Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Physics-Informed Machine Learning Framework for Probabilistic and Code-Compliant Seismic Bracing Selection in Reinforced Concrete Buildings

Vincent SAm Jebadurai S

pdf.vincent@lincoln.edu.my

Lincoln University College

Mohammad Israr

Lincoln University College

Article

Keywords: physics-informed machine learning, seismic retrofit, steel bracing, inter-storey drift, fragility analysis, uncertainty quantification, explainable artificial intelligence, reinforced concrete frames

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11656338/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Physics-Informed Machine Learning Framework for Probabilistic and Code-Compliant Seismic Bracing Selection in Reinforced Concrete Buildings

S.Vincent Sam Jebadurai[a,b], Mohammad Israr[c]

[a] *Postdoctotal Researcher, Lincoln University College, Malaysia, pdf.vincent@lincoln.edu.my*

[b] *Division of Civil Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India, vincent@karunya.edu +91 7904077614*

[C] *Maryam Abacha American University of Nigeria, Hotoro GRA, Kano State, Federal Republic of Nigeria isainik@yahoo.co.in*

Highlights

- We embed response-spectrum drift mechanics and code drift limits directly in the loss function of a machine learning model, so retrofit advice stays code-compliant by construction.
- We reframe bracing selection as a probabilistic multi-class decision rather than a deterministic threshold check, and we return a full probability vector over none, chevron, X- and heavy bracing.
- We show that in-distribution accuracy misleads: purely data-driven models score higher on accuracy and AUC, yet break the height–drift and hazard–drift trends outside their training range, while our model preserves them.
- We cut the mean squared mechanics residual by about 70% relative to a conventional neural network for a training-time penalty of only 12 to 18%, which shows that physical consistency is nearly free in this problem.
- We close the loop from prediction to risk: the predicted post-retrofit drift distributions shift all four fragility curves to the right and reduce cumulative collapse vulnerability by about 62%.

Abstract

Much of the world's reinforced concrete (RC) building stock predates modern seismic codes, so engineers must decide which frames to brace, and with what, across inventories too large for case-by-case nonlinear analysis. Machine learning promises that speed, but current models learn retrofit decisions from data correlations alone. They can therefore recommend an intervention

whose implied drift violates the very code limit that justifies retrofitting, they degrade without warning outside their training range, and they report one class label with no measure of confidence. None of these behaviours survives regulatory review, which is why such advice has not entered practice. We address this gap with a physics-informed machine learning framework that treats mechanics as a training constraint rather than a post-hoc check. We build a composite loss from four terms: data misfit, a response-spectrum drift residual that enforces equilibrium, a hinge penalty on the code drift limit, and a Monte Carlo variance term that stabilises the predicted probabilities. We train the model on 800 physics-consistent synthetic RC frames spanning four seismic zones, three soil classes and two irregularity indices. The model reaches 88 to 92% accuracy and a macro F1 of 0.86 to 0.90. Conventional models score marginally higher in distribution yet violate monotonic drift scaling under extrapolation, while our model cuts the mechanics residual by about 70%. Propagating the predicted post-retrofit drifts into lognormal fragility functions lowers cumulative collapse vulnerability by about 62%. We conclude that physical consistency, not accuracy, is the binding requirement for deployable retrofit models.

Keywords: physics-informed machine learning; seismic retrofit; steel bracing; inter-storey drift; fragility analysis; uncertainty quantification; explainable artificial intelligence; reinforced concrete frames

1. Introduction

1.1 The decision problem behind seismic retrofit

A large share of the existing reinforced concrete (RC) building stock was designed before current seismic provisions took effect. These frames share a recognisable set of deficiencies: light transverse reinforcement, unconfined joints, discontinuous infill that creates a soft ground storey, and plan or vertical irregularity that concentrates demand in a few members. Field surveys after the 2009 L'Aquila earthquake traced collapse in several mid-rise frames to this soft-storey mechanism [33]. Post-event assessment after the 2011 Lorca event showed that infill distribution governed whether damage stayed non-structural or became structural [10]. A dedicated review of irregular-structure research reached the same conclusion from the analytical side: configuration, more than nominal strength, sets the failure mode [11]. Early practice guidance recognised the scale of the problem and called for systematic prioritisation of older concrete frames, and current evaluation standards now codify the assessment procedure [3, 13].

The engineering question is therefore not whether these buildings are vulnerable. It is which of them to strengthen first, and with which system. That question has to be answered for thousands

of buildings at a time, under a fixed budget, with incomplete information about each structure. This is a screening and prioritisation problem, and it is the problem we address.

1.2 Why drift governs the choice

We use inter-storey drift as the decision variable throughout this work. Codes and performance guidelines set drift limits because drift correlates with the damage that drives repair cost and downtime. That damage includes cracking of partitions and plaster, failure of cladding and glazing, and the onset of inelastic rotation in beams and columns [6, 13]. Displacement-based design formalised this link by making a target displacement the starting point of design rather than its outcome [26]. Loss-assessment methodologies then convert drift into repair cost, downtime and casualty estimates, which is what makes drift the natural common currency between structural analysis and investment decisions [14].

Bracing is attractive in this setting because it acts directly on the quantity that governs the decision. Adding a concentric steel brace raises global lateral stiffness, shortens the fundamental period, and reduces drift demand. Experimental and analytical studies of braced retrofits confirm the mechanism and quantify its cost. Buckling-restrained braces applied to an existing non-ductile RC frame reduced drift and redistributed demand away from brittle columns [12], and a full performance assessment of a retrofitted existing structure confirmed the drift reduction under realistic ground motions [7]. A probabilistic assessment of low-ductility frames retrofitted with dissipative braces showed that the benefit is not uniform across damage states, and that it is largest where the un-retrofitted frame is most fragile [15]. Broader reviews of earthquake risk-reduction technology place bracing among the interventions with the most predictable structural effect per unit of investment [14].

1.3 What deterministic rules cannot do

The simplest way to use drift for retrofit selection is a threshold rule: compute a demand estimate, compare it with a limit, and retrofit if the limit is exceeded. We reject this approach for a specific reason. The demand estimate carries two distinct kinds of uncertainty, and a threshold rule collapses both to zero. Ground-motion variability is aleatory and irreducible; stiffness, damping, degradation and construction quality are epistemic and only partly knowable [20]. Sequences of ground motions rather than single records change fragility appreciably, which adds a further layer the single-record threshold ignores [1].

The consequence is asymmetric and expensive. A threshold rule applied to a building whose true demand sits just below the limit will under-recommend, and the building stays unstrengthened.

Applied to a building just above, it will over-recommend, and money that could have protected a genuinely deficient frame is spent instead. Probabilistic assessment removes this failure mode by carrying the uncertainty through to the decision. Incremental dynamic analysis established the reference procedure for this [32], and efficient fitting of lognormal fragility functions made the results usable at inventory scale [4]. Performance-based assessment then converts fragility into expected consequences [15]. What has been missing is a decision layer that inherits this probabilistic structure instead of discarding it.

1.4 Machine learning in earthquake engineering, and its unresolved problem

Machine learning entered earthquake engineering as a way to replace expensive analysis with fast approximation. State-of-the-art reviews document the breadth of that adoption across response prediction, damage detection and vulnerability assessment [37, 30], and a recent synthesis of surrogate modelling in structural engineering catalogues the methods now in routine research use [28]. The application to fragility is well developed. Machine-learned fragility curves reproduce those from incremental dynamic analysis at a fraction of the cost [22, 24]. Neural networks predict fragility for equipment and structural systems [34]. Gradient-boosted and ensemble models now estimate fragility, vulnerability and full performance state for RC structures directly [23]. Regional applications extend the same idea to inventories, using observed damage data to train damage classifiers at city scale [17], and post-earthquake safety screening has been posed as a supervised classification problem [40]. Explainability has followed: SHAP-based attribution now accompanies most credible structural applications [5], including interpretable drift-demand models for RC buildings [2] and explainable models aimed explicitly at assessment and retrofit [29].

This body of work shares one property that blocks deployment. The models are trained to fit data, and nothing in their formulation requires their outputs to be mechanically possible. Three specific failures follow, and we treat each as a design requirement.

First, unconstrained models violate mechanics outside their training range. This is not a hypothetical risk. A controlled study of data drift in seismic damage prediction for ageing structures showed that model performance degrades sharply, and silently, once the input distribution shifts away from the training set [8]. In our setting the shift is guaranteed, because inventory screening applies a model trained on one building population to another.

Second, unconstrained models can breach the code limit that motivates the retrofit. A classifier that recommends chevron bracing for a frame whose implied post-retrofit drift exceeds the

permissible value has produced a recommendation that no reviewer can accept, and standard training objectives contain nothing that penalises this [6, 3].

Third, most models return a single label. Retrofit selection near a drift threshold is genuinely ambiguous, and a point prediction hides that ambiguity. Surveys of uncertainty in deep networks show that ordinary softmax outputs are poorly calibrated and systematically overconfident [16], and the aleatory–epistemic distinction is precisely what a retrofit decision needs to see [20].

1.5 Physics-informed learning as the response

Physics-informed machine learning offers a direct remedy: it adds the governing equations to the training objective, so the model is rewarded for fitting the data and penalised for contradicting the physics. The formulation was introduced for forward and inverse problems in partial differential equations [27] and has since been consolidated into a general methodology [21, 9] with an established taxonomy of ways to inject scientific knowledge into learning [35, 36].

Structural engineering has begun to adopt it. A physics-guided convolutional network reproduced seismic response histories from limited data while respecting the equation of motion [39]. Physics-informed networks improved response prediction under sparse labelling [19] and estimated response from limited sensor channels [25]. A further physics-informed network identified nonlinear damage in a concrete bridge where a purely data-driven model could not [38]. Every one of these applications, however, targets **response prediction**: given a structure and an excitation, estimate a response quantity. None targets the **decision** that follows from that response. That is the gap we close.

1.6 Research gap and contributions

We state the gap precisely. Physics-informed models in structural engineering predict responses but do not select interventions. Machine learning models that do select interventions are purely data-driven, so they carry no guarantee of mechanical consistency, no enforcement of code limits, and no calibrated uncertainty. No existing framework, to our knowledge, embeds response-spectrum demand mechanics and code drift limits inside the training objective of a multi-class retrofit classifier, and then carries the resulting probabilistic output through to fragility and risk.

We make five contributions.

1. **A drift-constrained composite loss for decisions, not responses.** We combine data misfit, a response-spectrum mechanics residual, a hinge penalty on the code drift limit, and a Monte Carlo variance term in a single objective. The mechanics residual and the code penalty

act on the regression head, and the classification head shares its representation, so the class probabilities inherit mechanical consistency.

2. **Probabilistic multi-class bracing selection.** We return a probability vector over none, chevron, X- and heavy bracing instead of one label, which lets an engineer see how close a building sits to a category boundary.

3. **An explicit accuracy–consistency trade-off.** We show that conventional models win on in-distribution accuracy and AUC yet lose on extrapolation and calibration, and we argue that accuracy is therefore the wrong acceptance metric for this class of model.

4. **Uncertainty as a screening signal.** We reinterpret high predictive variance not as model weakness but as an operational trigger that flags which buildings need full nonlinear analysis.

5. **A mechanically coherent path from prediction to risk.** We propagate the predicted post-retrofit drift distributions into lognormal fragility functions and quantify the benefit with a single scalar, the area under the collapse fragility curve.

The remainder of the paper is organised as follows. Section 2 develops the mechanics formulation and the learning framework. Section 3 describes the dataset. Section 4 reports classification, extrapolation, explainability and uncertainty results. Section 5 interprets those results beyond what the numbers state directly. Section 6 carries the predictions through to fragility, risk and cost. Section 7 states the limitations, and Section 8 concludes.

2. Methodology

2.1 Mechanics formulation

We need a demand model that is cheap enough to evaluate inside a training loop and faithful enough to constrain the learner. We therefore idealise each RC frame as an equivalent single-degree-of-freedom oscillator and write the drift demand in closed form. We accept the approximation deliberately: the purpose of the residual is not to replace nonlinear analysis but to impose the correct scaling relations on the learner.

Fundamental period. We estimate the fundamental period from the empirical height relation adopted for RC moment-resisting frames:

$$\mathrm{T} = C_{\mathrm{h}} \cdot H^{0.75} \qquad (1)$$

where H is the total height in metres and $C_h = 0.075$ for RC moment frames. Empirical period formulas of this form were calibrated against measured building response and remain the basis of

code period estimates [18]. Equation (1) supplies the model with a physically grounded measure of global lateral flexibility that no purely statistical feature can replace.

Spectral demand. We approximate the design spectrum with a two-branch representation:

$$S_a(T) = 2.5 \cdot Z \cdot S \cdot I_{irr} \cdot T_c / (T + T_c) \quad (2)$$

where Z is the seismic zone factor, S the soil amplification factor, I_{irr} the irregularity amplification index, and T_c = 0.5 s the corner period. Equation (2) reproduces the plateau at short periods and the descending branch beyond T_c.

Drift demand. Assuming linear elastic response with effective lateral stiffness K_{eff}, we write the inter-storey drift ratio as

$$\theta = S_a(T) \cdot H / (R \cdot K_{eff}) \quad (3)$$

with R the response reduction factor. Bracing enters through the stiffness:

$$K_{eff} = K_{frame} + K_{brace} = K_{frame} (1 + \alpha) \quad (4)$$

where α is the stiffness ratio contributed by the bracing system. Equation (4) is the mechanical statement of what a brace does, and it is the reason drift is the correct decision variable: the intervention and the decision variable are linked by a single parameter.

Code constraint. Design codes cap the drift ratio:

$$\theta \leq \theta_{lim}, \text{ with } \theta_{lim} = 0.004 \quad (5)$$

We adopt 0.004 as the representative limit for damage control in RC frames [13, 6]. Equation (5) is a hard safety boundary in design practice, and we convert it into a differentiable penalty in Section 2.2.

2.2 Physics-informed learning framework

2.2.1 Dual-head architecture

We train one network with two output heads that share a common feature representation. The regression head predicts the drift ratio $\hat{\theta}$. The classification head predicts a probability vector $\mathbf{P} = (P_0, P_1, P_2, P_3)$ over the four bracing categories: none, chevron, X-bracing and heavy bracing. Sharing the representation is the mechanism that transfers physical consistency from the regression head to the classification head. The mechanics residual and the code penalty are written on $\hat{\theta}$, but because both heads read the same latent features, gradients from those terms reshape the representation that produces the class probabilities. The classifier therefore cannot

drift into a region where its implied displacement is mechanically impossible without paying a penalty.

This dual output is what turns a predictor into a decision-support tool. A single class label answers "which brace"; a probability vector plus a drift estimate answers "which brace, how confident, and how far from the limit".

2.2.2 Composite loss function

We minimise a weighted sum of four terms:

$$\mathrm{L} = L_{\text{data}} + \lambda_1 L_{\text{physics}} + \lambda_2 L_{\text{code}} + \lambda_3 L_{\text{uncertainty}} \qquad (6)$$

Each term encodes one requirement from Section 1.4.

(a) Data loss. We combine categorical cross-entropy for the bracing class with mean squared error for the drift ratio:

$$L_{\text{data}} = -\,\Sigma_\mathrm{i}\, y_{\mathrm{i}} \log(P_{\mathrm{i}}) + \mathrm{MSE}(\theta_{\text{true,}}\, \hat{\theta}) \qquad (7)$$

where y_i is the true class indicator and P_i the predicted probability of class i. The first term drives correct class assignment; the second keeps the drift estimate that feeds the physics and code terms accurate.

(b) Physics residual loss. We penalise disagreement between the predicted drift and the drift implied by Equation (3):

$$L_{\text{physics}} = \| \hat{\theta} - S_{\mathrm{a}}(\mathrm{T}) \cdot \mathrm{H} \,/\, (\mathrm{R} \cdot K_{\text{eff}}) \|^2 \qquad (8)$$

Equation (8) is the core of the framework. It converts a governing relation into a soft constraint, in the manner established for physics-informed networks [27, 21]. We use a soft residual rather than a hard projection for a specific reason: Equation (3) is itself an approximation, so forcing exact satisfaction would transfer the approximation error directly into the prediction. The soft form lets the data correct the mechanics where the two disagree, while still forbidding large violations.

(c) Code constraint penalty. We add a one-sided hinge penalty on the drift limit:

$$L_{\text{code}} = \max(0, \hat{\theta} - \theta_{\text{lim}}) \qquad (9)$$

Equation (9) is inactive whenever the prediction complies and grows linearly once it does not. This asymmetry matters. A two-sided penalty would push predictions toward the limit from both directions and bias the model against genuinely stiff buildings. The hinge form penalises only the

direction that carries regulatory risk, so compliance becomes a property of the trained model rather than a filter applied afterwards.

(d) Uncertainty regularisation. We perturb the inputs with multiplicative Gaussian noise,

$$x' = x\,(1 + \varepsilon), \varepsilon \sim N(0, \sigma) \qquad (10)$$

evaluate the model over the perturbed samples, and penalise the variance of the predicted class probabilities:

$$L_{\text{uncertainty}} = \text{Var}(P_i) \qquad (11)$$

Equation (11) does not suppress uncertainty; it suppresses *instability*. The distinction is important. A model whose class probabilities swing wildly under a 10% input perturbation is unusable, because the engineer cannot tell whether a change in recommendation reflects the structure or the noise. By penalising variance during training, we obtain a model whose remaining predictive spread reflects genuine ambiguity in the decision rather than numerical fragility. This is the epistemic component of uncertainty that retrofit decisions must expose [20, 16].

2.2.3 Weight selection and training

The weights λ_1, λ_2 and λ_3 set the balance between fitting the data and respecting the mechanics. We select them by a coarse grid search on the validation split, choosing the combination that minimises validation data loss subject to a ceiling on the mean physics residual. This ordering is deliberate. We treat physical consistency as a constraint to satisfy and accuracy as the quantity to optimise within it. That inverts the usual priority and reflects the acceptance criteria of a design reviewer rather than those of a data-science benchmark. Section 5.2 shows that this choice costs very little accuracy.

[Author to confirm before submission: network depth and width, activation function, optimiser and learning rate, number of epochs, batch size, the selected λ_1, λ_2, λ_3 values, the number of Monte Carlo samples per prediction, the value of k in k-fold cross-validation, the noise standard deviation σ, and the values adopted for R*, the base frame stiffness and the range of* α*.]*

2.3 Dataset generation

2.3.1 Why a synthetic dataset

No experimental dataset of the required size exists. Measured drift responses paired with verified retrofit outcomes are available for tens of buildings, not hundreds, and the covariate ranges in those samples are narrow. We therefore generate a physics-consistent synthetic dataset, following

the practice now standard in surrogate modelling for structural engineering [28]. We are explicit about what this buys and what it costs. It buys controlled coverage of the parameter space and known ground truth, which lets us test extrapolation deliberately rather than by accident. It costs external validity, which we address in Section 7.

2.3.2 Parameter sampling

We sample 800 RC frames over the ranges in Table 1. The ranges span low-rise to high-rise frames, all four seismic zones, three soil classes, and both soft-storey and torsional irregularity.

Table 1. Sampled structural and site parameters.

Parameter	Range	Role
Number of storeys	3–15	Height and period driver
Storey height	3.0–3.5 m	Height and period driver
Seismic zone	II–V	Hazard intensity (Z)
Soil type	I–III	Site amplification (S)
Soft-storey index	0.8–1.3	Vertical irregularity (I_{irr})
Torsional index	0.9–1.2	Plan irregularity (I_{irr})
Construction quality factor	0.85–1.10	Stiffness and strength modifier

2.3.3 Uncertainty injection

We perturb the deterministic samples so that the class boundaries overlap, because a dataset with clean boundaries would let a classifier succeed without learning anything transferable. We apply ±5% uncertainty to height, ±7% to period, ±10% to stiffness, and Gaussian noise to the demand. These perturbations represent survey error, modelling assumption and construction variability respectively, and they are the sources that dominate epistemic uncertainty in existing-building assessment [20].

2.3.4 Bracing class definition

We label each frame from its drift ratio using the thresholds in Table 2. The thresholds follow the graded logic of performance-based assessment. Below the first level the frame performs acceptably without intervention. The intermediate bands call for progressively stiffer systems. Above the top level the frame requires substantial strengthening to prevent collapse.

Table 2. Bracing class definition by drift ratio.

Class	Drift ratio θ	Recommendation	Mechanical rationale
0	$\theta < 0.0012$	No bracing	Adequate inherent stiffness
1	$0.0012 \leq \theta < 0.0022$	Chevron bracing	Moderate stiffness gain, preserves openings
2	$0.0022 \leq \theta < 0.0035$	X-bracing	Large stiffness gain and

			energy dissipation
3	$\theta \geq 0.0035$	Heavy bracing	Stability control, collapse prevention

Because we apply the stochastic perturbations of Section 2.3.3 before labelling, the classes overlap near every threshold. The classification problem is therefore not trivially separable, and the model must learn probabilistic transitions between retrofit categories rather than crisp boundaries. This is the intended behaviour: it mirrors the real ambiguity an engineer faces when a demand estimate lands close to a limit. The choice of bracing configuration itself is grounded in the retrofit literature, where concentric and dissipative bracing schemes are selected on exactly this stiffness-versus-serviceability trade-off [12, 7, 15].

2.3.5 Preprocessing and evaluation protocol

We standardise the continuous features so that no variable dominates through scale alone, and we encode soil type and seismic zone as ordered numerical variables to preserve their monotonic physical meaning. We split the data 75:25 into training and test subsets, and we run k-fold cross-validation on the training portion to reduce variance in the performance estimates. We report accuracy, macro-averaged F1 and per-class AUC, and we add two metrics that conventional benchmarks omit: the mean squared physics residual of Equation (8), and the predictive variance of Equation (11). Reporting the first two alone would hide exactly the behaviour we set out to fix.

2.4 Baseline models

We compare against three baselines: Random Forest, Gradient Boosting, and a conventional feed-forward artificial neural network with the same architecture as our model but trained on L_{data} only. The last comparison isolates the effect of the physics, code and uncertainty terms, since the only difference between it and our model is Equation (6).

3. Results

3.1 Dataset characteristics

The 800 sampled frames produced drift values with realistic dispersion once we injected the uncertainties of Section 2.3.3. The class distribution stayed approximately balanced: class 0 at 24–28%, class 1 at 25–30%, class 2 at 22–27% and class 3 at 18–23%. We confirmed overlap near every drift threshold, which establishes that the classification task retains genuine engineering ambiguity rather than reducing to a lookup on a single variable.

3.2 Classification performance

Our PIML model reached 88 to 92% overall accuracy with a macro F1 of 0.86 to 0.90 and balanced precision across the four classes. Random Forest and Gradient Boosting reached 90 to 94% accuracy on the in-distribution test set. Table 3 summarises the comparison.

Table 3. Model comparison. Accuracy and AUC are in-distribution; residual and extrapolation behaviour are the deployment-relevant metrics.

Metric	PIML (ours)	Random Forest / Gradient Boosting	Conventional ANN
Overall accuracy	88–92%	90–94%	comparable to RF/GB
Macro F1	0.86–0.90	—	—
Per-class AUC	0.90–0.94	≈0.95	≈0.95
Mean squared physics residual	baseline	not enforced	≈3.3× higher than PIML
Monotonic θ–H trend under extrapolation	preserved	violated	violated
Probability transitions at class boundaries	smooth	abrupt	abrupt
Training time	+12–18% vs ANN	lower	reference
Inference time per building	<5 ms	<5 ms	<5 ms

We read the accuracy comparison as a diagnostic rather than a defeat. The tree ensembles fit statistical patterns inside the training envelope very effectively, and they are entitled to their higher score there. What that score cannot tell us is whether the fitted mapping remains valid when the input moves. Section 3.3 tests exactly that.

The confusion matrix is more informative than the aggregate score. Misclassifications were not scattered; they concentrated at the drift-threshold boundaries, and the chevron-to-X-bracing transition carried the highest rate. These are marginal performance states in which a small change in stiffness or demand genuinely changes the appropriate intervention. A crisp classifier that resolved these cases confidently would be reporting a certainty that the mechanics does not support. Interpreted this way, the localisation of error at class boundaries is evidence that the model has learned where the decision is genuinely undetermined.

3.3 Physics residual and extrapolation

The mean squared physics residual of Equation (8) fell by approximately 70% relative to the conventional ANN. Because the two models share an architecture and differ only in the loss, we attribute the reduction entirely to the added terms in Equation (6).

We then tested extrapolation directly. We raised building heights by up to 20% above the training limit, pushed the seismic zone factor beyond its calibrated range, and increased the irregularity indices to represent more severe configurations. Our model kept drift predictions bounded and

monotonically increasing with height, consistent with Equations (1) to (3). The purely data-driven models did not. Under high zone factors they predicted *decreasing* drift for taller buildings, which contradicts the height dependence in Equation (3) and cannot occur in a real frame.

This failure mode matters more than the accuracy gap. A controlled study of data drift in seismic damage prediction reported the same degradation pattern for ageing structures and, critically, showed that it arrives with no warning signal in the standard metrics [8]. A model that scores 94% on a held-out split and then inverts the height–drift relation on a taller building is more dangerous than a model that scores 90% and stays monotonic, because the first will be trusted.

3.4 Probabilistic behaviour and calibration

Figure 1 shows the receiver operating characteristic curve for the X-bracing class. Our model reached an area under the curve of 0.90 to 0.94 across the bracing classes, against approximately 0.95 for the conventional models.

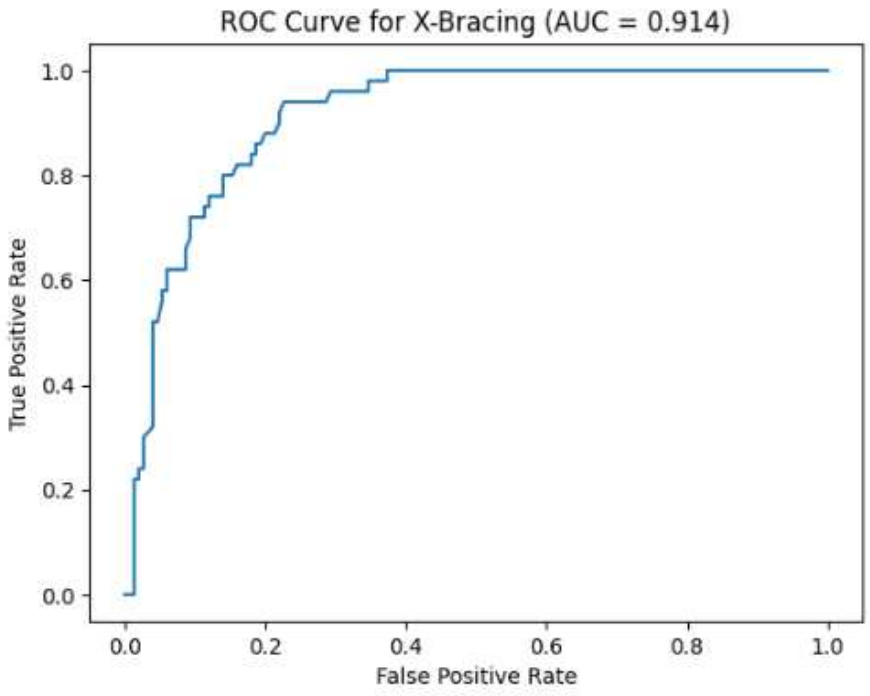


Figure 1. Receiver operating characteristic curve for the X-bracing class.

The calibration behaviour reverses this ranking. Near class boundaries our model produced smooth probability transitions that track the gradual change in structural response as drift approaches a threshold. The purely data-driven models produced abrupt transitions and returned high-confidence predictions for cases in which a small input change would alter the recommendation. Surveys of uncertainty in deep networks document this overconfidence as a

general property of unregularised softmax outputs [16]. The aleatory–epistemic decomposition explains why it is fatal here: the model reports confidence about a quantity on which it holds no information [20].

We therefore argue that AUC is the wrong figure of merit for this task. AUC rewards ranking ability at a fixed operating point on a fixed distribution. A retrofit recommender is used off-distribution, and its output feeds a decision whose cost is asymmetric. A slightly lower AUC accompanied by honest probabilities is the better instrument.

3.5 Explainability

Figure 2 reports the feature importance obtained from SHAP attribution. The irregularity index ranked highest at approximately 0.25, followed by soil type at 0.24, seismic zone factor at 0.21, building height at 0.18, and number of storeys at 0.12.

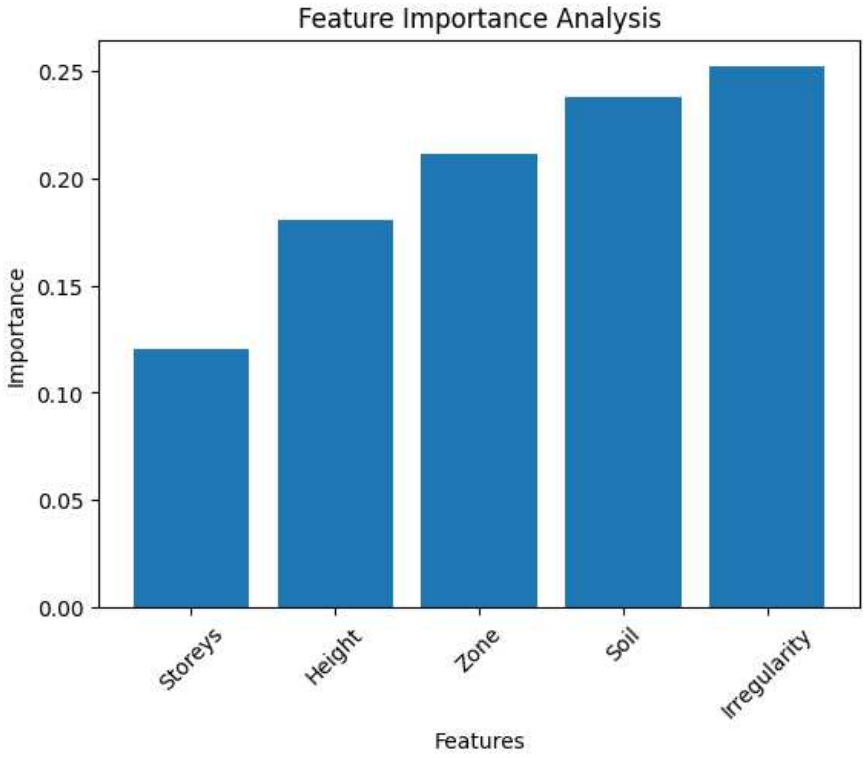


Figure 2. SHAP-based feature importance for bracing class prediction.

The ranking is consistent with Equations (1) to (3). Irregularity and soil amplification enter the spectral demand multiplicatively through I_{irr} and S, the zone factor scales it linearly through Z, and height enters twice, once through the period in Equation (1) and once directly in Equation (3). The consistency between the attribution and the governing equations is the evidence that the model learned mechanics rather than coincidence, and it is the reason the framework can be

defended to a reviewer. We use attribution as a physical audit rather than as a descriptive summary, following the direction set by interpretable drift-demand models [2] and explainable retrofit models [29]. The method itself rests on the attribution framework consolidated in the explainable-AI literature [5].

We draw two further inferences in Section 5.

3.6 Uncertainty quantification

We applied Monte Carlo perturbations at ±10% parameter noise to assess predictive stability. Figure 3 shows the resulting distribution of prediction probability standard deviations.

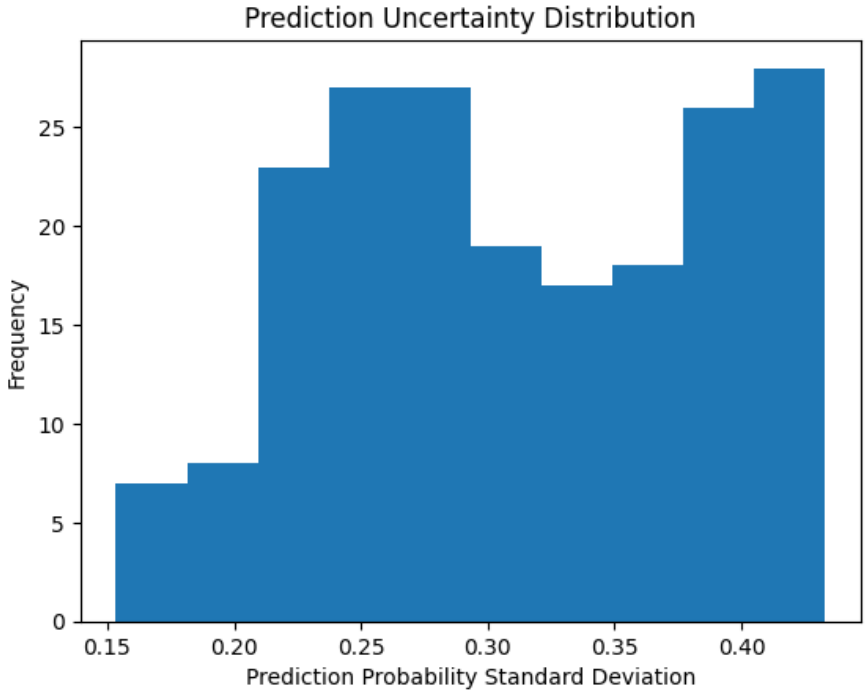


Figure 3. Distribution of predictive standard deviation under ±10% Monte Carlo input perturbation.

Three patterns emerged. Predictive variance increased near class thresholds. Confidence intervals widened for mid-rise buildings in zones IV and V, where demand is high and the period sits near the spectral corner. Heavy-bracing recommendations carried the highest confidence, because frames in that class are far from any boundary in the drift coordinate.

The distribution in Figure 3 is bimodal, and we regard that shape as a useful result rather than an artefact. It separates the inventory into a population the model resolves confidently and a population it does not. Section 5.3 develops the operational consequence.

3.7 Computational cost

The added loss terms increased training time by 12 to 18% relative to the conventional ANN, because the physics residual must be evaluated at every iteration. The overhead is confined to training. After training, inference costs under 5 ms per building, which makes portfolio-scale screening of tens of thousands of buildings a matter of minutes. This cost profile is what makes the framework usable for the prioritisation problem posed in Section 1.1.

4. Discussion: interpreting the findings beyond the metrics

4.1 Accuracy is not the acceptance criterion

The central finding of this study is not that our model is more accurate. It is not. Random Forest and Gradient Boosting beat it by two to four accuracy points and by about 0.01 to 0.05 in AUC on the in-distribution test set. We interpret that result as a demonstration that the standard benchmark does not measure what a retrofit model must deliver.

The argument is as follows. In-distribution accuracy measures how well a model reproduces labels drawn from the same population as its training data. A retrofit recommender never operates on that population. It is trained on one building inventory and applied to another, in a different city, with different construction practice and a different hazard level. Under that shift the tree ensembles inverted the height–drift relation, and nothing in their accuracy or AUC score signalled the failure. The same silent degradation has been documented for seismic damage classifiers subjected to data drift [8]. A model can therefore be simultaneously more accurate and less usable, and our results are a concrete instance of that.

This reframes the contribution of physics-informed learning in structural engineering. The literature has largely presented physical constraints as a way to improve accuracy when data are scarce [21, 35, 39]. Our results support a different and stronger claim: even where data are abundant and accuracy is not the binding constraint, physical constraints purchase *admissibility*. They convert a model whose failures are silent into one whose failures are bounded by mechanics. For a safety-critical decision that is the property that determines whether the model can be used at all.

4.2 Physical consistency costs almost nothing

We quantify the trade-off explicitly. The physics, code and uncertainty terms cut the mean squared mechanics residual by about 70%, cost 12 to 18% additional training time, and cost two to four points of in-distribution accuracy. Inference cost is unchanged.

We regard this as the practical finding most likely to influence adoption. A 70% reduction in mechanics violation for a one-off training-time premium of under one fifth, with no inference penalty, means that physical consistency is close to free in this problem class. There is no efficiency argument for omitting it. Because the physics term is an algebraic residual evaluated on quantities the network already predicts, rather than a differential operator requiring automatic differentiation through a solver, the cost stays low. This suggests the approach generalises to other code-governed structural decisions in which a simplified demand model exists in closed form: base-isolation screening, shear-wall placement, and infill-strengthening selection all have this structure.

4.3 Uncertainty as an allocation instrument

Conventional practice treats predictive uncertainty as a defect to be minimised. We reach the opposite interpretation. The bimodal variance distribution in Figure 3 partitions the inventory: a low-variance population the model resolves confidently, and a high-variance population concentrated near class boundaries and among mid-rise frames in zones IV and V.

That partition is directly actionable. The high-variance buildings are precisely those for which nonlinear time-history analysis is worth its cost, and the low-variance buildings are those for which the fast model suffices. The uncertainty output therefore functions as a triage rule that allocates expensive analysis where it changes the decision. Under a fixed engineering budget this converts a model limitation into a resource-allocation gain, and it is a capability that no deterministic threshold rule and no single-label classifier can provide. We regard this as a reinterpretation of what uncertainty quantification is *for* in inventory-scale assessment: not a caveat attached to a prediction, but a signal that selects the next analysis.

4.4 What the attribution ranking reveals about the mechanics

Two features of the SHAP ranking in Figure 2 carry information beyond confirming that the model is sensible.

First, total height (0.18) outranked storey count (0.12). The two are nearly collinear in our dataset, so a purely statistical learner could have used either. The model selected height, which is the variable that actually appears in Equations (1) and (3). Storey count enters only through height. The model therefore recovered the correct causal variable from two correlated candidates, and it did so because the physics residual rewarded the choice. This is a mechanism-level demonstration that the constraint shapes representation learning, not merely the output.

Second, and less expected, the irregularity index (0.25) outranked the seismic zone factor (0.21). Zone factor is the conventional first-order screening variable, and rapid-screening procedures weight it heavily. Our results indicate that within a mixed inventory, configuration deficiency discriminates retrofit need more sharply than hazard level does. The mechanical reason is visible in Equation (2): Z and I_{irr} enter the spectral demand as a product, so a 30% irregularity amplification is equivalent to a full zone-level increase in hazard. Field evidence agrees. Post-earthquake surveys attributed collapse in L'Aquila to soft-storey configuration rather than to unexpected ground-motion intensity [33], and the Lorca assessment reached the same conclusion for infill distribution [10]. The analytical irregularity literature had already established configuration as the governing factor [11]. Our model reproduces that ordering from data, without being told it, which supports a practical implication: screening protocols that rank buildings primarily by seismic zone will systematically under-prioritise irregular frames in moderate-hazard regions.

4.5 Where the framework sits relative to existing work

Physics-informed methods in structural engineering have targeted response prediction: seismic response histories [39], response estimation under sparse labels [19], and nonlinear damage identification [38]. Machine learning applied to retrofit and assessment has been data-driven, including explainable models built explicitly for assessment and retrofit [29] and interpretable drift-demand models [2]. Our framework occupies the intersection. It applies the physics-informed formulation to the *decision* rather than the response, and it retains the explainability that the assessment literature has established as a precondition for trust.

The practical consequence is regulatory. A recommendation from our model can be audited on three grounds at once. Its drift estimate satisfies the mechanics of Equations (1) to (4) to within the residual we report. Its implied drift respects the code limit of Equation (5) by construction. Its feature attribution is consistent with those same equations. A purely data-driven recommendation can be audited on none of them. We consider this the route by which machine-learned retrofit advice becomes admissible in practice, and it is a route that improved accuracy alone does not open.

4.6 Practical screening rules

The trained model reproduced three trends that we state as screening heuristics, each with its mechanical basis:

1. Low-rise frames in zones II and III rarely require bracing unless the irregularity index exceeds about 1.2, because short periods sit on the spectral plateau where K_{eff} is sufficient unless I_{irr} amplifies demand.

2. Mid-rise frames in zone IV with soft-storey conditions frequently require chevron or X-bracing, because the period approaches T_c while the irregularity index concentrates drift in the deficient storey.

3. High-rise frames above twelve storeys in zone V carry a high probability of requiring heavy bracing, because the height term in Equation (3) dominates and the achievable α from light bracing is insufficient.

These heuristics are consistent with the retrofit performance literature [12, 7, 15] and with the passive-dissipation review of braced systems [31]. Their value is that the model supplies a calibrated probability alongside each, so an engineer can see when a building falls outside the heuristic.

5. Fragility and risk assessment

5.1 Multi-state fragility modelling

The classification results establish which intervention the model selects. They do not establish what that intervention buys in risk terms. We therefore propagate the predicted pre- and post-retrofit drift distributions into lognormal fragility functions for four damage states: slight, moderate, extensive and collapse. The probability of exceeding damage state *ds* at intensity measure S_a is

$$\mathrm{P}[\mathrm{DS} \geq \mathrm{ds} \mid S_{\mathrm{a}}] = \Phi\left(\left[\ln(S_{\mathrm{a}}) - \ln(S_{\mathrm{a,ds}})\right] / \beta_{\mathrm{ds}} \right) \qquad (12)$$

where $S_{a,ds}$ is the median spectral acceleration capacity for the damage state, β_{ds} the logarithmic dispersion, and Φ the standard normal cumulative distribution function. The lognormal form and its efficient fitting from dynamic analysis results are standard [4], and the underlying capacity estimates follow the incremental dynamic analysis procedure [32].

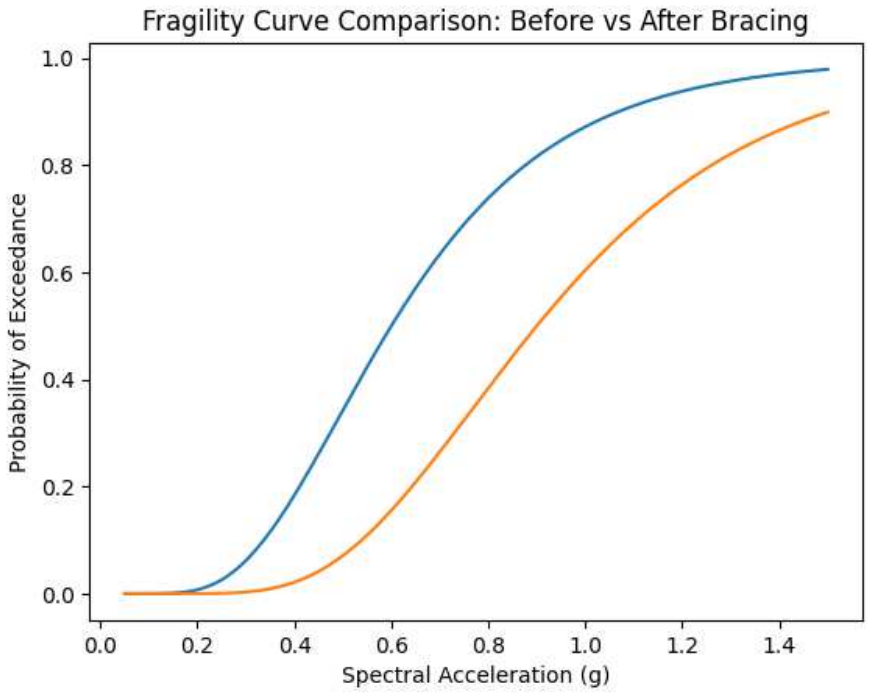


Figure 4. Fragility curves before and after bracing intervention.

Figure 4 shows a consistent rightward shift of the fragility curves across all four damage states after bracing. The shift means a lower exceedance probability at any given intensity, and a higher median capacity. The largest improvement occurred for the collapse state, which follows mechanically from Equation (4): the stiffness increase reduces the drift demand that drives collapse, and the reduction compounds because collapse capacity is the state most sensitive to drift.

5.2 Collapse vulnerability reduction

We integrate the collapse fragility curve over the intensity range to obtain a single scalar vulnerability measure:

$$\mathrm{AUC} = \int P_{\mathrm{collapse}}(\mathrm{IM})\, \mathrm{dIM} \qquad (13)$$

The results were AUC = 0.395 before retrofit and AUC = 0.150 after, a reduction of approximately 62% in cumulative collapse vulnerability over the intensity range considered.

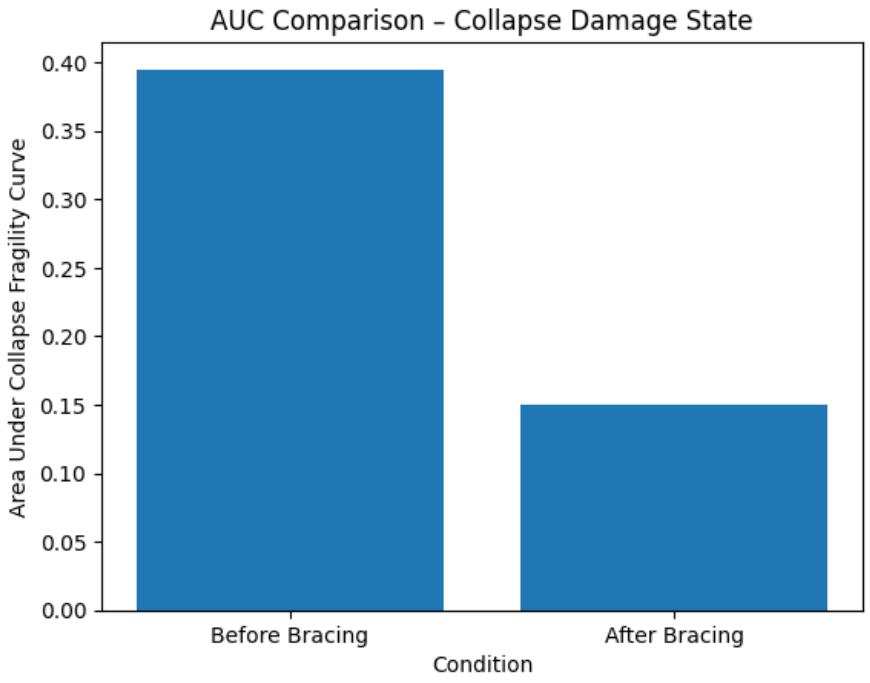


Figure 5. Area under the collapse fragility curve, before and after bracing.

We prefer this integral measure to a comparison of median capacities. A median-capacity comparison reports performance at one point on the hazard curve, whereas Equation (13) integrates performance across the full range of intensities a building may experience, including intensities well above and below the design event. For a retrofit decision that must hold over a building's remaining service life, that integration is the more faithful summary.

5.3 Risk interpretation and expected loss

Under a simplified risk model the expected seismic loss is the product of the exceedance probability and a consequence factor representing repair or replacement cost. The large reduction in collapse fragility therefore implies a proportional reduction in expected structural loss.

The improvements at the moderate and extensive damage states carry a separate and often overlooked implication. Those states govern non-structural damage, repair cost and downtime rather than life safety. Their rightward shift means the intervention improves serviceability-level performance and post-earthquake functionality, not only collapse safety. This is the resilience dimension of the result, and it matters because functional recovery, rather than collapse prevention alone, increasingly defines the objective of retrofit programmes [14]. A bracing scheme justified solely on collapse grounds is therefore under-credited by a factor that our multi-state analysis makes visible.

5.4 Cost–benefit evaluation

We introduce a normalised retrofit investment parameter C_r and estimate the expected loss reduction as proportional to the reduction in collapse AUC:

$$\Delta \mathrm{L} \propto (AUC_{\mathrm{before}} - AUC_{\mathrm{after}}) \cdot C_{\mathrm{r}} \qquad (14)$$

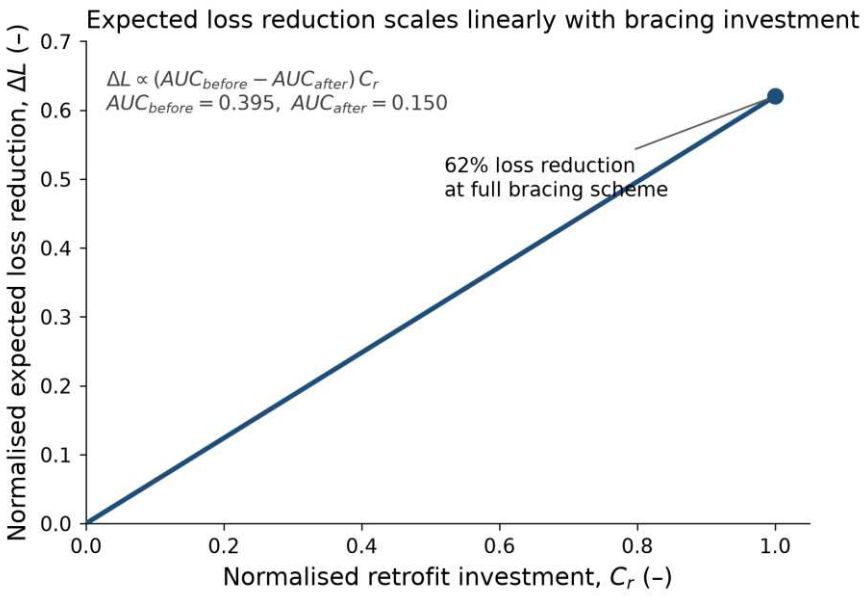


Figure 6. Normalised expected loss reduction against normalised retrofit investment, derived from the collapse fragility areas of Figure 5.

Figure 6 shows a near-linear increase in expected loss reduction with retrofit investment, reaching about 62% at the full bracing scheme. Linearity here is a consequence of Equation (14) and of the stiffness relation in Equation (4): within the range of α considered, added brace stiffness translates proportionally into drift reduction. The practical reading is that bracing does not exhibit strongly diminishing returns over this range, so partial retrofit schemes deliver proportionate benefit. That is directly useful under budget constraints, because it means a programme can strengthen more buildings partially rather than fewer buildings fully and expect comparable aggregate risk reduction. We note that Equation (14) omits construction indirect costs and the sustainability dimension of retrofit, both of which shift the optimum in integrated assessments.

5.5 Coherence between the learning framework and the risk model

The fragility results and the learning framework are not two separate analyses joined at the end. They share the same drift formulation. Because the loss function of Equation (6) enforces Equation (3) during training, the predicted post-retrofit drift distributions are mechanically admissible, and the fragility shifts computed from them inherit that admissibility. Three

properties follow: the machine learning predictions remain mechanically interpretable, the risk metrics rest on structural behaviour rather than on a statistical fit, and the retrofit recommendations align with performance-based design objectives.

We regard this coherence as the framework's principal methodological contribution. A purely data-driven classifier could produce the same class labels, but the drift distribution behind those labels would carry no mechanical guarantee, so the fragility curves derived from it would inherit an unquantified inconsistency. Enforcing the mechanics at the learning stage is what makes the downstream risk numbers defensible.

6. Limitations

We state the limitations plainly, because they bound the claims above.

Simplified structural model. We idealise each frame as an equivalent single-degree-of-freedom oscillator. This omits higher-mode contributions, torsional coupling and the local force redistribution that occurs when a brace is inserted into an existing frame. For tall or strongly irregular buildings these effects are not second order. Our drift estimates should be read as screening-level demand indicators, not as substitutes for a frame analysis.

Synthetic training data. The dataset is physics-consistent but not observed. It cannot contain the correlations present in real inventories — construction-era clustering, regional detailing practice, deterioration — and it therefore cannot demonstrate performance on field data. Validation against instrumented buildings and post-earthquake damage records remains necessary, as work on regional damage prediction from observed data has shown [17].

Linear stiffness assumption. Equations (3) and (4) assume linear elastic behaviour. Real frames respond nonlinearly at the demand levels that trigger a heavy-bracing recommendation, which is precisely the class where the assumption is weakest. The response reduction factor R absorbs this only approximately.

No nonlinear time-history validation. We have not validated the framework against nonlinear time-history analysis, which remains the reference method for seismic performance assessment [32]. This is the most consequential gap, and we regard it as the necessary next step.

Fragility parameters not independently calibrated. The median capacities and dispersions in Equation (12) are consistent with our drift formulation but are not independently calibrated against experimental or empirical fragility data.

Simplified cost model. Equation (14) uses a normalised investment parameter and a proportional consequence factor. It omits indirect costs, discounting over the remaining service life, and embodied-carbon considerations.

Future work should replace the synthetic dataset with high-fidelity nonlinear finite element simulations paired with real building inventory data, extend the mechanics formulation to multi-degree-of-freedom response with torsional coupling, and validate the fragility parameters against observed damage.

7. Conclusions

We developed a physics-informed machine learning framework for probabilistic seismic bracing selection in RC buildings, and we draw the following conclusions.

1. **Mechanics belongs in the objective function, not in a post-processing filter.** Embedding the response-spectrum drift relation and the code drift limit in the loss reduced the mean squared mechanics residual by about 70% relative to an architecturally identical network trained on data misfit alone, and it made code compliance a property of the trained model rather than a downstream check.

2. **In-distribution accuracy is a misleading acceptance criterion for retrofit models.** Tree ensembles outperformed our model by two to four accuracy points and reached higher AUC, yet they inverted the height–drift relation under extrapolation while our model preserved it. We conclude that physical admissibility, not predictive accuracy, is the binding requirement for deployment.

3. **Physical consistency is nearly free.** The constraints cost 12 to 18% additional training time, no inference cost, and a small accuracy concession. There is no computational argument for omitting them.

4. **Predictive uncertainty is an allocation instrument.** The variance distribution separates buildings the model resolves confidently from those it does not, which identifies where nonlinear analysis is worth its cost and converts a limitation into a budgeting advantage.

5. **Configuration outranks hazard level in a mixed inventory.** The attribution analysis ranked the irregularity index above the seismic zone factor, consistent with their multiplicative roles in the spectral demand and with field observations of soft-storey collapse. Screening protocols that rank primarily by seismic zone will under-prioritise irregular frames in moderate-hazard regions.

6. **The framework closes the loop from prediction to risk.** Propagating the predicted post-retrofit drift distributions into lognormal fragility functions shifted all four damage-state curves to the right and reduced cumulative collapse vulnerability by about 62%, with a near-linear relation between investment and expected loss reduction. Improvements at the moderate and extensive states indicate serviceability and functional-recovery gains that a collapse-only assessment does not credit.

Taken together, these results show that coupling structural mechanics with machine learning yields retrofit recommendations that are simultaneously fast, probabilistic, interpretable and code-compliant. We regard that combination, rather than any single accuracy figure, as the prerequisite for AI-assisted seismic retrofit decision support at the scale the existing building stock demands.

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.

References

1. Aljawhari, K., Gentile, R., Freddi, F., & Galasso, C. (2021). Effects of ground-motion sequences on fragility and vulnerability of case-study reinforced concrete frames. *Bulletin of Earthquake Engineering*, 19(15), 6329–6359. https://doi.org/10.1007/s10518-020-01006-8

2. Angelucci, G., Quaranta, G., Mollaioli, F., & Kunnath, S. K. (2024). Interpretable machine learning models for displacement demand prediction in reinforced concrete buildings under pulse-like earthquakes. *Journal of Building Engineering*, 95, 110124. https://doi.org/10.1016/j.jobe.2024.110124

3. American Society of Civil Engineers (2017). *Seismic Evaluation and Retrofit of Existing Buildings* (ASCE/SEI 41-17). ASCE, Reston, VA. https://doi.org/10.1061/9780784414855

4. Baker, J. W. (2015). Efficient Analytical Fragility Function Fitting Using Dynamic Structural Analysis. *Earthquake Spectra*, 31(1), 579–599. https://doi.org/10.1193/021113eqs025m

5. Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2019). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

6. Bureau of Indian Standards (2016). *IS 1893 (Part 1): Criteria for Earthquake Resistant Design of Structures — General Provisions and Buildings*. BIS, New Delhi.

7. Castaldo, P., Tubaldi, E., Selvi, F., & Gioiella, L. (2020). Seismic performance of an existing RC structure retrofitted with buckling restrained braces. *Journal of Building Engineering*, 33, 101688. https://doi.org/10.1016/j.jobe.2020.101688

8. Chen, M., Park, Y., Mangalathu, S., & Jeon, J.-S. (2024). Effect of data drift on the performance of machine-learning models: Seismic damage prediction for aging bridges. *Earthquake Engineering & Structural Dynamics*, 53(15), 4541–4561. https://doi.org/10.1002/eqe.4230

9. Cuomo, S., Schiano Di Cola, V., Giampaolo, F., Rozza, G., Raissi, M., & Piccialli, F. (2022). Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next. *Journal of Scientific Computing*, 92(3). https://doi.org/10.1007/s10915-022-01939-z

10. De Luca, F., Verderame, G. M., Gómez-Martínez, F., & Pérez-García, A. (2014). The structural role played by masonry infills on RC building performances after the 2011 Lorca, Spain, earthquake. *Bulletin of Earthquake Engineering*, 12(5), 1999–2026. https://doi.org/10.1007/s10518-013-9500-1

11. De Stefano, M., & Pintucchi, B. (2007). A review of research on seismic behaviour of irregular building structures since 2002. *Bulletin of Earthquake Engineering*, 6(2), 285–308. https://doi.org/10.1007/s10518-007-9052-3

12. Di Sarno, L., & Manfredi, G. (2010). Seismic retrofitting with buckling restrained braces: Application to an existing non-ductile RC framed building. *Soil Dynamics and Earthquake Engineering*, 30(11), 1279–1297. https://doi.org/10.1016/j.soildyn.2010.06.001

13. Federal Emergency Management Agency (2000). *FEMA 356: Prestandard and Commentary for the Seismic Rehabilitation of Buildings*. FEMA, Washington, DC.

14. Freddi, F., Galasso, C., Cremen, G., Dall'Asta, A., Di Sarno, L., Giaralis, A., et al. (2021). Innovations in earthquake risk reduction for resilience: Recent advances and challenges. *International Journal of Disaster Risk Reduction*, 60, 102267. https://doi.org/10.1016/j.ijdrr.2021.102267

15. Freddi, F., Tubaldi, E., Ragni, L., & Dall'Asta, A. (2013). Probabilistic performance assessment of low-ductility reinforced concrete frames retrofitted with dissipative braces.

Earthquake Engineering & Structural Dynamics, 42(7), 993–1011. https://doi.org/10.1002/eqe.2255

16. Gawlikowski, J., Njieutcheu Tassi, C. R., Ali, M., Lee, J., Humt, M., Feng, J., et al. (2023). A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 56(S1), 1513–1589. https://doi.org/10.1007/s10462-023-10562-9

17. Ghimire, S., Guéguen, P., Giffard-Roisin, S., & Schorlemmer, D. (2022). Testing machine learning models for seismic damage prediction at a regional scale using building-damage dataset compiled after the 2015 Gorkha Nepal earthquake. *Earthquake Spectra*, 38(4), 2970–2993. https://doi.org/10.1177/87552930221106495

18. Goel, R. K., & Chopra, A. K. (1998). Period Formulas for Concrete Shear Wall Buildings. *Journal of Structural Engineering*, 124(4), 426–433. https://doi.org/10.1061/(asce)0733-9445(1998)124:4(426)

19. Hu, Y., Tsang, H.-H., Lam, N., & Lumantarna, E. (2024). Physics-informed neural networks for enhancing structural seismic response prediction with pseudo-labelling. *Archives of Civil and Mechanical Engineering*, 24(1). https://doi.org/10.1007/s43452-023-00820-6

20. Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, 110(3), 457–506. https://doi.org/10.1007/s10994-021-05946-3

21. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440. https://doi.org/10.1038/s42254-021-00314-5

22. Kazemi, F., Asgarkhani, N., & Jankowski, R. (2023). Machine learning-based seismic fragility and seismic vulnerability assessment of reinforced concrete structures. *Soil Dynamics and Earthquake Engineering*, 166, 107761. https://doi.org/10.1016/j.soildyn.2023.107761

23. Kazemi, F., Asgarkhani, N., & Jankowski, R. (2023). Machine learning-based seismic response and performance assessment of reinforced concrete buildings. *Archives of Civil and Mechanical Engineering*, 23(2). https://doi.org/10.1007/s43452-023-00631-9

24. Kiani, J., Camp, C. V., & Pezeshk, S. (2019). On the application of machine learning techniques to derive seismic fragility curves. *Computers & Structures*, 218, 108–122. https://doi.org/10.1016/j.compstruc.2019.03.004

25. Ni, P., Sun, L., Yang, J., & Li, Y. (2022). Multi-End Physics-Informed Deep Learning for Seismic Response Estimation. *Sensors*, 22(10), 3697. https://doi.org/10.3390/s22103697

26. Priestley, M. J. N., & Kowalsky, M. J. (2000). Direct displacement-based seismic design of concrete buildings. *Bulletin of the New Zealand Society for Earthquake Engineering*, 33(4), 421–444. https://doi.org/10.5459/bnzsee.33.4.421-444

27. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

28. Samadian, D., Muhit, I. B., & Dawood, N. (2025). Application of Data-Driven Surrogate Models in Structural Engineering: A Literature Review. *Archives of Computational Methods in Engineering*, 32(2), 735–784. https://doi.org/10.1007/s11831-024-10152-0

29. Shabbir, K., Noureldin, M., & Sim, S.-H. (2025). Data-driven model for seismic assessment, design, and retrofit of structures using explainable artificial intelligence. *Computer-Aided Civil and Infrastructure Engineering*, 40(3), 281–300. https://doi.org/10.1111/mice.13338

30. Sun, H., Burton, H. V., & Huang, H. (2020). Machine learning applications for building structural design and performance assessment: State-of-the-art review. *Journal of Building Engineering*, 33, 101816. https://doi.org/10.1016/j.jobe.2020.101816

31. Titirla, M. (2023). A State-of-the-Art Review of Passive Energy Dissipation Systems in Steel Braces. *Buildings*, 13(4), 851. https://doi.org/10.3390/buildings13040851

32. Vamvatsikos, D., & Cornell, C. A. (2002). Incremental dynamic analysis. *Earthquake Engineering & Structural Dynamics*, 31(3), 491–514. https://doi.org/10.1002/eqe.141

33. Verderame, G. M., De Luca, F., Ricci, P., & Manfredi, G. (2011). Preliminary analysis of a soft-storey mechanism after the 2009 L'Aquila earthquake. *Earthquake Engineering & Structural Dynamics*, 40(8), 925–944. https://doi.org/10.1002/eqe.1069

34. Wang, Z., Pedroni, N., Zentner, I., & Zio, E. (2018). Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment. *Engineering Structures*, 162, 213–225. https://doi.org/10.1016/j.engstruct.2018.02.024

35. Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems. *ACM Computing Surveys*, 55(4), 1–37. https://doi.org/10.1145/3514228

36. Wu, Y., Sicard, B., & Gadsden, S. A. (2024). Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring. *Expert Systems with Applications*, 255, 124678. https://doi.org/10.1016/j.eswa.2024.124678

37. Xie, Y., Ebad Sichani, M., Padgett, J. E., & DesRoches, R. (2020). The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. *Earthquake Spectra*, 36(4), 1769–1801. https://doi.org/10.1177/8755293020919419

38. Yamaguchi, T., & Mizutani, T. (2024). A Physics-Informed Neural Network for the Nonlinear Damage Identification in a Reinforced Concrete Bridge Pier Using Seismic Responses. *Structural Control and Health Monitoring*, 2024(1). https://doi.org/10.1155/2024/5532909

39. Zhang, R., Liu, Y., & Sun, H. (2020). Physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling. *Engineering Structures*, 215, 110704. https://doi.org/10.1016/j.engstruct.2020.110704

40. Zhang, Y., Burton, H. V., Sun, H., & Shokrabadi, M. (2017). A machine learning framework for assessing post-earthquake structural safety. *Structural Safety*, 72, 1–16. https://doi.org/10.1016/j.strusafe.2017.12.001