

Noise-Resistant Training for Sequence-to-Expression Models

Systematic Evaluation on CAGI5 Clinical Variant Benchmark

74

Models

+1.2%

Best CAGI5
Improvement

-0.088

Best Noise
Correlation

 Rank Stability

 Contrastive

 Distributional

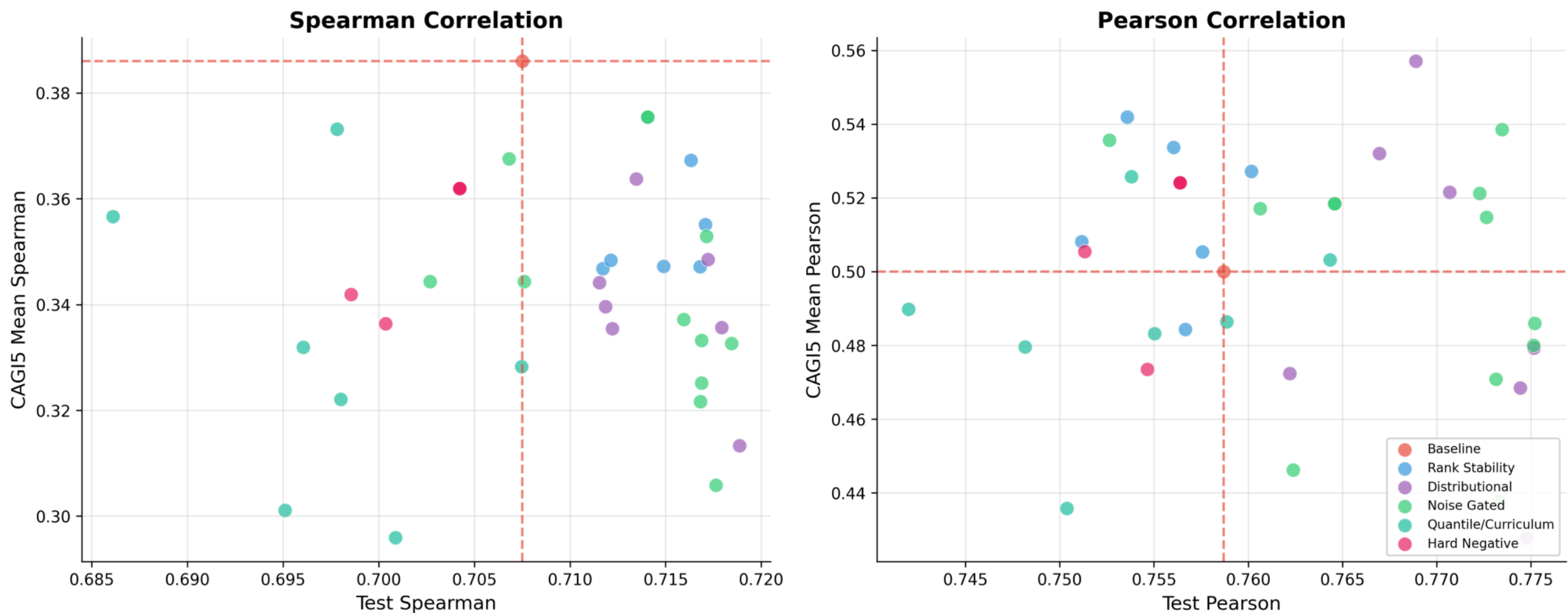
 Quantile/Curriculum

 Noise Gated

 Hard Negative

Figure 1: Test Performance vs CAGI5 Generalization

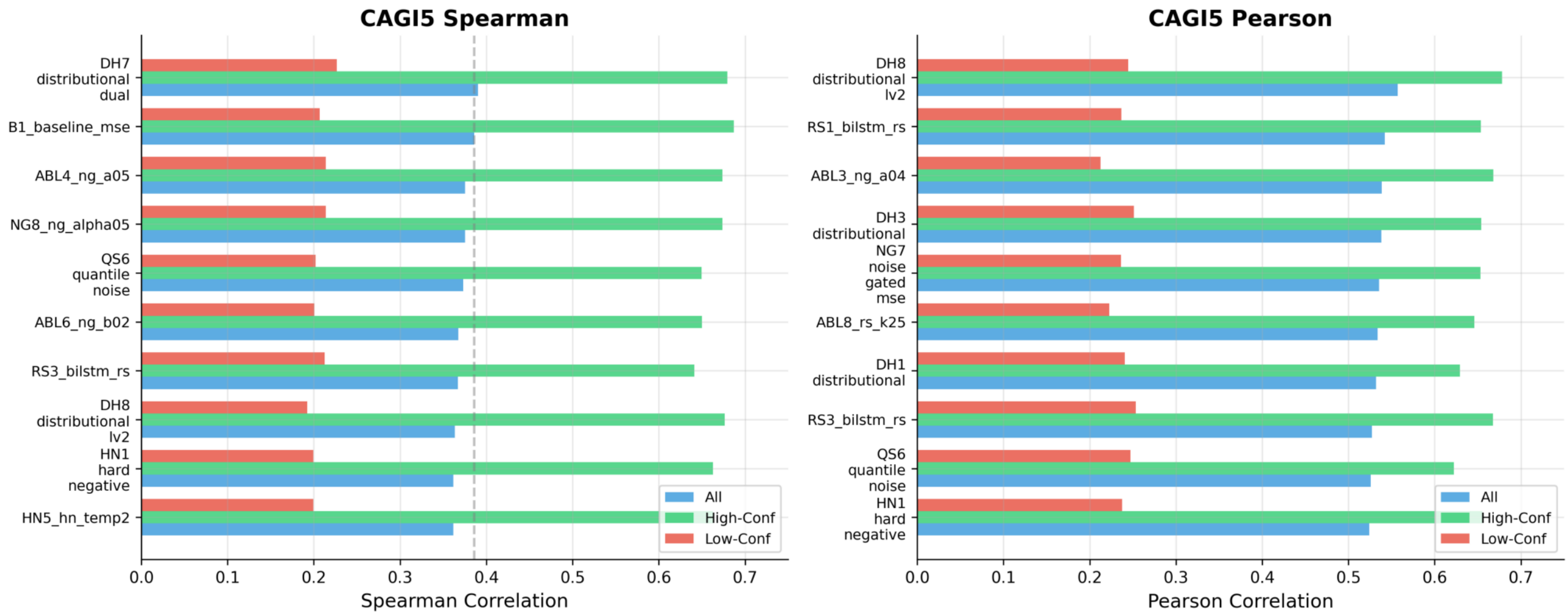
Test Performance vs CAGI5 Generalization



Both Spearman (left) and Pearson (right) correlations shown. Dashed red lines indicate baseline.
Key Finding: Higher test performance does not guarantee better CAGI5 generalization.

Figure 2: CAGI5 by Confidence Level

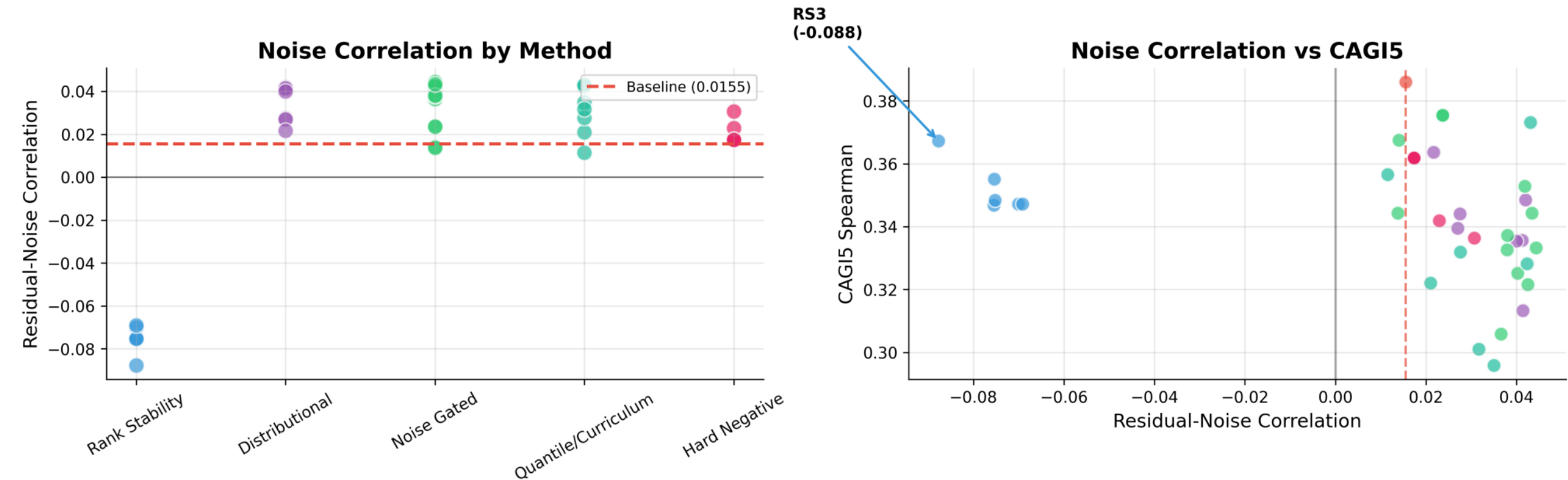
Top 10 Models: CAGI5 Performance by Confidence Level



Top 10 models ranked by Spearman (left) and Pearson (right). Bars show All/High-Conf/Low-Conf performance.
Key Finding: DH7 achieves best overall with +10% improvement on low-confidence variants.

Figure 3: Noise Avoidance Analysis

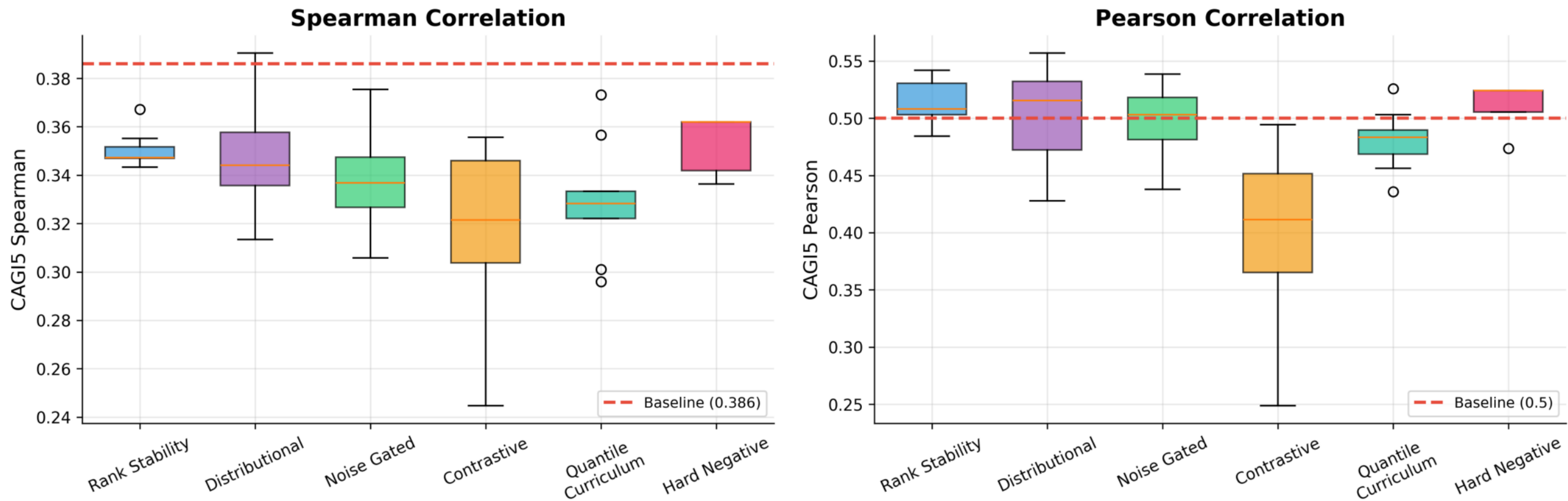
Noise Avoidance Analysis



Left: Distribution by method. Right: Noise correlation vs CAGI5 Spearman.
Key Finding: RS3 is the ONLY model with negative noise correlation (-0.088).

Figure 4: Method Category Comparison

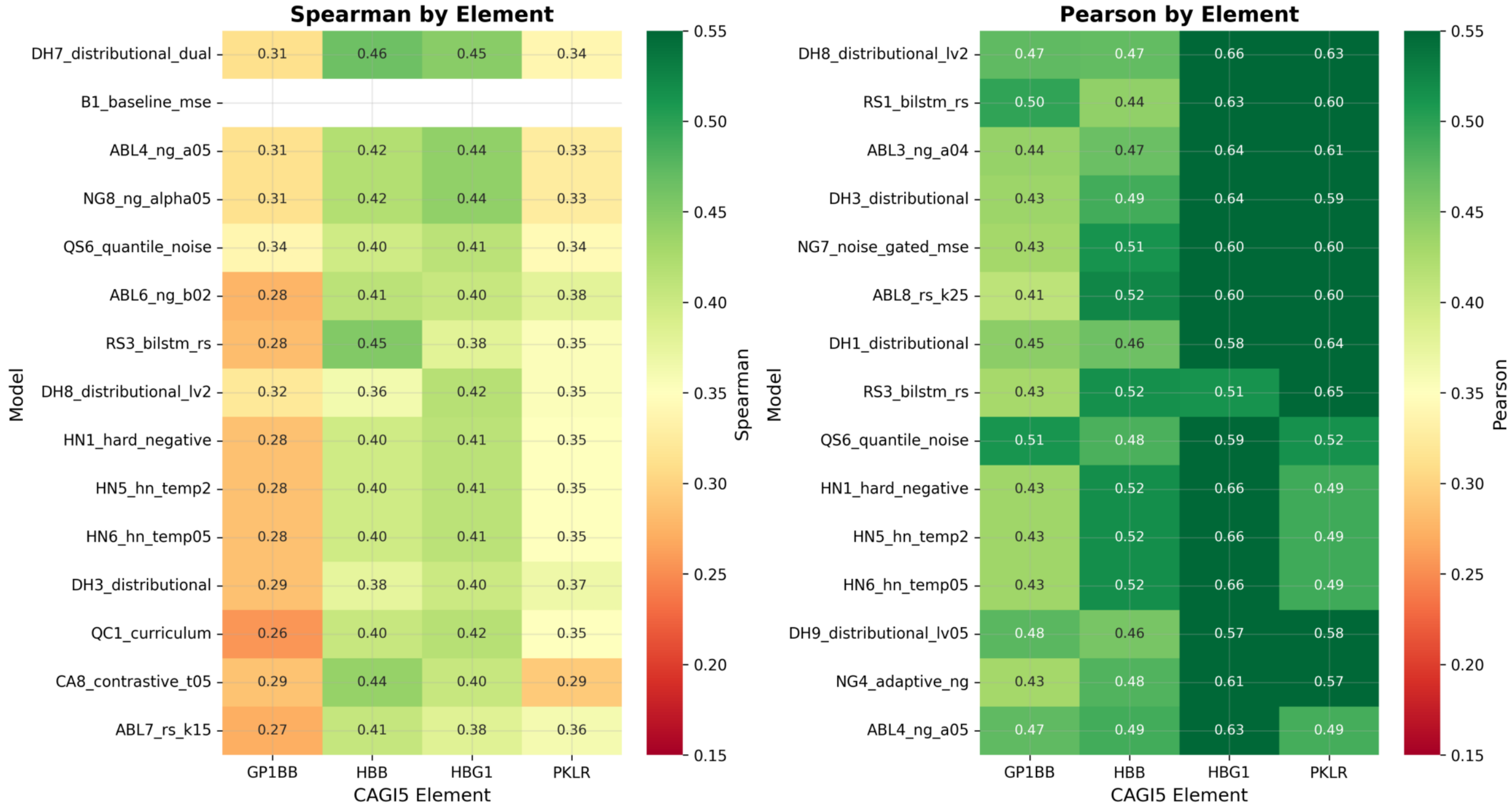
CAGI5 Performance by Method Category



Box plots for both Spearman (left) and Pearson (right). Red line = baseline.
Key Finding: Distributional methods achieve highest and most consistent improvements.

Figure 5: Per-Element CAGI5 Performance

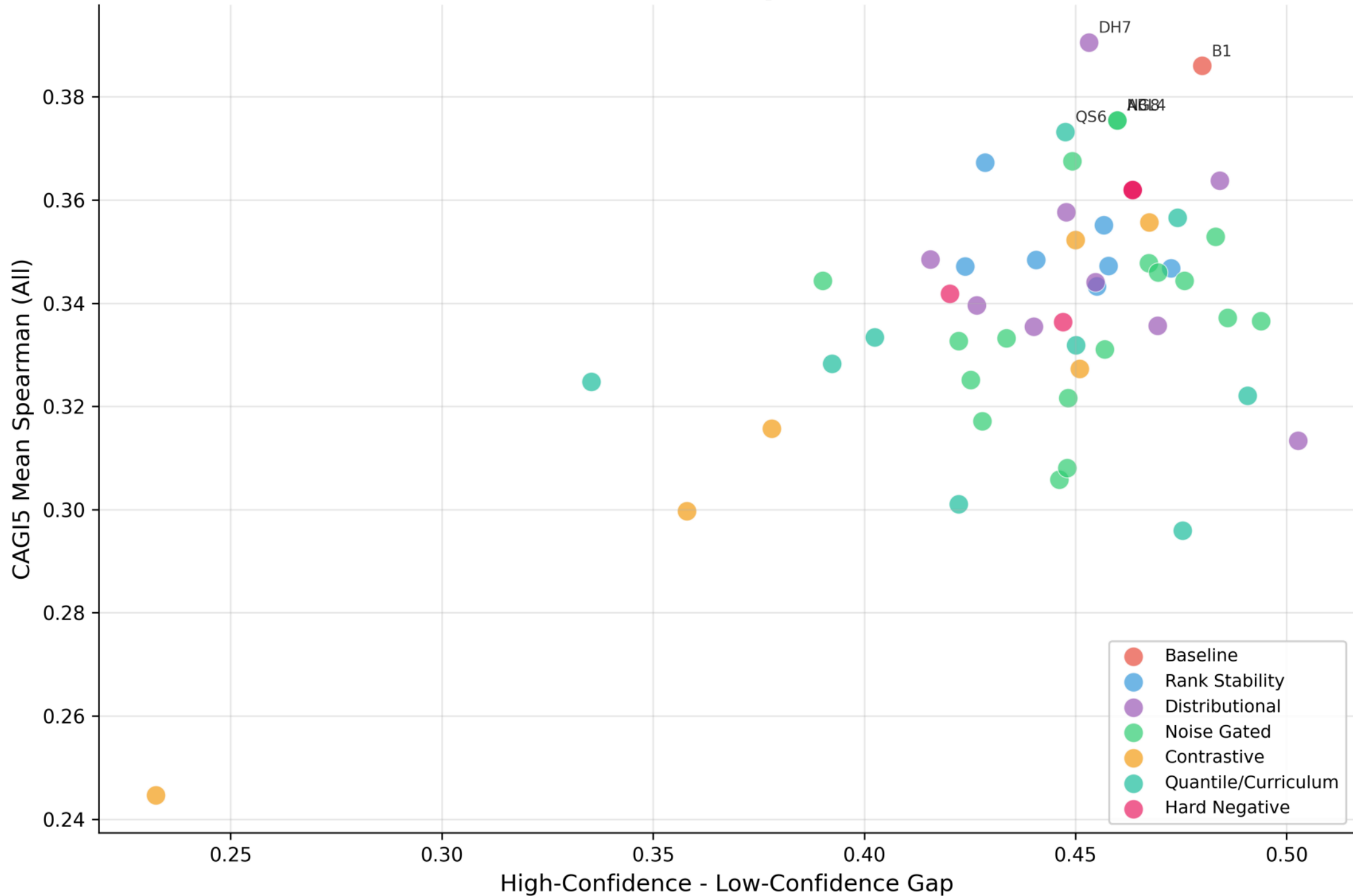
Per-Element CAGI5 Performance (Top 15 Models)



Heatmaps showing Spearman (left) and Pearson (right) for top 15 models across 4 CAGI5 elements.
Key Finding: GP1BB shows largest improvement (+108% for DH7). HBB consistently highest.

Figure 6: High-Confidence vs Low-Confidence Gap

HC-LC Performance Gap vs Overall CAGI5



Each point represents a model. X-axis: HC-LC gap. Y-axis: Overall CAGI5.
Key Finding: DH7 achieves smallest gap among top performers, indicating balanced performance.

Figure 7: Methods Architecture

Noise-Resistant Training Methods

Rank Stability (RS)

Weight pairs by noise reliability
 $w_{ij} = \text{sigmoid}(-k(\sigma_i^2 + \sigma_j^2))$

Distributional (DH)

Predict μ and σ^2 jointly
 $L = \text{MSE}(\mu, y) + \lambda \cdot \text{MSE}(\sigma^2, \text{noise})$

Noise Gated (NG)

Combine heteroscedastic +
rank stability losses

Contrastive (CA)

Noise-based similarity
for contrastive learning

Quantile Sampling (QS)

Stratified sampling across
activity + noise quantiles

Hard Negative (HN)

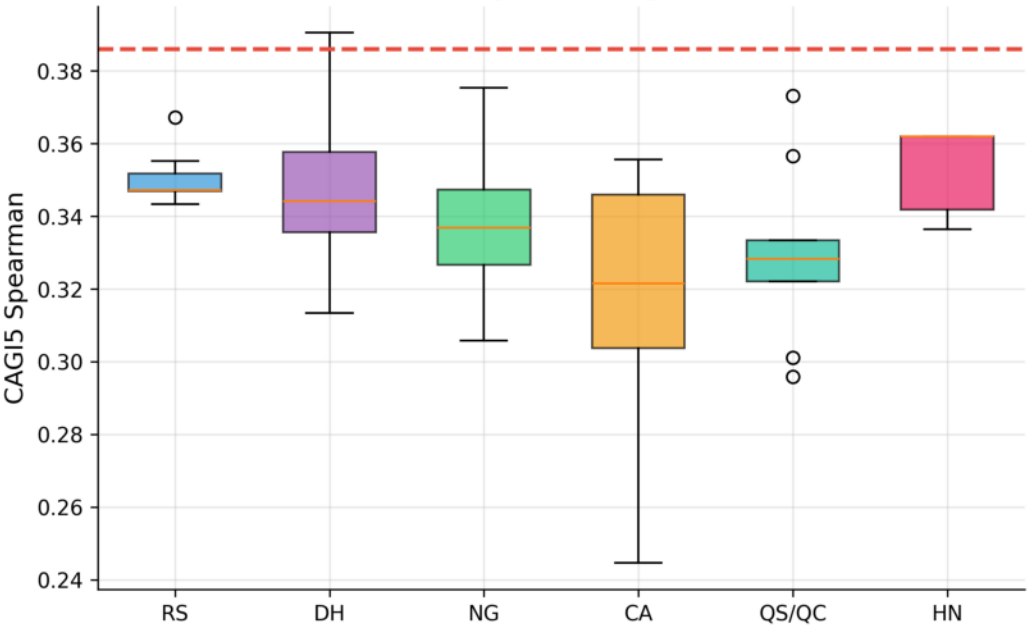
Mine informative pairs:
small $\Delta \text{activity}$ + low noise

Six noise-resistant training strategies: Rank Stability, Distributional, Noise Gated, Contrastive, Quantile Sampling, and Hard Negative Mining.

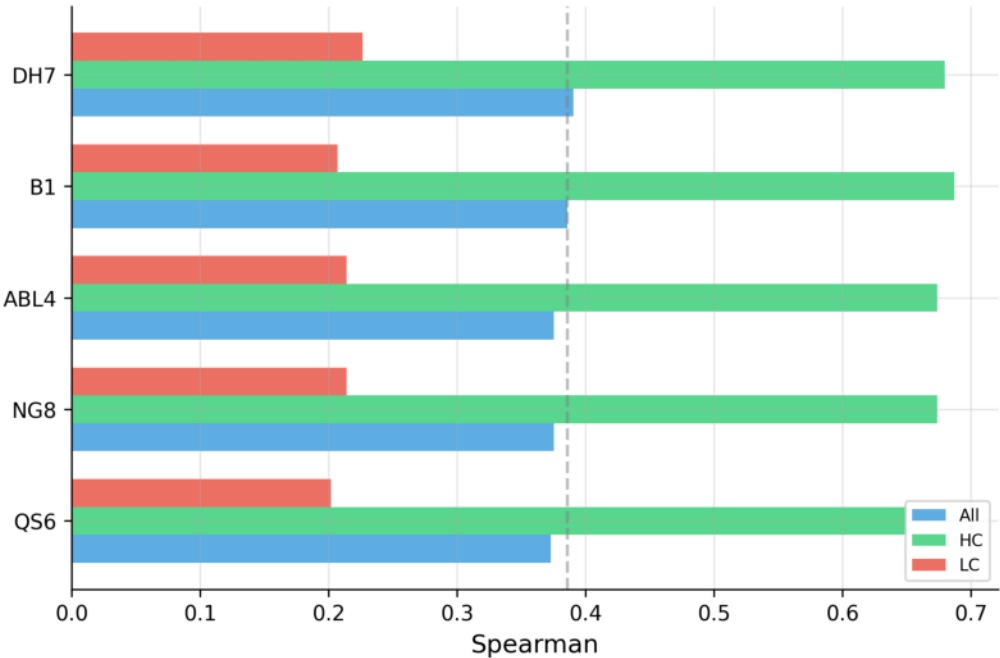
Figure 8: Executive Summary

Noise-Resistant Training: Executive Summary

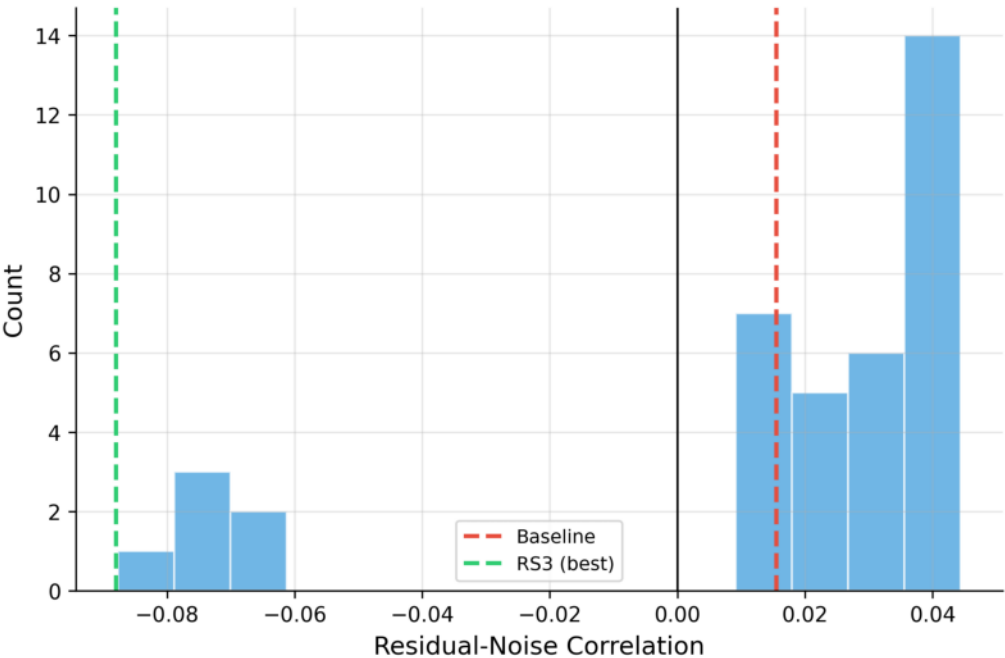
A. CAGI5 Spearman by Method



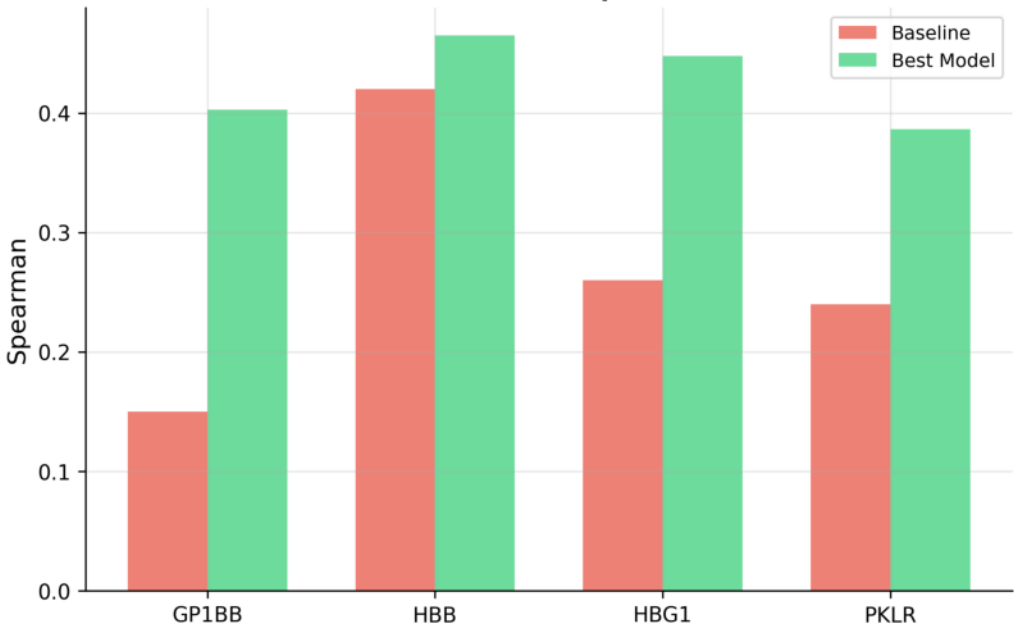
B. Top 5 Models



C. Noise Correlation Distribution



D. Per-Element Improvement



(A) Method comparison (B) Top 5 models (C) Noise correlation distribution (D) Per-element improvement.
Key Results: DH7 best CAGI5 (+1.2%), RS3 only negative noise correlation, GP1BB +108%.