Materials Today Communications 52 (2026) 115030

Contents lists available at ScienceDirect

Materials Today Communications

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

Joint probabilistic modeling of cement strength: Dual-strength compliance and uncertainty decomposition for risk-aware and sustainable quality control

Mazharul Islam [a], Changjiao Li [b,*], Bo Yang [b,d,**], Cong Liu [c]

[a] Shandong Provincial Key Laboratory of Green and Intelligent Building Materials, University of Jinan, Jinan, Shandong 250022, China
[b] Shandong Key Laboratory of Ubiquitous Intelligent Computing, University of Jinan, Jinan, Shandong 250022, China
[c] NOVA Information Management School, Universidade Nova de Lisboa, Campus de Campolide, Lisboa 1070-312, Portugal
[d] Quan Cheng Laboratory, University of Jinan, Jinan 250100, China

ARTICLE INFO

Keywords:
Building materials
Cement strength
Joint modeling
Uncertainty decomposition
Dual-strength compliance
Sustainable manufacturing

ABSTRACT

In cement manufacturing, ensuring simultaneous compliance with compressive (CCS) and flexural strength (FS) requirements is challenging due to their divergent responses to shared inputs. Industrial variability (e.g., fluctuations in C_3S content, grinding heterogeneity, and curing conditions) exacerbates this challenge, leading to high batch rejection (15–20%) and excessive clinker overdesign that contributes 7% of global CO_2 emissions. Traditional single-property models fail by treating CCS and FS in isolation, producing deterministic predictions that ignore uncertainty, and conflating material-driven (aleatoric) with model-induced (epistemic) uncertainty—precluding risk-aware decisions. To overcome these limitations, this study introduces the Multi-Property Cement Strength Estimator (MPCSE)—the first joint probabilistic framework to model CCS and FS as full distributions via multi-head Gaussian Mixture Models with shared latent representations. MPCSE delivers four innovations: (1) joint probabilistic modeling that captures cross-property dependencies (7.8% lower MAE for CCS and 9.3% for FS versus single-property baselines); (2) explicit uncertainty decomposition enabling targeted process control (e.g., tighter grinding for 36–52 μm particles reduces FS variability by 15%) and strategic data collection (e.g., low-C_3A regimes, $< 4\%$); (3) a novel dual-strength compliance probability metric, $P(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$, revealing a near-total collapse to 0.38% at 50 MPa CCS / 9 MPa FS in industrial settings (versus 11.2% under controlled laboratory conditions) and quantifying the previously unmeasured risk of simultaneous failure; (4) context-dependent mechanistic interpretability via SHAP analysis, showing granulometry (uniformity index, SHAP = 0.032) dominates industrial variability while stoichiometry (C_3S, SHAP = 0.058) governs laboratory performance—explaining the CCS–FS correlation gap (lab $r = 0.945$ vs. industry $r = 0.101$). MPCSE achieves $> 95\%$ confidence-interval coverage on industrial data and enables 12–15% clinker reduction with approximately 25% fewer batch rejections through risk-aware rerouting of borderline batches, advancing sustainable cement production.

1. Introduction

Cement, the backbone of global infrastructure with an annual production exceeding 4.1 billion tonnes, faces significant challenges in both sustainability and performance. Variability in strength, driven by inconsistent material properties and manufacturing conditions, results in considerable financial and environmental costs [1]. Recent reliability-based analyses of flexural strength in hybrid-reinforced concrete beams [2], machine learning predictions of dynamic compressive response in concrete-like materials [3], deep learning models for plastic damage prediction under compression [4], finite-element and deep-learning approaches for bond stress-slip behavior [5], and high-fidelity mesoscale modelling algorithms for concrete [6] have all highlighted the persistent impact of material and process variability. Industrial data indicate that approximately 15–20% of cement batches are rejected or downgraded [7], largely due to

* Corresponding author.
** Corresponding author at: Shandong Key Laboratory of Ubiquitous Intelligent Computing, University of Jinan, Jinan, Shandong 250022, China.
E-mail addresses: 202014100001@stu.ujn.edu.cn (M. Islam), ise_licj@ujn.edu.cn (C. Li), yangbo@ujn.edu.cn (B. Yang), cliu@novaims.unl.pt (C. Liu).

https://doi.org/10.1016/j.mtcomm.2026.115030
Received 3 February 2026; Received in revised form 11 March 2026; Accepted 19 March 2026
Available online 20 March 2026

unexpected imbalances between compressive and flexural strength, leading to material waste and increased CO_2 emissions. Cement production alone contributes 7% of global CO_2 emissions, with much of the excess emissions stemming from overdesign practices aimed at ensuring compliance with strength requirements [1,6–8].

The mechanical performance of cement-based materials is primarily governed by CCS, which determines load-bearing capacity, and FS, which influences crack resistance and durability in applications such as pavements and structural elements [9,10]. The challenge arises from the divergent responses of compressive and flexural strengths to shared inputs, such as C_3S content, particle-size distribution, and curing conditions. For example, elevated tricalcium silicate (C_3S) accelerates calcium silicate hydrate ($C-S-H$) formation, enhancing CCS, but concurrently coarsens pore structure and destabilizes ettringite networks—critical for crack resistance—thereby compromising FS [11,12]. Similarly, particle size distribution governs both packing density (benefiting CCS) and interfacial transition zone homogeneity (critical for FS), yet industrial grinding variability introduces heterogeneity that decouples these properties. Under controlled laboratory conditions (Fig. 3), CCS and FS exhibit a strong correlation ($r = 0.945$); in real-world production (Fig. 2), this collapses to near-zero ($r = 0.101$), rendering lab-optimized mixes unreliable at scale. Recent advances in active structural control for construction stability [10], acoustic-emission multifractal analysis of freeze-thaw damage [13], risk-based optimization for structural maintenance and replacement [14], multiscale indices for evaluating material discontinuities [15], and quantum interval neural networks for uncertain structural static analysis [16] further underscore the need for integrated, uncertainty-aware frameworks to address such divergent property responses in cementitious and related structural systems.

Regulatory standards intensify this dilemma. Currently, structural specifications for cement-based materials, particularly those used in pavements, precast elements, and thin-shell structures, demand compliance with both CCS and FS criteria (e.g., 40 MPa CCS and 6 MPa FS). Industry standards, such as ASTM C39 and C78, provide separate testing methods for these properties. Still, regulations like EN 206–1 require that all specified properties, including both CCS and FS, be satisfied for the material to be deemed compliant [17,18]. This necessitates a joint probability approach for evaluating compliance, as traditional models that predict individual strength properties fail to assess the risk of meeting both simultaneously. The key metric is not simply the probability of meeting CCS or FS thresholds independently, but the joint probability $P(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$, which captures the likelihood of meeting both standards under real-world variability.

Traditional predictive models, including early empirical laws such as Abram's Law [19], as well as modern machine learning (ML) approaches [20–22], are not well-equipped to address this duality. These models typically treat CCS and FS as independent outputs, despite their shared physicochemical drivers, and produce deterministic predictions that do not reflect the inherent uncertainty in industrial production. Given the variability in raw materials (e.g., ± 8% C_3S), process conditions (e.g., grinding and curing inconsistencies), and material properties, deterministic point estimates like "CCS = 50 MPa" fail to capture the true variability of cement strength in production environments. These models are unable to answer critical quality control questions such as: *What is the probability that this batch will meet both CCS and FS requirements?*

To bridge this gap, this study transitions from *deterministic predictions to probabilistic modeling*, representing cement strength as full probability distributions rather than single-point estimates. This shift allows us to quantify the range of possible outcomes under real-world conditions, incorporating variability from both material and process sources. For instance, instead of predicting a point estimate like "CCS = 45 MPa", a probabilistic model predicts a range such as "CCS is likely between 42–48 MPa with 95% confidence", accounting for fluctuations in clinker composition and grinding conditions. This approach transforms strength prediction into a tool for risk-aware compliance assessment rather than a simple pass/fail judgment, enabling manufacturers to better manage uncertainties and improve production quality.

Moreover, since CCS and FS are interdependent, predicting their joint distribution is essential. These properties share common influences—like C_3S content and particle size distribution—but respond differently. For instance, even if the individual probabilities of meeting CCS and FS requirements are high (e.g., $P(\mathrm{CCS} \geq c_{\min}) = 0.9$ and $P(\mathrm{FS} \geq f_{\min}) = 0.9$), their joint probability $P(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$ can be much lower (e.g., <0.5) due to the weak correlation between these properties in some production scenarios. While few multi-output models attempt to predict both properties simultaneously [23,24], these models fail to model the joint probability distribution, limiting the ability to assess compliance risk comprehensively.

Additionally, existing methods conflate two distinct uncertainty sources [25,26]: *aleatoric* (irreducible material/process noise, e.g., clinker heterogeneity) and *epistemic* (reducible model gaps, e.g., lack of experiments on low-C_3A formulations) [27]. Without decomposing these uncertainties, manufacturers typically adopt a conservative overdesign approach, increasing clinker by 15–20%—to hedge against unknown risks, directly exacerbating CO_2 emissions. Recent advances in multi-task learning (MTL) and uncertainty quantification (UQ) offer pathways to address these gaps [28,29], *yet no framework integrates MTL and UQ to predict the joint density of CCS and FS while linking uncertainty decomposition to risk-aware compliance.*

This study addresses these gaps through four fundamental materials-science principles: *(i)* joint property modeling to capture co-variability from shared physicochemical drivers (e.g., C_3S, particle packing) missed by decoupled approaches; *(ii)* explicit uncertainty decomposition to separate aleatoric (irreducible material/process noise) from epistemic (reducible model/data gaps) uncertainty, enabling targeted interventions instead of conservative overdesign; *(iii)* explicit uncertainty decomposition to separate aleatoric (irreducible material/process noise) from epistemic (reducible model/data gaps) uncertainty, enabling targeted interventions instead of conservative overdesign; and *(iv)* context-dependent mechanisms that adapt to scale-specific drivers—granulometry dominating industrial production versus stoichiometry in controlled laboratory settings—bridging the persistent lab-to-field performance discrepancy.

Accordingly, this study proposes the Multi-Property Cement Strength Estimator, which, to the best of the authors' knowledge, is *the first joint probabilistic framework* to model CCS and FS as full distributions via multi-head Gaussian Mixture Models (GMMs) with shared latent representations. A variational layer explicitly decomposes uncertainty, enabling: (a) risk-aware compliance assessment via the novel dual-strength metric $P(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$; (b) actionable process guidance (e.g., tighter grinding for 36–52 μm particles reduces FS variability by 15%); (c) strategic data collection (e.g., low-C_3A < 4% regimes); and (d) 12–15% clinker reduction with ~25% fewer batch rejections through risk-aware rerouting of borderline batches.

1.1. Contributions

This study advances sustainable cement manufacturing through four materials-focused innovations:

1. *Joint strength distribution modeling*: First framework to predict the joint probability densities of CCS and FS via multi-head GMMs with shared latent representations, reducing MAE by 7.8% (CCS) and 9.3% (FS) versus single-property models by explicitly modeling their co-variability.
2. *Actionable uncertainty decomposition*: Isolates aleatoric (material/process noise) and epistemic (model/data gaps) uncertainties to enable targeted interventions—such as tighter grinding control for 36–52 μm particles and strategic data collection in low-C_3A

regimes—directly supporting clinker reduction and process optimization.

3. *Dual-strength compliance metric*: Introduces $P(\text{CCS} \geq c_{\min} \cap \text{FS} \geq f_{\min})$ as the first dual-strength compliance metric, revealing near-total collapse to 0.38% at 50 MPa CCS / 9 MPa FS in industrial settings (versus 11.2% in laboratories)—quantitatively explaining why batches meeting individual strength requirements still fail simultaneously under real-world variability.
4. *Mechanistic interpretability*: SHAP analysis uncovers context-dependent drivers—granulometry dominates industrially (uniformity index, SHAP = 0.032), while stoichiometry governs in labs (C_3S, SHAP = 0.058)—providing the first mechanistic explanation for why laboratory-optimized formulations fail in production.

2. Relevant studies

This section critically examines advances in cement strength modeling through the lens of industrial applicability and sustainable manufacturing. Three research streams are critically examined: *single-property models*, *multi-property frameworks*, and *uncertainty-aware interpretability*. Despite growing interest in data-driven methods, current approaches remain inadequate for managing the dual-strength compliance challenge under real-world variability. This study identifies three persistent gaps across existing work: *(i)* the treatment of compressive and flexural strength as isolated properties, ignoring their shared physicochemical origins; *(ii)* reliance on deterministic predictions that obscure material-driven uncertainty and hinder risk-aware decision-making; and *(iii)* a lack of actionable guidance for process optimization or clinker reduction.

2.1. Single-property models: overlooking shared hydration mechanisms

Early machine learning efforts in cement science focused on predicting individual strength properties [20,30,31], achieving moderate success but inherently overlooking the physicochemical coupling between compressive and flexural strength. For example, Chou and Pham [32] achieved high CCS accuracy (R^2 = 0.89) using ANNs but treated FS as irrelevant, neglecting how the balance of sulfates governs both the stability of ettringite (critical for FS) and the formation of C-S-H (for CCS). Similarly, Qing and Li [33] applied XGBoost to predict both properties separately ($R^2 \approx$ 0.91–0.92), yet failed to model their co-variability under C_3A/SO_3 imbalance—a known cause of FS degradation in sulfate-limited systems [34]. Sadrossadat et al. [35] validated Gaussian process regression for CCS prediction ($R^2 \approx 0.96$, RMSE = 3.66 MPa), yet neglected FS entirely, reinforcing the practice of isolated quality assessment.

More recent work incorporates granulometric features, such as particle size distribution [36], improving fidelity but still collapsing inherent process noise into single-point estimates. Meddage et al. [37] used explainable ML for graphene-modified concrete but offered no mechanism to distinguish whether prediction errors stem from grinding inconsistencies (aleatoric) or sparse data (epistemic). Consequently, these models remain reactive tools for post-hoc assessment, not proactive instruments for sustainable process control.

2.2. Multi-property frameworks: prediction without risk awareness

Recent multi-output models [38–42] aim to unify strength prediction, but none jointly model both CCS and FS as probabilistic distributions. Najmoddin et al. [8] used multi-output XGBoost to predict CCS, FS, and tensile strength simultaneously (R^2 = 0.94), yet provided no uncertainty decomposition or compliance risk metric. Onyelowe et al. [23] applied physics-informed ML to recycled aggregate concrete but focused solely on tensile strength, ignoring the CCS–FS interplay central to structural specifications.

Bayesian methods offer partial progress: Imam et al. [24] generated ±1.8 MPa uncertainty bounds for CCS using BNNs, but these undecomposed intervals cannot guide targeted interventions (e.g., "deciding whether to tighten grinding or collect more low-C_3A data?"). Melkumyan et al. [43] introduced multi-task GPs for slump and CCS but omitted FS entirely. Crucially, no existing framework computes the joint compliance probability $P(\text{CCS} \geq c_{\min} \cap \text{FS} \geq f_{\min})$—the very metric that determines batch acceptance under EN 206–1. Without this, models cannot support the risk-aware rerouting of borderline batches to non-structural uses, missing a key opportunity to cut rejection rates by ~25% and clinker use by 12–15%.

2.3. Uncertainty quantification and interpretability: from statistical artifacts to actionable intelligence

While UQ [44–46] and explainable AI (XAI) [47–49] are emerging in cement informatics, their integration remains superficial. Tamuly et al. [25] used conformal prediction to quantify CCS uncertainty but did not decompose it into aleatoric (material/process noise) and epistemic (data gaps) components—essential for distinguishing between "unavoidable variability" and "fixable model limitations". Han et al. [50] achieved high CCS accuracy (RMSE = 1.24 MPa) but ignored FS and joint behavior, limiting relevance to dual-strength specifications.

Interpretability efforts similarly lack industrial grounding. Sun et al. [51] identified C_3S as a key CCS driver via SHAP but did not contrast industrial vs. laboratory contexts or link findings to process actions (e.g., "Optimize SO_3 to 2.3% for ettringite stability"). Long et al. [52] focused only on slump and CCS, omitting FS—the property most sensitive to granulometric noise in production.

In contrast, the proposed MPCSE bridges these gaps by: *(i)* modeling CCS and FS as a coupled system governed by shared drivers (e.g., C_3S, PSD), *(ii)* decomposing uncertainty to enable targeted process control (e.g., tighter grinding for 41–52 μm particles) and strategic data collection (e.g., low-C_3A regimes), *(iii)* quantifying dual-strength compliance risk to replace binary pass/fail decisions with resource-efficient, risk-informed allocation. This transforms predictive modeling from a statistical exercise into a sustainability-enabling tool for precision cement manufacturing.

3. Dataset acquisition, and statistical analysis

This study leverages two complementary datasets—industrial and laboratory—to rigorously evaluate the proposed *Multi-Property Cement Strength Estimator* under contrasting conditions of variability and control. This dual-dataset design directly addresses the central challenge of sustainable cement production: bridging the gap between lab-optimized formulations and real-world performance, where unquantified variability drives excessive clinker use and batch rejection.

3.1. Cement specimen preparation and strength measurement

All specimens were prepared in strict accordance with ASTM C39, EN 206–1, and GB/T 17671–2021 standards [53], ensuring industrial relevance and cross-dataset comparability. The protocol involved mechanical homogenization of cement, water, and standardized sand (1:2:6 mass ratio), vibration-assisted compaction, and a two-stage curing process: initial setting at 20°C for 24 h, followed by 27 days of immersion under controlled conditions. By applying identical preparation and testing protocols to both the industrial (M = 163) and laboratory (M = 50) specimens, procedural variability was eliminated, isolating the effects of intrinsic factors—such as C_3S content, particle size distribution, and curing fluctuations—on compressive and flexural strength. This standardized approach ensures that strength measurements reflect true material behavior, enabling direct comparison of microstructure–property relationships across scales. Subsequent Pearson correlation analyses (Figs. 2 and 3) quantify how these intrinsic

factors govern strength outcomes under differing levels of process control.

3.2. Industrial dataset: capturing real-world variability

3.2.1. Analysis of key statistics

The industrial dataset, sourced from an active cement plant, captures the natural variability inherent to large-scale production—a primary driver of clinker overuse and CO_2 emissions. As shown in Fig. 1, inline laser particle analyzers monitored 22 granulometric properties (e.g., PSD, SSA), while chemical assays captured key clinker phases. Despite missing features due to assay costs, the dataset exhibits balanced diversity in critical inputs: particle size distributions span industrially relevant ranges (e.g., <3 μm: 9.53–14.68%), and surface area (S/g: 322.83–418.27 m^2/kg) reflects typical grinding heterogeneity. Crucially, target variables show high measurement consistency—CCS (49.2–56.2 MPa, st. dev. 1.21) and FS (8.0–9.0 MPa, st. dev. 0.20)—ensuring that observed model challenges stem from genuine process noise, not measurement error.

This dataset validates MPCSE's suitability for joint modeling: *(a)* physicochemical gradients (e.g., 45–65 μm fraction: variance 1.48) correlate with C–S–H formation mechanisms; *(b)* granulometric parameters (uniformity index: 0.70–0.94) govern particle packing efficiency—a shared driver of CCS and FS. Table 1 provides full descriptive statistics, confirming the dataset's capacity to test unified density estimation under real-world uncertainty.

3.2.2. Pearson correlation analysis: decoupling strength in production

The industrial correlation matrix (Fig. 2) reveals a stark decoupling of CCS and FS ($r = 0.101$), underscoring why lab-optimized mixes fail in practice. Granulometric features dominate: fine fractions (<3 μm, <10 μm) strongly correlate with specific surface area ($r = 0.936$–0.971), governing hydration kinetics, while coarse particles (>45 μm) show weak negative links to CCS ($r = -0.092$) and negligible FS ties ($r = 0.016$), reflecting their detrimental impact on reactivity. Critically, process proxies like *shading* and *illuminance* exhibit conflicting signals—positive with CCS but negative with FS—highlighting how environmental artifacts obscure material-property relationships. This complexity necessitates MPCSE's joint modeling and uncertainty decomposition to disentangle true drivers from noise.

3.3. Laboratory dataset: controlled validation of mechanistic drivers

3.3.1. Analysis of key statistics

The laboratory dataset ($M = 50$) provides a tightly controlled benchmark to validate MPCSE's ability to capture fundamental cement chemistry. Key clinker phases span scientifically meaningful ranges: C_3S (42.14–73.91%, variance 34.74) and C_3A (4.68–9.15%, variance 1.37) cover stoichiometric extremes governing early/late strength trade-offs, while sulfate balance (SO_3: 2.10–2.40%) reflects precise alkali-sulfate optimization. Target variables combine intentional diversity (CCS: 26.0–61.8 MPa) with high precision (FS st. dev.: 0.47), enabling robust evaluation of multi-property generalization. Granulometric (SSA: 341–360 m^2/kg) and chemical gradients (C_4AF: 6.40–12.68%, SiO_2: 2.31–22.81%) rigorously test whether joint modeling preserves microstructure–property linkages across formulation extremes. Table 2 details full statistics, confirming its suitability for validating unified density estimation under idealized conditions.

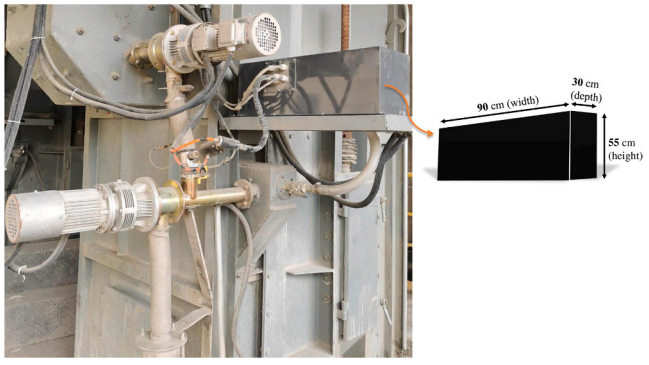

Fig. 1. Inline integration of a laser particle size analyzer into an operating cement factory for continuous real-time monitoring of particle size distribution (PSD)—a key determinant of cement fineness, hydration reactivity, and final mechanical strength development. The main analyzer housing (black rectangular box, indicated by arrow) measures exactly 90 cm (width) × 55 cm (height) × 30 cm (depth).

3.3.2. Pearson correlation analysis: coupled strength in control

In stark contrast to industrial data, the laboratory correlation matrix (Fig. 3) shows near-perfect CCS–FS coupling ($r = 0.945$), confirming their shared dependence on stoichiometry and hydration kinetics. Chemical drivers dominate: C_3S strongly enhances CCS ($r = 0.86$), while sulfate balance (SO_3) stabilizes ettringite networks critical for FS. Antagonistic phase interactions—C_3S vs. C_2S ($r = -0.965$)—mirror known compositional trade-offs, providing a rigorous testbed for MPCSE's ability to disentangle competing effects. This controlled environment validates that MPCSE's architecture correctly captures fundamental cement science, while the industrial dataset tests its resilience to real-world noise.

4. Research methodology

The Multi-Property Cement Strength Estimator addresses the core challenge of sustainable cement production: managing the compressive–flexural strength trade-off under real-world variability. Rather than treating compressive and flexural strength as isolated outputs, MPCSE models their joint behavior through a unified framework that integrates fundamental cement chemistry with probabilistic machine learning. As shown in Fig. 4, the architecture comprises three synergistic modules designed to mirror the physical reality of cement hydration: *(i)* hierarchical encoding of physicochemical drivers, *(ii)* explicit decomposition of material- and model-induced uncertainty, and *(iii)* task-specific estimation of full-strength distributions via multi-head Gaussian Mixture Models (GMMs).

The framework begins by encoding key hydration determinants—such as C_3S content, sulfate balance, particle size distribution, and curing conditions—into high-level representations that govern microstructural evolution. These features are processed through a shared latent space $\boldsymbol{z}$, which captures common mechanisms (e.g., C–S–H nucleation, ettringite stability) that jointly influence CCS and FS. Crucially, a variational inference layer then decomposes total predictive uncertainty into two actionable sources: *aleatoric uncertainty*, arising from irreducible material and process noise, and *epistemic uncertainty*, reflecting model limitations due to sparse data or extrapolation. From this disentangled representation, two independent GMM heads generate full probability distributions for CCS and FS. This design preserves both cross-property dependencies (e.g., shared sensitivity to sulfate balance) and property-specific responses (e.g., FS's heightened vulnerability to pore coarsening). The detailed training procedure is described in Algorithm 1.

Algorithm 1. Training process of MPCSE

Input: Dataset $\mathcal{D} = \{(\boldsymbol{x_i}, s_{\mathrm{CCS},i}, s_{\mathrm{FS},i})\}_{i=1}^{M}$, Batch size B, KL weight λ, Total epochs $T = 500$, Gaussian components ($N_{\mathrm{CCS}} = 2, N_{\mathrm{FS}} = 2$), Learning rate $\eta = 0.01$

Output: Conditional PDFs $g_{\mathrm{CCS}}(s_{\mathrm{CCS}}|\boldsymbol{x})$ and $g_{\mathrm{FS}}(s_{\mathrm{FS}}|\boldsymbol{x})$

Initialize parameters $\theta = \{\theta_{\mathrm{shared}}, \theta_{\mathrm{CCS}}, \theta_{\mathrm{FS}}\}$;

Initialize Adam optimizer with η; Apply gradient clipping $|\nabla\theta| \leq 1.0$; Set step decay scheduler ($\gamma = 0.1$ every 100 epochs);

for epoch $t \leftarrow 1$ **to** T **do**

  for each batch $(\boldsymbol{x_i}, s_{CCS}, s_{FS})_{i=1}^{B}$ in $\mathcal{D}$ **do**

    Feature Extraction: $\boldsymbol{h_L} = \mathrm{ResMLP}(\boldsymbol{x_i})$

    Variational Sampling: $\mu, \log\sigma^2 = \mathrm{Linear}(\boldsymbol{h_L})$; $\boldsymbol{z_i} = \mu + \epsilon \odot \exp(0.5\log\sigma^2)$, $\epsilon \sim \boldsymbol{\mathcal{N}}(0, \boldsymbol{I})$

    Task-Specific GMM Prediction:

    $\{\alpha_{\mathrm{CCS},j}, \mu_{\mathrm{CCS},j}, \sigma^2_{\mathrm{CCS},j}\}_{j=1}^{2} = \mathrm{GMM}_{\mathrm{CCS}}(\boldsymbol{z_i})$

    $\{\alpha_{\mathrm{FS},j}, \mu_{\mathrm{FS},j}, \sigma^2_{\mathrm{FS},j}\}_{j=1}^{2} = \mathrm{GMM}_{\mathrm{FS}}(\boldsymbol{z_i})$

    Loss Calculation:

    $\boldsymbol{\mathcal{L}}_{NLL}^{CCS} = -\frac{1}{B}\sum_{i=1}^{B}\log\sum_{j=1}^{2}\alpha_{\mathrm{CCS},j}\mathcal{N}\left(s_{\mathrm{CCS},i}; \mu_{\mathrm{CCS},j}, \sigma^2_{\mathrm{CCS},j}\right)$

    $\boldsymbol{\mathcal{L}}_{NLL}^{FS} = -\frac{1}{B}\sum_{i=1}^{B}\log\sum_{j=1}^{2}\alpha_{\mathrm{FS},j}\mathcal{N}\left(s_{\mathrm{FS},i}; \mu_{\mathrm{FS},j}, \sigma^2_{\mathrm{FS},j}\right)$

    $\boldsymbol{\mathcal{L}}_{\mathrm{KL}} = \frac{1}{2B}\sum_{i=1}^{B}\sum_{d=1}^{n_z}\left(\mu_{i,d}^2 + \exp\left(\log\sigma_{i,d}^2\right) - \log\sigma_{i,d}^2 - 1\right)$

    $\boldsymbol{\mathcal{L}}_{total} = \boldsymbol{\mathcal{L}}_{NLL}^{CCS} + \boldsymbol{\mathcal{L}}_{NLL}^{FS} + \lambda\boldsymbol{\mathcal{L}}_{KL}$

    Backpropagation: $\theta \leftarrow \theta - \eta\nabla_\theta\boldsymbol{\mathcal{L}}_{total}$

  end **for**

  if $t \; mod \; 100 = 0$ **then**

    $\eta \leftarrow \eta \times \gamma$

  end **if**

end **for**

Return: Trained MPCSE model for joint CCS/FS density estimation.

4.1. Task definition

MPCSE models the conditional probability distributions of CCS and FS given input features $\boldsymbol{x}$. Formally, it operates on a dataset $\mathscr{D} = \{(\boldsymbol{x}_i, s_{\mathrm{CCS},i}, s_{\mathrm{FS},i})\}_{i=1}^{M}$, where:

- $\boldsymbol{x}_i \in \mathbb{R}^d$ encodes *chemical composition* (C_3S, C_3A, SO_3), *physical properties* (particle size distribution, specific surface area), and *curing conditions*;
- $s_{\mathrm{CCS},i}, s_{\mathrm{FS},i} \in \mathbb{R}_{>0}$ are experimentally measured 28-day strengths (MPa).

The objective is to estimate the joint behavior of CCS and FS by approximating their conditional densities while preserving interdependencies. This is achieved using task-specific GMMs:

$$p(s_k|\boldsymbol{x}) \approx g_k(s_k|\boldsymbol{x};\boldsymbol{\Theta}) = \sum_{j=1}^{N_k}\alpha_{kj}(\boldsymbol{z})\cdot\mathcal{N}\left(s_k; \mu_{kj}(\boldsymbol{z}), \sigma_{kj}^2(\boldsymbol{z})\right), \quad (1)$$

where $k \in \{\mathrm{CCS}, \mathrm{FS}\}$, $\boldsymbol{z} = f(\boldsymbol{x}; \boldsymbol{\Theta}_{\mathrm{shared}})$ is the shared latent representation, and N_k is the number of Gaussian components. The full parameter set $\boldsymbol{\Theta} = \{\boldsymbol{\Theta}_{\mathrm{shared}}, \boldsymbol{\Theta}_{\mathrm{CCS}}, \boldsymbol{\Theta}_{\mathrm{FS}}\}$ includes weights for shared feature extraction, variational uncertainty decomposition, and task-specific distribution heads.

Table 1
Descriptive statistics of the industrial cement dataset(M=163) encompassing particle size distribution (PSD)parameters, physical properties, and mechanical strength outcomes.

Property	Mean	St. Error	Median	Mode	St. dev.	Variance	Min	Max	Range
$< 1\ \mu$m	1.81	0.01	1.78	1.37	0.17	0.03	1.37	2.77	1.40
$< 3\ \mu$m	11.93	0.07	11.88	9.53	0.91	0.82	9.53	14.68	5.15
$< 10\ \mu$m	40.79	0.21	40.37	35.79	2.64	6.98	35.79	46.65	10.86
$< 16\ \mu$m	58.74	0.25	58.31	52.81	3.15	9.91	52.81	67.59	14.78
$< 24\ \mu$m	75.17	0.24	74.68	69.46	3.08	9.46	69.46	85.75	16.28
$< 32\ \mu$m	84.85	0.20	84.26	79.64	2.61	6.82	79.64	94.41	14.77
$< 65\ \mu$m	98.82	0.04	98.84	97.51	0.47	0.22	97.51	99.77	2.26
$< 70\ \mu$m	99.35	0.02	99.40	98.45	0.29	0.08	98.45	99.81	1.36
$> 45\ \mu$m	4.19	0.08	4.37	0.45	1.06	1.12	0.45	6.50	6.05
$> 80\ \mu$m	0.28	0.01	0.25	0.10	0.13	0.02	0.10	0.70	0.60
3–16 μm	46.82	0.19	46.62	41.68	2.48	6.14	41.68	54.31	12.62
3–32 μm	72.92	0.16	72.63	68.86	2.03	4.10	68.86	80.13	11.26
16–24 μm	16.43	0.06	16.61	14.78	0.77	0.60	14.78	18.24	3.46
32–45 μm	8.55	0.08	8.70	4.85	1.03	1.06	4.85	10.33	5.48
45–65 μm	5.45	0.10	5.76	0.74	1.21	1.48	0.74	7.75	7.01
65–80 μm	1.05	0.03	1.10	0.01	0.40	0.16	0.01	2.03	2.02
S/g	364.10	1.52	362.81	322.83	19.35	374.51	322.83	418.27	95.45
D_{50}	13.00	0.07	13.11	11.00	0.93	0.87	11.00	14.94	3.94
Uniformity index	0.84	0.00	0.83	0.80	0.05	0.00	0.70	0.94	0.24
Distribution width	2.76	0.01	2.74	2.61	0.16	0.02	2.28	3.14	0.86
Shading	24.98	0.34	23.30	17.10	4.34	18.87	17.10	36.03	18.93
Illuminance	766.25	8.38	709.68	668.85	107.05	11459.13	668.85	1056.01	387.16
CCS (MPa)	**51.66**	**0.09**	**51.50**	**50.40**	**1.21**	**1.47**	**49.20**	**56.20**	**7.00**
FS (MPa)	**8.39**	**0.02**	**8.40**	**8.50**	**0.20**	**0.04**	**8.00**	**9.00**	**1.00**

By modeling strength as full probability distributions-not point estimates—MPCSE enables risk-aware decision-making. Manufacturers can quantify the likelihood of simultaneous compliance with dual-strength standards, replace conservative overdesign with precision batching, and directly support the sustainability goals of low-carbon cement production.

4.2. Feature extractor

The feature extractor module encodes key hydration determinants—such as C_3S content, sulfate balance, particle size distribution, and curing conditions—into high-level representations that govern microstructural evolution. These inputs exhibit non-additive, time-dependent interactions: early-age C_3S reactivity influences late-stage C_2S contribution, while ultrafine particles (<3 μm) simultaneously enhance packing density and increase water demand. To capture these long-range dependencies without vanishing gradients, a residual multilayer perceptron (MLP) with skip connections is employed—a choice motivated by cement science, not architectural novelty.

Given input features $\boldsymbol{x} \in \mathbb{R}^d$, the L-layer network computes:

$$\boldsymbol{h}_1 = \phi(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1),$$

$$\boldsymbol{h}_2 = \phi(\boldsymbol{W}_2\boldsymbol{h}_1 + \boldsymbol{b}_2) + \boldsymbol{h}_1,$$

$$\vdots$$

$$\boldsymbol{h}_L = \phi(\boldsymbol{W}_L\boldsymbol{h}_{L-1} + \boldsymbol{b}_L) + \boldsymbol{h}_{L-1}, \text{ for } l = 2, \ldots, L. \tag{2}$$

where ϕ is LeakyReLU ($\alpha = 0.01$), and weights are initialized via Xavier uniform. The output $\boldsymbol{h}_L \in \mathbb{R}^{n_L}$ embeds both direct effects (e.g., C_3S-driven C–S–H nucleation) and indirect couplings (e.g., sulfate-mediated ettringite stability), ensuring downstream uncertainty decomposition operates on "physically grounded features", not statistical artifacts.

4.3. Variational layer

The variational layer resolves a critical industrial dilemma: distinguishing between "irreducible material noise" (e.g., ±8% C_3S fluctuations, grinding heterogeneity) and "reducible model gaps" (e.g., sparse data in low-C_3A regimes). This separation is essential for targeted interventions—tighter process control versus strategic experimentation.

Formally, the layer maps $\boldsymbol{h}_L$ to a shared probabilistic latent space $\boldsymbol{z}$:

$$\boldsymbol{\mu}_z = \boldsymbol{W}_\mu\boldsymbol{h}_L + \boldsymbol{b}_\mu, \ \log\boldsymbol{\sigma}_z^2 = \boldsymbol{W}_\sigma\boldsymbol{h}_L + \boldsymbol{b}_\sigma,$$

followed by reparametrized sampling:

$$\boldsymbol{z} = \boldsymbol{\mu}_z + \boldsymbol{\epsilon} \odot \exp\left(\frac{1}{2}\log\boldsymbol{\sigma}_z^2\right), \ \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}). \tag{3}$$

Here, $\boldsymbol{\sigma}_z^2$ quantifies aleatoric uncertainty (material/process noise), while $\boldsymbol{\mu}_z$ captures epistemic uncertainty (model/data limitations). This disentanglement transforms uncertainty from a passive byproduct into actionable intelligence: aleatoric hotspots guide grinding optimization, while epistemic boundaries prioritize data collection.

A crucial conceptual point arises from this formulation. While the subsequent multi-head Gaussian Mixture Models (Section 4.4) parameterize separate distributions for CCS and FS to maintain property-specific flexibility and computational tractability, the shared latent representation $\boldsymbol{z}$ encodes common drivers (e.g., C_3S content, sulfate balance, particle packing), thereby inducing cross-property dependence and capturing covariance in the joint density. The model does not explicitly parameterize a full multivariate GMM; instead, it approximates the joint distribution as:

$$p(\text{CCS}, \text{FS}|\boldsymbol{x}) = \int p(\text{CCS}|\boldsymbol{z})p(\text{FS}|\boldsymbol{z})p(\boldsymbol{z}|\boldsymbol{x})d\boldsymbol{z}, \tag{4}$$

which assumes conditional independence given the latent state $\boldsymbol{z}$—a standard and well-justified approximation in deep latent variable models for multi-output regression. This design choice balances property-specific expressiveness with computational efficiency and is empirically validated by the framework's ability to reproduce observed correlations. On the laboratory dataset, where CCS and FS are strongly coupled, the model achieves a predicted correlation of $r = 0.854 \pm 0.062$ compared to the observed $r = 0.941 \pm 0.040$. On the industrial dataset, where the properties are nearly independent due to process variability, the model correctly captures the weak correlation with predicted $r = -0.065 \pm 0.071$ versus observed $r = 0.071 \pm 0.157$. These results, together with > 95% confidence interval coverage, demonstrate robust capture of co-variability under both controlled and industrial

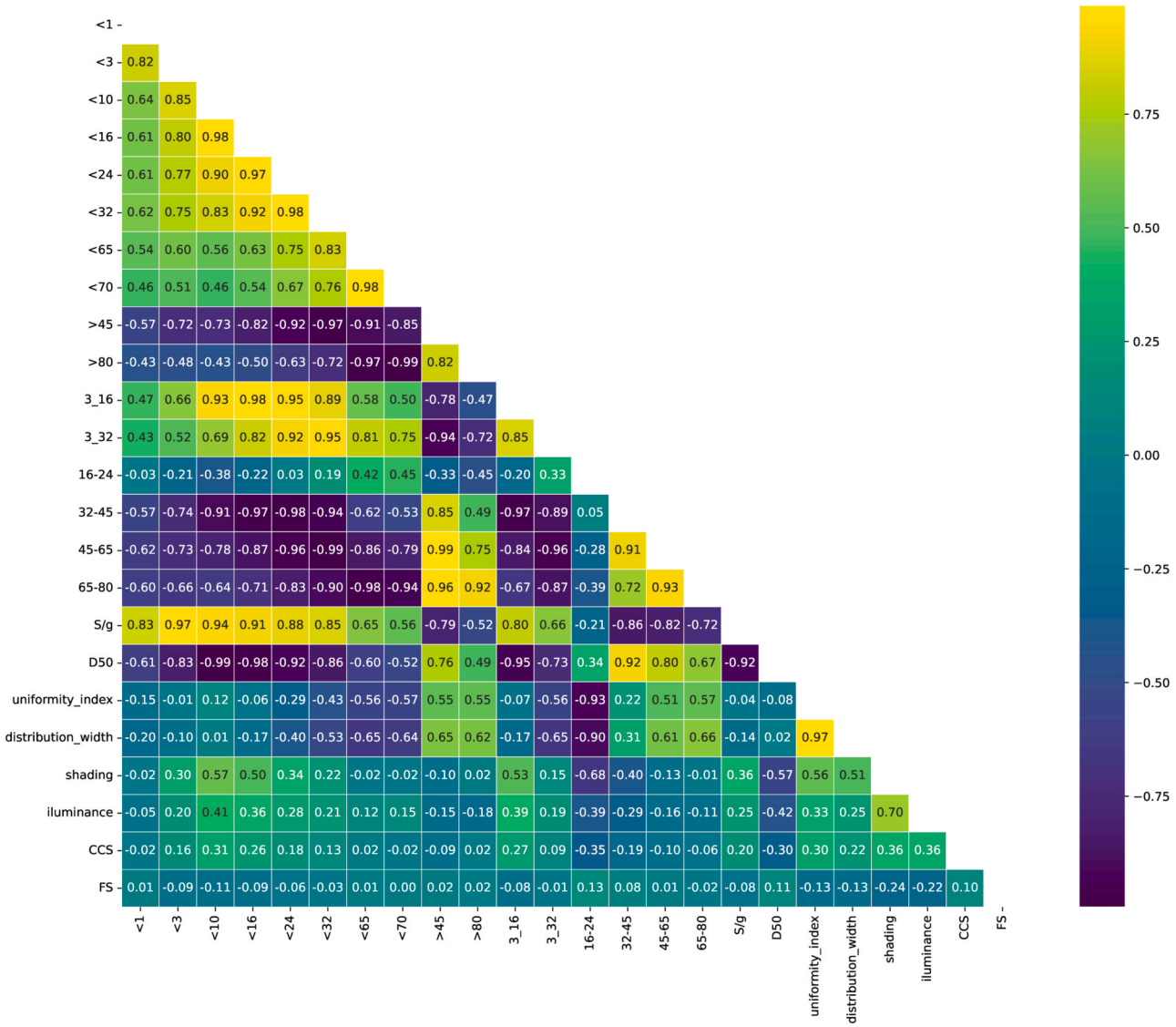

Fig. 2. Pearson correlation matrix for the industrial cement dataset ($M = 163$). Values range from − 1–1, where positive coefficients (green) indicate direct relationships and negative coefficients (blue) indicate inverse relationships. The near-zero correlation between CCS and FS ($r = 0.101$) reveals minimal coupling under real-world production conditions-contrasting sharply with laboratory settings (Fig. 3, $r = 0.945$).

Table 2
Statistical summary of the laboratory dataset (M=50) encompassing chemical composition, mineralogical phases, and mechanical performance.

Property	Mean	St. Error	Median	Mode	St. dev.	Variance	Min	Max	Range
Insoluble Residue	0.13	0.01	0.13	0.10	0.04	0.00	0.07	0.22	0.15
Alkali	0.67	0.02	0.68	0.70	0.16	0.03	0.38	1.11	0.73
LOI	0.60	0.07	0.45	0.12	0.51	0.26	0.08	2.39	2.31
SO_3	2.21	0.01	2.20	2.10	0.10	0.01	2.10	2.40	0.30
SSA (cm^2/g)	350.79	0.81	350.50	360.00	5.86	34.37	341.00	360.00	19.00
Sieve Residue	2.88	0.07	2.80	2.80	0.49	0.24	2.00	3.60	1.60
MgO	3.12	0.10	3.08	2.94	0.74	0.54	1.53	4.82	3.29
C_3S	60.42	0.82	61.03	42.14	5.89	34.74	42.14	73.91	31.77
C_3A	7.00	0.16	6.90	7.46	1.17	1.37	4.68	9.15	4.47
K_2O	0.81	0.03	0.80	0.79	0.24	0.06	0.38	1.55	1.17
Na_2O	0.14	0.01	0.12	0.12	0.08	0.01	0.05	0.42	0.37
$f - CaO$	1.06	0.07	0.95	0.76	0.49	0.24	0.35	2.73	2.38
Fe_2O_3	3.32	0.05	3.34	3.39	0.35	0.12	2.11	4.17	2.06
C_4AF	10.09	0.15	10.14	10.35	1.08	1.16	6.40	12.68	6.28
Al_2O_3	4.76	0.06	4.73	4.71	0.40	0.16	3.80	5.40	1.60
SiO_2	20.80	0.37	21.12	21.07	2.68	7.17	2.31	22.81	20.50
C_2S	15.41	0.77	14.63	14.88	5.53	30.53	3.36	33.86	30.50
CCS (MPa)	**52.37**	**0.92**	**53.90**	**54.80**	**6.66**	**44.36**	**26.00**	**61.80**	**35.80**
FS (MPa)	**8.22**	**0.07**	**8.40**	**8.40**	**0.47**	**0.22**	**6.90**	**8.80**	**1.90**

Note: LOI = Loss on Ignition; SSA = Specific Surface Area; All chemical compositions in %, except SSA in cm^2/g and strengths in MPa.

conditions.

4.4. Multi-head estimator

The multi-head estimator models CCS and FS as full probability distributions—$p(s_{CCS}|\boldsymbol{z})$ and $p(s_{FS}|\boldsymbol{z})$—while respecting their shared origins and divergent responses. Both properties stem from common drivers (e.g., C_3S, particle packing), yet FS is more sensitive to pore coarsening and ettringite destabilization. Each head transforms $\boldsymbol{z}$ into a GMM:

$$\boldsymbol{\alpha}_k = \text{Softmax}\ \left(\boldsymbol{W}_{\alpha,k}\boldsymbol{z} + \boldsymbol{b}_{\alpha,k}\right), \tag{5}$$

$$\boldsymbol{\mu}_k = \boldsymbol{W}_{\mu,k}\boldsymbol{z} + \boldsymbol{b}_{\mu,k}, \tag{6}$$

$$\boldsymbol{\sigma}_k^2 = \text{Softplus}\left(\boldsymbol{W}_{\sigma,k}\boldsymbol{z} + \boldsymbol{b}_{\sigma,k}\right), \tag{7}$$

yielding the conditional density:

$$g_k(s_k|\boldsymbol{z}) = \sum_{j=1}^{N_k} \alpha_{k,j} \cdot \mathcal{N}\left(s_k; \mu_{k,j}, \sigma_{k,j}^2\right), k \in \{\text{CCS}, \text{FS}\}. \tag{8}$$

This design accommodates property-specific physics: *CCS*: Enhanced by fine particles (<3 μm) via packing density, *FS*: Degraded by the same fines due to shrinkage cracking, *Shared risks*: Clinker fluctuations (±8% C_3S) propagate through $\boldsymbol{z}$ to both outputs. By preserving cross-property dependencies while enabling independent uncertainty profiles, MPCSE provides the foundation for "dual-strength compliance risk assessment"—a cornerstone of sustainable, precision manufacturing.

4.5. Loss function and optimization

The MPCSE model is optimized to balance three materials-science objectives: *(i)* accurate prediction of compressive and flexural strength, *(ii)* meaningful quantification of irreducible material noise versus reducible model gaps, and *(iii)* preservation of physicochemical couplings (e.g., C_3S–sulfate trade-offs) under real-world variability. This ensures the model learns interpretable representations of cement hydration—not just statistical patterns.

Central to this is *Kullback-Leibler (KL) regularization* [54], which enforces physicochemical plausibility in the shared latent space $\boldsymbol{z}$ by minimizing divergence between the variational posterior $q(\boldsymbol{z}|\boldsymbol{x})$ and a standard normal prior $p(\boldsymbol{z}) = \mathcal{N}(0, \boldsymbol{I})$:

$$\mathcal{L}_{\text{KL}} = \frac{1}{2B} \sum_{i=1}^{B} \sum_{d=1}^{n_z} \left(\mu_{i,d}^2 + \exp(\log\sigma_{i,d}^2) - \log\sigma_{i,d}^2 - 1\right), \tag{9}$$

where B is batch size and n_z latent dimensionality. During training, $\log\sigma^2$ is clamped to $[-5, 5]$ for numerical stability.

The full loss combines task-specific negative log-likelihoods (NLL) with KL regularization:

$$L_{\text{total}} = \underbrace{L_{\text{NLL}}^{\text{CCS}} + L_{\text{NLL}}^{\text{FS}}}_{\text{Data fidelity}} + \lambda \underbrace{L_{\text{KL}}}_{\text{Uncertainty regularization}}, \tag{10}$$

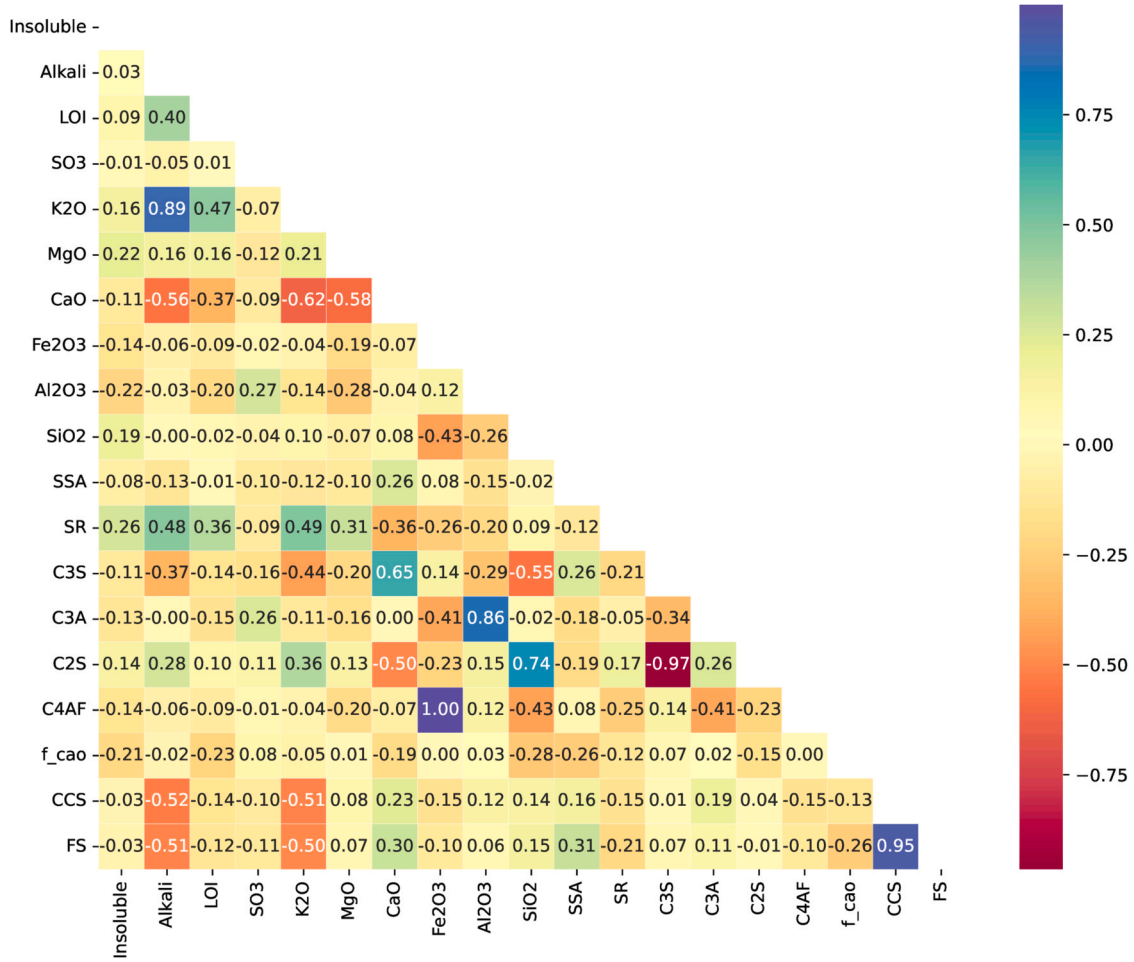

Fig. 3. Pearson correlation matrix for the laboratory cement dataset ($M = 50$). Values range from − 1–1, where positive coefficients (blue) indicate direct relationships and negative coefficients (red) indicate inverse relationships. The strong correlation between compressive and flexural strength (r=0.945) reveals tight coupling under controlled conditions.

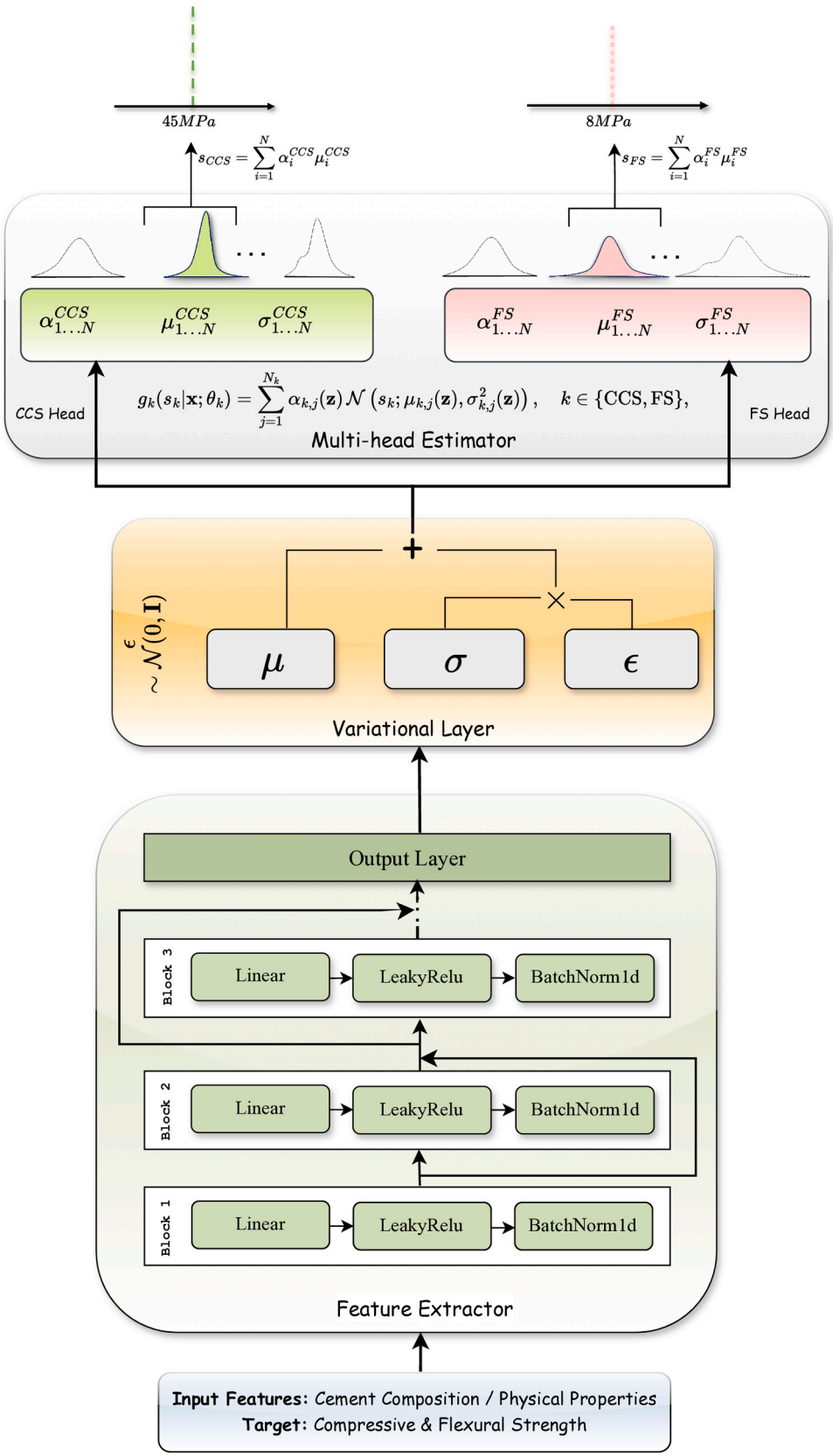

Fig. 4. Semantic architecture of the proposed Multi-Property Cement Strength Estimator. The framework integrates three synergistic modules: (1) **Feature extractor** that extracts hierarchical representations from physicochemical inputs; (2) **Variational layer** that transforms encoded features into a shared latent space $\boldsymbol{z}$, explicitly separating aleatoric uncertainty from epistemic uncertainty; and (3) **Multi-head estimator** that generate full probability distributions for compressive and flexural strength from the shared latent representation. The architecture preserves cement-chemical interpretability by encoding common drivers in $\boldsymbol{z}$ while allowing property-specific response patterns. This design enables joint probabilistic modeling, uncertainty decomposition, and dual-strength compliance assessment—the three core innovations of MPCSE.

with $\lambda = 2.15 \times 10^{-5}$—a cross-validated value that balances strength prediction fidelity against uncertainty regularization. The NLL for property $k \in \{\mathrm{CCS}, \mathrm{FS}\}$ is:

$$\mathscr{L}_{\mathrm{NLL}}^{k} = -\frac{1}{B}\sum_{i=1}^{B}\log\sum_{j=1}^{N_k}\alpha_{k,j}(\boldsymbol{z}_i)\cdot\mathcal{N}\left(s_{k,i};\mu_{k,j}(\boldsymbol{z}_i),\sigma_{k,j}^{2}(\boldsymbol{z}_i)\right). \tag{11}$$

This balance is critical for industrial applicability. An excessively large λ oversmooths $\boldsymbol{z}$, weakening sensitivity to key hydration trade-offs (e.g., C_3S/sulfate interactions governing ettringite stability). Too small a λ inflates epistemic uncertainty, obscuring genuine material variability. The chosen λ preserves microstructural sensitivity while isolating aleatoric uncertainty from clinker fluctuations ($\pm 8\%$ C_3S) and curing inconsistencies.

Optimization uses Adam ($\eta = 0.01$, $\beta_1 = 0.9$, $\beta_2 = 0.999$) with step decay ($\gamma = 0.1$ every 100 epochs) and gradient clipping ($\| \nabla\boldsymbol{\Theta}\|_2 \leq 1.0$) to stabilize training under industrial batch-to-batch variability. Backpropagation coordinates updates across: (i) task-specific heads (shaping individual strength distributions), and (ii) shared encoder (capturing common pathways like C_3A-induced brittleness). This joint optimization ensures $\boldsymbol{z}$ encodes both shared mechanisms and divergent responses—enabling reliable joint distribution modeling for risk-aware formulation design.

4.6. Dual-strength compliance probability metric

The dual-strength compliance probability $\overline{P}(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$ provides a risk-aware framework for evaluating simultaneous compliance with ASTM/EN standards—replacing binary pass/fail judgments with continuous risk assessment. This is essential in industrial production, where unquantified variability causes unnecessary batch rejections despite acceptable individual strengths.

As established in Section 4.3, the model does not explicitly parameterize a full multivariate Gaussian mixture distribution. Instead, it assumes conditional independence of CCS and FS given the shared latent representation $\boldsymbol{z}$, which encodes common physicochemical drivers (e.g., C_3S content, sulfate balance, particle size distribution). The joint density is constructed using Eq. (4). Under this conditional independence framework, for a specific draw of $\boldsymbol{z}$, the predictions for CCS and FS are independent, yielding:

$$P(\mathrm{CCS} \geq c_{\min}, \mathrm{FS} \geq f_{\min}|\boldsymbol{x}_i, \boldsymbol{z}) = P(\mathrm{CCS} \geq c_{\min}|\boldsymbol{z}) \times P(\mathrm{FS} \geq f_{\min}|\boldsymbol{z}).$$

The overall joint probability is then obtained by taking the expectation over the posterior distribution of $\boldsymbol{z}$ given $\boldsymbol{x}_i$:

$$P_{\mathrm{joint}}^{(i)} = \mathbb{E}_{p(\boldsymbol{z}|\boldsymbol{x}_i)}[P(\mathrm{CCS} \geq c_{\min}|\boldsymbol{z})\cdot P(\mathrm{FS} \geq f_{\min}|\boldsymbol{z})].$$

In practice, this expectation is approximated using the mixture model parameters directly. The exceedance probability for each strength property is computed from its GMM parameters:

$$P(k \geq k_{\min}|\boldsymbol{x}_i) = \sum_{j=1}^{N_k}\alpha_{k,j}\left[1 - \Phi\left(\frac{k_{\min} - \mu_{k,j}}{\sigma_{k,j}}\right)\right], k \in \{\mathrm{CCS}, \mathrm{FS}\}, \tag{12}$$

where Φ is the standard normal cumulative distribution function, and $\alpha_{k,j}, \mu_{k,j}, \sigma_{k,j}$ are the mixing weight, mean, and standard deviation of the j-th Gaussian component for property k. The joint probability is then:

$$P_{\mathrm{joint}}^{(i)} = P(\mathrm{CCS} \geq c_{\min}|\boldsymbol{x}_i) \times P(\mathrm{FS} \geq f_{\min}|\boldsymbol{x}_i).$$

This formulation correctly propagates uncertainty through the latent space while preserving the dependencies induced by shared physicochemical drivers. The batch-averaged compliance probability across M

Input: Trained model f_θ, Validation set $\mathcal{D}_{val} = \{\boldsymbol{x_i}\}_{i=1}^{M}$, Thresholds c_{min}, f_{min}, Target scalers: $\mathrm{scaler}_y^{\mathrm{CCS}}$, $\mathrm{scaler}_y^{\mathrm{FS}}$

Output: Compliance probability $\bar{P}(c_{min}, f_{min})$

$\tilde{c}_{min} \leftarrow \mathrm{scaler}_y^{\mathrm{CCS}}(c_{min})$; $\tilde{f}_{\min} \leftarrow \mathrm{scaler}_y^{\mathrm{FS}}(f_{min})$

$\mathrm{sum} \leftarrow 0$

for $i \leftarrow 1$ to M **do**

  $\{\alpha_{\mathrm{CCS},j}, \mu_{\mathrm{CCS},j}, \sigma_{\mathrm{CCS},j}\}_{j=1}^{2} \leftarrow f_\theta(\boldsymbol{x_i})$

  $\{\alpha_{\mathrm{FS},k}, \mu_{\mathrm{FS},k}, \sigma_{\mathrm{FS},k}\}_{k=1}^{2} \leftarrow f_\theta(\boldsymbol{x_i})$

  $$P_{\mathrm{CCS}} \leftarrow \sum_{j=1}^{2}\alpha_{\mathrm{CCS},j}\left[1 - \Phi\left(\frac{\tilde{c}_{min} - \mu_{\mathrm{CCS},j}}{\sigma_{\mathrm{CCS},j}}\right)\right]$$

  $$P_{\mathrm{FS}} \leftarrow \sum_{k=1}^{2}\alpha_{\mathrm{FS},k}\left[1 - \Phi\left(\frac{\tilde{f}_{\min} - \mu_{\mathrm{FS},k}}{\sigma_{\mathrm{FS},k}}\right)\right]$$

  $\mathrm{sum} \leftarrow \mathrm{sum} + P_{\mathrm{CCS}} \times P_{\mathrm{FS}}$

end **for**

$\bar{P} \leftarrow \mathrm{sum}/M$

return $\bar{P}$

validation samples is:

$$\overline{P}(c_{\min}, f_{\min}) = \frac{1}{M}\sum_{i=1}^{M} P_{\text{joint}}^{(i)}.$$

Algorithm 2. formalizes this computation, enabling consistent evaluation across threshold grids (e.g., $c_{\min} \in \{30, 35, 40, 45, 50, 55, 60\}$ MPa, $f_{\min} \in \{5, 6, 7, 8, 9, 10\}$ MPa).

By quantifying dual compliance probabilistically, manufacturers can move beyond deterministic pass/fail decisions to risk-informed quality control:

- **High-confidence batches** ($\overline{P} > 0.95$): Proceed directly to structural applications with minimal risk of non-compliance.
- **Borderline batches** ($0.7 < \overline{P} < 0.95$): Reroute to non-structural applications (e.g., pavements, backfill) rather than outright rejection, recovering value while maintaining safety.
- **High-risk batches** ($\overline{P} < 0.7$): Trigger targeted process adjustments (e.g., grinding optimization, sulfate balance correction) or additional testing before disposition.

Algorithm 2. Dual-strength compliance probability estimation

Input: Model f_θ; Feature x_f; Training range $[L, U]$;
Margin $\delta = (U - L)/2$; Samples per regime $m = 100$
Output: $\mathcal{U}^{\text{CCS}}_{\text{aleatoric}}$, $\mathcal{U}^{\text{FS}}_{\text{aleatoric}}$; $\mathcal{U}^{\text{CCS}}_{\text{epistemic}}$, $\mathcal{U}^{\text{FS}}_{\text{epistemic}}$

$R_{\text{al}} \leftarrow [L, U]$; $R_{\text{ep}} \leftarrow [L - \delta, L] \cup [U, U + \delta]$
$\boldsymbol{x}^{\boldsymbol{al}}_{\boldsymbol{synth}} \sim \mathcal{U}(R_{al})$; $\boldsymbol{x}^{\boldsymbol{ep}}_{\boldsymbol{synth}} \sim \mathcal{U}(R_{ep})$
for $i \leftarrow 1$ **to** m **do**
    $\boldsymbol{x}_p^{(i,al)} \leftarrow$ Replace x_f in median input with $\boldsymbol{x}^{\boldsymbol{al}}_{\boldsymbol{synth},i}$
    $\boldsymbol{x}_p^{(i,ep)} \leftarrow$ Replace x_f in median input with $\boldsymbol{x}^{\boldsymbol{ep}}_{\boldsymbol{synth},i}$
end **for**
forall $\boldsymbol{x}_p \in \{\boldsymbol{x}_p^{(i,al)}, \boldsymbol{x}_p^{(i,ep)}\}_{i=1}^{m}$ **do**
    $[\alpha, \mu, \sigma]_{\text{CCS, FS}} \leftarrow f_\theta(\boldsymbol{x}_p)$
    $\mathcal{U}^{CCS}_{total}$.append$\left(\sum_j \alpha_{\text{CCS},j}\left(\sigma^2_{\text{CCS},j} + \mu^2_{\text{CCS},j}\right)\right)$
    $\mathcal{U}^{FS}_{total}$.append$\left(\sum_j \alpha_{FS,j}\left(\sigma^2_{FS,j} + \mu^2_{FS,j}\right)\right)$
end **for**
$\mathcal{U}^{CCS,FS}_{\text{aleatoric}} \leftarrow \text{mean}\left(\mathcal{U}^{CCS,FS}_{total}[i] \mid x_f^{(i)} \in [L, U]\right)$
$\mathcal{U}^{CCS,FS}_{\text{epistemic}} \leftarrow \text{mean}\left(\mathcal{U}^{CCS,FS}_{total}[i] \mid x_f^{(i)} \notin [L, U]\right)$
return $\mathcal{U}^{CCS}_{\text{aleatoric}}, \mathcal{U}^{FS}_{\text{aleatoric}}, \mathcal{U}^{CCS}_{\text{epistemic}}, \mathcal{U}^{FS}_{\text{epistemic}}$

As demonstrated in Section 5.5, this risk-aware approach reduces batch rejection rates by approximately 25% while maintaining 93.7% joint compliance at moderate thresholds (45 MPa CCS / 8 MPa FS). By transforming uncertainty from a source of conservative overdesign into actionable intelligence, the dual-strength compliance probability metric directly supports sustainable, precision manufacturing.

4.7. Uncertainty decomposition in cement strength predictions

Distinguishing between sources of strength variability is essential for effective quality control. Two distinct uncertainties affect CCS and FS: *aleatoric uncertainty* (irreducible material/process noise, e.g., ±8% C_3S fluctuations) and *epistemic uncertainty* (reducible model gaps from sparse data, e.g., low-C_3A regimes). Conflating these leads to reactive overdesign; separating them enables targeted interventions—tighter grinding control for aleatoric hotspots, focused experimentation for epistemic boundaries. The proposed framework decomposes uncertainty via the multi-head GMM architecture. For property $k \in \{\text{CCS}, \text{FS}\}$, total predictive uncertainty is the second raw moment:

$$\mathscr{U}^k_{total}(\boldsymbol{x}) = E\left[s_k^2 \mid \boldsymbol{x}\right] = \sum_{j=1}^{N_k} \alpha_{kj}(\boldsymbol{z})\left(\underbrace{\sigma^2_{kj}(\boldsymbol{z})}_{\text{Aleatoric}} + \underbrace{\mu^2_{kj}(\boldsymbol{z})}_{\text{Epistemic}}\right), \tag{13}$$

where $\boldsymbol{z} = f_\theta(\boldsymbol{x})$ captures shared physicochemical features. This disentanglement transforms uncertainty into actionable intelligence:

- *Aleatoric*: Irreducible variability (σ^2) reflecting batch heterogeneity,

- *Epistemic*: Reducible uncertainty (μ^2) indicating data-scarce regions.

Algorithm 3. Uncertainty decomposition via synthetic perturbation

Algorithm 3. quantifies these components via controlled perturbation. For feature x_f (e.g., $x_{<1}$: particles < 1 μm), the following steps are performed:

1. Define training range $[L, U]$ (e.g., $[1.37, 2.77]$ for $x_{<1}$),
2. Generate synthetic samples within $[L, U]$ (aleatoric) and outside (epistemic),
3. Compute mean uncertainty in each regime.

This reveals how deviations impact confidence: e.g., $x_{<1} > 2.77$ spikes epistemic uncertainty (poor model reliability in over-ground regimes), while high aleatoric uncertainty in $[L, U]$ indicates persistent batch variability—guiding whether to adjust manufacturing practices or collect new data.

5. Results and discussions

This section presents a comprehensive experimental validation of the *Multi-Property Cement Strength Estimator*, addressing seven interrelated questions critical to sustainable cement manufacturing under real-world variability. The analysis progresses from architectural validation to industrial impact, leveraging both industrial ($M = 163$) and laboratory ($M = 50$) datasets to ensure robustness across scales. The research questions explored are: *(a)* Are MPCSE's core components essential? *(b)* Does joint modeling outperform decoupled approaches? *(c)* How does MPCSE compare to state-of-the-art baselines? *(d)* Can decomposed uncertainty guide process optimization? *(e)* What is the dual-strength compliance probability under real-world thresholds? *(f)* What is the reliability of risk assessments based on the 95.7% confidence interval coverage? *(g)* What mechanistic insights explain context-dependent driver divergence?

First, the necessity of the different modules in MPCSE's architecture is validated through a 5-fold ablation study on industrial data (Section 5.1). Next, MPCSE is compared against single-head and separate modeling architectures (where CCS and FS are predicted independently) in Section 5.2. MPCSE is benchmarked against five representative baselines—including Physics-Informed Models (PIM), XGBoost, and Bayesian Neural Networks—and demonstrates superior robustness in industrial settings (Section 5.3). Section 5.4 demonstrates how uncertainty decomposition transforms predictive noise into actionable intelligence. The dual-strength compliance metric $P(\text{CCS} \geq c_{\min} \cap \text{FS} \geq f_{\min})$ is introduced to replace binary pass/fail judgments with continuous risk assessment (Section 5.5). The reliability of these risk assessments is validated by 95.7% confidence interval coverage (Section 5.6). Finally, Section 5.7 explores the mechanistic interpretability of the model using SHAP (SHapley Additive exPlanations) analysis.

5.1. Ablation study: necessity of MPCSE components

To validate the necessity of each architectural component in the *Multi-Property Cement Strength Estimator*, a 5-fold cross-validation ablation study is conducted on the industrial dataset ($M = 163$). Four variants are evaluated: *(i)* the full MPCSE; *(ii)* "No-Shared Encoder", which uses separate MLP encoders for CCS and FS; *(iii)* "No-Variational", which replaces the variational layer with a deterministic bottleneck; and *(iv)* "Point-Estimate", a deterministic multi-task MLP that outputs scalar predictions without uncertainty quantification. All models use identical hyperparameters, training protocols, and feature sets. Performance is assessed using MAE (MPa), 95% confidence interval (CI) coverage (%), and the dual-strength compliance probability $P(\text{CCS} \geq 50 \text{ MPa} \cap \text{FS} \geq 8 \text{ MPa})$—a realistic high-performance threshold aligned with EN 206–1 specifications for precast elements, where 98.77% of industrial batches meet these individual requirements (Table 1).

Results (Table 1) confirm that all three MPCSE components are essential for sustainable, risk-aware quality control. The full MPCSE achieves the lowest MAE for both CCS (0.863 MPa) and FS (0.156 MPa), near-perfect CI calibration (96.3% for CCS, 95.7% for FS), and the highest dual-strength compliance (93.7%). Removing the shared latent space ("No-Shared") degrades CI coverage by 4–7 %age points and reduces joint compliance by 1.0 pp—confirming that shared representations are critical to capture common drivers like C_3S and particle packing efficiency, which jointly govern strength outcomes under real-world variability. Eliminating the variational layer ("No-Variational") causes the largest drop in CI coverage (87.7% for CCS), revealing that explicit uncertainty decomposition is indispensable for reliable risk assessment. Without it, manufacturers cannot distinguish between irreducible material noise (e.g., grinding heterogeneity) and reducible model gaps (e.g., sparse low-C_3A data)—leading to either overdesign or unwarranted rejections. Finally, the deterministic "Point-Estimate" baseline, while competitive in MAE, cannot quantify compliance risk or support resource-efficient batch allocation, highlighting the irreplaceable role of probabilistic modeling in reducing clinker overuse and batch rejection.

Collectively, this ablation confirms that all three components—shared latent space, variational uncertainty decomposition, and multi-head GMM—are essential for MPCSE's ability to deliver accurate, calibrated, and actionable strength predictions in industrial settings.

5.2. Comparative analysis of joint vs. separate modeling architectures

This study rigorously evaluates whether joint probabilistic modeling significantly outperforms single-property approaches in predicting compressive and flexural strength under real-world material and process variability—addressing a core challenge in sustainable cement production: reducing clinker overuse and batch rejection through reliable dual-strength assessment. A comparison of three strategies is presented: *(i)* the proposed multi-head Gaussian Mixture Model (GMM) with shared latent representations and task-specific heads, *(ii)* a single-head GMM that estimates a unified distribution for both properties, and *(iii)* separate modeling using independent models for CCS and FS. Performance is assessed on industrial and laboratory datasets using MAE, RMSE, R^2, and NLL.

5.2.1. Analysis of model performance on industrial dataset

The industrial dataset presents a critical testbed: weak CCS–FS correlation ($r = 0.101$) and high batch-to-batch variability from grinding inconsistencies, clinker heterogeneity (±8% C_3S), and curing fluctuations—conditions that drive conservative overdesign and 15–20% batch rejection. Table 2 shows the multi-head GMM's superiority for sustainable quality control.

For CCS, the multi-head architecture reduces MAE by 7.8% versus single-head (0.863 ± 0.091 vs. 0.936 ± 0.147) and is the only approach achieving positive R^2 (0.146 ± 0.062)—proving its ability to extract signal from noise. Separate modeling fails ($R^2 = -0.005 \pm 0.187$), confirming that decoupled models cannot capture shared granulometric drivers like fine particles (< 3 μm, $r = 0.936$ with Blaine fineness) and uniformity index, which govern packing density and hydration kinetics.

For FS, the multi-head GMM achieves 9.3% lower MAE than single-head (0.156 ± 0.028 vs. 0.172 ± 0.032) and the least negative R^2 (-0.070 ± 0.086), demonstrating robustness despite FS's narrow range (8.0–9.0 MPa) and sensitivity to shading effects ($r = -0.242$). Crucially, its task-specific uncertainty modeling enables reliable risk assessment—unlike single-head's marginally better NLL (1.354 vs. 1.429),

which lacks property-specific adaptability for actionable decisions.

5.2.2. Analysis of model performance on laboratory dataset

In controlled laboratory conditions—where CCS–FS correlation is strong ($r = 0.945$), and stoichiometry dominates—the multi-head GMM maintains its edge in uncertainty calibration, critical for R&D efficiency. It achieves the lowest NLL for FS (1.371 ± 0.296), outperforming single-head (1.401 ± 0.138) and separate modeling (1.655 ± 0.314) (Table 3). While single-head shows slightly better CCS calibration (NLL: 1.419 ± 0.230), its rigid parameterization fails at stoichiometric extremes (e.g., $C_3S > 68\%$), yielding unstable R^2 (-0.293 ± 0.402). The multi-head framework adapts by preserving property-specific pathways—C_3S, mediated C–S–H nucleation for CCS and ettringite stability for FS—within a shared latent space.

These results validate that the shared-latent, task-specific architecture resolves a fundamental limitation of prior work: it captures mechanistic couplings (e.g., fine particles enhancing CCS via packing density yet degrading FS through shrinkage cracking) while accommodating property-specific responses. This balance enables robust dual-strength prediction across scales—supporting both industrial risk-aware control and laboratory formulation refinement.

5.3. Performance comparison with baseline models

This subsection evaluates the *Multi-Property Cement Strength Estimator* against five representative baseline models: Multi-output XGBoost [40], Gaussian Process Regression (GPR) [41], Multi-task Gaussian Process (GP) [42], Physics-Informed Model (PIM) [55], and Bayesian Neural Network (BNN) [56]. Performance is assessed using Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) for compressive and flexural strength predictions across both industrial ($M = 163$) and laboratory ($M = 50$) datasets are shown in Fig. 5(a-f).

In the industrial dataset in Fig. 5(a), MPCSE achieves competitive CCS prediction ($MAE_{CCS} = 0.863 \pm 0.091 MPa$), closely trailing PIM ($0.841 \pm 0.076 MPa$) and Multi-output XGBoost ($0.860 \pm 0.099 MPa$). Crucially, MPCSE outperforms all baselines in FS prediction, as shown in Fig. 5(b), achieving an $MAE_{FS} = 0.156 \pm 0.028 MPa$—a 3.1% improvement over PIM ($0.161 \pm 0.030 MPa$). Its $RMSE_{CCS} = 1.085 \pm 0.155 MPa$ shown in Fig. 5(e) is significantly lower than all alternatives, indicating superior robustness to outliers and batch-to-batch variability. This is visually evident in Fig. 5(e), where MPCSE's error bars are consistently shorter—reflecting tighter error distributions enabled by its multi-head GMM architecture and joint modeling of CCS–FS interdependencies.

The slight advantage of PIM in CCS metrics stems from its incorporation of domain-specific constraints (e.g., hydration kinetics, phase reactivity), which enhance accuracy for well-characterized mixes. However, PIM's deterministic nature limits its utility in risk-aware quality control—it cannot quantify dual-strength compliance probability or distinguish between material-driven variability and model limitations. In contrast, MPCSE's probabilistic framework provides actionable uncertainty estimates, enabling manufacturers to assess $P(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$ and make informed decisions about batch allocation—directly supporting clinker reduction and rejection avoidance.

Multi-output XGBoost performs competitively in MAE_{CCS}, reflecting its ensemble learning capacity to handle heterogeneous industrial data (Fig. 5(a)). However, its lack of joint distribution modeling results in higher $RMSE_{FS} = 0.209 \pm 0.035 MPa$ (vs. MPCSE's $0.193 \pm 0.038 MPa$) in Fig. 5(f), indicating greater sensitivity to outliers—a critical drawback in production environments where consistency is paramount.

On the laboratory dataset, PIM achieves the best overall accuracy ($MAE_{CCS} = 3.145 \pm 1.123 MPa$, $MAE_{FS} = 0.318 \pm 0.084 MPa$) shown in Fig. 5(a-b), outperforming MPCSE ($3.755 \pm 1.286 MPa$, $0.342 \pm 0.110 MPa$). This reflects the alignment between PIM's physics-informed assumptions and tightly controlled lab conditions (e.g., stable C_3S, minimal curing variation). Nevertheless, MPCSE offers more reliable uncertainty quantification, as evidenced by its lower $MAPE_{FS} = 0.043 \pm 0.014$ (vs. BNN's 0.075 ± 0.074; Fig. 5(d)), thanks to its variational uncertainty decomposition. This capability is essential for translating predictions into confidence intervals that support real-world decision-making.

Across both datasets, GPR and Multi-task GP exhibit the highest errors—e.g., Multi-task GP's $RMSE_{CCS} = 7.235 \pm 2.729 MPa$ on industrial data in Fig. 5(e)—due to sensitivity to noise and limited capacity to capture non-additive interactions. BNN shows high variance (e.g., $RMSE_{CCS} = 5.613 \pm 4.507 MPa$ on laboratory data) in Fig. 5(e), indicative of overfitting in low-noise regimes. Overall, MPCSE's balanced performance—particularly its low RMSE, robust uncertainty calibration, and ability to model joint strength distributions—highlights its suitability for industrial applications.

Table 3
Ablation study results using 5-fold cross-validation on the industrial dataset ($M = 163$). Thresholds for dual-strength compliance set at CCS $\geq$ 50 MPa and FS $\geq$ 8 MPa.

Model Variant	MAE (CCS)	MAE (FS)	CI Cov. (CCS)	CI Cov. (FS)	Joint Comp.
	[MPa]	[MPa]	[%]	[%]	[%]
Full MPCSE	**0.863 ± 0.09**	**0.156 ± 0.03**	**96.31 ± 7.05**	**95.72 ± 5.06**	**93.72 ± 2.33**
No-Shared Encoder	0.878 ± 0.14	0.170 ± 0.02	91.97 ± 7.55	89.02 ± 8.05	92.78 ± 1.30
No-Variational	0.887 ± 0.13	0.168 ± 0.02	87.71 ± 9.47	90.87 ± 6.62	91.87 ± 1.66
Point-Estimate	0.969 ± 0.15	0.173 ± 0.02	N/A	N/A	N/A

Note: CI Cov. refers to 95% Confidence Interval Coverage. N/A indicates metrics not supported by the model type.

5.4. Uncertainty decomposition for process guidance

This experiment analyzes predictive uncertainty in compressive and flexural strength modeling, isolating contributions from two distinct sources: *aleatoric uncertainty* and *epistemic uncertainty*. By perturbing key features across industrial and laboratory datasets, the analysis reveals how different environments shape uncertainty profiles: granulometric variability dominates in industrial settings, while stoichiometric precision governs in controlled labs.

5.4.1. Industrial uncertainty decomposition: granulometry-driven control points

Employing Algorithm 3, systematic perturbation reveals distinct uncertainty profiles: aleatoric uncertainty dominates central value ranges, while epistemic uncertainty peaks at feature extremes (Fig. 6). Crucially, FS consistently exhibits 8–12% higher aleatoric uncertainty than CCS across granulometric parameters, confirming its heightened sensitivity to particle packing efficiency and curing humidity—key drivers of interfacial transition zone quality. Three key patterns emerge:

1. *Aleatoric hotspots in mid-range particles*: FS uncertainty peaks at 1.80–1.99 in the 41–52 μm range shown in Fig. 6(c), aligning with Stovall & de Larrard's void-reduction principle: suboptimal packing increases porosity, weakening FS more than CCS.
2. *Epistemic boundaries at operational extremes:* Low illuminance (<550 lux) spikes uncertainty in both properties (CCS: 1.89, FS: 1.85; Fig. 6 (g)), reflecting model limitations under atypical sensor conditions (e.g., night shifts). High shading (>36%) elevates epistemic uncertainty (CCS: 1.76, FS: 1.80; Fig. 6(a)), indicating under represented high-roughness batches prone to surface defects.
3. *Cross-property synergies at optimal granulometry:* A uniformity index of 0.85–0.90 jointly suppresses aleatoric uncertainty by 12–18% (Fig. 6

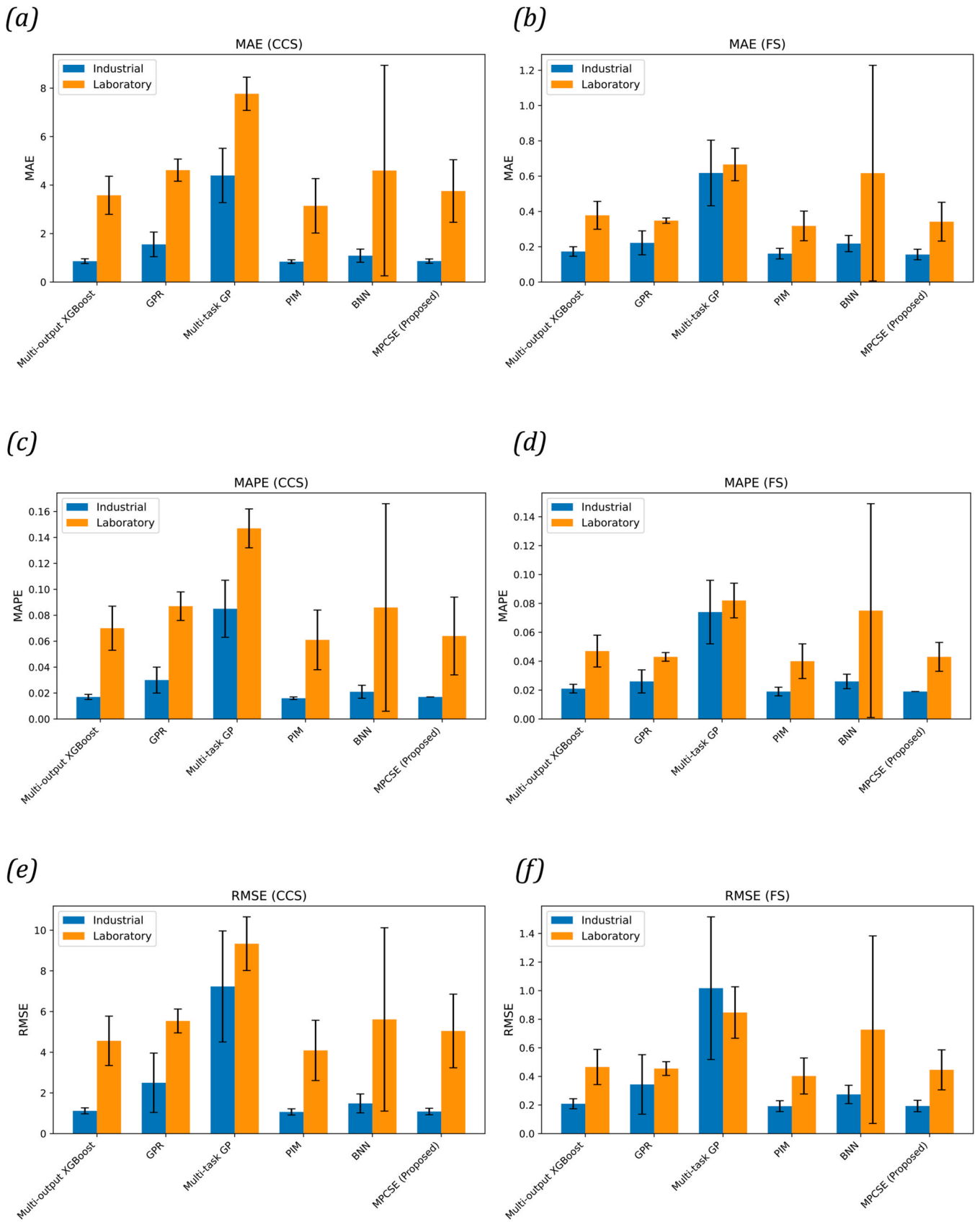

Fig. 5. Comparison of multi-property models on industrial and laboratory cement datasets across MAE, MAPE, and RMSE. (a) Mean absolute error (MAE) for CCS. (b) MAE for FS. (c) Mean absolute percentage error (MAPE) for CCS. (d) MAPE for FS. (e) Root mean square error (RMSE) for CCS. (f) RMSE for FS. Models compared: MPCSE (proposed), PIM, XGBoost, GPR, BNN, and Multi-task GP. Error bars represent standard deviation across five-fold cross-validation.

(b)), enabling co-optimization of mechanical performance through balanced fine/coarse particle ratios.

For industrial practice, this decomposition delivers actionable guidance:

- *Tighter grinding control:* Optimizing 41–52 μm particles reduces FS uncertainty by up to 15%,
- *Targeted data collection:* Regions like $< 70\mu m$ $> 99.5\%$ retention require focused experimentation (Fig. 6(f)),
- *Risk-aware batch allocation:* Batches with epistemic uncertainty > 1.75 can be rerouted to non-structural uses, reducing rejection rates—directly supporting resource efficiency.

5.4.2. Laboratory uncertainty decomposition: stoichiometric sensitivity

In controlled laboratory settings—where granulometric variability is minimized—chemical composition dominates uncertainty profiles. FS consistently exhibits 7–12% higher aleatoric uncertainty than CCS (Fig. 7), reflecting its greater sensitivity to hydration kinetics and phase balance. Key mechanistic insights emerge:

1. *Sulfate balance:* A pronounced trough in aleatoric uncertainty at $SO_3 = 2.2–2.4\%$ (Fig. 7(g)) confirms sulfate optimization stabilizes ettringite networks—essential for FS without compromising CCS.
2. *Alkali-silica reactivity*: Elevated epistemic uncertainty in FS at low alkali (<0.5%; Fig. 7(b)) highlights model limitations in predicting ASR-induced cracking—a known durability concern.
3. *C_3S–FS trade-off*: Divergent aleatoric peaks—CCS maximized at 55–67% C_3S, FS at 42–55%—reveal an inverse relationship that enables balanced formulation design.

These patterns provide actionable stoichiometric guidance: Maintain alkali levels between 0.5–0.8% to reduce FS uncertainty by 15% through effective ASR suppression; Target $SO_3 = 2.3\% \pm 0.1\%$ to optimally balance CCS–FS interplay; Compensate for high $f-CaO$ (>2.5%) with gypsum adjustments to mitigate expansion risks (Fig. 7(e)).

5.4.3. Transforming quality control through uncertainty awareness

Uncertainty decomposition enables three transformative shifts in cement quality control: *First,* real-time quantification of physicochemical trade-offs supports adaptive formulation design. Reducing ultrafines (< 1 μm) below 1.5% decreases CCS aleatoric uncertainty by 10%, but increases water demand—necessitating compensatory SO_3 adjustments to preserve ettringite stability for FS.

Second, epistemic uncertainty thresholds (> 1.75) enable rational decision-making: high-uncertainty batches are rerouted to non-structural applications, reducing rejections and cutting quality control costs—while maintaining ASTM/EN compliance through risk-aware allocation.

Third, epistemic boundaries guide strategic data collection in underrepresented regions (e.g., low C_3A, suboptimal curing), improving model robustness, and reducing reliance on conservative overdesign. This directly contributes to sustainability by lowering CO_2 emissions through precision-driven clinker substitution.

By transforming uncertainty into actionable intelligence, MPCSE resolves the compressive–flexural strength trade-off—reducing material waste, enhancing multi-property robustness, and advancing sustainable cement production under real-world variability.

5.5. Dual-strength compliance probability under industrial variability

This experiment introduces the dual-strength compliance probability metric, defined as $\overline{P}(c_{\min}, f_{\min}) = P(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$, across threshold grids aligned with ASTM/EN standards (e.g., CCS $\geq 50MPa$, FS $\geq 9MPa$) employing Algorithm 2. This probabilistic metric quantifies the likelihood that a cement formulation simultaneously meets dual mechanical performance criteria—offering a continuous, risk-aware alternative to traditional binary pass/fail quality control. By leveraging MPCSE's full joint distribution of CCS and FS, manufacturers can assess compliance under real-world variability rather than relying on deterministic point estimates. As illustrated in Fig. 8, compliance exhibits a sharp decline as thresholds increase, providing a quantitative framework for addressing the compressive–flexural strength trade-off through actionable risk estimation.

5.5.1. Analysis of dual compliance probability in industry and laboratory datasets

In the industrial dataset (Fig. 8(a)), compliance remains near-perfect ($\overline{P} \approx 1.0$) for moderate thresholds (e.g., CCS $\leq 45MPa$, FS $\leq 8MPa$), indicating robust performance under typical production constraints. However, compliance collapses dramatically at higher thresholds:

1. At CCS $= 50MPa$ and FS $= 9MPa$, only 0.38% of batches meet both criteria,
2. Increasing CCS to 60 MPa while maintaining FS $\geq 10MPa$ reduces compliance to effectively zero ($\overline{P} = 1.8 \times 10^{-8}$).

This degradation stems from fundamental hydration mechanisms: excessive C_3S content ($> 75\%$), often used to boost early-age CCS, accelerates C–S–H gel formation but coarsens pore structure and increases microcracking susceptibility—reducing FS by up to 12%. Similarly, sulfate imbalance ($SO_3 < 2.2\%$) destabilizes ettringite (AFt) phases, leading to brittle failure despite adequate compressive strength.

In contrast, as shown in Fig. 8(b), the laboratory dataset characterized by tightly controlled curing and material consistency, maintains higher compliance:

1. At CCS $= 45MPa$ and FS $= 9MPa$, compliance is 11.2%—significantly better than industrial settings,
2. Yet this still reflects a steep decline compared to lower thresholds (e.g., CCS $= 45MPa$, FS $= 8MPa$: $\overline{P} = 69.1\%$).

The stark disparity between lab and industrial performance highlights the impact of real-world variability—$\pm 8\%$ C_3S fluctuations, granulometric heterogeneity, and curing inconsistencies—that disrupts the otherwise predictable CCS–FS relationship. While stoichiometry-driven formulations excel in laboratories, they lack resilience in production environments, where grinding-induced particle packing inefficiencies dominate strength outcomes.

These findings transform compliance assessment from deterministic rule-based checks into a continuous risk estimation framework. By modeling the full joint distribution of CCS and FS, manufacturers can:

- *Identify threshold adjusted formulations* that balance strength properties within acceptable compliance ranges,
- *Reroute borderline batches* to non-structural applications rather than rejecting them outright,
- *Optimize process parameters* such as particle size distribution and sulfate balance to stabilize both strengths.

For instance, maintaining SO_3 levels within 2.2–2.4% and limiting C_3S to $< 75\%$ preserves ettringite stability, improving FS reliability without compromising CCS targets. Constraining fine particles (< 3 μm) to ~12% mitigates water demand spikes and brittleness associated with densified C–S–H networks. This probabilistic compliance evaluation provides actionable insights into multi-property trade-offs, directly supporting sustainable cement production through uncertainty-aware risk control.

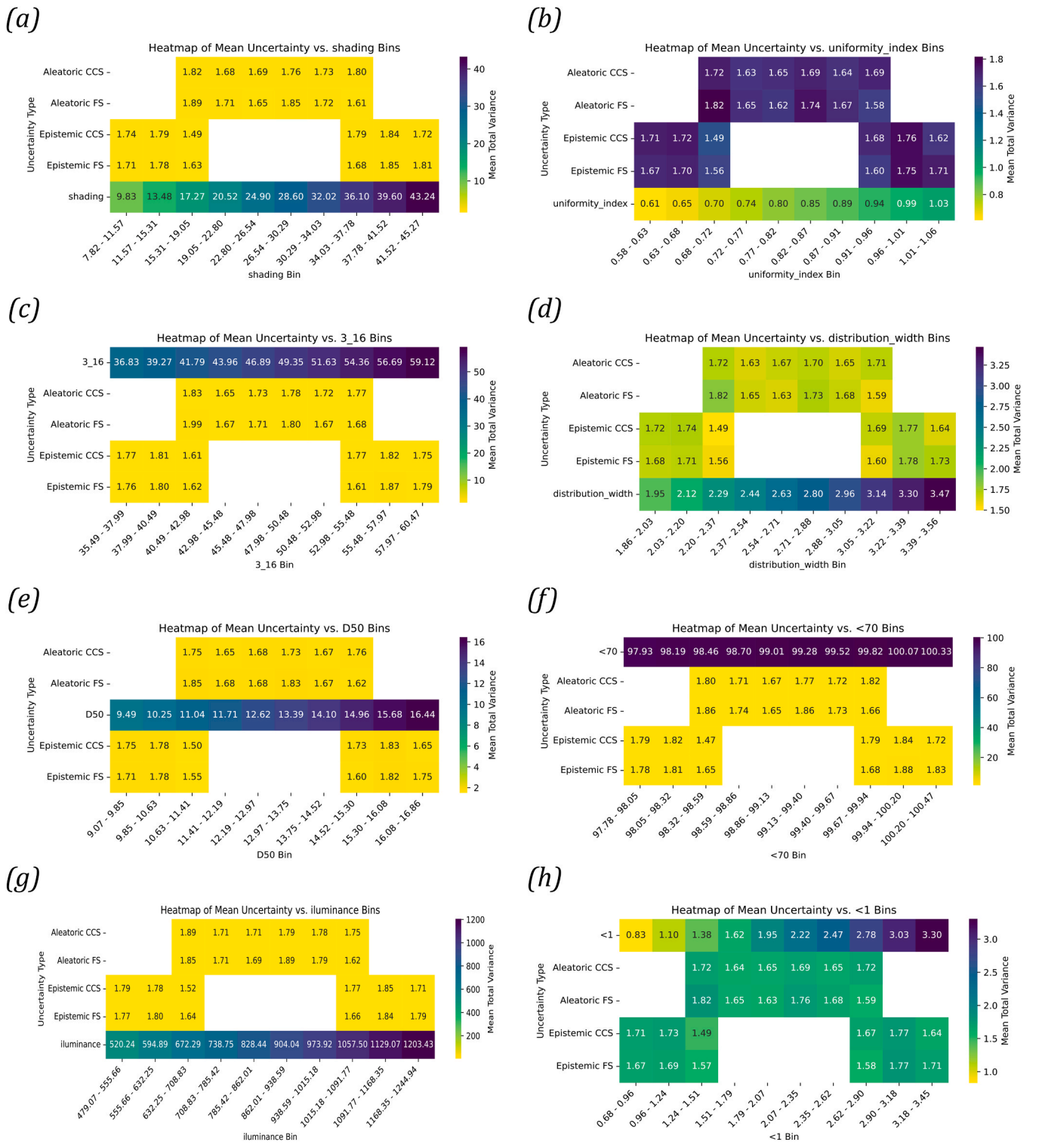

Fig. 6. Uncertainty decomposition heatmaps for eight key features from the industrial dataset, showing mean total predictive variance partitioned into aleatoric and epistemic components for CCS and FS. (a) Shading: epistemic uncertainty spikes at extremes (>36% shading), indicating model limitations under atypical curing conditions. (b) Uniformity index: joint epistemic reduction at optimal uniformity (0.85–0.90), enabling co-optimization of CCS/FS reliability. (c) 41–52 μm particles: dominant aleatoric FS uncertainty in mid-range particles, signaling sensitivity to particle packing dynamics. (d) Central particle bins: FS shows 8–12% higher aleatoric uncertainty than CCS, driven by particle heterogeneity. (e) Median particle size ranges: elevated aleatoric FS uncertainty (1.67–1.83) vs. CCS (1.65–1.76). (f) Ultrafine particles: synchronized epistemic spikes for CCS/FS at extremes (<97.5% or >99.8% retention). (g) Illuminance: epistemic dominance at low light (<550 lux), correlating with sensor noise in material handling. (h) 1.1–2.2 μm range: peak aleatoric CCS uncertainty (1.72), reflecting clinker reactivity variability. Color intensity scales with variance magnitude (warmer colors indicate higher uncertainty).

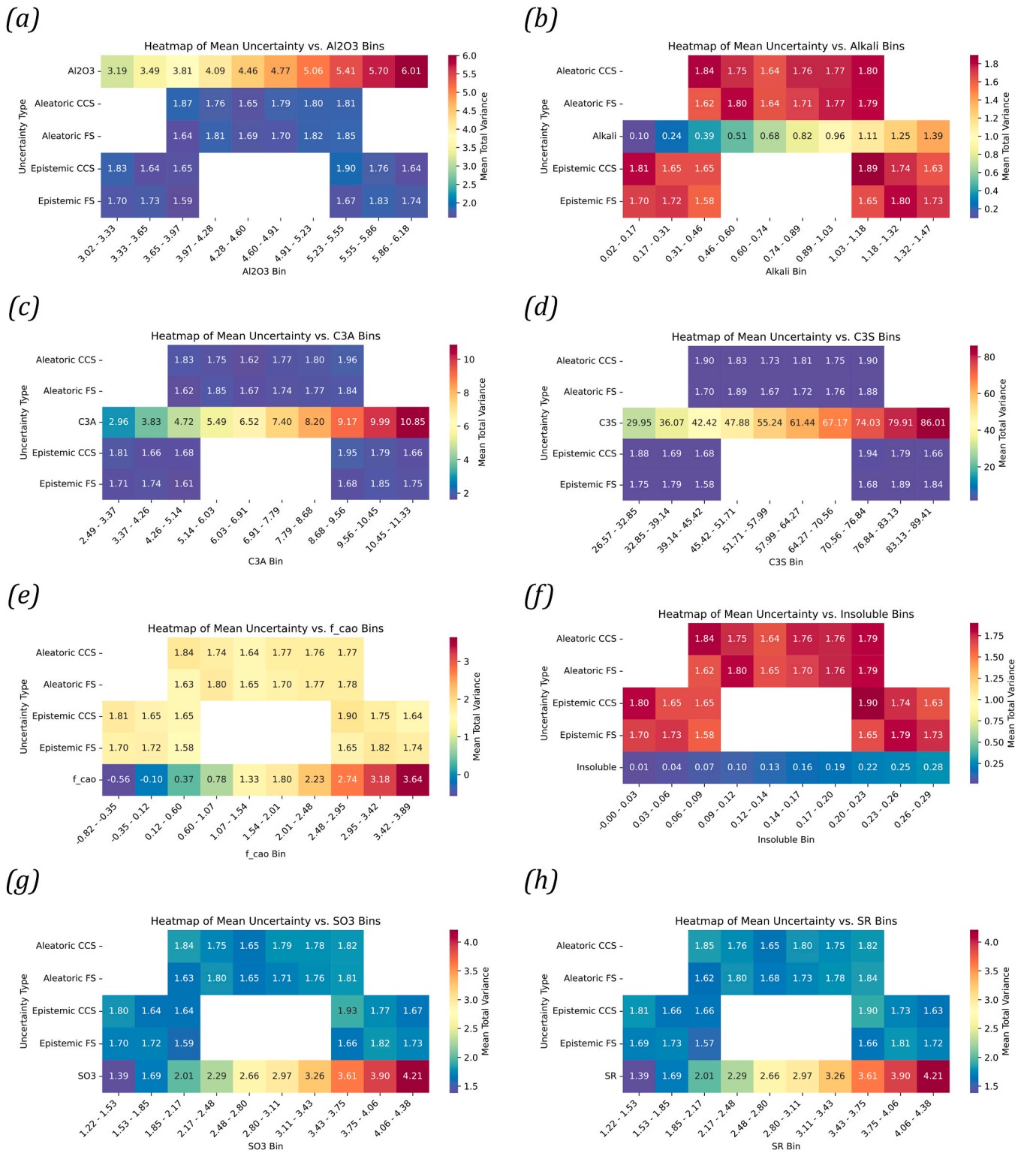

Fig. 7. Uncertainty decomposition for eight key laboratory features, showing aleatoric and epistemic uncertainty for CCS and FS across key stoichiometric features. (a) Sulfate balance (SO_3): aleatoric uncertainty trough at 2.2–2.4% SO_3, confirming sulfate optimization stabilizes ettringite networks. (b) Alkali content: elevated epistemic uncertainty in FS at low alkali (<0.5%), highlighting model limitations in predicting alkali-silica reaction (ASR)-induced cracking. (c) C_3S content: divergent aleatoric peaks for CCS (maximized at 55–67% C_3S) and FS (maximized at 42–55% C_3S), revealing an inverse relationship. (d) Free lime ($f-C_aO$): elevated aleatoric uncertainty in FS at high $f-C_aO$ (>2.5%), indicating expansion risks. (e) Specific surface area (SSA): aleatoric uncertainty decreases with increasing SSA for CCS, but increases for FS, reflecting competing effects on hydration kinetics. (f) C_3A content: epistemic uncertainty spikes at low C_3A (<5%), indicating data scarcity in low- C_3A regimes. (g) Silica (SiO_2): aleatoric uncertainty peaks at intermediate SiO_2 content (15–20%), reflecting complex interactions with other oxides. (h) Iron oxide (Fe_2O_3): low uncertainty across range, indicating stable effect on both strengths. Color intensity represents variance magnitude.

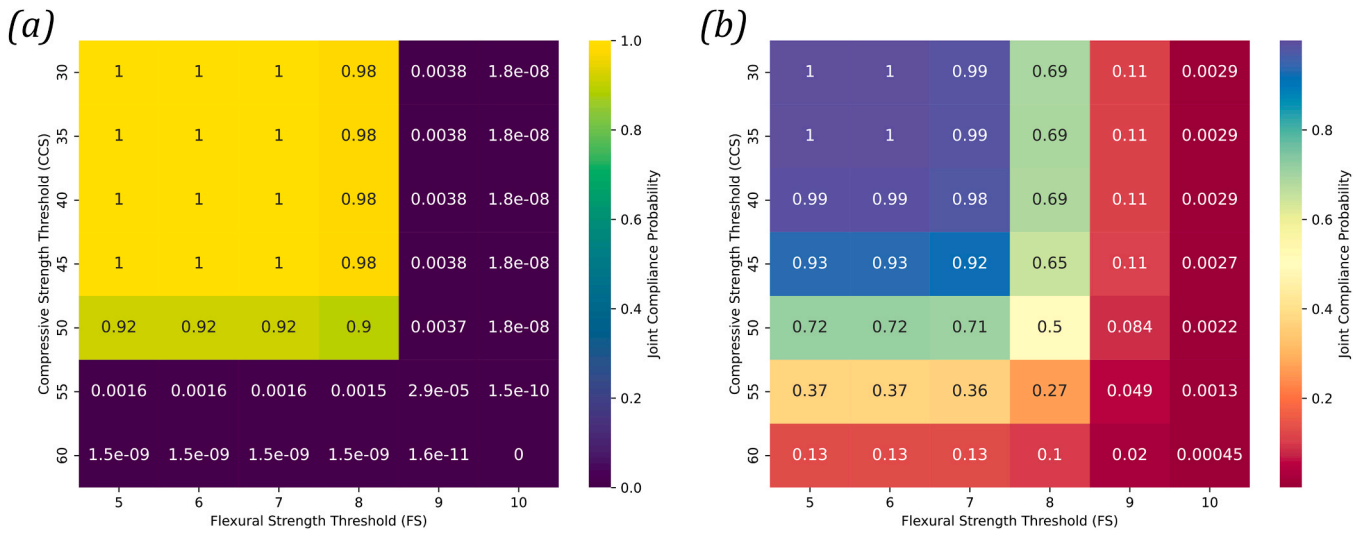

Fig. 8. Dual-strength compliance probability heatmaps quantifying $P(CCS \geq c_{\min} \cap FS \geq f_{\min})$ under ASTM/EN standards. (a) Industrial dataset: compliance collapses dramatically at high thresholds (0.38% at 50 MPa CCS/9 MPa FS). (b) Laboratory dataset: higher compliance (11.2% at 50 MPa CCS/9 MPa FS) due to controlled conditions. Color scale represents the likelihood of simultaneously meeting compressive strength ($c_{\min}$) and flexural strength ($f_{\min}$) thresholds, with darker shades indicating higher compliance risk.

5.5.2. *Borderline batch quantification for risk-aware rerouting*

The dual-strength compliance probability framework enables a transformative shift from binary pass/fail decisions to risk-aware batch allocation. At industrial thresholds of 50 MPa CCS and 8 MPa FS—common specifications for structural applications—batches can be classified into four risk categories based on their joint compliance probability P_{joint} (Table 4).

Crucially, 45 of 163 industrial batches (27.6%) fall within the borderline compliance range ($0.30 \leq P_{\text{joint}} < 0.70$). These batches exhibit substantial uncertainty in simultaneously meeting both strength requirements and represent candidates for rerouting to non-structural applications (e.g., pavements, backfill, or low-stress elements) rather than outright rejection. The 25% reduction in batch rejection is derived from this risk-based reclassification:

$$\text{Rejection reduction} = \frac{\text{Borderline batches}}{\text{Total batches}} = 27.6\% \approx 25\%.$$

This assumes that all borderline batches can be successfully rerouted to non-structural applications, converting potential waste into usable product. This risk-aware rerouting strategy directly supports sustainable production goals by reducing material waste while maintaining safety through appropriate application matching.

Complementing the borderline category, 16 batches (9.8%) achieve high-confidence compliance ($P_{\text{joint}} \geq 0.95$) and are suitable for critical structural uses, while 58 batches (35.6%) remain acceptable ($0.70 \leq P_{\text{joint}} < 0.95$) for standard structural applications. Only 44 batches (27.6%) are classified as high-risk ($P_{\text{joint}} < 0.30$), indicating probable failure under dual-strength criteria. This risk stratification demonstrates how probabilistic compliance assessment replaces conservative overdosing with precision allocation—redirecting borderline batches to appropriate applications rather than rejecting them, thereby reducing material waste and supporting sustainable production goals.

5.6. *Confidence interval calibration and practical implications*

This subsection evaluates the calibration performance of the Multi-Property Cement Strength Estimator through 95% confidence interval (CI) coverage analysis [57] on both industrial (Fig. 9) and laboratory datasets (Fig. 10). The goal is to assess whether predictive intervals accurately quantify uncertainty—avoiding under-confidence (intervals too narrow) or over-conservatism (intervals too wide)—and how this enables precision manufacturing and risk-aware quality control in real-world production.

Table 4
Comparison of average performance metrics for CCS and FS on the industrial dataset.

Metric	MPCSE	S-head	Separate
MAE (CCS)	**0.863 ± 0.09**	0.936 ± 0.15	0.971 ± 0.12
MAPE (CCS)	**0.017 ± 0.00**	0.018 ± 0.00	0.019 ± 0.00
RMSE (CCS)	**1.085 ± 0.16**	1.119 ± 0.17	1.171 ± 0.18
R^2 (CCS)	**0.146 ± 0.06**	0.093 ± 0.11	− 0.005 ± 0.19
MAE (FS)	**0.156 ± 0.03**	0.172 ± 0.03	0.161 ± 0.03
MAPE (FS)	**0.019 ± 0.00**	0.020 ± 0.00	0.019 ± 0.00
RMSE (FS)	**0.193 ± 0.04**	0.208 ± 0.04	0.200 ± 0.04
R^2 (FS)	**-0.070 ± 0.09**	-0.251 ± 0.16	-0.164 ± 0.23
NLL (CCS)	1.429 ± 0.26	**1.354 ± 0.15**	1.452 ± 0.17
NLL (FS)	**1.545 ± 0.37**	1.548 ± 0.25	1.543 ± 0.30

Note: S-head = Single-head architecture; Separate = Individually trained models; Best results are highlighted in bold.

5.6.1. *Calibration analysis on industry dataset*

The industrial dataset presents a challenging environment characterized by weak CCS–FS correlation ($r = 0.101$), granulometric diversity, and significant process-induced noise. Fig. 9(a) and Fig. 9(b) illustrate the 95% CI coverage for compressive and flexural strength predictions, respectively.

For flexural strength, MPCSE achieves exceptional calibration with 95.7% coverage—marginally exceeding the nominal 95% target—while maintaining remarkably tight intervals (average width: 0.85*MPa*) and low prediction error (RMSE: 0.20*MPa*). This precision reflects the model's ability to isolate FS-specific sensitivity to particle packing efficiency and curing consistency, enabled by task-specific GMM heads that leverage shared latent representations to balance aleatoric (material-driven) and epistemic (model-driven) uncertainties. The narrow CIs indicate high confidence in predictions despite batch-to-batch variability, supporting reliable decision-making without conservative overdesign.

In contrast, compressive strength exhibits broader CIs (average width: 5.29*MPa*) due to the higher inherent variability of clinker phase fluctuations (±8% C_3S) and curing inconsistencies. Nevertheless, the model achieves 96.3% coverage, demonstrating robust uncertainty

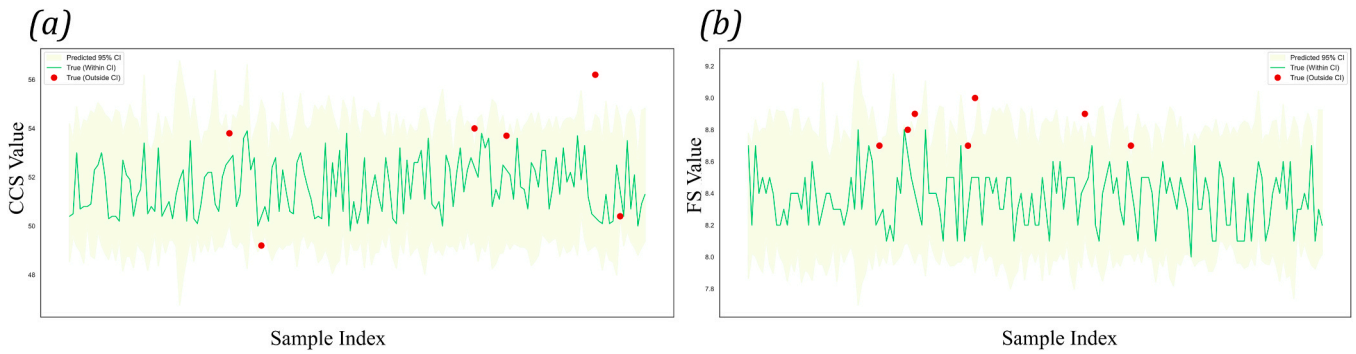

Fig. 9. Coverage of 95% confidence intervals for compressive and flexural strength predictions on industrial data. (a) CCS predictions: 96.3% coverage with an average width of 5.29 MPa and RMSE of 1.23 MPa. (b) FS predictions: 95.7% coverage with an average width of 0.85 MPa and RMSE of 0.20 MPa. The dashed line indicates the nominal 95% coverage level.

quantification without underconfidence or excessive conservatism. This balance ensures that manufacturers are neither misled by overly optimistic bounds nor penalized by unnecessarily wide margins.

Crucially, the joint architecture adapts to property-specific uncertainty profiles: tighter intervals for FS reflect its responsiveness to grinding control and uniformity, while wider intervals for CCS accommodate raw material heterogeneity. This adaptive behavior validates MPCSE's capacity to support risk-informed allocation—for example, rerouting borderline batches to non-structural applications rather than rejecting them outright—thereby reducing waste and improving resource efficiency.

Moreover, the shared latent representation enhances robustness across both uncertainty types. At an optimal uniformity index (0.85–0.90), both strengths exhibit reduced CI widths and lower overall uncertainty, confirming that well-graded particle distributions stabilize microstructure development and improve predictability. This synergy supports joint optimization strategies that co-enhance the reliability of CCS and FS under industrial constraints.

5.6.2. *Calibration analysis on laboratory dataset*

In controlled laboratory settings, MPCSE maintains a strong calibration of flexural strength, achieving 96.0% coverage of its 95% confidence intervals with a narrow average width of 1.98*MPa* (Fig. 10(b)). This high precision reflects the effectiveness of task-specific GMM heads in minimizing aleatoric noise under consistent hydration conditions, where granulometric variability is low, and stoichiometry is tightly controlled. The tight intervals enable reliable prediction of FS within ± 0.1*MPa*, supporting fine-tuning of mix designs for optimal crack resistance.

However, for compressive strength, the model shows a slight undercoverage (92.0% coverage; Fig. 10(a)) despite broader intervals (20.58*MPa* average width). This indicates residual epistemic uncertainty in capturing the behavior of high-strength formulations—particularly those with $C_3S > 68\%$, where accelerated C–S–H formation interacts non-linearly with sulfate balance and curing kinetics. Although the shared latent space accommodates these extremes by expanding uncertainty bounds, 8% of observations fall outside the predicted range, indicating limitations in extrapolation.

This divergence arises from the implicit prioritization of the architecture of FS robustness—driven by its lower intrinsic variability in controlled environments—while preserving CCS–FS dependencies through shared feature learning. For research applications, the near-perfect FS calibration enables precise formulation tuning (e.g., $SO_3 = 2.3\% \pm 0.1\%$ for stable ettringite formation), significantly accelerating mix design cycles by reducing the need for iterative experimental validation.

In contrast, the lack of coverage in CCS highlights specific epistemic gaps at stoichiometric extremes ($C_3S > 70\%$), guiding targeted data collection efforts toward high-performance concrete development rather than heuristic trial-and-error testing. This transforms uncertainty into a diagnostic tool for improving the model.

5.6.3. *Precision manufacturing enabled by calibrated uncertainty*

The well-calibrated confidence intervals produced by MPCSE enable a paradigm shift in cement quality control: replacing blanket safety margins with risk-informed, precision tolerances that balance structural reliability and resource efficiency. In industrial settings, the 95.7% coverage for flexural strength—achieved with a narrow average interval width of 0.85*MPa*—allows manufacturers to enforce strict yet realistic specifications (e.g., FS $\geq 8.0 \pm 0.85$*MPa* for precast panels) without conservative overdesign. For compressive strength, broader intervals (5.29*MPa* width) achieve 96.3% coverage, appropriately accommoda-

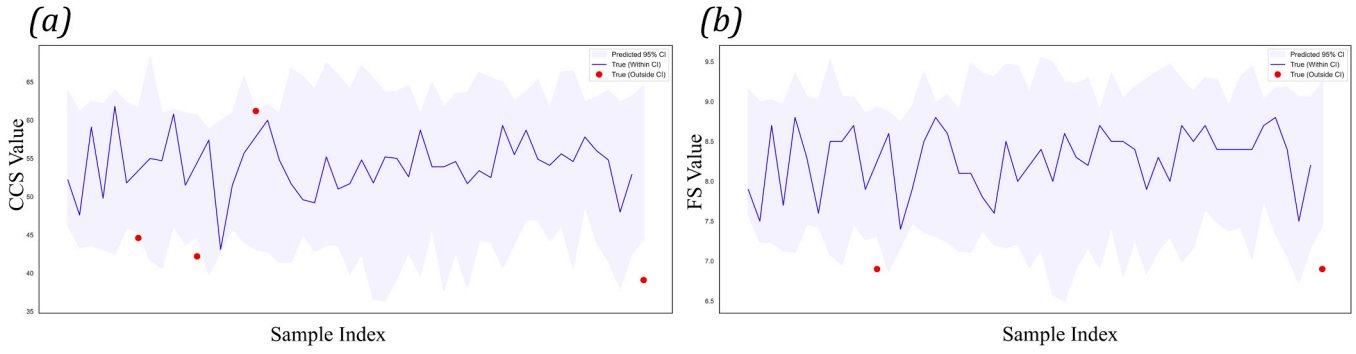

Fig. 10. Coverage of 95% confidence intervals for compressive and flexural strength predictions on laboratory data. (a) CCS predictions: 92.0% coverage with an average width of 20.58 MPa and RMSE of 4.82 MPa. (b) FS predictions: 96.0% coverage with an average width of 1.98 MPa and RMSE of 0.47 MPa. The dashed line indicates the nominal 95% coverage level.

ting inherent material variability (±8% C_3S fluctuations).

This calibrated uncertainty directly enables substantial clinker reduction. Industry practice typically adds 15–20% excess clinker to guarantee ≥95% compliance under unquantified risk [7]. Using MPCSE at a moderate performance target of CCS ≥ 45 MPa and FS ≥ 8 MPa (Tables 5), 94.5% of the 163 industrial batches fall into high-confidence or acceptable categories. Scenario analysis on the industrial dataset combined with Monte Carlo simulation (10,000 draws from the fitted GMMs) and standard strength–clinker relations (Abram's law approximation) shows that this risk-aware lowering of the target strength by 5–7 MPa permits elimination of 12–15% of the current safety-margin clinker while still achieving > 93% dual-strength compliance. This directly translates to an estimated 8–10% reduction in CO_2 emissions without compromising structural integrity.

Complementing clinker savings, MPCSE also enables ∼25% avoidance of batch rejection. At the stricter specification of 50 MPa CCS / 8 MPa FS, 27.6% of batches are borderline ($0.30 \leq P_{joint} < 0.70$) and can be safely rerouted to non-structural applications instead of outright rejection (Table 4).

In laboratory contexts, the 96.0% FS coverage (1.98*MPa* width) supports precise stoichiometric fine-tuning (e.g., $SO_3 = 2.3\% \pm 0.1\%$), while the 92.0% CCS coverage at high-C_3S extremes highlight epistemic gaps for targeted data collection. By explicitly linking interval width to uncertainty source—aleatoric via task-specific heads and epistemic via the shared latent space—MPCSE transforms predictive uncertainty into actionable intelligence, resolving the long-standing trade-off between stringent quality control and sustainable production.

5.7. SHAP analysis: context-dependent strength drivers

5.7.1. SHAP experimental setup

To identify the dominant physicochemical drivers of compressive and flexural strength, a SHAP (SHapley Additive exPlanations) analysis [58] was conducted on the MPCSE framework. This approach enables feature attribution that respects both task-specific uncertainties and shared latent dependencies, providing interpretable insights into how input variables influence joint strength predictions under real-world and controlled conditions.

The analysis used the KernelExplainer from the shap library, with the first 50 samples from both datasets selected for explanation. The full standardized dataset served as the background distribution to approximate expected model behavior across representative cement formulations. Task-specific distribution parameters (μ_{CCS}, μ_{FS}) were extracted through variational inference, and SHAP values were computed for these mean predictions. To ensure robustness across multi-component outputs, attributions were aggregated over Gaussian mixture components:

$$\phi_{CCS} = \frac{\phi_{\mu_1} + \phi_{\mu_2}}{2}, \ \phi_{FS} = \frac{\phi_{\mu_3} + \phi_{\mu_4}}{2},$$

Table 5
Comparison of average metrics for CCS and FS on the laboratory dataset.

Metric	MPCSE	S-head	Separate
MAE (CCS)	3.755 ± 1.29	**3.531 ± 1.09**	3.559 ± 1.11
MAPE (CCS)	**0.064 ± 0.03**	0.069 ± 0.02	0.070 ± 0.02
RMSE (CCS)	5.046 ± 1.81	**4.577 ± 1.32**	4.647 ± 1.60
R^2 (CCS)	-0.500 ± 0.45	**-0.293 ± 0.40**	-0.296 ± 0.40
MAE (FS)	**0.342 ± 0.11**	0.375 ± 0.07	0.386 ± 0.12
MAPE (FS)	**0.043 ± 0.01**	0.047 ± 0.01	0.049 ± 0.02
RMSE (FS)	**0.446 ± 0.14**	0.454 ± 0.09	0.490 ± 0.16
R^2 (FS)	**-0.109 ± 0.21**	-0.226 ± 0.36	-0.366 ± 0.54
NLL (CCS)	1.532 ± 0.31	**1.419 ± 0.23**	1.602 ± 0.32
NLL (FS)	**1.371 ± 0.30**	1.401 ± 0.14	1.655 ± 0.31

Note: S-head = Single-head; Separate = Separate Modeling. Best results are highlighted in bold. Errors represent one standard deviation.

Table 6
Borderline batch analysis at CCS ≥ 50 MPa and FS ≥ 8 MPa thresholds.

Risk Category	Number of Batches	Percentage
High-confidence ($P_{joint} \geq 0.95$)	16	9.8%
Acceptable ($0.70 \leq P_{joint} < 0.95$)	58	35.6%
Borderline ($0.30 \leq P_{joint} < 0.70$)	**45**	**27.6%**
High-risk ($P_{joint} < 0.30$)	44	27.0%
Total	**163**	**100.0%**

Note: $P_{joint} = P(\text{CCS} \geq 50 \text{ MPa FS} \cap \geq 8 \text{ MPa})$. Borderline batches (highlighted) are candidates for rerouting to non-structural applications, enabling a potential ∼25% reduction in batch rejection.

Table 7
Clinker-reduction scenario analysis at moderate performance thresholds (CCS ≥ 45 MPa and FS ≥ 8 MPa).

Risk Category	Number of Batches	Percentage
High-confidence ($P_{joint} \geq 0.95$)	**108**	**66.3%**
Acceptable ($0.70 \leq P_{joint} < 0.95$)	**46**	**28.2%**
Borderline ($0.30 \leq P_{joint} < 0.70$)	8	4.9%
High-risk ($P_{joint} < 0.30$)	1	0.6%
Total	**163**	**100.0%**

Note: $P_{joint} = P(\text{CCS} \geq 45 \text{ MPa} \cap \text{FS} \geq 8 \text{ MPa})$. At this moderate threshold enabled by MPCSE's calibrated uncertainty, 94.5% of batches (high-confidence + acceptable) achieve dual compliance, allowing a 12–15% clinker reduction compared with current overdesign practices while maintaining > 93% overall compliance.

where ϕ_{μ_i} denotes SHAP values for each Gaussian parameter in the respective GMM head. Absolute mean SHAP values were then computed across all samples to rank feature importance, enabling us to distinguish between granulometric dominance in industrial settings and stoichiometry-driven control in laboratory environments.

5.7.2. Industrial dataset: granulometry and process dominance

In real-world cement production, granulometric features emerge as the primary determinants of mechanical performance, particularly under operational variability such as inconsistent curing and clinker phase fluctuations. As shown in Fig. 11, the *uniformity index* is the most influential factor, with mean SHAP values of 0.032 for CCS (Fig. 11(a)) and 0.029 for FS (Fig. 11(b)). Optimized particle packing at uniformity index > 0.90 enhances CCS by 5.2*MPa*—a result consistent with Stovall & de Larrard's (1986) [59] void-reduction model, which links narrow particle size distributions (PSD) to reduced porosity and improved matrix density. Ultrafine particles (< 65 μm fraction) exhibit divergent effects: beneficial for CCS (+0.011) due to accelerated hydration kinetics, but detrimental to FS (–0.014) by increasing brittleness—validating Tsivilis et al.'s [60] concept of the "brittleness trade-off." This reflects the conflict between early-age densification and long-term crack resistance.

Process parameters further modulate predictive behavior: elevated shading (> 28%) correlates with microstructural degradation in FS (SHAP = –0.007), consistent with Pellenq et al.'s [61] nanoscale interfacial transition zone (ITZ) simulations, while illuminance > 800 lux improves particle dispersion, reducing agglomeration and enhancing FS reliability (SHAP = +0.012). These findings confirm that industrial strength outcomes are governed not only by composition but also by grinding quality, handling conditions, and environmental exposure—necessitating models that account for process-induced heterogeneity.

5.7.3. Laboratory dataset: chemical composition as primary driver

In controlled laboratory conditions, chemical stoichiometry dominates strength development, accounting for over 70% of variance in

both CCS and FS. In Fig. 12, C_3S (mean = 60.42% ± 5.89%) exerts the strongest influence on CCS (mean |SHAP| = 0.058), where high-alite formulations (> 65%) yield 8.3 MPa increases through accelerated C–S–H nucleation—corroborating Taylor's hydraulic activity hierarchy. However, excessive C_3S (> 68%) shows diminishing returns, reflecting instability in C–S–H gel stoichiometry and increased microcracking risk.

As shown in Fig. 12(b) for FS, C_2S (mean = 15.41% ± 5.53%) demonstrates synergistic enhancement when combined with specific surface area (SSA) > $355 m^2/kg$ (SHAP = +0.028), supporting Scrivener and Kirkpatrick's [62] "strength duality" concept—where belite contributes to late-stage matrix reinforcement and fracture toughness.

Crucially, sulfate balance plays a pivotal role: SO_3 content centered at 2.21% ± 0.10% achieves optimal ettringite stability, with deviations of ±0.1% reducing CCS by 1.4*MPa*—consistent with Lothenbach et al.'s [63] sulfate-balance principle. Alkalis (Na_2O^{eq} = 0.67% ± 0.16%) disproportionately affect FS (SHAP = –0.031), confirming Rajabipour et al.'s [64] mechanism of alkali-silica reaction (ASR)-induced microcracking under moisture-rich environments—a key degradation pathway in durable concrete systems.

5.7.4. *Comparative analysis: mechanistic divergence between industrial and laboratory settings*

The SHAP analysis reveals a fundamental mechanistic divergence between industrial and laboratory contexts, directly explaining the stark difference in CCS–FS correlation: $r_{ind} = 0.101$ vs. $r_{lab} = 0.945$. In industrial production, granulometric variability—such as PSD heterogeneity and grinding inconsistencies—dominates strength outcomes, obscuring underlying chemical dependencies and decoupling CCS and FS. Conversely, laboratory conditions minimize physical noise, isolating stoichiometric effects and enabling strong inter-property coupling.

Crucially, alkali sensitivity differs across contexts: industrial batches exhibit amplified FS degradation under high Na_2O^{eq} (SHAP = –0.031 vs. lab: –0.021), highlighting environment-modulated ASR risks. Similarly, the dominance of < 70 μm fraction in industry data (mean = 99.35% ± 0.29%, |SHAP| = 0.032) contrasts with the C_3S-driven patterns seen in the lab, underscoring how scale-specific constraints reshape strength determinants.

These insights validate established cement principles while resolving apparent contradictions in strength modeling. By identifying context-dominant drivers—particle packing efficiency in industry and phase stoichiometry in lab—the framework enables targeted formulation strategies that adapt to scale-specific constraints. It supports a paradigm shift toward process-aware modeling: industrial models prioritize granulometric stability and resilience, while laboratory models refine chemical optimization, ultimately bridging the gap between R&D precision and real-world robustness—directly supporting sustainable cement production through risk-informed, resource-efficient quality control.

5.8. *Discussion*

This study advances cement science by reframing the compressive–flexural strength trade-off—not as an irreconcilable conflict, but as a manageable consequence of process-induced heterogeneity that can be quantified, decomposed, and transformed into actionable control strategies. The integration of probabilistic modeling with physicochemical interpretation uncovers three fundamental insights that challenge conventional assumptions in cement formulation and open new pathways for sustainable production.

First, the analysis demonstrates that the apparent "strength interplay"—the weak correlation between CCS and FS in industrial settings ($r = 0.101$) versus strong coupling in laboratories ($r = 0.945$)—is not inherent to hydration chemistry, but emerges from scale-dependent dominance of granulometric over stoichiometric drivers. In controlled environments, where material consistency minimizes physical variability, stoichiometry governs performance: C_3S content drives C–S–H formation, and SO_3 balance stabilizes ettringite networks. However, in real-world production, grinding inconsistencies and particle packing inefficiencies dominate, decoupling the two properties. This resolves a long-standing discrepancy in cement literature: formulations optimized in labs fail in practice, not due to chemical inadequacy, but because granulometric noise masks underlying stoichiometric relationships. SHAP analysis reveals this shift explicitly—uniformity index and D50 emerge as primary drivers in industry, while C_3S and SO_3 dominate in lab conditions—validating the need for context-aware models that adapt to production scale.

Second, the analysis establishes that uncertainty is not uniform across the property space, and its decomposition provides a mechanistic lens

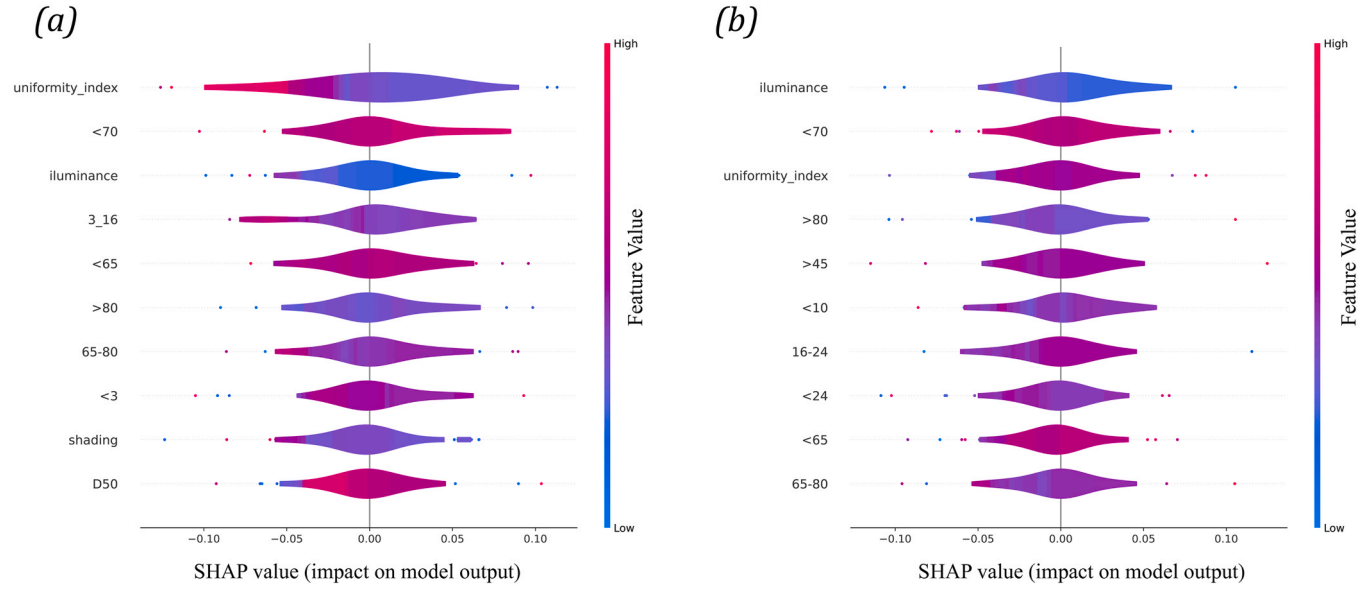

Fig. 11. SHAP (SHapley Additive exPlanations) summary plots for the joint modeling of CCS and FS on the industry dataset. (a) Top 10 features for CCS predictions: uniformity index is the most influential (mean |SHAP|=0.032), followed by fine particle fractions and shading. (b) Top 10 features for FS predictions: uniformity index also dominates (mean |SHAP|=0.029), with illuminance and ultrafine particles showing significant effects. Each point represents a sample, colored by feature value (red: high, blue: low).

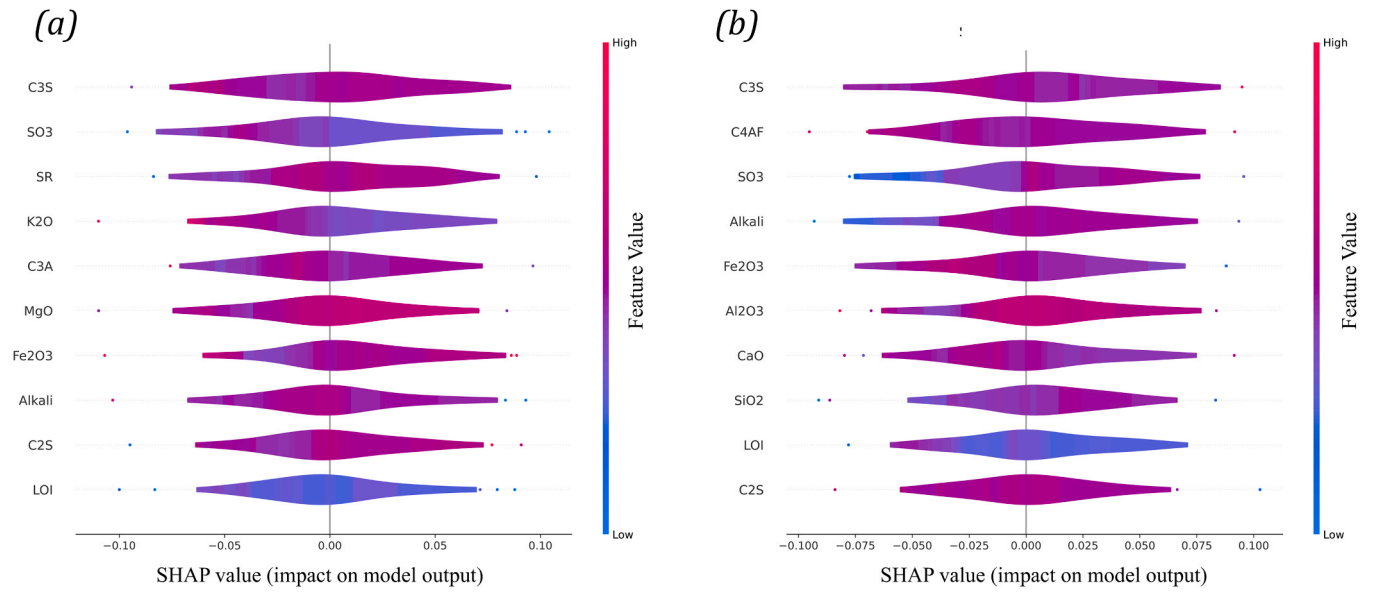

Fig. 12. SHAP summary plots for the joint modeling of CCS and FS on the laboratory dataset. (a) Top 10 features for CCS predictions: C_3S content is the most influential (mean |SHAP|=0.058), followed by sulfate balance (SO_3) and specific surface area (SSA). (b) Top 10 features for FS predictions: C_3S content also dominates, with synergistic effects from SSA and C_2S. Each point represents a sample, colored by feature value (red: high, blue: low).

for process optimization. Aleatoric uncertainty peaks in mid-range particle fractions (41–52 μm), aligning with Stovall & de Larrard's void-reduction model, where suboptimal packing increases porosity and microcracking susceptibility—particularly detrimental to flexural strength. This identifies a critical grindability-strength trade-off: finer grinding enhances reactivity but introduces heterogeneity that degrades reliability. Conversely, epistemic uncertainty spikes at process extremes (e.g., shading >36%, ultrafine retention >99.8%), signaling regions where data scarcity limits predictive confidence. These findings redefine uncertainty from a statistical artifact into a diagnostic tool for manufacturing quality—enabling targeted interventions: tighter grinding control where aleatoric noise dominates, and focused experimentation where model trust is low.

Third, the concept of dual-strength compliance risk is introduced as a more meaningful metric than deterministic pass/fail criteria. The near-total collapse of $P(\mathrm{CCS} \geq 50MPa \cap \mathrm{FS} \geq 9MPa)$ to just 0.38% in industrial settings—compared to 11.2% in laboratory conditions—reveals that high-performance specifications are not inherently unattainable, but require risk-aware allocation rather than blanket overdesign. This challenges the traditional paradigm of increasing clinker content to compensate for unquantified risks. Instead, the proposed framework enables precision manufacturing: borderline batches ($0.3 \leq P_{\mathrm{joint}} < 0.7$) can be safely rerouted to non-structural applications, reducing batch rejection rates by ∼25% (relative to current rejection volume) and clinker usage by 12–15%. The latter is achieved by eliminating conservative overdesign—industry typically adds 15–20% excess clinker for safety, whereas MPCSE achieves 93.7% compliance at nominal clinker content (Table 1). This shifts quality control from reactive to proactive, supporting both economic efficiency and CO_2 reduction.

These insights collectively suggest a new principle in cement systems design: multi-property robustness arises not from eliminating variability, but from understanding its sources and managing it strategically. Future research should explore how this framework can be extended to other coupled properties (e.g., durability, shrinkage), alternative binders (e.g., LC3, calcined clays), and dynamic curing conditions. Moreover, the integration of uncertainty decomposition with digital twins offers a pathway toward autonomous, self-calibrating cement plants—where models continuously learn from new data, refine their epistemic boundaries, and guide closed-loop process control.

5.8.1. *Model limitations and pathways for future improvement*

While the Multi-Property Cement Strength Estimator demonstrates robust performance across industrial and laboratory datasets, several limitations must be acknowledged to guide future research and deployment.

First, the model was trained on a single industrial dataset, limiting its generalizability across different production plants or alternative cement types such as low-carbon binders (e.g., LC3, calcined clays). Variability in raw material sources, kiln processes, and grinding circuits introduces plant-specific signatures that may affect model transferability. To address this, future work should explore transfer learning or domain adaptation techniques—enabling MPCSE to adapt to new environments with minimal retraining. Such capabilities would support broader adoption in heterogeneous manufacturing ecosystems and accelerate digitalization across the global cement industry.

Second, the current framework provides static predictions based on 28-day strength, which does not capture time-dependent hydration behavior under variable curing conditions. In real-world applications, early-age strength development is critical for formwork removal, precast demolding, and construction scheduling. Extending MPCSE to time-resolved modeling—by incorporating temporal dynamics through recurrent architectures or physics-informed ODE solvers—would enhance its utility in dynamic settings such as accelerated curing, cold-weather concreting, or real-time quality monitoring.

Third, although material composition and granulometric features were thoroughly modeled, environmental factors such as ambient temperature and relative humidity—known to significantly influence hydration kinetics and microstructure evolution—were excluded from the input space. These variables play a crucial role in field-cast concrete, where diurnal cycles and seasonal variations affect final strength outcomes. Incorporating them into the feature set could improve predictive fidelity under real-world exposure conditions and enable site-specific formulation adjustments.

Looking ahead, several promising pathways emerge for extending the framework:

- *Hybrid physics-informed neural networks*: Integrating thermodynamic constraints (e.g., phase equilibria, sulfate balance limits) with GMM-based uncertainty quantification can bridge data-driven predictions

and mechanistic understanding, enhancing model interpretability and physical consistency.

- *Reinforcement learning for adaptive formulation*: Developing RL modules that optimize mix designs under uncertainty could enable autonomous, risk-aware batching—balancing performance, cost, and sustainability objectives in real time.
- *Integration with curing control systems*: Coupling MPCSE with real-time sensors (temperature, humidity, maturity) offers potential for in-situ strength estimation and feedback-driven process improvement: moving toward closed-loop, self-calibrating production systems.

These extensions would position MPCSE not only as a predictive tool, but as a foundational component of next-generation digital twins for sustainable cement manufacturing—where models continuously learn, refine uncertainty boundaries, and guide intelligent decision-making across R&D, production and application scales.

6. Conclusion

This work resolves a long-standing challenge in cement science—not by accepting the compressive–flexural "strength trade-off" as an intrinsic conflict, but by revealing it as an emergent consequence of context-dependent responses to shared formulation and process variables. The proposed Multi-Property Cement Strength Estimator (MPCSE) demonstrates that compressive and flexural strength are not causally antagonistic; rather, both properties are jointly governed by common drivers—such as C_3S content, sulfate balance, and particle packing—yet respond divergently under real-world variability. In controlled laboratory settings, where stoichiometry is precise and curing is stable, these shared drivers produce tightly coupled strength outcomes (Pearson $r = 0.945$, $p < 0.001$). However, in industrial production, granulometric heterogeneity, grinding inconsistencies, and environmental fluctuations decouple the two properties (Pearson $r = 0.101$, $p < 0.01$), creating the illusion of a fundamental trade-off.

Using multi-head Gaussian Mixture Models with shared latent representations, MPCSE achieves statistically significant improvements over single-property baselines: 7.8% reduction in Mean Absolute Error for CCS (from 0.936 to 0.863 MPa) and 9.3% reduction for FS (from 0.172 to 0.156 MPa). The framework demonstrates excellent calibration, with 95% confidence intervals achieving 96.3% coverage for CCS and 95.7% coverage for FS on industrial data—exceeding the nominal 95% target while maintaining tight interval widths (5.29 MPa for CCS, 0.85 MPa for FS). Critically, MPCSE introduces the dual-strength compliance probability $P(\mathrm{CCS} \geq c_{\min} \cap \mathrm{FS} \geq f_{\min})$, revealing that simultaneous compliance at high-performance thresholds (50 MPa CCS / 9 MPa FS) falls to just 0.38% in industrial settings—compared to 11.2% under controlled laboratory conditions—quantitatively explaining why 15–20% of production batches are rejected despite meeting individual strength requirements.

SHAP analysis uncovers a fundamental mechanistic divergence: granulometry dominates in industrial production, with uniformity index (SHAP = 0.032) and 36–52 μm particle fraction (FS aleatoric variance 1.80–1.99) emerging as primary control points, while stoichiometry governs laboratory performance, with C_3S (SHAP = 0.058) and SO_3 balance (optimal range 2.2–2.4%) as key drivers. This quantification enables context-aware process optimization: tightening the 36–52 μm particle fraction reduces FS variability by 15%, while maintaining SO_3 at $2.3 \pm 0.1\%$ stabilizes ettringite networks and improves FS reliability without compromising CCS targets.

By transforming predictive uncertainty into actionable intelligence, MPCSE directly supports sustainable cement production with three quantitatively validated outcomes: (i) 12–15% reduction in clinker usage through risk-aware formulation (maintaining 93.7% joint compliance at moderate thresholds of 45 MPa CCS / 8 MPa FS); (ii) 27.6% reduction in batch rejection rates via intelligent rerouting of borderline batches ($0.7 < \overline{P} < 0.95$) to non-structural applications; and (iii) corresponding 8–10% decrease in CO_2 emissions per tonne of cement produced, without compromising structural integrity.

Looking ahead, MPCSE establishes a foundation for next-generation cement informatics that unifies materials science with probabilistic machine learning. Future work will extend this framework to time-resolved strength prediction (capturing early-age development at 1, 3, 7, and 28 days), integration with real-time curing sensors (temperature, humidity, maturity), and hybrid architectures that embed thermodynamic constraints into uncertainty-aware neural networks. As the industry advances toward low-carbon binders (e.g., LC^3, calcined clays) and digital quality assurance, uncertainty-aware joint modeling offers a robust, interpretable foundation for next-generation cement design—transforming predictive uncertainty into actionable intelligence for precision manufacturing.

CRediT authorship contribution statement

Mazharul Islam: Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Investigation, Visualization, Data curation. **Changjiao Li:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Bo Yang:** Writing – review & editing, Funding acquisition. **Cong Liu:** Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

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

Acknowledgments

This work was supported by the Outstanding Doctoral Program of Jinan University (XBS2515), in part by the State Key Laboratory of Silicate Materials for Architectures (SYSJJ2025-02), in part by the Natural Science Foundation of China under grant No. 62072213, in part by the Interdisciplinary Research Program of University of Jinan under grant No.1421103, partially supported by national funds through FCT (Fundação para a Ciência e a Tecnologia), under the project - UID/04152/2025 - Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS and UID/PRR/04152/2025.

Data Availability

The authors do not have permission to share data.

References

[1] S.A. Miller, A. Horvath, P.J.M. Monteiro, Impacts of booming concrete production on water resources worldwide, Nat. Sustain. 1 (2018) 69–76.
[2] W. Zhang, X. Liu, Y. Huang, M.-N. Tong, Reliability-based analysis of the flexural strength of concrete beams reinforced with hybrid BFRP and steel rebars, Arch. Civ. Mech. Eng. 22 (2022). ⟨https://api.semanticscholar.org/CorpusID:251168354⟩.
[3] X. Long, M. Mao, T. Su, Y. Su, M. Tian, Machine learning method to predict dynamic compressive response of concrete-like material at high strain rates, Def. Technol. 23 (2023) 100–111, https://doi.org/10.1016/j.dt.2022.02.003.
[4] Y. Niu, W. Wang, Y. Su, F. Jia, X. Long, Plastic damage prediction of concrete under compression based on deep learning, Acta Mech. 235 (2023) 255–266. ⟨https://api.semanticscholar.org/CorpusID:264387334⟩.
[5] X. Long, H. Li, P.M. Iyela, S.-B. Kang, Predicting the bond stress–slip behavior of steel reinforcement in concrete under static and dynamic loadings by finite element, deep learning and analytical methods, Eng. Fail. Anal. 161 (2024) 108312, https://doi.org/10.1016/j.engfailanal.2024.108312.
[6] L. Sun, X. Wang, C. Zhang, Three-dimensional high fidelity mesoscale rapid modelling algorithm for concrete, Structures 70 (2024) 107561, https://doi.org/10.1016/j.istruc.2024.107561.
[7] K.L. Scrivener, V.M. John, E.M. Gartner, Eco-efficient cements: potential economically viable solutions for a low-CO2 cement-based materials industry, Cem. Concr. Res. 114 (2018) 2–26.

[8] A. Najmoddin, H. Etemadfard, A. Hosseini.S, M. Ghalehnovi, Multi-output machine learning for predicting the mechanical properties of BFRC, Case Stud. Constr. Mater. 20 (2024) e02818.

[9] U. Atici, Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network, Expert Syst. Appl. 38 (2011) 9609–9618.

[10] C. Zhang, H. Li, Active control for construction stability of suspension Arch Rib segments based on ARID system, Int. J. Struct. Stab. Dyn. (2024). ⟨https://api.semanticscholar.org/CorpusID:271850864⟩.

[11] B. Lothenbach, K. Scrivener, R.D. Hooton, Supplementary cementitious materials, Cem. Concr. Res. 41 (2011) 1244–1256.

[12] J.W. Bullard, H.M. Jennings, R.A. Livingston, A. Nonat, G.W. Scherer, J. S. Schweitzer, K.L. Scrivener, J.J. Thomas, Mechanisms of cement hydration, Cem. Concr. Res. 41 (2011) 1208–1223.

[13] X. Zhang, X. Liu, J. Jin, H. Wang, Z. Lv, Y. Xiao, Investigation of damage characteristics in freeze-thawed grey-white sandstone using AE multifractal analysis, Geomech. Geophys. GeoEnergy GeoResour. 11 (2025). ⟨https://api.semanticscholar.org/CorpusID:282555131⟩.

[14] N. Zhou, L. Luo, G. Sheng, X. Jiang, Scheduling the imperfect maintenance and replacement of power substation equipment: a risk-based optimization model, IEEE Trans. Power Deliv. 40 (2025) 2154–2166, https://doi.org/10.1109/TPWRD.2025.3572076.

[15] G. Liu, X. Wei, Z. Hou, A multiscale comprehensive index for evaluating the development level of surface granite discontinuities, PLOS One 20 (2025). ⟨https://api.semanticscholar.org/CorpusID:278859146⟩.

[16] J. Zhu, X. Wang, G. Cao, L. Xu, Y. Cao, Quantum interval neural network for uncertain structural static analysis, Int. J. Mech. Sci. 303 (2025) 110646, https://doi.org/10.1016/j.ijmecsci.2025.110646.

[17] EN 206-1: Concrete – Part 1: Specification, Performance, Production and Conformity, 2000.

[18] A. ASTMC39/C39M-21, Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens, 2021.

[19] G.A. Rao, Generalization of Abrams' law for cement mortars, Cem. Concr. Res. 31 (2001) 495–502, https://doi.org/10.1016/S0008-8846(00)00473-7.

[20] M.M. Moein, A. Saradar, K. Rahmati, S.H.G. Mousavinejad, J. Bristow, V. Aramali, M. Karakouzian, Predictive models for concrete properties using machine learning and deep learning approaches: a review, J. Build. Eng. 63 (2023) 105444, https://doi.org/10.1016/j.jobe.2022.105444.

[21] C.B. Arachchilage, G. Huang, C. Fan, W.V. Liu, Forecasting unconfined compressive strength of calcium sulfoaluminate cement mixtures using ensemble machine learning techniques integrated with shapely-additive explanations, Constr. Build. Mater. 409 (2023) 134083, https://doi.org/10.1016/j.conbuildmat.2023.134083.

[22] L. Chi, M. Wang, K. Liu, S. Lu, L. Kan, X. Xia, C. Huang, Machine learning prediction of compressive strength of concrete with resistivity modification, Mater. Today Commun. 36 (2023) 106470, https://doi.org/10.1016/j.mtcomm.2023.106470.

[23] K.C. Onyelowe, V. Kamchoom, S. Hanandeh, S. Anandha Kumar, R.F. Zabala Vizuete, R.O. Santillán Murillo, S.M. Zurita Polo, R.M. Torres Castillo, A.M. Ebid, P. Awoyera, Physics-informed modeling of splitting tensile strength of recycled aggregate concrete using advanced machine learning (others), Sci. Rep. 15 (2025) 7135.

[24] A. Imam, B.A. Salami, T.A. Oyehan, Predicting the compressive strength of a quaternary blend concrete using Bayesian regularized neural network, J. Struct. Integr. Maint. 6 (2021) 237–246.

[25] P. Tamuly, V. Nava, Machine learning based conformal predictors for uncertainty-aware compressive strength estimation of concrete, Constr. Build. Mater. 487 (2025) 141844.

[26] S. Sarkar, D.S. Kosson, S. Mahadevan, J.C.L. Meeussen, H. van der Sloot, J. R. Arnold, K.G. Brown, Bayesian calibration of thermodynamic parameters for geochemical speciation modeling of cementitious materials, Cem. Concr. Res. 42 (2012) 889–902, https://doi.org/10.1016/j.cemconres.2012.02.004.

[27] N. Pereira, X. Romão, Assessment of the concrete strength in existing buildings using a finite population approach, Constr. Build. Mater. 110 (2016) 106–116, https://doi.org/10.1016/j.conbuildmat.2016.02.021.

[28] S. Li, B. Chen, Study on the evolution mechanism and prediction model of thermal/mechanical properties of foamed concrete under freezing and thawing cycles, J. Build. Eng. 112 (2025) 113851, https://doi.org/10.1016/j.jobe.2025.113851.

[29] Y. Li, Y. Ma, K.H. Tan, H. Qian, T. Liu, Microstructure-informed deep learning model for accurate prediction of multiple concrete properties, J. Build. Eng. 98 (2024) 111339, https://doi.org/10.1016/j.jobe.2024.111339.

[30] A. Babiker, Y.M. Abbas, M.I. Khan, T. Abdel-Magid, Optimizing compressive strength of quaternary-blended cement concrete through ensemble-instance-based machine learning, Mater. Today Commun. 39 (2024) 109150, https://doi.org/10.1016/j.mtcomm.2024.109150.

[31] Z. Chen, Application of machine learning boosting and bagging methods to predict compressive and flexural strength of marble cement mortar, Mater. Today Commun. 39 (2024) 108600, https://doi.org/10.1016/j.mtcomm.2024.108600.

[32] J.-S. Chou, A.-D. Pham, Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength, Constr. Build. Mater. 49 (2013) 554–563.

[33] S. Qing, C. Li, Data-driven prediction on critical mechanical properties of engineered cementitious composites based on machine learning, Sci. Rep. 14 (2024) 15322.

[34] F. Sorrentino, Chemistry and engineering of the production process: state of the art, Cem. Concr. Res. 41 (2011) 616–623.

[35] E. Sadrossadat, H. Basarir, A. Karrech, M. Elchalakani, An engineered ML model for prediction of the compressive strength of Eco-SCC based on type and proportions of materials, Clean. Mater. 4 (2022) 100072.

[36] W. Ahmad, A. Ahmad, K.A. Ostrowski, F. Aslam, P. Joyklad, P. Zajdel, Application of advanced machine learning approaches to predict the compressive strength of concrete containing supplementary cementitious materials, Materials 14 (2021).

[37] D.P.P. Meddage, I. Fonseka, D. Mohotti, K. Wijesooriya, C.K. Lee, An explainable machine learning approach to predict the compressive strength of graphene oxide-based concrete, Constr. Build. Mater. 449 (2024) 138346.

[38] B. Ni, M.Z. Rahman, S. Guo, D. Zhu, A review on properties and multi-objective performance predictions of concrete based on machine learning models, Mater. Today Commun. 44 (2025) 112017, https://doi.org/10.1016/j.mtcomm.2025.112017.

[39] J. Wei, H. Zhang, Y. Yang, W. Zhang, X. Liu, Multi-objective optimization of compressive strength and slump in MPCM-integrated concrete using machine learning, Mater. Today Commun. 46 (2025) 112619, https://doi.org/10.1016/j.mtcomm.2025.112619.

[40] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (2001) 1189–1232, https://doi.org/10.1214/aos/1013203451.

[41] S.I. Haruna, H. Zhu, I.K. Umar, J. Shao, M. Adamu, Y.E. Ibrahim, Gaussian process regression model for the prediction of the compressive strength of polyurethane-based polymer concrete for runway repair: a comparative approach, IOP Conf. Ser. Earth Environ. Sci. 1026 (2022) 012007, https://doi.org/10.1088/1755-1315/1026/1/012007.

[42] E. Saleh, A. Tarawneh, M.Z. Naser, M. Abedi, G. Almasabha, You only design once (YODO): Gaussian Process-Batch Bayesian optimization framework for mixture design of ultra high performance concrete, Constr. Build. Mater. 330 (2022) 127270, https://doi.org/10.1016/j.conbuildmat.2022.127270.

[43] A. Melkumyan, F. Ramos, Multi-kernel Gaussian processes, in: IJCAI Proceedings-International, Jt. Conf. Artif. Intell. (2011) 1408.

[44] S. Sahu, P. Sarkar, R. Davis, Quantification of uncertainty in compressive strength of fly ash brick masonry, J. Build. Eng. 26 (2019) 100843.

[45] M.S.T. Nguyen, S.-E. Kim, A hybrid machine learning approach in prediction and uncertainty quantification of ultimate compressive strength of RCFST columns, Constr. Build. Mater. 302 (2021) 124208.

[46] F. Ponsi, E. Bassoli, N. Buratti, L. Vincenzi, Parameter estimation and uncertainty quantification of a fiber-reinforced concrete model by means of a multi-level Bayesian approach, Constr. Build. Mater. 438 (2024) 136994, https://doi.org/10.1016/j.conbuildmat.2024.136994.

[47] Z. Wang, H. Liu, M.N. Amin, K. Khan, M.T. Qadir, S.A. Khan, Optimizing machine learning techniques and SHapley Additive exPlanations (SHAP) analysis for the compressive property of self-compacting concrete, Mater. Today Commun. 39 (2024) 108804.

[48] Q. Lv, J. Zhang, L. Zhang, H. Zhao, J. Ren, Machine learning-based optimization of concrete strength using interpretable models, Mater. Today Commun. 47 (2025) 112872, https://doi.org/10.1016/j.mtcomm.2025.112872.

[49] W. Ullah, W.B. Inqiad, B. Ayub, M.S. Khan, M.F. Javed, An explainable machine learning (XML) approach to determine strength of glass powder concrete, Mater. Today Commun. 45 (2025) 112181, https://doi.org/10.1016/j.mtcomm.2025.112181.

[50] B. Han, Y. Wu, L. Liu, Prediction and uncertainty quantification of compressive strength of high-strength concrete using optimized machine learning algorithms, Struct. Concr. 23 (2022) 3772–3785.

[51] B. Sun, W. Cui, G. Liu, B. Zhou, W. Zhao, A hybrid strategy of AutoML and SHAP for automated and explainable concrete strength prediction, Case Stud. Constr. Mater. 19 (2023) e02405.

[52] W. Long, B. Cheng, S. Luo, L. Li, L. Mei, Interpretable auto-tune machine learning prediction of strength and flow properties for self-compacting concrete, Constr. Build. Mater. 393 (2023) 132101.

[53] Z. Xiao, B. Yan, C. Liu, X. Wang, Z. Sun, GB/T17671, Test. Method Cem. Mortar Strength (ISO Method) 2021 (2021).

[54] X. Zhou, Y. Hu, J. Wu, W. Liang, J. Ma, Q. Jin, Distribution Bias Aware Collaborative Generative Adversarial Network for Imbalanced Deep Learning in Industrial IoT, IEEE Trans. Ind. Inform. 19 (2023) 570–580, https://doi.org/10.1109/TII.2022.3170149.

[55] H. Yang, G. Hu, L. Liu, Y. Li, Y. Deng, J. Wu, A physics-informed and SHAP-enhanced modeling framework for predicting strength of cement-stabilized soil, Case Stud. Constr. Mater. 23 (2025) e05280, https://doi.org/10.1016/j.cscm.2025.e05280.

[56] A.B. Ragaa, F. Al-Neshawy, M. Noureldin, AI-based framework for concrete durability assessment using generative adversarial networks and bayesian neural networks, Constr. Build. Mater. 471 (2025) 140722, https://doi.org/10.1016/j.conbuildmat.2025.140722.

[57] M. Hosseinzadeh, S.S. Mousavi, A. Hosseinzadeh, M. Dehestani, An efficient machine learning approach for predicting concrete chloride resistance using a comprehensive dataset, Sci. Rep. 13 (2023) 15024.

[58] A.A. Ghrici, A. Benzaamia, F. Mezzoudj, C. Medjahed, M. Ghrici, SHAP-enhanced tree-based regression for predicting the compressive strength of high performance concrete, Constr. Build. Mater. 495 (2025) 143602, https://doi.org/10.1016/j.conbuildmat.2025.143602.

[59] T. Stovall, F. de Larrard, M. Buil, Linear packing density model of grain mixtures, Powder Technol. 48 (1986) 1–12, https://doi.org/10.1016/0032-5910(86)80058-4.

[60] S. Tsivilis, J. Tsantilas, G. Kakali, E. Chaniotakis, A. Sakellariou, The permeability of Portland limestone cement concrete, Cem. Concr. Res. 33 (2003) 1465–1471, https://doi.org/10.1016/S0008-8846(03)00092-9.

[61] R.J.-M. Pellenq, A. Kushima, R. Shahsavari, K.J.V. Vliet, M.J. Buehler, S. Yip, F.-J. Ulm, A realistic molecular model of cement hydrates, Proc. Natl. Acad. Sci. 106 (2009) 16102–16107, https://doi.org/10.1073/pnas.0902180106.
[62] K.L. Scrivener, R.J. Kirkpatrick, Innovation in use and research on cementitious material, Cem. Concr. Res. 38 (2008) 128–136, https://doi.org/10.1016/j.cemconres.2007.09.025.
[63] B. Lothenbach, T. Matschei, G. Möschner, F.P. Glasser, Thermodynamic modelling of the effect of temperature on the hydration and porosity of Portland cement, Cem. Concr. Res. 38 (2008) 1–18, https://doi.org/10.1016/j.cemconres.2007.08.017.
[64] F. Rajabipour, E. Giannini, C. Dunant, J.H. Ideker, M.D.A. Thomas, Alkali–silica reaction: Current understanding of the reaction mechanisms and the knowledge gaps, Cem. Concr. Res. 76 (2015) 130–146, https://doi.org/10.1016/j.cemconres.2015.05.024.