.; Lourenço, P.B. Durability of Natural Hydraulic Lime–Pozzolan Mortars for TRM Strengthening Systems under Acidic Aging: Linking Microstructural Degradation to Mechanical Performance via Chemo-Mechanical Modeling. *Constr. Build. Mater.* **2025**, *500*, 144244. [\[CrossRef\]](#)
43. Azimi, N.; Hassanshahi, O.; Bakhshi, M.; Mehdipour, Z.; Sangdehi, S.M.S.; Bajare, D. A Leaching-Index-Driven Framework for Durability-Oriented Design of Mineral Binders: Validation on Acid-Induced Degradation of an NHL–Pozzolan System and Prospective Extensions to Circular Materials. *Appl. Sci.* **2026**, *16*, 7030. [\[CrossRef\]](#)
44. ASTM E178-21; Standard Practice for Dealing with Outlying Observations. ASTM International: West Conshohocken, PA, USA, 2021.
45. Taffese, W.Z.; Hilloulou, B.; Villagran Zaccardi, Y.; Marani, A.; Nehdi, M.L.; Hanif, M.U.; Kamath, M.; Nunes, S.; von Greve-Dierfeld, S.; Kanellopoulos, A. Machine Learning in Concrete Durability: Challenges and Pathways Identified by RILEM TC 315-DCS towards Enhanced Predictive Models. *Mater. Struct.* **2025**, *58*, 145. [\[CrossRef\]](#)
46. Tibshirani, R.J.; Efron, B. *An Introduction to the Bootstrap*; Monographs on Statistics and Applied Probability; Springer: Berlin/Heidelberg, Germany, 1993; Volume 57, pp. 1–436.
47. Dong, B.; Zhao, S.; Zhang, Y.; Zhang, Y.; Kitipornchai, S.; Yang, J. Chloride Ingress Mechanisms-Informed Virtual Reliability Assessment of Corroded Reinforced Concrete Structures. *Struct. Saf.* **2026**, *121*, 102716. [\[CrossRef\]](#)
48. Austin, G.I.; Pe'er, I.; Korem, T. Distributional Bias Compromises Leave-One-out Cross-Validation. *Sci. Adv.* **2025**, *11*, ead6976. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Khallaf, R.; Khallaf, L.; Anumba, C.J.; Madubuike, O.C. Review of Digital Twins for Constructed Facilities. *Buildings* **2022**, *12*, 2029. [\[CrossRef\]](#)
50. de Felice, G.; Aiello, M.A.; Caggegi, C.; Ceroni, F.; De Santis, S.; Garbin, E.; Gattesco, N.; Hojdis, L.; Krajewski, P.; Kwiecień, A.; et al. Recommendation of RILEM Technical Committee 250-CSM: Test Method for Textile Reinforced Mortar to Substrate Bond Characterization. *Mater. Struct./Mater. Constr.* **2018**, *51*, 95. [\[CrossRef\]](#)
51. Gunes, M.E.; Pekmezci, B.Y.; Girgin, Z.C. Durability of Natural Hydraulic Lime (NHL)-Based TRM Composites through Hot-Water Immersion Method. *Mater. Struct.* **2021**, *54*, 24. [\[CrossRef\]](#)
52. Younis, A.; Ebead, U. Bond Characteristics of Different FRCM Systems. *Constr. Build. Mater.* **2018**, *175*, 610–620. [\[CrossRef\]](#)
53. Carloni, C.; Verre, S.; Sneed, L.H.; Ombres, L. Loading Rate Effect on the Debonding Phenomenon in Fiber Reinforced Cementitious Matrix-Concrete Joints. *Compos. B Eng.* **2017**, *108*, 301–314. [\[CrossRef\]](#)
54. Carloni, C.; D'Antino, T.; Sneed, L.H.; Pellegrino, C. Role of the Matrix Layers in the Stress-Transfer Mechanism of FRCM Composites Bonded to a Concrete Substrate. *J. Eng. Mech.* **2015**, *141*, 1–11. [\[CrossRef\]](#)
55. D'Antino, T.; Carloni, C.; Sneed, L.H.; Pellegrino, C. Matrix-Fiber Bond Behavior in PBO FRCM Composites: A Fracture Mechanics Approach. *Eng. Fract. Mech.* **2014**, *117*, 94–111. [\[CrossRef\]](#)
56. De Santis, S.; Ceroni, F.; de Felice, G.; Fagone, M.; Ghiassi, B.; Kwiecień, A.; Lignola, G.P.; Morganti, M.; Santandrea, M.; Valluzzi, M.R.; et al. Round Robin Test on Tensile and Bond Behaviour of Steel Reinforced Grout Systems. *Compos. B Eng.* **2017**, *127*, 100–120. [\[CrossRef\]](#)
57. Pei, X.; Zhou, C.; Zhang, T.; Liu, S.; Wang, J.; Luo, S.; Xu, Z. Durability Properties of Steel Textiles and Bonding Properties at the Interface of Steel Textiles with Reinforced Mortar-Reinforced Concrete Systems. *Langmuir* **2024**, *40*, 21010–21023. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Hasanshahi, O.; Salati, M.; Azmi, N. Hybrid Steel Fibers and RCA in RCC Pavements: Mechanical Recovery, Freeze–Thaw Durability, and Eco-Efficiency. *Infrastructures* **2025**, *10*, 321. [\[CrossRef\]](#)
59. Bakshi, M.; Azmi, N.; Hasanshahi, O. Effect of Strain Rate on the Energy Dissipation Capacity of Steel Fiber Reinforced Concrete Beams. *J. Build. Eng.* **2026**, 115527. [\[CrossRef\]](#)
60. Hasanshahi, O.; Karimzadeh, S.; Bakshi, M.; Azmi, N. Synergistic Role of Recycled Concrete Aggregates and Hybrid Steel Fibers in Roller-Compacted Concrete Pavements: A Multi-Criteria Assessment for Eco-Efficiency Optimization. *Buildings* **2025**, *15*, 4279. [\[CrossRef\]](#)
61. Li, B.; Xiong, H.; Jiang, J.; Dou, X. Tensile Behavior of Basalt Textile Grid Reinforced Engineering Cementitious Composite. *Compos. B Eng.* **2019**, *156*, 185–200. [\[CrossRef\]](#)
62. Contamine, R.; Si Larbi, A.; Hamelin, P. Contribution to Direct Tensile Testing of Textile Reinforced Concrete (TRC) Composites. *Mater. Sci. Eng. A* **2011**, *528*, 8589–8598. [\[CrossRef\]](#)

63. Carlo Rampini, M.; Zani, G.; Colombo, M.; di Prisco, M. Stiffness Degradation of Textile-Reinforced Mortar under Unloading–Reloading Tensile Cycles. *J. Compos. Constr.* **2023**, *27*, 04022083. [[CrossRef](#)]
64. Butler, M.; Mechtcherine, V.; Hempel, S. Durability of Textile Reinforced Concrete Made with AR Glass Fibre: Effect of the Matrix Composition. *Mater. Struct./Mater. Constr.* **2010**, *43*, 1351–1368. [[CrossRef](#)]
65. Rasmussen, C.E. Gaussian Processes in Machine Learning. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 63–71.

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.