VARIANCE-GUIDED TIME-SERIES CORRECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Post-hoc correction offers a lightweight paradigm to mitigate residual errors in frozen models. The central challenge is deciding which relationships finite calibration data can support, as similar residual energies often correspond to fundamentally different corrective values. Through multi-seed spectral diagnostics, we reveal that residual errors conflate persistent compensable patterns with stochastic training-run variations, motivating a correction rule that dynamically adapts its dependency structure. To resolve this, we formulate this choice as *covariance-resolution selection*, ranging from coordinate-wise scaling to channel, phase, and neighboring-mode interactions. Theoretically, our analysis characterizes linearly recoverable error, decomposes the benefit of expanding structure into additional recoverable signal and fitted-operator excess, and bounds corrective-gain estimation error uniformly over norm-bounded classes. The resulting framework fits closed-form candidates and selects their resolution, ridge multiplier, and fitting source using cross-fitted validation. Across an extensive 320-case evaluation matrix, our method improves $279/320$ settings with an average relative MSE reduction of 6.34%, compared with 3.72% for the best fixed resolution. It also improves outputs already enhanced by post-hoc adapters or training losses. These results establish structural resolution as an effective decision variable, showing that adapting complexity to calibration support enables reliable correction from a single frozen checkpoint.

1 INTRODUCTION

Modern time-series forecasting predominantly scales model capacity—spanning deep linear, transformer, and frequency-native architectures (Zeng et al., 2023; Nie et al., 2023; Ye et al., 2024)—to minimize prediction errors. Yet, trained models inevitably leave structured errors in their residuals. Post-hoc correction (Liang et al., 2026; Liu et al., 2025) offers a lightweight paradigm to mitigate this deficiency: it compensates for residual errors in frozen checkpoints without altering internal representations or retraining.

However, post-hoc correction confronts a fundamental identification challenge: *residual energy alone does not reflect corrective value*. Because models optimize over finite batches, similar residual magnitudes can conceal fundamentally different mixtures of compensable error and stochastic training-run variation. We term this conflict *forecast residual ambiguity*. To decouple these entangled patterns, we perform multi-seed spectral analysis across eight benchmark datasets. In Figure 1a, high-skill modes exhibit tight error distributions across seeds, whereas weak-skill modes show substantially greater cross seed variance relative to forecast energy. The median energy-weighted normalized seed variance $v_{\mathcal{B}}$ is 29% in the weak-skill band versus 0.2% in the high-skill band (Appendix B). Thus, spectral components possess fundamentally different training stability, and residual magnitude alone cannot determine whether an error pattern reflects transferable structure.

In deployment, a single checkpoint lacks cross-seed variance and equal-energy modes can conceal fundamentally different forecast skill and training stability. Existing correctors (Liang et al., 2026; Liu et al., 2025) rely on fixed architectures, risking underfitting multivariate structures or learning spurious correlations that fail to generalize. Thus, expanding a correction structure is valuable only when its additional recoverable signal exceeds its added fitting cost. This leads to our central question: *which forecast–error dependencies are useful for correction, and at what structural resolution can finite calibration data support them?*

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

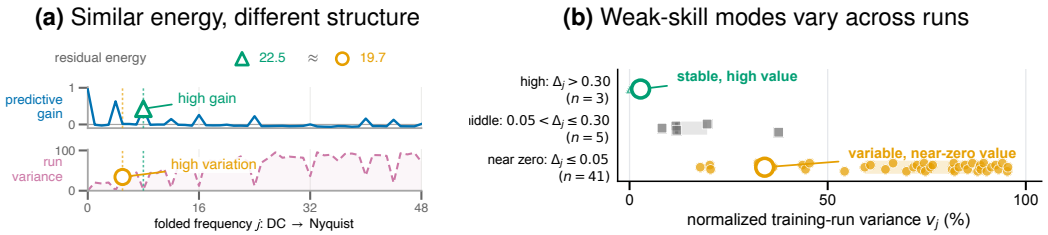


Figure 1: Forecast residual ambiguity on ETTh2/TimeXer at $H = 96$ (seeds $2/3/4$). **(a)** Modes with residual energies 22.5 and 19.7 differ in zero-baseline forecast skill and normalized seed variance. **(b)** Normalized cross-seed variance over 49 rFFT modes, grouped by forecast skill Δ_j. The weak-skill band ($\Delta_j \leq 5\%$) has greater variation relative to forecast energy.

We answer this question by treating the allowed dependency structure, or *covariance resolution*, as a selection variable. A weighted projection identity characterizes the error recoverable within each structure. Expanding a nested structure adds recoverable directions, while a uniform moment-perturbation bound describes the estimation error relevant to those directions. This yields a design principle: additional dependencies are valuable when their recoverable signal exceeds their added fitting cost. We implement this principle through closed-form candidates on the nested deployment chain, stabilized by ridge fitting and ranked by cross-fitted validation. The contributions are:

- **Problem Discovery and Diagnosis:** We formalize forecast residual ambiguity, showing that equal-energy residuals conceal different mixtures of compensable bias and stochastic training variation. Controlled multi-seed spectral analysis across 74 cells confirms that cross-seed variance concentrates predominantly in near-zero-skill modes.
- **Theoretical Foundations:** We identify linearly recoverable error and the exact risk increment between nested fitted structures. A seed-averaged decomposition explains training variation, while a checkpoint-specific uniform gain bound characterizes calibration-moment error.
- **Adaptive Framework and Validation:** We develop a closed-form pipeline over the G_0–G_∞ hierarchy that tailors capacity to each checkpoint via cross-fitted validation and dimension guards. Across 320 evaluation cells (8 datasets, 10 backbones, 4 horizons), our method improves $279/320$ cases with an average MSE reduction of 6.34% ($+7.61\%$ at H=96), while consistently delivering additional gains on top of existing adapters and loss enhancements.

2 Related Work

Backbone-independent post-hoc correction. We focus on post-hoc methods that fit a corrector from labeled calibration data and freeze both the backbone and the corrector during deployment. Existing approaches are typically categorized by their correction mechanisms: 1) *magnitude-bounded methods*, such as δ-Adapter (Liang et al., 2026), which constrain the norm of modified forecasts; 2) *selective revision methods* determine where to correct, with PIR (Liu et al., 2025) revising error-prone instances and Socket+Plug (Wang et al., 2025a) calibrating specific variables or horizons; and 3) *structured residual correctors* isolate predefined patterns, such as trend–seasonal splits in UEC-STD (Hoang Nguyen et al., 2026) or validation-guarded correctors in CRC (Xie et al., 2025). While controlling update magnitude, instances, or safety, existing correctors rely on prescribed capacities, failing to adapt their dependency structures to finite calibration support.

Estimator variation and statistical shrinkage. Statistical learning theory classically decomposes the expected error of a predictor into systematic bias and variation across fitted estimators (Geman et al., 1992). In deep forecasting, random initialization and stochastic optimization induce significant variation across independently trained checkpoints (Adlam & Pennington, 2020; Summers & Dinneen, 2021). While neural-network ensembles can expose and marginalize this variation (Krogh & Vedelsby, 1994), they require the prohibitive cost of training and deploying multiple models. Classical shrinkage provides a complementary theoretical foundation to control this variability: Wiener filtering attenuates components based on their signal-to-noise ratios (Wiener, 1949), James–Stein estimation improves noisy multivariate estimates by shrinking them toward a shared reference (James & Stein, 1961; Efron & Morris, 1973), and forecast reconciliation regularizes forecast-error covariance under coherence constraints (Wickramasuriya et al., 2019). While these classical approaches theoretically justify

the attenuation of unstable estimates, they lack a deployable mechanism to dynamically adapt the correction's structural capacity to the finite calibration data of a single deployed checkpoint.

3 FROM MULTI-SEED DIAGNOSIS TO CHECKPOINT-SPECIFIC CORRECTION

3.1 POST-HOC CORRECTION IN SPECTRAL COORDINATES

Let $S = f_\theta(X) \in \mathbb{R}^{H\times C}$ and $Y \in \mathbb{R}^{H\times C}$ denote the forecast and target of a frozen backbone. A post-hoc corrector transforms S into $\widetilde{S}$ without updating the backbone or requiring another backbone pass. For the even horizons considered here, an rFFT along the H predicted steps yields $M = H/2+1$ modes. Write $s_j, y_j, e_j \in \mathbb{C}^C$ for the forecast, target, and residual coefficients, with $e_j = s_j - y_j$. A correction subtracts c_j:

$$\widetilde{s}_j = s_j - c_j, \qquad \widetilde{e}_j = e_j - c_j. \tag{1}$$

Inverse rFFT returns the corrected time-domain forecast.

Parseval's identity makes spectral squared error additive across modes. This supports both mode-level diagnosis and structured correction in the same coordinates. For theory, under orthonormal FFT normalization, we stack $\sqrt{q_j}\,\mathrm{rv}(s_j)$ into δ and $\sqrt{q_j}\,\mathrm{rv}(e_j)$ into ϵ, where $\mathrm{rv}(z) = [\mathrm{Re}(z)^\top, \mathrm{Im}(z)^\top]^\top$, $q_j = 2$ for interior modes and $q_j = 1$ at DC and Nyquist. After removing identically zero endpoint imaginary coordinates, these vectors lie in a common real signal space. Its squared norm divided by HC gives time-domain MSE; the diagnostic ratios below are dimensionless (Appendix A.1).

3.2 TWO DIAGNOSTIC AXES: FORECAST SKILL AND TRAINING-RUN INSTABILITY

To understand which residual components are genuinely compensable, we first diagnose the backbone's spectral behavior across independent training seeds along two continuous metrics: predictive value and estimator instability.

Normalized forecast skills. Let w index evaluation windows and $k \in \{1, \dots, N_s\}$ index independent training runs. The normalized forecast value

$$\Delta_j = 1 - \frac{\mathbb{E}_{w,k}\|F_j(S_w^{(k)}) - F_j(Y_w)\|^2}{\mathbb{E}_w\|F_j(Y_w)\|^2} \tag{2}$$

measures error reduction relative to a zero-prediction baseline. A higher Δ_j (≤ 1) reflects dependable predictive signal, whereas $\Delta_j < 0$ indicates performance worse than predicting zero.

Normalized cross-seed variance. With $\widehat{\mu}_{wj} = \frac{1}{N_s}\sum_{k=1}^{N_s} F_j(S_w^{(k)})$ denoting the empirical seed-averaged forecast, the normalized cross-seed variance v_j quantifies checkpoint volatility relative to forecast energy:

$$\sigma_j^2 = \mathbb{E}_w\left[\frac{1}{N_s - 1}\sum_{k=1}^{N_s}\|F_j(S_w^{(k)}) - \widehat{\mu}_{wj}\|^2\right], \qquad v_j = \frac{\sigma_j^2}{\mathbb{E}_{w,k}\|F_j(S_w^{(k)})\|^2}. \tag{3}$$

A large v_j flags a mode that is highly sensitive to stochastic optimization noise. These metrics uncover the fundamental tension of post-hoc correction: *preserving high-Δ_j predictive signal while attenuating high-v_j noise*. Appendix B specifies numerical denominator conventions and reports these diagnostics over 74 cells. For the 65 cells containing both weak- and high-skill bands, the median band-level normalized variance $v_{\mathcal{B}}$ is 29% versus 0.2%. The bands summarize diagnosis. Deployment instead uses checkpoint-specific forecast–error moments (Appendix B.2).

3.3 SINGLE-CHECKPOINT RISK AND RECOVERABLE ERROR

While multi-seed diagnostics reveal latent training instability across runs, real-world deployment provides only a single frozen checkpoint and finite labeled calibration data. This available data allows us to formulate and optimize the correction risk directly. In the unified Parseval-weighted spectral space, a linear corrector produces the corrected residual $\epsilon - K\delta$. Here, K represents a linear operator acting on the common real signal space, while mode-wise and neighboring-mode corrections form

structured block restrictions of this operator (Appendix C.4). Under this formulation, diagonal entries of K govern coordinate-wise shrinkage, while off-diagonals capture cross-channel and cross-phase dependencies. Setting $K = 0$ leaves the uncorrected forecast untouched ($\widetilde{s} = s$), whereas $K = I$ completely cancels it ($\widetilde{s} = 0$, reverting to the zero baseline).

Fixing checkpoint k and a reference window distribution P, we define the quadratic correction risk:

$$\mathcal{R}(K) = \mathcal{R}_P^{(k)}(K) = \mathbb{E}_{w\sim P}\|\epsilon - K\delta\|^2. \tag{4}$$

Assuming finite joint second moments, let $\Sigma = \mathbb{E}_w[\delta\delta^\top]$ denote the uncentered input second moment and $\Gamma = \mathbb{E}_w[\epsilon\delta^\top]$ the error–forecast cross-moment. The risk expands analytically as:

$$\mathcal{R}(K) = \mathcal{R}(0) - 2\,\mathrm{tr}(K\Gamma^\top) + \mathrm{tr}(K\Sigma K^\top). \tag{5}$$

Physically, Σ registers which forecast directions carry energy and correlation, while Γ identifies which directions systematically correlate with residual errors. Setting the gradient with respect to K to zero yields the normal equation $K\Sigma = \Gamma$. Its minimum-Frobenius-norm minimizer is given by:

$$K^\star = \Gamma\Sigma^+, \tag{6}$$

where Σ^+ denotes the Moore–Penrose pseudoinverse (Appendix A). Substituting $K^\star$ back into Eq. (5) yields the maximum *linearly recoverable error*:

$$\mathcal{C}_{\mathrm{lin}} := \mathcal{R}(0) - \mathcal{R}(K^\star) = \mathrm{tr}(\Gamma\Sigma^+\Gamma^\top). \tag{7}$$

Here, it is the cross-moment Γ, rather than the residual energy $\mathcal{R}(0)$ alone, that determines this corrective value. Appendix A.4 derives this identity and provides two concrete distributions with identical forecast and residual marginals but drastically different recoverable errors.

3.4 Training Variation and Gain Estimation

Moving from multi-seed diagnosis to deployment introduces a natural distinction: training randomness explains model instability across runs, whereas calibration sample size limits estimation accuracy on a single checkpoint. Our risk analysis formalizes both aspects.

Let k vary over training randomness independently of w. Write y_w for the target in the same Parseval-weighted real coordinates, fixed across runs. Define $\mu_w = \mathbb{E}_k\delta_w^{(k)}$, $\xi_w^{(k)} = \delta_w^{(k)} - \mu_w$, and $\bar{\epsilon}_w = \mu_w - y_w$. Then $\epsilon_w^{(k)} = \bar{\epsilon}_w + \xi_w^{(k)}$. Write $\Sigma_\mu = \mathbb{E}_w[\mu_w\mu_w^\top]$, $\Gamma_\mu = \mathbb{E}_w[\bar{\epsilon}_w\mu_w^\top]$, and $\Sigma_\xi = \mathbb{E}_{w,k}[\xi_w^{(k)}\xi_w^{(k)\top}]$.

Theorem 1 (Seed-averaged correction-risk decomposition)**.** *Assume finite joint second moments. For every deterministic linear operator K shared across windows and training runs,*

$$\overline{\mathcal{R}}(K) := \mathbb{E}_k\mathcal{R}^{(k)}(K) = \mathbb{E}_w\|\bar{\epsilon}_w - K\mu_w\|^2 + \mathrm{tr}\big((I-K)\Sigma_\xi(I-K)^\top\big). \tag{8}$$

Moreover, $\mathbb{E}_k[\Sigma^{(k)}] = \Sigma_\mu + \Sigma_\xi$ *and* $\mathbb{E}_k[\Gamma^{(k)}] = \Gamma_\mu + \Sigma_\xi$.

The two terms are remaining seed-mean error and retained training variation. Their values before correction do not determine how they change with K. The corresponding shared optimum and its distinction from checkpoint-specific optima are developed in Appendix A.6.

For deployment, define *corrective gain* relative to the same checkpoint's uncorrected forecast:

$$g(K) = \mathcal{R}(0) - \mathcal{R}(K) = 2\,\mathrm{tr}(K\Gamma^\top) - \mathrm{tr}(K\Sigma K^\top). \tag{9}$$

Replacing the moments with averages over labeled calibration windows gives $\widehat{g}(K)$. For a structural linear space $\mathcal{K}_g$, let Π_g^F be Frobenius projection onto that space and define

$$a_g = \|\Pi_g^F(\widehat{\Gamma} - \Gamma)\|_F, \qquad b_g = \sup_{\substack{A\in\mathcal{K}_g \\ \|A\|_F\le 1}} \left|\mathrm{tr}\big(A(\widehat{\Sigma} - \Sigma)A^\top\big)\right|. \tag{10}$$

Theorem 2 (Reliability of checkpoint-specific corrective gain)**.** *For any calibration sample with finite empirical moments and fixed $0 \le B_K < \infty$, let $\mathcal{C}_g(B_K) = \{K \in \mathcal{K}_g : \|K\|_F \le B_K\}$. Then*

$$\sup_{K\in\mathcal{C}_g(B_K)} |\widehat{g}(K) - g(K)| \le \eta_g(B_K) := 2B_K a_g + B_K^2 b_g. \tag{11}$$

For any fixed $B_K < \infty$, the bound holds uniformly over $\mathcal{C}_g(B_K)$ and therefore applies to data-dependent choices that remain within this prescribed norm-bounded class. If $a_g, b_g \to 0$ in probability, the gain estimates converge uniformly in probability on $\mathcal{C}_g(B_K)$.

Expanding nested spaces enlarges the directions available for correction and cannot decrease a_g, b_g, or their worst-case bound at a common B_K. Resolution controls both representational capacity and the moment errors relevant to estimating gain. Appendix A.2 proves the bound, and Appendix A.6 gives consistency conditions for dependent windows. The bound characterizes estimation error; the deployment rule uses validation scores to choose a fitted operator.

4 Checkpoint-Specific Estimation and Covariance Resolution

To deploy post-hoc correction without underfitting complex interactions or overfitting limited calibration data, we dynamically adapt the operator's structural capacity, or *covariance resolution*, balancing representational power against estimation variance.

4.1 Spectral Correction as Structured Linear Regression

We constrain the correction operator block K_j to a structured linear subspace $\mathcal{K}$ (acting on $\mathbb{R}^{2C}$ for single-mode maps, or $\mathbb{R}^{6C}$ for neighboring-mode $G_5(3)$) defined by specific dependency masks and parameter ties. Given n_f calibration windows, empirical moments $\widehat{\Sigma}_j = \frac{1}{n_f}\sum_{w=1}^{n_f} \delta_{wj}\delta_{wj}^\top$ and $\widehat{\Gamma}_j = \frac{1}{n_f}\sum_{w=1}^{n_f} \epsilon_{wj}\delta_{wj}^\top$ yield the empirical risk $\widehat{\mathcal{R}}(K_j) = \widehat{\mathcal{R}}(0) - 2\,\mathrm{tr}(K_j\widehat{\Gamma}_j^\top) + \mathrm{tr}(K_j\widehat{\Sigma}_j K_j^\top)$, whose unconstrained minimizer is $\widehat{K}_j^\star = \widehat{\Gamma}_j\widehat{\Sigma}_j^+$. When restricting $K_j \in \mathcal{K}$, the following projection identity characterizes the structural risk penalty.

Theorem 3 (Weighted projection onto operator subspaces)**.** *For a linear operator subspace $\mathcal{K}$, define the Σ-seminorm $\|A\|_\Sigma^2 = \mathrm{tr}(A\Sigma A^\top)$. Let $\Pi_{\mathcal{K}}^\Sigma K^\star$ denote the Σ-projection of population optimum $K^\star$ onto $\mathcal{K}$. Every $K \in \mathcal{K}$ satisfies:*

$$\mathcal{R}(K) = \mathcal{R}(\Pi_{\mathcal{K}}^\Sigma K^\star) + \|K - \Pi_{\mathcal{K}}^\Sigma K^\star\|_\Sigma^2. \tag{12}$$

Thus, $\Pi_{\mathcal{K}}^\Sigma K^\star$ achieves optimal population risk within $\mathcal{K}$, while differences in the null space of Σ leave risk invariant (proof in Appendix A.3).

4.2 When Additional Dependencies Are Worth Fitting

For class $\mathcal{K}_g$, write $K_g^\star$ for its population optimum and $\widehat{K}_g \in \mathcal{K}_g$ for a fitted operator. Theorem 3 yields

$$\mathcal{R}(\widehat{K}_g) - \mathcal{R}(K^\star) = \underbrace{\mathcal{R}(K_g^\star) - \mathcal{R}(K^\star)}_{\text{Approximation error}} + \underbrace{\|\widehat{K}_g - K_g^\star\|_\Sigma^2}_{\text{Fitted-operator excess } E_g}. \tag{13}$$

The fitted-operator excess includes moment estimation, ridge shrinkage, and gating effects. Comparing nested structures makes their decision consequence explicit.

Corollary 4 (Incremental value of a wider structure)**.** *For $\mathcal{K}_g \subseteq \mathcal{K}_h$ under the same reference risk, and fitted operators in their respective spaces,*

$$\mathcal{R}(\widehat{K}_g) - \mathcal{R}(\widehat{K}_h) = \underbrace{\|K_h^\star - K_g^\star\|_\Sigma^2}_{\textit{Additional recoverable error}} - (E_h - E_g). \tag{14}$$

A wider structure lowers risk exactly when its additional recoverable error exceeds its added fitted-operator excess. The latter difference can have either sign. This identity and the restricted moment errors in Theorem 2 motivate selecting resolution from calibration evidence rather than parameter count alone (Appendix A.5). We select from the nested deployment chain $G_2 \subseteq G_3 \subseteq G_4 \subseteq G_5(3)$:

- G_2: one real scalar per mode–channel pair, shared by real and imaginary parts.
- G_3: a $C \times C$ channel map per mode, shared by both parts.
- G_4: a full $2C \times 2C$ map per mode, allowing channel and real–imaginary coupling.
- $G_5(3)$: an output-mode map using three adjacent input modes, with nominal block size $2C \times 6C$.

These classes belong to a larger, partially ordered family ending at the full-spectrum reference G_∞ (Appendices C.4, E.5). Boundary coordinates reduce the effective dimensions.

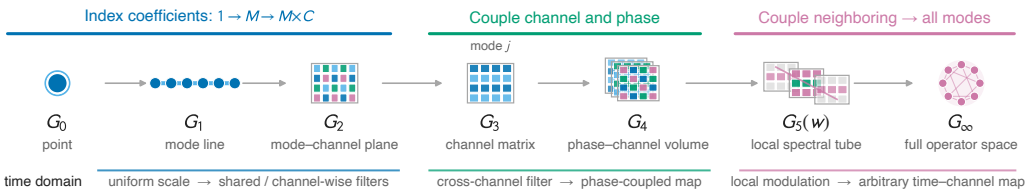


Figure 2: Covariance-resolution hierarchy. Each step from G_0 to G_∞ resolves more temporal-scale, channel, phase, or neighboring-mode dependence at greater finite-sample estimation cost.

For G_2, a scalar κ_{jc} specifies the fraction subtracted from the forecast. With numerical stabilizer $\zeta = 10^{-12}$, the fitted coefficient is

$$\hat{b}_{jc} = \frac{1}{n_f}\sum_w \mathrm{Re}(e_{wjc}\bar{s}_{wjc}), \quad \hat{a}_{jc} = \frac{1}{n_f}\sum_w |s_{wjc}|^2, \quad \widehat{\kappa}_{jc} = \Pi_{[0,1]}\left(\frac{\hat{b}_{jc}}{\hat{a}_{jc} + \zeta/n_f}\right). \tag{15}$$

This minimizes empirical residual error plus $(\zeta/n_f)\kappa_{jc}^2$ on $[0,1]$ and gives zero when the fitting input is zero. The retained fraction lies in $[0,1]$. Matrix classes use the data-dependent trace-scaled ridge penalties specified in Algorithm 1. The perturbation analysis in Eq. (20) accounts for the empirical moment errors under a fixed regularizer τ.

4.3 Cross-Fitted Selection for the Deployed Operator

Because appropriate capacity varies by dataset and horizon, we select the correction operator per checkpoint via a three-stage validation protocol:

Cross-Fitted Evaluation. We construct stride-one validation windows and split them chronologically at the index midpoint into halves A and B. To prevent temporal leakage, all forecast horizons in both folds terminate before the first test target. Validation-fitted candidates alternate fitting and scoring between the two halves ($A \leftrightarrow B$), while training-fitted candidates fit on the training split and evaluate across both halves (Appendix C.6, Table 10).

Dimension Guards and Regularized Gating. As an empirical capacity screen against severe overparameterization, a nominal dimension guard excludes candidate classes lacking positive residual degrees of freedom ($2n_f \le C$ for G_3, $n_f \le 2C$ for G_4, and $n_f \le 6C$ for $G_5(3)$). For each admissible candidate, we select a shared ridge penalty from 64 values in $[10^{-4}, 10^5]$ using the ungated validation score (Eq. 23), scaled per mode by the trace of its design matrix. We then apply selective block gates (mode–channel coordinates for G_2, whole modes for matrix classes), preserving the correction only where it outperforms the uncorrected forecast on validation windows.

Model Selection and Refitting. Candidates are ranked by their gated validation scores (Eq. 23), favoring more parsimonious classes in the event of ties. While Eq. 23 uses the unweighted spectral score, interior modes account for $> 95\%$ of all frequencies ($H \ge 96$), rendering boundary Parseval discrepancies practically negligible in candidate selection (Appendix C.6). The selected configuration is refitted on its entire source (all validation windows, or the complete training split) with its selected gates frozen, producing $\widehat{K}_{\mathrm{deploy}}$ for test deployment (Algorithm 1). This mechanism effectively bridges the statistical volume of training data with the temporal proximity of validation residuals.

5 Experiments

5.1 Experimental Setup

Datasets and Backbones: The complete evaluation matrix covers 320 dataset–backbone–horizon cells: eight datasets, ten backbones, and four horizons. A "cell" is defined as one dataset–backbone–horizon combination. The eight datasets encompass various domains: ETTh1, ETTh2, ETTm1, ETTm2 (electricity-transformer operation); Weather (meteorology); ECL (electricity consumption); Solar (photovoltaic output); and Traffic (freeway occupancy). We evaluate our method on ten

forecasting backbones covering four broad modeling paradigms: 1) lightweight linear forecasters: DLinear (Zeng et al., 2023) and MixLinear (Ma et al., 2026); 2) frequency-domain models: FITS (Xu et al., 2024) and PaiFilter, the first variant introduced in FilterNet (Yi et al., 2024); 3) Transformer-based models: PatchTST (Nie et al., 2023), iTransformer (Liu et al., 2024), and TimeXer (Wang et al., 2024b); and 4) multivariate or multiscale models: DUET (Qiu et al., 2025), FACT (Wang et al., 2026), and TimeMixer (Wang et al., 2024a). Appendix C.3 gives concise descriptions of all ten backbones.

Implementation Details: Backbones are trained with an input length of 96 and forecast horizons $H \in \{96, 192, 336, 720\}$, optimized via Adam with early stopping (patience 3) on validation MSE. The spectral transformation uses the real-valued fast Fourier transform (rFFT), producing $M = H/2 + 1$ nonredundant modes. Full dataset, split, and implementation details are provided in Appendix C.

Evaluation Protocol: The primary metric is the paired relative test-MSE reduction, calculated as $100(\text{MSE}_{\text{base}} - \text{MSE}_{\text{corr}})/\text{MSE}_{\text{base}}$, where positive values indicate lower error (improvement). To account for training-run variation, the correction is applied independently to checkpoints from seeds 2, 3, and 4, yielding 960 paired seed-level evaluations; each cell reports the mean and sample standard deviation (SD) of its three paired outcomes. Exact two-sided Wilcoxon signed-rank tests and dataset-resampled bootstrap confidence intervals assess statistical significance (Appendix C.7).

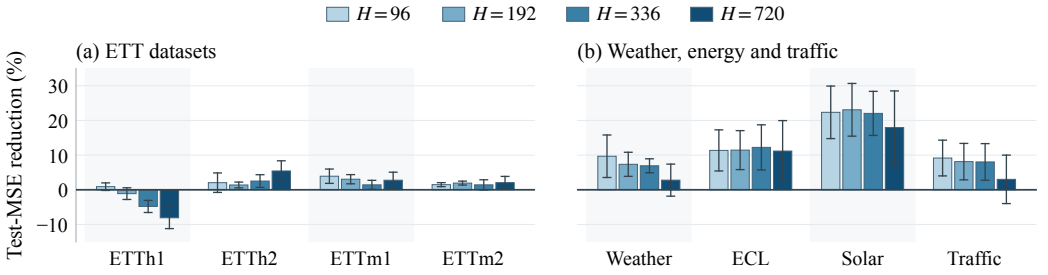


Figure 3: Averaged Test-MSE reduction (%, ↑) by dataset and horizon across ten backbones.

5.2 Main Results

Overall Performance. As illustrated in Fig. 3, our framework consistently reduces test MSE across the 320-cell benchmark, improving 279/320 cells with an average reduction of +6.34%. As detailed in Table 1 and Table 2, this performance gain is robust across both temporal lookaheads and modeling paradigms: 1) Across horizons, aggregate gains remain consistently positive from short to extended lookaheads (+7.61%, +6.91%, +6.22%, and +4.64%); 2) Across backbones, every single architecture achieves positive aggregate improvement over its 32 cells. This confirms that systematic post-hoc error is pervasive across diverse model inductive biases (linear, spectral, transformer, and multiscale) rather than an artifact of specific architectures.

Table 1: Dataset–horizon paired test-MSE reduction (%, ↑), averaged over seeds 2/3/4 and reported as mean ± SD across the same ten backbone-cell means at every horizon.

H	ETTh1	ETTh2	ETTm1	ETTm2	Weather	ECL	Solar	Traffic	all 80
96	$+0.90_{\pm 1.09}{}^{\dagger}$	$+2.05_{\pm 2.81}{}^{\dagger}$	$+3.92_{\pm 2.07}{}^{\dagger}$	$+1.48_{\pm 0.57}{}^{\dagger}$	$+9.68_{\pm 6.13}{}^{\dagger}$	$+11.36_{\pm 5.93}{}^{\dagger}$	$+22.34_{\pm 7.57}{}^{\dagger}$	$+9.16_{\pm 5.15}{}^{\dagger}$	$+7.61_{\pm 8.10}{}^{\dagger}$
192	$-1.11_{\pm 1.69}$	$+1.39_{\pm 0.84}{}^{\dagger}$	$+3.05_{\pm 1.32}{}^{\dagger}$	$+1.93_{\pm 0.56}{}^{\dagger}$	$+7.35_{\pm 3.47}{}^{\dagger}$	$+11.45_{\pm 5.62}{}^{\dagger}$	$+23.07_{\pm 7.60}{}^{\dagger}$	$+8.12_{\pm 5.25}{}^{\dagger}$	$+6.91_{\pm 8.26}{}^{\dagger}$
336	$-4.80_{\pm 1.75}{}^{\dagger}$	$+2.52_{\pm 1.83}{}^{\dagger}$	$+1.43_{\pm 1.30}{}^{\dagger}$	$+1.42_{\pm 1.46}{}^{\dagger}$	$+6.92_{\pm 2.02}{}^{\dagger}$	$+12.23_{\pm 6.51}{}^{\dagger}$	$+22.04_{\pm 6.35}{}^{\dagger}$	$+8.03_{\pm 5.28}{}^{\dagger}$	$+6.22_{\pm 8.60}{}^{\dagger}$
720	$-8.08_{\pm 3.12}{}^{\dagger}$	$+5.43_{\pm 2.93}{}^{\dagger}$	$+2.77_{\pm 2.30}{}^{\dagger}$	$+2.06_{\pm 1.79}{}^{\dagger}$	$+2.78_{\pm 4.62}{}^{\dagger}$	$+11.18_{\pm 8.78}{}^{\dagger}$	$+17.94_{\pm 10.57}{}^{\dagger}$	$+3.01_{\pm 6.99}$	$+4.64_{\pm 9.11}{}^{\dagger}$

SD is computed across the backbone-cell means named by each column: ten for a dataset and 80 for the aggregate. † * marks an exact two-sided Wilcoxon signed-rank test with $p < 0.05$ across backbone means. Positive values indicate lower test MSE. To account for cross-backbone dependence on shared datasets, cluster-robust bootstrap confidence intervals are reported in Table 11 (overall 95% CI: [1.98%, 11.59%]).

Long Horizons and Boundaries. Across 240 long-horizon settings ($H \geq 192$), our approach improves 202/240cells by +5.92% on average, with the largest gains on ECL and Solar (Appendix Tables 25–31). At H=720, however, distribution drift weakens calibration transfer on non-stationary regimes like ETTh1, delineating an empirical boundary (Table 1). Finally, optimizing quadratic risk on skewed series (e.g., Weather) substantially reduces MSE (+9.68% at H=96) while naturally trading off absolute deviation metrics.

Peer Composition. Table 3 benchmarks our method against post-hoc output adapters (δ-Adapter, UEC-STD, Socket+Plug) and spectral training objectives (FreDF (Wang et al., 2025b), PS Loss (Ku-

Table 2: Test-MSE reduction by backbone and horizon. Overall averages 32 cells per backbone.

backbone	overall	improves	H=96	H=192	H=336	H=720
DLinear	+7.24±8.54	26/32	+10.18±8.57	+7.37±8.70	+6.61±8.54	+4.82±9.17
DUET	+7.28±9.56	29/32	+8.55±9.98	+7.85±9.56	+7.11±10.35	+5.59±10.06
FACT	+2.35±3.88	27/32	+3.10±2.61	+3.01±2.59	+2.86±4.85	+0.44±4.88
FITS	+8.77±10.94	29/32	+10.74±10.68	+9.33±10.90	+8.34±11.37	+6.69±12.57
MixLinear	+10.50±11.78	29/32	+11.90±11.37	+11.03±12.15	+9.56±12.29	+9.52±13.49
PaiFilter	+7.00±9.42	27/32	+8.07±8.99	+7.50±9.75	+6.92±10.10	+5.50±10.55
PatchTST	+5.89±7.73	28/32	+7.14±7.47	+6.34±7.58	+5.47±8.47	+4.59±8.67
TimeMixer	+5.57±6.93	29/32	+6.84±6.60	+6.60±7.37	+5.57±7.43	+3.28±7.08
TimeXer	+3.82±5.92	26/32	+3.95±5.34	+4.28±5.69	+4.50±6.42	+2.55±7.09
iTransformer	+5.02±5.68	29/32	+5.64±4.87	+5.75±5.94	+5.30±6.47	+3.39±6.15
All backbones	+6.34±8.56	279/320	+7.61±8.10	+6.91±8.26	+6.22±8.60	+4.64±9.11

Improves counts cells with lower test MSE. Horizon-specific counts appear in Table 27.

drat et al., 2025)). Our framework delivers consistent secondary error reductions when applied to peer outputs across most cells. This composition evidence shows that selecting output dependencies can complement the correction mechanisms and training objectives already used by a forecaster.

Table 3: Standalone performance and the additional benefit of ours at H=96 (%, ↑). Entries are mean ± SD across 80 dataset–backbone cell means after averaging paired seeds 2/3/4.

		Output correctors			Training objectives	
comparison	Plain forecast	δ-Adapter	UEC-STD	Socket+Plug	FreDF	PS Loss
Alone vs. plain ↑	0.00±0.00	+3.17±3.97	+2.16±3.83	+6.41±7.10	−11.72±33.79	−0.80±5.07
+ours vs. plain ↑	**+7.61**±8.10	**+8.13**±9.55	**+6.57**±6.67	**+10.34**±10.77	**+6.25**±7.91	**+7.74**±7.36
Extra vs. input ↑	**+7.61**±8.10	**+5.35**±6.64	**+4.55**±4.87	**+4.43**±6.74	**+12.14**±15.37	**+8.11**±9.53
Improved cells	77/80	72/80	76/80	74/80	76/80	77/80

Extra reduction averages paired percentages relative to the input MSE; improves counts positive cell means. The plain column is ours alone. Absolute errors and wins vs. plain training are in Appendix D.4.

5.3 Analysis and Ablation

Resolution Dynamics: Dual Mechanism. Figure 4 illustrates how selected resolutions shift across benchmarks, uncovering a dual mechanism of post-hoc correction:

- **Variance Damping**: In low-dimensional benchmarks with heavy training-run volatility at weak modes (e.g., ETT), the selector predominantly contracts to coordinate-wise G_2. In these noise-dominated settings, attempting to fit cross-variable interactions is counterproductive; reducing error relies on safely damping volatile modes without inflating estimation variance.
- **Spatial Reconstruction & Cross-Mode Coupling**: In contrast, benchmarks with rich multi-variable or temporal correlations benefit from expanded structures. Notably, while Weather exhibits many weak-skill modes, it selects $G_5(3)$ in 88 of 120 runs, demonstrating that wide structures can capture useful dependencies even in noisy regimes. On Solar, the adaptive rule achieves a 21.35% MSE reduction (compared with only 2.66% for fixed G_2), and removing $G_5(3)$ drops the full-matrix mean from 6.34% to 5.51%.

The framework thus dynamically navigates this transition, allocating capacity where multivariate dependencies dominate while enforcing compact shrinkage where noise prevails.

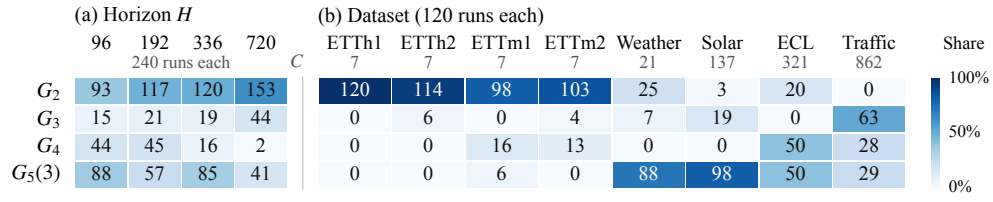


Figure 4: Covariance resolutions selected across 960 runs, grouped by **(a)** horizon (240 runs each) and **(b)** dataset (120 runs each, headers show channel count C).

Choosing a Structure rather than Fixing It. As shown in Table 4, our adaptive rule (6.34%) outperforms every fixed resolution across all horizons, outperform the best static class ($G_5(3)$ at

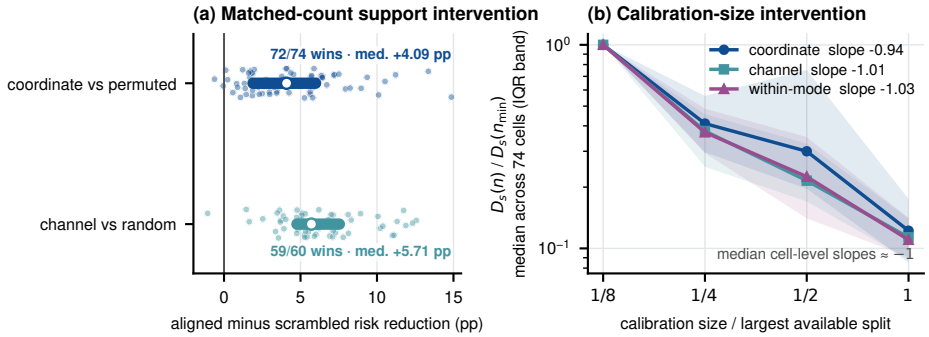


Figure 5: Structural alignment and sample scaling. **(a)** In-sample error-reduction margin of aligned over count-matched scrambled patterns (positive favors alignment). **(b)** Normalized held-out disagreement $D_s(n)$ between operators fitted on disjoint calibration splits.

3.72%) by 1.70x. At H=720, fixed wide classes degrade severely, whereas adaptive selection and its dimension guard preserve a robust $+4.64\%$ improvement (Appendix Table 24).

Table 4: Covariance resolution across four horizons. Fixed classes tune source and penalty; ours also selects the resolution. Entries show mean ± SD reduction (%, ↑) and win counts.

H	coordinate G_2 MC		channel G_3 $\sim MC^2$		within-mode G_4 $\sim 4MC^2$		local-mode $G_5(3)$ $\sim 12MC^2$		**selected (ours)**		test-ranked upper reference	
	mean	wins	mean	wins	mean	wins	mean	wins	mean	wins	mean	wins
96	$+1.83_{\pm 2.49}$	69	$+5.37_{\pm 5.80}$	74	$+6.02_{\pm 6.51}$	75	$+6.99_{\pm 8.47}$	73	$+7.61_{\pm 8.10}$	77	$+7.96_{\pm 7.89}$	80
192	$+2.07_{\pm 3.06}$	64	$+4.53_{\pm 7.87}$	69	$+5.07_{\pm 8.03}$	68	$+5.61_{\pm 9.69}$	69	$+6.91_{\pm 8.26}$	73	$+7.68_{\pm 7.78}$	80
336	$+1.66_{\pm 3.98}$	58	$+2.92_{\pm 12.17}$	67	$+3.29_{\pm 12.63}$	69	$+3.50_{\pm 13.92}$	69	$+6.22_{\pm 8.60}$	68	$+7.62_{\pm 7.52}$	78
720	$+0.12_{\pm 4.68}$	49	$-4.38_{\pm 30.28}$	58	$-2.67_{\pm 27.05}$	61	$-1.22_{\pm 21.71}$	59	$+4.64_{\pm 9.11}$	61	$+7.93_{\pm 7.43}$	73
all	$+1.42_{\pm 3.71}$	240	$+2.11_{\pm 17.39}$	268	$+2.93_{\pm 16.08}$	273	$+3.72_{\pm 14.68}$	270	$+6.34_{\pm 8.56}$	279	$+7.80_{\pm 7.62}$	311

Parameter counts are per $M = H/2 + 1$ modes. The final column group ranks already fitted candidates by test MSE only to measure selection headroom; it is not deployed and does not enter any headline result. Fixed classes retain their named class even when its nominal width exceeds the fitting-row count, thereby isolating the consequence of committing to that resolution.

Structural Alignment vs. Estimation. Matched-support controls hold coefficient count fixed. Aligned dependencies achieve in-sample error reductiosn than scrambled permutations on $72/74$ cells and random channel supports on $59/60$, with median margins of 4.09 and 5.71 pp (Fig. 5a). Separately, operator disagreement across disjoint calibration splits falls with sample size at an empirical log–log slope near -1.0 (Fig. 5b). Thus corrective value follows structurally aligned dependencies, while calibration volume determines how wide a structure can be estimated reliably.

Complementary Fitting Sources. Training offers abundant calibration windows, whereas validation offers temporal proximity to deployment. Joint selection achieves 6.34%, versus 3.85% training-only and 3.32% validation-only; at H=720, it retains a $+4.64\%$ gain versus -0.93% for training-only (Appendix Tables 34–36). Evaluating both sources under one cross-validation criterion therefore balances sample volume against temporal distribution shift.

6 Conclusion

This paper identifies *forecast residual ambiguity*—the conflation of compensable systematic bias and stochastic training-run variation—as a core challenge in post-hoc time-series correction. To address this, we formalize post-hoc correction as a structural approximation–estimation tradeoff over a nested covariance-resolution hierarchy ($G_0 \subset \cdots \subset G_\infty$). Theoretically, we decouple population diagnosis from single-checkpoint deployment, combining a seed-averaged risk decomposition with uniform gain perturbation bounds. By solving closed-form operators and selecting structural complexity via cross-fitted validation, our framework dynamically matches capacity to calibration support. Extensive evaluations across 320 cells show that this adaptive approach consistently improves diverse backbones and horizons while unlocking complementary gains on top of existing adapters and loss enhancements. Future work will extend resolution selection to nonlinear correctors and explore theoretical guarantees under distribution shifts.

LLM USAGE DISCLOSURE

Generative AI tools were used only to polish the manuscript's language and improve its readability. They were not used to generate data, develop theoretical models, formulate or prove mathematical claims, design or interpret experiments, or implement methods. The authors reviewed all AI-assisted edits and take responsibility for the final content.

REFERENCES

Ben Adlam and Jeffrey Pennington. Understanding double descent requires a fine-grained bias-variance decomposition. *Advances in neural information processing systems*, 33:11022–11032, 2020.

Bradley Efron and Carl Morris. Stein's estimation rule and its competitors—an empirical bayes approach. *Journal of the American Statistical Association*, 68(341):117–130, 1973.

Stuart Geman, Elie Bienenstock, and René Doursat. Neural networks and the bias/variance dilemma. *Neural Computation*, 4(1):1–58, 1992. doi: 10.1162/neco.1992.4.1.1.

Minh Hoang Nguyen, Dai Do, Huu Hiep Nguyen, Dung Nguyen, Kien Do, and Hung Le. Reviving error correction in modern deep time-series forecasting. In *International Conference on Machine Learning*, PMLR, 2026.

William James and Charles Stein. Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pp. 361–379, 1961.

Anders Krogh and Jesper Vedelsby. Neural network ensembles, cross validation, and active learning. *Advances in neural information processing systems*, 7, 1994.

Dilfira Kudrat, Zongxia Meng, Yifan Yang, Xinyang Wu, Ke Zhang, and Shuo Wang. Patch-wise structural loss for time series forecasting. In *International Conference on Machine Learning*, volume 267 of *PMLR*, pp. 31841–31859, 2025.

Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long- and short-term temporal patterns with deep neural networks. In *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 95–104, 2018. doi: 10.1145/3209978.3210006.

Daojun Liang, Qi Li, Yinglong Wang, Jing Chen, Hu Zhang, Xiaoxiao Cui, Qizheng Wang, and Shuo Li. The forecast after the forecast: A post-processing shift in time series. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=syfWdclGE1.

Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. iTransformer: Inverted transformers are effective for time series forecasting. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=JePfAI8fah.

Zhiding Liu, Mingyue Cheng, Guanhao Zhao, Jiqian Yang, Qi Liu, and Enhong Chen. Improving time series forecasting via instance-aware post-hoc revision. In *Advances in Neural Information Processing Systems*, volume 38, pp. 35604–35629, 2025. doi: 10.52202/085713-1196.

Aitian Ma, Dongsheng Luo, and Mo Sha. MixLinear: Extreme low resource multivariate time series forecasting with 0.1k parameters. In *International Conference on Learning Representations*, 2026.

Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=Jbdc0vTOcol.

Xiangfei Qiu, Xingjian Wu, Yan Lin, Chenjuan Guo, Jilin Hu, and Bin Yang. DUET: Dual clustering enhanced multivariate time series forecasting. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1185–1196, 2025. doi: 10.1145/3690624.3709325.

Cecilia Summers and Michael J Dinneen. Nondeterminism and instability in neural network optimization. In *International Conference on Machine Learning*, pp. 9913–9922. PMLR, 2021.

Artur Trindade. ElectricityLoadDiagrams20112014. UCI Machine Learning Repository, 2015. URL https://doi.org/10.24432/C58C86.

Bin Wang, Yongqi Han, Minbo Ma, Tianrui Li, Junbo Zhang, Feng Hong, and Yanwei Yu. Non-collective calibrating strategy for time series forecasting. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, IJCAI '25, 2025a. ISBN 978-1-956792-06-5. doi: 10.24963/ijcai.2025/371. URL https://doi.org/10.24963/ijcai.2025/371.

Hao Wang, Licheng Pan, Zhichao Chen, Degui Yang, Sen Zhang, Yifei Yang, Xinggao Liu, Haoxuan Li, and Dacheng Tao. FreDF: Learning to forecast in the frequency domain. In *International Conference on Learning Representations*, 2025b.

Huiqiang Wang, Jieming Shi, and Qing Li. FACT: Fine-grained across-variable convolution for multivariate time-series forecasting. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=j3gNYqrHtl.

Shiyu Wang, Haixu Wu, Xiaoming Shi, Tengge Hu, Huakun Luo, Lintao Ma, James Y. Zhang, and Jun Zhou. TimeMixer: Decomposable multiscale mixing for time series forecasting. In *International Conference on Learning Representations*, 2024a. URL https://openreview.net/forum?id=7oLshfEIC2.

Yuxuan Wang, Haixu Wu, Jiaxiang Dong, Guo Qin, Haoran Zhang, Yong Liu, Yunzhong Qiu, Jianmin Wang, and Mingsheng Long. TimeXer: Empowering transformers for time series forecasting with exogenous variables. In *Advances in Neural Information Processing Systems*, volume 37, pp. 469–498, 2024b. doi: 10.52202/079017-0015.

Shanika L. Wickramasuriya, George Athanasopoulos, and Rob J. Hyndman. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 114(526):804–819, 2019. doi: 10.1080/01621459.2018.1448825.

Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. MIT Press, 1949.

Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. In *Advances in Neural Information Processing Systems*, volume 34, pp. 22419–22430, 2021.

Jianxiang Xie, Yuncheng Hua, Mingyue Cheng, Flora Salim, and Hao Xue. Causality-inspired safe residual correction for multivariate time series. *arXiv preprint arXiv:2512.22428*, 2025. doi: 10.48550/arXiv.2512.22428. URL https://arxiv.org/abs/2512.22428.

Zhijian Xu, Ailing Zeng, and Qiang Xu. FITS: Modeling time series with $10k$ parameters. In *International Conference on Learning Representations*, 2024.

Weiwei Ye, Songgaojun Deng, Qiaosha Zou, and Ning Gui. Frequency adaptive normalization for non-stationary time series forecasting. In *Advances in Neural Information Processing Systems*, volume 37, pp. 31350–31379, 2024. doi: 10.52202/079017-0985.

Kun Yi, Jingru Fei, Qi Zhang, Hui He, Shufeng Hao, Defu Lian, and Wei Fan. FilterNet: Harnessing frequency filters for time series forecasting. In *Advances in Neural Information Processing Systems*, 2024.

Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? In *AAAI Conference on Artificial Intelligence*, volume 37, pp. 11121–11128, 2023. doi: 10.1609/AAAI.V37I9.26317.

Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *AAAI Conference on Artificial Intelligence*, volume 35, pp. 11106–11115, 2021. doi: 10.1609/AAAI.V35I12.17325.

APPENDIX

This appendix is organized into six sections:

- **Appendix A: Theoretical Proofs and Mathematical Foundations** detailing Parseval weighting, seed-averaged risk decomposition (Theorem 1), checkpoint-specific gain estimation (Theorem 2), ridge sensitivity, coordinate clipping, and weighted projection (Theorem 3).
- **Appendix B: Full Multi-Seed Diagnostic Evidence** presenting spectral variance distributions across all cells, the deterministic cell/mode selection protocol for Figure 1, and training-split verification of near-zero predictive gain.
- **Appendix C: Experimental Setup, Covariance-Resolution Family, and Protocols** specifying benchmark datasets, 10 forecasting backbones, the complete G_0–G_∞ operator family, nominal dimension guards, dual fitting sources, and statistical testing procedures.
- **Appendix D: Detailed Forecasting Benchmarks and Peer Compositions** reporting full-matrix MSE/MAE results across four horizons ($H \in \{96, 192, 336, 720\}$), complete Traffic tables, and serial compositions with both post-hoc adapters and training losses.
- **Appendix E: Empirical Mechanism and Finite-Sample Analyses** investigating aligned vs. scrambled structure, empirical calibration-size trends, temporal-distance controls, resolution lattice inclusions, and singular second moments.
- **Appendix F.1: Algorithmic Implementation and Computational Complexity Analysis** providing the formal execution workflow for one-checkpoint covariance-resolution selection (Algorithm 1) alongside runtime, memory footprint, and scalability analyses.

A THEORETICAL PROOFS AND TECHNICAL DETAILS

This section collects technical details supporting the theoretical results in the main text, establishes the Parseval spectral weighting convention, and details the proofs of Theorems 1, 2, and 3.

A.1 PARSEVAL WEIGHTING AND SPECTRAL RISK

For an even horizon H, an rFFT retains $M = H/2 + 1$ bins. Write $q_1 = q_M = 1$ for DC and Nyquist modes, and $q_j = 2$ for interior bins. Under orthonormal Fourier normalization, a real residual $E \in \mathbb{R}^{H\times C}$ satisfies:

$$\|E\|_F^2 = \sum_{j=1}^{M} q_j \|F_j(E)\|^2, \qquad \mathrm{MSE}(E) = \frac{1}{HC}\sum_{j=1}^{M} q_j \|F_j(E)\|^2.$$

An unnormalized forward FFT requires an additional factor $1/H$ on the spectral side. The ratio of unweighted rFFT energy to full energy depends on endpoint energy and is not universally $(H/2 + 1)/H$.

Our theoretical formulation stacks $\sqrt{q_j}\,\mathrm{rv}(s_j)$ and $\sqrt{q_j}\,\mathrm{rv}(e_j)$, suppressing the global $1/(HC)$ factor. Removing the zero endpoint imaginary coordinates leaves HC real coordinates. Let $J \in \mathbb{R}^{2MC\times HC}$ insert these zero coordinates, so $J^\top J = I$, and let $T \in \mathbb{R}^{HC\times HC}$ apply the weights $\sqrt{q_j}$. A nominal implementation map L induces the theoretical operator

$$K = TJ^\top LJT^{-1}.$$

Here J embeds the input and $J^\top$ discards endpoint imaginary outputs, as inverse rFFT does. Thus risk analysis uses the effective map on real signals even when the implementation retains nominal blocks.

At fixed nominal dimension, rescaling both predictor and response within one mode multiplies their second moments and an ideal trace-scaled ridge penalty by the same factor, leaving the fit unchanged. The numerical floor in Eq. 24 is applied in the raw implementation coordinates. Removing zero coordinates does not recompute its nominal dimension or penalty. Neighboring-mode inputs can have different Parseval weights; their existing raw-coordinate fit is transported by the map above, rather than refitted with an isotropic penalty in weighted coordinates.

Candidate selection uses the unweighted one-sided spectral score in Eq. 23, whereas test MSE uses the Parseval weights above. The projection and gain identities also hold for unweighted spectral risk

by using the corresponding unweighted vectors and moments consistently. Throughout the main risk analysis, $\mathcal{R}$ denotes the Parseval-weighted reference risk; a candidate's validation score and its evaluated time-domain risk are distinct quantities.

A.2 Proofs and Scope of the Risk Results

Proof of Theorem 1 (Seed-averaged Risk Decomposition). At any evaluation window w, the target y_w is fixed across training seeds k. Decomposing the forecast and residual into seed-mean and seed-dependent fluctuations ($\delta_w^{(k)} = \mu_w + \xi_w^{(k)}$ and $\epsilon_w^{(k)} = \bar{\epsilon}_w + \xi_w^{(k)}$ with $\mathbb{E}_k[\xi_w^{(k)}] = 0$), the corrected error expands as:

$$\epsilon_w^{(k)} - K\delta_w^{(k)} = (\bar{\epsilon}_w - K\mu_w) + (I - K)\xi_w^{(k)}.$$

For any linear operator K held fixed across training runs, the cross term vanishes under the conditional expectation over seeds $\mathbb{E}_{k|w}[\cdot]$ since $\mathbb{E}_k[\xi_w^{(k)}] = 0$:

$$\mathbb{E}_k\|\epsilon_w^{(k)} - K\delta_w^{(k)}\|^2 = \|\bar{\epsilon}_w - K\mu_w\|^2 + \operatorname{tr}\left((I - K)\mathbb{E}_k[\xi_w^{(k)}\xi_w^{(k)\top}](I - K)^\top\right).$$

Averaging over windows w yields Eq. 8. Furthermore, direct expansion of second moments at each window gives:

$$\mathbb{E}_k[\delta_w^{(k)}\delta_w^{(k)\top}] = \mu_w\mu_w^\top + \mathbb{E}_k[\xi_w^{(k)}\xi_w^{(k)\top}], \tag{16}$$

$$\mathbb{E}_k[\epsilon_w^{(k)}\delta_w^{(k)\top}] = \bar{\epsilon}_w\mu_w^\top + \mathbb{E}_k[\xi_w^{(k)}\xi_w^{(k)\top}]. \tag{17}$$

Averaging Eqs. 16–17 over w establishes $\mathbb{E}_k[\Sigma^{(k)}] = \Sigma_\mu + \Sigma_\xi$ and $\mathbb{E}_k[\Gamma^{(k)}] = \Gamma_\mu + \Sigma_\xi$. Finite joint second moments ensure integrability. □

Proof of Theorem 2 (Reliability of Checkpoint-Specific Gain). Fix the checkpoint k and suppress its index. Let $\Delta\Gamma = \widehat{\Gamma} - \Gamma$ and $\Delta\Sigma = \widehat{\Sigma} - \Sigma$. The residual-only term cancels when taking the difference between empirical and population gains:

$$\widehat{g}(K) - g(K) = 2\operatorname{tr}(K\Delta\Gamma^\top) - \operatorname{tr}(K\Delta\Sigma K^\top). \tag{18}$$

For $K \in \mathcal{K}_g$, Frobenius orthogonality ensures $\langle K, \Delta\Gamma\rangle_F = \langle K, \Pi_g^F\Delta\Gamma\rangle_F$. Applying Cauchy–Schwarz:

$$|\operatorname{tr}(K\Delta\Gamma^\top)| \le \|K\|_F\|\Pi_g^F\Delta\Gamma\|_F = \|K\|_F a_g.$$

For nonzero K, define $A = K/\|K\|_F \in \mathcal{K}_g$ with $\|A\|_F = 1$. By the definition of b_g:

$$|\operatorname{tr}(K\Delta\Sigma K^\top)| = \|K\|_F^2|\operatorname{tr}(A\Delta\Sigma A^\top)| \le \|K\|_F^2 b_g.$$

Combining both terms via the triangle inequality yields the single-operator bound:

$$|\widehat{g}(K) - g(K)| \le 2\|K\|_F a_g + \|K\|_F^2 b_g. \tag{19}$$

Taking the supremum over $\mathcal{C}_g(B_K) = \{K \in \mathcal{K}_g : \|K\|_F \le B_K\}$ directly proves Eq. 11. Because the bound holds uniformly for all $K \in \mathcal{C}_g(B_K)$ simultaneously, choosing K based on the calibration sample does not invalidate the guarantee. If $a_g, b_g \to 0$ in probability, then $\eta_g(B_K) \to 0$ in probability. □

Nested Classes and Capacity Trade-off. The projection norm has the dual representation

$$a_g = \sup_{\substack{A\in\mathcal{K}_g \\ \|A\|_F\le 1}} |\langle A, \Delta\Gamma\rangle_F|.$$

When classes are nested ($\mathcal{K}_g \subseteq \mathcal{K}_h$), the supremum domains for both a_g and b_g expand, meaning $a_g \le a_h$ and $b_g \le b_h$. Simultaneously, $\mathcal{C}_g(B_K) \subseteq \mathcal{C}_h(B_K)$ guarantees that population approximation error cannot increase: $\inf_{K\in\mathcal{C}_h(B_K)}\mathcal{R}(K) \le \inf_{K\in\mathcal{C}_g(B_K)}\mathcal{R}(K)$. This formalizes the core approximation-estimation tension in Eq. 13 Theorem 2 provides a uniform gain perturbation bound over any pre-specified norm ball $\mathcal{C}_g(B_K)$, characterizing the estimation complexity across structural classes rather than serving as an unconstrained finite-sample uniform guarantee for the final unclipped estimators .

Ridge Sensitivity to Moment Error. For any regularized linear solve with positive penalty τ, let $K_\tau = \Gamma(\Sigma + \tau I)^{-1}$ and $\widehat{K}_\tau = \widehat{\Gamma}(\widehat{\Sigma} + \tau I)^{-1}$. The normal equations imply the exact algebraic identity:

$$\widehat{K}_\tau - K_\tau = (\Delta\Gamma - K_\tau \Delta\Sigma)(\widehat{\Sigma} + \tau I)^{-1}.$$

Since $\widehat{\Sigma}$ is positive semidefinite, $\|(\widehat{\Sigma} + \tau I)^{-1}\|_{\mathrm{op}} \le 1/\tau$. Therefore:

$$\|\widehat{K}_\tau - K_\tau\|_F \le \frac{\|\Delta\Gamma\|_F + \|K_\tau\|_F \|\Delta\Sigma\|_{\mathrm{op}}}{\tau}. \tag{20}$$

This pointwise bound compares empirical and population ridge solutions at the same positive penalty τ. It quantifies the effect of moment perturbations, separately from the difference between K_τ and the unregularized optimum $K^\star$.

Coordinate Clipping. For a decoupled scalar coordinate, the empirical risk is $\widehat{\mathcal{R}}(\kappa) = \widehat{\mathcal{R}}(0) - 2\kappa\widehat{b} + \kappa^2\widehat{a}$. The implementation adds $\zeta = 10^{-12}$ to the accumulated input energy. With moments averaged over n_f windows, it minimizes

$$\widehat{\mathcal{R}}(\kappa) + \frac{\zeta}{n_f}\kappa^2 = \left(\widehat{a} + \frac{\zeta}{n_f}\right)\left(\kappa - \frac{\widehat{b}}{\widehat{a} + \zeta/n_f}\right)^2 + \text{constant}.$$

The quadratic coefficient is strictly positive. Projection onto $[0, 1]$ therefore proves Eq. 15. If $\widehat{a} = 0$, all fitting inputs and $\widehat{b}$ vanish, giving $\widehat{\kappa} = 0$. Without the stabilizer, the unregularized constrained solution is $\Pi_{[0,1]}(\widehat{b}/\widehat{a})$ for $\widehat{a} > 0$.

A.3 Weighted Projection and Its Consequences

Proof of Theorem 3. Work in the common finite-dimensional signal space with finite second moments. For $v \in \ker\Sigma$, $\mathbb{E}(\delta^\top v)^2 = v^\top \Sigma v = 0$, hence $\delta^\top v = 0$ almost surely and $\Gamma v = 0$. Thus $\ker\Sigma \subseteq \ker\Gamma$, so $K^\star = \Gamma\Sigma^+$ satisfies $K^\star\Sigma = \Gamma$. For $A = K - K^\star$, the risk expansion gives

$$\begin{aligned}\mathcal{R}(K) - \mathcal{R}(K^\star) &= -2\,\mathrm{tr}(A\Gamma^\top) + 2\,\mathrm{tr}(A\Sigma K^{\star\top}) + \mathrm{tr}(A\Sigma A^\top) \\ &= \|A\|_\Sigma^2. \end{aligned} \tag{21}$$

The normal equation cancels the first-order terms. All unrestricted minimizers have the form $K^\star + B(I - \Sigma\Sigma^+)$ for an arbitrary matrix B. Since $K^\star(I - \Sigma\Sigma^+) = 0$, it is Frobenius-orthogonal to every added term and has the smallest Frobenius norm, as asserted in Eq. 6.

To establish existence when Σ is singular, consider the image $\{K\Sigma^{1/2} : K \in \mathcal{K}\}$. It is a finite-dimensional linear subspace and is therefore closed. Euclidean projection of $K^\star\Sigma^{1/2}$ onto this image exists and is unique. Any preimage $K_{\mathcal{K}}^\star \in \mathcal{K}$ minimizes risk. Two such preimages differ by an operator A satisfying $A\Sigma^{1/2} = 0$, which acts as zero on δ almost surely.

The first-order condition for this projection is $\mathrm{tr}((K_{\mathcal{K}}^\star - K^\star)\Sigma A^\top) = 0$ for every $A \in \mathcal{K}$. For $K \in \mathcal{K}$, take $A = K - K_{\mathcal{K}}^\star$ and expand:

$$\|K - K^\star\|_\Sigma^2 = \|K - K_{\mathcal{K}}^\star\|_\Sigma^2 + \|K_{\mathcal{K}}^\star - K^\star\|_\Sigma^2.$$

Combining this with Eq. 21 proves the theorem. The identity is independent of the representative chosen in each zero-risk equivalence class. □

Fitted operators and the deployed structures. The projection benchmark is the unregularized population optimum in the underlying linear space. The deployed fits remain elements of that space: G_2 clips one shared real–imaginary coefficient per mode–channel pair and gates the pair together; G_3 gates the whole output-mode map, preserving its shared real–imaginary channel map; and G_4 and $G_5(3)$ gate whole output-mode maps without adding dependencies. Ridge fitting and source-specific refitting likewise preserve the class's coefficient constraints. These statements use the actual block gates in Algorithm 1; arbitrary entrywise gates need not preserve coefficient ties. Consequently, Eq. 13 holds for the resulting fitted operator. Clipping, ridge, gating, and source selection affect $\|\widehat{K}_g - K_g^\star\|_\Sigma^2$ rather than changing the population projection benchmark.

A.4 Recoverable Error and Residual Ambiguity

Setting $K = 0$ in Theorem 3, using $0 \in \mathcal{K}$, gives

$$\mathcal{R}(0) - \mathcal{R}(K^\star_{\mathcal{K}}) = \|K^\star_{\mathcal{K}}\|^2_\Sigma.$$

For the unrestricted class, substituting $K^\star = \Gamma\Sigma^+$ and $\Sigma^+\Sigma\Sigma^+ = \Sigma^+$ yields $\mathcal{C}_{\mathrm{lin}} = \mathrm{tr}(\Gamma\Sigma^+\Gamma^\top)$, proving Eq. 7.

Residual energy, and even the two separate marginals, cannot determine this quantity. Let $\delta = U$, where U is uniform on $\{-1, 1\}$, and let $\epsilon = UZ$ with $Z \in \{-1, 1\}$ independent of U, with $\Pr(Z = 1) = p$ and $\Pr(Z = -1) = 1 - p$ for $0 \le p \le 1$. Define the target as $y = \delta - \epsilon$. For every p, both δ and ϵ are uniform on $\{-1, 1\}$, so $\mathbb{E}\epsilon^2 = \Sigma = 1$. However, $\Gamma = 2p - 1$, $K^\star = 2p - 1$, and $\mathcal{C}_{\mathrm{lin}} = (2p - 1)^2$. Thus $p = 1/2$ and $p = 3/4$ have identical forecast and residual marginals but recoverable errors 0 and $1/4$. Their joint forecast–error relationships distinguish the useful corrections.

A.5 Incremental Value of Nested Structures

Proof of Corollary 4. Fix the same reference distribution and common operator space for $\mathcal{K}_g \subseteq \mathcal{K}_h$. Apply Theorem 3 in $\mathcal{K}_h$ to $K^\star_g \in \mathcal{K}_h$:

$$\mathcal{R}(K^\star_g) - \mathcal{R}(K^\star_h) = \|K^\star_g - K^\star_h\|^2_\Sigma.$$

Apply the theorem separately to $\widehat{K}_g$ and $\widehat{K}_h$ in their own spaces, then subtract the identities. This gives

$$\mathcal{R}(\widehat{K}_g) - \mathcal{R}(\widehat{K}_h) = \|K^\star_h - K^\star_g\|^2_\Sigma + \|\widehat{K}_g - K^\star_g\|^2_\Sigma - \|\widehat{K}_h - K^\star_h\|^2_\Sigma.$$

All terms are invariant to zero-risk representatives, so the result also holds for singular Σ. □

A.6 Training-Run Averages and Checkpoint-Specific Estimation

The seed-averaged least-squares solution. Write $\overline{\Sigma} = \Sigma_\mu + \Sigma_\xi$ and $\overline{\Gamma} = \Gamma_\mu + \Sigma_\xi$. The risk change is

$$\overline{\mathcal{R}}(K) - \overline{\mathcal{R}}(0) = -2\,\mathrm{tr}(K\overline{\Gamma}^\top) + \mathrm{tr}(K\overline{\Sigma}K^\top).$$

For $v \in \ker\overline{\Sigma}$, $\mathbb{E}_{k,w}(\delta^{(k)\top}_w v)^2 = 0$, so $\delta^{(k)\top}_w v = 0$ almost surely and $\overline{\Gamma}v = 0$. Thus $\overline{K}^\star = \overline{\Gamma}\,\overline{\Sigma}^+$ solves $K\overline{\Sigma} = \overline{\Gamma}$ and is the minimum-Frobenius-norm minimizer of the unrestricted seed-averaged risk. This is the best operator shared across training runs. The corresponding checkpoint-specific minimizer is $K^{(k)\star} = \Gamma^{(k)}(\Sigma^{(k)})^+$; these operators need not agree.

For example, let two equally likely checkpoints predict $\delta^{(1)} = 1$ and $\delta^{(2)} = 2$ for the constant target $y = 1$. Their optimal correction coefficients are 0 and $1/2$, whereas the best shared coefficient is $\overline{K}^\star = 2/5$. Additional windows from either fixed checkpoint do not change its optimum into the shared one.

Training-run variation and calibration estimation error are different quantities. The seed-averaged risk contains the retained-variation term $(I - K)\Sigma_\xi(I - K)^\top$, whereas ridge penalizes the correction coefficients K to stabilize finite-sample fitting. These are not the same penalty.

Unbiasedness and realized checkpoints. For a fixed operator, empirical risk also satisfies a conditional unbiasedness identity. Suppose, conditional on k, each calibration window W_i has the reference marginal distribution P, and K is fixed independently of the calibration sample. Then linearity of expectation gives

$$\mathbb{E}_{D|k}\widehat{\mathcal{R}}^{(k)}_n(K) = \mathcal{R}^{(k)}(K).$$

For a common fixed K, taking the seed expectation yields $\mathbb{E}_{k,D}\widehat{\mathcal{R}}^{(k)}_n(K) = \overline{\mathcal{R}}(K)$. Independence between windows is not necessary for these expectation identities. They do not identify a realized checkpoint's risk with the seed-averaged risk, nor apply automatically to a K fitted on those same windows.

Moment consistency and temporal dependence. For a fixed checkpoint, a sufficient condition for consistency of empirical moments is that the calibration forecast–residual sequence is stationary and ergodic with the reference marginal distribution and finite second moments. The ergodic theorem then applies to each entry of $\delta_w \delta_w^\top$ and $\epsilon_w \delta_w^\top$. In fixed dimension, entrywise convergence implies convergence in the matrix norms used above. For any fixed K, the same argument gives $\widehat{\mathcal{R}}_n^{(k)}(K) \to \mathcal{R}^{(k)}(K)$ almost surely, not convergence to $\overline{\mathcal{R}}(K)$ in general. This is an additional sampling condition, not a consequence of merely collecting more overlapping windows. Finite-sample rates require further tail and temporal-dependence assumptions.

The reference distribution in Theorem 2 must be specified. If it is the deployment distribution P_{dep} and the fitting data follow P_{cal}, the total matrix errors include both sampling error and distribution change, for example

$$\widehat{\Sigma}_{\mathrm{cal}} - \Sigma_{\mathrm{dep}} = (\widehat{\Sigma}_{\mathrm{cal}} - \Sigma_{\mathrm{cal}}) + (\Sigma_{\mathrm{cal}} - \Sigma_{\mathrm{dep}}).$$

The deterministic gain bound still holds with these total errors. More calibration windows alone do not eliminate the second term. Calibration windows reused from backbone fitting or checkpoint selection do not automatically satisfy the conditional sampling assumptions in the unbiasedness observation; the gain perturbation bound does not require those assumptions.

B Full multi-seed diagnostic evidence

The multi-seed study establishes the empirical phenomenon that motivates correction. It measures forecast skill and training-run variation using independently trained checkpoints; it is not part of the deployed fitting procedure. Using the same 74-cell subset as the measurement and mechanism diagnostics, each cell's three independently seeded checkpoints are evaluated with frozen forward passes. The diagnostic subset contains four Traffic backbones. The complete 80-cell outcome matrix adds six Traffic backbones. Table 5 reports the two measurements defined in Eqs. 2 and 3: zero-baseline relative MSE reduction Δ_j and normalized seed variance v_j. Among the 65 cells containing both reporting bands, every cell has higher energy-weighted $v_{\mathcal{B}}$ in the weak-skill band. The matrix medians are 29% and 0.2%. For a reporting band $\mathcal{B}$, define $v_{\mathcal{B}} = \sum_{j\in\mathcal{B}} \sigma_j^2 / \sum_{j\in\mathcal{B}} \mathbb{E}\|F_j(S)\|^2$. This is variance relative to that band's forecast energy. Dataset medians for the weak-skill mode fraction range from 0% on Solar to 90% on Weather. We therefore report support count alongside the variance statistic.

Numerical diagnostic conventions. The ratios in Section 3 are defined for positive target or forecast energy, respectively, and seed variance uses $N_s \geq 2$. Reporting uses unnormalized FFT coefficients and a denominator floor $\rho = 10^{-30}$. For skill, each seed's residual energy is divided by the target energy summed over windows and channels, floored at ρ, before averaging over seeds. For v_j, the forecast energy is first averaged over seeds and windows and summed over channels, then floored at ρ. A nonempty band's variance uses the summed mode variances divided by the summed forecast energies, with the same floor in the diagnostic routine; an empty band is undefined and excluded from its median. All exported mode-level forecast energies in the 74-cell diagnostic are positive and exceed this floor, so nonempty-band aggregation uses the ordinary ratio. These numerical conventions give zero normalized variance for an identically zero forecast. A zero-target-energy mode has no unregularized relative skill; its recorded numerical value uses the floored denominator.

Table 5: Spectral diagnostics by dataset, median over cells, three seeds each. "Modes in $\mathcal{B}_{\text{near}}$" is the weak-skill fraction. "Cells with any" counts cells containing at least one such mode. The last column reports energy-weighted band variance $v_{\mathcal{B}}$ in $\mathcal{B}_{\text{near}}$ and $\mathcal{B}_{\text{high}}$. Cells with empty weak-skill support are counted but do not enter that band's median.

dataset	C	cells	modes in $\mathcal{B}_{\text{near}}$	cells with any	$v_{\mathcal{B}}$: weak / high
ETTh1	7	10	82%	10/10	43% / 0.2%
ETTh2	7	10	88%	10/10	37% / 0.1%
ETTm1	7	10	80%	10/10	56% / 0.7%
ETTm2	7	10	88%	10/10	34% / 0.0%
Weather	21	10	90%	10/10	10% / 0.1%
Solar	137	10	0%	1/10	21% / 0.5%
ECL	321	10	45%	10/10	15% / 0.2%
Traffic	862	4	31%	4/4	9% / 0.6%

Table 6: Zero-baseline relative MSE reduction and normalized seed variance for iTransformer with H=96, 49 rFFT modes, and seeds 2/3/4. The reporting bands use the single margin $t_{\text{skill}} = 0.05$: near-zero modes satisfy $\Delta_j \leq t_{\text{skill}}$ and the separated reference band satisfies $\Delta_j > 6t_{\text{skill}}$. The normalized seed variance is energy-weighted within each band.

	ETTh1	ETTh2	ETTm1
modes in $\mathcal{B}_{\text{near}}$	40/49	43/49	39/49
$v_{\mathcal{B}}$ in $\mathcal{B}_{\text{near}}$ [†]	**56**%	**60**%	**77**%
$v_{\mathcal{B}}$ in $\mathcal{B}_{\text{high}}$	1.0%	0.2%	2.5%
mean training-split Δ_j in $\mathcal{B}_{\text{near}}$	−0.032	−0.061	−0.015

[†] Energy-weighted as $\sum_j \sigma_j^2 / \sum_j \mathbb{E}\|F_j(S)\|^2$, using the unbiased sample variance across seeds on uncorrected forecasts. The unweighted ETTh1 mean is 78%.

Table 6 gives representative mode-level measurements for iTransformer, and Table 7 checks the reporting-threshold sensitivity.

Table 7: Sensitivity to forecast-skill cutoffs. Weak-band cutoffs range from 0% to 10% in 2.5-point steps; high-band cutoffs range from 20% to 40% in 5-point steps. For all 25 cutoff pairs, every cell with both bands has higher energy-weighted normalized seed variance in its weak-skill band.

near cutoff	median near modes	near higher at five high cutoffs	v_{near}	v_{high} range
0.0%	53.1%	$5 \times 48/48$	57.80%	0.151–0.163%
2.5%	78.6%	$5 \times 59/59$	35.58%	0.135–0.154%
5.0%	82.7%	$5 \times 65/65$	28.80%	0.149–0.155%
7.5%	82.7%	$5 \times 65/65$	25.06%	0.149–0.155%
10.0%	82.7%	$5 \times 65/65$	24.34%	0.149–0.155%

These multi-seed measurements show where training-run variation concentrates. One-checkpoint deployment then uses held-out forecast–target moments to fit and select the correction directly.

Deterministic selection of Figure 1. Among the 65 cells containing both reporting bands, we rank each cell by its weak-skill mode fraction and by its normalized seed variance in the near-zero- and high-skill bands. ETTh2/TimeXer minimizes the summed distance of these three empirical ranks from their median ranks. Panel 1(a) shows all 49 rFFT modes in that cell and highlights the selected pair. The pair minimizes the absolute log ratio of residual energy across the high- and weak-skill sets. These bands use the single reporting margin $t_{\text{skill}} = 0.05$. These fixed rules specify the reported cell and mode pair.

B.1 Training-split verification of weak forecast skill

We evaluate on training data the same rFFT mode indices identified on the test split. Across the 65 cells with weak-skill test modes, their mean training-split Δ_j has median 0.005 and maximum 0.041. At coordinate level, the 2804 test-near-zero coordinates have energy-weighted mean training Δ_{jc} of -0.020, compared with $+0.605$ for high-skill coordinates. No dataset-level mean exceeds $+0.024$. The weakness is therefore already present on the training split. This matched-index mechanism analysis is distinct from the cross-fitted deployment rule.

B.2 From multi-seed diagnosis to one-checkpoint deployment

Independent seeds are used only to expose how forecast skill and training-run variation differ across spectral modes. Deployment does not require repeated training: each selection unit consists of one frozen checkpoint and its labeled calibration windows. Theorem 1 separates seed-mean error from retained training variation in the risk averaged over runs. Theorem 2 instead bounds the error in estimating corrective gain for the realized checkpoint. More windows can improve its moment estimates under suitable sampling conditions; they do not identify cross-seed covariance. Held-out validation supplies empirical evidence of improvement beyond the fitting windows.

C Experimental setup and implementation details

This section specifies the data, forecasting backbones, correction classes, selection protocol, and paired evaluation procedure used in all deployment experiments.

C.1 Notation and evaluation units

We first fix notation. Let H denote the forecast horizon, C the number of channels, and $M = H/2+1$ the number of modes produced by the real-valued fast Fourier transform (rFFT). A *spectral coordinate* is one mode–channel pair. A *cell* is one dataset–backbone–horizon combination whose reported statistics average seeds 2, 3, and 4. The *pass-through baseline* is the forecast from the same frozen checkpoint before correction. Relative mean-squared-error (MSE) reduction is $100(\mathrm{MSE}_{\mathrm{base}} - \mathrm{MSE}_{\mathrm{corr}})/\mathrm{MSE}_{\mathrm{base}}$. Positive values indicate lower error.

C.2 Datasets and chronological splits

Table 8 lists each dataset's measurements, dimensionality, sampling interval, and source. The ETT datasets use the benchmark's first 12 months for training, next 4 for validation, and final 4 for test. Weather, ECL, Solar, and Traffic use the first 70%, next 10%, and final 20%, respectively. Validation and test windows may read the preceding input context, but every forecast target remains inside its split. Standardization is fitted on training observations and applied unchanged to validation and test data.

Experiments use input length 96, horizons $H \in \{96, 192, 336, 720\}$, Adam optimization, and early stopping with patience 3 on validation MSE. A *calibration split* is a labeled training or validation split used to fit and score a post-hoc operator before test evaluation. The method uses the zero reference and selects operator class, ridge penalty, and fitting source with held-out validation error. It requires one frozen checkpoint and no additional backbone forward pass. Training and validation data determine every fitted operator and deployed choice before test evaluation. Test labels evaluate that frozen choice. The explicitly marked non-deployable oracle is reported only as a reference for validation-based selection. Released scripts reproduce tables from run outputs.

C.3 Forecasting backbones and competing methods

Every horizon covers ten forecasting backbones, grouped here by their main modeling mechanism:

- The linear and lightweight group contains DLinear, which applies linear prediction after trend–remainder decomposition (Zeng et al., 2023), and MixLinear, a low-resource multivariate linear model (Ma et al., 2026).

Table 8: Datasets used in the forecasting matrices. C is the number of channels forecast jointly. The sampling interval is the time between consecutive rows in the files used here.

dataset	measured variables	C	interval	source
ETTh1/ETTh2	transformer oil temperature and six power-load measurements	7	1 h	ETT benchmark (Zhou et al., 2021; Zeng et al., 2023)
ETTm1/ETTm2	the same transformer measurements at finer resolution	7	15 min	ETT benchmark (Zhou et al., 2021; Zeng et al., 2023)
Weather	temperature, humidity, pressure, wind, rainfall, and related Jena meteorology	21	10 min	Autoformer benchmark (Wu et al., 2021)
ECL	electricity consumption of 321 clients, aggregated from 15-minute load diagrams	321	1 h	UCI and Informer (Trindade, 2015; Zhou et al., 2021)
Solar	photovoltaic production from 137 plants in Alabama	137	10 min	LSTNet benchmark (Lai et al., 2018)
Traffic	road occupancy from 862 sensors on San Francisco Bay Area freeways	862	1 h	LSTNet benchmark (Lai et al., 2018)

- The frequency-domain group contains FITS, which interpolates a truncated spectral representation (Xu et al., 2024), and PaiFilter, the first of the two variants introduced in FilterNet, which learns frequency-wise filters (Yi et al., 2024).
- The Transformer group contains PatchTST, which represents each channel through temporal patches (Nie et al., 2023); iTransformer, which treats variables as tokens (Liu et al., 2024); and TimeXer, which models interactions between endogenous and exogenous variables (Wang et al., 2024b).
- The remaining three backbones emphasize multivariate structure or multiscale mixing: DUET uses dual clustering (Qiu et al., 2025), FACT uses fine-grained across-variable convolution (Wang et al., 2026), and TimeMixer mixes decomposed temporal patterns across scales (Wang et al., 2024a).

Published correction and training objectives are evaluated at their original interfaces. δ-Adapter, UEC-STD, and Socket+Plug are output-side post-hoc corrections, whereas FreDF and PS Loss modify backbone training. For the composed variants, our selector is fitted on the exact cached output of the corresponding peer method.

C.4 Complete covariance-resolution family

The complete family includes the four deployed candidates and the intermediate structural classes. Each class below specifies a linear subspace of operators acting on the real-stacked spectrum. Its class-optimal population operator minimizes the same reference risk over that subspace. Table 9 states each class's allowed dependencies. The four with a check mark form the deployed candidate set.

C.5 Closed-form fits and deployed candidates

Fits use the unnormalized forward rFFT and enforce each class's coefficient ties. For G_3, write $x_{R,wj} = \mathrm{Re}(s_{wj})$, $x_{I,wj} = \mathrm{Im}(s_{wj})$ and $r_{R,wj} = \mathrm{Re}(e_{wj})$, $r_{I,wj} = \mathrm{Im}(e_{wj})$, all in $\mathbb{R}^C$. The shared map B_j minimizes

$$\frac{1}{n_f}\sum_{w}\left(\|r_{R,wj} - B_j x_{R,wj}\|^2 + \|r_{I,wj} - B_j x_{I,wj}\|^2\right) + \tau\|B_j\|_F^2.$$

Table 9: The class family. $M{=}H/2{+}1$ is the number of rFFT modes and C the channel count, so a count is per mode times M and includes nominal zero or repeated input coordinates. Three branches leave G_2: $G_2^{\mathbb{C}}$ and G_2' both enter G_2'', while G_3 enters G_3'. All three branches enter G_4, followed by $G_5(3)$, $G_5(5)$, and G_∞. Class G_4 is unrestricted *within* a mode but diagonal *in* modes. $G_5(w)$ locally approaches the highest-resolution linear reference G_∞ with odd neighborhood width w.

class	nominal parameters	allowed action at one mode	candidate
G_0	1	one scalar for all modes and channels	
G_1	M	one scalar per mode	
G_2	MC	one real scalar per (mode, channel)	✓
$G_2^{\mathbb{C}}$	$2MC$	one complex scalar per coordinate: rescale and rotate phase	
G_2'	$2MC$	separate real scalars on the real and imaginary parts	
G_2''	$4MC$	any real 2×2 per coordinate, including shear	
G_3	MC^2	real $C\times C$ channel map per mode, shared by both parts	✓
G_3'	$2MC^2$	independent $C\times C$ maps on the two parts	
G_4	$4MC^2$	real $2C\times 2C$ per mode: unrestricted *within* a mode	✓
$G_5(3)$	$12MC^2$	G_4 plus the 3 modes centered on each	✓
$G_5(5)$	$20MC^2$	G_4 plus the 5 modes centered on each	
G_∞	$(2MC)^2$	highest-resolution linear reference on the spectrum	

Its $C\times C$ fitting moments are

$$\widehat{S}_{3,j}=\frac{1}{n_f}\sum_w (x_{R,wj}x_{R,wj}^\top + x_{I,wj}x_{I,wj}^\top),$$
$$\widehat{T}_{3,j}=\frac{1}{n_f}\sum_w (r_{R,wj}x_{R,wj}^\top + r_{I,wj}x_{I,wj}^\top). \tag{22}$$

Differentiating gives $B_j(\widehat{S}_{3,j}+\tau I_C)=\widehat{T}_{3,j}$. Thus one regression pools the real and imaginary parts as $2n_f$ rows with C predictors, using normalization $1/n_f$ throughout. Its prediction block is $K_{3,j}=\mathrm{diag}(B_j,B_j)$.

For $g\in\{G_4,G_5(3)\}$, let $x_{g,wj}=\delta_{g,wj}$ be the real-stacked same-mode or three-mode input and $r_{wj}=\mathrm{rv}(e_{wj})$. Then

$$\widehat{S}_{g,j}=\frac{1}{n_f}\sum_w x_{g,wj}x_{g,wj}^\top,\qquad \widehat{T}_{g,j}=\frac{1}{n_f}\sum_w r_{wj}x_{g,wj}^\top.$$

The fitted matrix $W_{g,j}$ satisfies $W_{g,j}(\widehat{S}_{g,j}+\tau I)=\widehat{T}_{g,j}$ and is already the prediction block $K_{g,j}$. For uniform solver notation, set $W_{3,j}=B_j$; its full prediction block remains $\mathrm{diag}(W_{3,j},W_{3,j})$. The predictor dimensions are $d_3=C$, $d_4=2C$, and $d_5=6C$, including nominal zero or repeated input coordinates. The positive numerical penalty and solver convention are given in Eq. 24.

For the unregularized structural references, scalar ratios give G_0, G_1, G_2, $G_2^{\mathbb{C}}$, and G_2', while G_2'' uses a 2×2 solve per coordinate. Class G_0 pools modes with the weights of the chosen risk. Class G_3' fits separate $C\times C$ maps to the real and imaginary parts. Singular unregularized solves use the Moore–Penrose solution. Deployment instead uses the stabilized, clipped G_2 fit in Eq. 15 and the ridge fits above.

Deployment selects among $G_2, G_3, G_4, G_5(3)$. Its G_3–$G_5(3)$ candidates use ridge regression toward zero with one shared ridge multiplier per class–source candidate and the gates defined in Section C.6. The remaining ladder classes and G_∞ are used only for structural checks. Both fitting sources use the same two-half validation score. Training and validation data determine fitting and ranking.

C.6 Exact covariance-resolution selection protocol

Table 10 identifies the data used by each stage. Training and validation determine the correction coefficients, ridge penalty, block gates, and class–source choice. The resulting fixed operator acts on test forecasts; test labels supply evaluation metrics and the separately reported oracle.

Candidates and common score. Each frozen checkpoint is one selection unit. The four covariance resolutions are coordinate G_2, channel G_3, within-mode G_4, and local three-mode $G_5(3)$, defined in Table 9. Each class has a validation-fitted candidate and a training-fitted candidate.

Table 10: Data dependencies of one-checkpoint correction. Fitting, ridge selection, gates, and class–source selection use training and validation data. Test forecasts enter the frozen correction at inference; test labels are used for metric reporting and the separately reported oracle. The oracle does not determine the deployed choice.

stage	forecast and label source	quantity determined	test labels used?
Calibration rFFT	Training and validation forecasts and labels	Forecast and residual spectra	No
Moment estimation and coefficient fit	Training, or the fitting half of validation	Second moments and candidate coefficients	No
Global ridge selection	Unweighted spectral error on validation halves	One ridge multiplier per candidate	No
Block gates	Held-out validation halves	Coordinate or output-mode mask	No
Class and source selection	Gated validation scores	Operator class and fitting source	No
Full-source refit	Selected source: training or full validation	Final correction coefficients	No
Frozen inference	Test forecasts; no labels	Corrected forecasts	No
Metric and oracle reporting	Test forecasts and test labels	MSE, MAE, and diagnostic oracle	Yes

An rFFT acts along the H predicted steps of each window, not across a concatenation of dataset splits. Stride-one validation windows are constructed over the full segment and then divided chronologically into halves. Adjacent windows across the division shift by one time step, sharing input observations and $H-1$ target time points. Every validation forecast horizon ends before the first test target; the A/B overlap remains within validation. Training-fitted candidates score on both validation halves; validation-fitted candidates exchange their fitting and scoring halves. Evaluation of cached test forecasts is a separate dataflow branch and supplies no fitting moment, selection score, or gate. Any rFFT of test labels belongs to that evaluation branch, not calibration.

Te construct stride-one validation windows and split them chronologically at the index midpoint into halves A and B. To prevent temporal leakage, all forecast horizons in both folds terminate before the first test target. Validation-fitted candidates alternate fitting and scoring between the two halves ($A \leftrightarrow B$), while training-fitted candidates fit on the training split and evaluate across both halves The validation/test boundary instead limits window construction: each validation window, including its complete H-step target interval, ends at or before the validation endpoint. Test targets begin after that endpoint. A validation-fitted candidate fits on A and scores on B, then exchanges the roles of the halves. After selection it is refitted on all validation windows. A training-fitted candidate fits on the training split and scores on both A and B. Every candidate is ranked on the same validation halves. For the fitting/scoring pairs $\mathcal{P}_r$ of source r, its block score is

$$L_{r,g,\lambda,b} = \sum_{(\mathcal{D}_f,\mathcal{D}_s)\in\mathcal{P}_r} \sum_{w\in\mathcal{D}_s} \sum_{(j,c)\in b} \left| e_{wjc} - c_{wjc}^{(\mathcal{D}_f,g,\lambda)} \right|^2 . \tag{23}$$

Here $c^{(\mathcal{D}_f,g,\lambda)}$ is the correction fitted on $\mathcal{D}_f$, and a block is one mode–channel pair for G_2 or a full output mode for the matrix classes. This is the implemented unweighted one-sided spectral score, not Parseval-weighted time-domain MSE. Moment-consistency conditions for dependent windows are given in Appendix A.6.

G_2 uses the closed-form coefficient in Eq. 15. Classes G_3–$G_5(3)$ each evaluate 64 ridge penalties log-spaced from 10^{-4} to 10^5 and scaled per mode by the mean diagonal of the design second-moment matrix. One multiplier is shared across modes. Within each class–source candidate, the multiplier minimizes the ungated validation score; held-out block gates are then applied. The gated scores rank the candidates, as detailed in Algorithm 1. Ties favor the narrower class. Each seed is processed independently, producing one corrected forecast per checkpoint.

Dimension guard. Before comparing validation scores, we exclude designs whose predictor count leaves no positive nominal residual degrees of freedom. Let n_f be the number of fitting windows. In real-stacked form, G_3 has $2n_f$ stacked design rows for C channel predictors, G_4 has n_f rows for $2C$ predictors, and $G_5(3)$ has n_f rows for $6C$ predictors. The guard therefore excludes G_3 when $2n_f \le C$, G_4 when $n_f \le 2C$, and $G_5(3)$ when $n_f \le 6C$. We call this the *nominal dimension-admissibility rule*. It depends only on fitting-set size and operator width, not on validation outcomes

or test labels. Table 24 compares the unguarded variant fixed before the long-horizon experiments with the deployed dimension-guarded rule.

Complementary fitting sources. The fitting source determines which data are used to estimate the second moments, whereas the validation score determines which fitted operator is retained. Fitting on training data supplies many more labeled windows, providing more design rows relative to predictors and admitting wider covariance resolutions. Every training-fitted operator is evaluated on held-out validation windows. Fitting on validation data instead supplies later residuals from data unseen during backbone training, at the cost of fewer fitting windows and making fewer wide classes admissible. The $A \leftrightarrow B$ cross-fit evaluates each validation-fitted operator on the opposite half. After selection, a validation-fitted candidate is refitted on all validation windows; a training-fitted candidate uses the full training source. The selected gates are retained. The two sources therefore trade sample size against temporal proximity within one selection rule.

C.7 Evaluation metrics and paired uncertainty

The evaluation unit is a dataset–backbone cell. For cell i and seed s, corrected and baseline errors come from the same frozen checkpoint, and the cell statistic is the mean paired relative reduction $\bar{r}_i = 3^{-1}\sum_s 100(\mathrm{MSE}_{is}^{\mathrm{base}} - \mathrm{MSE}_{is}^{\mathrm{corr}})/\mathrm{MSE}_{is}^{\mathrm{base}}$. Pairing controls for differences between checkpoints while retaining training-run uncertainty. Every dataset–horizon group contains $B_{dh} = 10$ backbones, and Table 1 reports $B_{dh}^{-1}\sum_{i\in(d,h)}\bar{r}_i$ together with the sample standard deviation (SD) across those ten matched backbone-cell means. The signed-rank test uses the same ten means.

The *cell bootstrap* resamples these paired cell statistics as independent units. The *dataset-resampled bootstrap* samples eight complete datasets and retains all ten backbones within each sampled dataset; for the full-matrix interval, it also retains all four horizons. This grouping accounts for dependence among models and horizons evaluated on the same data. All bootstraps use 10^4 percentile resamples. The two-sided sign test uses only the direction of each cell reduction. The exact two-sided Wilcoxon signed-rank test also uses magnitude and enumerates signs over the observed absolute ranks. The final column counts cells whose mean reduction exceeds twice the standard error across their three paired seed-level changes.

D Additional experimental results

This section reports complete deployment results. All choices are fixed by training and validation data before test evaluation.

D.1 Overall paired results and uncertainty

Table 11: Paired uncertainty and exact tests for the complete matrix. The effect column gives mean ± SD across 80 dataset–backbone cell means per horizon and across all 320 cell means in the final row. Positive values indicate lower test MSE.

scope	improves	mean ± SD (%)	cell 95%	dataset 95%	sign p	signed-rank p	$>$ 2SE
H=96	77/80	$+7.61_{\pm 8.10}$	$[5.91, 9.38]$	$[3.48, 12.80]$	1.4×10^{-19}	2.3×10^{-22}	75/80
H=192	73/80	$+6.91_{\pm 8.26}$	$[5.17, 8.77]$	$[2.42, 12.43]$	5.8×10^{-15}	1.2×10^{-15}	70/80
H=336	68/80	$+6.22_{\pm 8.60}$	$[4.38, 8.14]$	$[1.36, 11.92]$	1.2×10^{-10}	7.8×10^{-10}	68/80
H=720	61/80	$+4.64_{\pm 9.11}$	$[2.69, 6.59]$	$[-0.27, 9.70]$	2.7×10^{-6}	9.3×10^{-6}	57/80
all four horizons	279/320	$+6.34_{\pm 8.56}$	$[5.43, 7.30]$	$[1.98, 11.59]$	1.2×10^{-44}	1.1×10^{-41}	270/320

Both bootstraps use 10^4 percentile resamples. The dataset-resampled bootstrap samples eight whole datasets and retains their ten backbones; in the final row it also retains all four horizons. $>$ 2SE counts cell means larger than twice their paired-seed standard error.

D.2 Complete results on Traffic

Tables 13 and 14 expand Traffic across all ten backbones, four horizons, and three seeds. They report the three independently selected source–class pairs, absolute MSE, and paired uncertainty for each of the 40 Traffic cells. Table 37 separately isolates unregularized pseudoinverse estimation to quantify the finite-sample behavior addressed by ridge estimation, mode-wise gates, and resolution selection.

Table 12: Paired uncertainty for the five peer inputs and their compositions with ours on the complete 80-cell $H{=}96$ matrix. Every percentage is paired to the baseline named in the first column. Panel (a) reports the effect size as mean $\pm$ SD across the 80 cell means; panel (b) reports the corresponding bootstrap intervals and exact tests.

(a) Effect size and directional consistency

comparison	improves	mean $\pm$ SD (%)	$>$ 2SE
δ-Adapter, standalone	$62/80$	$+3.17_{\pm 3.97}$	$55/80$
δ-Adapter+ours vs. plain	$71/80$	$+8.13_{\pm 9.55}$	$68/80$
ours after δ-Adapter	$72/80$	$+5.35_{\pm 6.64}$	$71/80$
UEC-STD input vs. plain	$67/80$	$+2.16_{\pm 3.83}$	$46/80$
UEC-STD+ours vs. plain	$76/80$	$+6.57_{\pm 6.67}$	$71/80$
ours after UEC-STD	$76/80$	$+4.55_{\pm 4.87}$	$67/80$
Socket+Plug input vs. plain	$75/80$	$+6.41_{\pm 7.10}$	$69/80$
Socket+Plug+ours vs. plain	$80/80$	$+10.34_{\pm 10.77}$	$77/80$
ours after Socket+Plug	$74/80$	$+4.43_{\pm 6.74}$	$72/80$
FreDF input vs. plain	$39/80$	$-11.72_{\pm 33.79}$	$26/80$
FreDF+ours vs. plain	$66/80$	$+6.25_{\pm 7.91}$	$61/80$
ours after FreDF	$76/80$	$+12.14_{\pm 15.37}$	$73/80$
PS Loss input vs. plain	$40/80$	$-0.80_{\pm 5.07}$	$33/80$
PS Loss+ours vs. plain	$77/80$	$+7.74_{\pm 7.36}$	$73/80$
ours after PS Loss	$77/80$	$+8.11_{\pm 9.53}$	$76/80$

(b) Bootstrap intervals and exact tests

comparison	cell bootstrap 95%	dataset-resampled 95%	sign p	Wilcoxon p
δ-Adapter, standalone	$[2.33, 4.06]$	$[0.94, 5.67]$	8.1×10^{-7}	1.1×10^{-10}
δ-Adapter+ours vs. plain	$[6.09, 10.30]$	$[3.02, 14.45]$	4.4×10^{-13}	1.5×10^{-16}
ours after δ-Adapter	$[3.96, 6.88]$	$[2.05, 9.78]$	5.4×10^{-14}	9.5×10^{-17}
UEC-STD input vs. plain	$[1.38, 3.08]$	$[0.70, 4.22]$	6.4×10^{-10}	1.3×10^{-14}
UEC-STD+ours vs. plain	$[5.17, 8.09]$	$[3.60, 9.64]$	2.8×10^{-18}	2.3×10^{-22}
ours after UEC-STD	$[3.52, 5.62]$	$[2.30, 6.92]$	2.8×10^{-18}	1.3×10^{-20}
Socket+Plug input vs. plain	$[4.94, 8.02]$	$[3.91, 9.14]$	4.3×10^{-17}	8.2×10^{-21}
Socket+Plug+ours vs. plain	$[8.10, 12.80]$	$[5.13, 16.66]$	1.7×10^{-24}	1.7×10^{-24}
ours after Socket+Plug	$[3.05, 5.99]$	$[1.15, 9.45]$	5.4×10^{-16}	6.1×10^{-18}
FreDF input vs. plain	$[-19.55, -5.31]$	$[-21.92, -3.08]$	0.9111	0.0433
FreDF+ours vs. plain	$[4.57, 7.96]$	$[2.55, 10.38]$	3.1×10^{-9}	3.7×10^{-11}
ours after FreDF	$[8.97, 15.53]$	$[5.40, 20.05]$	2.8×10^{-18}	2.4×10^{-20}
PS Loss input vs. plain	$[-1.96, 0.24]$	$[-3.85, 1.41]$	1.0000	0.9259
PS Loss+ours vs. plain	$[6.19, 9.40]$	$[4.04, 12.14]$	1.4×10^{-19}	1.5×10^{-19}
ours after PS Loss	$[6.11, 10.24]$	$[2.98, 14.46]$	1.4×10^{-19}	2.4×10^{-20}

D.3 Comparison with existing corrections and training losses

Table 15 next compares the selected correction, published alternatives, and their serial combinations with ours under a common plain-training baseline.

Among the four standalone post-hoc methods, ours has the largest mean reduction (7.61%) and coverage ($77/80$). Fitting our selector after δ-Adapter, UEC-STD, and Socket+Plug further reduces their respective inputs by $+5.35\%$, $+4.55\%$, and $+4.43\%$ on average. Socket+Plug+ours is the strongest composition, lowering MSE relative to plain training in $80/80$ cells by $+10.34\%$. FreDF and PS Loss cover the same complete matrix; ours further reduces their cached forecasts by $+12.14\%$ and $+8.11\%$. Tables 16–19, 22, and 23 report the matched raw MSE and MAE.

D.4 Per-cell comparisons

Table 16 reports the complete matrix for δ-Adapter, an output-side post-hoc correction, by dataset and backbone. This configuration freezes each plain-training forecaster, fits a 512-hidden-unit output adapter on the training split, and selects the epoch by validation MSE. Each change is paired with that run's unedited forecast.

Table 13: Complete Traffic results, part 1 of 2. Together the two parts cover ten backbones, four horizons, and seeds $2/3/4$. The selected column lists the independently chosen fitting source (T: training, V: validation) and covariance class. Every MSE and reduction entry is mean $\pm$ sample SD over the three paired seeds; reductions are percentages (↑).

horizon	backbone	selected at seeds $2/3/4$	baseline MSE	corrected MSE	reduction (%) ↑
H=96	DLinear	T:$G_5(3)$ (3/3)	$0.6484_{\pm 0.0003}$	$0.5556_{\pm 0.0005}$	$+14.32_{\pm 0.04}$
H=96	DUET	V:G_3 (3/3)	$0.4580_{\pm 0.0034}$	$0.3999_{\pm 0.0028}$	$+12.67_{\pm 0.04}$
H=96	FACT	V:G_3 (3/3)	$0.4042_{\pm 0.0018}$	$0.3878_{\pm 0.0011}$	$+4.06_{\pm 0.14}$
H=96	FITS	T:$G_5(3)$ (3/3)	$0.6540_{\pm 0.0003}$	$0.5504_{\pm 0.0003}$	$+15.83_{\pm 0.02}$
H=96	MixLinear	T:$G_5(3)$ (3/3)	$0.6795_{\pm 0.0053}$	$0.5676_{\pm 0.0073}$	$+16.47_{\pm 0.75}$
H=96	PaiFilter	T:$G_5(3)$ (3/3)	$0.5462_{\pm 0.0010}$	$0.4983_{\pm 0.0006}$	$+8.77_{\pm 0.08}$
H=96	PatchTST	T:$G_5(3)$ (3/3)	$0.4465_{\pm 0.0012}$	$0.4229_{\pm 0.0010}$	$+5.29_{\pm 0.05}$
H=96	TimeMixer	T:G_4 (3/3)	$0.4937_{\pm 0.0049}$	$0.4657_{\pm 0.0041}$	$+5.66_{\pm 0.12}$
H=96	TimeXer	V:G_3 (3/3)	$0.4450_{\pm 0.0018}$	$0.4238_{\pm 0.0017}$	$+4.77_{\pm 0.10}$
H=96	iTransformer	T:$G_5(3)$ (3/3)	$0.3931_{\pm 0.0007}$	$0.3784_{\pm 0.0007}$	$+3.75_{\pm 0.04}$
H=192	DLinear	T:G_4 (3/3)	$0.5984_{\pm 0.0001}$	$0.5318_{\pm 0.0001}$	$+11.13_{\pm 0.01}$
H=192	DUET	V:G_3 (3/3)	$0.4850_{\pm 0.0012}$	$0.4300_{\pm 0.0011}$	$+11.33_{\pm 0.02}$
H=192	FACT	V:G_3 (3/3)	$0.4233_{\pm 0.0031}$	$0.4095_{\pm 0.0017}$	$+3.24_{\pm 0.29}$
H=192	FITS	T:G_4 (3/3)	$0.6052_{\pm 0.0005}$	$0.5235_{\pm 0.0002}$	$+13.50_{\pm 0.04}$
H=192	MixLinear	T:G_4 (3/3)	$0.6604_{\pm 0.0057}$	$0.5412_{\pm 0.0015}$	$+18.05_{\pm 0.50}$
H=192	PaiFilter	T:G_4 (3/3)	$0.5318_{\pm 0.0003}$	$0.4894_{\pm 0.0007}$	$+7.98_{\pm 0.13}$
H=192	PatchTST	T:$G_5(3)$ (3/3)	$0.4556_{\pm 0.0006}$	$0.4323_{\pm 0.0004}$	$+5.12_{\pm 0.06}$
H=192	TimeMixer	T:G_4 (3/3)	$0.4985_{\pm 0.0058}$	$0.4709_{\pm 0.0036}$	$+5.54_{\pm 0.40}$
H=192	TimeXer	V:G_3 (3/3)	$0.4547_{\pm 0.0004}$	$0.4378_{\pm 0.0004}$	$+3.73_{\pm 0.05}$
H=192	iTransformer	V:G_3 (3/3)	$0.4124_{\pm 0.0004}$	$0.4057_{\pm 0.0002}$	$+1.61_{\pm 0.06}$

Table 16: Full-matrix output-side δ-Adapter results at H=96. Each cell first averages paired seed-level reductions over seeds $2/3/4$. Group effects are mean $\pm$ sample SD across the listed cells (ten backbones per dataset or eight datasets per backbone); positive is better.

By dataset

group	improves	mean $\pm$ SD ↑	worst ↑
ETTh1	1/10	-0.84 ± 0.85	-3.03
ETTh2	6/10	-0.13 ± 0.55	-1.53
ETTm1	10/10	$+1.89 \pm 0.95$	$+0.10$
ETTm2	5/10	$+0.26 \pm 1.63$	-2.35
Weather	10/10	$+5.51 \pm 1.71$	$+2.39$
ECL	10/10	$+5.49 \pm 3.45$	$+0.10$
Solar	10/10	$+9.87 \pm 3.57$	$+3.79$
Traffic	10/10	$+3.33 \pm 1.61$	$+0.60$

By backbone

group	improves	mean $\pm$ SD ↑	worst ↑
DLinear	6/8	$+4.18 \pm 5.68$	-3.03
DUET	6/8	$+4.43 \pm 5.35$	-0.98
FACT	5/8	$+0.47 \pm 1.88$	-2.35
FITS	7/8	$+4.28 \pm 4.39$	-0.32
MixLinear	7/8	$+4.45 \pm 4.43$	-0.72
PaiFilter	6/8	$+3.71 \pm 4.01$	-0.75
PatchTST	6/8	$+3.32 \pm 3.70$	-0.92
TimeMixer	7/8	$+3.08 \pm 3.55$	-0.94
TimeXer	6/8	$+1.88 \pm 2.79$	-1.08
iTransformer	6/8	$+1.92 \pm 2.30$	-0.80

Table 14: Complete Traffic results, part 2 of 2. Together the two parts cover ten backbones, four horizons, and seeds $2/3/4$. The selected column lists the independently chosen fitting source (T: training, V: validation) and covariance class. Every MSE and reduction entry is mean $\pm$ sample SD over the three paired seeds; reductions are percentages (↑).

horizon	backbone	selected at seeds $2/3/4$	baseline MSE	corrected MSE	reduction (%) ↑
H=336	DLinear	T:$G_5(3)$ (3/3)	$0.6051_{\pm 0.0000}$	$0.5335_{\pm 0.0003}$	$+11.83_{\pm 0.05}$
H=336	DUET	V:G_3 (3/3)	$0.5035_{\pm 0.0006}$	$0.4475_{\pm 0.0009}$	$+11.14_{\pm 0.26}$
H=336	FACT	V:G_3 (3/3)	$0.4460_{\pm 0.0040}$	$0.4306_{\pm 0.0027}$	$+3.46_{\pm 0.26}$
H=336	FITS	T:G_4 (3/3)	$0.6116_{\pm 0.0002}$	$0.5272_{\pm 0.0001}$	$+13.81_{\pm 0.02}$
H=336	MixLinear	T:G_4/T:$G_5(3)$/T:G_4	$0.6427_{\pm 0.0017}$	$0.5319_{\pm 0.0014}$	$+17.25_{\pm 0.08}$
H=336	PaiFilter	T:$G_5(3)$/T:G_4/T:G_4	$0.5436_{\pm 0.0010}$	$0.4985_{\pm 0.0008}$	$+8.29_{\pm 0.07}$
H=336	PatchTST	V:G_3 (3/3)	$0.4720_{\pm 0.0007}$	$0.4541_{\pm 0.0005}$	$+3.79_{\pm 0.05}$
H=336	TimeMixer	T:G_4 (3/3)	$0.5137_{\pm 0.0031}$	$0.4839_{\pm 0.0008}$	$+5.79_{\pm 0.45}$
H=336	TimeXer	V:G_3 (3/3)	$0.4721_{\pm 0.0029}$	$0.4543_{\pm 0.0013}$	$+3.78_{\pm 0.32}$
H=336	iTransformer	V:G_3 (3/3)	$0.4231_{\pm 0.0020}$	$0.4180_{\pm 0.0020}$	$+1.20_{\pm 0.02}$
H=720	DLinear	V:G_3 (3/3)	$0.6456_{\pm 0.0001}$	$0.6279_{\pm 0.0001}$	$+2.74_{\pm 0.02}$
H=720	DUET	V:G_3 (3/3)	$0.5377_{\pm 0.0017}$	$0.4911_{\pm 0.0006}$	$+8.67_{\pm 0.23}$
H=720	FACT	V:G_3 (3/3)	$0.4837_{\pm 0.0028}$	$0.4713_{\pm 0.0012}$	$+2.57_{\pm 0.34}$
H=720	FITS	V:G_3 (3/3)	$0.6494_{\pm 0.0001}$	$0.6358_{\pm 0.0001}$	$+2.09_{\pm 0.01}$
H=720	MixLinear	T:$G_5(3)$ (3/3)	$0.7368_{\pm 0.0031}$	$0.5933_{\pm 0.0048}$	$+19.48_{\pm 0.47}$
H=720	PaiFilter	V:G_3 (3/3)	$0.5813_{\pm 0.0022}$	$0.5821_{\pm 0.0018}$	$-0.15_{\pm 0.17}$
H=720	PatchTST	V:G_3 (3/3)	$0.5064_{\pm 0.0010}$	$0.4934_{\pm 0.0005}$	$+2.58_{\pm 0.19}$
H=720	TimeMixer	V:G_3 (3/3)	$0.5490_{\pm 0.0074}$	$0.5537_{\pm 0.0013}$	$-0.86_{\pm 1.17}$
H=720	TimeXer	V:G_3 (3/3)	$0.5109_{\pm 0.0013}$	$0.5468_{\pm 0.0041}$	$-7.04_{\pm 1.06}$
H=720	iTransformer	V:G_3 (3/3)	$0.4586_{\pm 0.0003}$	$0.4585_{\pm 0.0003}$	$+0.03_{\pm 0.05}$

Table 15: Complete published-method comparison at H=96, averaging seeds $2/3/4$ over all 80 cells. A *post-hoc* method edits a frozen forecast, a *training-loss* method changes optimization, and a *combined* method applies our selector to that peer's cached forecast. Percentage columns use the matched plain-training forecast as baseline and report relative test-MSE reduction (↑); SD is computed across the 80 cell means.

interface	method	cells	improves ↑	mean $\pm$ SD ↑	median ↑	worst ↑
post-hoc	ours	80	$\mathbf{77/80}$	$\mathbf{+7.61_{\pm 8.10\%}}$	$+3.91\%$	$\mathbf{-0.73\%}$
post-hoc	δ-Adapter	80	$62/80$	$+3.17_{\pm 3.97\%}$	$+2.46\%$	-3.03%
combined	δ-Adapter + ours	80	$71/80$	$\mathbf{+8.13_{\pm 9.55\%}}$	$\mathbf{+4.34\%}$	-3.80%
post-hoc	UEC-STD	80	$67/80$	$+2.16_{\pm 3.83\%}$	$+0.53\%$	-0.52%
combined	UEC-STD + ours	80	$\mathbf{76/80}$	$\mathbf{+6.57_{\pm 6.67\%}}$	$\mathbf{+3.78\%}$	$\mathbf{-0.49\%}$
post-hoc	Socket+Plug	80	$75/80$	$+6.41_{\pm 7.10\%}$	$+3.23\%$	-1.53%
combined	Socket+Plug + ours	80	$\mathbf{80/80}$	$\mathbf{+10.34_{\pm 10.77\%}}$	$\mathbf{+4.96\%}$	$\mathbf{+0.08\%}$
training loss	FreDF	80	$39/80$	$-11.72_{\pm 33.79\%}$	-0.10%	-195.90%
combined	FreDF + ours	80	$\mathbf{66/80}$	$\mathbf{+6.25_{\pm 7.91\%}}$	$\mathbf{+4.87\%}$	$\mathbf{-9.90\%}$
training loss	PS Loss	80	$40/80$	$-0.80_{\pm 5.07\%}$	$+0.04\%$	-19.30%
combined	PS Loss + ours	80	$\mathbf{77/80}$	$\mathbf{+7.74_{\pm 7.36\%}}$	$\mathbf{+5.79\%}$	$\mathbf{-3.53\%}$

Table 17 applies the two stages to the same 80 cells. Incremental changes are measured from the fitted δ-Adapter output, and total changes from the plain forecast.

Table 17: Serial composition of output-side δ-Adapter and ours on the complete $H{=}96$ matrix. Dataset entries are mean $\pm$ sample SD across ten backbone-cell means after seed averaging. Incremental reduction is relative to the fitted δ-Adapter output; total reduction is relative to the plain forecast. SD is shown as a subscript.

dataset	plain MSE	δ MSE	δ+ours MSE	incr. wins	incr. ↑	total ↑
ETTh1	$0.3952_{\pm 0.0177}$	$0.3986_{\pm 0.0186}$	$0.3988_{\pm 0.0213}$	4/10	$-0.03_{\pm 1.10}$	$-0.88_{\pm 1.51}$
ETTh2	$0.3033_{\pm 0.0177}$	$0.3038_{\pm 0.0192}$	$0.2997_{\pm 0.0186}$	10/10	$+1.34_{\pm 0.72}$	$+1.20_{\pm 0.74}$
ETTm1	$0.3413_{\pm 0.0155}$	$0.3348_{\pm 0.0135}$	$0.3276_{\pm 0.0118}$	10/10	$+2.14_{\pm 0.81}$	$+3.98_{\pm 1.41}$
ETTm2	$0.1813_{\pm 0.0085}$	$0.1808_{\pm 0.0064}$	$0.1777_{\pm 0.0062}$	10/10	$+1.71_{\pm 0.47}$	$+1.94_{\pm 1.50}$
Weather	$0.1800_{\pm 0.0165}$	$0.1698_{\pm 0.0127}$	$0.1621_{\pm 0.0071}$	8/10	$+4.31_{\pm 4.30}$	$+9.54_{\pm 5.47}$
ECL	$0.1745_{\pm 0.0263}$	$0.1645_{\pm 0.0228}$	$0.1517_{\pm 0.0174}$	10/10	$+7.49_{\pm 3.65}$	$+12.47_{\pm 6.38}$
Solar	$0.2648_{\pm 0.0585}$	$0.2372_{\pm 0.0457}$	$0.1901_{\pm 0.0208}$	10/10	$+18.72_{\pm 6.96}$	$+26.56_{\pm 8.96}$
Traffic	$0.5169_{\pm 0.1083}$	$0.4982_{\pm 0.0972}$	$0.4589_{\pm 0.0685}$	10/10	$+7.16_{\pm 4.47}$	$+10.21_{\pm 5.71}$

UEC-STD (Hoang Nguyen et al., 2026) uses a seasonal–trend error corrector with hidden width 32, kernel size 25, dropout 0.5, and SmoothL1 loss. Its correction strength is selected from $\{0.01, 0.03, 0.05, 0.07, 0.1, 0.3, 0.5, 0.7, 1.0\}$ using validation data. Socket+Plug (Wang et al., 2025a) uses its variable-wise H–256–H GELU Plug and validation-based checkpoint selection. Both methods run independently for seeds $2/3/4$; our second stage uses their cached calibration and test forecasts.

Table 18: Complete 80-cell $H{=}96$ composition with UEC-STD. Each backbone has paired MSE/MAE columns. Rows compare plain, UEC-STD (peer), and peer+ours. Entries are mean $\pm$ sample SD over seeds $2/3/4$, with smaller SD and leading zeros omitted. Δ is incremental MSE reduction (%, $\uparrow$) against the same cached peer forecast. Bold marks the lowest error among the three forecasts.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	.3963$_{\pm.0004}$	**.4110**$_{\pm.0004}$	.3822$_{\pm.0018}$	**.3998**$_{\pm.0012}$	.3821$_{\pm.0029}$	**.4014**$_{\pm.0025}$	.3910$_{\pm.0030}$	**.4012**$_{\pm.0031}$	.3928$_{\pm.0029}$	**.3990**$_{\pm.0027}$
	peer	**.3958**$_{\pm.0014}$	.4128$_{\pm.0013}$	**.3789**$_{\pm.0013}$	.4062$_{\pm.0034}$	.3792$_{\pm.0028}$	.4023$_{\pm.0007}$	.3890$_{\pm.0016}$	.4102$_{\pm.0055}$	.3909$_{\pm.0064}$	.4026$_{\pm.0060}$
	+ours	.3966$_{\pm.0003}$	.4165$_{\pm.0005}$	.3801$_{\pm.0022}$	.4073$_{\pm.0038}$	**.3737**$_{\pm.0017}$	.4026$_{\pm.0025}$	**.3875**$_{\pm.0005}$	.4102$_{\pm.0034}$	**.3904**$_{\pm.0050}$	.4051$_{\pm.0067}$
	$\Delta\uparrow$ (%)	-0.20$_{\pm.33}$		-0.32$_{\pm.24}$		+1.45$_{\pm.67}$		+0.39$_{\pm.28}$		+0.12$_{\pm.48}$	
ETTh2	plain	.3458$_{\pm.0061}$	.3997$_{\pm.0041}$	.2926$_{\pm.0009}$	.3435$_{\pm.0010}$	.2936$_{\pm.0039}$	.3421$_{\pm.0030}$	.2992$_{\pm.0009}$	.3459$_{\pm.0007}$	.3001$_{\pm.0039}$	.3478$_{\pm.0037}$
	peer	.3440$_{\pm.0055}$	.3982$_{\pm.0037}$	.2928$_{\pm.0005}$	.3443$_{\pm.0007}$	.2941$_{\pm.0043}$	.3432$_{\pm.0034}$	.2990$_{\pm.0006}$	.3462$_{\pm.0002}$	.3005$_{\pm.0023}$	.3484$_{\pm.0024}$
	+ours	**.3073**$_{\pm.0019}$	**.3658**$_{\pm.0018}$	**.2898**$_{\pm.0015}$	**.3426**$_{\pm.0009}$	**.2872**$_{\pm.0018}$	**.3379**$_{\pm.0015}$	**.2967**$_{\pm.0002}$	**.3449**$_{\pm.0008}$	**.2969**$_{\pm.0013}$	**.3454**$_{\pm.0015}$
	$\Delta\uparrow$ (%)	+10.65$_{\pm.87}$		+1.02$_{\pm.44}$		+2.36$_{\pm.86}$		+0.77$_{\pm.15}$		+1.22$_{\pm.35}$	
ETTm1	plain	.3456$_{\pm.0002}$	.3721$_{\pm.0001}$	.3325$_{\pm.0026}$	.3684$_{\pm.0008}$	.3211$_{\pm.0042}$	.3606$_{\pm.0023}$	.3547$_{\pm.0003}$	.3751$_{\pm.0003}$	.3695$_{\pm.0029}$	.3848$_{\pm.0019}$
	peer	.3329$_{\pm.0020}$	**.3680**$_{\pm.0041}$	.3169$_{\pm.0028}$	.3609$_{\pm.0029}$	.3188$_{\pm.0032}$	.3586$_{\pm.0014}$	.3399$_{\pm.0043}$	.3763$_{\pm.0023}$	.3476$_{\pm.0024}$	.3803$_{\pm.0080}$
	+ours	**.3278**$_{\pm.0010}$	.3688$_{\pm.0042}$	**.3121**$_{\pm.0037}$	**.3594**$_{\pm.0032}$	**.3105**$_{\pm.0010}$	**.3571**$_{\pm.0016}$	**.3320**$_{\pm.0021}$	**.3730**$_{\pm.0039}$	**.3380**$_{\pm.0039}$	**.3769**$_{\pm.0090}$
	$\Delta\uparrow$ (%)	+1.54$_{\pm.30}$		+1.54$_{\pm.42}$		+2.58$_{\pm1.08}$		+2.30$_{\pm1.06}$		+2.74$_{\pm.50}$	
ETTm2	plain	.1926$_{\pm.0025}$	.2912$_{\pm.0039}$	.1780$_{\pm.0029}$	.2602$_{\pm.0031}$	.1676$_{\pm.0004}$	.2504$_{\pm.0004}$	.1831$_{\pm.0002}$	.2663$_{\pm.0000}$	.1887$_{\pm.0014}$	.2718$_{\pm.0016}$
	peer	.1889$_{\pm.0052}$	.2872$_{\pm.0068}$	.1771$_{\pm.0023}$	.2596$_{\pm.0027}$	.1667$_{\pm.0003}$	.2504$_{\pm.0005}$	.1819$_{\pm.0005}$	.2656$_{\pm.0004}$	.1878$_{\pm.0015}$	.2712$_{\pm.0017}$
	+ours	**.1876**$_{\pm.0052}$	**.2855**$_{\pm.0068}$	**.1754**$_{\pm.0018}$	**.2587**$_{\pm.0014}$	**.1644**$_{\pm.0003}$	**.2485**$_{\pm.0002}$	**.1793**$_{\pm.0004}$	**.2627**$_{\pm.0003}$	**.1836**$_{\pm.0015}$	**.2672**$_{\pm.0017}$
	$\Delta\uparrow$ (%)	+0.66$_{\pm.10}$		+0.93$_{\pm.26}$		+1.42$_{\pm.11}$		+1.43$_{\pm.06}$		+2.21$_{\pm.08}$	
Weather	plain	.1958$_{\pm.0003}$	.2560$_{\pm.0011}$	.1671$_{\pm.0001}$	**.2131**$_{\pm.0004}$	.1566$_{\pm.0024}$	.2006$_{\pm.0025}$	.1969$_{\pm.0001}$	.2371$_{\pm.0001}$	.1994$_{\pm.0010}$	**.2393**$_{\pm.0005}$
	peer	.1956$_{\pm.0002}$	.2555$_{\pm.0008}$	.1670$_{\pm.0005}$	.2141$_{\pm.0010}$	.1562$_{\pm.0022}$	**.2002**$_{\pm.0022}$	.1963$_{\pm.0002}$	**.2370**$_{\pm.0002}$	.1984$_{\pm.0015}$	.2399$_{\pm.0016}$
	+ours	**.1624**$_{\pm.0007}$	**.2436**$_{\pm.0009}$	**.1627**$_{\pm.0005}$	.2388$_{\pm.0002}$	**.1523**$_{\pm.0046}$	.2158$_{\pm.0072}$	**.1653**$_{\pm.0004}$	.2480$_{\pm.0005}$	**.1662**$_{\pm.0007}$	.2493$_{\pm.0012}$
	$\Delta\uparrow$ (%)	+16.97$_{\pm.27}$		+2.60$_{\pm.05}$		+2.56$_{\pm1.74}$		+15.80$_{\pm.11}$		+16.24$_{\pm.32}$	
ECL	plain	.1946$_{\pm.0001}$	.2774$_{\pm.0006}$	.1668$_{\pm.0002}$	.2846$_{\pm.0002}$	.1372$_{\pm.0004}$	**.2336**$_{\pm.0005}$	.2033$_{\pm.0002}$	.2835$_{\pm.0004}$	.2172$_{\pm.0191}$	.2972$_{\pm.0268}$
	peer	.1867$_{\pm.0032}$	.2759$_{\pm.0013}$	.1574$_{\pm.0042}$	.2638$_{\pm.0056}$	.1377$_{\pm.0006}$	.2344$_{\pm.0010}$	.1962$_{\pm.0042}$	.2832$_{\pm.0038}$	.2161$_{\pm.0234}$	.3032$_{\pm.0326}$
	+ours	**.1714**$_{\pm.0051}$	**.2671**$_{\pm.0039}$	**.1440**$_{\pm.0062}$	**.2456**$_{\pm.0074}$	**.1367**$_{\pm.0005}$	.2350$_{\pm.0011}$	**.1722**$_{\pm.0055}$	**.2715**$_{\pm.0039}$	**.1774**$_{\pm.0039}$	**.2773**$_{\pm.0063}$
	$\Delta\uparrow$ (%)	+8.19$_{\pm3.75}$		+8.54$_{\pm2.40}$		+0.75$_{\pm.36}$		+12.25$_{\pm1.34}$		+17.38$_{\pm7.43}$	
Solar	plain	.2900$_{\pm.0001}$	.3775$_{\pm.0002}$	.2345$_{\pm.0099}$	.3070$_{\pm.0029}$	.1939$_{\pm.0022}$	.2410$_{\pm.0017}$	.3415$_{\pm.0012}$	.3699$_{\pm.0010}$	.3646$_{\pm.0146}$	.3734$_{\pm.0078}$
	peer	.2584$_{\pm.0031}$	.3217$_{\pm.0012}$	.2063$_{\pm.0136}$	.2568$_{\pm.0076}$	.1941$_{\pm.0030}$	.2417$_{\pm.0026}$	.2847$_{\pm.0033}$	.3362$_{\pm.0067}$	.2904$_{\pm.0241}$	.3316$_{\pm.0099}$
	+ours	**.2308**$_{\pm.0197}$	**.2993**$_{\pm.0037}$	**.2047**$_{\pm.0131}$	**.2541**$_{\pm.0050}$	**.1861**$_{\pm.0081}$	**.2372**$_{\pm.0023}$	**.2465**$_{\pm.0196}$	**.3101**$_{\pm.0053}$	**.2462**$_{\pm.0234}$	**.3081**$_{\pm.0039}$
	$\Delta\uparrow$ (%)	+10.73$_{\pm6.92}$		+0.77$_{\pm.30}$		+4.06$_{\pm5.49}$		+13.36$_{\pm7.62}$		+15.03$_{\pm7.51}$	
Traffic	plain	.6484$_{\pm.0003}$	.3961$_{\pm.0004}$	.4580$_{\pm.0034}$	.3142$_{\pm.0006}$	.4042$_{\pm.0018}$	**.2593**$_{\pm.0010}$	.6540$_{\pm.0003}$	.3926$_{\pm.0008}$	.6795$_{\pm.0053}$	.4101$_{\pm.0032}$
	peer	.5974$_{\pm.0063}$	.3870$_{\pm.0045}$	.4491$_{\pm.0050}$	.2998$_{\pm.0118}$	.4043$_{\pm.0018}$	.2595$_{\pm.0011}$	.6262$_{\pm.0465}$	.3837$_{\pm.0148}$	.6011$_{\pm.0167}$	.3840$_{\pm.0077}$
	+ours	**.5605**$_{\pm.0079}$	**.3700**$_{\pm.0027}$	**.4023**$_{\pm.0046}$	**.2780**$_{\pm.0066}$	**.3878**$_{\pm.0012}$	.2617$_{\pm.0011}$	**.5571**$_{\pm.0107}$	**.3704**$_{\pm.0047}$	**.5524**$_{\pm.0069}$	**.3607**$_{\pm.0069}$
	$\Delta\uparrow$ (%)	+6.18$_{\pm.35}$		+10.39$_{\pm1.97}$		+4.07$_{\pm.13}$		+10.60$_{\pm8.72}$		+8.05$_{\pm3.12}$	

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	.4409$_{\pm.0041}$	**.4408**$_{\pm.0027}$	**.3833**$_{\pm.0007}$	**.3982**$_{\pm.0006}$	.4067$_{\pm.0023}$	**.4161**$_{\pm.0022}$	.3869$_{\pm.0011}$	**.4045**$_{\pm.0006}$	.3898$_{\pm.0021}$	**.4076**$_{\pm.0017}$
	peer	**.4379**$_{\pm.0063}$	.4502$_{\pm.0082}$	.3853$_{\pm.0059}$	.4042$_{\pm.0012}$	.4053$_{\pm.0068}$	.4202$_{\pm.0102}$	.3866$_{\pm.0037}$	.4114$_{\pm.0074}$	.3868$_{\pm.0008}$	.4129$_{\pm.0006}$
	+ours	.4430$_{\pm.0068}$	.4560$_{\pm.0081}$	.3852$_{\pm.0053}$	.4057$_{\pm.0005}$	**.3997**$_{\pm.0053}$	.4199$_{\pm.0075}$	**.3865**$_{\pm.0043}$	.4126$_{\pm.0069}$	**.3832**$_{\pm.0019}$	.4131$_{\pm.0017}$
	$\Delta\uparrow$ (%)	-1.16$_{\pm.09}$		+0.02$_{\pm.36}$		+1.36$_{\pm.68}$		+0.03$_{\pm.37}$		+0.94$_{\pm.42}$	
ETTh2	plain	.3190$_{\pm.0149}$	.3650$_{\pm.0128}$	.2871$_{\pm.0011}$	.3399$_{\pm.0006}$	.2969$_{\pm.0043}$	.3471$_{\pm.0032}$	.2899$_{\pm.0023}$	.3419$_{\pm.0022}$	.3091$_{\pm.0083}$	.3547$_{\pm.0041}$
	peer	.3141$_{\pm.0108}$	.3605$_{\pm.0091}$	.2869$_{\pm.0010}$	.3406$_{\pm.0008}$	.2965$_{\pm.0042}$	.3469$_{\pm.0029}$	.2906$_{\pm.0023}$	.3439$_{\pm.0013}$	.3085$_{\pm.0085}$	.3544$_{\pm.0042}$
	+ours	**.3095**$_{\pm.0072}$	**.3569**$_{\pm.0050}$	**.2838**$_{\pm.0010}$	**.3381**$_{\pm.0010}$	**.2926**$_{\pm.0033}$	**.3439**$_{\pm.0022}$	**.2862**$_{\pm.0015}$	**.3404**$_{\pm.0006}$	**.3017**$_{\pm.0082}$	**.3485**$_{\pm.0042}$
	$\Delta\uparrow$ (%)	+1.44$_{\pm1.60}$		+1.06$_{\pm.13}$		+1.31$_{\pm.28}$		+1.51$_{\pm.24}$		+2.21$_{\pm.45}$	
ETTm1	plain	.3569$_{\pm.0040}$	.3797$_{\pm.0007}$	.3337$_{\pm.0038}$	**.3690**$_{\pm.0003}$	.3339$_{\pm.0038}$	**.3712**$_{\pm.0014}$	.3233$_{\pm.0005}$	.3604$_{\pm.0010}$	.3426$_{\pm.0047}$	.3752$_{\pm.0037}$
	peer	.3432$_{\pm.0071}$	**.3780**$_{\pm.0012}$	.3280$_{\pm.0073}$	.3699$_{\pm.0004}$	.3293$_{\pm.0016}$	.3764$_{\pm.0018}$	.3171$_{\pm.0022}$	.3599$_{\pm.0007}$	.3358$_{\pm.0048}$	.3745$_{\pm.0047}$
	+ours	**.3407**$_{\pm.0063}$	.3794$_{\pm.0011}$	**.3217**$_{\pm.0042}$	.3690$_{\pm.0010}$	**.3237**$_{\pm.0015}$	.3755$_{\pm.0013}$	**.3126**$_{\pm.0003}$	**.3592**$_{\pm.0027}$	**.3282**$_{\pm.0044}$	**.3731**$_{\pm.0046}$
	$\Delta\uparrow$ (%)	+0.71$_{\pm.25}$		+1.91$_{\pm.90}$		+1.72$_{\pm.28}$		+1.42$_{\pm.62}$		+2.25$_{\pm.43}$	
ETTm2	plain	.1930$_{\pm.0007}$	.2772$_{\pm.0008}$	.1756$_{\pm.0006}$	.2604$_{\pm.0009}$	.1775$_{\pm.0018}$	.2623$_{\pm.0021}$	.1726$_{\pm.0002}$	.2570$_{\pm.0011}$	.1848$_{\pm.0026}$	.2669$_{\pm.0025}$
	peer	.1909$_{\pm.0026}$	.2755$_{\pm.0021}$	.1750$_{\pm.0006}$	.2600$_{\pm.0008}$	.1760$_{\pm.0024}$	.2617$_{\pm.0021}$	.1718$_{\pm.0004}$	.2569$_{\pm.0012}$	.1841$_{\pm.0027}$	.2665$_{\pm.0025}$
	+ours	**.1882**$_{\pm.0024}$	**.2727**$_{\pm.0019}$	**.1729**$_{\pm.0008}$	**.2597**$_{\pm.0010}$	**.1726**$_{\pm.0019}$	**.2599**$_{\pm.0019}$	**.1702**$_{\pm.0011}$	**.2563**$_{\pm.0009}$	**.1800**$_{\pm.0033}$	**.2638**$_{\pm.0031}$
	$\Delta\uparrow$ (%)	+1.41$_{\pm.11}$		+1.23$_{\pm.18}$		+1.90$_{\pm.27}$		+0.94$_{\pm.47}$		+2.25$_{\pm.32}$	
Weather	plain	.1975$_{\pm.0024}$	.2405$_{\pm.0037}$	.1768$_{\pm.0011}$	.2180$_{\pm.0007}$	.1781$_{\pm.0023}$	**.2240**$_{\pm.0023}$	.1587$_{\pm.0007}$	.2075$_{\pm.0005}$	.1732$_{\pm.0011}$	.2132$_{\pm.0009}$
	peer	.1968$_{\pm.0023}$	**.2400**$_{\pm.0036}$	.1758$_{\pm.0010}$	**.2174**$_{\pm.0003}$	.1770$_{\pm.0018}$	.2247$_{\pm.0037}$	**.1582**$_{\pm.0007}$	**.2070**$_{\pm.0005}$	.1728$_{\pm.0010}$	**.2127**$_{\pm.0006}$
	+ours	**.1690**$_{\pm.0048}$	.2512$_{\pm.0058}$	**.1577**$_{\pm.0001}$	.2362$_{\pm.0006}$	**.1638**$_{\pm.0062}$	.2436$_{\pm.0054}$	.1589$_{\pm.0009}$	.2331$_{\pm.0058}$	**.1593**$_{\pm.0019}$	.2347$_{\pm.0032}$
	$\Delta\uparrow$ (%)	+14.17$_{\pm1.44}$		+10.33$_{\pm.53}$		+7.44$_{\pm3.37}$		-0.39$_{\pm.12}$		+7.78$_{\pm.57}$	
ECL	plain	.1864$_{\pm.0008}$	.2738$_{\pm.0015}$	.1768$_{\pm.0004}$	.2710$_{\pm.0006}$	.1698$_{\pm.0006}$	.2609$_{\pm.0006}$	.1450$_{\pm.0004}$	.2436$_{\pm.0003}$	.1481$_{\pm.0003}$	.2400$_{\pm.0003}$
	peer	.1849$_{\pm.0000}$	.2751$_{\pm.0027}$	.1754$_{\pm.0014}$	.2686$_{\pm.0021}$	.1698$_{\pm.0010}$	.2637$_{\pm.0029}$	.1454$_{\pm.0003}$	.2433$_{\pm.0003}$	.1475$_{\pm.0004}$	.2394$_{\pm.0006}$
	+ours	**.1626**$_{\pm.0029}$	**.2635**$_{\pm.0005}$	**.1545**$_{\pm.0102}$	**.2526**$_{\pm.0077}$	**.1557**$_{\pm.0037}$	**.2546**$_{\pm.0018}$	**.1373**$_{\pm.0008}$	**.2347**$_{\pm.0010}$	**.1381**$_{\pm.0005}$	**.2369**$_{\pm.0004}$
	$\Delta\uparrow$ (%)	+12.05$_{\pm1.58}$		+11.94$_{\pm5.14}$		+8.31$_{\pm2.04}$		+5.54$_{\pm.63}$		+6.39$_{\pm.09}$	
Solar	plain	.3088$_{\pm.0015}$	.3160$_{\pm.0058}$	.2398$_{\pm.0033}$	.2885$_{\pm.0019}$	.2506$_{\pm.0024}$	.2939$_{\pm.0021}$	.2123$_{\pm.0027}$	.2709$_{\pm.0027}$	.2117$_{\pm.0041}$	.2545$_{\pm.0023}$
	peer	.2777$_{\pm.0039}$	.3037$_{\pm.0079}$	.2275$_{\pm.0052}$	**.2635**$_{\pm.0019}$	.2457$_{\pm.0015}$	.2830$_{\pm.0039}$	.2061$_{\pm.0014}$	**.2499**$_{\pm.0008}$	.1987$_{\pm.0036}$	.2500$_{\pm.0045}$
	+ours	**.2599**$_{\pm.0045}$	**.2950**$_{\pm.0049}$	**.2257**$_{\pm.0018}$	.2638$_{\pm.0025}$	**.2426**$_{\pm.0033}$	**.2789**$_{\pm.0016}$	**.2059**$_{\pm.0015}$	.2501$_{\pm.0008}$	**.1936**$_{\pm.0027}$	**.2477**$_{\pm.0021}$
	$\Delta\uparrow$ (%)	+6.42$_{\pm.84}$		+0.79$_{\pm1.61}$		+1.26$_{\pm1.60}$		+0.09$_{\pm1.02}$		+2.51$_{\pm2.71}$	
Traffic	plain	.5462$_{\pm.0010}$	.3555$_{\pm.0015}$	.4465$_{\pm.0012}$	.2908$_{\pm.0005}$	.4937$_{\pm.0049}$	.3253$_{\pm.0015}$	.4450$_{\pm.0018}$	.2929$_{\pm.0011}$	.3931$_{\pm.0007}$	.2684$_{\pm.0006}$
	peer	.5449$_{\pm.0033}$	.3591$_{\pm.0041}$	.4464$_{\pm.0011}$	.2912$_{\pm.0009}$	.4938$_{\pm.0049}$	.3269$_{\pm.0038}$	.4461$_{\pm.0023}$	.2949$_{\pm.0024}$	.3936$_{\pm.0013}$	.2703$_{\pm.0037}$
	+ours	**.5057**$_{\pm.0077}$	**.3436**$_{\pm.0035}$	**.4234**$_{\pm.0017}$	**.2812**$_{\pm.0063}$	**.4688**$_{\pm.0061}$	**.3201**$_{\pm.0088}$	**.4242**$_{\pm.0018}$	**.2899**$_{\pm.0040}$	**.3812**$_{\pm.0054}$	**.2670**$_{\pm.0076}$
	$\Delta\uparrow$ (%)	+7.20$_{\pm1.50}$		+5.16$_{\pm.26}$		+5.06$_{\pm1.04}$		+4.92$_{\pm.10}$		+3.15$_{\pm1.08}$	

Table 19: Complete 80-cell $H{=}96$ composition with Socket+Plug. Each backbone has paired MSE/MAE columns. Rows compare plain, Socket+Plug (peer), and peer+ours. Entries are mean $\pm$ sample SD over seeds $2/3/4$, with smaller SD and leading zeros omitted. Δ is incremental MSE reduction (%, $\uparrow$) against the same cached peer forecast. Bold marks the lowest error among the three forecasts.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	.3963$_{\pm.0004}$	.4110$_{\pm.0004}$	.3822$_{\pm.0018}$	.3997$_{\pm.0012}$	.3821$_{\pm.0029}$	.4014$_{\pm.0025}$	.3910$_{\pm.0030}$	.4012$_{\pm.0031}$	.3928$_{\pm.0029}$	.3990$_{\pm.0027}$
	peer	**.3860**$_{\pm.0004}$	**.4015**$_{\pm.0006}$	.3745$_{\pm.0010}$	**.3937**$_{\pm.0002}$	**.3759**$_{\pm.0039}$	**.4004**$_{\pm.0053}$	.3828$_{\pm.0011}$	**.3946**$_{\pm.0005}$	**.3810**$_{\pm.0007}$	**.3951**$_{\pm.0012}$
	+ours	.3893$_{\pm.0004}$	.4081$_{\pm.0004}$	**.3733**$_{\pm.0009}$	.3969$_{\pm.0003}$	.3759$_{\pm.0033}$	.4040$_{\pm.0047}$	**.3824**$_{\pm.0011}$	.3994$_{\pm.0005}$	.3812$_{\pm.0017}$	.3997$_{\pm.0003}$
	$\Delta\uparrow$ (%)	-0.85$_{\pm.01}$		+0.33$_{\pm.03}$		-0.01$_{\pm.15}$		+0.12$_{\pm.08}$		-0.06$_{\pm.29}$	
ETTh2	plain	.3458$_{\pm.0061}$	.3997$_{\pm.0041}$	.2926$_{\pm.0009}$	.3435$_{\pm.0010}$	.2936$_{\pm.0039}$	.3421$_{\pm.0030}$	.2992$_{\pm.0009}$	.3459$_{\pm.0007}$	.3001$_{\pm.0039}$	.3478$_{\pm.0037}$
	peer	.3318$_{\pm.0069}$	**.3842**$_{\pm.0040}$	.2905$_{\pm.0012}$	.3404$_{\pm.0022}$	.2873$_{\pm.0017}$	.3353$_{\pm.0007}$	.2893$_{\pm.0007}$	.3394$_{\pm.0004}$	.2909$_{\pm.0021}$	.3401$_{\pm.0013}$
	+ours	**.3310**$_{\pm.0064}$	.3846$_{\pm.0040}$	**.2890**$_{\pm.0008}$	**.3394**$_{\pm.0020}$	**.2854**$_{\pm.0018}$	**.3337**$_{\pm.0006}$	**.2879**$_{\pm.0011}$	**.3384**$_{\pm.0006}$	**.2901**$_{\pm.0021}$	**.3397**$_{\pm.0012}$
	$\Delta\uparrow$ (%)	+0.24$_{\pm.15}$		+0.54$_{\pm.17}$		+0.66$_{\pm.04}$		+0.46$_{\pm.30}$		+0.26$_{\pm.01}$	
ETTm1	plain	.3456$_{\pm.0002}$	.3721$_{\pm.0001}$	.3324$_{\pm.0027}$	.3684$_{\pm.0007}$	.3211$_{\pm.0042}$	.3606$_{\pm.0023}$	.3547$_{\pm.0003}$	.3751$_{\pm.0003}$	.3695$_{\pm.0029}$	.3848$_{\pm.0019}$
	peer	.3253$_{\pm.0004}$	**.3619**$_{\pm.0007}$	.3269$_{\pm.0020}$	**.3586**$_{\pm.0002}$	.3150$_{\pm.0057}$	**.3549**$_{\pm.0039}$	.3318$_{\pm.0001}$	**.3657**$_{\pm.0004}$	.3367$_{\pm.0034}$	**.3694**$_{\pm.0022}$
	+ours	**.3224**$_{\pm.0003}$	.3642$_{\pm.0034}$	**.3201**$_{\pm.0030}$	.3611$_{\pm.0008}$	**.3082**$_{\pm.0070}$	.3574$_{\pm.0036}$	**.3286**$_{\pm.0015}$	.3695$_{\pm.0021}$	**.3315**$_{\pm.0060}$	.3716$_{\pm.0081}$
	$\Delta\uparrow$ (%)	+0.88$_{\pm.14}$		+2.06$_{\pm.35}$		+2.18$_{\pm.60}$		+0.97$_{\pm.45}$		+1.54$_{\pm.80}$	
ETTm2	plain	.1926$_{\pm.0025}$	.2912$_{\pm.0039}$	.1780$_{\pm.0029}$	.2602$_{\pm.0031}$	.1676$_{\pm.0004}$	.2504$_{\pm.0004}$	.1831$_{\pm.0002}$	.2663$_{\pm.0000}$	.1887$_{\pm.0014}$	.2718$_{\pm.0016}$
	peer	.1893$_{\pm.0054}$	**.2796**$_{\pm.0058}$	.1731$_{\pm.0011}$	**.2521**$_{\pm.0013}$	.1655$_{\pm.0005}$	.2471$_{\pm.0008}$	.1766$_{\pm.0006}$	**.2559**$_{\pm.0007}$	.1770$_{\pm.0008}$	**.2564**$_{\pm.0008}$
	+ours	**.1889**$_{\pm.0054}$	.2799$_{\pm.0059}$	**.1720**$_{\pm.0011}$	.2528$_{\pm.0028}$	**.1646**$_{\pm.0005}$	**.2468**$_{\pm.0006}$	**.1756**$_{\pm.0005}$	.2593$_{\pm.0010}$	**.1759**$_{\pm.0008}$	.2598$_{\pm.0008}$
	$\Delta\uparrow$ (%)	+0.21$_{\pm.05}$		+0.68$_{\pm.13}$		+0.56$_{\pm.02}$		+0.55$_{\pm.04}$		+0.67$_{\pm.04}$	
Weather	plain	.1958$_{\pm.0003}$	.2560$_{\pm.0011}$	.1671$_{\pm.0001}$	.2131$_{\pm.0004}$	.1566$_{\pm.0024}$	.2006$_{\pm.0025}$	.1969$_{\pm.0001}$	.2371$_{\pm.0001}$	.1994$_{\pm.0010}$	.2393$_{\pm.0005}$
	peer	.1645$_{\pm.0010}$	**.2228**$_{\pm.0023}$	.1627$_{\pm.0006}$	**.2078**$_{\pm.0009}$	.1562$_{\pm.0012}$	**.1989**$_{\pm.0011}$	.1644$_{\pm.0003}$	**.2081**$_{\pm.0003}$	.1635$_{\pm.0003}$	**.2077**$_{\pm.0002}$
	+ours	**.1599**$_{\pm.0008}$	.2393$_{\pm.0010}$	**.1598**$_{\pm.0006}$	.2358$_{\pm.0007}$	**.1515**$_{\pm.0034}$	.2141$_{\pm.0047}$	**.1574**$_{\pm.0004}$	.2358$_{\pm.0007}$	**.1569**$_{\pm.0009}$	.2352$_{\pm.0010}$
	$\Delta\uparrow$ (%)	+2.77$_{\pm.71}$		+1.81$_{\pm.55}$		+3.00$_{\pm1.50}$		+4.26$_{\pm.20}$		+4.04$_{\pm.47}$	
ECL	plain	.1946$_{\pm.0001}$	.2774$_{\pm.0006}$	.1668$_{\pm.0002}$	.2846$_{\pm.0002}$	.1372$_{\pm.0004}$	**.2336**$_{\pm.0005}$	.2033$_{\pm.0002}$	.2835$_{\pm.0004}$	.2172$_{\pm.0191}$	.2972$_{\pm.0268}$
	peer	.1569$_{\pm.0001}$	.2535$_{\pm.0002}$	.1343$_{\pm.0001}$	.2306$_{\pm.0001}$	.1371$_{\pm.0003}$	.2340$_{\pm.0003}$	.1605$_{\pm.0001}$	.2555$_{\pm.0001}$	.1633$_{\pm.0025}$	.2573$_{\pm.0037}$
	+ours	**.1456**$_{\pm.0009}$	**.2464**$_{\pm.0010}$	**.1304**$_{\pm.0002}$	**.2292**$_{\pm.0001}$	**.1365**$_{\pm.0003}$	.2362$_{\pm.0004}$	**.1469**$_{\pm.0001}$	**.2490**$_{\pm.0007}$	**.1488**$_{\pm.0020}$	**.2506**$_{\pm.0036}$
	$\Delta\uparrow$ (%)	+7.19$_{\pm.65}$		+2.86$_{\pm.03}$		+0.41$_{\pm.09}$		+8.48$_{\pm.10}$		+8.89$_{\pm.16}$	
Solar	plain	.2900$_{\pm.0001}$	.3775$_{\pm.0002}$	.2345$_{\pm.0099}$	.3070$_{\pm.0029}$	.1939$_{\pm.0022}$	.2410$_{\pm.0017}$	.3415$_{\pm.0012}$	.3699$_{\pm.0010}$	.3646$_{\pm.0146}$	.3734$_{\pm.0078}$
	peer	.2371$_{\pm.0003}$	.3021$_{\pm.0002}$	.2080$_{\pm.0073}$	.2517$_{\pm.0012}$	.1915$_{\pm.0019}$	**.2330**$_{\pm.0016}$	.2497$_{\pm.0008}$	.2780$_{\pm.0005}$	.2479$_{\pm.0076}$	.2763$_{\pm.0047}$
	+ours	**.1866**$_{\pm.0003}$	**.2746**$_{\pm.0003}$	**.1703**$_{\pm.0020}$	**.2427**$_{\pm.0034}$	**.1742**$_{\pm.0013}$	.2334$_{\pm.0021}$	**.1862**$_{\pm.0008}$	**.2656**$_{\pm.0005}$	**.1839**$_{\pm.0026}$	**.2622**$_{\pm.0045}$
	$\Delta\uparrow$ (%)	+21.31$_{\pm.11}$		+18.10$_{\pm2.00}$		+9.06$_{\pm.23}$		+25.46$_{\pm.10}$		+25.79$_{\pm1.29}$	
Traffic	plain	.6484$_{\pm.0003}$	.3961$_{\pm.0004}$	.4580$_{\pm.0034}$	.3142$_{\pm.0006}$	.4042$_{\pm.0018}$	.2593$_{\pm.0010}$	.6540$_{\pm.0003}$	.3926$_{\pm.0008}$	.6795$_{\pm.0053}$	.4101$_{\pm.0032}$
	peer	.5487$_{\pm.0007}$	.3396$_{\pm.0003}$	.4185$_{\pm.0034}$	**.2646**$_{\pm.0004}$	.4103$_{\pm.0013}$	**.2589**$_{\pm.0007}$	.5441$_{\pm.0008}$	.3385$_{\pm.0001}$	.5536$_{\pm.0052}$	.3424$_{\pm.0014}$
	+ours	**.5198**$_{\pm.0006}$	**.3364**$_{\pm.0003}$	**.4030**$_{\pm.0026}$	.2668$_{\pm.0005}$	**.3961**$_{\pm.0009}$	.2610$_{\pm.0007}$	**.5138**$_{\pm.0004}$	**.3250**$_{\pm.0001}$	**.5214**$_{\pm.0038}$	**.3287**$_{\pm.0037}$
	$\Delta\uparrow$ (%)	+5.27$_{\pm.02}$		+3.72$_{\pm.17}$		+3.46$_{\pm.10}$		+5.57$_{\pm.05}$		+5.82$_{\pm.68}$	

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	.4409$_{\pm.0041}$	.4408$_{\pm.0027}$	.3833$_{\pm.0007}$	.3982$_{\pm.0006}$	.4067$_{\pm.0023}$	.4161$_{\pm.0022}$	.3869$_{\pm.0011}$	**.4045**$_{\pm.0006}$	.3898$_{\pm.0021}$	.4076$_{\pm.0017}$
	peer	**.4027**$_{\pm.0024}$	**.4107**$_{\pm.0015}$	**.3780**$_{\pm.0015}$	**.3972**$_{\pm.0006}$	**.3818**$_{\pm.0016}$	**.3998**$_{\pm.0004}$	.3882$_{\pm.0006}$	.4069$_{\pm.0007}$	.3795$_{\pm.0022}$	**.4020**$_{\pm.0017}$
	+ours	.4076$_{\pm.0040}$	.4190$_{\pm.0033}$	.3803$_{\pm.0017}$	.4029$_{\pm.0006}$	.3831$_{\pm.0010}$	.4048$_{\pm.0002}$	**.3858**$_{\pm.0008}$	.4121$_{\pm.0023}$	**.3771**$_{\pm.0011}$	.4051$_{\pm.0020}$
	$\Delta\uparrow$ (%)	-1.21$_{\pm.51}$		-0.61$_{\pm.07}$		-0.35$_{\pm.16}$		+0.62$_{\pm.31}$		+0.64$_{\pm.62}$	
ETTh2	plain	.3190$_{\pm.0149}$	.3650$_{\pm.0128}$	.2871$_{\pm.0011}$	.3399$_{\pm.0006}$	.2969$_{\pm.0043}$	.3471$_{\pm.0032}$	.2899$_{\pm.0023}$	.3419$_{\pm.0022}$	.3091$_{\pm.0083}$	.3547$_{\pm.0041}$
	peer	.3124$_{\pm.0129}$	.3576$_{\pm.0109}$	.2836$_{\pm.0003}$	.3365$_{\pm.0009}$	.2891$_{\pm.0005}$	.3392$_{\pm.0005}$	.2910$_{\pm.0015}$	.3398$_{\pm.0014}$	.3003$_{\pm.0112}$	.3458$_{\pm.0061}$
	+ours	**.3073**$_{\pm.0089}$	**.3538**$_{\pm.0076}$	**.2813**$_{\pm.0002}$	**.3344**$_{\pm.0005}$	**.2885**$_{\pm.0005}$	**.3388**$_{\pm.0003}$	**.2891**$_{\pm.0013}$	**.3377**$_{\pm.0012}$	**.2979**$_{\pm.0106}$	**.3437**$_{\pm.0058}$
	$\Delta\uparrow$ (%)	+1.62$_{\pm1.31}$		+0.82$_{\pm.16}$		+0.21$_{\pm.04}$		+0.68$_{\pm.07}$		+0.82$_{\pm.16}$	
ETTm1	plain	.3569$_{\pm.0040}$	.3797$_{\pm.0007}$	.3337$_{\pm.0038}$	.3690$_{\pm.0003}$	.3339$_{\pm.0038}$	.3712$_{\pm.0014}$	.3233$_{\pm.0005}$	.3604$_{\pm.0010}$	.3426$_{\pm.0047}$	.3752$_{\pm.0037}$
	peer	.3490$_{\pm.0024}$	**.3781**$_{\pm.0014}$	.3220$_{\pm.0029}$	**.3611**$_{\pm.0016}$	.3234$_{\pm.0040}$	**.3634**$_{\pm.0026}$	.3185$_{\pm.0017}$	**.3569**$_{\pm.0015}$	.3361$_{\pm.0052}$	**.3696**$_{\pm.0024}$
	+ours	**.3445**$_{\pm.0028}$	.3820$_{\pm.0028}$	**.3166**$_{\pm.0039}$	.3612$_{\pm.0049}$	**.3193**$_{\pm.0031}$	.3674$_{\pm.0014}$	**.3143**$_{\pm.0016}$	.3576$_{\pm.0057}$	**.3263**$_{\pm.0033}$	.3701$_{\pm.0010}$
	$\Delta\uparrow$ (%)	+1.31$_{\pm.11}$		+1.67$_{\pm.54}$		+1.29$_{\pm.25}$		+1.31$_{\pm.56}$		+2.90$_{\pm.57}$	
ETTm2	plain	.1930$_{\pm.0007}$	.2772$_{\pm.0008}$	.1756$_{\pm.0006}$	.2604$_{\pm.0009}$	.1775$_{\pm.0018}$	.2623$_{\pm.0021}$	.1726$_{\pm.0002}$	.2570$_{\pm.0011}$	.1848$_{\pm.0026}$	.2669$_{\pm.0025}$
	peer	.1829$_{\pm.0003}$	**.2632**$_{\pm.0007}$	.1712$_{\pm.0011}$	**.2533**$_{\pm.0014}$	.1705$_{\pm.0007}$	**.2527**$_{\pm.0003}$	.1702$_{\pm.0024}$	**.2526**$_{\pm.0015}$	.1769$_{\pm.0014}$	**.2561**$_{\pm.0014}$
	+ours	**.1817**$_{\pm.0002}$	.2660$_{\pm.0006}$	**.1700**$_{\pm.0011}$	.2536$_{\pm.0013}$	**.1700**$_{\pm.0008}$	.2558$_{\pm.0005}$	**.1693**$_{\pm.0023}$	.2531$_{\pm.0012}$	**.1759**$_{\pm.0015}$	.2563$_{\pm.0015}$
	$\Delta\uparrow$ (%)	+0.66$_{\pm.07}$		+0.66$_{\pm.06}$		+0.30$_{\pm.13}$		+0.57$_{\pm.06}$		+0.58$_{\pm.03}$	
Weather	plain	.1975$_{\pm.0024}$	.2405$_{\pm.0037}$	.1768$_{\pm.0011}$	.2180$_{\pm.0007}$	.1781$_{\pm.0023}$	.2240$_{\pm.0023}$	.1587$_{\pm.0007}$	.2075$_{\pm.0005}$	.1732$_{\pm.0011}$	.2132$_{\pm.0009}$
	peer	.1728$_{\pm.0020}$	**.2206**$_{\pm.0040}$	.1625$_{\pm.0007}$	**.2061**$_{\pm.0003}$	.1646$_{\pm.0013}$	**.2097**$_{\pm.0007}$	.1588$_{\pm.0008}$	**.2065**$_{\pm.0006}$	.1613$_{\pm.0007}$	**.2032**$_{\pm.0005}$
	+ours	**.1672**$_{\pm.0007}$	.2415$_{\pm.0063}$	**.1576**$_{\pm.0006}$	.2362$_{\pm.0006}$	**.1606**$_{\pm.0038}$	.2385$_{\pm.0054}$	**.1585**$_{\pm.0015}$	.2352$_{\pm.0018}$	**.1564**$_{\pm.0010}$	.2303$_{\pm.0015}$
	$\Delta\uparrow$ (%)	+3.23$_{\pm1.30}$		+2.97$_{\pm.57}$		+2.41$_{\pm1.55}$		+0.14$_{\pm.46}$		+3.05$_{\pm.34}$	
ECL	plain	.1864$_{\pm.0008}$	.2738$_{\pm.0015}$	.1768$_{\pm.0004}$	.2710$_{\pm.0006}$	.1698$_{\pm.0006}$	.2609$_{\pm.0006}$	.1450$_{\pm.0004}$	.2436$_{\pm.0003}$	.1481$_{\pm.0003}$	.2400$_{\pm.0003}$
	peer	.1602$_{\pm.0004}$	.2550$_{\pm.0006}$	.1562$_{\pm.0002}$	.2493$_{\pm.0003}$	.1552$_{\pm.0002}$	.2497$_{\pm.0002}$	.1380$_{\pm.0003}$	.2345$_{\pm.0001}$	.1380$_{\pm.0002}$	.2343$_{\pm.0004}$
	+ours	**.1467**$_{\pm.0004}$	**.2493**$_{\pm.0006}$	**.1440**$_{\pm.0001}$	**.2452**$_{\pm.0001}$	**.1430**$_{\pm.0002}$	**.2447**$_{\pm.0010}$	**.1335**$_{\pm.0002}$	**.2329**$_{\pm.0001}$	**.1341**$_{\pm.0003}$	**.2335**$_{\pm.0002}$
	$\Delta\uparrow$ (%)	+8.48$_{\pm.04}$		+7.77$_{\pm.08}$		+7.89$_{\pm.02}$		+3.20$_{\pm.07}$		+2.80$_{\pm.04}$	
Solar	plain	.3088$_{\pm.0015}$	.3160$_{\pm.0058}$	.2398$_{\pm.0033}$	.2885$_{\pm.0019}$	.2506$_{\pm.0024}$	.2939$_{\pm.0021}$	.2123$_{\pm.0027}$	.2709$_{\pm.0027}$	.2117$_{\pm.0041}$	.2545$_{\pm.0023}$
	peer	.2858$_{\pm.0034}$	.2924$_{\pm.0020}$	.2364$_{\pm.0020}$	.2692$_{\pm.0016}$	.2469$_{\pm.0009}$	.2770$_{\pm.0004}$	.2105$_{\pm.0024}$	.2555$_{\pm.0024}$	.2053$_{\pm.0047}$	.2480$_{\pm.0046}$
	+ours	**.2061**$_{\pm.0029}$	**.2855**$_{\pm.0019}$	**.1853**$_{\pm.0009}$	**.2646**$_{\pm.0015}$	**.1883**$_{\pm.0012}$	**.2679**$_{\pm.0005}$	**.1777**$_{\pm.0028}$	**.2518**$_{\pm.0022}$	**.1756**$_{\pm.0058}$	**.2441**$_{\pm.0051}$
	$\Delta\uparrow$ (%)	+27.89$_{\pm.27}$		+21.63$_{\pm.50}$		+23.73$_{\pm.45}$		+15.57$_{\pm.51}$		+14.47$_{\pm.95}$	
Traffic	plain	.5462$_{\pm.0010}$	.3555$_{\pm.0015}$	.4465$_{\pm.0012}$	.2908$_{\pm.0005}$	.4937$_{\pm.0049}$	.3253$_{\pm.0015}$	.4450$_{\pm.0018}$	.2929$_{\pm.0011}$	.3931$_{\pm.0007}$	.2684$_{\pm.0006}$
	peer	.5078$_{\pm.0019}$	.3220$_{\pm.0008}$	.4426$_{\pm.0006}$	**.2815**$_{\pm.0004}$	.4773$_{\pm.0027}$	**.3032**$_{\pm.0011}$	.4401$_{\pm.0021}$	**.2769**$_{\pm.0006}$	.3959$_{\pm.0005}$	**.2645**$_{\pm.0005}$
	+ours	**.4843**$_{\pm.0020}$	**.3119**$_{\pm.0020}$	**.4312**$_{\pm.0006}$	.2827$_{\pm.0004}$	**.4611**$_{\pm.0021}$	.3043$_{\pm.0009}$	**.4262**$_{\pm.0017}$	.2787$_{\pm.0006}$	**.3888**$_{\pm.0006}$	.2661$_{\pm.0005}$
	$\Delta\uparrow$ (%)	+4.63$_{\pm.24}$		+2.58$_{\pm.08}$		+3.39$_{\pm.18}$		+3.17$_{\pm.08}$		+1.82$_{\pm.03}$	

Table 20 places the same post-hoc interface beside training-time peers on the 24-cell matched subset. For the complete peer matrices, FreDF uses α=0.8 and PS Loss uses λ_{PS}=3.0, their published-default strengths. Each peer is trained with the same dataset split, backbone, and seed as its plain checkpoint. We cache that peer forecast once and use the identical realization both as the reported input and as the input to our fitted second stage.

Table 20: Compact raw-MSE comparison on a 24-cell subset at H=96. Every entry is mean ± sample SD over paired seeds 2/3/4. Columns identify dataset/backbone cases. Ours selects resolution, ridge penalty, and fitting source by cross-fitted validation error. AdaY is the output-side δ-Adapter (Liang et al., 2026); FreDF and PS Loss alter training, and each +ours entry refits our selector on the exact cached peer forecast. Bold marks the minimum across methods for each case. Smaller type gives SD; leading zeros are omitted.

method	ECL DLinear	ECL PatchTST	ECL iTransformer	ETTh1 DLinear	ETTh1 PatchTST	ETTh1 TimeMixer	ETTh1 iTransformer	ETTh2 DLinear
base	.1946$_{\pm.0001}$	.1768$_{\pm.0004}$	.1481$_{\pm.0003}$	.3963$_{\pm.0004}$	.3833$_{\pm.0007}$	.4067$_{\pm.0023}$	.3898$_{\pm.0021}$	.3458$_{\pm.0061}$
ours	.1735$_{\pm.0000}$	**.1495**$_{\pm.0003}$	.1386$_{\pm.0004}$	.3976$_{\pm.0003}$	**.3803**$_{\pm.0011}$	.3964$_{\pm.0027}$	.3813$_{\pm.0015}$	.3117$_{\pm.0016}$
AdaY	.1831$_{\pm.0001}$	.1633$_{\pm.0002}$	.1452$_{\pm.0007}$	.4083$_{\pm.0006}$	.3869$_{\pm.0003}$	.4105$_{\pm.0016}$	.3908$_{\pm.0025}$	.3511$_{\pm.0072}$
FreDF	.1957$_{\pm.0000}$	.3208$_{\pm.0014}$	.1466$_{\pm.0003}$	.3914$_{\pm.0005}$	.4077$_{\pm.0019}$	.4218$_{\pm.0180}$	.3771$_{\pm.0007}$	.2997$_{\pm.0008}$
FreDF+ours	.1741$_{\pm.0001}$	.1613$_{\pm.0020}$	**.1364**$_{\pm.0002}$	.3943$_{\pm.0004}$	.3970$_{\pm.0026}$	.4196$_{\pm.0178}$	**.3759**$_{\pm.0008}$	.2992$_{\pm.0006}$
PS Loss	.1961$_{\pm.0000}$	.1803$_{\pm.0004}$	.1467$_{\pm.0001}$	**.3911**$_{\pm.0003}$	.3844$_{\pm.0005}$	.3966$_{\pm.0005}$	.3797$_{\pm.0014}$	.2989$_{\pm.0005}$
PS+ours	**.1735**$_{\pm.0001}$	.1521$_{\pm.0010}$	.1366$_{\pm.0001}$	.3927$_{\pm.0004}$	.3832$_{\pm.0001}$	**.3915**$_{\pm.0012}$	.3760$_{\pm.0012}$	**.2984**$_{\pm.0002}$

method	ETTh2 PatchTST	ETTh2 TimeMixer	ETTh2 iTransformer	ETTm1 DLinear	ETTm1 PatchTST	ETTm1 TimeMixer	ETTm1 iTransformer	ETTm2 DLinear
base	.2871$_{\pm.0011}$	.2969$_{\pm.0043}$	.3091$_{\pm.0083}$	.3456$_{\pm.0002}$	.3337$_{\pm.0038}$	.3339$_{\pm.0038}$	.3426$_{\pm.0047}$	.1926$_{\pm.0025}$
ours	.2845$_{\pm.0012}$	.2934$_{\pm.0031}$	.3023$_{\pm.0079}$	.3302$_{\pm.0003}$	.3250$_{\pm.0074}$	.3253$_{\pm.0035}$	.3301$_{\pm.0034}$	.1915$_{\pm.0024}$
AdaY	.2868$_{\pm.0015}$	.2965$_{\pm.0045}$	.3089$_{\pm.0073}$	.3400$_{\pm.0004}$	.3254$_{\pm.0039}$	.3298$_{\pm.0036}$	.3340$_{\pm.0041}$	.1863$_{\pm.0029}$
FreDF	.2935$_{\pm.0035}$	.2995$_{\pm.0066}$	.2891$_{\pm.0010}$	.3425$_{\pm.0001}$	.4342$_{\pm.0032}$	.3293$_{\pm.0007}$	.3343$_{\pm.0059}$	.1846$_{\pm.0002}$
FreDF+ours	.2917$_{\pm.0039}$	.2958$_{\pm.0035}$	**.2862**$_{\pm.0012}$	**.3291**$_{\pm.0000}$	.3306$_{\pm.0010}$	**.3226**$_{\pm.0005}$	.3257$_{\pm.0074}$	.1833$_{\pm.0003}$
PS Loss	.2804$_{\pm.0002}$	.2924$_{\pm.0020}$	.2944$_{\pm.0038}$	.3507$_{\pm.0002}$	.3359$_{\pm.0028}$	.3414$_{\pm.0013}$	.3304$_{\pm.0022}$	.1822$_{\pm.0003}$
PS+ours	**.2794**$_{\pm.0003}$	**.2913**$_{\pm.0016}$	.2907$_{\pm.0041}$	.3309$_{\pm.0004}$	**.3185**$_{\pm.0024}$	.3317$_{\pm.0011}$	**.3204**$_{\pm.0032}$	**.1810**$_{\pm.0003}$

method	ETTm2 iTransformer	Solar DLinear	Solar iTransformer	Traffic DLinear	Traffic PatchTST	Traffic iTransformer	Weather DLinear	Weather iTransformer
base	.1848$_{\pm.0026}$	.2900$_{\pm.0001}$	.2117$_{\pm.0041}$	.6484$_{\pm.0003}$	.4465$_{\pm.0012}$	.3931$_{\pm.0007}$	.1958$_{\pm.0003}$	.1732$_{\pm.0011}$
ours	.1806$_{\pm.0032}$	.2178$_{\pm.0001}$	.1770$_{\pm.0062}$	.5556$_{\pm.0005}$	**.4229**$_{\pm.0010}$	.3784$_{\pm.0007}$	.1628$_{\pm.0006}$	.1591$_{\pm.0015}$
AdaY	.1863$_{\pm.0018}$	.2459$_{\pm.0000}$	.2001$_{\pm.0048}$	.6174$_{\pm.0002}$	.4338$_{\pm.0012}$	.3870$_{\pm.0007}$	.1811$_{\pm.0010}$	.1648$_{\pm.0017}$
FreDF	.1793$_{\pm.0007}$	.3024$_{\pm.0001}$	.2153$_{\pm.0089}$	.6562$_{\pm.0003}$	.5139$_{\pm.0012}$	.3906$_{\pm.0004}$	.1956$_{\pm.0000}$	.1719$_{\pm.0027}$
FreDF+ours	.1754$_{\pm.0004}$	.2240$_{\pm.0001}$	**.1725**$_{\pm.0056}$	.5612$_{\pm.0003}$	.4353$_{\pm.0024}$	**.3767**$_{\pm.0003}$	.1642$_{\pm.0001}$	.1574$_{\pm.0024}$
PS Loss	.1760$_{\pm.0013}$	.2849$_{\pm.0005}$	.2236$_{\pm.0126}$	.6456$_{\pm.0001}$	.4536$_{\pm.0000}$	.3934$_{\pm.0010}$	.1981$_{\pm.0012}$	.1690$_{\pm.0007}$
PS+ours	**.1728**$_{\pm.0007}$	**.2168**$_{\pm.0003}$	.1769$_{\pm.0060}$	**.5536**$_{\pm.0002}$	.4271$_{\pm.0002}$	.3770$_{\pm.0007}$	**.1626**$_{\pm.0022}$	**.1529**$_{\pm.0011}$

Within this subset, a column containing our correction attains the lowest MSE on 23/24 cells. The complete second-stage experiments improve 72/80 δ-Adapter outputs, 76/80 FreDF outputs, and 77/80 PS Loss outputs. The following tables report MSE and MAE for every FreDF and PS Loss cell; their incremental percentages pair the corrected output with the exact cached input used to fit it.

Table 21: Per-cell serial composition on all 80 $H{=}96$ dataset–backbone cells. Entries are mean $\pm$ sample SD over seeds $2/3/4$, with smaller SD and leading zeros omitted. The first three rows give MSE (↓); the last three give relative reductions (%, ↑). δ and total reductions use plain forecasts as baseline; incremental reduction uses δ-Adapter outputs. iTransf. denotes iTransformer.

dataset	statistic	DLinear	DUET	FACT	FITS	MixLinear	PaiFilter	PatchTST	TimeMixer	TimeXer	iTransf.
ETTh1	plain	.3963$_{\pm.0004}$	.3821$_{\pm.0019}$	.3821$_{\pm.0029}$	.3910$_{\pm.0030}$	.3928$_{\pm.0029}$	.4409$_{\pm.0041}$	.3833$_{\pm.0007}$	.4067$_{\pm.0023}$	.3869$_{\pm.0011}$	.3898$_{\pm.0021}$
	δ	.4083$_{\pm.0006}$	.3818$_{\pm.0015}$	.3844$_{\pm.0014}$	.3922$_{\pm.0017}$	.3956$_{\pm.0026}$	.4443$_{\pm.0042}$	.3869$_{\pm.0003}$	.4105$_{\pm.0016}$	.3908$_{\pm.0003}$	.3908$_{\pm.0025}$
	δ+ours	.4114$_{\pm.0007}$	.3849$_{\pm.0010}$	.3797$_{\pm.0023}$	.3930$_{\pm.0011}$	.3980$_{\pm.0034}$	.4528$_{\pm.0043}$	.3877$_{\pm.0005}$	.4048$_{\pm.0020}$	.3902$_{\pm.0005}$	.3853$_{\pm.0008}$
	Δ_{δ}	-3.03$_{\pm.17}$	+0.10$_{\pm.25}$	-0.61$_{\pm.39}$	-0.32$_{\pm.47}$	-0.72$_{\pm.09}$	-0.75$_{\pm.17}$	-0.92$_{\pm.24}$	-0.94$_{\pm.20}$	-1.00$_{\pm.31}$	-0.25$_{\pm.11}$
	$\Delta_{\text{incr.}}$	-0.75$_{\pm.03}$	-0.81$_{\pm.15}$	+1.23$_{\pm.63}$	-0.22$_{\pm.15}$	-0.59$_{\pm.39}$	-1.92$_{\pm.03}$	-0.22$_{\pm.10}$	+1.39$_{\pm.63}$	+0.16$_{\pm.12}$	+1.40$_{\pm.53}$
	Δ_{total}	-3.80$_{\pm.20}$	-0.71$_{\pm.37}$	+0.63$_{\pm.92}$	-0.53$_{\pm.57}$	-1.32$_{\pm.42}$	-2.69$_{\pm.20}$	-1.15$_{\pm.32}$	+0.46$_{\pm.78}$	-0.84$_{\pm.41}$	+1.16$_{\pm.43}$
ETTh2	plain	.3458$_{\pm.0061}$	.2925$_{\pm.0009}$	.2936$_{\pm.0039}$	.2992$_{\pm.0009}$	.3001$_{\pm.0039}$	.3190$_{\pm.0149}$	.2871$_{\pm.0011}$	.2969$_{\pm.0043}$	.2899$_{\pm.0023}$	.3091$_{\pm.0083}$
	δ	.3511$_{\pm.0072}$	.2941$_{\pm.0021}$	.2944$_{\pm.0034}$	.2983$_{\pm.0004}$	.2998$_{\pm.0044}$	.3198$_{\pm.0149}$	.2868$_{\pm.0015}$	.2965$_{\pm.0045}$	.2887$_{\pm.0013}$	.3089$_{\pm.0073}$
	δ+ours	.3437$_{\pm.0060}$	.2907$_{\pm.0020}$	.2873$_{\pm.0017}$	.2952$_{\pm.0013}$	.2968$_{\pm.0031}$	.3198$_{\pm.0150}$	.2828$_{\pm.0012}$	.2929$_{\pm.0031}$	.2860$_{\pm.0008}$	.3022$_{\pm.0078}$
	Δ_{δ}	-1.53$_{\pm.41}$	-0.54$_{\pm.41}$	-0.27$_{\pm.27}$	+0.30$_{\pm.26}$	+0.12$_{\pm.18}$	-0.25$_{\pm.04}$	+0.08$_{\pm.14}$	+0.13$_{\pm.09}$	+0.41$_{\pm.54}$	+0.06$_{\pm.39}$
	$\Delta_{\text{incr.}}$	+2.11$_{\pm.32}$	+1.14$_{\pm.17}$	+2.42$_{\pm.55}$	+1.04$_{\pm.51}$	+0.98$_{\pm.44}$	+0.01$_{\pm.15}$	+1.39$_{\pm.15}$	+1.21$_{\pm.51}$	+0.95$_{\pm.21}$	+2.18$_{\pm.31}$
	Δ_{total}	+0.61$_{\pm.42}$	+0.61$_{\pm.38}$	+2.16$_{\pm.77}$	+1.34$_{\pm.42}$	+1.10$_{\pm.27}$	-0.24$_{\pm.11}$	+1.47$_{\pm.09}$	+1.35$_{\pm.46}$	+1.35$_{\pm.54}$	+2.25$_{\pm.08}$
ETTm1	plain	.3456$_{\pm.0002}$	.3321$_{\pm.0024}$	.3211$_{\pm.0042}$	.3547$_{\pm.0003}$	.3695$_{\pm.0029}$	.3569$_{\pm.0040}$	.3337$_{\pm.0038}$	.3339$_{\pm.0038}$	.3233$_{\pm.0005}$	.3426$_{\pm.0047}$
	δ	.3400$_{\pm.0004}$	.3229$_{\pm.0018}$	.3208$_{\pm.0049}$	.3450$_{\pm.0003}$	.3586$_{\pm.0023}$	.3514$_{\pm.0036}$	.3254$_{\pm.0039}$	.3298$_{\pm.0036}$	.3204$_{\pm.0013}$	.3340$_{\pm.0041}$
	δ+ours	.3304$_{\pm.0011}$	.3173$_{\pm.0012}$	.3128$_{\pm.0036}$	.3362$_{\pm.0000}$	.3460$_{\pm.0020}$	.3460$_{\pm.0030}$	.3206$_{\pm.0030}$	.3234$_{\pm.0037}$	.3181$_{\pm.0020}$	.3252$_{\pm.0030}$
	Δ_{δ}	+1.62$_{\pm.14}$	+2.77$_{\pm.21}$	+0.10$_{\pm.43}$	+2.72$_{\pm.01}$	+2.94$_{\pm.12}$	+1.56$_{\pm.35}$	+2.47$_{\pm.08}$	+1.23$_{\pm.14}$	+0.90$_{\pm.33}$	+2.51$_{\pm.31}$
	$\Delta_{\text{incr.}}$	+2.83$_{\pm.26}$	+1.74$_{\pm.19}$	+2.48$_{\pm.89}$	+2.56$_{\pm.09}$	+3.52$_{\pm.07}$	+1.52$_{\pm.14}$	+1.48$_{\pm.31}$	+1.93$_{\pm.45}$	+0.71$_{\pm.30}$	+2.63$_{\pm.37}$
	Δ_{total}	+4.41$_{\pm.37}$	+4.46$_{\pm.36}$	+2.58$_{\pm.47}$	+5.21$_{\pm.08}$	+6.36$_{\pm.18}$	+3.05$_{\pm.43}$	+3.92$_{\pm.30}$	+3.14$_{\pm.58}$	+1.61$_{\pm.61}$	+5.07$_{\pm.61}$
ETTm2	plain	.1926$_{\pm.0025}$	.1780$_{\pm.0029}$	.1676$_{\pm.0004}$	.1831$_{\pm.0002}$	.1887$_{\pm.0014}$	.1930$_{\pm.0007}$	.1756$_{\pm.0006}$	.1775$_{\pm.0018}$	.1726$_{\pm.0002}$	.1848$_{\pm.0026}$
	δ	.1863$_{\pm.0029}$	.1796$_{\pm.0017}$	.1715$_{\pm.0004}$	.1809$_{\pm.0003}$	.1866$_{\pm.0012}$	.1905$_{\pm.0004}$	.1757$_{\pm.0009}$	.1763$_{\pm.0020}$	.1743$_{\pm.0024}$	.1863$_{\pm.0018}$
	δ+ours	.1846$_{\pm.0035}$	.1759$_{\pm.0014}$	.1687$_{\pm.0003}$	.1774$_{\pm.0004}$	.1820$_{\pm.0014}$	.1879$_{\pm.0004}$	.1733$_{\pm.0012}$	.1731$_{\pm.0012}$	.1721$_{\pm.0030}$	.1822$_{\pm.0025}$
	Δ_{δ}	+3.25$_{\pm.28}$	-0.90$_{\pm.85}$	-2.35$_{\pm.49}$	+1.16$_{\pm.13}$	+1.10$_{\pm.16}$	+1.28$_{\pm.32}$	-0.07$_{\pm.62}$	+0.67$_{\pm.13}$	-1.01$_{\pm1.28}$	-0.80$_{\pm.85}$
	$\Delta_{\text{incr.}}$	+0.90$_{\pm.35}$	+2.01$_{\pm.13}$	+1.63$_{\pm.09}$	+1.96$_{\pm.12}$	+2.48$_{\pm.14}$	+1.40$_{\pm.03}$	+1.39$_{\pm.28}$	+1.81$_{\pm.66}$	+1.30$_{\pm.39}$	+2.20$_{\pm.46}$
	Δ_{total}	+4.12$_{\pm.61}$	+1.14$_{\pm.91}$	-0.68$_{\pm.44}$	+3.10$_{\pm.24}$	+3.56$_{\pm.15}$	+2.66$_{\pm.32}$	+1.33$_{\pm.61}$	+2.46$_{\pm.54}$	+0.30$_{\pm1.58}$	+1.42$_{\pm.54}$
Weather	plain	.1958$_{\pm.0003}$	.1671$_{\pm.0001}$	.1566$_{\pm.0024}$	.1969$_{\pm.0001}$	.1994$_{\pm.0010}$	.1975$_{\pm.0024}$	.1768$_{\pm.0011}$	.1781$_{\pm.0023}$	.1587$_{\pm.0007}$	.1732$_{\pm.0011}$
	δ	.1811$_{\pm.0010}$	.1596$_{\pm.0007}$	.1527$_{\pm.0020}$	.1828$_{\pm.0006}$	.1851$_{\pm.0014}$	.1843$_{\pm.0030}$	.1666$_{\pm.0012}$	.1675$_{\pm.0016}$	.1534$_{\pm.0004}$	.1648$_{\pm.0017}$
	δ+ours	.1632$_{\pm.0011}$	.1633$_{\pm.0005}$	.1482$_{\pm.0029}$	.1666$_{\pm.0004}$	.1683$_{\pm.0013}$	.1743$_{\pm.0054}$	.1583$_{\pm.0007}$	.1628$_{\pm.0007}$	.1566$_{\pm.0010}$	.1592$_{\pm.0012}$
	Δ_{δ}	+7.52$_{\pm.55}$	+4.46$_{\pm.47}$	+2.45$_{\pm.88}$	+7.14$_{\pm.33}$	+7.17$_{\pm.42}$	+6.65$_{\pm.39}$	+5.73$_{\pm.13}$	+5.95$_{\pm.70}$	+3.33$_{\pm.52}$	+4.84$_{\pm.51}$
	$\Delta_{\text{incr.}}$	+9.87$_{\pm.43}$	-2.32$_{\pm.34}$	+2.96$_{\pm.83}$	+8.88$_{\pm.27}$	+9.08$_{\pm.79}$	+5.48$_{\pm1.37}$	+5.00$_{\pm.53}$	+2.78$_{\pm.95}$	-2.08$_{\pm.38}$	+3.42$_{\pm.48}$
	Δ_{total}	+16.65$_{\pm.55}$	+2.24$_{\pm.30}$	+5.34$_{\pm.56}$	+15.39$_{\pm.23}$	+15.60$_{\pm.86}$	+11.76$_{\pm1.64}$	+10.45$_{\pm.37}$	+8.57$_{\pm.93}$	+1.32$_{\pm.77}$	+8.09$_{\pm.14}$
ECL	plain	.1946$_{\pm.0001}$	.1668$_{\pm.0002}$	.1372$_{\pm.0004}$	.2033$_{\pm.0002}$	.2172$_{\pm.0191}$	.1864$_{\pm.0008}$	.1768$_{\pm.0004}$	.1698$_{\pm.0006}$	.1450$_{\pm.0004}$	.1481$_{\pm.0003}$
	δ	.1831$_{\pm.0001}$	.1457$_{\pm.0001}$	.1371$_{\pm.0004}$	.1924$_{\pm.0003}$	.2017$_{\pm.0133}$	.1753$_{\pm.0008}$	.1633$_{\pm.0002}$	.1608$_{\pm.0007}$	.1409$_{\pm.0005}$	.1452$_{\pm.0007}$
	δ+ours	.1701$_{\pm.0032}$	.1305$_{\pm.0002}$	.1365$_{\pm.0005}$	.1757$_{\pm.0007}$	.1773$_{\pm.0041}$	.1565$_{\pm.0003}$	.1469$_{\pm.0002}$	.1492$_{\pm.0002}$	.1356$_{\pm.0005}$	.1386$_{\pm.0008}$
	Δ_{δ}	+5.95$_{\pm.03}$	+12.64$_{\pm.08}$	+0.10$_{\pm.01}$	+5.36$_{\pm.21}$	+7.02$_{\pm1.96}$	+5.96$_{\pm.03}$	+7.64$_{\pm.19}$	+5.31$_{\pm.08}$	+2.89$_{\pm.18}$	+1.99$_{\pm.25}$
	$\Delta_{\text{incr.}}$	+7.09$_{\pm1.74}$	+10.44$_{\pm.10}$	+0.43$_{\pm.13}$	+8.71$_{\pm.21}$	+11.93$_{\pm3.62}$	+10.75$_{\pm.27}$	+10.03$_{\pm.23}$	+7.19$_{\pm.44}$	+3.72$_{\pm.06}$	+4.58$_{\pm.07}$
	Δ_{total}	+12.62$_{\pm1.63}$	+21.76$_{\pm.15}$	+0.53$_{\pm.13}$	+13.60$_{\pm.39}$	+18.07$_{\pm5.05}$	+16.07$_{\pm.24}$	+16.90$_{\pm.34}$	+12.13$_{\pm.35}$	+6.50$_{\pm.12}$	+6.48$_{\pm.31}$
Solar	plain	.2900$_{\pm.0001}$	.2345$_{\pm.0099}$	.1939$_{\pm.0022}$	.3415$_{\pm.0012}$	.3646$_{\pm.0146}$	.3088$_{\pm.0015}$	.2398$_{\pm.0033}$	.2506$_{\pm.0024}$	.2123$_{\pm.0027}$	.2117$_{\pm.0041}$
	δ	.2459$_{\pm.0000}$	.2064$_{\pm.0087}$	.1865$_{\pm.0015}$	.2970$_{\pm.0011}$	.3191$_{\pm.0124}$	.2749$_{\pm.0010}$	.2184$_{\pm.0026}$	.2269$_{\pm.0018}$	.1966$_{\pm.0017}$	.2001$_{\pm.0048}$
	δ+ours	.1947$_{\pm.0001}$	.1662$_{\pm.0031}$	.1726$_{\pm.0011}$	.2164$_{\pm.0007}$	.2225$_{\pm.0067}$	.2147$_{\pm.0010}$	.1803$_{\pm.0010}$	.1850$_{\pm.0006}$	.1706$_{\pm.0019}$	.1779$_{\pm.0015}$
	Δ_{δ}	+15.19$_{\pm.04}$	+12.01$_{\pm.39}$	+3.80$_{\pm.31}$	+13.03$_{\pm.02}$	+12.47$_{\pm.10}$	+10.98$_{\pm.22}$	+8.94$_{\pm.23}$	+9.44$_{\pm.14}$	+7.41$_{\pm.43}$	+5.52$_{\pm.46}$
	$\Delta_{\text{incr.}}$	+20.85$_{\pm.02}$	+19.43$_{\pm1.83}$	+7.45$_{\pm.26}$	+27.13$_{\pm.09}$	+30.25$_{\pm.60}$	+21.89$_{\pm.32}$	+17.43$_{\pm.62}$	+18.45$_{\pm.90}$	+13.22$_{\pm.29}$	+11.08$_{\pm1.37}$
	Δ_{total}	+32.87$_{\pm.04}$	+29.11$_{\pm1.63}$	+10.97$_{\pm.53}$	+36.63$_{\pm.06}$	+38.95$_{\pm.59}$	+30.47$_{\pm.46}$	+24.81$_{\pm.64}$	+26.15$_{\pm.92}$	+19.65$_{\pm.17}$	+15.98$_{\pm.93}$
Traffic	plain	.6484$_{\pm.0003}$	.4580$_{\pm.0034}$	.4042$_{\pm.0018}$	.6540$_{\pm.0003}$	.6795$_{\pm.0053}$	.5462$_{\pm.0010}$	.4465$_{\pm.0012}$	.4937$_{\pm.0049}$	.4450$_{\pm.0018}$	.3931$_{\pm.0007}$
	δ	.6174$_{\pm.0002}$	.4360$_{\pm.0034}$	.4017$_{\pm.0017}$	.6216$_{\pm.0000}$	.6432$_{\pm.0053}$	.5251$_{\pm.0008}$	.4338$_{\pm.0012}$	.4798$_{\pm.0046}$	.4360$_{\pm.0018}$	.3870$_{\pm.0007}$
	δ+ours	.5414$_{\pm.0003}$	.3972$_{\pm.0034}$	.3869$_{\pm.0018}$	.5440$_{\pm.0017}$	.5534$_{\pm.0067}$	.4889$_{\pm.0008}$	.4174$_{\pm.0009}$	.4596$_{\pm.0043}$	.4185$_{\pm.0016}$	.3818$_{\pm.0007}$
	Δ_{δ}	+4.79$_{\pm.01}$	+4.80$_{\pm.04}$	+0.60$_{\pm.02}$	+4.94$_{\pm.05}$	+5.35$_{\pm.05}$	+3.85$_{\pm.04}$	+2.84$_{\pm.01}$	+2.80$_{\pm.04}$	+2.02$_{\pm.02}$	+1.55$_{\pm.02}$
	$\Delta_{\text{incr.}}$	+12.31$_{\pm.02}$	+8.89$_{\pm.12}$	+3.69$_{\pm.10}$	+12.48$_{\pm.28}$	+13.95$_{\pm.54}$	+6.90$_{\pm.07}$	+3.79$_{\pm.05}$	+4.22$_{\pm.05}$	+4.01$_{\pm.05}$	+1.34$_{\pm.04}$
	Δ_{total}	+16.51$_{\pm.01}$	+13.26$_{\pm.14}$	+4.27$_{\pm.08}$	+16.81$_{\pm.30}$	+18.55$_{\pm.51}$	+10.48$_{\pm.10}$	+6.51$_{\pm.05}$	+6.90$_{\pm.08}$	+5.95$_{\pm.06}$	+2.87$_{\pm.05}$

Table 22: Complete 80-cell $H{=}96$ composition with FreDF. Each backbone has paired MSE/MAE columns. Rows compare plain, FreDF (peer), and peer+ours. Entries are mean ± sample SD over seeds $2/3/4$, with smaller SD and leading zeros omitted. Δ is incremental MSE reduction (%, ↑) against the same cached peer forecast. Bold marks the lowest error among the three forecasts.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	.3963$_{\pm.0004}$	.4110$_{\pm.0004}$	**.3822**$_{\pm.0019}$	**.3997**$_{\pm.0012}$	.3821$_{\pm.0029}$	.4014$_{\pm.0025}$	.3910$_{\pm.0030}$	.4012$_{\pm.0031}$	.3928$_{\pm.0029}$	.3990$_{\pm.0027}$
	peer	**.3914**$_{\pm.0005}$	**.4046**$_{\pm.0005}$	.4729$_{\pm.0017}$	.4515$_{\pm.0013}$	.3705$_{\pm.0021}$	**.3928**$_{\pm.0014}$	.3849$_{\pm.0012}$	**.3961**$_{\pm.0017}$	**.3884**$_{\pm.0036}$	**.3951**$_{\pm.0035}$
	+ours	.3943$_{\pm.0004}$	**.4046**$_{\pm.0005}$	.3956$_{\pm.0008}$	.4122$_{\pm.0010}$	**.3695**$_{\pm.0012}$	.3956$_{\pm.0005}$	**.3847**$_{\pm.0012}$	.3994$_{\pm.0010}$	.3892$_{\pm.0047}$	.4013$_{\pm.0047}$
	$\Delta\uparrow$ (%)	-0.75$_{\pm.05}$		+16.35$_{\pm.20}$		+0.27$_{\pm.24}$		+0.04$_{\pm.06}$		-0.20$_{\pm.30}$	
ETTh2	plain	.3458$_{\pm.0061}$	.3997$_{\pm.0041}$	**.2926**$_{\pm.0009}$	**.3435**$_{\pm.0010}$	.2936$_{\pm.0039}$	.3421$_{\pm.0030}$	.2992$_{\pm.0009}$	.3459$_{\pm.0007}$	.3001$_{\pm.0039}$	.3478$_{\pm.0037}$
	peer	.2997$_{\pm.0008}$	**.3532**$_{\pm.0009}$	.3099$_{\pm.0002}$	.3557$_{\pm.0002}$	.2828$_{\pm.0022}$	**.3314**$_{\pm.0014}$	.2963$_{\pm.0009}$	**.3428**$_{\pm.0006}$	.2955$_{\pm.0012}$	.3404$_{\pm.0017}$
	+ours	**.2992**$_{\pm.0006}$	**.3532**$_{\pm.0009}$	.3034$_{\pm.0002}$	.3496$_{\pm.0001}$	**.2815**$_{\pm.0019}$	.3314$_{\pm.0013}$	**.2958**$_{\pm.0008}$	.3429$_{\pm.0006}$	**.2942**$_{\pm.0015}$	**.3399**$_{\pm.0019}$
	$\Delta\uparrow$ (%)	+0.16$_{\pm.07}$		+2.12$_{\pm.03}$		+0.45$_{\pm.18}$		+0.18$_{\pm.04}$		+0.45$_{\pm.11}$	
ETTm1	plain	.3456$_{\pm.0002}$	.3721$_{\pm.0001}$	.3323$_{\pm.0025}$	**.3683**$_{\pm.0007}$	.3211$_{\pm.0042}$	.3606$_{\pm.0023}$	.3547$_{\pm.0003}$	.3751$_{\pm.0003}$	.3695$_{\pm.0029}$	.3848$_{\pm.0019}$
	peer	.3425$_{\pm.0001}$	**.3661**$_{\pm.0001}$	.5220$_{\pm.0046}$	.4697$_{\pm.0022}$	.3087$_{\pm.0005}$	**.3458**$_{\pm.0010}$	.3460$_{\pm.0002}$	.3681$_{\pm.0002}$	.3626$_{\pm.0006}$	.3781$_{\pm.0006}$
	+ours	**.3291**$_{\pm.0000}$	**.3661**$_{\pm.0001}$	**.3313**$_{\pm.0007}$	.3718$_{\pm.0002}$	**.2997**$_{\pm.0007}$	.3466$_{\pm.0009}$	**.3342**$_{\pm.0001}$	**.3674**$_{\pm.0001}$	**.3386**$_{\pm.0011}$	**.3633**$_{\pm.0005}$
	$\Delta\uparrow$ (%)	+3.91$_{\pm.03}$		+36.53$_{\pm.44}$		+2.92$_{\pm.38}$		+3.42$_{\pm.04}$		+6.61$_{\pm.26}$	
ETTm2	plain	.1926$_{\pm.0025}$	.2912$_{\pm.0039}$	**.1780**$_{\pm.0030}$	**.2602**$_{\pm.0032}$	.1676$_{\pm.0004}$	.2504$_{\pm.0004}$	.1831$_{\pm.0002}$	.2663$_{\pm.0000}$	.1887$_{\pm.0014}$	.2718$_{\pm.0016}$
	peer	.1846$_{\pm.0002}$	**.2671**$_{\pm.0005}$	.2058$_{\pm.0002}$	.2877$_{\pm.0002}$	.1637$_{\pm.0017}$	**.2443**$_{\pm.0009}$	.1824$_{\pm.0000}$	.2600$_{\pm.0000}$	.1884$_{\pm.0002}$	.2643$_{\pm.0004}$
	+ours	**.1833**$_{\pm.0003}$	**.2671**$_{\pm.0005}$	.1956$_{\pm.0002}$	.2759$_{\pm.0003}$	**.1634**$_{\pm.0014}$	.2448$_{\pm.0008}$	**.1795**$_{\pm.0000}$	**.2595**$_{\pm.0000}$	**.1845**$_{\pm.0005}$	**.2631**$_{\pm.0004}$
	$\Delta\uparrow$ (%)	+0.70$_{\pm.02}$		+4.95$_{\pm.03}$		+0.21$_{\pm.16}$		+1.58$_{\pm.01}$		+2.08$_{\pm.15}$	
Weather	plain	.1958$_{\pm.0003}$	.2560$_{\pm.0011}$	**.1671**$_{\pm.0001}$	**.2131**$_{\pm.0004}$	.1566$_{\pm.0024}$	.2006$_{\pm.0025}$	.1969$_{\pm.0001}$	.2371$_{\pm.0001}$	.1994$_{\pm.0010}$	**.2393**$_{\pm.0005}$
	peer	.1956$_{\pm.0000}$	**.2417**$_{\pm.0001}$	.1956$_{\pm.0002}$	.2491$_{\pm.0004}$	.1503$_{\pm.0030}$	**.1936**$_{\pm.0031}$	.1984$_{\pm.0001}$	**.2360**$_{\pm.0001}$	.2048$_{\pm.0014}$	.2448$_{\pm.0022}$
	+ours	**.1642**$_{\pm.0001}$	**.2417**$_{\pm.0001}$	.1711$_{\pm.0012}$	.2538$_{\pm.0006}$	**.1479**$_{\pm.0039}$	.2141$_{\pm.0076}$	**.1644**$_{\pm.0002}$	.2465$_{\pm.0002}$	**.1751**$_{\pm.0041}$	.2601$_{\pm.0050}$
	$\Delta\uparrow$ (%)	+16.04$_{\pm.03}$		+12.52$_{\pm.56}$		+1.60$_{\pm.70}$		+17.15$_{\pm.07}$		+14.52$_{\pm1.40}$	
ECL	plain	.1946$_{\pm.0001}$	.2774$_{\pm.0006}$	.1668$_{\pm.0002}$	.2846$_{\pm.0002}$	.1372$_{\pm.0004}$	.2336$_{\pm.0005}$	.2033$_{\pm.0002}$	.2835$_{\pm.0004}$	.2172$_{\pm.0191}$	.2972$_{\pm.0268}$
	peer	.1957$_{\pm.0000}$	.2740$_{\pm.0005}$	.4936$_{\pm.0007}$	.5672$_{\pm.0005}$	**.1295**$_{\pm.0002}$	**.2237**$_{\pm.0003}$	.2037$_{\pm.0000}$	.2824$_{\pm.0001}$	.2084$_{\pm.0008}$	.2803$_{\pm.0007}$
	+ours	**.1741**$_{\pm.0001}$	**.2731**$_{\pm.0002}$	**.1532**$_{\pm.0002}$	**.2621**$_{\pm.0002}$	.1297$_{\pm.0002}$	.2268$_{\pm.0003}$	**.1790**$_{\pm.0001}$	**.2800**$_{\pm.0001}$	**.1813**$_{\pm.0006}$	**.2792**$_{\pm.0006}$
	$\Delta\uparrow$ (%)	+11.06$_{\pm.04}$		+68.96$_{\pm.01}$		-0.18$_{\pm.04}$		+12.14$_{\pm.02}$		+13.00$_{\pm.13}$	
Solar	plain	.2900$_{\pm.0001}$	**.3775**$_{\pm.0002}$	.2345$_{\pm.0099}$	**.3070**$_{\pm.0029}$	.1939$_{\pm.0022}$	.2410$_{\pm.0017}$	.3415$_{\pm.0012}$	.3699$_{\pm.0010}$	.3646$_{\pm.0146}$	.3734$_{\pm.0078}$
	peer	.3024$_{\pm.0001}$	.3777$_{\pm.0002}$	.6147$_{\pm.0027}$	.5651$_{\pm.0025}$	.1896$_{\pm.0014}$	**.2269**$_{\pm.0008}$	.3435$_{\pm.0012}$	.3759$_{\pm.0017}$	.3740$_{\pm.0112}$	.3777$_{\pm.0073}$
	+ours	**.2240**$_{\pm.0001}$	.3777$_{\pm.0002}$	**.2143**$_{\pm.0011}$	.3296$_{\pm.0019}$	**.1665**$_{\pm.0022}$	.2331$_{\pm.0025}$	**.2336**$_{\pm.0005}$	**.3173**$_{\pm.0006}$	**.2484**$_{\pm.0069}$	**.3351**$_{\pm.0093}$
	$\Delta\uparrow$ (%)	+25.90$_{\pm.03}$		+65.13$_{\pm.18}$		+12.18$_{\pm.61}$		+31.98$_{\pm.11}$		+33.59$_{\pm.17}$	
Traffic	plain	.6484$_{\pm.0003}$	.3961$_{\pm.0004}$	.4580$_{\pm.0034}$	.3142$_{\pm.0006}$	.4042$_{\pm.0018}$	.2593$_{\pm.0010}$	.6540$_{\pm.0003}$	.3926$_{\pm.0008}$	.6795$_{\pm.0053}$	.4101$_{\pm.0032}$
	peer	.6562$_{\pm.0003}$	.4009$_{\pm.0004}$	.6928$_{\pm.0026}$	.5124$_{\pm.0012}$	.3891$_{\pm.0023}$	**.2513**$_{\pm.0002}$	.6580$_{\pm.0000}$	.3924$_{\pm.0011}$	.6708$_{\pm.0043}$	.3981$_{\pm.0065}$
	+ours	**.5612**$_{\pm.0003}$	**.3714**$_{\pm.0001}$	**.4420**$_{\pm.0025}$	**.2869**$_{\pm.0029}$	**.3784**$_{\pm.0014}$	.2560$_{\pm.0005}$	**.5562**$_{\pm.0015}$	**.3720**$_{\pm.0012}$	**.5589**$_{\pm.0023}$	**.3743**$_{\pm.0038}$
	$\Delta\uparrow$ (%)	+14.48$_{\pm.01}$		+36.21$_{\pm.46}$		+2.76$_{\pm.24}$		+15.48$_{\pm.22}$		+16.67$_{\pm.70}$	

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	.4409$_{\pm.0041}$	.4408$_{\pm.0027}$	**.3833**$_{\pm.0007}$	**.3982**$_{\pm.0006}$	**.4067**$_{\pm.0023}$	**.4161**$_{\pm.0022}$	**.3869**$_{\pm.0011}$	.4045$_{\pm.0006}$	.3898$_{\pm.0021}$	.4076$_{\pm.0017}$
	peer	**.4406**$_{\pm.0211}$	**.4385**$_{\pm.0141}$	.4077$_{\pm.0019}$	.4125$_{\pm.0013}$	.4218$_{\pm.0180}$	.4259$_{\pm.0116}$	.4023$_{\pm.0004}$	.4067$_{\pm.0001}$	.3771$_{\pm.0007}$	**.3942**$_{\pm.0003}$
	+ours	.4447$_{\pm.0213}$	.4467$_{\pm.0139}$	.3970$_{\pm.0026}$	.4125$_{\pm.0013}$	.4196$_{\pm.0178}$	.4259$_{\pm.0116}$	.3883$_{\pm.0029}$	**.4035**$_{\pm.0025}$	**.3759**$_{\pm.0008}$	**.3942**$_{\pm.0003}$
	$\Delta\uparrow$ (%)	-0.93$_{\pm.01}$		+2.64$_{\pm.66}$		+0.52$_{\pm.10}$		+3.48$_{\pm.68}$		+0.32$_{\pm.06}$	
ETTh2	plain	.3190$_{\pm.0149}$	.3650$_{\pm.0128}$	**.2871**$_{\pm.0011}$	**.3399**$_{\pm.0006}$	.2969$_{\pm.0043}$	.3471$_{\pm.0032}$	.2899$_{\pm.0023}$	.3419$_{\pm.0022}$	.3091$_{\pm.0083}$	.3547$_{\pm.0041}$
	peer	.3056$_{\pm.0032}$	.3505$_{\pm.0028}$	.2935$_{\pm.0035}$	.3410$_{\pm.0021}$	.2995$_{\pm.0066}$	**.3470**$_{\pm.0060}$	.2897$_{\pm.0032}$	**.3402**$_{\pm.0026}$	.2891$_{\pm.0010}$	**.3388**$_{\pm.0010}$
	+ours	**.3046**$_{\pm.0031}$	**.3502**$_{\pm.0028}$	.2917$_{\pm.0039}$	.3410$_{\pm.0021}$	**.2958**$_{\pm.0035}$	**.3470**$_{\pm.0060}$	**.2892**$_{\pm.0023}$	.3408$_{\pm.0028}$	**.2862**$_{\pm.0012}$	**.3388**$_{\pm.0010}$
	$\Delta\uparrow$ (%)	+0.30$_{\pm.02}$		+0.61$_{\pm.13}$		+1.23$_{\pm1.05}$		+0.15$_{\pm.28}$		+1.00$_{\pm.08}$	
ETTm1	plain	.3569$_{\pm.0040}$	.3797$_{\pm.0007}$	.3337$_{\pm.0038}$	**.3690**$_{\pm.0003}$	.3339$_{\pm.0038}$	.3712$_{\pm.0014}$	**.3233**$_{\pm.0005}$	**.3604**$_{\pm.0010}$	.3426$_{\pm.0047}$	.3752$_{\pm.0037}$
	peer	.3569$_{\pm.0020}$	**.3743**$_{\pm.0014}$	.4342$_{\pm.0032}$	.4277$_{\pm.0017}$	.3293$_{\pm.0007}$	**.3583**$_{\pm.0010}$	.4194$_{\pm.0006}$	.4193$_{\pm.0012}$	.3343$_{\pm.0059}$	**.3668**$_{\pm.0024}$
	+ours	**.3470**$_{\pm.0018}$	.3751$_{\pm.0012}$	**.3306**$_{\pm.0010}$	.3701$_{\pm.0012}$	**.3226**$_{\pm.0005}$	.3615$_{\pm.0010}$	.3315$_{\pm.0045}$	.3715$_{\pm.0034}$	**.3257**$_{\pm.0074}$	**.3668**$_{\pm.0024}$
	$\Delta\uparrow$ (%)	+2.77$_{\pm.06}$		+23.85$_{\pm.62}$		+2.03$_{\pm.10}$		+20.96$_{\pm1.17}$		+2.58$_{\pm.51}$	
ETTm2	plain	.1930$_{\pm.0007}$	.2772$_{\pm.0008}$	**.1756**$_{\pm.0006}$	**.2604**$_{\pm.0009}$	.1774$_{\pm.0018}$	.2623$_{\pm.0021}$	**.1726**$_{\pm.0002}$	**.2570**$_{\pm.0011}$	.1848$_{\pm.0026}$	.2669$_{\pm.0025}$
	peer	.1918$_{\pm.0018}$	.2688$_{\pm.0015}$	.1919$_{\pm.0002}$	.2740$_{\pm.0004}$	.1775$_{\pm.0019}$	**.2565**$_{\pm.0013}$	.1872$_{\pm.0020}$	.2708$_{\pm.0010}$	.1793$_{\pm.0007}$	**.2586**$_{\pm.0026}$
	+ours	**.1899**$_{\pm.0019}$	**.2671**$_{\pm.0015}$	.1854$_{\pm.0001}$	.2657$_{\pm.0004}$	**.1760**$_{\pm.0017}$	.2567$_{\pm.0012}$	.1816$_{\pm.0020}$	.2640$_{\pm.0008}$	**.1754**$_{\pm.0004}$	**.2586**$_{\pm.0026}$
	$\Delta\uparrow$ (%)	+0.97$_{\pm.07}$		+3.39$_{\pm.03}$		+0.82$_{\pm.10}$		+2.98$_{\pm.12}$		+2.17$_{\pm.17}$	
Weather	plain	.1975$_{\pm.0024}$	.2405$_{\pm.0037}$	.1768$_{\pm.0011}$	**.2180**$_{\pm.0007}$	.1781$_{\pm.0023}$	.2240$_{\pm.0023}$	**.1587**$_{\pm.0007}$	**.2075**$_{\pm.0005}$	.1732$_{\pm.0011}$	.2132$_{\pm.0009}$
	peer	.1968$_{\pm.0010}$	**.2381**$_{\pm.0016}$	.1895$_{\pm.0004}$	.2392$_{\pm.0004}$	.1787$_{\pm.0063}$	**.2239**$_{\pm.0062}$	.1790$_{\pm.0007}$	.2333$_{\pm.0016}$	.1719$_{\pm.0027}$	**.2109**$_{\pm.0039}$
	+ours	**.1648**$_{\pm.0012}$	.2459$_{\pm.0017}$	**.1644**$_{\pm.0004}$	.2462$_{\pm.0009}$	**.1595**$_{\pm.0002}$	.2388$_{\pm.0001}$	.1635$_{\pm.0020}$	.2458$_{\pm.0026}$	**.1574**$_{\pm.0024}$	**.2109**$_{\pm.0039}$
	$\Delta\uparrow$ (%)	+16.26$_{\pm.20}$		+13.28$_{\pm.24}$		+10.69$_{\pm3.05}$		+8.65$_{\pm.78}$		+8.44$_{\pm.04}$	
ECL	plain	.1864$_{\pm.0008}$	.2738$_{\pm.0015}$	.1768$_{\pm.0004}$	.2710$_{\pm.0006}$	.1698$_{\pm.0006}$	.2609$_{\pm.0006}$	**.1450**$_{\pm.0004}$	**.2436**$_{\pm.0003}$	.1481$_{\pm.0003}$	.2400$_{\pm.0003}$
	peer	.1906$_{\pm.0004}$	.2712$_{\pm.0005}$	.3208$_{\pm.0014}$	.4340$_{\pm.0019}$	.1717$_{\pm.0011}$	.2586$_{\pm.0009}$	.2683$_{\pm.0040}$	.3935$_{\pm.0043}$	.1466$_{\pm.0003}$	.2357$_{\pm.0004}$
	+ours	**.1634**$_{\pm.0002}$	**.2641**$_{\pm.0001}$	**.1613**$_{\pm.0020}$	**.2673**$_{\pm.0019}$	**.1544**$_{\pm.0016}$	**.2556**$_{\pm.0014}$	.1473$_{\pm.0028}$	.2511$_{\pm.0027}$	**.1364**$_{\pm.0002}$	**.2346**$_{\pm.0005}$
	$\Delta\uparrow$ (%)	+14.29$_{\pm.11}$		+49.71$_{\pm.78}$		+10.08$_{\pm1.01}$		+45.07$_{\pm1.62}$		+6.96$_{\pm.06}$	
Solar	plain	.3088$_{\pm.0015}$	.3160$_{\pm.0058}$	.2398$_{\pm.0033}$	**.2885**$_{\pm.0019}$	.2506$_{\pm.0024}$	.2939$_{\pm.0021}$	.2123$_{\pm.0027}$	**.2709**$_{\pm.0027}$	.2117$_{\pm.0041}$	.2545$_{\pm.0023}$
	peer	.3156$_{\pm.0032}$	.3152$_{\pm.0055}$	.4040$_{\pm.0021}$	.4581$_{\pm.0030}$	.2728$_{\pm.0069}$	.2894$_{\pm.0021}$	.3980$_{\pm.0032}$	.4579$_{\pm.0041}$	.2153$_{\pm.0089}$	**.2469**$_{\pm.0042}$
	+ours	**.2336**$_{\pm.0010}$	**.2967**$_{\pm.0019}$	**.2216**$_{\pm.0013}$	.3072$_{\pm.0028}$	**.2083**$_{\pm.0046}$	**.2809**$_{\pm.0023}$	**.2005**$_{\pm.0024}$	.2957$_{\pm.0042}$	**.1725**$_{\pm.0056}$	**.2469**$_{\pm.0042}$
	$\Delta\uparrow$ (%)	+25.96$_{\pm.47}$		+45.14$_{\pm.58}$		+23.64$_{\pm.30}$		+49.61$_{\pm.38}$		+19.87$_{\pm.69}$	
Traffic	plain	.5462$_{\pm.0010}$	.3555$_{\pm.0015}$	.4465$_{\pm.0012}$	.2908$_{\pm.0005}$	.4937$_{\pm.0049}$	.3253$_{\pm.0015}$	.4450$_{\pm.0018}$	**.2929**$_{\pm.0011}$	.3931$_{\pm.0007}$	.2684$_{\pm.0006}$
	peer	.5452$_{\pm.0028}$	.3491$_{\pm.0018}$	.5139$_{\pm.0012}$	.3749$_{\pm.0011}$	.4864$_{\pm.0066}$	.3178$_{\pm.0031}$	.5100$_{\pm.0033}$	.3707$_{\pm.0004}$	.3906$_{\pm.0004}$	.2612$_{\pm.0004}$
	+ours	**.4994**$_{\pm.0006}$	**.3423**$_{\pm.0006}$	**.4353**$_{\pm.0024}$	**.2868**$_{\pm.0029}$	**.4601**$_{\pm.0050}$	**.3121**$_{\pm.0021}$	**.4285**$_{\pm.0023}$	.2930$_{\pm.0006}$	**.3767**$_{\pm.0003}$	**.2583**$_{\pm.0004}$
	$\Delta\uparrow$ (%)	+8.41$_{\pm.41}$		+15.30$_{\pm.59}$		+5.39$_{\pm.29}$		+15.97$_{\pm.12}$		+3.56$_{\pm.04}$	

Table 23: Complete 80-cell $H{=}96$ composition with PS Loss. Each backbone has paired MSE/MAE columns. Rows compare plain, PS Loss (peer), and peer+ours. Entries are mean ± sample SD over seeds $2/3/4$, with smaller SD and leading zeros omitted. Δ is incremental MSE reduction (%, ↑) against the same cached peer forecast. Bold marks the lowest error among the three forecasts.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	.3963$_{\pm.0004}$	.4110$_{\pm.0004}$	**.3822**$_{\pm.0019}$	.3997$_{\pm.0012}$	.3821$_{\pm.0029}$	.4014$_{\pm.0025}$	.3910$_{\pm.0030}$	**.4012**$_{\pm.0031}$	.3928$_{\pm.0029}$	.3990$_{\pm.0027}$
	peer	**.3911**$_{\pm.0003}$	**.4045**$_{\pm.0002}$	.3842$_{\pm.0012}$	**.3983**$_{\pm.0004}$	.3692$_{\pm.0012}$	**.3926**$_{\pm.0007}$	.3936$_{\pm.0043}$	.4033$_{\pm.0035}$	.3895$_{\pm.0045}$	**.3971**$_{\pm.0037}$
	+ours	.3927$_{\pm.0004}$	**.4045**$_{\pm.0002}$	.3857$_{\pm.0009}$	.4012$_{\pm.0002}$	**.3664**$_{\pm.0005}$	.3943$_{\pm.0006}$	**.3892**$_{\pm.0030}$	.4031$_{\pm.0021}$	**.3879**$_{\pm.0044}$	.4011$_{\pm.0042}$
	$\Delta\uparrow$ (%)	-0.41$_{\pm.04}$		-0.38$_{\pm.08}$		+0.76$_{\pm.19}$		+1.09$_{\pm.34}$		+0.41$_{\pm.10}$	
ETTh2	plain	.3458$_{\pm.0061}$	.3997$_{\pm.0041}$	.2926$_{\pm.0009}$	.3435$_{\pm.0010}$	.2936$_{\pm.0039}$	.3421$_{\pm.0030}$	.2992$_{\pm.0009}$	.3459$_{\pm.0007}$	.3001$_{\pm.0039}$	.3478$_{\pm.0037}$
	peer	.2989$_{\pm.0005}$	**.3510**$_{\pm.0005}$	.2858$_{\pm.0003}$	.3353$_{\pm.0005}$	.2898$_{\pm.0017}$	.3327$_{\pm.0011}$	.2982$_{\pm.0009}$	.3437$_{\pm.0007}$	.3015$_{\pm.0037}$	.3452$_{\pm.0036}$
	+ours	**.2984**$_{\pm.0002}$	**.3510**$_{\pm.0005}$	**.2854**$_{\pm.0012}$	**.3352**$_{\pm.0016}$	**.2873**$_{\pm.0005}$	**.3323**$_{\pm.0006}$	**.2963**$_{\pm.0008}$	**.3426**$_{\pm.0007}$	**.2993**$_{\pm.0029}$	**.3436**$_{\pm.0027}$
	$\Delta\uparrow$ (%)	+0.16$_{\pm.13}$		+0.15$_{\pm.30}$		+0.85$_{\pm.43}$		+0.63$_{\pm.03}$		+0.73$_{\pm.27}$	
ETTm1	plain	.3456$_{\pm.0002}$	**.3721**$_{\pm.0001}$	.3323$_{\pm.0025}$	.3683$_{\pm.0007}$	.3211$_{\pm.0042}$	.3606$_{\pm.0023}$	.3547$_{\pm.0003}$	.3751$_{\pm.0003}$	.3695$_{\pm.0029}$	.3848$_{\pm.0019}$
	peer	.3507$_{\pm.0002}$	.3746$_{\pm.0001}$	.3312$_{\pm.0007}$	.3687$_{\pm.0004}$	.3157$_{\pm.0007}$	.3540$_{\pm.0006}$	.3647$_{\pm.0021}$	.3834$_{\pm.0013}$	.3916$_{\pm.0101}$	.4007$_{\pm.0065}$
	+ours	**.3309**$_{\pm.0004}$	.3746$_{\pm.0001}$	**.3129**$_{\pm.0010}$	**.3573**$_{\pm.0014}$	**.3062**$_{\pm.0007}$	**.3531**$_{\pm.0022}$	**.3333**$_{\pm.0006}$	**.3632**$_{\pm.0005}$	**.3562**$_{\pm.0070}$	**.3805**$_{\pm.0052}$
	$\Delta\uparrow$ (%)	+5.65$_{\pm.16}$		+5.55$_{\pm.52}$		+2.99$_{\pm.19}$		+8.60$_{\pm.37}$		+9.04$_{\pm.70}$	
ETTm2	plain	.1926$_{\pm.0025}$	.2912$_{\pm.0039}$	.1780$_{\pm.0030}$	.2602$_{\pm.0032}$	.1676$_{\pm.0004}$	.2504$_{\pm.0004}$	.1831$_{\pm.0002}$	.2663$_{\pm.0000}$	.1887$_{\pm.0014}$	.2718$_{\pm.0016}$
	peer	.1822$_{\pm.0003}$	**.2656**$_{\pm.0010}$	.1744$_{\pm.0005}$	**.2547**$_{\pm.0010}$	.1663$_{\pm.0007}$	.2458$_{\pm.0005}$	.1838$_{\pm.0001}$	.2629$_{\pm.0001}$	.1884$_{\pm.0014}$	.2671$_{\pm.0018}$
	+ours	**.1810**$_{\pm.0003}$	**.2656**$_{\pm.0010}$	**.1741**$_{\pm.0005}$	.2548$_{\pm.0010}$	**.1652**$_{\pm.0007}$	**.2454**$_{\pm.0006}$	**.1817**$_{\pm.0000}$	**.2605**$_{\pm.0000}$	**.1859**$_{\pm.0014}$	**.2643**$_{\pm.0017}$
	$\Delta\uparrow$ (%)	+0.70$_{\pm.01}$		+0.12$_{\pm.01}$		+0.67$_{\pm.08}$		+1.15$_{\pm.01}$		+1.35$_{\pm.05}$	
Weather	plain	.1958$_{\pm.0003}$	.2560$_{\pm.0011}$	.1671$_{\pm.0001}$	.2131$_{\pm.0004}$	.1566$_{\pm.0024}$	.2006$_{\pm.0025}$	.1969$_{\pm.0001}$	.2371$_{\pm.0001}$	.1994$_{\pm.0010}$	**.2393**$_{\pm.0005}$
	peer	.1981$_{\pm.0012}$	**.2405**$_{\pm.0037}$	.1616$_{\pm.0001}$	**.2033**$_{\pm.0002}$	.1508$_{\pm.0038}$	**.1926**$_{\pm.0037}$	.2008$_{\pm.0001}$	**.2370**$_{\pm.0002}$	.2071$_{\pm.0004}$	.2418$_{\pm.0010}$
	+ours	**.1626**$_{\pm.0022}$	**.2405**$_{\pm.0037}$	**.1569**$_{\pm.0005}$	.2323$_{\pm.0010}$	**.1480**$_{\pm.0041}$	.2152$_{\pm.0077}$	**.1648**$_{\pm.0002}$	.2472$_{\pm.0002}$	**.1655**$_{\pm.0014}$	.2480$_{\pm.0019}$
	$\Delta\uparrow$ (%)	+17.95$_{\pm.60}$		+2.94$_{\pm.22}$		+1.87$_{\pm.38}$		+17.94$_{\pm.05}$		+20.11$_{\pm.59}$	
ECL	plain	.1946$_{\pm.0001}$	.2774$_{\pm.0006}$	.1668$_{\pm.0002}$	.2846$_{\pm.0002}$	.1372$_{\pm.0004}$	.2336$_{\pm.0005}$	.2033$_{\pm.0002}$	.2835$_{\pm.0004}$	.2172$_{\pm.0191}$	.2972$_{\pm.0268}$
	peer	.1961$_{\pm.0000}$	**.2738**$_{\pm.0001}$	.1642$_{\pm.0005}$	.2719$_{\pm.0004}$	.1339$_{\pm.0006}$	**.2291**$_{\pm.0005}$	.2049$_{\pm.0004}$	.2836$_{\pm.0007}$	.2106$_{\pm.0034}$	.2842$_{\pm.0030}$
	+ours	**.1735**$_{\pm.0001}$	**.2738**$_{\pm.0001}$	**.1337**$_{\pm.0001}$	**.2332**$_{\pm.0001}$	**.1336**$_{\pm.0005}$	.2316$_{\pm.0004}$	**.1789**$_{\pm.0000}$	**.2796**$_{\pm.0000}$	**.1811**$_{\pm.0012}$	**.2795**$_{\pm.0011}$
	$\Delta\uparrow$ (%)	+11.54$_{\pm.03}$		+18.58$_{\pm.23}$		+0.17$_{\pm.09}$		+12.67$_{\pm.15}$		+13.99$_{\pm.79}$	
Solar	plain	.2900$_{\pm.0001}$	.3775$_{\pm.0002}$	.2345$_{\pm.0099}$	.3070$_{\pm.0029}$	.1939$_{\pm.0022}$	**.2410**$_{\pm.0017}$	.3415$_{\pm.0012}$	.3699$_{\pm.0010}$	.3646$_{\pm.0146}$	.3734$_{\pm.0078}$
	peer	.2849$_{\pm.0005}$	.3465$_{\pm.0011}$	.2545$_{\pm.0030}$	.2908$_{\pm.0020}$	.2186$_{\pm.0028}$	.2597$_{\pm.0018}$	.4041$_{\pm.0055}$	.3877$_{\pm.0022}$	.3988$_{\pm.0273}$	.4045$_{\pm.0222}$
	+ours	**.2168**$_{\pm.0003}$	**.3087**$_{\pm.0007}$	**.1813**$_{\pm.0012}$	**.2632**$_{\pm.0009}$	**.1734**$_{\pm.0012}$	.2493$_{\pm.0007}$	**.2345**$_{\pm.0020}$	**.3334**$_{\pm.0012}$	**.2503**$_{\pm.0092}$	**.3412**$_{\pm.0104}$
	$\Delta\uparrow$ (%)	+23.89$_{\pm.03}$		+28.77$_{\pm.60}$		+20.69$_{\pm1.53}$		+41.94$_{\pm1.23}$		+37.15$_{\pm2.19}$	
Traffic	plain	.6484$_{\pm.0003}$	.3961$_{\pm.0004}$	.4580$_{\pm.0034}$	.3142$_{\pm.0006}$	.4042$_{\pm.0018}$	.2593$_{\pm.0010}$	.6540$_{\pm.0003}$	.3926$_{\pm.0008}$	.6795$_{\pm.0053}$	.4101$_{\pm.0032}$
	peer	.6456$_{\pm.0001}$	**.3840**$_{\pm.0000}$	.4636$_{\pm.0015}$	.2950$_{\pm.0001}$	.3954$_{\pm.0010}$	**.2548**$_{\pm.0003}$	.6608$_{\pm.0004}$	.3933$_{\pm.0021}$	.7064$_{\pm.0124}$	.4244$_{\pm.0099}$
	+ours	**.5536**$_{\pm.0002}$	**.3840**$_{\pm.0000}$	**.4079**$_{\pm.0007}$	**.2787**$_{\pm.0001}$	**.3805**$_{\pm.0009}$	.2588$_{\pm.0003}$	**.5552**$_{\pm.0009}$	**.3728**$_{\pm.0003}$	**.5957**$_{\pm.0108}$	**.3993**$_{\pm.0084}$
	$\Delta\uparrow$ (%)	+14.25$_{\pm.02}$		+12.01$_{\pm.14}$		+3.77$_{\pm.09}$		+15.98$_{\pm.17}$		+15.67$_{\pm.11}$	

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	plain	**.4409**$_{\pm.0041}$	**.4408**$_{\pm.0027}$	.3833$_{\pm.0007}$	.3982$_{\pm.0006}$	.4067$_{\pm.0023}$	.4161$_{\pm.0022}$	.3869$_{\pm.0011}$	.4045$_{\pm.0006}$	.3898$_{\pm.0021}$	.4076$_{\pm.0017}$
	peer	.4494$_{\pm.0029}$	.4445$_{\pm.0018}$	.3844$_{\pm.0005}$	**.3952**$_{\pm.0004}$	.3966$_{\pm.0005}$	**.4072**$_{\pm.0005}$	.3800$_{\pm.0003}$	**.3936**$_{\pm.0007}$	.3797$_{\pm.0014}$	**.3980**$_{\pm.0012}$
	+ours	.4532$_{\pm.0027}$	.4516$_{\pm.0017}$	**.3832**$_{\pm.0001}$	**.3952**$_{\pm.0004}$	**.3915**$_{\pm.0012}$	**.4072**$_{\pm.0005}$	**.3791**$_{\pm.0003}$	.3963$_{\pm.0010}$	**.3760**$_{\pm.0012}$	**.3980**$_{\pm.0012}$
	$\Delta\uparrow$ (%)	-0.86$_{\pm.05}$		+0.31$_{\pm.12}$		+1.31$_{\pm.29}$		+0.26$_{\pm.12}$		+0.98$_{\pm.07}$	
ETTh2	plain	.3190$_{\pm.0149}$	.3650$_{\pm.0128}$	.2871$_{\pm.0011}$	.3399$_{\pm.0006}$	.2969$_{\pm.0043}$	.3471$_{\pm.0032}$	.2899$_{\pm.0023}$	.3419$_{\pm.0022}$	.3091$_{\pm.0083}$	.3547$_{\pm.0041}$
	peer	.3039$_{\pm.0024}$	.3483$_{\pm.0018}$	.2804$_{\pm.0002}$	**.3304**$_{\pm.0003}$	.2924$_{\pm.0020}$	**.3402**$_{\pm.0014}$	.2806$_{\pm.0026}$	.3320$_{\pm.0020}$	.2944$_{\pm.0038}$	**.3410**$_{\pm.0015}$
	+ours	**.3032**$_{\pm.0023}$	**.3483**$_{\pm.0016}$	**.2794**$_{\pm.0003}$	**.3304**$_{\pm.0003}$	**.2913**$_{\pm.0016}$	**.3402**$_{\pm.0014}$	**.2792**$_{\pm.0015}$	**.3319**$_{\pm.0017}$	**.2907**$_{\pm.0041}$	**.3410**$_{\pm.0015}$
	$\Delta\uparrow$ (%)	+0.23$_{\pm.05}$		+0.34$_{\pm.05}$		+0.40$_{\pm.13}$		+0.49$_{\pm.40}$		+1.26$_{\pm.20}$	
ETTm1	plain	**.3569**$_{\pm.0040}$	**.3797**$_{\pm.0007}$	.3337$_{\pm.0038}$	.3690$_{\pm.0003}$	.3339$_{\pm.0038}$	.3712$_{\pm.0014}$	.3233$_{\pm.0005}$	.3604$_{\pm.0010}$	.3426$_{\pm.0047}$	.3752$_{\pm.0037}$
	peer	.3872$_{\pm.0076}$	.3972$_{\pm.0052}$	.3359$_{\pm.0028}$	**.3671**$_{\pm.0022}$	.3414$_{\pm.0013}$	**.3702**$_{\pm.0008}$	.3268$_{\pm.0008}$	.3594$_{\pm.0007}$	.3304$_{\pm.0022}$	**.3657**$_{\pm.0008}$
	+ours	.3696$_{\pm.0071}$	.3892$_{\pm.0062}$	**.3185**$_{\pm.0024}$	**.3671**$_{\pm.0022}$	**.3317**$_{\pm.0011}$	**.3702**$_{\pm.0008}$	**.3167**$_{\pm.0039}$	**.3561**$_{\pm.0038}$	**.3204**$_{\pm.0032}$	**.3657**$_{\pm.0008}$
	$\Delta\uparrow$ (%)	+4.55$_{\pm.31}$		+5.19$_{\pm.26}$		+2.85$_{\pm.08}$		+3.09$_{\pm.99}$		+3.02$_{\pm.98}$	
ETTm2	plain	.1930$_{\pm.0007}$	.2772$_{\pm.0008}$	.1756$_{\pm.0006}$	.2604$_{\pm.0009}$	.1774$_{\pm.0018}$	.2623$_{\pm.0021}$	.1726$_{\pm.0002}$	.2570$_{\pm.0011}$	.1848$_{\pm.0026}$	.2669$_{\pm.0025}$
	peer	.1932$_{\pm.0014}$	.2713$_{\pm.0012}$	.1759$_{\pm.0004}$	.2556$_{\pm.0001}$	.1753$_{\pm.0014}$	.2560$_{\pm.0012}$	.1679$_{\pm.0008}$	**.2485**$_{\pm.0008}$	.1760$_{\pm.0013}$	**.2556**$_{\pm.0016}$
	+ours	**.1910**$_{\pm.0014}$	**.2689**$_{\pm.0012}$	**.1751**$_{\pm.0002}$	**.2556**$_{\pm.0001}$	**.1730**$_{\pm.0012}$	**.2554**$_{\pm.0008}$	**.1678**$_{\pm.0007}$	.2489$_{\pm.0008}$	**.1728**$_{\pm.0007}$	**.2556**$_{\pm.0016}$
	$\Delta\uparrow$ (%)	+1.12$_{\pm.01}$		+0.47$_{\pm.13}$		+1.34$_{\pm.12}$		+0.04$_{\pm.02}$		+1.82$_{\pm.36}$	
Weather	plain	.1975$_{\pm.0024}$	.2405$_{\pm.0037}$	.1768$_{\pm.0011}$	.2180$_{\pm.0007}$	.1781$_{\pm.0023}$	.2240$_{\pm.0023}$	.1587$_{\pm.0007}$	.2075$_{\pm.0005}$	.1732$_{\pm.0011}$	.2132$_{\pm.0009}$
	peer	.1953$_{\pm.0011}$	**.2354**$_{\pm.0012}$	.1750$_{\pm.0007}$	**.2130**$_{\pm.0012}$	.1689$_{\pm.0009}$	**.2120**$_{\pm.0007}$	.1570$_{\pm.0005}$	**.2001**$_{\pm.0009}$	.1690$_{\pm.0007}$	**.2068**$_{\pm.0010}$
	+ours	**.1638**$_{\pm.0003}$	.2452$_{\pm.0004}$	**.1575**$_{\pm.0005}$	.2351$_{\pm.0008}$	**.1602**$_{\pm.0043}$	.2371$_{\pm.0014}$	**.1551**$_{\pm.0004}$	.2287$_{\pm.0051}$	**.1529**$_{\pm.0011}$	**.2068**$_{\pm.0010}$
	$\Delta\uparrow$ (%)	+16.13$_{\pm.29}$		+9.99$_{\pm.44}$		+5.11$_{\pm2.34}$		+1.19$_{\pm.39}$		+9.50$_{\pm1.02}$	
ECL	plain	.1864$_{\pm.0008}$	.2738$_{\pm.0015}$	.1768$_{\pm.0004}$	.2710$_{\pm.0006}$	.1698$_{\pm.0006}$	.2609$_{\pm.0006}$	.1450$_{\pm.0004}$	.2436$_{\pm.0003}$	.1481$_{\pm.0003}$	.2400$_{\pm.0003}$
	peer	.1964$_{\pm.0006}$	.2777$_{\pm.0011}$	.1803$_{\pm.0004}$	**.2674**$_{\pm.0003}$	.1718$_{\pm.0007}$	.2596$_{\pm.0009}$	.1466$_{\pm.0005}$	.2416$_{\pm.0006}$	.1467$_{\pm.0001}$	**.2362**$_{\pm.0001}$
	+ours	**.1670**$_{\pm.0008}$	**.2691**$_{\pm.0011}$	**.1521**$_{\pm.0010}$	**.2674**$_{\pm.0003}$	**.1523**$_{\pm.0004}$	**.2546**$_{\pm.0002}$	**.1367**$_{\pm.0002}$	**.2351**$_{\pm.0002}$	**.1366**$_{\pm.0001}$	**.2362**$_{\pm.0001}$
	$\Delta\uparrow$ (%)	+14.99$_{\pm.30}$		+15.64$_{\pm.37}$		+11.34$_{\pm.19}$		+6.73$_{\pm.20}$		+6.89$_{\pm.07}$	
Solar	plain	.3088$_{\pm.0015}$	**.3160**$_{\pm.0058}$	.2398$_{\pm.0033}$	.2885$_{\pm.0019}$	.2506$_{\pm.0024}$	.2939$_{\pm.0021}$	.2123$_{\pm.0027}$	.2709$_{\pm.0027}$	.2117$_{\pm.0041}$	**.2545**$_{\pm.0023}$
	peer	.3685$_{\pm.0259}$	.3605$_{\pm.0162}$	.2538$_{\pm.0003}$	.2799$_{\pm.0012}$	.2826$_{\pm.0024}$	.3044$_{\pm.0007}$	.2385$_{\pm.0029}$	.2710$_{\pm.0022}$	.2236$_{\pm.0126}$	.2550$_{\pm.0078}$
	+ours	**.2449**$_{\pm.0014}$	.3162$_{\pm.0090}$	**.1971**$_{\pm.0011}$	**.2699**$_{\pm.0008}$	**.2152**$_{\pm.0024}$	**.2887**$_{\pm.0023}$	**.1904**$_{\pm.0019}$	**.2677**$_{\pm.0022}$	**.1769**$_{\pm.0060}$	.2550$_{\pm.0078}$
	$\Delta\uparrow$ (%)	+33.33$_{\pm4.49}$		+22.33$_{\pm.41}$		+23.85$_{\pm.65}$		+20.18$_{\pm.71}$		+20.80$_{\pm1.77}$	
Traffic	plain	.5462$_{\pm.0010}$	.3555$_{\pm.0015}$	.4465$_{\pm.0012}$	.2908$_{\pm.0005}$	.4937$_{\pm.0049}$	.3253$_{\pm.0015}$	.4450$_{\pm.0018}$	.2929$_{\pm.0011}$	.3931$_{\pm.0007}$	.2684$_{\pm.0006}$
	peer	.5641$_{\pm.0028}$	.3643$_{\pm.0019}$	.4536$_{\pm.0000}$	**.2900**$_{\pm.0002}$	.4906$_{\pm.0026}$	.3258$_{\pm.0015}$	.4522$_{\pm.0013}$	.2902$_{\pm.0005}$	.3934$_{\pm.0010}$	**.2670**$_{\pm.0008}$
	+ours	**.5095**$_{\pm.0023}$	**.3482**$_{\pm.0019}$	**.4271**$_{\pm.0002}$	**.2900**$_{\pm.0002}$	**.4603**$_{\pm.0023}$	**.3169**$_{\pm.0009}$	**.4278**$_{\pm.0010}$	**.2886**$_{\pm.0004}$	**.3770**$_{\pm.0007}$	**.2670**$_{\pm.0008}$
	$\Delta\uparrow$ (%)	+9.67$_{\pm.06}$		+5.83$_{\pm.04}$		+6.17$_{\pm.10}$		+5.41$_{\pm.12}$		+4.18$_{\pm.10}$	

D.5 Additional horizons and selection-rule variants

Rule variants. Table 24 compares four candidate restrictions at every horizon. Each entry gives the number of improved cells, mean relative test-MSE reduction, and worst reduction. Higher percentage values are better.

Table 24: Selection-rule variants on the same 80 cells at every horizon. Each entry lists mean ± SD across cell means on the first line, followed by improved cells and the worst reduction together on the second line (%, ↑). The nominal dimension-admissibility guard uses only fitting-set width and channel count. The final row ranks already fitted candidates by test MSE and is an upper reference, not a deployable rule.

	H=96	H=192	H=336	H=720
first-run rule	$+7.62_{\pm 8.10}$ 77/80; −0.73	$+6.91_{\pm 8.26}$ 73/80; −3.45	$+6.22_{\pm 8.60}$ 68/80; −7.43	$+4.04_{\pm 9.52}$ 60/80; −18.08
dimension-guarded rule	$+7.61_{\pm 8.10}$ 77/80; −0.73	$+6.91_{\pm 8.26}$ 73/80; −3.45	$+6.22_{\pm 8.60}$ 68/80; −7.43	$+4.64_{\pm 9.11}$ 61/80; −11.68
train-fitted only	$+7.28_{\pm 8.25}$ 75/80; −2.13	$+5.72_{\pm 9.75}$ 70/80; −14.76	$+3.34_{\pm 14.13}$ 68/80; −36.86	$-0.93_{\pm 22.82}$ 59/80; −67.57
no $G_5(3)$	$+6.28_{\pm 6.28}$ 78/80; −0.73	$+5.93_{\pm 6.79}$ 73/80; −3.45	$+5.65_{\pm 7.85}$ 68/80; −7.43	$+4.16_{\pm 8.69}$ 61/80; −11.68
test-ranked upper reference	$+7.96_{\pm 7.89}$ 80/80; +0.20	$+7.68_{\pm 7.78}$ 80/80; +0.42	$+7.62_{\pm 7.52}$ 78/80; −1.18	$+7.93_{\pm 7.43}$ 73/80; −4.05

The unguarded row preserves the rule fixed before the long-horizon experiments; the dimension-guarded row is the deployed rule. For $H \leq 336$, they have the same improvement counts and worst reductions. Their reported means are identical except for a 0.01-point difference at H=96. At H=720, the guard changes the mean from 4.04% to 4.64% and the worst reduction from -18.08% to -11.68%. Excluding $G_5(3)$ retains $61/80$ improvements at H=720 with a 4.16% mean, while the training-fitted restriction has a -0.93% mean. These comparisons use all 80 cells at every horizon.

Complete four-horizon matrix. Tables 25 and 26 resolve the aggregate rows above by dataset and backbone. Each entry repeats validation selection for the three paired checkpoints and reports their mean relative reduction and sample standard deviation. The two parts contain all 80 matched cells at each of the four horizons.

Complete MSE and MAE. Tables 28–31 report baseline and corrected MSE and MAE for every dataset–backbone–horizon cell. Every metric and reduction is a paired-seed mean with its sample standard deviation. Candidate operators are selected by the validation spectral score in Eq. 23; both test MSE and test MAE evaluate the resulting forecast.

On Weather at $H = 96$, correction reduces MSE by 9.68% while increasing mean absolute error (MAE) by 6.91%. The two metrics weight residual magnitudes differently, and candidate selection uses squared error. We retain both measurements in the full tables to distinguish the selected objective from absolute-error performance.

Table 25: Complete four-horizon matrix, part 1 of 2 (ETT benchmarks). Each entry is the mean ± sample SD of the paired test-MSE reduction (%, ↑) over seeds $2/3/4$. Bold marks lower test MSE. Datasets are paired horizontally; leading zeros are omitted.

backbone	ETTh1				ETTh2			
	H=96	H=192	H=336	H=720	H=96	H=192	H=336	H=720
DLinear	-0.32$_{\pm.01}$	-3.45$_{\pm.02}$	-6.95$_{\pm.00}$	-10.74$_{\pm.13}$	**+9.82**$_{\pm1.67}$	**+1.48**$_{\pm.16}$	**+0.77**$_{\pm.07}$	**+10.28**$_{\pm.03}$
DUET	**+0.01**$_{\pm.10}$	-3.22$_{\pm.28}$	-7.43$_{\pm.58}$	-10.60$_{\pm.44}$	**+0.85**$_{\pm.24}$	**+1.38**$_{\pm.63}$	**+3.98**$_{\pm.54}$	**+5.00**$_{\pm1.34}$
FACT	**+2.11**$_{\pm.65}$	**+0.66**$_{\pm.30}$	-3.84$_{\pm1.68}$	-4.44$_{\pm1.08}$	**+2.20**$_{\pm.85}$	**+2.54**$_{\pm.43}$	**+4.79**$_{\pm.18}$	**+6.92**$_{\pm1.92}$
FITS	**+0.89**$_{\pm.17}$	-1.25$_{\pm.13}$	-4.99$_{\pm.09}$	-11.68$_{\pm.30}$	**+0.79**$_{\pm.05}$	**+0.39**$_{\pm.02}$	**+0.40**$_{\pm.03}$	**+0.61**$_{\pm.02}$
MixLinear	**+0.73**$_{\pm.20}$	-0.78$_{\pm.63}$	-4.91$_{\pm.36}$	-11.43$_{\pm1.03}$	**+1.22**$_{\pm.39}$	**+0.86**$_{\pm.18}$	**+0.53**$_{\pm.04}$	**+0.59**$_{\pm.10}$
PaiFilter	-0.73$_{\pm.05}$	-2.38$_{\pm.05}$	-5.05$_{\pm.02}$	-7.37$_{\pm1.78}$	**+0.06**$_{\pm.15}$	**+0.10**$_{\pm.11}$	**+0.81**$_{\pm.47}$	**+5.26**$_{\pm.58}$
PatchTST	**+0.78**$_{\pm.32}$	-2.16$_{\pm.18}$	-6.12$_{\pm.11}$	-9.35$_{\pm1.45}$	**+0.90**$_{\pm.06}$	**+1.34**$_{\pm.38}$	**+3.59**$_{\pm1.80}$	**+6.09**$_{\pm.20}$
TimeMixer	**+2.52**$_{\pm.68}$	**+1.24**$_{\pm.31}$	-2.22$_{\pm.76}$	-7.20$_{\pm.51}$	**+1.18**$_{\pm.37}$	**+1.11**$_{\pm.15}$	**+1.76**$_{\pm.20}$	**+6.27**$_{\pm.24}$
TimeXer	**+0.86**$_{\pm.15}$	-0.62$_{\pm.36}$	-4.19$_{\pm1.55}$	-3.73$_{\pm4.13}$	**+1.22**$_{\pm.29}$	**+2.60**$_{\pm.15}$	**+4.73**$_{\pm1.97}$	**+7.22**$_{\pm1.42}$
iTransformer	**+2.17**$_{\pm.72}$	**+0.87**$_{\pm.28}$	-2.34$_{\pm.04}$	-4.25$_{\pm2.98}$	**+2.23**$_{\pm.42}$	**+2.07**$_{\pm.53}$	**+3.86**$_{\pm1.90}$	**+6.07**$_{\pm.39}$
backbone	ETTm1				ETTm2			
	H=96	H=192	H=336	H=720	H=96	H=192	H=336	H=720
DLinear	**+4.44**$_{\pm.08}$	**+2.56**$_{\pm.13}$	**+1.50**$_{\pm.03}$	-0.16$_{\pm.08}$	**+0.57**$_{\pm.07}$	**+1.51**$_{\pm.45}$	**+5.03**$_{\pm4.29}$	**+5.83**$_{\pm6.04}$
DUET	**+4.26**$_{\pm.53}$	**+3.04**$_{\pm.44}$	**+0.61**$_{\pm.28}$	**+2.37**$_{\pm1.62}$	**+1.00**$_{\pm.35}$	**+2.03**$_{\pm.20}$	**+0.58**$_{\pm.21}$	**+0.68**$_{\pm.08}$
FACT	**+2.63**$_{\pm.93}$	**+2.96**$_{\pm.60}$	**+2.02**$_{\pm.71}$	**+5.15**$_{\pm.11}$	**+1.51**$_{\pm.17}$	**+2.30**$_{\pm.46}$	**+1.07**$_{\pm.15}$	**+2.46**$_{\pm1.15}$
FITS	**+7.07**$_{\pm.06}$	**+4.04**$_{\pm.02}$	**+3.10**$_{\pm.01}$	**+4.68**$_{\pm.01}$	**+1.52**$_{\pm.03}$	**+2.08**$_{\pm.07}$	**+0.56**$_{\pm.01}$	**+0.50**$_{\pm.01}$
MixLinear	**+7.89**$_{\pm.18}$	**+4.94**$_{\pm.42}$	**+3.38**$_{\pm.05}$	**+4.92**$_{\pm.31}$	**+2.28**$_{\pm.04}$	**+2.51**$_{\pm.08}$	**+0.67**$_{\pm.04}$	**+0.58**$_{\pm.03}$
PaiFilter	**+2.19**$_{\pm.14}$	**+2.16**$_{\pm.16}$	**+1.12**$_{\pm.09}$	**+1.92**$_{\pm.27}$	**+1.50**$_{\pm.05}$	**+0.82**$_{\pm.75}$	**+0.31**$_{\pm.09}$	**+0.24**$_{\pm.17}$
PatchTST	**+2.59**$_{\pm1.12}$	**+2.16**$_{\pm.46}$	-0.60$_{\pm.07}$	**+0.27**$_{\pm.88}$	**+1.22**$_{\pm.16}$	**+1.68**$_{\pm.09}$	**+0.80**$_{\pm.12}$	**+2.44**$_{\pm1.50}$
TimeMixer	**+2.59**$_{\pm.30}$	**+2.11**$_{\pm.10}$	**+0.39**$_{\pm.21}$	**+1.28**$_{\pm.13}$	**+1.95**$_{\pm.23}$	**+1.54**$_{\pm.54}$	**+0.92**$_{\pm.20}$	**+1.11**$_{\pm.14}$
TimeXer	**+1.88**$_{\pm.76}$	**+1.23**$_{\pm.61}$	**+0.28**$_{\pm.13}$	**+0.92**$_{\pm.15}$	**+0.95**$_{\pm.34}$	**+2.09**$_{\pm1.64}$	**+1.38**$_{\pm.33}$	**+3.18**$_{\pm.80}$
iTransformer	**+3.63**$_{\pm.68}$	**+5.32**$_{\pm.72}$	**+2.44**$_{\pm1.73}$	**+6.31**$_{\pm.51}$	**+2.32**$_{\pm.37}$	**+2.76**$_{\pm.17}$	**+2.93**$_{\pm.14}$	**+3.62**$_{\pm.30}$

Table 26: Complete four-horizon matrix, part 2 of 2 (Weather, ECL, Solar, and Traffic). Each entry is the mean ± sample SD of the paired test-MSE reduction (%, ↑) over seeds $2/3/4$. Bold marks lower test MSE. Datasets are paired horizontally; leading zeros are omitted.

backbone	Weather				ECL			
	H=96	H=192	H=336	H=720	H=96	H=192	H=336	H=720
DLinear	**+16.84**$_{\pm.23}$	**+9.66**$_{\pm.11}$	**+7.05**$_{\pm.23}$	-1.02$_{\pm.14}$	**+10.86**$_{\pm.02}$	**+11.93**$_{\pm.00}$	**+13.05**$_{\pm.05}$	**+13.47**$_{\pm.03}$
DUET	**+3.10**$_{\pm.53}$	**+4.53**$_{\pm.22}$	**+4.77**$_{\pm.06}$	**+0.17**$_{\pm.19}$	**+21.15**$_{\pm.18}$	**+20.74**$_{\pm.45}$	**+21.27**$_{\pm.06}$	**+19.76**$_{\pm.59}$
FACT	**+2.87**$_{\pm1.42}$	**+2.87**$_{\pm.55}$	**+5.73**$_{\pm.30}$	**+1.29**$_{\pm.88}$	**+0.43**$_{\pm.16}$	**+0.59**$_{\pm.99}$	-2.10$_{\pm.42}$	-3.54$_{\pm.78}$
FITS	**+15.59**$_{\pm.10}$	**+10.49**$_{\pm.12}$	**+8.74**$_{\pm.26}$	**+11.06**$_{\pm.15}$	**+12.23**$_{\pm.08}$	**+13.09**$_{\pm.05}$	**+14.07**$_{\pm.04}$	**+15.76**$_{\pm.01}$
MixLinear	**+15.97**$_{\pm.31}$	**+10.85**$_{\pm.16}$	**+9.72**$_{\pm.28}$	**+11.53**$_{\pm.20}$	**+16.36**$_{\pm5.28}$	**+16.20**$_{\pm.39}$	**+17.26**$_{\pm.76}$	**+19.95**$_{\pm.78}$
PaiFilter	**+14.22**$_{\pm1.30}$	**+10.57**$_{\pm.65}$	**+8.36**$_{\pm.96}$	**+1.70**$_{\pm1.25}$	**+13.91**$_{\pm.55}$	**+13.38**$_{\pm.17}$	**+14.88**$_{\pm.22}$	**+17.17**$_{\pm.13}$
PatchTST	**+10.67**$_{\pm.64}$	**+8.08**$_{\pm.23}$	**+7.15**$_{\pm.26}$	**+1.45**$_{\pm.09}$	**+15.42**$_{\pm.29}$	**+13.91**$_{\pm.15}$	**+15.90**$_{\pm.10}$	**+17.85**$_{\pm.09}$
TimeMixer	**+9.59**$_{\pm.49}$	**+7.63**$_{\pm.47}$	**+5.15**$_{\pm.49}$	**+0.99**$_{\pm.27}$	**+10.68**$_{\pm.26}$	**+11.10**$_{\pm.20}$	**+12.58**$_{\pm.11}$	**+7.96**$_{\pm6.84}$
TimeXer	-0.25$_{\pm.32}$	**+0.89**$_{\pm.30}$	**+3.70**$_{\pm.54}$	-1.35$_{\pm.79}$	**+6.15**$_{\pm.02}$	**+7.18**$_{\pm.12}$	**+8.98**$_{\pm.35}$	**+5.53**$_{\pm4.36}$
iTransformer	**+8.16**$_{\pm.43}$	**+7.94**$_{\pm.12}$	**+8.86**$_{\pm.65}$	**+1.96**$_{\pm1.18}$	**+6.43**$_{\pm.41}$	**+6.32**$_{\pm.27}$	**+6.41**$_{\pm2.41}$	-2.15$_{\pm.25}$
backbone	Solar				Traffic			
	H=96	H=192	H=336	H=720	H=96	H=192	H=336	H=720
DLinear	**+24.88**$_{\pm.00}$	**+24.13**$_{\pm.03}$	**+20.58**$_{\pm.01}$	**+18.14**$_{\pm.02}$	**+14.32**$_{\pm.04}$	**+11.13**$_{\pm.01}$	**+11.83**$_{\pm.05}$	**+2.74**$_{\pm.02}$
DUET	**+25.39**$_{\pm.74}$	**+22.98**$_{\pm.17}$	**+21.99**$_{\pm.17}$	**+18.69**$_{\pm.51}$	**+12.67**$_{\pm.04}$	**+11.33**$_{\pm.02}$	**+11.14**$_{\pm.26}$	**+8.67**$_{\pm.23}$
FACT	**+9.02**$_{\pm.53}$	**+8.90**$_{\pm.54}$	**+11.75**$_{\pm.30}$	-6.90$_{\pm.16}$	**+4.06**$_{\pm.14}$	**+3.24**$_{\pm.29}$	**+3.46**$_{\pm.26}$	**+2.57**$_{\pm.34}$
FITS	**+31.99**$_{\pm.05}$	**+32.29**$_{\pm.07}$	**+30.99**$_{\pm.13}$	**+30.48**$_{\pm.13}$	**+15.83**$_{\pm.02}$	**+13.50**$_{\pm.04}$	**+13.81**$_{\pm.02}$	**+2.09**$_{\pm.01}$
MixLinear	**+34.27**$_{\pm.83}$	**+35.62**$_{\pm.04}$	**+32.60**$_{\pm.46}$	**+30.54**$_{\pm.37}$	**+16.47**$_{\pm.75}$	**+18.05**$_{\pm.50}$	**+17.25**$_{\pm.08}$	**+19.48**$_{\pm.47}$
PaiFilter	**+24.63**$_{\pm.47}$	**+27.38**$_{\pm.44}$	**+26.63**$_{\pm.38}$	**+25.26**$_{\pm.43}$	**+8.77**$_{\pm.08}$	**+7.98**$_{\pm.13}$	**+8.29**$_{\pm.07}$	-0.15$_{\pm.17}$
PatchTST	**+20.23**$_{\pm.70}$	**+20.63**$_{\pm.29}$	**+19.29**$_{\pm1.04}$	**+15.38**$_{\pm.41}$	**+5.29**$_{\pm.05}$	**+5.12**$_{\pm.06}$	**+3.79**$_{\pm.05}$	**+2.58**$_{\pm.19}$
TimeMixer	**+20.56**$_{\pm.63}$	**+22.53**$_{\pm.83}$	**+20.19**$_{\pm.51}$	**+16.68**$_{\pm.31}$	**+5.66**$_{\pm.12}$	**+5.54**$_{\pm.40}$	**+5.79**$_{\pm.45}$	-0.86$_{\pm1.17}$
TimeXer	**+16.03**$_{\pm.64}$	**+17.15**$_{\pm.85}$	**+17.30**$_{\pm.24}$	**+15.66**$_{\pm.60}$	**+4.77**$_{\pm.10}$	**+3.73**$_{\pm.05}$	**+3.78**$_{\pm.32}$	-7.04$_{\pm1.06}$
iTransformer	**+16.43**$_{\pm1.29}$	**+19.10**$_{\pm.63}$	**+19.03**$_{\pm.40}$	**+15.49**$_{\pm.45}$	**+3.75**$_{\pm.04}$	**+1.61**$_{\pm.06}$	**+1.20**$_{\pm.02}$	**+0.03**$_{\pm.05}$

Table 27: Horizon-specific win counts corresponding to Table 2. A win is a positive mean paired test-MSE reduction over seeds $2/3/4$. Each horizon contains eight datasets per backbone.

H	DLinear	DUET	FACT	FITS	MixLinear	PaiFilter	PatchTST	TimeMixer	TimeXer	iTransformer	all
96	7/8	8/8	8/8	8/8	8/8	7/8	8/8	8/8	7/8	8/8	77/80
192	7/8	7/8	8/8	7/8	7/8	7/8	7/8	8/8	7/8	8/8	73/80
336	7/8	7/8	6/8	7/8	7/8	7/8	6/8	7/8	7/8	7/8	68/80
720	5/8	7/8	5/8	7/8	7/8	6/8	7/8	6/8	5/8	6/8	61/80
all	26/32	29/32	27/32	29/32	29/32	27/32	28/32	29/32	26/32	29/32	279/320

Table 28: Complete 80-cell results at $H{=}96$. Each backbone has paired MSE/MAE columns; the two panels cover all ten backbones. Entries are mean $\pm$ sample SD over seeds $2/3/4$; SD is smaller and leading zeros are omitted. Δ is the seed-wise relative reduction (%, $\uparrow$), computed from unrounded metrics. Bold marks lower base/ours error and positive reductions.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.3963**$_{\pm.0004}$	**.4110**$_{\pm.0004}$	.3822$_{\pm.0019}$	**.3997**$_{\pm.0012}$	.3821$_{\pm.0029}$	.4014$_{\pm.0025}$	.3910$_{\pm.0030}$	**.4012**$_{\pm.0031}$	.3928$_{\pm.0029}$	**.3990**$_{\pm.0027}$
	ours	.3976$_{\pm.0003}$	.4163$_{\pm.0004}$	**.3821**$_{\pm.0016}$	.4010$_{\pm.0011}$	**.3741**$_{\pm.0016}$	**.4014**$_{\pm.0037}$	**.3875**$_{\pm.0024}$	.4020$_{\pm.0018}$	**.3900**$_{\pm.0031}$	.4025$_{\pm.0034}$
	$\Delta\uparrow$ (%)	-0.32$_{\pm.01}$	-1.28$_{\pm.02}$	**+0.01**$_{\pm.10}$	-0.31$_{\pm.09}$	**+2.11**$_{\pm.65}$	**+0.01**$_{\pm.41}$	**+0.89**$_{\pm.17}$	-0.19$_{\pm.31}$	**+0.73**$_{\pm.20}$	-0.88$_{\pm.19}$
ETTh2	base	.3458$_{\pm.0061}$	.3997$_{\pm.0041}$	.2926$_{\pm.0009}$	.3435$_{\pm.0010}$	.2936$_{\pm.0039}$	.3421$_{\pm.0030}$	.2992$_{\pm.0009}$	.3459$_{\pm.0007}$	.3001$_{\pm.0039}$	.3478$_{\pm.0037}$
	ours	**.3117**$_{\pm.0016}$	**.3703**$_{\pm.0017}$	**.2901**$_{\pm.0016}$	**.3426**$_{\pm.0017}$	**.2871**$_{\pm.0018}$	**.3374**$_{\pm.0008}$	**.2968**$_{\pm.0008}$	**.3444**$_{\pm.0007}$	**.2964**$_{\pm.0027}$	**.3450**$_{\pm.0025}$
	$\Delta\uparrow$ (%)	**+9.82**$_{\pm1.67}$	**+7.36**$_{\pm1.09}$	**+0.85**$_{\pm.24}$	**+0.26**$_{\pm.23}$	**+2.20**$_{\pm.85}$	**+1.38**$_{\pm.63}$	**+0.79**$_{\pm.05}$	**+0.43**$_{\pm.02}$	**+1.22**$_{\pm.39}$	**+0.79**$_{\pm.35}$
ETTm1	base	.3456$_{\pm.0002}$	.3721$_{\pm.0001}$	.3323$_{\pm.0025}$	.3683$_{\pm.0007}$	.3211$_{\pm.0042}$	**.3606**$_{\pm.0023}$	.3547$_{\pm.0003}$	.3751$_{\pm.0003}$	.3695$_{\pm.0029}$	.3848$_{\pm.0019}$
	ours	**.3302**$_{\pm.0003}$	**.3620**$_{\pm.0003}$	**.3181**$_{\pm.0007}$	**.3603**$_{\pm.0013}$	**.3126**$_{\pm.0016}$	.3633$_{\pm.0026}$	**.3296**$_{\pm.0001}$	**.3599**$_{\pm.0001}$	**.3404**$_{\pm.0027}$	**.3681**$_{\pm.0022}$
	$\Delta\uparrow$ (%)	**+4.44**$_{\pm.08}$	**+2.72**$_{\pm.05}$	**+4.26**$_{\pm.53}$	**+2.20**$_{\pm.27}$	**+2.63**$_{\pm.93}$	-0.75$_{\pm.09}$	**+7.07**$_{\pm.06}$	**+4.05**$_{\pm.04}$	**+7.89**$_{\pm.18}$	**+4.34**$_{\pm.18}$
ETTm2	base	.1926$_{\pm.0025}$	.2912$_{\pm.0039}$	.1780$_{\pm.0030}$	.2602$_{\pm.0032}$	.1676$_{\pm.0004}$	.2504$_{\pm.0004}$	.1831$_{\pm.0002}$	.2663$_{\pm.0000}$	.1887$_{\pm.0014}$	.2718$_{\pm.0016}$
	ours	**.1915**$_{\pm.0024}$	**.2898**$_{\pm.0037}$	**.1762**$_{\pm.0023}$	**.2592**$_{\pm.0016}$	**.1650**$_{\pm.0005}$	**.2485**$_{\pm.0005}$	**.1803**$_{\pm.0002}$	**.2633**$_{\pm.0001}$	**.1844**$_{\pm.0015}$	**.2677**$_{\pm.0016}$
	$\Delta\uparrow$ (%)	**+0.57**$_{\pm.07}$	**+0.49**$_{\pm.07}$	**+1.00**$_{\pm.35}$	**+0.38**$_{\pm.65}$	**+1.51**$_{\pm.17}$	**+0.75**$_{\pm.12}$	**+1.52**$_{\pm.03}$	**+1.15**$_{\pm.04}$	**+2.28**$_{\pm.04}$	**+1.51**$_{\pm.02}$
Weather	base	.1958$_{\pm.0003}$	.2560$_{\pm.0011}$	.1671$_{\pm.0001}$	**.2131**$_{\pm.0004}$	.1566$_{\pm.0024}$	**.2006**$_{\pm.0025}$	.1969$_{\pm.0001}$	**.2371**$_{\pm.0001}$	.1994$_{\pm.0010}$	**.2393**$_{\pm.0005}$
	ours	**.1628**$_{\pm.0006}$	**.2441**$_{\pm.0009}$	**.1619**$_{\pm.0008}$	.2385$_{\pm.0008}$	**.1521**$_{\pm.0045}$	.2150$_{\pm.0068}$	**.1662**$_{\pm.0001}$	.2491$_{\pm.0002}$	**.1676**$_{\pm.0002}$	.2508$_{\pm.0003}$
	$\Delta\uparrow$ (%)	**+16.84**$_{\pm.23}$	**+4.65**$_{\pm.12}$	**+3.10**$_{\pm.53}$	-11.91$_{\pm.17}$	**+2.87**$_{\pm1.42}$	-7.15$_{\pm2.02}$	**+15.59**$_{\pm.10}$	-5.07$_{\pm.09}$	**+15.97**$_{\pm.31}$	-4.81$_{\pm.16}$
ECL	base	.1946$_{\pm.0001}$	.2774$_{\pm.0006}$	.1668$_{\pm.0002}$	.2846$_{\pm.0002}$	.1372$_{\pm.0004}$	**.2336**$_{\pm.0005}$	.2033$_{\pm.0002}$	.2835$_{\pm.0004}$	.2172$_{\pm.0191}$	.2972$_{\pm.0268}$
	ours	**.1735**$_{\pm.0000}$	**.2736**$_{\pm.0001}$	**.1315**$_{\pm.0002}$	**.2314**$_{\pm.0003}$	**.1367**$_{\pm.0005}$	.2360$_{\pm.0005}$	**.1785**$_{\pm.0000}$	**.2796**$_{\pm.0000}$	**.1810**$_{\pm.0040}$	**.2818**$_{\pm.0065}$
	$\Delta\uparrow$ (%)	**+10.86**$_{\pm.02}$	**+1.36**$_{\pm.16}$	**+21.15**$_{\pm.18}$	**+18.71**$_{\pm.15}$	**+0.43**$_{\pm.16}$	-0.99$_{\pm.06}$	**+12.23**$_{\pm.08}$	**+1.36**$_{\pm.12}$	**+16.36**$_{\pm5.28}$	**+4.84**$_{\pm6.10}$
Solar	base	.2900$_{\pm.0001}$	.3775$_{\pm.0002}$	.2345$_{\pm.0099}$	.3070$_{\pm.0029}$	.1939$_{\pm.0022}$	.2410$_{\pm.0017}$	.3415$_{\pm.0012}$	.3699$_{\pm.0010}$	.3646$_{\pm.0146}$	.3734$_{\pm.0078}$
	ours	**.2178**$_{\pm.0001}$	**.3251**$_{\pm.0002}$	**.1750**$_{\pm.0063}$	**.2531**$_{\pm.0057}$	**.1764**$_{\pm.0011}$	**.2381**$_{\pm.0016}$	**.2322**$_{\pm.0010}$	**.3172**$_{\pm.0008}$	**.2395**$_{\pm.0066}$	**.3277**$_{\pm.0055}$
	$\Delta\uparrow$ (%)	**+24.88**$_{\pm.00}$	**+13.89**$_{\pm.03}$	**+25.39**$_{\pm.74}$	**+17.56**$_{\pm1.18}$	**+9.02**$_{\pm.53}$	**+1.22**$_{\pm.19}$	**+31.99**$_{\pm.05}$	**+14.25**$_{\pm.08}$	**+34.27**$_{\pm.83}$	**+12.22**$_{\pm.41}$
Traffic	base	.6484$_{\pm.0003}$	.3961$_{\pm.0004}$	.4580$_{\pm.0034}$	.3142$_{\pm.0006}$	.4042$_{\pm.0018}$	**.2593**$_{\pm.0010}$	.6540$_{\pm.0003}$	.3926$_{\pm.0008}$	.6795$_{\pm.0053}$	.4101$_{\pm.0032}$
	ours	**.5556**$_{\pm.0005}$	**.3707**$_{\pm.0003}$	**.3999**$_{\pm.0028}$	**.2740**$_{\pm.0012}$	**.3878**$_{\pm.0011}$	.2614$_{\pm.0010}$	**.5504**$_{\pm.0003}$	**.3730**$_{\pm.0003}$	**.5676**$_{\pm.0073}$	**.3822**$_{\pm.0056}$
	$\Delta\uparrow$ (%)	**+14.32**$_{\pm.04}$	**+6.43**$_{\pm.01}$	**+12.67**$_{\pm.04}$	**+12.78**$_{\pm.22}$	**+4.06**$_{\pm.14}$	-0.79$_{\pm.03}$	**+15.83**$_{\pm.02}$	**+4.99**$_{\pm.11}$	**+16.47**$_{\pm.75}$	**+6.81**$_{\pm.80}$

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.4409**$_{\pm.0041}$	**.4408**$_{\pm.0027}$	.3833$_{\pm.0007}$	**.3982**$_{\pm.0006}$	.4067$_{\pm.0023}$	**.4161**$_{\pm.0022}$	.3869$_{\pm.0011}$	**.4045**$_{\pm.0006}$	.3898$_{\pm.0021}$	**.4076**$_{\pm.0017}$
	ours	.4442$_{\pm.0040}$	.4485$_{\pm.0027}$	**.3803**$_{\pm.0011}$	.4007$_{\pm.0007}$	**.3964**$_{\pm.0027}$	.4173$_{\pm.0006}$	**.3836**$_{\pm.0005}$	.4052$_{\pm.0003}$	**.3813**$_{\pm.0015}$	.4078$_{\pm.0018}$
	$\Delta\uparrow$ (%)	-0.73$_{\pm.05}$	-1.76$_{\pm.05}$	**+0.78**$_{\pm.32}$	-0.61$_{\pm.34}$	**+2.52**$_{\pm.68}$	-0.28$_{\pm.59}$	**+0.86**$_{\pm.15}$	-0.16$_{\pm.17}$	**+2.17**$_{\pm.72}$	-0.06$_{\pm.09}$
ETTh2	base	.3190$_{\pm.0149}$	**.3650**$_{\pm.0128}$	.2871$_{\pm.0011}$	.3399$_{\pm.0006}$	.2969$_{\pm.0043}$	.3471$_{\pm.0032}$	.2899$_{\pm.0023}$	.3419$_{\pm.0022}$	.3091$_{\pm.0083}$	.3547$_{\pm.0041}$
	ours	**.3188**$_{\pm.0153}$	.3656$_{\pm.0135}$	**.2845**$_{\pm.0012}$	**.3382**$_{\pm.0008}$	**.2934**$_{\pm.0031}$	**.3451**$_{\pm.0022}$	**.2863**$_{\pm.0015}$	**.3392**$_{\pm.0018}$	**.3023**$_{\pm.0079}$	**.3489**$_{\pm.0039}$
	$\Delta\uparrow$ (%)	**+0.06**$_{\pm.15}$	-0.17$_{\pm.18}$	**+0.90**$_{\pm.06}$	**+0.49**$_{\pm.07}$	**+1.18**$_{\pm.37}$	**+0.57**$_{\pm.38}$	**+1.22**$_{\pm.29}$	**+0.77**$_{\pm.26}$	**+2.23**$_{\pm.42}$	**+1.63**$_{\pm.24}$
ETTm1	base	.3569$_{\pm.0040}$	**.3797**$_{\pm.0007}$	.3337$_{\pm.0038}$	.3690$_{\pm.0003}$	.3339$_{\pm.0038}$	**.3712**$_{\pm.0014}$	.3233$_{\pm.0005}$	.3604$_{\pm.0010}$	.3426$_{\pm.0047}$	.3752$_{\pm.0037}$
	ours	**.3491**$_{\pm.0035}$	.3824$_{\pm.0003}$	**.3250**$_{\pm.0074}$	**.3686**$_{\pm.0069}$	**.3253**$_{\pm.0035}$	.3727$_{\pm.0017}$	**.3172**$_{\pm.0026}$	**.3603**$_{\pm.0045}$	**.3301**$_{\pm.0034}$	**.3736**$_{\pm.0030}$
	$\Delta\uparrow$ (%)	**+2.19**$_{\pm.14}$	-0.70$_{\pm.13}$	**+2.59**$_{\pm1.12}$	**+0.11**$_{\pm1.80}$	**+2.59**$_{\pm.30}$	-0.41$_{\pm.31}$	**+1.88**$_{\pm.76}$	**+0.03**$_{\pm1.23}$	**+3.63**$_{\pm.68}$	**+0.43**$_{\pm.22}$
ETTm2	base	.1930$_{\pm.0007}$	.2772$_{\pm.0008}$	.1756$_{\pm.0006}$	.2604$_{\pm.0009}$	.1774$_{\pm.0018}$	.2623$_{\pm.0021}$	.1726$_{\pm.0002}$	.2570$_{\pm.0011}$	.1848$_{\pm.0026}$	.2669$_{\pm.0025}$
	ours	**.1901**$_{\pm.0006}$	**.2742**$_{\pm.0007}$	**.1735**$_{\pm.0009}$	**.2602**$_{\pm.0011}$	**.1740**$_{\pm.0015}$	**.2609**$_{\pm.0016}$	**.1709**$_{\pm.0007}$	**.2565**$_{\pm.0009}$	**.1806**$_{\pm.0032}$	**.2640**$_{\pm.0032}$
	$\Delta\uparrow$ (%)	**+1.50**$_{\pm.05}$	**+1.05**$_{\pm.02}$	**+1.22**$_{\pm.16}$	**+0.07**$_{\pm.11}$	**+1.95**$_{\pm.23}$	**+0.52**$_{\pm.31}$	**+0.95**$_{\pm.34}$	**+0.19**$_{\pm.16}$	**+2.32**$_{\pm.37}$	**+1.05**$_{\pm.30}$
Weather	base	.1975$_{\pm.0024}$	**.2405**$_{\pm.0037}$	.1768$_{\pm.0011}$	**.2180**$_{\pm.0007}$	.1781$_{\pm.0023}$	**.2240**$_{\pm.0023}$	**.1587**$_{\pm.0007}$	**.2075**$_{\pm.0005}$	.1732$_{\pm.0011}$	**.2132**$_{\pm.0009}$
	ours	**.1694**$_{\pm.0046}$	.2517$_{\pm.0059}$	**.1579**$_{\pm.0002}$	.2367$_{\pm.0005}$	**.1610**$_{\pm.0012}$	.2414$_{\pm.0012}$	.1591$_{\pm.0012}$	.2364$_{\pm.0015}$	**.1591**$_{\pm.0015}$	.2342$_{\pm.0025}$
	$\Delta\uparrow$ (%)	**+14.22**$_{\pm1.30}$	-4.64$_{\pm.80}$	**+10.67**$_{\pm.64}$	-8.57$_{\pm.25}$	**+9.59**$_{\pm.49}$	-7.77$_{\pm.59}$	-0.25$_{\pm.32}$	-13.90$_{\pm.51}$	**+8.16**$_{\pm.43}$	-9.89$_{\pm.71}$
ECL	base	.1864$_{\pm.0008}$	.2738$_{\pm.0015}$	.1768$_{\pm.0004}$	.2710$_{\pm.0006}$	.1698$_{\pm.0006}$	.2609$_{\pm.0006}$	.1450$_{\pm.0004}$	.2436$_{\pm.0003}$	.1481$_{\pm.0003}$	.2400$_{\pm.0003}$
	ours	**.1605**$_{\pm.0006}$	**.2632**$_{\pm.0015}$	**.1495**$_{\pm.0003}$	**.2504**$_{\pm.0004}$	**.1517**$_{\pm.0008}$	**.2542**$_{\pm.0020}$	**.1361**$_{\pm.0004}$	**.2355**$_{\pm.0002}$	**.1386**$_{\pm.0004}$	**.2375**$_{\pm.0003}$
	$\Delta\uparrow$ (%)	**+13.91**$_{\pm.55}$	**+3.89**$_{\pm.93}$	**+15.42**$_{\pm.29}$	**+7.59**$_{\pm.30}$	**+10.68**$_{\pm.26}$	**+2.56**$_{\pm.53}$	**+6.15**$_{\pm.02}$	**+3.31**$_{\pm.08}$	**+6.43**$_{\pm.41}$	**+1.04**$_{\pm.17}$
Solar	base	.3088$_{\pm.0015}$	.3160$_{\pm.0058}$	.2398$_{\pm.0033}$	.2885$_{\pm.0019}$	.2506$_{\pm.0024}$	.2939$_{\pm.0021}$	.2123$_{\pm.0027}$	.2709$_{\pm.0027}$	.2117$_{\pm.0041}$	.2545$_{\pm.0023}$
	ours	**.2327**$_{\pm.0010}$	**.2982**$_{\pm.0018}$	**.1913**$_{\pm.0009}$	**.2665**$_{\pm.0005}$	**.1991**$_{\pm.0003}$	**.2825**$_{\pm.0039}$	**.1783**$_{\pm.0036}$	**.2548**$_{\pm.0029}$	**.1770**$_{\pm.0062}$	**.2479**$_{\pm.0029}$
	$\Delta\uparrow$ (%)	**+24.63**$_{\pm.47}$	**+5.61**$_{\pm1.23}$	**+20.23**$_{\pm.70}$	**+7.64**$_{\pm.45}$	**+20.56**$_{\pm.63}$	**+3.86**$_{\pm1.99}$	**+16.03**$_{\pm.64}$	**+5.93**$_{\pm.95}$	**+16.43**$_{\pm1.29}$	**+2.58**$_{\pm.45}$
Traffic	base	.5462$_{\pm.0010}$	.3555$_{\pm.0015}$	.4465$_{\pm.0012}$	.2908$_{\pm.0005}$	.4937$_{\pm.0049}$	.3253$_{\pm.0015}$	.4450$_{\pm.0018}$	.2929$_{\pm.0011}$	.3931$_{\pm.0007}$	.2684$_{\pm.0006}$
	ours	**.4983**$_{\pm.0006}$	**.3413**$_{\pm.0007}$	**.4229**$_{\pm.0010}$	**.2777**$_{\pm.0005}$	**.4657**$_{\pm.0041}$	**.3154**$_{\pm.0010}$	**.4238**$_{\pm.0017}$	**.2865**$_{\pm.0013}$	**.3784**$_{\pm.0007}$	**.2628**$_{\pm.0006}$
	$\Delta\uparrow$ (%)	**+8.77**$_{\pm.08}$	**+4.00**$_{\pm.28}$	**+5.29**$_{\pm.05}$	**+4.48**$_{\pm.11}$	**+5.66**$_{\pm.12}$	**+3.04**$_{\pm.16}$	**+4.77**$_{\pm.10}$	**+2.20**$_{\pm.17}$	**+3.75**$_{\pm.04}$	**+2.08**$_{\pm.06}$

Table 29: Complete 80-cell results at $H=192$. Each backbone has paired MSE/MAE columns; the two panels cover all ten backbones. Entries are mean ± sample SD over seeds $2/3/4$; SD is smaller and leading zeros are omitted. Δ is the seed-wise relative reduction (%, ↑), computed from unrounded metrics. Bold marks lower base/ours error and positive reductions.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.4450**$_{\pm.0004}$	**.4404**$_{\pm.0005}$	**.4281**$_{\pm.0003}$	**.4284**$_{\pm.0008}$	.4452$_{\pm.0029}$	**.4389**$_{\pm.0047}$	**.4391**$_{\pm.0039}$	**.4297**$_{\pm.0044}$	**.4424**$_{\pm.0046}$	**.4281**$_{\pm.0044}$
	ours	.4604$_{\pm.0004}$	.4609$_{\pm.0004}$	.4419$_{\pm.0015}$	.4429$_{\pm.0002}$	**.4422**$_{\pm.0016}$	.4511$_{\pm.0033}$	.4446$_{\pm.0037}$	.4453$_{\pm.0020}$	.4458$_{\pm.0025}$	.4465$_{\pm.0026}$
	$\Delta\uparrow$ (%)	-3.45$_{\pm.02}$	-4.67$_{\pm.02}$	-3.22$_{\pm.28}$	-3.36$_{\pm.22}$	**+0.66**$_{\pm.30}$	-2.78$_{\pm.37}$	-1.25$_{\pm.13}$	-3.64$_{\pm.62}$	-0.78$_{\pm.63}$	-4.28$_{\pm.58}$
ETTh2	base	.4760$_{\pm.0019}$	.4766$_{\pm.0016}$	.3740$_{\pm.0008}$	.3932$_{\pm.0009}$	.3734$_{\pm.0032}$	.3926$_{\pm.0030}$	.3846$_{\pm.0008}$	.3962$_{\pm.0005}$	.3869$_{\pm.0036}$	.3994$_{\pm.0031}$
	ours	**.4689**$_{\pm.0025}$	**.4765**$_{\pm.0016}$	**.3688**$_{\pm.0030}$	**.3902**$_{\pm.0012}$	**.3639**$_{\pm.0019}$	**.3865**$_{\pm.0023}$	**.3831**$_{\pm.0007}$	**.3949**$_{\pm.0005}$	**.3836**$_{\pm.0033}$	**.3968**$_{\pm.0028}$
	$\Delta\uparrow$ (%)	**+1.48**$_{\pm.16}$	**+0.02**$_{\pm.11}$	**+1.38**$_{\pm.63}$	**+0.76**$_{\pm.25}$	**+2.54**$_{\pm.43}$	**+1.56**$_{\pm.60}$	**+0.39**$_{\pm.02}$	**+0.34**$_{\pm.01}$	**+0.86**$_{\pm.18}$	**+0.64**$_{\pm.14}$
ETTm1	base	.3819$_{\pm.0003}$	.3911$_{\pm.0006}$	.3690$_{\pm.0032}$	.3900$_{\pm.0019}$	.3662$_{\pm.0053}$	.3866$_{\pm.0028}$	.3923$_{\pm.0001}$	**.3930**$_{\pm.0001}$	.4146$_{\pm.0032}$	.4094$_{\pm.0020}$
	ours	**.3721**$_{\pm.0003}$	**.3899**$_{\pm.0047}$	**.3578**$_{\pm.0017}$	**.3839**$_{\pm.0010}$	**.3554**$_{\pm.0043}$	**.3862**$_{\pm.0022}$	**.3764**$_{\pm.0001}$	.3941$_{\pm.0001}$	**.3941**$_{\pm.0017}$	**.4087**$_{\pm.0012}$
	$\Delta\uparrow$ (%)	**+2.56**$_{\pm.13}$	**+0.32**$_{\pm1.34}$	**+3.04**$_{\pm.44}$	**+1.56**$_{\pm.34}$	**+2.96**$_{\pm.60}$	**+0.12**$_{\pm.33}$	**+4.04**$_{\pm.02}$	-0.27$_{\pm.02}$	**+4.94**$_{\pm.42}$	**+0.19**$_{\pm.25}$
ETTm2	base	.2805$_{\pm.0029}$	.3576$_{\pm.0040}$	.2496$_{\pm.0021}$	.3079$_{\pm.0015}$	.2357$_{\pm.0015}$	.2967$_{\pm.0018}$	.2472$_{\pm.0002}$	.3051$_{\pm.0002}$	.2519$_{\pm.0006}$	.3100$_{\pm.0005}$
	ours	**.2763**$_{\pm.0041}$	**.3566**$_{\pm.0046}$	**.2445**$_{\pm.0021}$	**.3055**$_{\pm.0011}$	**.2303**$_{\pm.0023}$	**.2940**$_{\pm.0021}$	**.2421**$_{\pm.0001}$	**.3030**$_{\pm.0002}$	**.2456**$_{\pm.0004}$	**.3073**$_{\pm.0004}$
	$\Delta\uparrow$ (%)	**+1.51**$_{\pm.45}$	**+0.29**$_{\pm.18}$	**+2.03**$_{\pm.20}$	**+0.76**$_{\pm.37}$	**+2.30**$_{\pm.46}$	**+0.91**$_{\pm.29}$	**+2.08**$_{\pm.07}$	**+0.67**$_{\pm.03}$	**+2.51**$_{\pm.08}$	**+0.86**$_{\pm.06}$
Weather	base	.2365$_{\pm.0003}$	**.2948**$_{\pm.0007}$	.2158$_{\pm.0005}$	**.2556**$_{\pm.0002}$	.2030$_{\pm.0008}$	**.2445**$_{\pm.0001}$	.2416$_{\pm.0000}$	**.2725**$_{\pm.0000}$	.2449$_{\pm.0001}$	**.2753**$_{\pm.0003}$
	ours	**.2136**$_{\pm.0001}$	.2949$_{\pm.0001}$	**.2060**$_{\pm.0007}$	.2835$_{\pm.0003}$	**.1972**$_{\pm.0010}$	.2616$_{\pm.0013}$	**.2163**$_{\pm.0003}$	.2991$_{\pm.0002}$	**.2184**$_{\pm.0003}$	.3018$_{\pm.0003}$
	$\Delta\uparrow$ (%)	**+9.66**$_{\pm.11}$	-0.05$_{\pm.24}$	**+4.53**$_{\pm.22}$	-10.88$_{\pm.11}$	**+2.87**$_{\pm.55}$	-7.01$_{\pm.56}$	**+10.49**$_{\pm.12}$	-9.73$_{\pm.08}$	**+10.85**$_{\pm.16}$	-9.63$_{\pm.21}$
ECL	base	.1940$_{\pm.0000}$	.2806$_{\pm.0001}$	.1840$_{\pm.0006}$	.2973$_{\pm.0008}$	.1565$_{\pm.0013}$	**.2512**$_{\pm.0018}$	.2026$_{\pm.0002}$	.2854$_{\pm.0003}$	.2126$_{\pm.0026}$	.2908$_{\pm.0015}$
	ours	**.1708**$_{\pm.0000}$	**.2723**$_{\pm.0000}$	**.1459**$_{\pm.0004}$	**.2449**$_{\pm.0009}$	**.1556**$_{\pm.0028}$	.2533$_{\pm.0046}$	**.1761**$_{\pm.0001}$	**.2779**$_{\pm.0001}$	**.1782**$_{\pm.0023}$	**.2795**$_{\pm.0017}$
	$\Delta\uparrow$ (%)	**+11.93**$_{\pm.00}$	**+2.96**$_{\pm.01}$	**+20.74**$_{\pm.45}$	**+17.62**$_{\pm.52}$	**+0.59**$_{\pm.99}$	-0.81$_{\pm1.13}$	**+13.09**$_{\pm.05}$	**+2.61**$_{\pm.08}$	**+16.20**$_{\pm.39}$	**+3.90**$_{\pm.32}$
Solar	base	.3198$_{\pm.0001}$	.3978$_{\pm.0001}$	.2574$_{\pm.0008}$	.3234$_{\pm.0014}$	.2224$_{\pm.0001}$	.2637$_{\pm.0010}$	.3825$_{\pm.0003}$	.3942$_{\pm.0010}$	.3983$_{\pm.0012}$	.3944$_{\pm.0023}$
	ours	**.2427**$_{\pm.0002}$	**.3383**$_{\pm.0002}$	**.1983**$_{\pm.0010}$	**.2758**$_{\pm.0006}$	**.2026**$_{\pm.0012}$	**.2613**$_{\pm.0011}$	**.2590**$_{\pm.0001}$	**.3613**$_{\pm.0005}$	**.2564**$_{\pm.0006}$	**.3599**$_{\pm.0008}$
	$\Delta\uparrow$ (%)	**+24.13**$_{\pm.03}$	**+14.97**$_{\pm.06}$	**+22.98**$_{\pm.17}$	**+14.69**$_{\pm.33}$	**+8.90**$_{\pm.54}$	**+0.91**$_{\pm.38}$	**+32.29**$_{\pm.07}$	**+8.35**$_{\pm.33}$	**+35.62**$_{\pm.04}$	**+8.75**$_{\pm.55}$
Traffic	base	.5984$_{\pm.0001}$	.3707$_{\pm.0001}$	.4850$_{\pm.0012}$	.3246$_{\pm.0003}$	.4233$_{\pm.0031}$	**.2710**$_{\pm.0004}$	.6052$_{\pm.0005}$	.3678$_{\pm.0008}$	.6604$_{\pm.0057}$	.4132$_{\pm.0055}$
	ours	**.5318**$_{\pm.0001}$	**.3489**$_{\pm.0001}$	**.4300**$_{\pm.0011}$	**.2860**$_{\pm.0002}$	**.4095**$_{\pm.0017}$	.2738$_{\pm.0003}$	**.5235**$_{\pm.0002}$	**.3481**$_{\pm.0002}$	**.5412**$_{\pm.0015}$	**.3634**$_{\pm.0009}$
	$\Delta\uparrow$ (%)	**+11.13**$_{\pm.01}$	**+5.90**$_{\pm.02}$	**+11.33**$_{\pm.02}$	**+11.88**$_{\pm.12}$	**+3.24**$_{\pm.29}$	-1.06$_{\pm.05}$	**+13.50**$_{\pm.04}$	**+5.34**$_{\pm.17}$	**+18.05**$_{\pm.50}$	**+12.06**$_{\pm.94}$

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.4884**$_{\pm.0079}$	**.4658**$_{\pm.0054}$	**.4314**$_{\pm.0011}$	**.4275**$_{\pm.0004}$	.4744$_{\pm.0114}$	**.4553**$_{\pm.0072}$	**.4370**$_{\pm.0005}$	**.4342**$_{\pm.0010}$	.4462$_{\pm.0025}$	**.4391**$_{\pm.0023}$
	ours	.5001$_{\pm.0079}$	.4895$_{\pm.0049}$	.4407$_{\pm.0004}$	.4460$_{\pm.0003}$	**.4685**$_{\pm.0098}$	.4677$_{\pm.0067}$	.4397$_{\pm.0010}$	.4457$_{\pm.0018}$	**.4423**$_{\pm.0014}$	.4508$_{\pm.0019}$
	$\Delta\uparrow$ (%)	-2.38$_{\pm.05}$	-5.08$_{\pm.18}$	-2.16$_{\pm.18}$	-4.33$_{\pm.07}$	**+1.24**$_{\pm.31}$	-2.72$_{\pm.16}$	-0.62$_{\pm.36}$	-2.65$_{\pm.44}$	**+0.87**$_{\pm.28}$	-2.67$_{\pm.21}$
ETTh2	base	.3982$_{\pm.0157}$	**.4109**$_{\pm.0117}$	.3766$_{\pm.0054}$	.3960$_{\pm.0007}$	.3828$_{\pm.0044}$	.4000$_{\pm.0035}$	.3733$_{\pm.0009}$	.3962$_{\pm.0010}$	.3826$_{\pm.0091}$	.4001$_{\pm.0054}$
	ours	**.3978**$_{\pm.0161}$	.4114$_{\pm.0117}$	**.3716**$_{\pm.0041}$	**.3917**$_{\pm.0008}$	**.3785**$_{\pm.0042}$	**.3975**$_{\pm.0039}$	**.3636**$_{\pm.0012}$	**.3918**$_{\pm.0011}$	**.3747**$_{\pm.0084}$	**.3937**$_{\pm.0047}$
	$\Delta\uparrow$ (%)	**+0.10**$_{\pm.11}$	-0.11$_{\pm.02}$	**+1.34**$_{\pm.38}$	**+1.08**$_{\pm.38}$	**+1.11**$_{\pm.15}$	**+0.63**$_{\pm.11}$	**+2.60**$_{\pm.15}$	**+1.11**$_{\pm.24}$	**+2.07**$_{\pm.53}$	**+1.60**$_{\pm.42}$
ETTm1	base	.3936$_{\pm.0022}$	**.3949**$_{\pm.0006}$	.3718$_{\pm.0012}$	**.3910**$_{\pm.0010}$	.3741$_{\pm.0009}$	**.3881**$_{\pm.0018}$	.3660$_{\pm.0003}$	**.3851**$_{\pm.0005}$	.3957$_{\pm.0050}$	**.4029**$_{\pm.0020}$
	ours	**.3850**$_{\pm.0015}$	.4022$_{\pm.0004}$	**.3638**$_{\pm.0007}$	.3953$_{\pm.0008}$	**.3662**$_{\pm.0010}$	.3943$_{\pm.0015}$	**.3615**$_{\pm.0019}$	.3896$_{\pm.0006}$	**.3746**$_{\pm.0019}$	.4036$_{\pm.0023}$
	$\Delta\uparrow$ (%)	**+2.16**$_{\pm.16}$	-1.84$_{\pm.05}$	**+2.16**$_{\pm.46}$	-1.10$_{\pm.18}$	**+2.11**$_{\pm.10}$	-1.58$_{\pm.09}$	**+1.23**$_{\pm.61}$	-1.17$_{\pm.07}$	**+5.32**$_{\pm.72}$	-0.19$_{\pm.22}$
ETTm2	base	.2564$_{\pm.0016}$	.3154$_{\pm.0014}$	.2419$_{\pm.0004}$	.3043$_{\pm.0005}$	.2422$_{\pm.0009}$	.3046$_{\pm.0004}$	.2395$_{\pm.0008}$	.3015$_{\pm.0006}$	.2519$_{\pm.0018}$	.3104$_{\pm.0009}$
	ours	**.2543**$_{\pm.0009}$	**.3147**$_{\pm.0012}$	**.2378**$_{\pm.0004}$	**.3040**$_{\pm.0007}$	**.2385**$_{\pm.0022}$	**.3028**$_{\pm.0008}$	**.2345**$_{\pm.0043}$	**.2996**$_{\pm.0025}$	**.2450**$_{\pm.0022}$	**.3071**$_{\pm.0009}$
	$\Delta\uparrow$ (%)	**+0.82**$_{\pm.75}$	**+0.21**$_{\pm.71}$	**+1.68**$_{\pm.09}$	**+0.08**$_{\pm.07}$	**+1.54**$_{\pm.54}$	**+0.60**$_{\pm.16}$	**+2.09**$_{\pm1.64}$	**+0.62**$_{\pm.84}$	**+2.76**$_{\pm.17}$	**+1.06**$_{\pm.16}$
Weather	base	.2405$_{\pm.0016}$	**.2733**$_{\pm.0024}$	.2245$_{\pm.0003}$	**.2586**$_{\pm.0003}$	.2263$_{\pm.0011}$	**.2642**$_{\pm.0007}$	.2070$_{\pm.0005}$	**.2513**$_{\pm.0009}$	.2243$_{\pm.0014}$	**.2568**$_{\pm.0010}$
	ours	**.2151**$_{\pm.0030}$	.2954$_{\pm.0028}$	**.2064**$_{\pm.0005}$	.2864$_{\pm.0005}$	**.2090**$_{\pm.0004}$	.2891$_{\pm.0005}$	**.2052**$_{\pm.0002}$	.2793$_{\pm.0017}$	**.2065**$_{\pm.0010}$	.2801$_{\pm.0017}$
	$\Delta\uparrow$ (%)	**+10.57**$_{\pm.65}$	-8.09$_{\pm.11}$	**+8.08**$_{\pm.23}$	-10.75$_{\pm.26}$	**+7.63**$_{\pm.47}$	-9.42$_{\pm.13}$	**+0.89**$_{\pm.30}$	-11.16$_{\pm.29}$	**+7.94**$_{\pm.12}$	-9.06$_{\pm.26}$
ECL	base	.1938$_{\pm.0004}$	.2806$_{\pm.0006}$	.1854$_{\pm.0002}$	.2790$_{\pm.0004}$	.1805$_{\pm.0010}$	.2703$_{\pm.0007}$	.1611$_{\pm.0002}$	.2569$_{\pm.0002}$	.1650$_{\pm.0005}$	.2563$_{\pm.0007}$
	ours	**.1679**$_{\pm.0001}$	**.2696**$_{\pm.0002}$	**.1596**$_{\pm.0001}$	**.2607**$_{\pm.0002}$	**.1604**$_{\pm.0006}$	**.2618**$_{\pm.0006}$	**.1496**$_{\pm.0000}$	**.2477**$_{\pm.0008}$	**.1546**$_{\pm.0009}$	**.2536**$_{\pm.0010}$
	$\Delta\uparrow$ (%)	**+13.38**$_{\pm.17}$	**+3.91**$_{\pm.15}$	**+13.91**$_{\pm.15}$	**+6.57**$_{\pm.18}$	**+11.10**$_{\pm.20}$	**+3.16**$_{\pm.08}$	**+7.18**$_{\pm.12}$	**+3.58**$_{\pm.37}$	**+6.32**$_{\pm.27}$	**+1.06**$_{\pm.12}$
Solar	base	.3471$_{\pm.0033}$	**.3355**$_{\pm.0072}$	.2703$_{\pm.0008}$	.3076$_{\pm.0016}$	.2933$_{\pm.0076}$	**.3171**$_{\pm.0018}$	.2427$_{\pm.0010}$	.2935$_{\pm.0016}$	.2412$_{\pm.0025}$	.2761$_{\pm.0011}$
	ours	**.2521**$_{\pm.0009}$	.3435$_{\pm.0012}$	**.2145**$_{\pm.0004}$	**.2938**$_{\pm.0008}$	**.2272**$_{\pm.0035}$	.3176$_{\pm.0057}$	**.2011**$_{\pm.0029}$	**.2794**$_{\pm.0035}$	**.1951**$_{\pm.0009}$	**.2707**$_{\pm.0014}$
	$\Delta\uparrow$ (%)	**+27.38**$_{\pm.44}$	-2.40$_{\pm1.90}$	**+20.63**$_{\pm.29}$	**+4.48**$_{\pm.75}$	**+22.53**$_{\pm.83}$	-0.17$_{\pm1.55}$	**+17.15**$_{\pm.85}$	**+4.79**$_{\pm1.06}$	**+19.10**$_{\pm.63}$	**+1.95**$_{\pm.33}$
Traffic	base	.5318$_{\pm.0003}$	.3463$_{\pm.0007}$	.4556$_{\pm.0006}$	.2954$_{\pm.0002}$	.4985$_{\pm.0058}$	.3264$_{\pm.0053}$	.4547$_{\pm.0004}$	.2967$_{\pm.0003}$	.4124$_{\pm.0004}$	**.2770**$_{\pm.0003}$
	ours	**.4894**$_{\pm.0007}$	**.3311**$_{\pm.0002}$	**.4323**$_{\pm.0004}$	**.2780**$_{\pm.0003}$	**.4709**$_{\pm.0036}$	**.3140**$_{\pm.0033}$	**.4378**$_{\pm.0004}$	**.2915**$_{\pm.0004}$	**.4057**$_{\pm.0002}$	.2781$_{\pm.0002}$
	$\Delta\uparrow$ (%)	**+7.98**$_{\pm.13}$	**+4.40**$_{\pm.13}$	**+5.12**$_{\pm.06}$	**+5.88**$_{\pm.09}$	**+5.54**$_{\pm.40}$	**+3.79**$_{\pm.60}$	**+3.73**$_{\pm.05}$	**+1.74**$_{\pm.04}$	**+1.61**$_{\pm.06}$	-0.39$_{\pm.05}$

Table 30: Complete 80-cell results at $H=336$. Each backbone has paired MSE/MAE columns; the two panels cover all ten backbones. Entries are mean ± sample SD over seeds $2/3/4$; SD is smaller and leading zeros are omitted. Δ is the seed-wise relative reduction (%, ↑), computed from unrounded metrics. Bold marks lower base/ours error and positive reductions.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.4870**$_{\pm.0001}$	**.4650**$_{\pm.0000}$	**.4705**$_{\pm.0035}$	**.4533**$_{\pm.0045}$	**.4924**$_{\pm.0030}$	**.4585**$_{\pm.0031}$	**.4807**$_{\pm.0007}$	**.4533**$_{\pm.0009}$	**.4822**$_{\pm.0019}$	**.4487**$_{\pm.0017}$
	ours	.5209$_{\pm.0000}$	.5041$_{\pm.0000}$	.5054$_{\pm.0011}$	.4867$_{\pm.0023}$	.5112$_{\pm.0058}$	.4894$_{\pm.0012}$	.5047$_{\pm.0011}$	.4861$_{\pm.0009}$	.5058$_{\pm.0017}$	.4878$_{\pm.0011}$
	$\Delta\uparrow$ (%)	-6.95$_{\pm.00}$	-8.42$_{\pm.00}$	-7.43$_{\pm.58}$	-7.37$_{\pm.56}$	-3.84$_{\pm1.68}$	-6.73$_{\pm.54}$	-4.99$_{\pm.09}$	-7.24$_{\pm.12}$	-4.91$_{\pm.36}$	-8.71$_{\pm.32}$
ETTh2	base	.5925$_{\pm.0119}$	.5404$_{\pm.0064}$	.4160$_{\pm.0055}$	.4275$_{\pm.0036}$	.4194$_{\pm.0007}$	.4287$_{\pm.0021}$	.4221$_{\pm.0005}$	**.4296**$_{\pm.0003}$	.4227$_{\pm.0005}$	**.4297**$_{\pm.0004}$
	ours	**.5879**$_{\pm.0122}$	**.5395**$_{\pm.0067}$	**.3994**$_{\pm.0032}$	**.4162**$_{\pm.0024}$	**.3993**$_{\pm.0009}$	**.4155**$_{\pm.0014}$	**.4204**$_{\pm.0006}$	.4306$_{\pm.0005}$	**.4204**$_{\pm.0006}$	.4302$_{\pm.0004}$
	$\Delta\uparrow$ (%)	**+0.77**$_{\pm.07}$	**+0.16**$_{\pm.05}$	**+3.98**$_{\pm.54}$	**+2.65**$_{\pm.28}$	**+4.79**$_{\pm.18}$	**+3.08**$_{\pm.15}$	**+0.40**$_{\pm.03}$	-0.22$_{\pm.04}$	**+0.53**$_{\pm.04}$	-0.11$_{\pm.03}$
ETTm1	base	.4133$_{\pm.0002}$	**.4133**$_{\pm.0004}$	.4033$_{\pm.0021}$	**.4120**$_{\pm.0018}$	.4067$_{\pm.0054}$	**.4103**$_{\pm.0019}$	.4244$_{\pm.0001}$	**.4140**$_{\pm.0002}$	.4342$_{\pm.0019}$	**.4214**$_{\pm.0015}$
	ours	**.4071**$_{\pm.0003}$	.4177$_{\pm.0004}$	**.4008**$_{\pm.0018}$	.4138$_{\pm.0011}$	**.3984**$_{\pm.0029}$	.4122$_{\pm.0034}$	**.4112**$_{\pm.0002}$	.4196$_{\pm.0001}$	**.4195**$_{\pm.0020}$	.4265$_{\pm.0016}$
	$\Delta\uparrow$ (%)	**+1.50**$_{\pm.03}$	-1.07$_{\pm.01}$	**+0.61**$_{\pm.28}$	-0.45$_{\pm.33}$	**+2.02**$_{\pm.71}$	-0.46$_{\pm.66}$	**+3.10**$_{\pm.01}$	-1.35$_{\pm.02}$	**+3.38**$_{\pm.05}$	-1.21$_{\pm.07}$
ETTm2	base	.3840$_{\pm.0022}$	.4278$_{\pm.0022}$	.3092$_{\pm.0011}$	**.3464**$_{\pm.0014}$	.2935$_{\pm.0006}$	.3339$_{\pm.0008}$	.3071$_{\pm.0001}$	.3420$_{\pm.0001}$	.3109$_{\pm.0004}$	.3461$_{\pm.0004}$
	ours	**.3647**$_{\pm.0167}$	**.4175**$_{\pm.0153}$	**.3074**$_{\pm.0009}$	.3464$_{\pm.0017}$	**.2904**$_{\pm.0010}$	**.3333**$_{\pm.0022}$	**.3054**$_{\pm.0001}$	**.3420**$_{\pm.0001}$	**.3088**$_{\pm.0003}$	**.3457**$_{\pm.0004}$
	$\Delta\uparrow$ (%)	**+5.03**$_{\pm4.29}$	**+2.40**$_{\pm3.72}$	**+0.58**$_{\pm.21}$	-0.00$_{\pm.08}$	**+1.07**$_{\pm.15}$	**+0.20**$_{\pm.44}$	**+0.56**$_{\pm.01}$	**+0.01**$_{\pm.01}$	**+0.67**$_{\pm.04}$	**+0.12**$_{\pm.02}$
Weather	base	.2823$_{\pm.0005}$	**.3323**$_{\pm.0011}$	.2708$_{\pm.0003}$	**.2947**$_{\pm.0002}$	.2619$_{\pm.0014}$	**.2893**$_{\pm.0015}$	.2934$_{\pm.0000}$	**.3084**$_{\pm.0000}$	.2967$_{\pm.0003}$	**.3113**$_{\pm.0002}$
	ours	**.2624**$_{\pm.0002}$	.3338$_{\pm.0001}$	**.2579**$_{\pm.0003}$	.3174$_{\pm.0004}$	**.2469**$_{\pm.0014}$	.3077$_{\pm.0016}$	**.2678**$_{\pm.0008}$	.3394$_{\pm.0007}$	**.2678**$_{\pm.0009}$	.3407$_{\pm.0010}$
	$\Delta\uparrow$ (%)	**+7.05**$_{\pm.23}$	-0.45$_{\pm.34}$	**+4.77**$_{\pm.06}$	-7.69$_{\pm.08}$	**+5.73**$_{\pm.30}$	-6.36$_{\pm.95}$	**+8.74**$_{\pm.26}$	-10.04$_{\pm.21}$	**+9.72**$_{\pm.28}$	-9.45$_{\pm.28}$
ECL	base	.2070$_{\pm.0001}$	.2963$_{\pm.0003}$	.2007$_{\pm.0002}$	.3123$_{\pm.0002}$	**.1716**$_{\pm.0023}$	**.2683**$_{\pm.0021}$	.2168$_{\pm.0001}$	.2995$_{\pm.0002}$	.2229$_{\pm.0023}$	.3016$_{\pm.0018}$
	ours	**.1799**$_{\pm.0001}$	**.2825**$_{\pm.0001}$	**.1580**$_{\pm.0001}$	**.2593**$_{\pm.0001}$	.1752$_{\pm.0016}$	.2761$_{\pm.0014}$	**.1863**$_{\pm.0000}$	**.2898**$_{\pm.0001}$	**.1844**$_{\pm.0004}$	**.2862**$_{\pm.0004}$
	$\Delta\uparrow$ (%)	**+13.05**$_{\pm.05}$	**+4.67**$_{\pm.10}$	**+21.27**$_{\pm.06}$	**+16.96**$_{\pm.10}$	-2.10$_{\pm.42}$	-2.90$_{\pm.30}$	**+14.07**$_{\pm.04}$	**+3.25**$_{\pm.05}$	**+17.26**$_{\pm.76}$	**+5.11**$_{\pm.45}$
Solar	base	.3523$_{\pm.0000}$	.4150$_{\pm.0000}$	.2816$_{\pm.0042}$	.3372$_{\pm.0036}$	.2380$_{\pm.0022}$	.2776$_{\pm.0009}$	.4219$_{\pm.0002}$	.4078$_{\pm.0012}$	.4355$_{\pm.0015}$	.4081$_{\pm.0017}$
	ours	**.2798**$_{\pm.0000}$	**.3637**$_{\pm.0000}$	**.2197**$_{\pm.0031}$	**.2972**$_{\pm.0040}$	**.2100**$_{\pm.0021}$	**.2730**$_{\pm.0010}$	**.2911**$_{\pm.0007}$	**.3811**$_{\pm.0007}$	**.2935**$_{\pm.0010}$	**.3858**$_{\pm.0019}$
	$\Delta\uparrow$ (%)	**+20.58**$_{\pm.01}$	**+12.36**$_{\pm.01}$	**+21.99**$_{\pm.17}$	**+11.86**$_{\pm.27}$	**+11.75**$_{\pm.30}$	**+1.65**$_{\pm.23}$	**+30.99**$_{\pm.13}$	**+6.55**$_{\pm.15}$	**+32.60**$_{\pm.46}$	**+5.47**$_{\pm.78}$
Traffic	base	.6051$_{\pm.0000}$	.3735$_{\pm.0001}$	.5035$_{\pm.0006}$	.3317$_{\pm.0004}$	.4460$_{\pm.0040}$	**.2798**$_{\pm.0009}$	.6116$_{\pm.0002}$	.3692$_{\pm.0003}$	.6427$_{\pm.0017}$	.3918$_{\pm.0018}$
	ours	**.5335**$_{\pm.0003}$	**.3482**$_{\pm.0001}$	**.4475**$_{\pm.0009}$	**.2944**$_{\pm.0006}$	**.4306**$_{\pm.0027}$	.2822$_{\pm.0009}$	**.5272**$_{\pm.0001}$	**.3478**$_{\pm.0001}$	**.5319**$_{\pm.0014}$	**.3498**$_{\pm.0016}$
	$\Delta\uparrow$ (%)	**+11.83**$_{\pm.05}$	**+6.76**$_{\pm.05}$	**+11.14**$_{\pm.26}$	**+11.24**$_{\pm.24}$	**+3.46**$_{\pm.26}$	-0.85$_{\pm.02}$	**+13.81**$_{\pm.02}$	**+5.79**$_{\pm.06}$	**+17.25**$_{\pm.08}$	**+10.71**$_{\pm.12}$

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.5228**$_{\pm.0032}$	**.4816**$_{\pm.0022}$	**.4697**$_{\pm.0013}$	**.4493**$_{\pm.0011}$	**.5098**$_{\pm.0062}$	**.4711**$_{\pm.0037}$	**.4852**$_{\pm.0051}$	**.4538**$_{\pm.0040}$	**.4879**$_{\pm.0006}$	**.4593**$_{\pm.0004}$
	ours	.5492$_{\pm.0034}$	.5199$_{\pm.0021}$	.4984$_{\pm.0010}$	.4865$_{\pm.0007}$	.5211$_{\pm.0026}$	.5015$_{\pm.0012}$	.5055$_{\pm.0036}$	.4853$_{\pm.0030}$	.4993$_{\pm.0006}$	.4876$_{\pm.0005}$
	$\Delta\uparrow$ (%)	-5.05$_{\pm.02}$	-7.96$_{\pm.07}$	-6.12$_{\pm.11}$	-8.28$_{\pm.13}$	-2.22$_{\pm.76}$	-6.46$_{\pm.64}$	-4.19$_{\pm1.55}$	-6.94$_{\pm.41}$	-2.34$_{\pm.04}$	-6.18$_{\pm.07}$
ETTh2	base	.4338$_{\pm.0152}$	.4417$_{\pm.0108}$	.4167$_{\pm.0018}$	.4300$_{\pm.0008}$	.4267$_{\pm.0074}$	.4369$_{\pm.0049}$	.4272$_{\pm.0084}$	.4327$_{\pm.0065}$	.4306$_{\pm.0178}$	.4346$_{\pm.0097}$
	ours	**.4303**$_{\pm.0171}$	**.4403**$_{\pm.0128}$	**.4017**$_{\pm.0065}$	**.4183**$_{\pm.0045}$	**.4192**$_{\pm.0069}$	**.4322**$_{\pm.0051}$	**.4069**$_{\pm.0006}$	**.4211**$_{\pm.0003}$	**.4138**$_{\pm.0115}$	**.4224**$_{\pm.0057}$
	$\Delta\uparrow$ (%)	**+0.81**$_{\pm.47}$	**+0.31**$_{\pm.48}$	**+3.59**$_{\pm1.80}$	**+2.74**$_{\pm1.21}$	**+1.76**$_{\pm.20}$	**+1.09**$_{\pm.17}$	**+4.73**$_{\pm1.97}$	**+2.66**$_{\pm1.40}$	**+3.86**$_{\pm1.90}$	**+2.80**$_{\pm1.22}$
ETTm1	base	.4245$_{\pm.0010}$	**.4139**$_{\pm.0003}$	**.4004**$_{\pm.0008}$	**.4102**$_{\pm.0022}$	.4059$_{\pm.0016}$	**.4090**$_{\pm.0005}$	.3998$_{\pm.0013}$	**.4091**$_{\pm.0010}$	.4253$_{\pm.0095}$	**.4223**$_{\pm.0049}$
	ours	**.4197**$_{\pm.0006}$	.4282$_{\pm.0004}$	.4028$_{\pm.0005}$	.4203$_{\pm.0016}$	**.4043**$_{\pm.0008}$	.4229$_{\pm.0008}$	**.3987**$_{\pm.0017}$	.4154$_{\pm.0015}$	**.4148**$_{\pm.0032}$	.4287$_{\pm.0046}$
	$\Delta\uparrow$ (%)	**+1.12**$_{\pm.09}$	-3.44$_{\pm.01}$	-0.60$_{\pm.07}$	-2.48$_{\pm.15}$	**+0.39**$_{\pm.21}$	-3.40$_{\pm.08}$	**+0.28**$_{\pm.13}$	-1.54$_{\pm.16}$	**+2.44**$_{\pm1.73}$	-1.54$_{\pm.21}$
ETTm2	base	.3158$_{\pm.0019}$	**.3516**$_{\pm.0011}$	.3050$_{\pm.0016}$	.3448$_{\pm.0011}$	.3026$_{\pm.0006}$	**.3431**$_{\pm.0007}$	.3002$_{\pm.0024}$	.3414$_{\pm.0007}$	.3174$_{\pm.0017}$	.3518$_{\pm.0025}$
	ours	**.3148**$_{\pm.0021}$	.3518$_{\pm.0014}$	**.3026**$_{\pm.0017}$	**.3437**$_{\pm.0009}$	**.2998**$_{\pm.0006}$	.3432$_{\pm.0009}$	**.2960**$_{\pm.0019}$	**.3397**$_{\pm.0012}$	**.3081**$_{\pm.0016}$	**.3484**$_{\pm.0022}$
	$\Delta\uparrow$ (%)	**+0.31**$_{\pm.09}$	-0.07$_{\pm.10}$	**+0.80**$_{\pm.12}$	**+0.33**$_{\pm.10}$	**+0.92**$_{\pm.20}$	-0.03$_{\pm.21}$	**+1.38**$_{\pm.33}$	**+0.49**$_{\pm.39}$	**+2.93**$_{\pm.14}$	**+0.97**$_{\pm.16}$
Weather	base	.2909$_{\pm.0016}$	**.3075**$_{\pm.0022}$	.2790$_{\pm.0006}$	**.2984**$_{\pm.0007}$	.2774$_{\pm.0016}$	**.2986**$_{\pm.0024}$	.2630$_{\pm.0004}$	**.2918**$_{\pm.0005}$	.2833$_{\pm.0040}$	**.2999**$_{\pm.0034}$
	ours	**.2666**$_{\pm.0043}$	.3358$_{\pm.0043}$	**.2591**$_{\pm.0013}$	.3214$_{\pm.0011}$	**.2632**$_{\pm.0007}$	.3244$_{\pm.0013}$	**.2533**$_{\pm.0011}$	.3142$_{\pm.0017}$	**.2582**$_{\pm.0047}$	.3231$_{\pm.0078}$
	$\Delta\uparrow$ (%)	**+8.36**$_{\pm.96}$	-9.20$_{\pm.63}$	**+7.15**$_{\pm.26}$	-7.73$_{\pm.17}$	**+5.15**$_{\pm.49}$	-8.63$_{\pm1.08}$	**+3.70**$_{\pm.54}$	-7.67$_{\pm.74}$	**+8.86**$_{\pm.65}$	-7.71$_{\pm1.60}$
ECL	base	.2116$_{\pm.0014}$	.2982$_{\pm.0015}$	.2018$_{\pm.0001}$	.2949$_{\pm.0002}$	.1964$_{\pm.0010}$	.2856$_{\pm.0008}$	.1789$_{\pm.0013}$	.2744$_{\pm.0006}$	.1788$_{\pm.0002}$	.2713$_{\pm.0006}$
	ours	**.1801**$_{\pm.0007}$	**.2823**$_{\pm.0012}$	**.1697**$_{\pm.0002}$	**.2709**$_{\pm.0003}$	**.1717**$_{\pm.0006}$	**.2739**$_{\pm.0014}$	**.1628**$_{\pm.0007}$	**.2627**$_{\pm.0002}$	**.1673**$_{\pm.0043}$	**.2679**$_{\pm.0029}$
	$\Delta\uparrow$ (%)	**+14.88**$_{\pm.22}$	**+5.33**$_{\pm.35}$	**+15.90**$_{\pm.10}$	**+8.14**$_{\pm.11}$	**+12.58**$_{\pm.11}$	**+4.10**$_{\pm.26}$	**+8.98**$_{\pm.35}$	**+4.25**$_{\pm.27}$	**+6.41**$_{\pm2.41}$	**+1.25**$_{\pm.89}$
Solar	base	.3897$_{\pm.0012}$	**.3544**$_{\pm.0061}$	.2978$_{\pm.0066}$	.3206$_{\pm.0042}$	.3205$_{\pm.0035}$	.3388$_{\pm.0023}$	.2686$_{\pm.0011}$	.3119$_{\pm.0020}$	.2640$_{\pm.0014}$	.2920$_{\pm.0002}$
	ours	**.2860**$_{\pm.0011}$	.3660$_{\pm.0005}$	**.2403**$_{\pm.0022}$	**.3102**$_{\pm.0025}$	**.2558**$_{\pm.0038}$	**.3342**$_{\pm.0026}$	**.2221**$_{\pm.0010}$	**.2965**$_{\pm.0007}$	**.2138**$_{\pm.0003}$	**.2866**$_{\pm.0015}$
	$\Delta\uparrow$ (%)	**+26.63**$_{\pm.38}$	-3.29$_{\pm1.74}$	**+19.29**$_{\pm1.04}$	**+3.25**$_{\pm1.05}$	**+20.19**$_{\pm.51}$	**+1.35**$_{\pm1.06}$	**+17.30**$_{\pm.24}$	**+4.92**$_{\pm.51}$	**+19.03**$_{\pm.40}$	**+1.85**$_{\pm.45}$
Traffic	base	.5436$_{\pm.0010}$	.3511$_{\pm.0003}$	.4720$_{\pm.0007}$	.3023$_{\pm.0003}$	.5137$_{\pm.0031}$	.3324$_{\pm.0044}$	.4721$_{\pm.0029}$	.3026$_{\pm.0005}$	.4231$_{\pm.0020}$	**.2824**$_{\pm.0006}$
	ours	**.4985**$_{\pm.0008}$	**.3328**$_{\pm.0010}$	**.4541**$_{\pm.0005}$	**.2965**$_{\pm.0002}$	**.4839**$_{\pm.0008}$	**.3151**$_{\pm.0028}$	**.4543**$_{\pm.0013}$	**.2988**$_{\pm.0008}$	**.4180**$_{\pm.0020}$	.2841$_{\pm.0005}$
	$\Delta\uparrow$ (%)	**+8.29**$_{\pm.07}$	**+5.21**$_{\pm.36}$	**+3.79**$_{\pm.05}$	**+1.93**$_{\pm.06}$	**+5.79**$_{\pm.45}$	**+5.19**$_{\pm.44}$	**+3.78**$_{\pm.32}$	**+1.27**$_{\pm.14}$	**+1.20**$_{\pm.02}$	-0.62$_{\pm.05}$

Table 31: Complete 80-cell results at $H=720$. Each backbone has paired MSE/MAE columns; the two panels cover all ten backbones. Entries are mean ± sample SD over seeds $2/3/4$; SD is smaller and leading zeros are omitted. Δ is the seed-wise relative reduction (%, ↑), computed from unrounded metrics. Bold marks lower base/ours error and positive reductions.

dataset	forecast	DLinear		DUET		FACT		FITS		MixLinear	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.5132**$_{\pm.0022}$	**.5107**$_{\pm.0019}$	**.4753**$_{\pm.0081}$	**.4785**$_{\pm.0040}$	**.5237**$_{\pm.0104}$	**.4941**$_{\pm.0070}$	**.4769**$_{\pm.0043}$	**.4778**$_{\pm.0032}$	**.4779**$_{\pm.0108}$	**.4704**$_{\pm.0077}$
	ours	.5684$_{\pm.0018}$	.5576$_{\pm.0016}$	.5257$_{\pm.0080}$	.5204$_{\pm.0040}$	.5469$_{\pm.0054}$	.5250$_{\pm.0038}$	.5326$_{\pm.0035}$	.5218$_{\pm.0025}$	.5324$_{\pm.0071}$	.5201$_{\pm.0050}$
	$\Delta\uparrow$ (%)	-10.74$_{\pm.13}$	-9.17$_{\pm.11}$	-10.60$_{\pm.44}$	-8.75$_{\pm.15}$	-4.44$_{\pm1.08}$	-6.26$_{\pm.74}$	-11.68$_{\pm.30}$	-9.20$_{\pm.22}$	-11.43$_{\pm1.03}$	-10.59$_{\pm.75}$
ETTh2	base	.8395$_{\pm.0023}$	.6608$_{\pm.0011}$	.4244$_{\pm.0059}$	.4423$_{\pm.0036}$	.4306$_{\pm.0041}$	.4449$_{\pm.0026}$	.4248$_{\pm.0005}$	**.4414**$_{\pm.0005}$	.4288$_{\pm.0038}$	**.4445**$_{\pm.0032}$
	ours	**.7532**$_{\pm.0023}$	**.6260**$_{\pm.0011}$	**.4032**$_{\pm.0028}$	**.4290**$_{\pm.0010}$	**.4008**$_{\pm.0067}$	**.4244**$_{\pm.0048}$	**.4222**$_{\pm.0005}$	.4425$_{\pm.0004}$	**.4263**$_{\pm.0033}$	.4456$_{\pm.0027}$
	$\Delta\uparrow$ (%)	**+10.28**$_{\pm.03}$	**+5.27**$_{\pm.00}$	**+5.00**$_{\pm1.34}$	**+3.01**$_{\pm1.00}$	**+6.92**$_{\pm1.92}$	**+4.60**$_{\pm1.32}$	**+0.61**$_{\pm.02}$	-0.26$_{\pm.02}$	**+0.59**$_{\pm.10}$	-0.25$_{\pm.12}$
ETTm1	base	**.4732**$_{\pm.0006}$	**.4514**$_{\pm.0009}$	.4661$_{\pm.0009}$	**.4496**$_{\pm.0005}$	.4705$_{\pm.0025}$	.4536$_{\pm.0026}$	.4856$_{\pm.0001}$	**.4479**$_{\pm.0001}$	.5054$_{\pm.0056}$	**.4622**$_{\pm.0034}$
	ours	.4739$_{\pm.0009}$	.4650$_{\pm.0008}$	**.4551**$_{\pm.0071}$	.4505$_{\pm.0016}$	**.4463**$_{\pm.0029}$	**.4483**$_{\pm.0020}$	**.4628**$_{\pm.0000}$	.4509$_{\pm.0000}$	**.4806**$_{\pm.0037}$	.4648$_{\pm.0017}$
	$\Delta\uparrow$ (%)	-0.16$_{\pm.08}$	-3.01$_{\pm.02}$	**+2.37**$_{\pm1.62}$	-0.20$_{\pm.38}$	**+5.15**$_{\pm.11}$	**+1.17**$_{\pm.17}$	**+4.68**$_{\pm.01}$	-0.65$_{\pm.01}$	**+4.92**$_{\pm.31}$	-0.57$_{\pm.38}$
ETTm2	base	.5415$_{\pm.0079}$	.5162$_{\pm.0056}$	.4125$_{\pm.0033}$	.4041$_{\pm.0016}$	.3988$_{\pm.0086}$	.3950$_{\pm.0044}$	.4073$_{\pm.0001}$	.3974$_{\pm.0001}$	.4114$_{\pm.0002}$	.4017$_{\pm.0002}$
	ours	**.5096**$_{\pm.0262}$	**.5027**$_{\pm.0192}$	**.4096**$_{\pm.0030}$	**.4028**$_{\pm.0014}$	**.3890**$_{\pm.0058}$	**.3917**$_{\pm.0036}$	**.4053**$_{\pm.0001}$	**.3971**$_{\pm.0001}$	**.4090**$_{\pm.0002}$	**.4010**$_{\pm.0001}$
	$\Delta\uparrow$ (%)	**+5.83**$_{\pm6.04}$	**+2.57**$_{\pm4.70}$	**+0.68**$_{\pm.08}$	**+0.33**$_{\pm.09}$	**+2.46**$_{\pm1.15}$	**+0.85**$_{\pm.80}$	**+0.50**$_{\pm.01}$	**+0.08**$_{\pm.01}$	**+0.58**$_{\pm.03}$	**+0.19**$_{\pm.02}$
Weather	base	**.3461**$_{\pm.0015}$	**.3828**$_{\pm.0026}$	.3466$_{\pm.0002}$	**.3436**$_{\pm.0004}$	.3525$_{\pm.0045}$	**.3475**$_{\pm.0028}$	.3662$_{\pm.0001}$	**.3549**$_{\pm.0000}$	.3691$_{\pm.0003}$	**.3583**$_{\pm.0004}$
	ours	.3497$_{\pm.0017}$	.4040$_{\pm.0021}$	**.3460**$_{\pm.0006}$	.3848$_{\pm.0001}$	**.3479**$_{\pm.0022}$	.3856$_{\pm.0026}$	**.3257**$_{\pm.0006}$	.3762$_{\pm.0005}$	**.3265**$_{\pm.0009}$	.3772$_{\pm.0010}$
	$\Delta\uparrow$ (%)	-1.02$_{\pm.14}$	-5.54$_{\pm.19}$	**+0.17**$_{\pm.19}$	-11.99$_{\pm.15}$	**+1.29**$_{\pm.88}$	-10.97$_{\pm1.36}$	**+11.06**$_{\pm.15}$	-6.00$_{\pm.13}$	**+11.53**$_{\pm.20}$	-5.27$_{\pm.20}$
ECL	base	.2423$_{\pm.0001}$	.3291$_{\pm.0004}$	.2275$_{\pm.0019}$	.3359$_{\pm.0014}$	**.2034**$_{\pm.0060}$	**.2986**$_{\pm.0047}$	.2580$_{\pm.0000}$	.3312$_{\pm.0001}$	.2736$_{\pm.0041}$	.3425$_{\pm.0059}$
	ours	**.2097**$_{\pm.0001}$	**.3127**$_{\pm.0002}$	**.1825**$_{\pm.0014}$	**.2877**$_{\pm.0021}$	.2106$_{\pm.0078}$	.3123$_{\pm.0059}$	**.2173**$_{\pm.0000}$	**.3190**$_{\pm.0000}$	**.2190**$_{\pm.0012}$	**.3212**$_{\pm.0005}$
	$\Delta\uparrow$ (%)	**+13.47**$_{\pm.03}$	**+4.99**$_{\pm.09}$	**+19.76**$_{\pm.59}$	**+14.36**$_{\pm.45}$	-3.54$_{\pm.78}$	-4.57$_{\pm.33}$	**+15.76**$_{\pm.01}$	**+3.70**$_{\pm.02}$	**+19.95**$_{\pm.78}$	**+6.20**$_{\pm1.53}$
Solar	base	.3565$_{\pm.0000}$	.4128$_{\pm.0000}$	.2832$_{\pm.0029}$	.3374$_{\pm.0029}$	**.2427**$_{\pm.0008}$	**.2809**$_{\pm.0007}$	.4210$_{\pm.0007}$	.3987$_{\pm.0011}$	.4331$_{\pm.0039}$	.4019$_{\pm.0057}$
	ours	**.2918**$_{\pm.0000}$	**.3644**$_{\pm.0000}$	**.2303**$_{\pm.0009}$	**.3058**$_{\pm.0019}$	.2595$_{\pm.0006}$	.3626$_{\pm.0021}$	**.2927**$_{\pm.0008}$	**.3741**$_{\pm.0015}$	**.3008**$_{\pm.0014}$	**.3826**$_{\pm.0015}$
	$\Delta\uparrow$ (%)	**+18.14**$_{\pm.02}$	**+11.74**$_{\pm.01}$	**+18.69**$_{\pm.51}$	**+9.36**$_{\pm.47}$	-6.90$_{\pm.16}$	-29.09$_{\pm.66}$	**+30.48**$_{\pm.13}$	**+6.17**$_{\pm.40}$	**+30.54**$_{\pm.37}$	**+4.79**$_{\pm1.08}$
Traffic	base	.6456$_{\pm.0001}$	**.3952**$_{\pm.0001}$	.5377$_{\pm.0017}$	.3468$_{\pm.0006}$	.4837$_{\pm.0028}$	**.3006**$_{\pm.0005}$	.6494$_{\pm.0001}$	**.3885**$_{\pm.0002}$	.7368$_{\pm.0031}$	.4518$_{\pm.0012}$
	ours	**.6279**$_{\pm.0001}$	.4174$_{\pm.0009}$	**.4911**$_{\pm.0006}$	**.3174**$_{\pm.0001}$	**.4713**$_{\pm.0012}$	.3050$_{\pm.0005}$	**.6358**$_{\pm.0001}$	.4187$_{\pm.0002}$	**.5933**$_{\pm.0048}$	**.3916**$_{\pm.0040}$
	$\Delta\uparrow$ (%)	**+2.74**$_{\pm.02}$	-5.62$_{\pm.22}$	**+8.67**$_{\pm.23}$	**+8.48**$_{\pm.16}$	**+2.57**$_{\pm.34}$	-1.44$_{\pm.04}$	**+2.09**$_{\pm.01}$	-7.78$_{\pm.04}$	**+19.48**$_{\pm.47}$	**+13.32**$_{\pm.67}$

dataset	forecast	PaiFilter		PatchTST		TimeMixer		TimeXer		iTransformer	
		MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
ETTh1	base	**.5250**$_{\pm.0076}$	**.5023**$_{\pm.0052}$	**.4801**$_{\pm.0136}$	**.4801**$_{\pm.0104}$	**.5175**$_{\pm.0041}$	**.4937**$_{\pm.0029}$	**.5111**$_{\pm.0184}$	**.4905**$_{\pm.0073}$	**.5119**$_{\pm.0080}$	**.4952**$_{\pm.0039}$
	ours	.5636$_{\pm.0033}$	.5408$_{\pm.0018}$	.5248$_{\pm.0080}$	.5205$_{\pm.0073}$	.5547$_{\pm.0017}$	.5322$_{\pm.0014}$	.5297$_{\pm.0052}$	.5175$_{\pm.0033}$	.5335$_{\pm.0101}$	.5235$_{\pm.0049}$
	$\Delta\uparrow$ (%)	-7.37$_{\pm1.78}$	-7.66$_{\pm.89}$	-9.35$_{\pm1.45}$	-8.44$_{\pm.84}$	-7.20$_{\pm.51}$	-7.80$_{\pm.34}$	-3.73$_{\pm4.13}$	-5.52$_{\pm2.10}$	-4.25$_{\pm2.98}$	-5.72$_{\pm1.45}$
ETTh2	base	.4409$_{\pm.0121}$	.4540$_{\pm.0087}$	.4260$_{\pm.0029}$	.4467$_{\pm.0009}$	.4354$_{\pm.0025}$	.4512$_{\pm.0020}$	.4275$_{\pm.0048}$	.4449$_{\pm.0029}$	.4427$_{\pm.0057}$	.4520$_{\pm.0034}$
	ours	**.4178**$_{\pm.0140}$	**.4405**$_{\pm.0111}$	**.4001**$_{\pm.0032}$	**.4270**$_{\pm.0017}$	**.4081**$_{\pm.0017}$	**.4334**$_{\pm.0016}$	**.3966**$_{\pm.0031}$	**.4242**$_{\pm.0022}$	**.4158**$_{\pm.0047}$	**.4331**$_{\pm.0029}$
	$\Delta\uparrow$ (%)	**+5.26**$_{\pm.58}$	**+2.98**$_{\pm.59}$	**+6.09**$_{\pm.20}$	**+4.41**$_{\pm.27}$	**+6.27**$_{\pm.24}$	**+3.93**$_{\pm.15}$	**+7.22**$_{\pm1.42}$	**+4.65**$_{\pm.97}$	**+6.07**$_{\pm.39}$	**+4.17**$_{\pm.39}$
ETTm1	base	.4846$_{\pm.0010}$	**.4466**$_{\pm.0006}$	.4599$_{\pm.0006}$	**.4421**$_{\pm.0007}$	.4698$_{\pm.0017}$	**.4452**$_{\pm.0016}$	.4562$_{\pm.0029}$	**.4431**$_{\pm.0015}$	.5026$_{\pm.0038}$	.4659$_{\pm.0012}$
	ours	**.4753**$_{\pm.0005}$	.4613$_{\pm.0009}$	**.4586**$_{\pm.0038}$	.4549$_{\pm.0054}$	**.4638**$_{\pm.0015}$	.4577$_{\pm.0020}$	**.4520**$_{\pm.0023}$	.4541$_{\pm.0017}$	**.4709**$_{\pm.0015}$	**.4635**$_{\pm.0011}$
	$\Delta\uparrow$ (%)	**+1.92**$_{\pm.27}$	-3.31$_{\pm.07}$	**+0.27**$_{\pm.88}$	-2.90$_{\pm1.14}$	**+1.28**$_{\pm.13}$	-2.81$_{\pm.14}$	**+0.92**$_{\pm.15}$	-2.49$_{\pm.07}$	**+6.31**$_{\pm.51}$	**+0.52**$_{\pm.24}$
ETTm2	base	.4126$_{\pm.0028}$	**.4042**$_{\pm.0015}$	.4047$_{\pm.0014}$	.4032$_{\pm.0025}$	.3987$_{\pm.0004}$	**.3986**$_{\pm.0009}$	.3976$_{\pm.0017}$	.3985$_{\pm.0008}$	.4180$_{\pm.0016}$	.4104$_{\pm.0014}$
	ours	**.4116**$_{\pm.0023}$	.4073$_{\pm.0018}$	**.3949**$_{\pm.0056}$	**.4011**$_{\pm.0017}$	**.3943**$_{\pm.0002}$	.3988$_{\pm.0005}$	**.3850**$_{\pm.0045}$	**.3969**$_{\pm.0023}$	**.4028**$_{\pm.0026}$	**.4045**$_{\pm.0018}$
	$\Delta\uparrow$ (%)	**+0.24**$_{\pm.17}$	-0.75$_{\pm.75}$	**+2.44**$_{\pm1.50}$	**+0.50**$_{\pm.99}$	**+1.11**$_{\pm.14}$	-0.04$_{\pm.19}$	**+3.18**$_{\pm.80}$	**+0.40**$_{\pm.77}$	**+3.62**$_{\pm.30}$	**+1.44**$_{\pm.24}$
Weather	base	.3613$_{\pm.0010}$	**.3519**$_{\pm.0014}$	.3545$_{\pm.0003}$	**.3483**$_{\pm.0004}$	.3529$_{\pm.0012}$	**.3474**$_{\pm.0006}$	**.3431**$_{\pm.0014}$	**.3432**$_{\pm.0012}$	.3600$_{\pm.0014}$	**.3508**$_{\pm.0004}$
	ours	**.3551**$_{\pm.0055}$	.3926$_{\pm.0058}$	**.3493**$_{\pm.0006}$	.3869$_{\pm.0009}$	**.3495**$_{\pm.0009}$	.3874$_{\pm.0009}$	.3477$_{\pm.0038}$	.3881$_{\pm.0040}$	**.3530**$_{\pm.0040}$	.3916$_{\pm.0042}$
	$\Delta\uparrow$ (%)	**+1.70**$_{\pm1.25}$	-11.55$_{\pm1.21}$	**+1.45**$_{\pm.09}$	-11.09$_{\pm.13}$	**+0.99**$_{\pm.27}$	-11.51$_{\pm.09}$	-1.35$_{\pm.79}$	-13.07$_{\pm1.03}$	**+1.96**$_{\pm1.18}$	-11.65$_{\pm1.10}$
ECL	base	.2523$_{\pm.0005}$	.3305$_{\pm.0006}$	.2436$_{\pm.0015}$	.3278$_{\pm.0004}$	.2388$_{\pm.0013}$	.3202$_{\pm.0013}$	.2153$_{\pm.0038}$	.3059$_{\pm.0022}$	**.2097**$_{\pm.0018}$	**.2995**$_{\pm.0017}$
	ours	**.2090**$_{\pm.0002}$	**.3112**$_{\pm.0003}$	**.2001**$_{\pm.0011}$	**.3005**$_{\pm.0011}$	**.2198**$_{\pm.0153}$	**.3142**$_{\pm.0091}$	**.2034**$_{\pm.0115}$	**.3019**$_{\pm.0080}$	.2142$_{\pm.0014}$	.3127$_{\pm.0015}$
	$\Delta\uparrow$ (%)	**+17.17**$_{\pm.13}$	**+5.82**$_{\pm.12}$	**+17.85**$_{\pm.09}$	**+8.32**$_{\pm.24}$	**+7.96**$_{\pm6.84}$	**+1.84**$_{\pm3.16}$	**+5.53**$_{\pm4.36}$	**+1.33**$_{\pm2.09}$	-2.15$_{\pm.25}$	-4.41$_{\pm.12}$
Solar	base	.3924$_{\pm.0012}$	**.3542**$_{\pm.0057}$	.2938$_{\pm.0025}$	.3151$_{\pm.0010}$	.3217$_{\pm.0036}$	**.3420**$_{\pm.0040}$	.2755$_{\pm.0048}$	.3145$_{\pm.0039}$	.2757$_{\pm.0087}$	.3000$_{\pm.0058}$
	ours	**.2933**$_{\pm.0008}$	.3668$_{\pm.0003}$	**.2486**$_{\pm.0032}$	**.3111**$_{\pm.0028}$	**.2680**$_{\pm.0039}$	.3435$_{\pm.0049}$	**.2323**$_{\pm.0027}$	**.3009**$_{\pm.0034}$	**.2330**$_{\pm.0078}$	**.2968**$_{\pm.0075}$
	$\Delta\uparrow$ (%)	**+25.26**$_{\pm.43}$	-3.58$_{\pm1.63}$	**+15.38**$_{\pm.41}$	**+1.27**$_{\pm1.19}$	**+16.68**$_{\pm.31}$	-0.44$_{\pm.29}$	**+15.66**$_{\pm.60}$	**+4.34**$_{\pm.41}$	**+15.49**$_{\pm.45}$	**+1.05**$_{\pm.72}$
Traffic	base	**.5813**$_{\pm.0022}$	**.3694**$_{\pm.0015}$	.5064$_{\pm.0010}$	.3200$_{\pm.0010}$	**.5490**$_{\pm.0074}$	**.3466**$_{\pm.0026}$	**.5109**$_{\pm.0013}$	**.3187**$_{\pm.0007}$	.4586$_{\pm.0003}$	**.3006**$_{\pm.0004}$
	ours	.5821$_{\pm.0018}$	.3915$_{\pm.0008}$	**.4934**$_{\pm.0005}$	**.3176**$_{\pm.0007}$	.5537$_{\pm.0013}$	.3726$_{\pm.0019}$	.5468$_{\pm.0041}$	.3631$_{\pm.0012}$	**.4585**$_{\pm.0003}$	.3039$_{\pm.0003}$
	$\Delta\uparrow$ (%)	-0.15$_{\pm.17}$	-6.00$_{\pm.37}$	**+2.58**$_{\pm.19}$	**+0.73**$_{\pm.09}$	-0.86$_{\pm1.17}$	-7.49$_{\pm.53}$	-7.04$_{\pm1.06}$	-13.92$_{\pm.29}$	**+0.03**$_{\pm.05}$	-1.11$_{\pm.02}$

Non-deployable test-ranked reference. The deployable variants quantify candidate restrictions and validation selection. As a non-deployable reference, the upper reference chooses, for each cell, the lowest test MSE among candidates fitted without test labels. Test MSE performs this ranking only after candidate fitting; the final row of Table 24 reports this upper reference over the complete matrix. It enters neither validation choice nor the headline result.

E Further analyses and ablations

The following experiments explain which dependencies are useful, how finite calibration data limit their estimation, and why the selected resolution varies across datasets and checkpoints.

E.1 Structural alignment and estimation reliability

The central empirical question is whether a covariance class captures aligned residual dependencies and can still be estimated reliably on held-out calibration windows. Figure 5 isolates these two questions on the 74-cell subset.

Panel (a) changes dependency support while holding coefficient count fixed, whereas panel (b) changes calibration size while holding the class fixed. These controls separate what a class can represent from how reliably it can be estimated. They motivate selecting covariance resolution by held-out error rather than parameter count or in-sample fit.

E.2 Representation capacity and finite-sample estimation

Each synthetic test specifies the true relation between correction input δ and forecast error ϵ at a fixed signal-to-noise ratio. This setup tests whether a class can represent that relation separately from whether it can estimate the relation on finite data. Every operator is fitted on one half of the samples and scored on the other, then the halves are swapped. Entries report the fraction of the maximum population risk reduction recovered by the class. A value of 100% denotes complete recovery.

Table 32: Held-out recovery of specified forecast–error dependencies. Each row names the smallest operator class that can represent the specified relation. Entries are percentages of the maximum population risk reduction. Bold marks that smallest matching class.

specified structure	minimal	G0	G1	G2	G2$^{\mathbb{C}}$	G2″	G3	G3′	G4	G5(3)
independent channels	G2	3.6	20.4	**100.0**	100.0	100.0	100.0	100.0	100.0	100.0
dense channel map	G3	0.3	4.9	17.5	17.5	17.5	**100.0**	100.0	100.0	99.9
complex scalar	G2$^{\mathbb{C}}$	2.0	11.4	55.6	**100.0**	100.0	55.6	55.5	100.0	99.9
dense phase–channel map	G4	0.0	0.5	3.7	8.0	16.0	25.1	48.4	**100.0**	99.9
neighboring modes	G5(3)	0.0	0.2	1.3	2.9	5.6	8.6	16.2	34.1	**93.5**

The dense-channel and complex-scalar cases lead to different winners. For the dense channel map, every coordinate-separable class recovers 17.5%, whereas G_3 recovers 100.0%. For the complex-scalar relation, $G_2^{\mathbb{C}}$ reaches 100.0% with fewer coefficients than G_3, which reaches 55.6%. The hierarchy therefore branches according to allowed dependencies rather than following one ordering by parameter count.

Coordinate-level alternatives. Three intermediate classes add freedom within each coordinate without mixing channels or modes. Two of them use $2MC$ parameters but relax G_2 differently: G_2' assigns separate real scalars to the real and imaginary parts, whereas $G_2^{\mathbb{C}}$ assigns one *complex* scalar to each coordinate. Its real form is $aI + bJ$, where I is the identity and $J = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ is the 90-degree rotation matrix. This map can scale and rotate but cannot shear, meaning that it cannot change the angle between coordinate axes. Class G_2'' assigns a general real 2×2 map to each coordinate, uses $4MC$ parameters, and can also shear. Because $\mathrm{tr}(J\Sigma) = \Sigma_{12} - \Sigma_{21} = 0$ for a symmetric second moment and $\mathrm{tr}(J\Sigma J^\top) = \mathrm{tr}(\Sigma)$, the two basis directions of $G_2^{\mathbb{C}}$ are Σ-orthogonal. This orthogonality yields separate closed-form coefficients for scaling and rotation:

$$a_c = \frac{(\Gamma^{rr} + \Gamma^{ii})_{cc}}{(\Sigma^{rr} + \Sigma^{ii})_{cc}}, \qquad b_c = \frac{(\Gamma^{ir} - \Gamma^{ri})_{cc}}{(\Sigma^{rr} + \Sigma^{ii})_{cc}},$$

where superscripts rr, ri, ir, and ii identify real–imaginary block pairs. The coefficient a_c equals the G_2 coefficient, while b_c uses $\Gamma^{ir} - \Gamma^{ri}$, the difference between the two real–imaginary cross blocks. When the regularized operators are fitted on training windows and evaluated on validation windows, $G_2^{\mathbb{C}}$ exceeds G_2 by 0.03, 0.01, 0.09, 0.26, and 0.53 percentage points across the five channel-count groups. At C=862, their reductions are 1.95% and 1.42%. These values isolate the contribution of phase rotation from cross-channel mixing.

How class width affects held-out error. This finite-sample analysis fits unregularized pseudoinverses on training windows and evaluates them on validation windows. It isolates class width before ridge, gates, and selection. Entries are mean spectral-error reduction by channel count. Positive values reduce error. "G_3 or wider helps" counts cells where at least one of G_3, G_4, or $G_5(3)$ exceeds G_2. The G_4 in-sample column reports reduction measured on its fitting moments rather than held-out performance.

	cells	G2	G3	G4	G5(3)	G4 in-sample	G3 or wider helps
$C \leq 7$	40	-7.2	-45.2	-45.4	-56.0	17.2	1/40
C=21	10	$+8.0$	$+8.3$	$+8.6$	$\mathbf{+11.4}$	26.7	10/10
C=137	10	$+4.4$	$+17.2$	$\mathbf{+17.8}$	$+14.7$	60.7	10/10
C=321	10	$\mathbf{+5.7}$	-71.6	-87.8	-143.1	89.0	0/10
C=862	4	$\mathbf{+1.4}$	-156.4	-235.2	-286.4	99.9	0/4

The held-out ordering depends on channel count: wider classes help every cell at C=21 and C=137, whereas G_2 remains best at C=321 and C=862 despite the much larger in-sample reduction of G_4. This reversal shows why representation and estimation must be tested separately. The matched-support controls below hold coefficient count fixed to isolate the representation side of that distinction.

Matched coefficient counts with scrambled support. Three controls preserve coefficient count and change only which inputs the linear map may use. G2$_\pi$ keeps one real scalar per output coordinate but applies it to a different input coordinate under a fixed permutation in which no coordinate maps to itself. G2$_Q$ is diagonal in a random orthogonal basis of the $2C$ space rather than in the channel basis. G3$_\rho$ keeps C inputs per output row, chosen at random rather than as the two aligned channel blocks. Its cost grows as C^4, so it is run up to 137 channels and omitted above that width. The in-sample columns fit and score on the same moments. The held-out columns fit on training and score on validation. All numeric entries are spectral-error reductions in percentage points, so positive values are better.

	in sample: can it represent?			out of split: can it be estimated?		
control	aligned	scrambled	aligned ahead	aligned	scrambled	aligned ahead
G2$_\pi$ (permuted support)	6.51	2.08	$\mathbf{72/74}$	-1.32	-0.96	43/74
G2$_Q$ (random basis)	6.51	5.53	$\mathbf{59/74}$	-1.32	-0.40	27/74
G3$_\rho$ (random support)	22.11	15.79	$\mathbf{59/60}$	-25.84	-8.46	19/60

On the fitting moments, aligned support outperforms its matched permutation on $72/74$ cells, the random basis on $59/74$, and random channel support on $59/60$. Median margins for the first and third controls are 4.09 and 5.71 points. Their two-sided exact sign-test values are $p = 2.9\times10^{-19}$ and $p = 1.1\times10^{-16}$. Out of split, the corresponding counts are $43/74$, $27/74$, and $19/60$. The same aligned support can therefore fit more of the error on the fitting moments yet perform worse on held-out data. This empirical tension mirrors Eq. 13. The selector consequently ranks fitted classes on held-out error rather than choosing support from coefficient count or in-sample fit. A temporal-distance control tests whether the time gap between fitting and evaluation explains the held-out difference.

Temporal-distance control. The in-sample versus held-out gap can reflect both finite-sample estimation and changes in the moments over time. We split the training windows into contiguous halves A and B, fit on A, and compare scores on nearby B with scores on the later validation split.

The fitting set is identical. Only temporal distance changes. Over all 74 cells, the nearby split scores better on 37, 39, and 37 cells for G_2, G_3, and G_4. Let $\hat{K}_A$, $\hat{K}_B$, and $\hat{K}_{\mathrm{val}}$ denote operators fitted on A, B, and validation, and let Σ_{val} be the validation input second moment. The distance $\|\hat{K}_A - \hat{K}_B\|^2_{\Sigma_{\mathrm{val}}}$ exceeds $\|\hat{K}_A - \hat{K}_{\mathrm{val}}\|^2_{\Sigma_{\mathrm{val}}}$ on 31, 37, and 36 cells. Neither score nor operator disagreement consistently favors the temporally closer split. The calibration-size control varies the fitting budget directly.

Calibration-size control. Two disjoint random subsets of the training windows, each of size $n/2$, span the full training period and are fitted independently. Their normalized disagreement is $\|\hat{K}_A - \hat{K}_B\|^2_{\Sigma_{\mathrm{val}}} / \|\hat{K}_{\mathrm{val}}\|^2_{\Sigma_{\mathrm{val}}}$. For each fit, we define the held-out estimation gap as its in-sample reduction minus its validation reduction. Four values of n span an eightfold range over all 74 cells and three seeds. In the table, ρ_{Sp} denotes the per-cell Spearman correlation with n:

Table 33: Calibration-size control over 74 cells and three seeds. Disagreement compares two operators fitted on disjoint samples; "falls" counts cells in which the corresponding quantity decreases as the calibration size grows.

class	cells	$\rho_{\mathrm{Sp}}(n, \text{disagr.})$	falls	$\rho_{\mathrm{Sp}}(n, \text{gap})$	falls	log-log slope	disagr. $n_{\min} \to n_{\max}$
G2	74	-0.80	$\mathbf{74/74}$	$+0.10$	$33/74$	-0.94	$0.147 \to 0.009$
G3	74	-1.00	$\mathbf{74/74}$	-0.80	$\mathbf{61/74}$	$\mathbf{-1.01}$	$1.202 \to 0.109$
G4	74	-1.00	$\mathbf{74/74}$	-1.00	$\mathbf{62/74}$	$\mathbf{-1.03}$	$1.264 \to 0.185$

Disagreement decreases with n in every cell for each class. For G_3 and G_4, the per-cell Spearman correlation is -1.00, median log–log slopes are -1.01 and -1.03, and the held-out gap decreases in $61/74$ and $62/74$ cells. For G_2, disagreement also decreases, but the held-out gap decreases in $33/74$ cells and has Spearman correlation $+0.10$ with n. The approximately $1/n$ disagreement scaling is an empirical result. Equation 13 decomposes risk but does not assert this rate. Together with the matched-support control, these results are consistent with stronger finite-sample sensitivity in wider classes and more restrictive approximation in G_2.

At C=321, a G_4 input has dimension $2C = 642$ and each mode uses a 642×642 map. With 4555 training windows, its held-out gap remains 153 percentage points. This train-to-validation gap motivates ridge regularization and held-out selection. The nominal dimension-admissibility guard addresses the more extreme cases in which the predictor dimension meets or exceeds the fitting-row count (Section C.6). These results separate the dependencies a class can represent from those that finite calibration data can estimate reliably.

E.3 Fitting-source and selection analyses

Train-only, validation-only, and joint choices. Tables 34, 35, and 36 show this tradeoff under the same score. The tables report training-fitted candidates, validation-fitted candidates, and their union, respectively. In all three cases, the nominal dimension-admissibility guard uses the fitting-window count of that source, and the class and ridge penalty minimize the same summed held-out validation error. Accordingly, "joint validation-best" means the lowest validation score over source, class, and penalty. The non-deployable test-ranked reference is reported separately with the selection-rule variants.

Why both fitting sources are needed. Across all 960 runs, training-only selection chooses $G_5(3)$ 452 times, whereas validation-only selection chooses the narrower G_2 630 times. Joint selection retains a training-fitted operator in 582 runs and a validation-fitted one in 378. The benefit is evaluated independently for every checkpoint: joint selection improves $279/320$ cells by 6.34%, compared with $272/320$ and 3.85% for training-only or $258/320$ and 3.32% for validation-only. At H=720, the corresponding means are $+4.64\%$, -0.93%, and $+0.90\%$.

Target-use audit. A target-permutation audit checks the selection path. On ETTh1/iTransformer, seed 2, permuting fit and score targets while leaving test targets unchanged makes the G_2, G_3, and G_4 test MSE increase by 161.97%, 234.28%, and 235.83%. Conversely, permuting only test targets

Table 34: Training-fitted choices over all four horizons. Every candidate fits on training windows and is ranked on held-out validation residuals. Class counts are checkpoint–seed selections; performance first averages seeds within each cell and reports mean ± SD across the resulting cell means.

scope	G_2	G_3	G_4	$G_5(3)$	selections	improves ↑	mean ± SD ↑
H=96	71	6	45	118	240	75/80	$+7.28_{\pm 8.25\%}$
H=192	74	11	53	102	240	70/80	$+5.72_{\pm 9.75\%}$
H=336	71	6	26	137	240	68/80	$+3.34_{\pm 14.13\%}$
H=720	92	43	10	95	240	59/80	$-0.93_{\pm 22.82\%}$
all four horizons	308	66	134	452	960	272/320	$+3.85_{\pm 15.11\%}$

Table 35: Validation-fitted choices over all four horizons. Each candidate is cross-fitted between contiguous validation halves before refitting. Class counts are checkpoint–seed selections; performance first averages seeds within each cell and reports mean ± SD across the resulting cell means.

scope	G_2	G_3	G_4	$G_5(3)$	selections	improves ↑	mean ± SD ↑
H=96	128	34	48	30	240	76/80	$+5.24_{\pm 5.61\%}$
H=192	164	37	18	21	240	70/80	$+4.35_{\pm 5.45\%}$
H=336	167	38	11	24	240	60/80	$+2.81_{\pm 5.49\%}$
H=720	171	45	9	15	240	52/80	$+0.90_{\pm 5.43\%}$
all four horizons	630	154	86	90	960	258/320	$+3.32_{\pm 5.71\%}$

Table 36: Joint choices over all four horizons. Each T/V entry counts training-fitted and validation-fitted selections under their source-specific dimension guards. All candidates are ranked by the same held-out validation criterion. Performance first averages the three paired seeds within each cell and reports mean ± SD across cell means.

scope	G_2 (T/V)	G_3 (T/V)	G_4 (T/V)	$G_5(3)$ (T/V)	total T/V	improves ↑	mean ± SD ↑
H=96	44/49	0/15	25/19	87/1	156/84	77/80	$+7.61_{\pm 8.10\%}$
H=192	43/74	9/12	42/3	57/0	151/89	73/80	$+6.91_{\pm 8.26\%}$
H=336	57/63	4/15	15/1	85/0	161/79	68/80	$+6.22_{\pm 8.60\%}$
H=720	58/95	13/31	2/0	41/0	114/126	61/80	$+4.64_{\pm 9.11\%}$
all four horizons	202/281	26/73	84/23	270/1	582/378	279/320	$+6.34_{\pm 8.56\%}$

leaves every validation score and selected candidate unchanged. Thus fitting and selection depend on training and validation targets, while test targets enter final evaluation.

E.4 Beyond coordinate correction

Residual ambiguity motivates correction, while covariance resolution determines the required operator. Across Solar's 40 backbone–horizon cells, fixed coordinate correction reduces MSE by 2.66% on average, whereas the selected class reduces it by 21.35%. This gap shows recoverable dependence beyond coordinate-wise scaling. Excluding $G_5(3)$ changes the complete-matrix mean from 6.34% to 5.51% (Section C.6). Thus the frequency of weak modes and the complexity required for correction are separate questions: the multi-seed diagnostic identifies where ambiguity appears, whereas covariance-resolution selection determines which dependencies remain useful for correction.

E.5 Numerical checks of the covariance-resolution hierarchy

Singular Σ and the Moore–Penrose solution. The nominal real-stacked spectrum contains redundant endpoint imaginary coordinates. Its measured second-moment rank is 672 of 686 on ETTh1 and 30769 of 31458 on ECL, with an infinite condition number in every checked cell. On ETTh1, the deficit of 14 equals the $2C$ structurally zero endpoint coordinates; on ECL, 642 of the 689 missing directions are explained this way. The nominal ranks therefore combine endpoint redundancy with any additional rank loss. The theoretical risk is defined on the nonredundant signal space, where further singularity is still permitted. A minimizer still exists, because the columns of $\Gamma^\top$ lie in the range of Σ, but it is unique only on that range and the null directions contribute nothing to the risk. With $K^\star = \Gamma\Sigma^+$ the identity of Theorem 3 holds to 10^{-14}. Adding a small diagonal ridge term instead moves $K^\star$ away from the minimizer by an amount that grows when Σ is poorly conditioned. The same check then deviates by 2.3×10^{-4}, tracking that perturbation. Ridge minimizes a penalized empirical objective, whereas Theorem 3 characterizes the unpenalized population risk.

Partially ordered covariance-resolution lattice. Channel structure and the treatment of the real and imaginary parts are independent restrictions, so G_2'' and G_3 are incomparable. Their largest shared subclass, called their meet, is G_2: a matrix that is both diagonal in channels and equal across the two parts is exactly a tied real scalar per coordinate. Across cells, the median relative difference between the real and imaginary blocks is 2.55×10^{-2}, confirming that the two restrictions are not identical.

The branches form three inclusion chains: $G_0 \subset G_1 \subset G_2 \subset G_2^{\mathbb{C}} \subset G_2'' \subset G_4$; $G_2 \subset G_2' \subset G_2'' \subset G_4$; and $G_2 \subset G_3 \subset G_3' \subset G_4$. All three numerical orderings agree with Theorem 3. The smallest class containing both G_2'' and G_3, called their join, has dimension $C^2 + 3C$ in each nominal $2C$-coordinate mode block and lies inside G_4. At DC and Nyquist, removing the zero imaginary coordinates reduces the join dimension to C^2; the full effective dimension is $(M-2)(C^2+3C)+2C^2$. Class G_3' is block diagonal in the two parts, whereas G_2'' includes off-diagonal entries that rotate phase. Empirically, G_3' exceeds G_2'' on 74 of 74 cells, capturing 86.4 versus 26.5 points of G_4's gain on average.

Nominal sample size. Rolling windows share observations, so their count N is a nominal rather than an independent sample size. The dimension guard therefore serves only as an admissibility screen; held-out error selects the class.

Numerical checks. We verify the proof and inclusion relations with `verify_derivation.py` on cached validation forecasts from real cells, using the zero reference. Table 37 reports this check over the whole matrix. It pairs the ordering predicted by Theorem 3 with a cross-fitted finite-sample reference.

A single global scalar removes 1.2% of the spectral error at $C{=}321$, versus 8.4% for the coordinate-wise class and 89.0% for G_4. This comparison supplies the low-resolution endpoint of the approximation–estimation trade-off quantified below.

$G_5(w)$ and G_∞ are linear approximations at different resolutions. Class G_4 treats modes independently and therefore omits the cross-mode dependence induced by spectral leakage. Removing

that restriction within the linear family gives G_∞, a single map on the whole spectrum with $(2MC)^2$ parameters, which is 4.7×10^5 on ETT at H=96 but 7.1×10^9 on Traffic. Class $G_5(w)$ removes the restriction locally, letting each mode read the w modes centered on it. At spectral boundaries, out-of-range input indices are clamped to the nearest endpoint. The implementation therefore repeats endpoint inputs in its nominal block, reducing the number of distinct directions without shortening the block. The resulting support means that $G_5(w)$ remains distinct from G_∞ even at $w = M$. We compute the full-spectrum G_∞ reference wherever its $2MC$-dimensional pseudoinverse is computationally feasible: at H=96, $C \leq 224$ covers 60 of the 74 cells. Both classes approximate the unknown optimal post-hoc correction within a linear family. G_∞ supplies the highest-resolution reference.

In-sample fit and cross-fitted transfer diverge with resolution. This highest-resolution reference sharpens the tradeoff between approximation and estimation. Table 37 gives every class at the zero reference over all 74 cells and three seeds, as a percentage of that cell's spectral error in two columns: *in*, fitted and scored on the same second moments, and *xf*, fitted on one half of validation and scored on the other. The *in* column rises monotonically with inclusion at every width, as Theorem 3 predicts, reaching 99.8% at C=137 where G_∞ has more free directions than the split has windows. Cross-fitted performance instead deteriorates as the classes widen and becomes negative once cross-channel mixing begins. For example, at C=137, diagonal G_2 achieves $+4.2$ under cross-fitting, whereas G_3 falls to -47.3 and G_∞ to -1032.8. The in-sample column records how much each class can fit on the empirical moments and may include in-sample overfit. The cross-fitted column evaluates that fitted operator on held-out moments. Their gap provides an empirical finite-sample reference. Both columns use unregularized empirical pseudoinverses; deployed candidates add ridge, gates, and validation selection.

Table 37: Each class at the zero reference, mean spectral-error reduction over cell-seeds (%). Positive values indicate lower error. Cell counts per width in parentheses, three seeds each. *in* is fitted and scored on the same half, *xf* fitted on one half and scored on the other. These reference columns use unregularized pseudoinverses. Deployed candidates use ridge, gates, and validation selection. G_∞ is evaluated through C=137.

Unregularized finite-sample reference. Deployed candidates use ridge, gates, and validation selection.

		C=7 (40)		C=21 (10)		C=137 (10)		C=321 (10)		C=862 (4)	
class	dim	in	xf	in	xf	in	xf	in	xf	in	xf
G0	1	1.5	0.5	2.6	2.2	2.7	2.2	1.2	1.2	0.2	0.2
G1	M	3.7	2.0	3.3	2.6	4.6	4.1	3.2	3.1	1.0	0.7
G2	MC	5.5	−1.7	9.3	7.4	5.0	4.2	8.4	3.4	5.8	−6.1
G2$^{\mathbb{C}}$	$2MC$	5.7	−2.0	9.5	6.6	5.2	4.2	9.1	3.2	7.6	−12.9
G2′	$2MC$	5.6	−1.7	9.4	7.1	5.4	4.5	8.6	3.5	5.8	−6.4
G2″	$4MC$	5.9	−2.0	9.9	6.7	6.0	4.9	9.4	3.2	7.8	−14.3
G3	MC^2	14.7	−264.7	21.5	−106.3	47.3	−47.3	74.8	−174.7	90.0	−2964.6
G3′	$2MC^2$	15.0	−265.9	22.3	−113.4	50.3	−59.8	79.3	−219.8	94.3	−8225.3
G4	$4MC^2$	17.2	−269.0	26.7	−134.5	60.7	−119.5	89.0	−272.6	99.9	−479.0
G5(3)	$12MC^2$	24.9	−451.8	44.3	−305.5	86.5	−504.0	98.3	−1283.5	100.0	−54.0
G5(5)	$20MC^2$	27.2	−508.9	49.2	−429.4	91.5	−562.4	99.4	−2133.4	100.0	−43.1
G∞	$(2MC)^2$	47.5	−1165.4	78.9	−968.2	99.8	−1032.8	–	–	–	–

F Algorithmic Implementation and Computational complexity

F.1 Algorithm for One-Checkpoint Deployment

Algorithm 1 restates the selection rule in Section C.6. It operates on one frozen checkpoint and its labeled training and validation windows. For a class g, $\mathrm{FIT}_g(\mathcal{D}, \lambda)$ returns the prediction blocks $K_{g,j}$. Class G_2 uses Eq. 15, with $K_{2,j} = \mathrm{diag}(D_j, D_j)$ and $D_j = \mathrm{diag}(\widehat{\kappa}_{j1}, \ldots, \widehat{\kappa}_{jC})$. For G_3–$G_5(3)$, form the fitting moments $\widehat{S}_{g,j}$ and $\widehat{T}_{g,j}$ from Appendix C.5 and solve

$$\widehat{W}_{g,\lambda,j}\left(\widetilde{S}_{g,j} + \lambda\eta_{g,j} I\right) = \widehat{T}_{g,j}, \qquad \eta_{g,j} = \max\left\{\frac{\mathrm{tr}(\widehat{S}_{g,j})}{d_g}, \frac{\zeta}{n_f}\right\}, \tag{24}$$

where $n_f = |\mathcal{D}|$, $\zeta = 10^{-12}$, and $d_g = C, 2C, 6C$ for $G_3, G_4, G_5(3)$. A shared $\lambda > 0$ gives a positive penalty at every mode. The implementation accumulates unnormalized moment sums, so its trace floor ζ becomes ζ/n_f in this averaged notation. If the eigensolver returns $\widehat{S}_{g,j} = V\,\mathrm{diag}(\nu)V^\top$, it uses $\widetilde{S}_{g,j} = V\,\mathrm{diag}(\max\{\nu, 0\})V^\top$. In exact arithmetic $\widehat{S}_{g,j} \succeq 0$, so this changes nothing. Map the solver output to $K_{3,j} = \mathrm{diag}(\widehat{W}_{3,\lambda,j}, \widehat{W}_{3,\lambda,j})$ for G_3 and $K_{g,j} = \widehat{W}_{g,\lambda,j}$ for $G_4, G_5(3)$.

Algorithm 1: One-checkpoint covariance-resolution selection and correction

Input: frozen forecaster f_θ; labeled windows $\mathcal{D}_{\mathrm{tr}}$ and $\mathcal{D}_{\mathrm{val}}$; test inputs X_{te}; classes $\mathcal{G} = \{G_2, G_3, G_4, G_5(3)\}$; 64 log-spaced ridge multipliers $\Lambda \subset [10^{-4}, 10^5]$.

Output: corrected test forecast $\widetilde{S}_{\mathrm{te}}$.

1. Run f_θ once on each split. For every labeled calibration window, compute $s_{wj} = F_j(f_\theta(X_w))$ and $e_{wj} = s_{wj} - F_j(Y_w)$ for all rFFT modes j.
2. Construct stride-one validation windows with complete targets inside the validation segment. Split their chronological index list at its midpoint into A and B. Adjacent windows across the division share observations, including $H - 1$ target time points; neither half crosses the validation/test target boundary.
3. Define the two fitting-source recipes

$$\mathcal{P}_{\mathrm{tr}} = \{(\mathcal{D}_{\mathrm{tr}}, A), (\mathcal{D}_{\mathrm{tr}}, B)\}, \quad \mathcal{D}_{\mathrm{tr}}^{\mathrm{refit}} = \mathcal{D}_{\mathrm{tr}},$$
$$\mathcal{P}_{\mathrm{val}} = \{(A, B), (B, A)\}, \qquad \mathcal{D}_{\mathrm{val}}^{\mathrm{refit}} = \mathcal{D}_{\mathrm{val}}.$$

   Each ordered pair contains a fitting split followed by its scoring split.
4. For each source $r \in \{\mathrm{tr}, \mathrm{val}\}$, set $n_{\mathrm{guard}} = |\mathcal{D}_{\mathrm{tr}}|$ for $r = \mathrm{tr}$ and $n_{\mathrm{guard}} = \min(|A|, |B|)$ for $r = \mathrm{val}$. Retain G_2 unconditionally; retain G_3, G_4, and $G_5(3)$ only when $2n_{\mathrm{guard}} > C$, $n_{\mathrm{guard}} > 2C$, and $n_{\mathrm{guard}} > 6C$, respectively. Each fit uses its own $n_f = |\mathcal{D}_f|$ in Eq. 24.
5. For every retained class g, set $\Lambda_g = \{0\}$ for G_2 and $\Lambda_g = \Lambda$ otherwise. For each $\lambda \in \Lambda_g$ and each $(\mathcal{D}_f, \mathcal{D}_s) \in \mathcal{P}_r$, fit $\mathrm{FIT}_g(\mathcal{D}_f, \lambda)$ and evaluate the spectral score in Eq. 23 on $\mathcal{D}_s$. Accumulate the error $L_{r,g,\lambda,b}$ separately for each gate block b, where a block is one mode–channel coordinate for G_2 and one output mode for G_3–$G_5(3)$. Let L_b^0 be the corresponding pass-through error on the same scoring windows.
6. Select one global ridge multiplier for the class, $\lambda_{r,g} = \arg\min_{\lambda \in \Lambda_g} \sum_b L_{r,g,\lambda,b}$, and define its held-out gate $m_{r,g,b} = \mathbf{1}[L_{r,g,\lambda_{r,g},b} < L_b^0]$. The gated candidate score is

$$Q_{r,g} = \sum_b \left[m_{r,g,b} L_{r,g,\lambda_{r,g},b} + (1 - m_{r,g,b}) L_b^0\right].$$

7. Select the fitting source and covariance class by

$$(r^\star, g^\star) = \arg\min_{r,g} Q_{r,g}.$$

   Ties favor the narrower covariance class.
8. Refit $\widehat{K}_{\mathrm{deploy}} = \mathrm{FIT}_{g^\star}(\mathcal{D}_{r^\star}^{\mathrm{refit}}, \lambda_{r^\star,g^\star})$ and set every block with $m_{r^\star,g^\star,b} = 0$ to zero.
9. Compute $s_{\mathrm{te},j} = F_j(f_\theta(X_{\mathrm{te}}))$ and the class-specific correction $c_{\mathrm{te},j} = \widehat{K}_{\mathrm{deploy},j}\delta_{g^\star,j}(s_{\mathrm{te}})$. Return

$$\widetilde{S}_{\mathrm{te}} = \mathrm{irFFT}\left(\{s_{\mathrm{te},j} - c_{\mathrm{te},j}\}_{j=1}^M\right).$$

F.2 Complexity Analysis

The implementation separates coefficient storage, moment storage, and working memory during fitting. Table 38 gives the Traffic tensor sizes and the memory organization used by the optimized solver. Its four rows are the deployed candidate classes; the other structural classes remain defined in Table 9.

Moment construction and ridge-path reuse. For G_3, G_4, and $G_5(3)$, the predictor and output dimensions (d, q) are respectively (C, C), $(2C, 2C)$, and $(6C, 2C)$. Given cached spectra from n_f fitting windows, forming the moments costs $\mathcal{O}(Mn_f(d^2 + dq))$, and their eigendecomposition costs $\mathcal{O}(Md^3)$. The 64 ridge multipliers reuse that decomposition by changing the eigenvalue-dependent gains. Materializing each dense correction map costs $\mathcal{O}(Md^2q)$, followed by its application to validation forecasts. Coordinate class G_2 instead accumulates scalar numerator and denominator tensors in $\mathcal{O}(n_fMC)$ time, without a matrix factorization.

Working memory and the streamed path. Sample-wise accumulation avoids constructing a full real-stacked design over all fitting windows. The eigensolver processes mode batches with preallocated outputs. When a training-side $G_5(3)$ Gram tensor would exceed 8 GiB, the solver processes two center modes and 256 fitting windows at a time. For Traffic at $H = 720$, this changes the active Gram tensor from 35.974 GiB to 0.199 GiB. The same training fit serves both validation halves, and ridge scoring uses eigenbasis projections without materializing the full correction tensor. Large training-spectrum caches are CPU-backed when their estimated GPU residency exceeds the implementation's memory threshold. Peak fitting memory also includes eigenvectors, cross-moments, cached spectra, temporary projections, and solver workspace; it is not bounded by one Gram tensor.

The $H = 96$ benchmark uses the same two-mode, 256-window accumulation for $G_5(3)$, retaining the eigenbasis and selected dense map. Its coefficient tensor has shape $[49, 1724, 5172]$; the full and active Gram shapes are $[49, 5172, 5172]$ and $[2, 5172, 5172]$. This memory-bounded implementation reproduces the original training-side $G_5(3)$ selection and ridge multiplier 7.197 for all three DLinear seeds, with validation scores agreeing at the precision of the stored experiment records.

Fitting source and admissible classes. At $H = 96$, Traffic supplies 12,089 training windows and at least 830 windows per validation half. The validation-side guard excludes G_4 because $830 \leq 2C = 1{,}724$ and $G_5(3)$ because $830 \leq 6C = 5{,}172$. Both classes remain admissible on training data. Across the 30 Traffic backbone–seed runs at this horizon, selection retains training-fitted $G_5(3)$ in 18 runs, training-fitted G_4 in 3, and validation-fitted G_3 in 9.

Inference and timing boundaries. For each frozen backbone forecast $\widehat{Y} \in \mathbb{R}^{H\times C}$, inference consists of its rFFT, application of the fixed correction, and an inverse rFFT. The transforms cost $\mathcal{O}(CH \log H)$ per window. Spectral correction costs $\mathcal{O}(MC)$ for G_2 and $\mathcal{O}(MC^2)$ for G_3–$G_5(3)$ at fixed neighborhood width. Test labels are unnecessary for these operations. As specified in Table 10, moment fitting, ridge selection, gates, and class–source selection depend on training and validation data; test errors and the diagnostic oracle are computed in a separate evaluation branch. Fitting time excludes this branch, and inference latency includes the forecast transforms and correction rather than label-based scoring. For the measured Traffic/DLinear checkpoints at $H = 96$, all four candidate classes together take 112.24 ± 1.02 seconds from cached spectra; calibration rFFT preprocessing adds 0.55 ± 0.08 seconds. Measurements use FP32/complex64, PyTorch 2.13 with CUDA 13.0, and eight CPU threads on an AMD EPYC 9654 host. Table 38 reports peak PyTorch-allocated GPU memory rather than allocator-reserved memory or one tensor's size.

Table 38: Traffic coefficient storage and measured computational costs ($C = 862$). (a) Nominal counts and dense FP32 tensor sizes at $H = 96$, $M = 49$; the Gram column counts one design second-moment tensor (its diagonal for G_2). (b) Measurements on real DLinear checkpoints, seeds $2, 3, 4$, each on an otherwise idle A100-SXM4-40GB GPU. Entries are mean $\pm$ sample SD across the three runs. Peak GPU memory is PyTorch-allocated memory, including the 8.665 GiB resident calibration cache and baseline tensors; the extra-peak column subtracts that baseline. (c) Calculated $G_5(3)$ Gram storage when processing two center modes at a time. MiB and GiB denote 2^{20} and 2^{30} bytes; tensor storage is distinct from measured peak memory.

(a) Coefficients and full Gram, $H = 96$

class	nominal count	Traffic count	coefficients (MiB)	full Gram (MiB)
G_2	MC	42,238	0.161	0.161
G_3	MC^2	36,409,156	138.890	138.890
G_4	$4MC^2$	145,636,624	555.560	555.560
$G_5(3)$	$12MC^2$	436,909,872	1,666.679	5,000.037

(b) Measured fit, selection, refit, and frozen inference, $H = 96$

class	peak GPU (GiB)	extra peak (GiB)	fit + select + refit (s)	inference (ms/window)
G_2	$16.54_{\pm 0.00}$	$7.87_{\pm 0.00}$	$0.30_{\pm 0.12}$	$0.089_{\pm 0.035}$
G_3	$17.08_{\pm 0.00}$	$8.41_{\pm 0.00}$	$8.74_{\pm 0.24}$	$0.376_{\pm 0.001}$
G_4	$18.71_{\pm 0.00}$	$10.04_{\pm 0.00}$	$11.88_{\pm 0.11}$	$0.575_{\pm 0.001}$
$G_5(3)$	$23.01_{\pm 0.00}$	$14.35_{\pm 0.00}$	$91.33_{\pm 0.94}$	$1.641_{\pm 0.002}$

(c) Calculated $G_5(3)$ Gram storage across horizons

H	M	full Gram (GiB)	two-mode Gram (GiB)
96	49	4.883	0.199
192	97	9.666	0.199
336	169	16.841	0.199
720	361	35.974	0.199

At $H = 96$, training provides $12{,}089$ fitting windows; the smaller validation half contains 830. Fit + select + refit sums all admissible sources: train and validation for G_2/G_3, train for $G_4/G_5(3)$. It includes the two-fold scores, global 64-point ridge path for G_3–$G_5(3)$, gates, and full-source refit from cached spectra. The $G_5(3)$ measurement uses two-mode, 256-window moment accumulation and retains the eigenbasis and selected dense map. Inference includes rFFT, fixed correction, and inverse rFFT for one cached forecast, with 30 warm-ups and 200 synchronized repetitions per seed. Backbone forward passes, CPU input transfers at inference, and test metric evaluation are excluded.