Machine Learning-Based Prediction on Normalized Cumulative Dissipated Energy (NCDE) of Reinforced Concrete Shear Walls

Negar Dokhtmirzahasanvahid

Research Objective

This project aims to **predict the normalized cumulative dissipated energy (NCDE)** of **reinforced concrete shear walls** using machine learning models. NCDE is a measure of how much energy a wall can absorb before failure — which is crucial in **earthquake-resistant design**.

To do this, the project:

1. Uses **three interpretable models**: Neural Additive Model (NAM), Random Forest (RF), and XGBoost (XGB).
2. Trains them on **18 structural and material features** extracted from a database of wall specimens.
3. Applies **SHAP analysis** to interpret feature contributions and compare models.
4. Evaluates which wall properties most influence energy dissipation, and whether models agree on the direction (positive/negative) of impact.

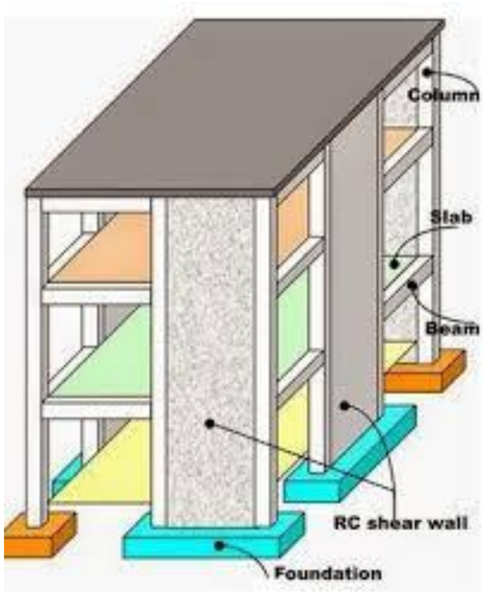
Outline

- Dataset and Feature Exploration
- NAM Configuration and Training
- Baseline Model Configuration
- Model Comparison
- SHAP Analysis on Baseline Models
- Comparing NAM Feature Importance and SHAP
- Conclusion

Data Understanding and Exploration

What is a Shear Wall?

A shear wall is a structural support element that resists shear forces, like heavy winds and seismic activity.



Feature Analysis and Statistical Insights

Feature Name	Description
`hw`	Wall height
`lw`	Wall length
`b0`	Boundary width
`tw`	Wall thickness
`fysh`	Transverse reinforcement yield strength in boundary
`ρsh`	Transverse reinforcement ratio
`M/Vlw`	Moment-to-shear ratio (flexural vs shear)
`AR`	Aspect ratio (height/length)
`P/(Agf′c)`	Axial load ratio
`f′c`	Concrete compressive strength
`fyt`, `fyl`, `fybl`	Steel yield strengths
`s/db`	Hoop spacing to bar diameter
`ρbl`	Plastic hinge length

Data Cleaning, Normalization, and Splitting

Cleaning & Preprocessing Steps

- **Read Excel file** containing raw shear wall data
- **Drop metadata columns** and retain only 18 features + 1 target
- **Handle missing values** by filling with **training set median**
- **Ensure valid target** by raising errors if NCDE is missing
- **Input: (312 x 18)**
- **Output (312 x 1)**

Normalization Strategy

- Apply MinMaxScaler(feature_range=(-1, 1)) to:
  - Feature matrix X
  - Target variable y=NCDE

🔀 **Final Output (20/80)**

- Xtrain ,Xtest, ytrain, ytestX

Feature Analysis and Statistical Insights

	Feature	Mean	Std Dev	Min	Max	Skew	Kurtosis
1	lw	-0.352885	0.388193	-1.0	1.0	1.216663	1.717661
2	hw	-0.33073	0.479448	-1.0	3.6	2.327667	13.667403
3	tw	-0.230067	0.330108	-1.0	1.0	0.449187	-0.501191
4	f'c	-0.460248	0.375978	-1.0	1.0	1.795692	3.689697
5	fyt	-0.298218	0.333924	-1.0	1.0	0.846475	1.537292
6	fysh	-0.494004	0.297745	-1.0	1.0	2.76544	10.768604
7	fyl	-0.278302	0.329516	-1.0	1.0	0.999973	1.909749
8	fybl	-0.626374	0.199719	-1.0	1.0	2.214792	13.880487
9	ρt	-0.611	0.279614	-1.002641	1.0	1.69637	5.556095
10	ρsh	-0.822886	0.159468	-1.0	1.0	6.190322	59.501334
11	ρl	-0.715882	0.243084	-1.0	1.0	2.584816	11.774368
12	ρbl	-0.509668	0.457474	-1.0	1.0	1.529157	1.392031
13	P/(Agf'c)	-0.593558	0.465544	-1.0	1.0	1.736983	3.178958
14	b0	-0.784042	0.222026	-1.0	1.0	4.308101	27.970313
15	db	-0.3068	0.391542	-1.0	1.325984	0.675563	1.612299
16	s/db	-0.848991	0.283247	-1.0	1.0	3.311556	15.589486
17	AR	-0.288611	0.420676	-1.0	2.742502	1.349713	7.910063
18	M/Vlw	-0.498012	0.319806	-1.0	1.486772	1.316477	5.473075

Feature Analysis and Statistical Insights

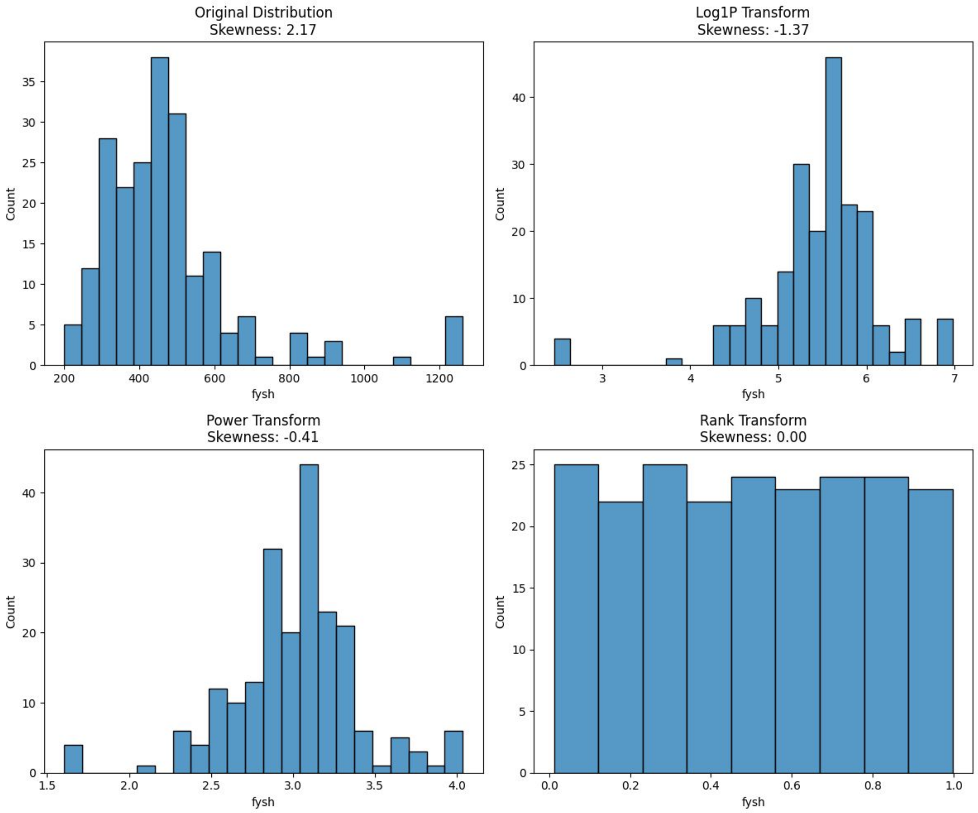
Transverse boundary reinforcement yield strength (fysh)
→ Highly skewed (2.76): most values are low, rare extreme highs

Transverse boundary reinforcement ratio (ρsh)
→ Extremely skewed (6.19): indicates rare but extreme reinforcement cases

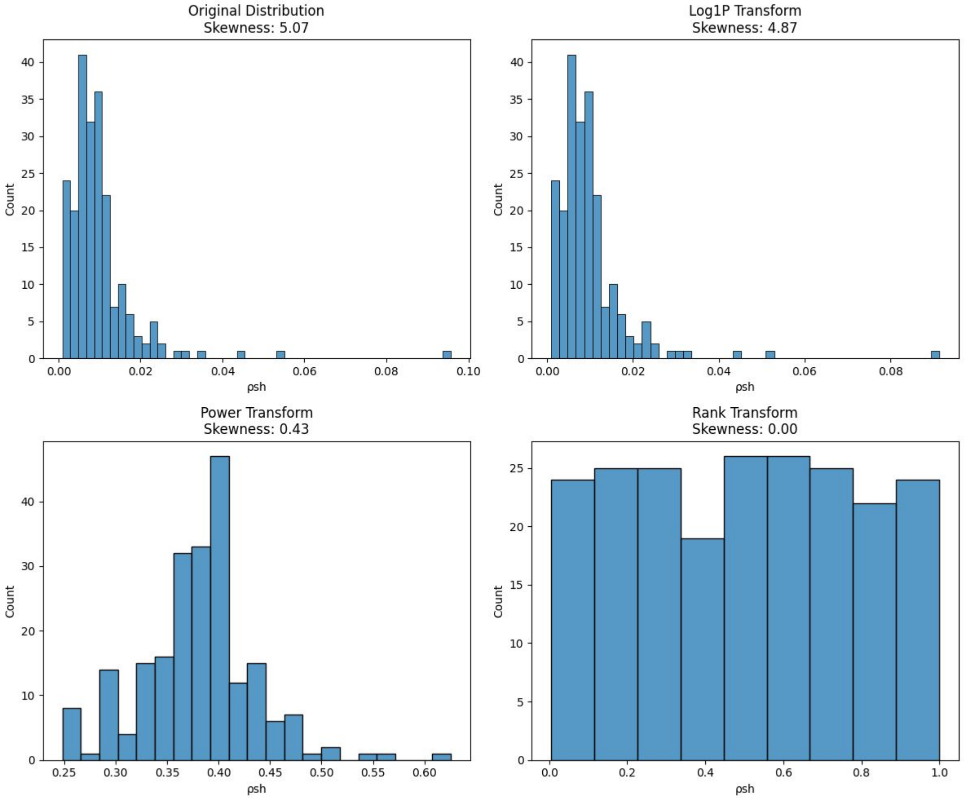
Boundary element depth (b0)
→ Skewed (4.31): extreme depth values are rare but impactful

Hoop spacing / boundary element length (s/db)
→ Skewed (3.12): indicates most configurations have closely spaced hoops

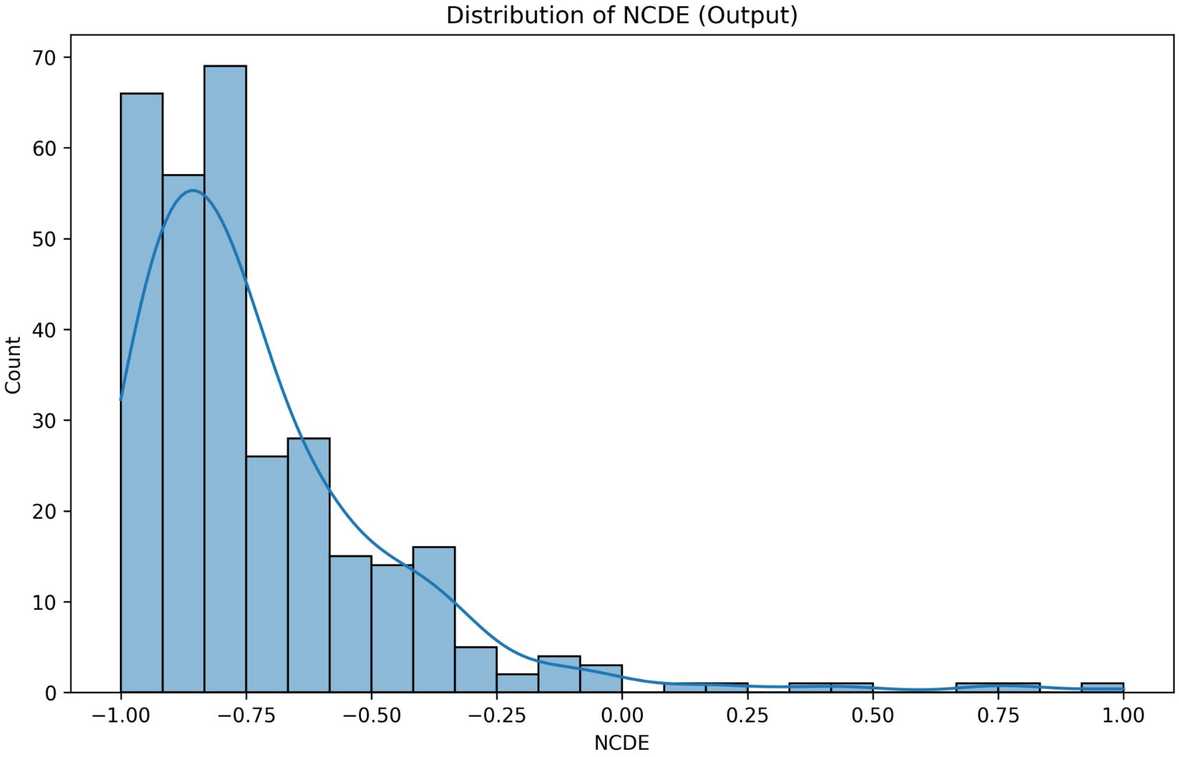
Fysh Transverse boundary: Reinforcement yield strength Transformations



Psh Transverse boundary reinforcement ratio Analysis



Feature Analysis and Statistical Insights



Feature Analysis and Statistical Insights

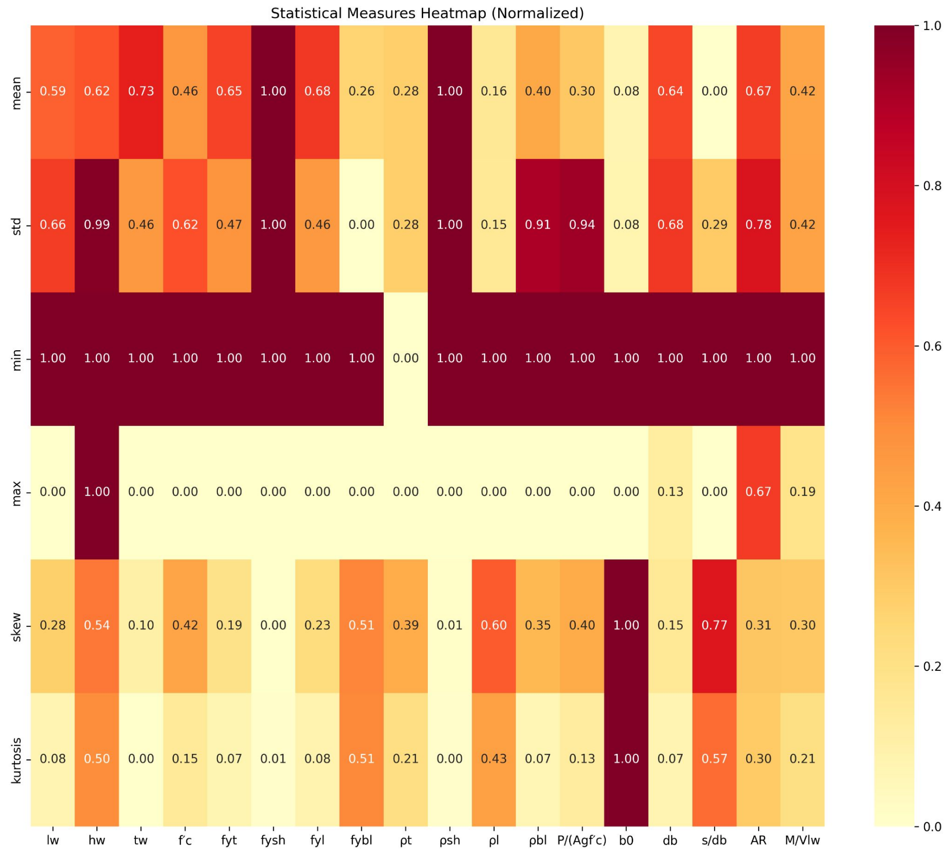
1. Heavily Left-Skewed Output

- Most samples cluster tightly around **–1.0 to –0.6**, meaning the majority of walls have **low normalized energy dissipation**.
- Very few samples exhibit higher NCDE, and values closer to 0 or above are **rare outliers**.

2. Interpretation in Engineering Terms

- **NCDE** reflects the wall's ability to dissipate energy under seismic or cyclic loading.
- A concentration at low NCDE values implies most designs are **less energy-absorptive** — possibly brittle or minimally ductile.
- High-NCDE walls (those to the right side of the plot) are likely **well-reinforced, ductile designs** that can absorb more seismic energy.

Feature Analysis and Statistical Insights



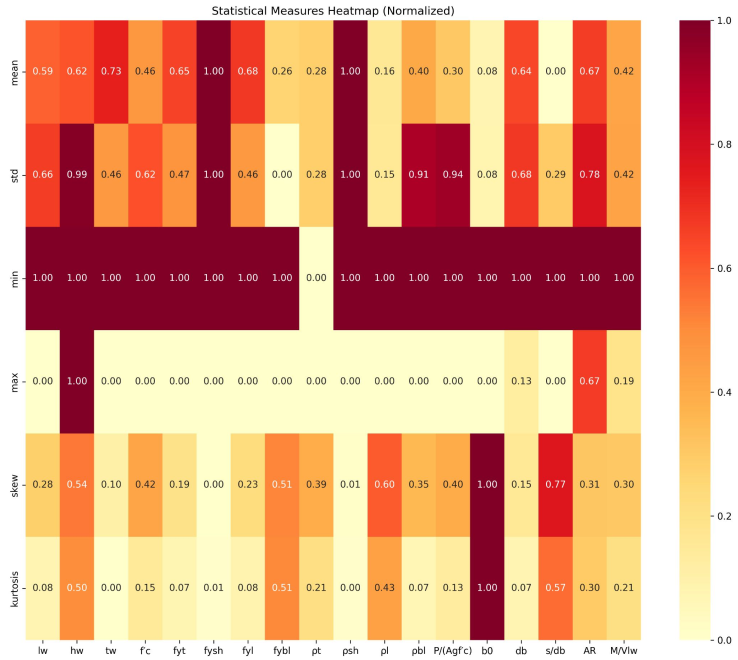
Feature Analysis and Statistical Insights

Wall height (hw) and **boundary element length (db)** show the highest variability among all features, indicating that structural designs vary significantly in overall size and edge configuration.

Hoop spacing ratio (s/db) and **boundary depth (b0)** stand out for having strong skewness and kurtosis, meaning their distributions are dominated by rare extreme values — this could impact how models learn about confinement effects.

Axial load ratio (P/(Agf′c)) has a highly peaked distribution, with most walls clustered around similar axial load values, and a few rare outliers that can dominate learning if not treated carefully.

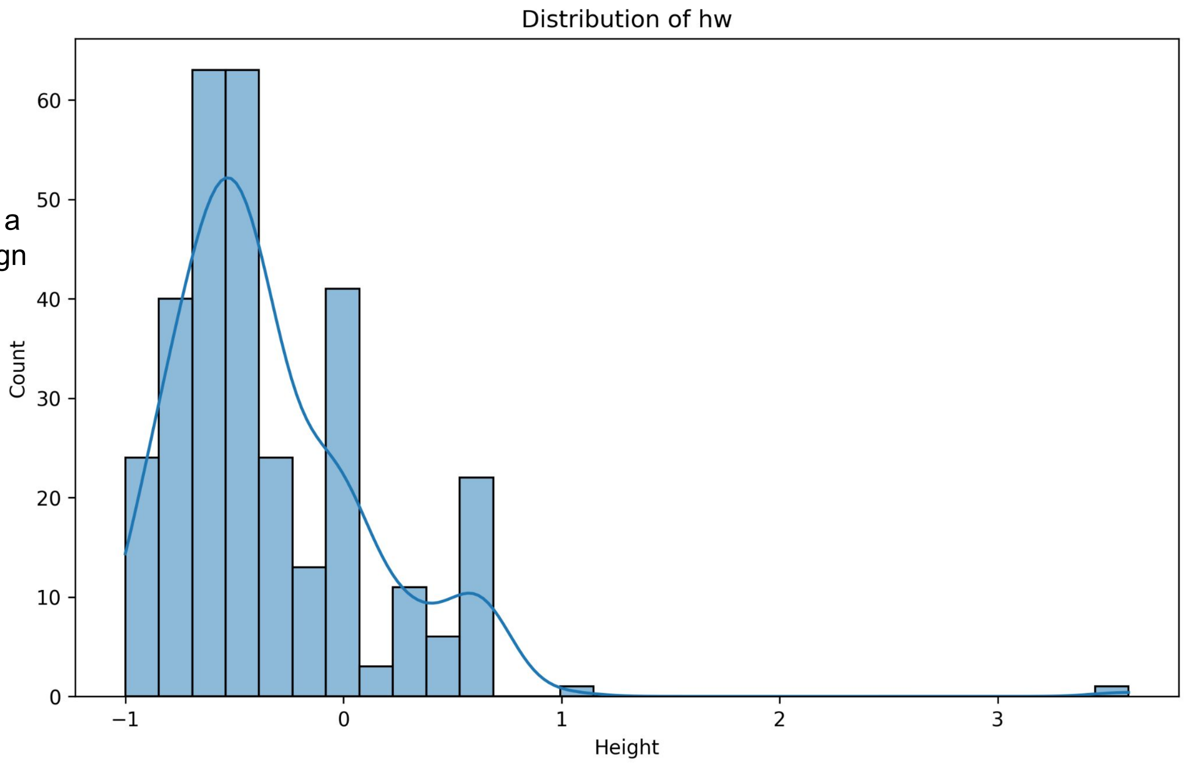
Wall thickness (tw) and **transverse web reinforcement yield strength (fyt)** are among the most balanced features — their distributions are symmetrical and consistent, making them easier for models to learn from reliably.



Feature Analysis and Statistical Insights

There's a sharp peak around **–0.6**, suggesting a strong cluster of walls around a typical normalized height — indicating design consistency in wall sizing.

A small number of walls reach extreme values (up to **+3.6**), representing outlier configurations or high-rise structural cases.



Feature Analysis and Statistical Insights

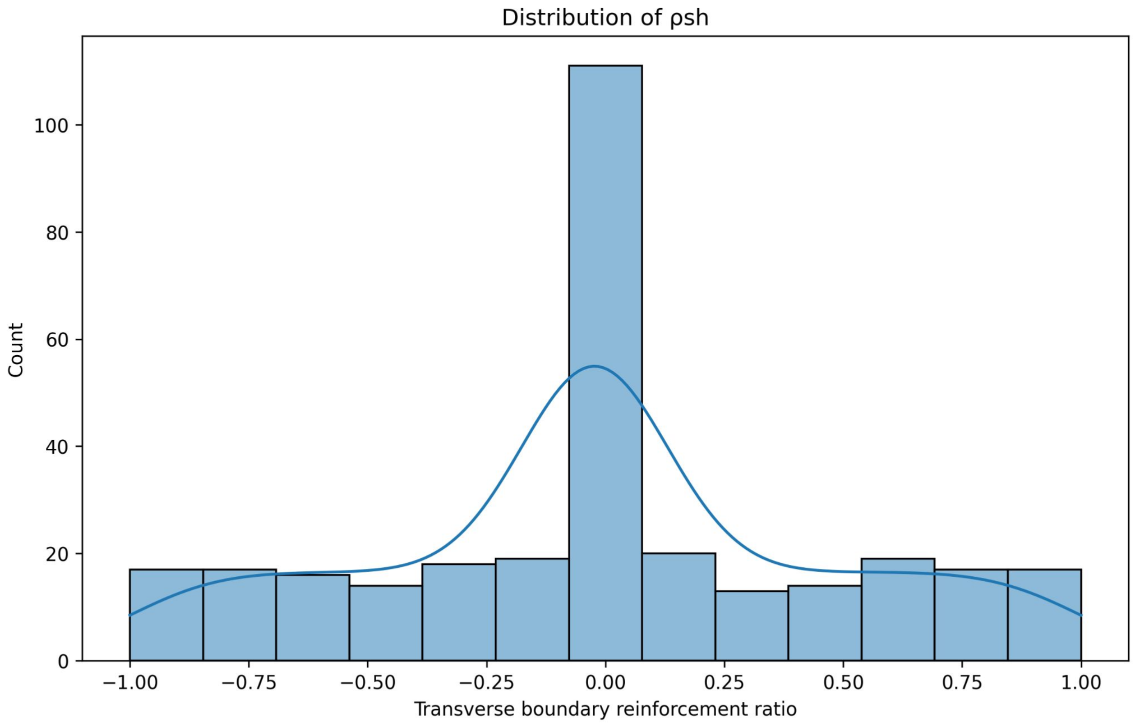
ρsh (Transverse boundary reinforcement ratio) refers to the **amount of horizontal reinforcement** placed in the boundary region of the wall, relative to the concrete volume. It reflects how well-confined the wall edges are.

The distribution is **extremely peaked around 0**, meaning that **most walls have little to no transverse reinforcement in their boundary elements**.

There's a sharp spike in the middle, with very few examples on either side — indicating **very limited variation** in this feature across the dataset.

Why it matters:
Models may treat this feature as a constant due to its lack of variability

The heavy peak also contributes to high kurtosis — making the dataset sensitive to small changes in this variable.



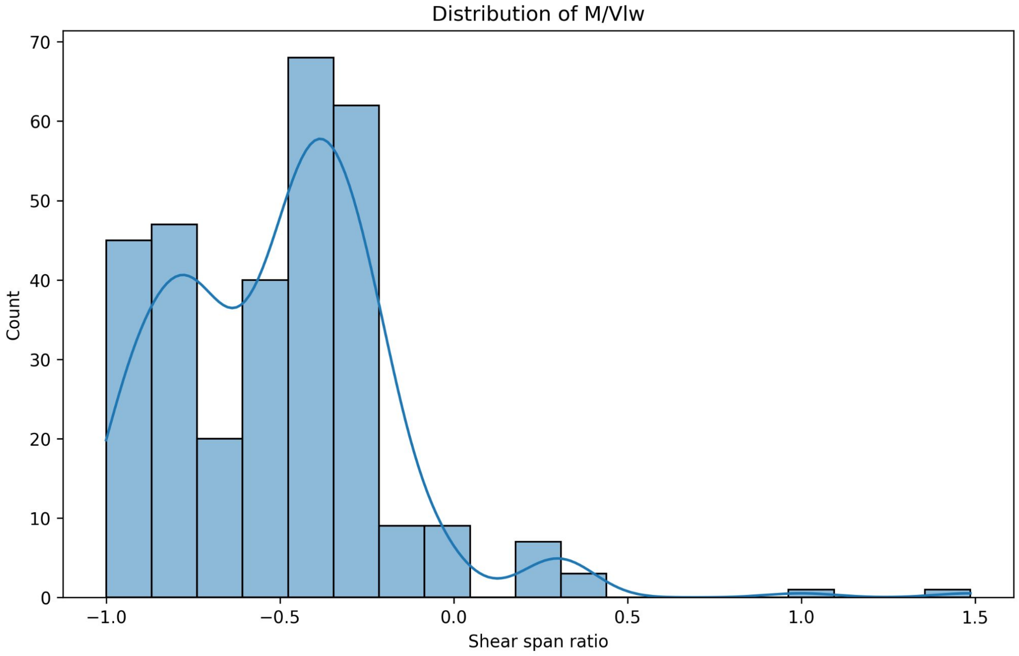
Feature Analysis and Statistical Insights

M/Vlw (Shear span ratio) reflects the ratio between **bending moment (M)** and **shear force (V)** scaled by wall length — essentially indicating how a wall behaves under lateral loading: **flexural (bending)** vs **shear-dominated** behavior.

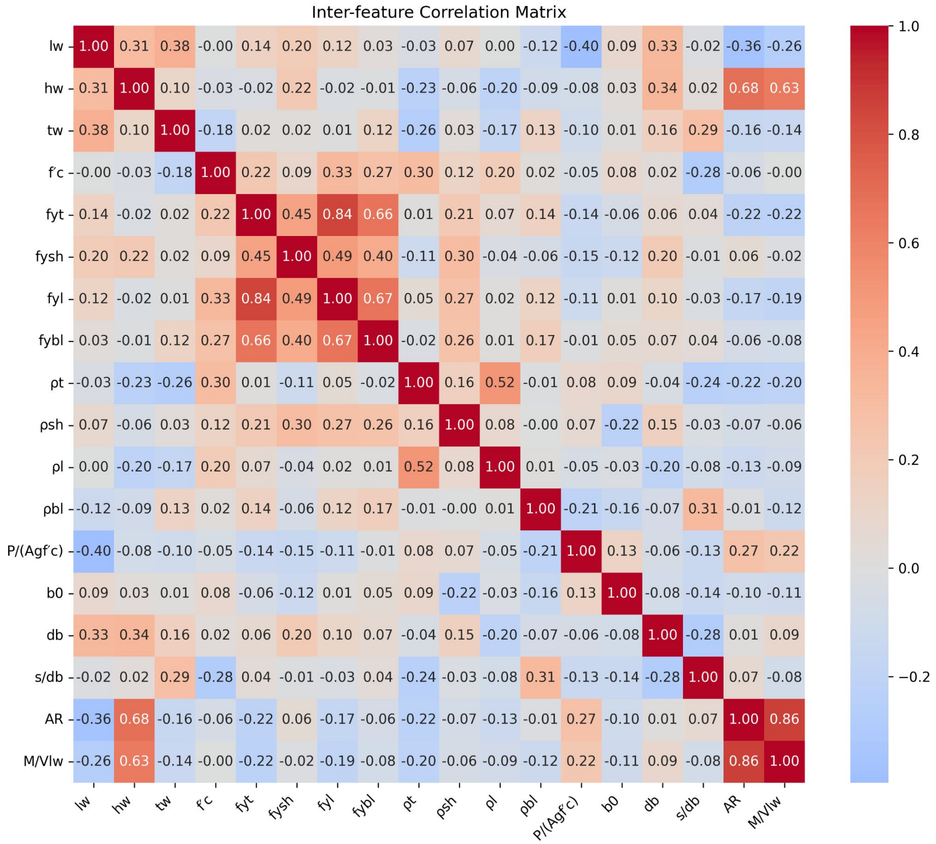
The distribution is **sharply left-skewed**, with most walls concentrated between **–1.0 and –0.25** in the normalized range.
This shows that **most walls are shear-dominated**, especially under seismic loading.

A small number of samples fall in the flexure-dominant region (closer to 0 or positive values), making those cases **rare and potentially underrepresented**.

Why it matters:
The model will likely learn shear-dominated behavior well.



Feature Analysis and Statistical Insights



Highly Correlated Feature Pairs

This lists pairs of features whose **Pearson correlation coefficient** |r| **|r|** |r| exceeds **0.7**, which is generally considered **strong correlation**.

Feature Correlation (|r|)

Feature Pair	|r|
fyt – fyl	0.842
AR – M/Vlw	0.861

Highly Correlated Feature Pairs

1. fyt – fyl (|r| = 0.842)

- **fyt**: Transverse web reinforcement yield strength
- **fyl**: Vertical web reinforcement yield strength
- These two types of steel are typically selected and detailed together in design codes, leading to strong correlation.

Implication: They may be **redundant in the model**, carrying overlapping information. Dimensionality reduction or regularization could help.

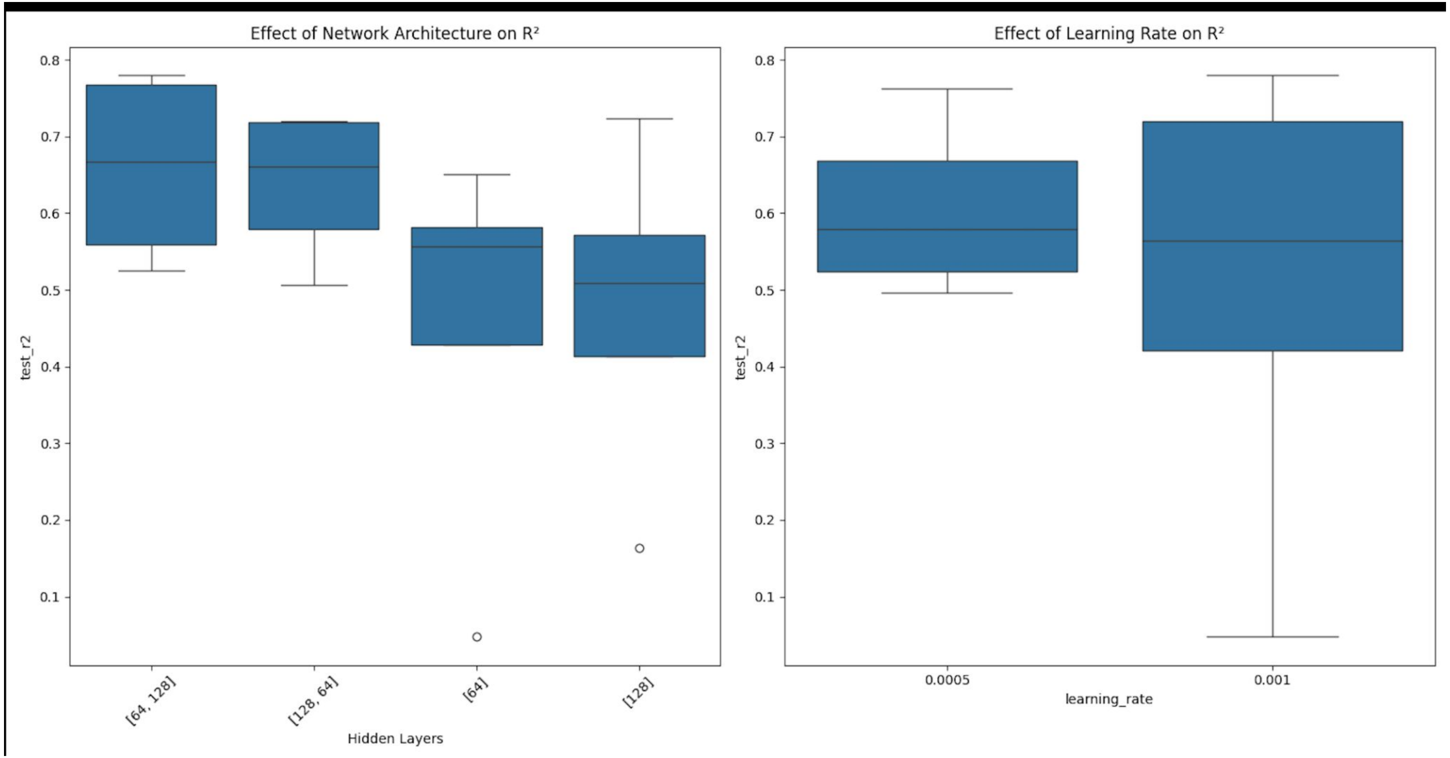
2. AR – M/Vlw (|r| = 0.861)

- **AR**: Aspect ratio = height/length
- **M/Vlw**: Shear span ratio = moment / (shear × length)
- Both express how **geometry and loading interact** to affect wall behavior — taller, more slender walls usually correspond to higher shear spans.

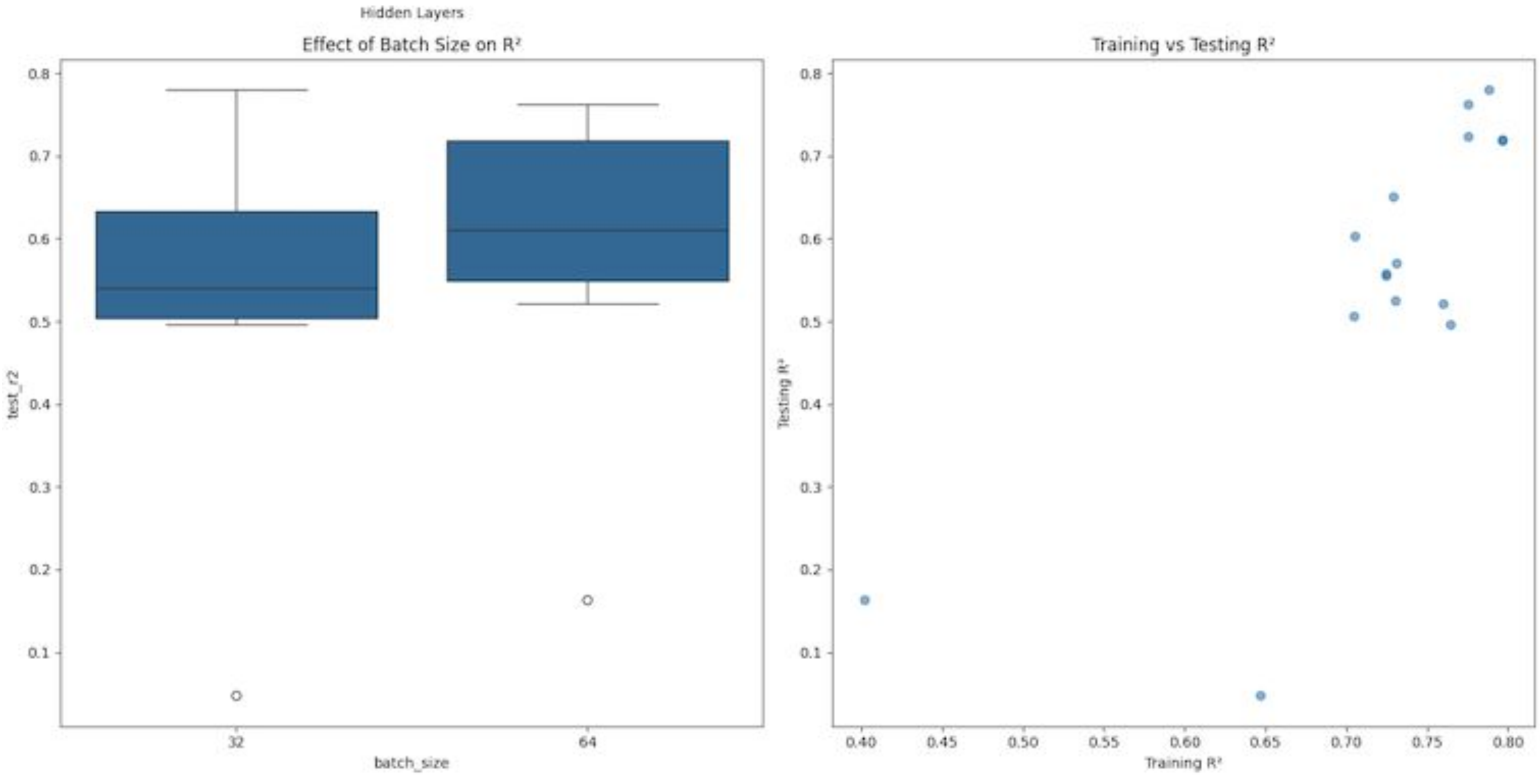
Implication: These features describe **related structural effects**. Including both may be unnecessary unless their individual roles are clearly interpretable.

NAM Generation

NAM Configuration



NAM Configuration

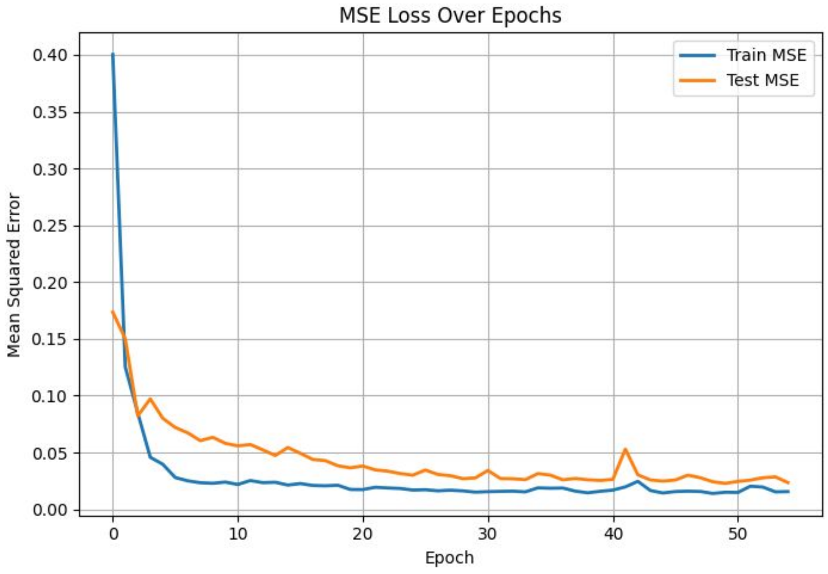


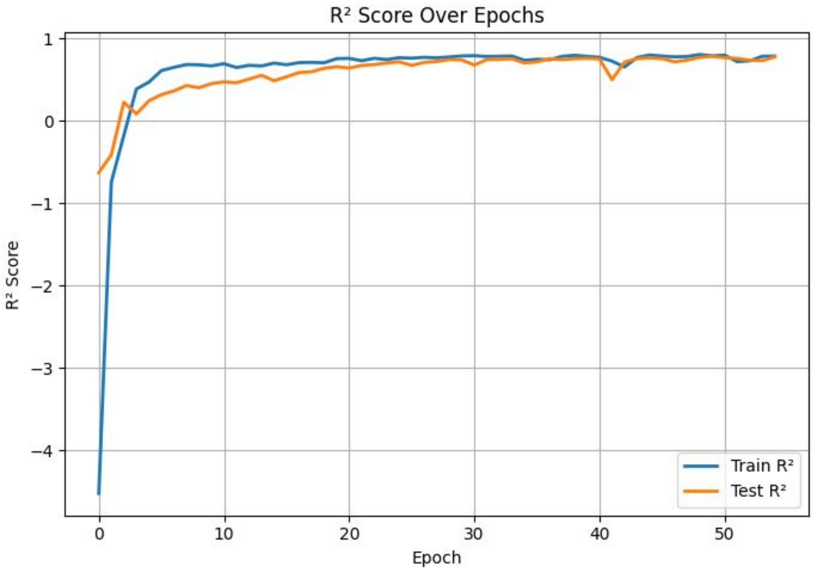
NAM Configuration

✅ Best NAM Model Summary

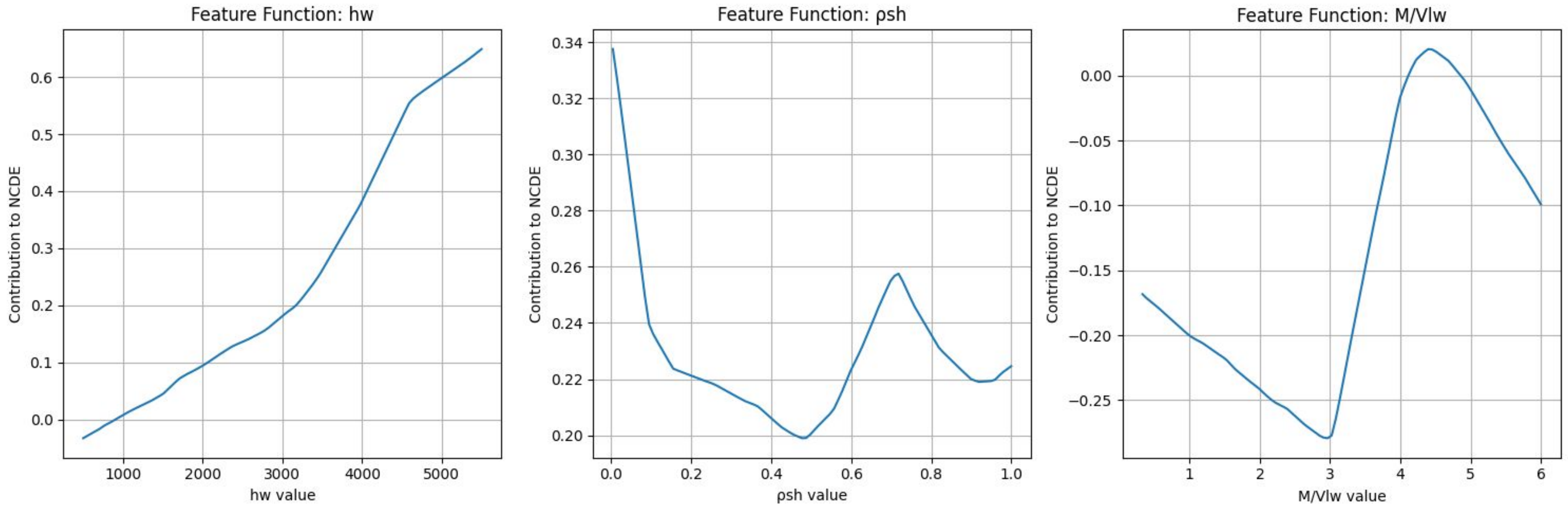
Category	Metric/Parameter	Value
Configuration	Hidden Layers	[64, 128]
	Learning Rate	0.001
	Batch Size	32
Training Performance	RMSE	177.4806
	MAE	114.3431
	R^2	0.7880
	Explained Variance	0.7881
Testing Performance	RMSE	217.9604
	MAE	147.7334
	R^2	0.7796

NAM Configuration





NAM Configuration



Feature Analysis and Statistical Insights

Analysis for hw:
- Function is monotonic
- Function is linear
- Overall positive effect on NCDE:
Contribution range: 0.6814

Taller walls enhance ductility and capacity to absorb seismic energy, making height a reliable design factor.

Analysis for ρsh:
- Function is non-monotonic
- Function is non-linear
- Overall negative effect on NCDE:
Contribution range: 0.1385

Indicates that simply increasing horizontal reinforcement in boundaries doesn't always improve energy performance — detailing and placement likely matter more than quantity.

Analysis for M/Vlw:
- Function is non-monotonic
- Function is non-linear
- Overall positive effect on NCDE
- Contribution range: 0.2998

Suggests that the wall's failure mode plays a role — walls that are too shear- or too flexure-dominant may perform poorly, while those in a balanced zone dissipate more energy.

NAM Configuration

Wall Height (hw)

- Function is monotonic

- Function is linear

- Overall positive effect on NCDE: Contribution range: 0.6814

The model has learned that taller walls are generally more ductile or energy-absorbing, making height a strongly **positive predictor** of NCDE.

Transverse Boundary Reinforcement Ratio (ρsh)

- Function is non-monotonic

- Function is non-linear

- Overall negative effect on NCDE:

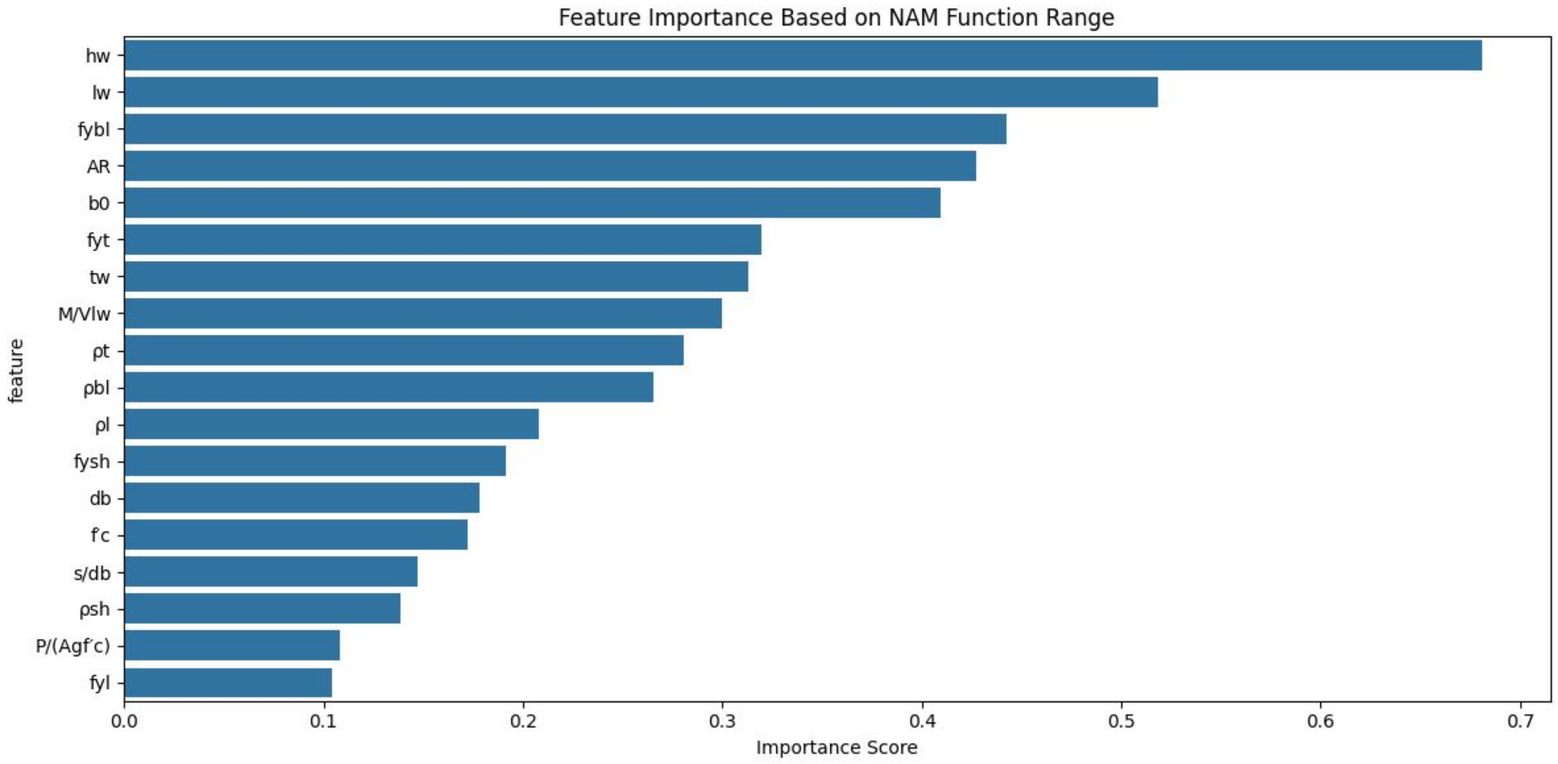
Contribution range: 0.1385

The effect of transverse reinforcement in boundaries is **complex and nonlinear**. It may be beneficial at very low or moderate levels, but oversaturation or poor detailing could reduce its efficiency.

Shear Span Ratio (M/Vlw): - Contribution range: 0.2998

Walls with moderate shear span ratios reduce energy dissipation, while extreme flexural or shear cases may actually perform better — likely due to detailing or failure mode differences.

NAM Configuration



NAM Configuration

- **1. Most Influential Features**

  - **hw (Wall height)** and **lw (Wall length)** dominate, with the highest importance scores.

    These geometry features directly impact deformation capacity and energy dissipation.
    Their strong signal confirms that **larger walls are more ductile** and structurally significant.

  - **fybl (Vertical boundary reinforcement yield strength)** and **AR (Aspect ratio)** follow closely.

    Reinforcement quality in the boundary and overall wall slenderness are key to understanding energy behavior.

- **2. Moderately Important Features**

  - **b0 (Boundary depth)**, **fyt (Transverse web steel strength)**, and **tw (Thickness)** show balanced influence.

    They contribute to stiffness, confinement, and detailing but are less dominant than global geometry.

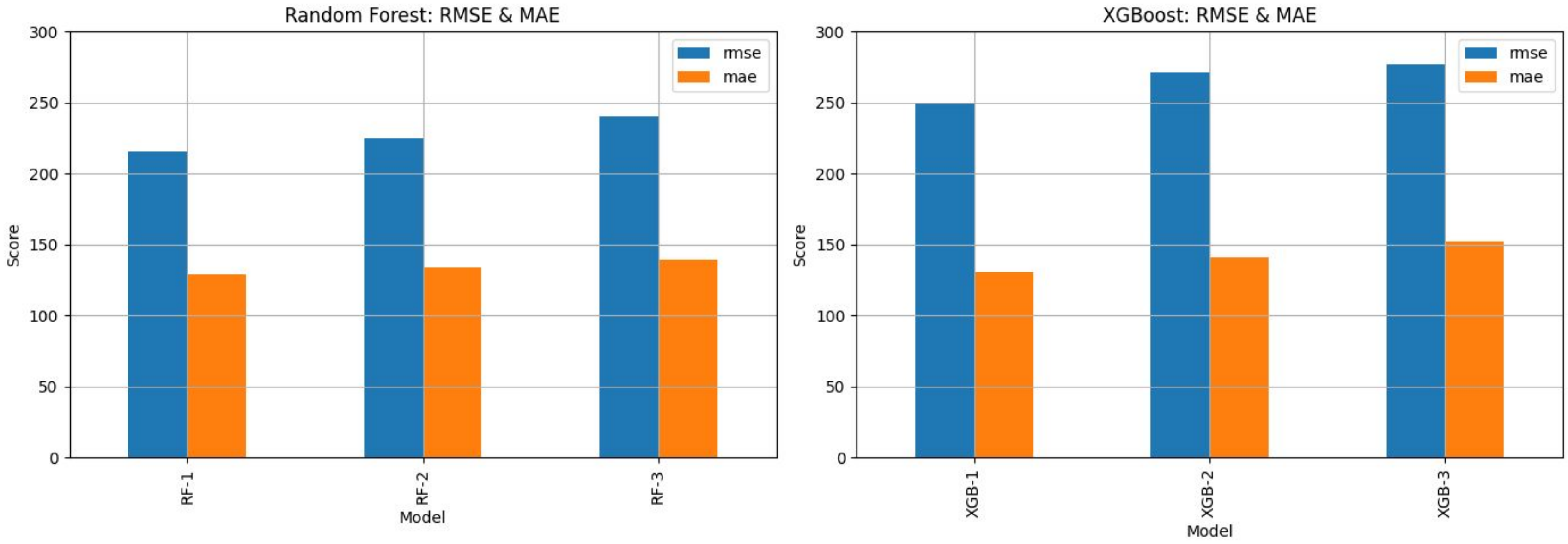
  - **M/Vlw (Shear span ratio)** shows a non-trivial effect — reaffirming the role of failure mode (flexure vs shear) in dissipated energy.

- **3. Less Influential Features**

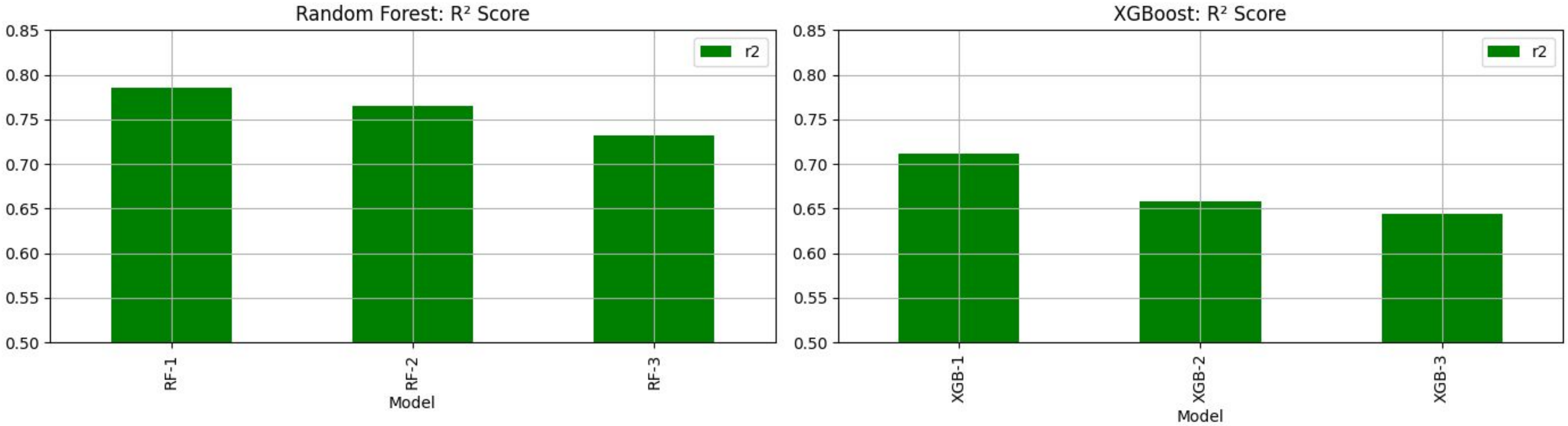
  - **P/(Agf′c) (Axial load ratio)**, **ρsh (Transverse boundary reinforcement ratio)**, and **fyl (Vertical web steel strength)** score lowest.

Baseline Models

Baseline Models



Baseline Models

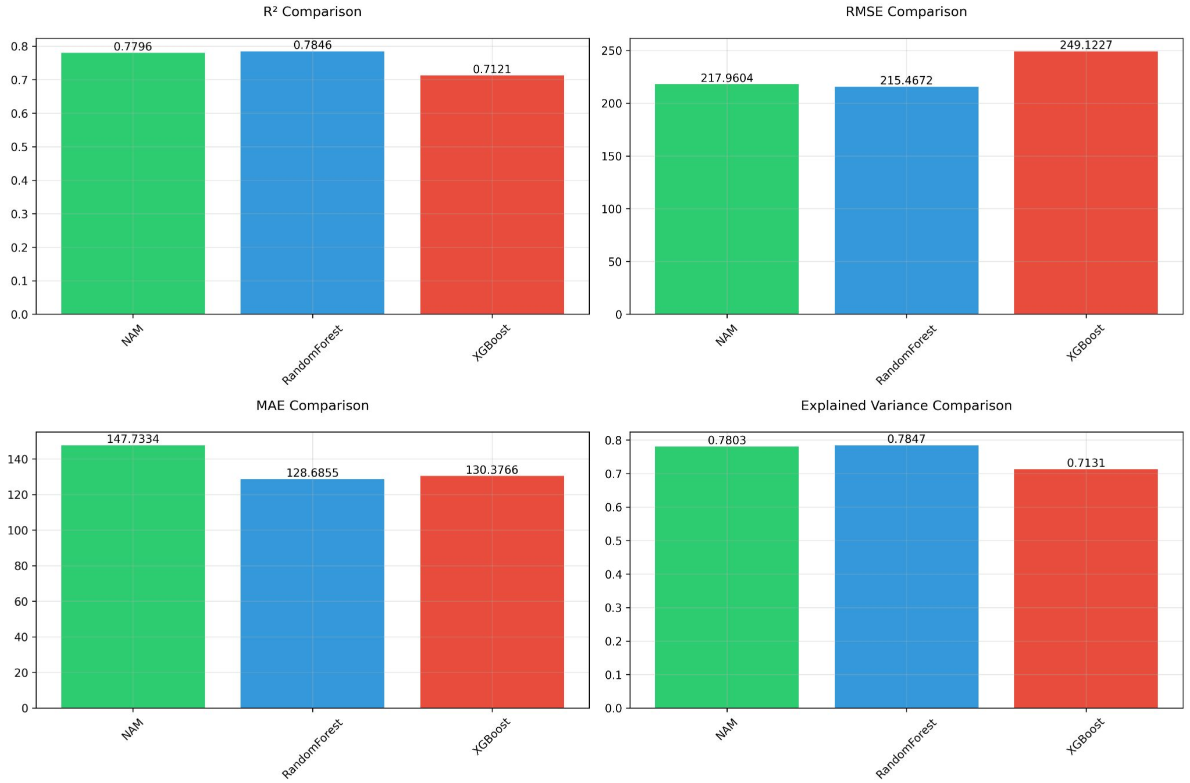


Baseline Models

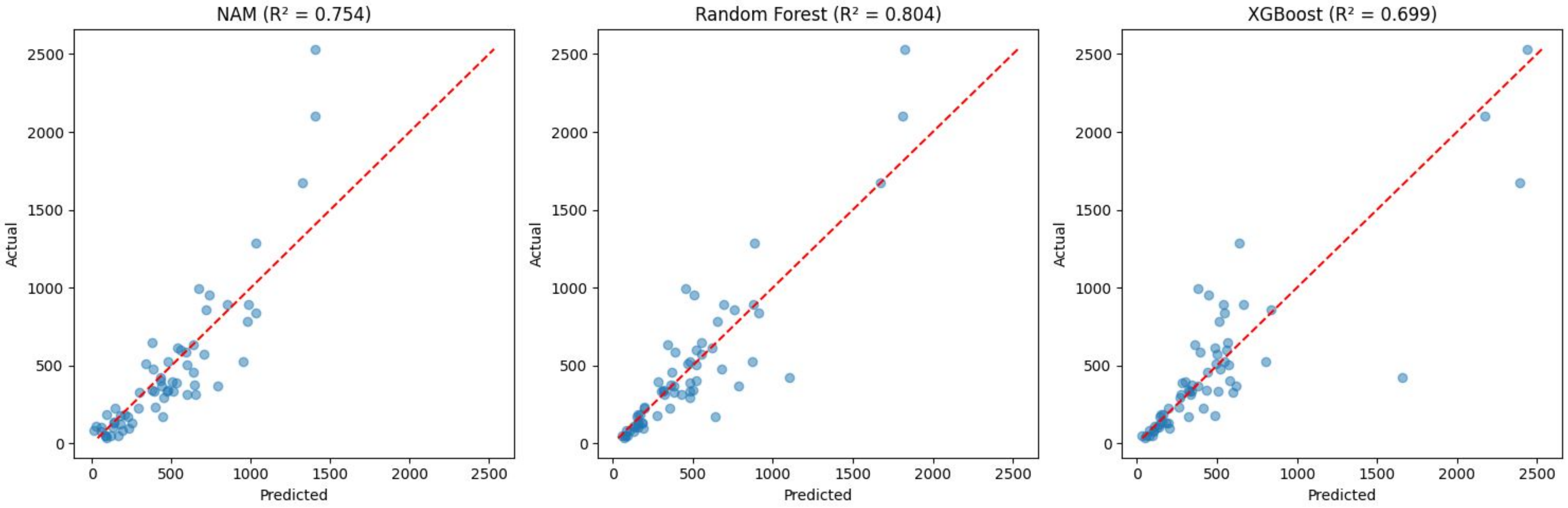
```
RandomForest:
Parameters:
{
  "n_estimators": 100,
  "max_depth": 10,
  "min_samples_split": 2
}

Metrics:
R²: 0.7846
RMSE: 215.4672
MAE: 128.6855
Explained Variance: 0.7847

XGBoost:
Parameters:
{
  "n_estimators": 100,
  "max_depth": 6,
  "learning_rate": 0.1
}

Metrics:
R²: 0.7121
RMSE: 249.1227
MAE: 130.3766
Explained Variance: 0.7131
```

Compare Models

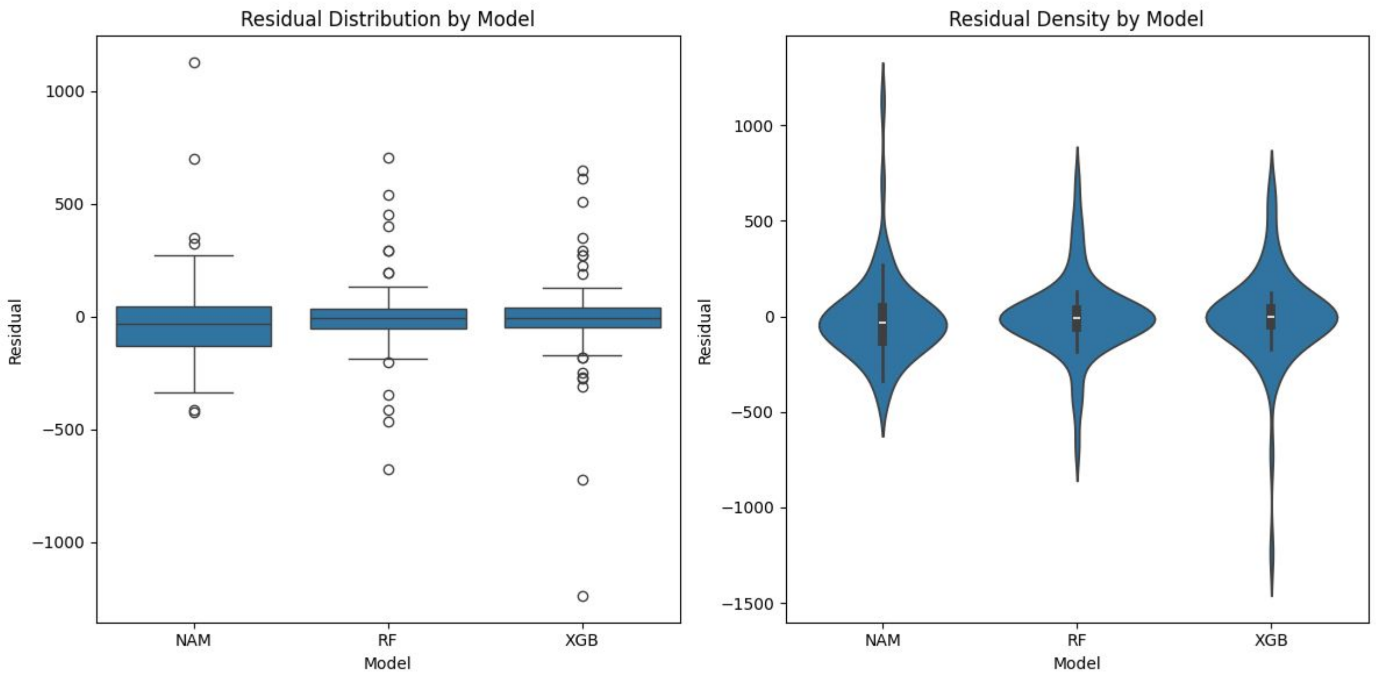
Comparing Models



Comparing Models



Comparing Models



Comparing Models

Model Comparison Summary

ANOVA Results:
F-statistic = 0.0486, p-value = 0.9526

Model	RMSE	MAE	R^2
NAM	230.146235	146.824781	0.754276
RF	205.631512	125.483154	0.803836
XGB	254.609167	140.471791	0.699263

Comparing Models

Tukey's HSD Test Results

Group 1	Group 2	Mean Diff	p-value	Lower	Upper	Reject
NAM	RF	12.9057	0.948	-85.07	110.8813	False
NAM	XGB	7.0769	0.9841	-90.8987	105.0526	False
RF	XGB	-5.8287	0.9892	-103.8044	92.1469	False

SHAP Analysis

Random Forest

lw (Wall Length)

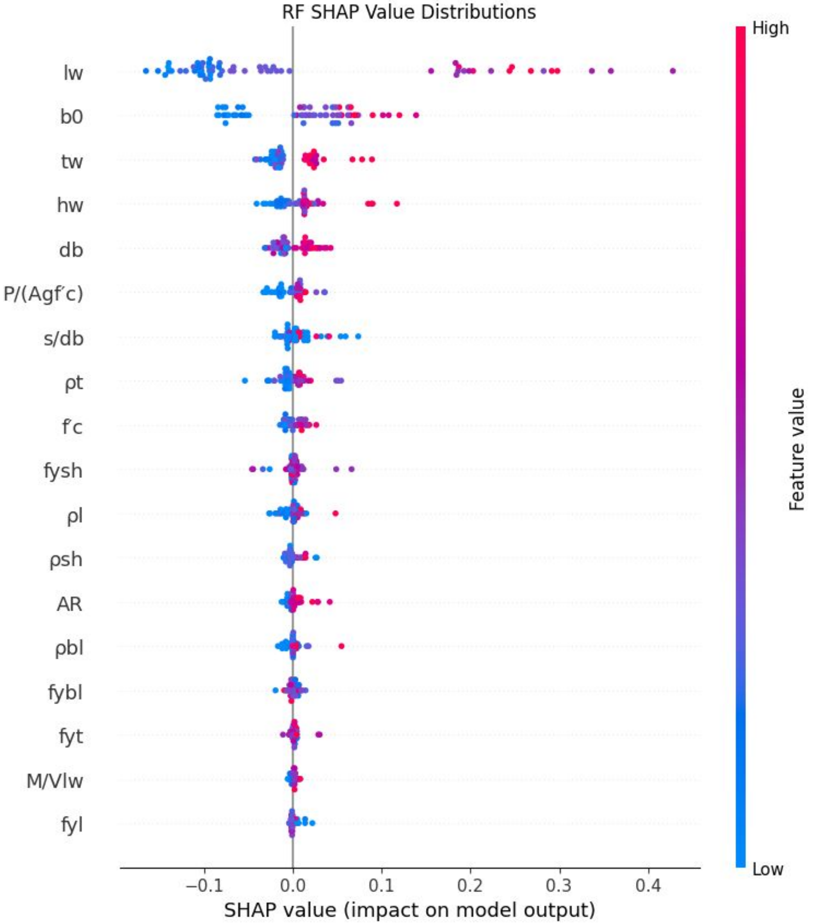
The horizontal span of the wall.

- **Higher wall length** (pink) leads to **positive SHAP values** → increases NCDE.
- **Interpretation**: Longer walls can dissipate more seismic energy due to larger in-plane stiffness and force distribution.

b0 (Boundary Element Depth)

Width of the confined zone at wall edges.

- **Deeper boundary elements** result in **higher predicted NCDE**.
- **Interpretation**: Indicates better confinement, enhancing ductility and energy dissipation.



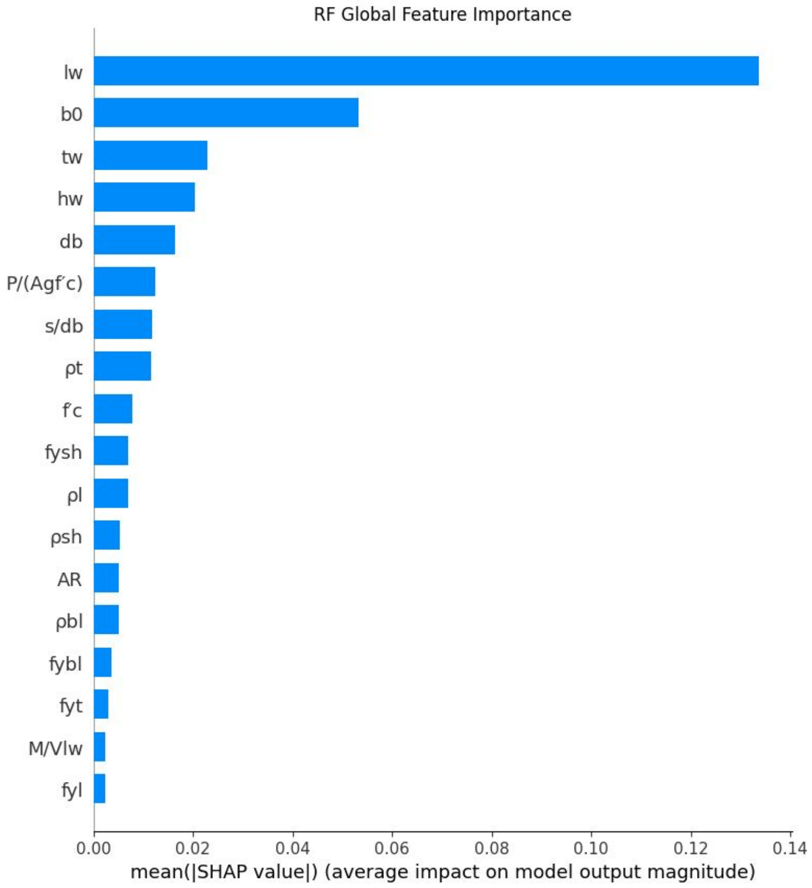
Random Forest

Wall length (lw) is the most influential feature in predicting NCDE according to the Random Forest model. Longer walls typically improve lateral load resistance and energy dissipation capacity.

Boundary element width (b0) and **wall thickness (tw)**
Both relate to the wall's cross-sectional robustness and its ability to resist bending and shear forces.

Wall height (hw) and **longitudinal bar diameter (db)** geometric slenderness and reinforcement size significantly affect structural performance.

Features like **reinforcement yield strengths (fyl, fyt, fybl)** and **axial load ratio (M/Vlw)** have minimal contribution, suggesting that geometry and reinforcement size dominate over material strength in this context.

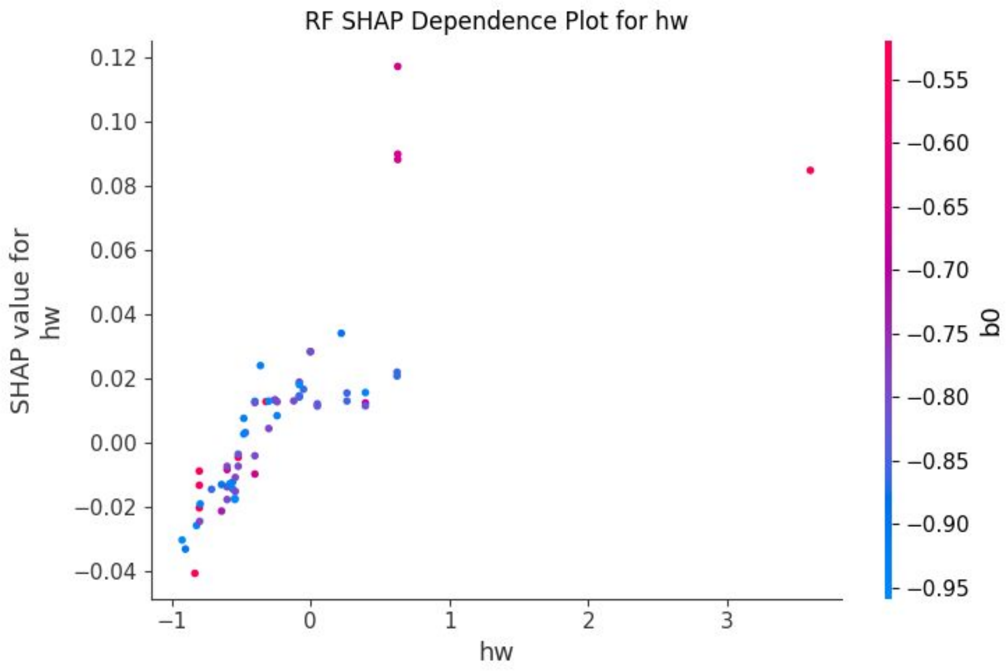


Random Forest

Feature analyzed: hw (*Wall height*)
This plot shows how different values of hw influence the model's output. The color represents the value of b0 (*boundary element width*), another important feature.

Interpretation:
As hw increases, its SHAP value tends to increase—indicating a **positive and roughly monotonic relationship**. Taller walls generally have a stronger positive impact on the predicted NCDE.

Interaction insight:
The color gradient suggests an interaction:
Smaller b0 Boundary width values (bluer) are associated -> higher SHAP values for hw
Wider boundary elements (redder) slightly suppress the contribution of hw.
<u>This implies that **tall walls with narrow boundaries** contribute the most positively.</u>

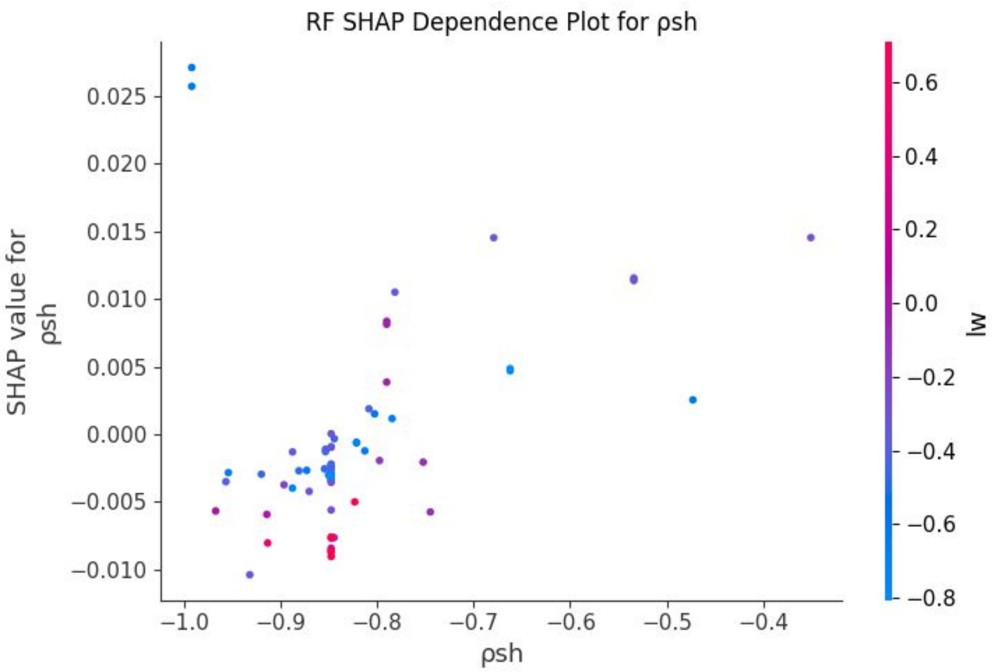


Random Forest

Feature analyzed: `psh` (*Transverse reinforcement ratio*)
This plot shows how changes in `psh` affect the prediction output, with color indicating the value of `lw` (*wall length*).

Interpretation:
The effect of `psh` on the model output is mostly small (SHAP values near zero), but becomes slightly more positive for higher values of `psh`. However, this impact is **modulated by wall length** (`lw`). Longer walls (warmer colors) slightly reduce the effect of `psh`.

Interaction insight:
The color gradient shows that **shorter walls** (bluer dots) tend to amplify the positive contribution of `psh`, whereas longer walls suppress it. This suggests a mild interaction between `psh` and `lw`.

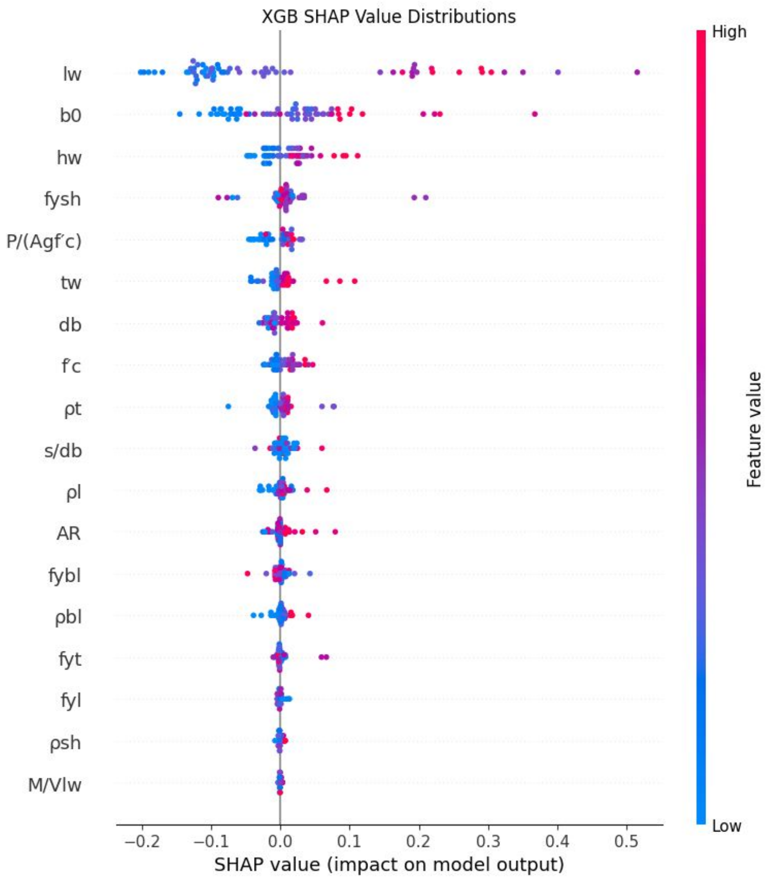


XGboost

Top features like wall length (lw) and boundary width (b0) have the widest SHAP value spread, meaning they strongly affect model predictions. Higher values (red) generally increase predicted NCDE.

Color gradients show interaction: for example, high `fysh` (transverse reinforcement strength) pushes predictions upward, supporting better energy dissipation.

Bottom features like `psh`, `fyt`, and `M/Vlw` cluster near zero SHAP values, indicating little influence on the model across the dataset.



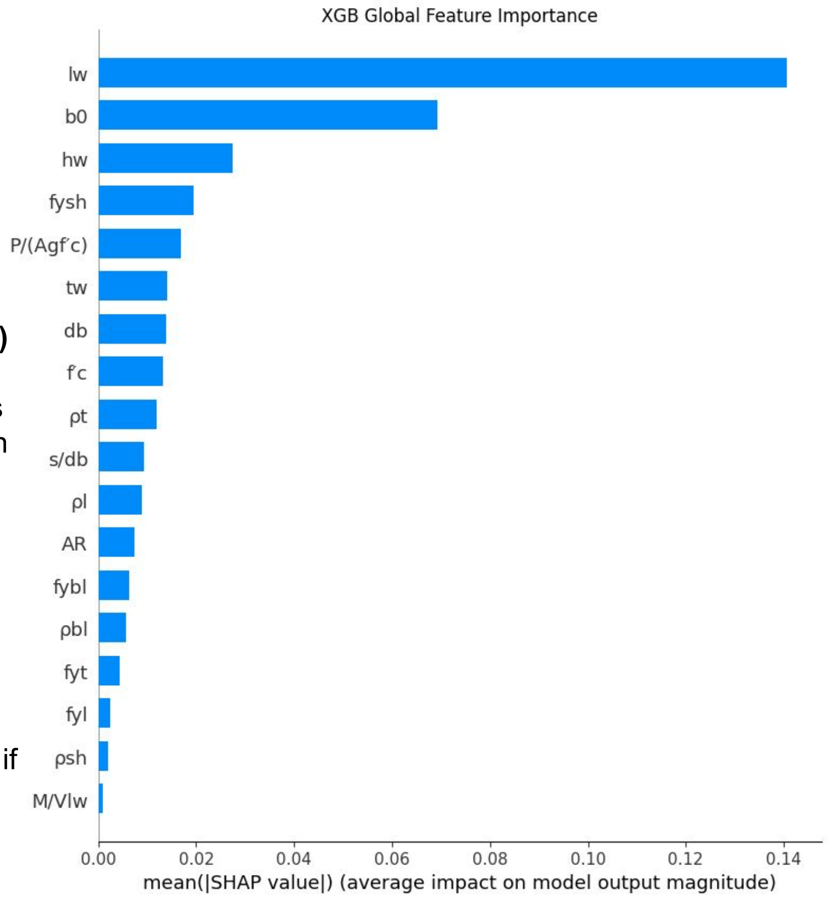
XGboost

Wall length (lw) and **boundary width (b0)** are the most impactful features in the XGBoost model, strongly influencing the prediction of NCDE. Their dominance suggests that the wall's geometry governs energy dissipation behavior.

Wall height (hw) and **transverse reinforcement yield strength (fysh)** also contribute substantially. A high **fysh** means the transverse (horizontal) reinforcement in boundary zones can endure greater stress before yielding, enhancing ductility and energy absorption. Low fysh, on the other hand, may lead to early yielding and reduced seismic performance.

Mid-level features such as **axial load ratio (P/(Agf′c))**, **wall thickness (tw)**, and **bar diameter (db)** play moderate roles, capturing load conditions and reinforcement scale.

The bottom features (e.g., **M/Vlw**, **psh**, **fyl**) show very low average SHAP values, suggesting limited influence in the learned model—even if they may be physically relevant.

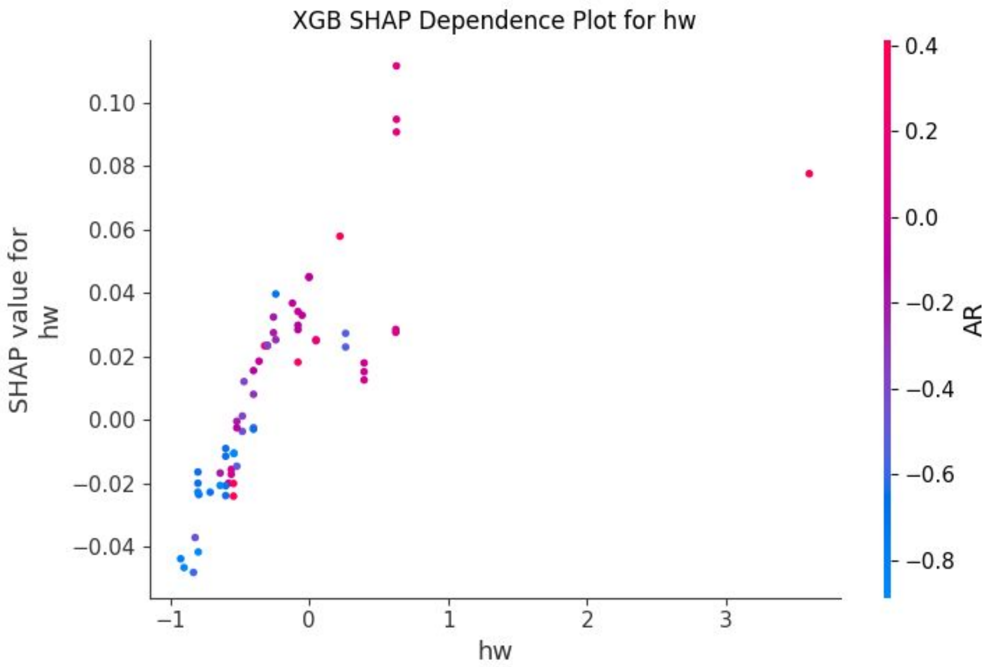


XGboost

Positive effect: As hw (wall height) increases, its SHAP value also increases, meaning **taller walls generally contribute positively** to predicted NCDE. The function is **monotonic and mostly linear**.

Interaction with AR (Aspect Ratio): Color represents AR, and we see that **lower AR values (blue)** slightly reduce the contribution of hw. Higher AR (pink-red) points tend to have a **stronger positive SHAP impact**.

Key insight: Wall height is a consistent positive contributor to energy dissipation, especially when aspect ratio is moderate to high — confirming the importance of geometry in structural performance.

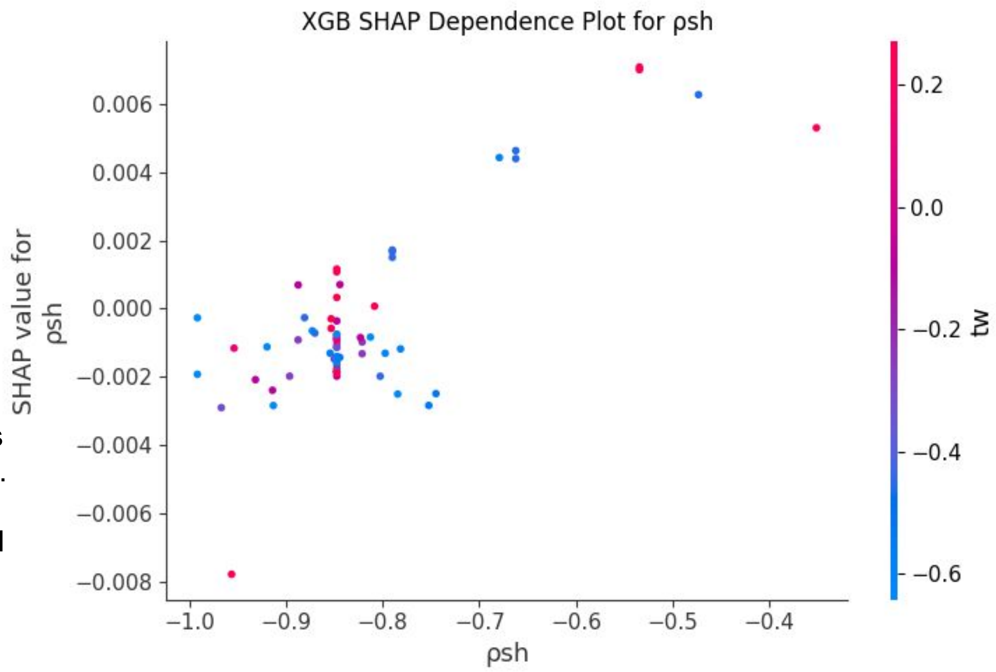


XGboost

Low importance of psh (transverse reinforcement ratio in shear zone): This variable represents the amount of horizontal reinforcement resisting shear cracks. Across its range, SHAP values stay near zero, meaning psh has minimal impact on the XGBoost model's predictions of NCDE.

Slight positive effect for high psh: At higher values of psh, the SHAP value increases slightly — suggesting that increasing shear reinforcement may contribute weakly to energy dissipation, especially in specific cases.

Interaction with `tw` (wall thickness): The color bar shows `tw`, which measures the thickness of the wall cross-section. Redder dots (thicker walls) align with higher SHAP values for psh, indicating that psh may matter more when the wall is thick enough to anchor and mobilize that reinforcement effectively.

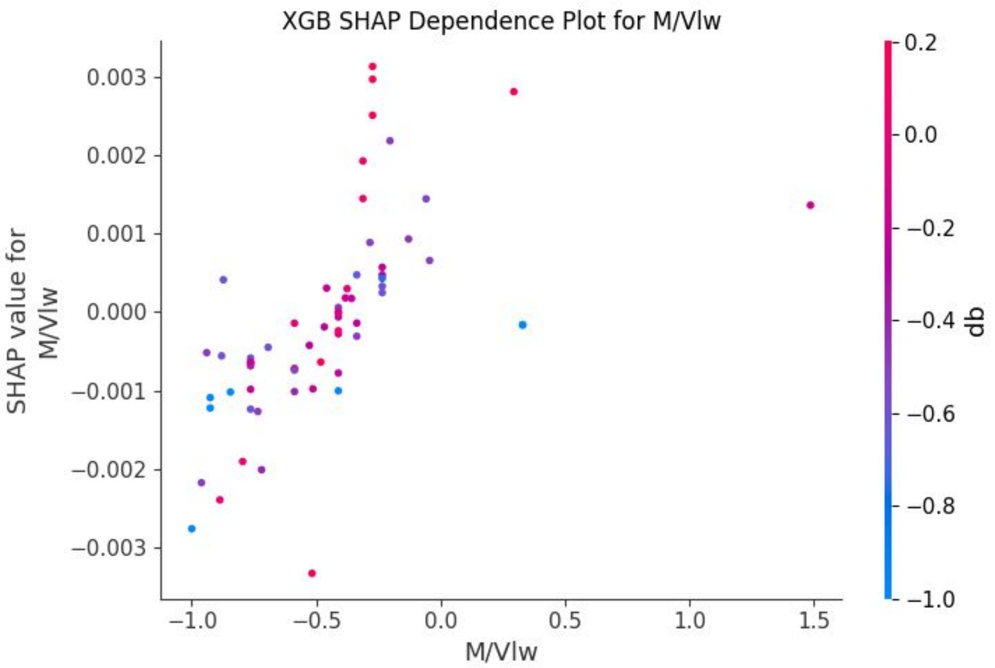


XGboost

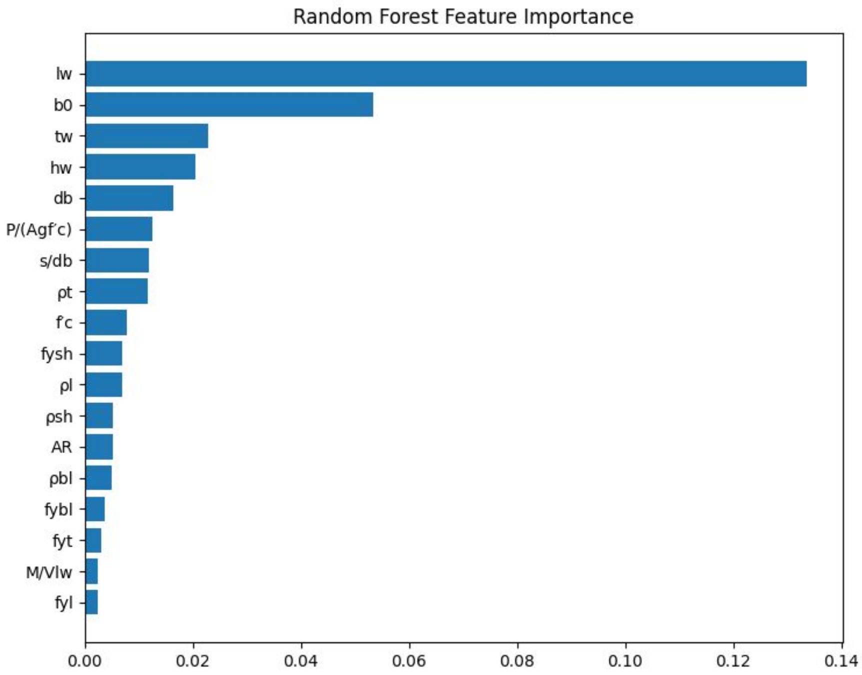
- **M/Vlw (Moment-to-shear ratio normalized by wall length)** reflects the balance between bending and shear demands on the wall. Higher values imply flexure-dominated behavior, while lower values suggest shear dominance.

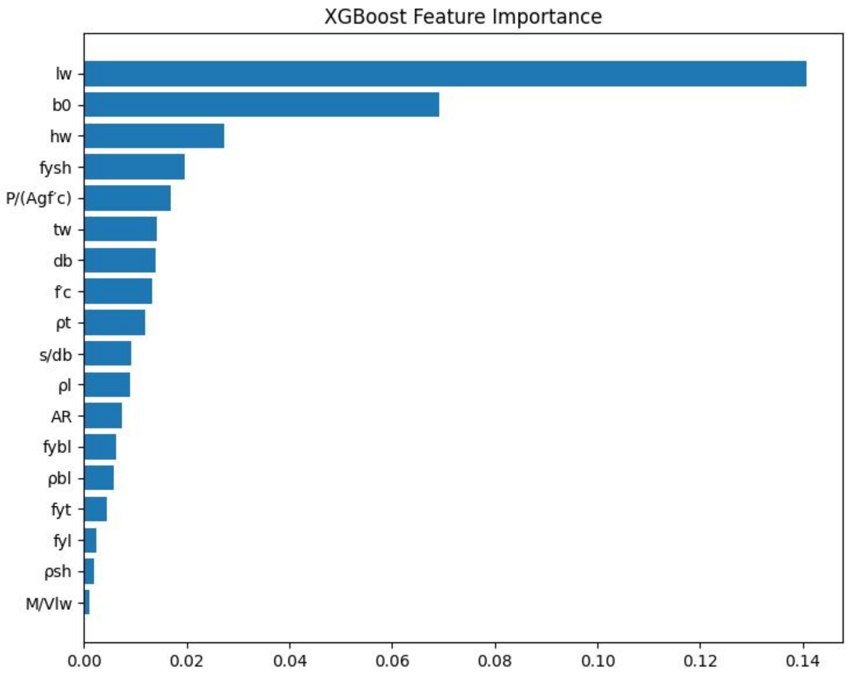
- **Positive trend:** As `M/Vlw` increases, SHAP values also rise — meaning **walls under more flexural demand tend to have slightly higher predicted NCDE**, though the effect is weak overall.

- **Interaction with db (diameter of longitudinal bars):** The color bar shows that **larger db values (redder points)** tend to amplify the positive effect of `M/Vlw`. This makes sense, as thicker longitudinal bars are more effective in flexural resistance, boosting energy dissipation in bending-dominated regimes.



Comparison


Random Forest Feature Importance
lw
b0
tw
hw
db
P/(Agf'c)
s/db
ρt
f'c
fysh
ρl
ρsh
AR
ρbl
fybl
fyt
M/Vlw
fyl
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14


XGBoost Feature Importance
lw
b0
hw
fysh
P/(Agf'c)
tw
db
f'c
ρt
s/db
ρl
AR
fybl
ρbl
fyt
fyl
ρsh
M/Vlw
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14

XGboost

Consistent top features: Both models rank **wall length (lw)** and **boundary element width (b0)** as the most important features. These geometric parameters heavily influence structural behavior and energy dissipation capacity (NCDE).

Variation in mid-tier importance: XGBoost gives more weight to **wall height (hw)** and **transverse reinforcement yield strength (fysh)**, suggesting it captures more complex nonlinear interactions related to wall aspect and material strength.

Random Forest emphasizes tw (wall thickness) and **db (diameter of longitudinal bars)** slightly more, reflecting its tendency to favor direct and localized feature splits.

Low-impact features across both models include **fyt**, **fyl**, **psh**, and **M/Vlw** — indicating these parameters have a minimal average influence on NCDE within this dataset's range, even if they are physically meaningful.

Effect Comparison

Effect Comparison

Top 5 Important Features

	NAM	RF	XGB
1	hw	lw	lw
2	AR	b0	b0
3	lw	tw	hw
4	ρbl	hw	fysh
5	tw	db	P/(Agf′c)

Effect Comparison

NAM prioritizes `hw` (wall height) and `AR` (aspect ratio), showing its focus on global wall proportions. `pbl` (plastic hinge length at the base) also ranks high, highlighting NAM's sensitivity to deformation mechanisms and ductility modeling.

Random Forest emphasizes `lw` (wall length) and `b0` (boundary element width), consistent with its greedy split-based structure. It also includes `tw` (wall thickness) and `db` (diameter of longitudinal bars), suggesting a focus on geometric and reinforcement cross-section parameters.

XGBoost aligns with RF at the top (`lw`, `b0`), but includes more material-sensitive features like `fysh` (transverse reinforcement yield strength) and `P/(Agf'c)` (normalized axial load). This reflects XGBoost's ability to capture nuanced interactions between loading and reinforcement capacity.

Overall, **geometric features dominate across all models**, but **NAM captures structural performance patterns**, while **tree-based models focus on split-efficient predictors**.

Effect Comparison

Directional Disagreement

	Feature	NAM	RF	XGB
1	f′c	negative	positive	positive
2	fyt	positive	positive	negative
3	fyl	positive	negative	negative
4	ρsh	negative	positive	positive
5	ρl	negative	positive	positive
6	db	negative	positive	positive
7	s/db	negative	positive	negative
8	AR	negative	positive	positive

Effect Comparison

Concrete Strength (f′c):
RF & XGB see it as helpful. NAM sees it as harmful — maybe because stronger concrete is less ductile during earthquakes.

Yield Strength (fyt, fyl) & ρsh:
Mixed signals. NAM and RF say fyt helps; XGB says it hurts. Stiffer steel might limit energy dissipation in some cases.

Geometry (AR, s/db):
RF & XGB prefer tall walls and tight stirrups. NAM disagrees — may link these to failure under stress.

Effect Comparison

Directional Agreement

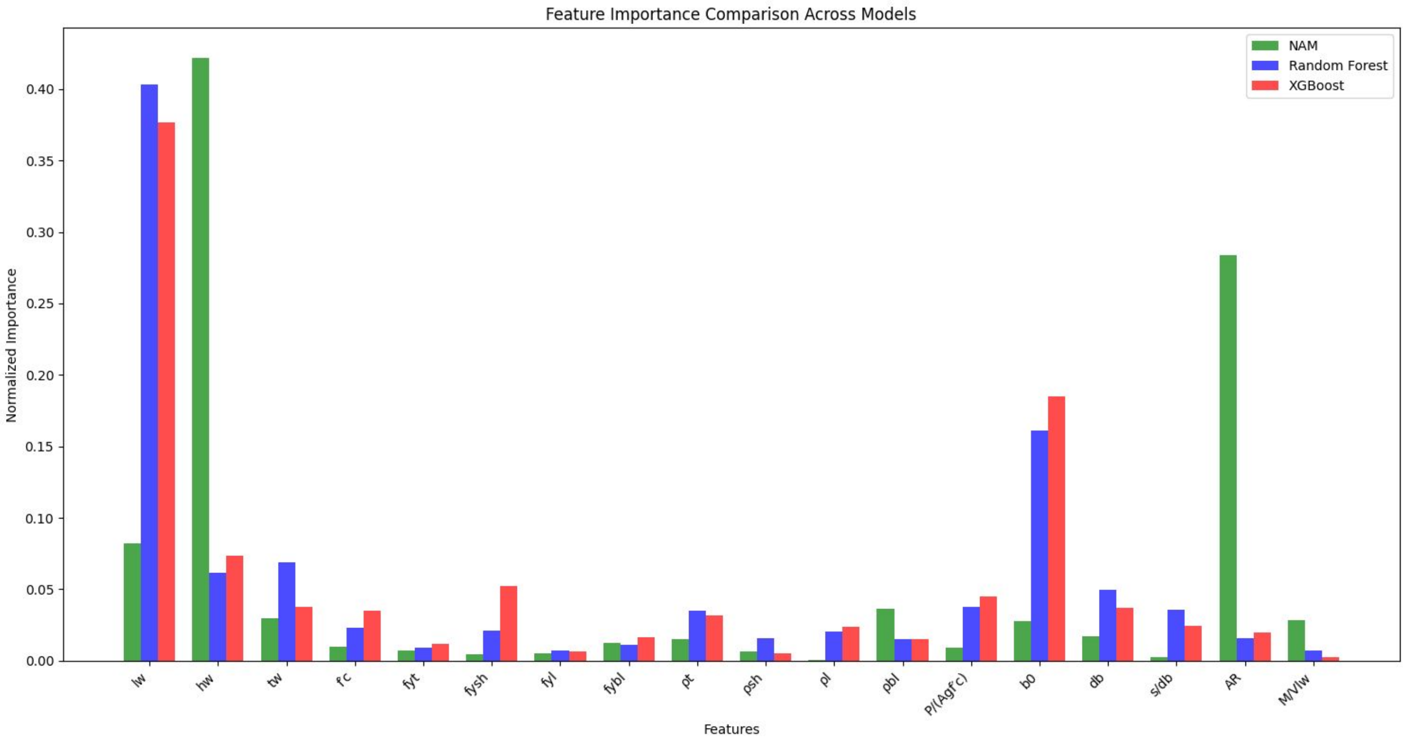
	Feature	Effect	Agreement
1	lw	positive	Agree
2	hw	positive	Agree
3	tw	positive	Agree
4	fysh	positive	Agree
5	fybl	negative	Agree
6	ρt	positive	Agree
7	ρbl	positive	Agree
8	P/(Agf'c)	positive	Agree
9	b0	positive	Agree
10	M/Vlw	positive	Agree

Effect Comparison

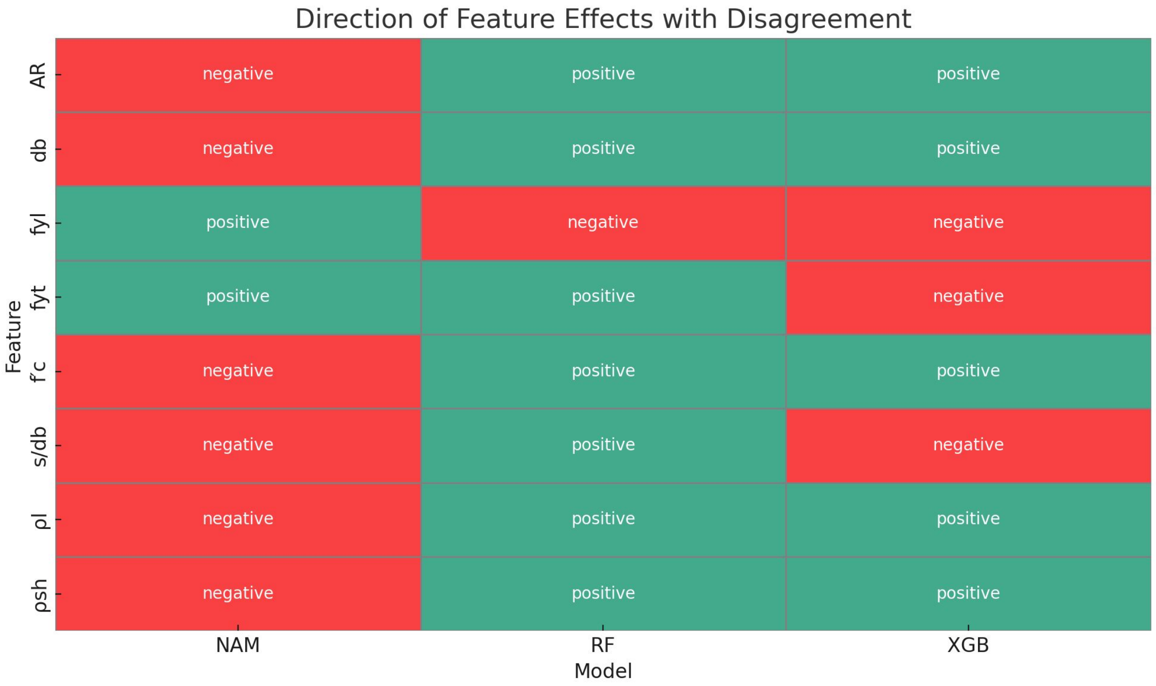
Across all models, there is **consistent directional agreement** on the most influential features. Geometric variables like **lw** **(wall length)**, **hw** **(wall height)**, and **tw** **(wall thickness)** all show a **positive effect** on NCDE, indicating that <u>larger wall dimensions enhance energy dissipation</u>. Material properties such as **fysh** (transverse reinforcement strength) and **pt** (axial reinforcement ratio) also contribute positively.

Interestingly, **fybl** (longitudinal boundary reinforcement strength) shows a **negative effect**, suggesting that beyond a certain threshold, stiffer boundary elements may reduce ductility. All models agree not only on the direction but also the relevance of these key features, reinforcing the robustness of the findings.

Effect Comparison



Effect Comparison

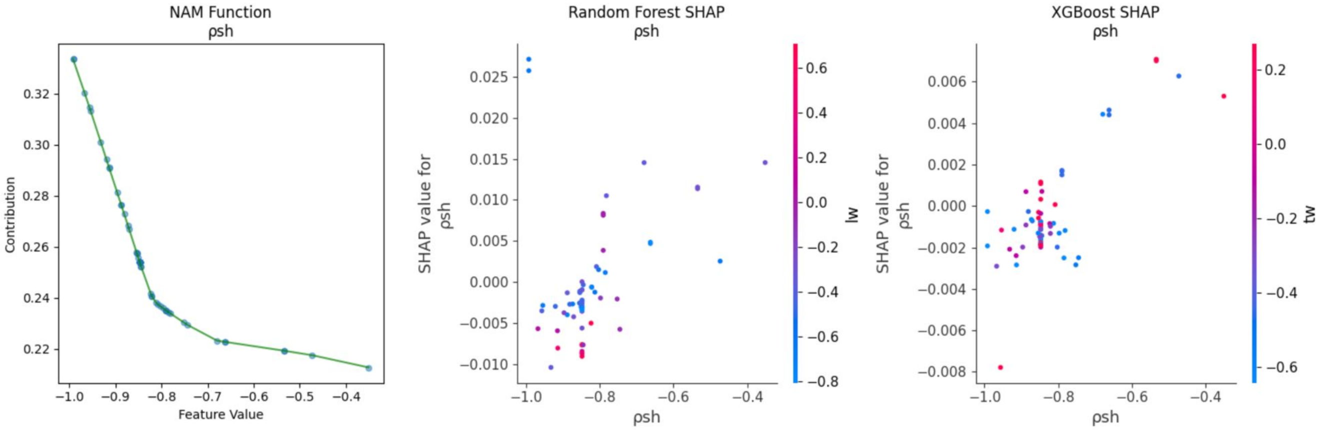


Effect Comparison: transverse shear reinforcement ratio

a measure of how much steel reinforcement is placed transversely (i.e., sideways) across a concrete element to resist shear forces

The NAM model shows a **clear negative effect** of ρsh on the output, while both SHAP explanations from Random Forest and XGBoost suggest a **positive effect**, meaning the models **disagree on the direction** of influence.

Feature: ρsh



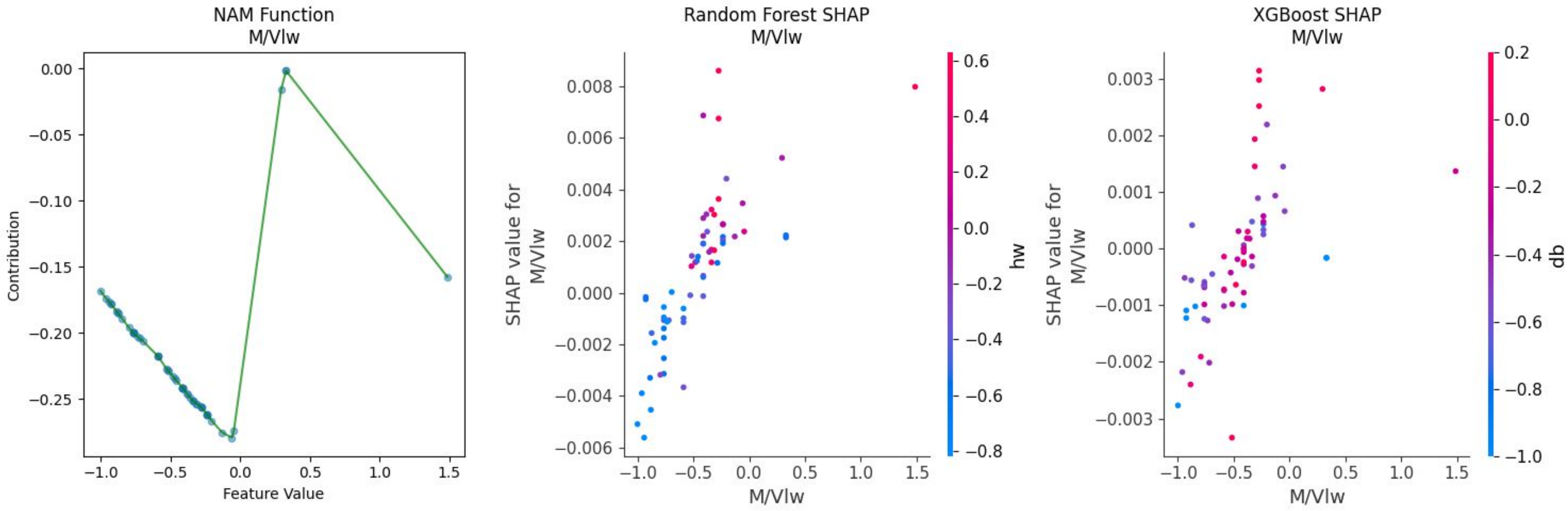
Direction agreement: No

Effect direction:

- NAM: negative
- RF: positive
- XGB: positive

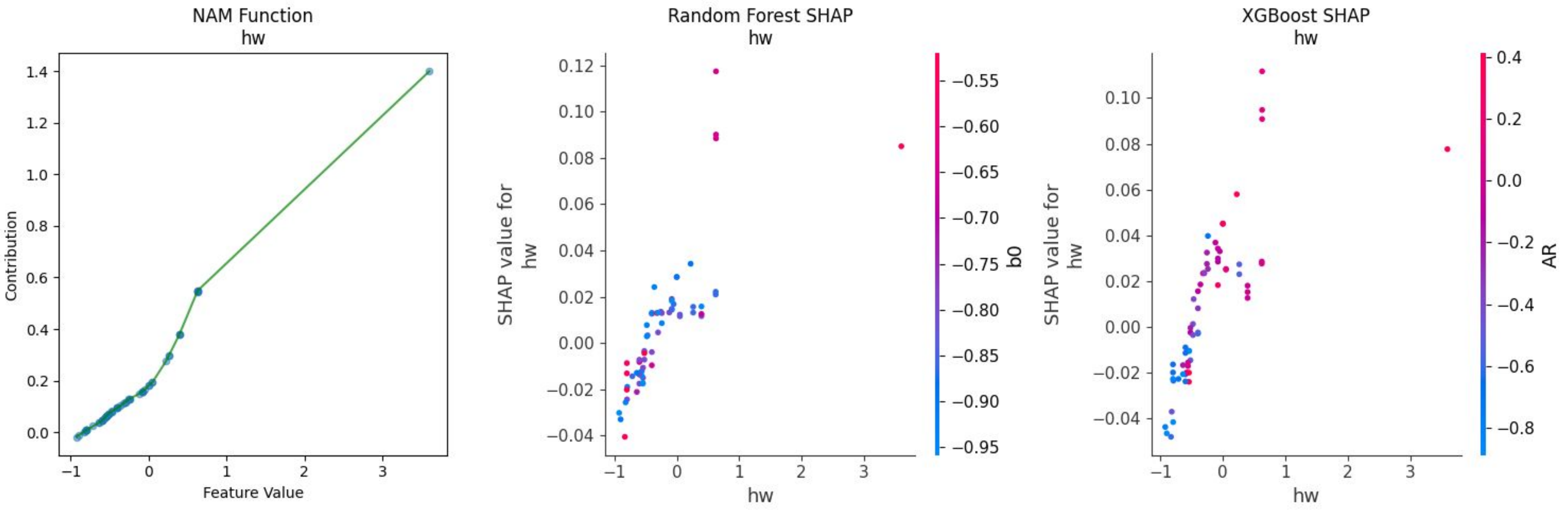
Effect Comparison:

All three models show that **higher M/Vlw values** (more flexure-dominated behavior) tend to **slightly increase energy dissipation (NCDE)**, but the effect is **weak and model-dependent** — with NAM showing a **nonlinear trend**, and RF/XGB showing a **mild positive correlation** modulated by wall height (hw).



Effect Comparison

All three models agree that **increasing wall height (hw) strongly boosts energy dissipation**, with NAM showing a clearly **monotonic and nonlinear increase**, while Random Forest and XGBoost show a **positive SHAP trend** influenced by boundary width (b0) and aspect ratio (AR) respectively.



Conclusion

Conclusion

Wall geometry matters most.
All models (NAM, RF, XGBoost) agree that wall length (`lw`), wall height (`hw`), and boundary width (`b0`) are the most important features for predicting NCDE.

Bigger walls = better performance.
Longer, taller, and thicker walls tend to **increase energy dissipation**, which matches engineering intuition.

Models agree on most effects.
For 10 key features, all models agree whether they have a positive or negative effect on NCDE — this shows strong directional consistency.

Some features are more complex.
Features like concrete strength (`f'c`), reinforcement yield strengths (`fyt`, `fyl`), and aspect ratio (`AR`) cause disagreement between models, likely due to how each one learns patterns differently.

Conclusion

Wall geometry is the most decisive factor in energy dissipation.

NAMs are better at modeling *deformation* and *failure modes*.

Tree models (RF/XGB) excel at identifying *threshold-based rules* from geometry and materials.

Next Steps

- **Investigating why XGBoost behaved worse**
- **Focusing more on Feature Engineering**

Thank you!