COMPLEMENTARY RETRIEVAL GEOMETRIES FOR ENHANCING MULTIVARIATE TIME SERIES FORECASTING

Anonymous authors
Paper under double-blind review

ABSTRACT

Retrieval-augmented forecasting provides multivariate time-series forecasters with additional evidence by reusing trajectories that follow similar historical states; its effectiveness depends on how predictive historical similarity is defined. Existing methods typically organize retrieval around one principal criterion, causing a target variable's temporal shape and its dynamic role in a multivariate system to share the same similarity definition. We propose CoReG, a complementary retrieval framework for multivariate time-series forecasting. CoReG constructs two semantically distinct retrieval geometries: TimeGeo matches the target variable's multiscale temporal shape, while RoleGeo matches its system role using stable cross-variate relations and time-aligned role responses. The two paths independently retrieve the target variable's own historical future segments, form separate views relative to the same frozen backbone forecast, and produce the final prediction through symmetric merging. Across ten datasets and six forecasting backbones, CoReG improves both MSE and MAE in 58 of 60 paired settings, with mean relative reductions of 14.94% and 10.52%, respectively. Anonymous code is available at `https://anonymous.4open.science/r/CoReg-3CCD/`.

1 INTRODUCTION

Multivariate time-series forecasting generates multiple future steps from historical observations of multiple variables and has been applied to settings such as energy-load and traffic forecasting (Lai et al., 2018; Song et al., 2020). Modern forecasters learn temporal dependencies from the current history window and explicitly model cross-variate dependencies to generate the future sequence (Zhang & Yan, 2023; Liu et al., 2024b). Retrieval-augmented forecasting adds another source of information by searching the training history for reference segments and using their subsequent values as external evidence for the current prediction. MQ-ReTCNN, ReTime, RATD, RAFT, and PFRP all demonstrate the predictive value of such historical evidence (Yang et al., 2022; Jing et al., 2022; Liu et al., 2024a; Han et al., 2025; Du et al., 2026). Retrieval-augmented forecasting therefore faces a basic question: ***what makes a historical state similar for future prediction?***

Existing retrieval methods typically organize candidate segments according to historical patterns, representational similarity, or global temporal structure (Yang et al., 2022; Jing et al., 2022; Liu et al., 2024a; Han et al., 2025; Cao et al., 2026). This single-geometry retrieval paradigm uses one retrieval space to determine both which candidates are selected and what their evidence means. For multivariate series, whether one geometry can express every predictive historical analogy remains an open question.

① ***Temporal resemblance is only one source of predictive analogy.*** Historical analogies can arise from the target variable's temporal shape, including trends, periodic phases, local fluctuations, and transitions at different time scales. Multi-scale forecasting studies decompose or aggregate dynamics over different temporal ranges to extract local and global temporal structure (Wu et al., 2021; Wang et al., 2023). Whether the target variable has exhibited a similar waveform in the past therefore provides a natural basis for retrieving future responses. However, this view describes only how the target variable evolves in isolation; it does not characterize its relative state within the surrounding multivariate system.

② ***Predictive analogy can also arise from system role.*** In a multivariate system, synchronized changes, cross-variate dependencies, and lead–lag relations between the target and other variables

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.

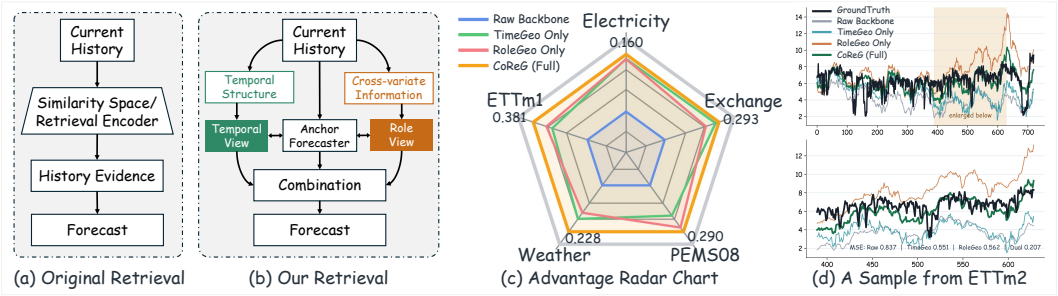


Figure 1: From a single retrieval space to complementary retrieval geometries. (a) A single retrieval paradigm selects historical evidence using one notion of similarity. (b) CoReG retrieves evidence from temporal-shape and system-role views to enhance a shared forecaster. (c) The relative strengths of the two views vary across datasets, while their combination delivers more consistently low errors. (d) A representative window shows their distinct corrections and complementary result. A larger radar radius indicates lower error.

also carry predictive information. Studies of spatiotemporal synchronization, cross-dimension dependence, and leading indicators show that these relations can improve target prediction (Song et al., 2020; Zhang & Yan, 2023; Liu et al., 2024b; Zhao & Shen, 2024). Consequently, two target windows can have different local waveforms yet correspond to useful future responses when they occupy similar dynamic roles. We refer to these two sources of similarity as temporal shape and system role.

③ ***No single retrieval geometry is sufficient across forecasting conditions.*** Figure 1(c) shows that the relative strengths of the temporal-shape and system-role views vary across datasets, with no fixed ordering between the two branches. This variation indicates that predictive historical analogies depend on the dataset and forecasting condition: some conditions favor matching the target variable's temporal shape, whereas others rely more on its dynamic role in the system. A single retrieval geometry therefore cannot reliably cover all forecasting states, motivating complementary retrieval geometries.

These observations indicate that the central challenge is to represent heterogeneous predictive analogies while retaining reliable retrieval matches. Temporal shape and system role correspond to different predictive geometries; placing them in one space mixes the sources of historical evidence and weakens state-specific evidence reading. This motivates the central question of this paper: ***can independent retrieval geometries preserve the semantics of different predictive similarities and use them as complementary historical evidence for joint forecasting?***

To answer this question, we propose CoReG (Complementary Retrieval Geometries), which decomposes predictive historical similarity into two semantically independent retrieval geometries: temporal shape and system role. TimeGeo searches for historical references according to the target variable's multiscale temporal shape, whereas RoleGeo matches historical states according to the target's dynamic role in the multivariate system. The two paths retrieve historical future segments of the target variable, form independent enhanced views relative to the same frozen Backbone forecast, and produce the final prediction through symmetric fusion. Experiments across ten datasets and six forecasting Backbones show that both geometries provide useful predictive evidence, their relative effects vary with the dataset, forecasting horizon, and prediction state, and their combination yields more consistent forecasting gains. Based on this design and these empirical findings, our main contributions are summarized as follows:

① ***Conceptual Perspective.*** We introduce a complementary retrieval-geometry perspective for multivariate retrieval-augmented forecasting, organizing historical similarity into temporal shape and system role.

② ***Methodological Framework.*** We construct two independent retrieval paths, TimeGeo and RoleGeo, so that target-local temporal structure and cross-variate dynamics separately participate in selecting target-owned future-value evidence, while joint forecasting preserves their complementarity.

Table 1: Concise comparison between CoReG and directly related retrieval-augmented forecasting methods. Methods are ordered by publication year. Cross-variate retrieval evidence indicates that retrieval explicitly uses structural or role information across variables.

Method	Retrieval criterion	Multivariate Forecasting	Plug-and-Play	Cross-variate
RATD (2024)	Pre-trained representation distance	✓	×	×
RAFT (2025)	Historical-pattern similarity	✓	×	×
PFRP (2026)	Representation-feature similarity	×	✓	×
GTR (2026)	Global temporal representation	✓	✓	×
CoReG (Ours)	Temporal shape + system role	✓	✓	✓

③ ***Empirical Findings.*** We show that TimeGeo and RoleGeo provide differentiated predictive evidence whose relative strengths vary across forecasting conditions, while their combination yields reliable improvements across datasets and forecasting Backbones.

2 Related Works

2.1 Multivariate Time Series Forecasting and Cross-Variate Modeling

Multivariate time series forecasting builds predictors through temporal representations and cross-variate interactions. DLinear shows that simple linear mappings can form strong forecasting baselines (Zeng et al., 2023); PatchTST extends the effective history through patching and channel-independent modeling (Nie et al., 2023); iTransformer treats variables as tokens and learns their dependencies with attention (Liu et al., 2024b); and Crossformer jointly captures temporal and cross-variate dependencies (Zhang & Yan, 2023). Recent work continues this line: PatchMLP and CSformer coordinate channel-independent processing with cross-variate mixing (Tang & Zhang, 2025; Wang et al., 2025), PHAT models heterogeneous periodicity across variables (Ma et al., 2026), and ChannelTokenFormer jointly addresses channel dependencies, asynchronous sampling, and missing observations (Jang et al., 2026). These methods mainly forecast from the joint representation of the current multivariate input.

2.2 Retrieval-Augmented Forecasting

Retrieval-augmented forecasting extracts reference evidence from historical data or external knowledge to complement a local history window. RATD injects complete reference series with similar representations into diffusion denoising (Liu et al., 2024a); RAFT retrieves pattern-similar historical candidates and their subsequent values (Han et al., 2025); PFRP uses a fixed-size global memory bank to provide retrieval forecasts for arbitrary local univariate forecasters (Du et al., 2026); and GTR attaches global-periodic representations to multivariate forecasting backbones as an add-on module (Cao et al., 2026). As summarized in Table 1, CoReG further uses temporal shape and system role as complementary retrieval criteria, providing cross-variate retrieval evidence for multivariate forecasting backbones.

3 Method

3.1 Overview and Paired Anchor

CoReG targets multivariate time-series forecasting. Given a historical window $X \in \mathbb{R}^{N\times L\times C}$, the task is to produce $\widehat{Y} \in \mathbb{R}^{N\times H\times C}$ over a horizon of length H, where N, L, and C denote batch size, history length, and the number of variables, respectively. For a target variable, predictive historical analogies arise from both its temporal shape and its dynamic role in the multivariate system. CoReG therefore constructs two complementary predictive retrieval geometries, TimeGeo and RoleGeo, and uses them as an attachable source of forecasting evidence for any paired backbone. As shown in Figure 2, the framework combines a shared Anchor forecast, two independent retrieval-enhancement paths, and a final symmetric merger.

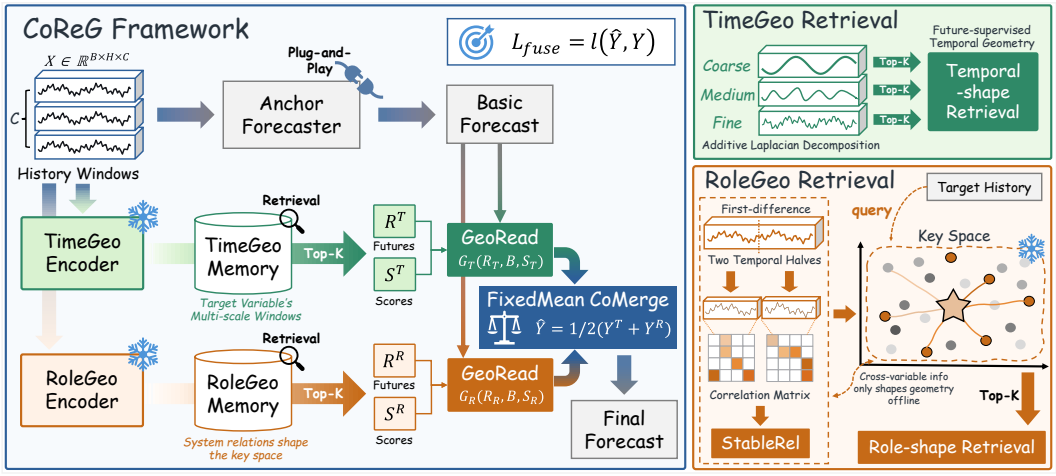


Figure 2: Overview of CoReG. A frozen Anchor provides the shared forecast for both paths. TimeGeo and RoleGeo retrieve historical predictive evidence according to temporal shape and system role, and independent GeoRead modules form two enhanced views. FixedMean CoMerge combines these views into the final prediction. The right panels summarize the construction of the two retrieval geometries.

CoReG uses the trained checkpoint of the paired backbone as an Anchor Forecaster A. The Anchor remains frozen during fusion training and inference, and supplies the shared Anchor forecast

$$B_{\text{anc}} = A(X), \qquad B_{\text{anc}} \in \mathbb{R}^{N\times H\times C}. \tag{1}$$

For target variable c, TimeGeo and RoleGeo produce retrieval forecasts R_T, R_R and their adjusted Top-K scores S_T, S_R, respectively. Two independent GeoRead modules form complete forecast views, and FixedMean CoMerge averages these views symmetrically:

$$Y_T = G_T(R_T, B_{\text{anc}}, S_T), \qquad Y_R = G_R(R_R, B_{\text{anc}}, S_R), \tag{2}$$

$$\widehat{Y} = \frac{1}{2}(Y_T + Y_R). \tag{3}$$

Both paths use target-local future ownership. When forecasting variable c, memory values come from training-period future windows of variable c; cross-variate information enters the retrieval key space of RoleGeo and provides system-relation evidence for the target variable.

3.2 TimeGeo: Multiscale Temporal-Shape Retrieval

TimeGeo retrieves multiscale temporal shapes from the target variable itself. Slow evolution, intermediate transitions, and local perturbations in a historical window can carry distinct predictive implications. TimeGeo establishes an independent similarity comparison for each component, then reconstructs their forecasting evidence into one future window.

Let $P_s(u)$ denote non-overlapping block averaging of a one-dimensional window u with block length s, with each average repeated to the original length. With the scale set $\mathcal{S} = (4, 2, 1)$, TimeGeo constructs three additive components:

$$\ell_4(u) = P_4(u), \qquad \ell_2(u) = P_2(u) - P_4(u), \qquad \ell_1(u) = u - P_2(u). \tag{4}$$

They satisfy

$$u = \ell_4(u) + \ell_2(u) + \ell_1(u). \tag{5}$$

Here, ℓ_4 retains slow variation, ℓ_2 expresses intermediate transitions, and ℓ_1 captures local residuals. For each scale $s \in \mathcal{S}$, TimeGeo encodes the normalized $\ell_s(x_{b,c})$ as a query key and retrieves Top-K historical candidates from the TimeGeo memory of variable c. It applies the same decomposition to

the training-period future value of each candidate and uses adjusted scores to obtain a scale forecast $\tilde{r}_{b,c}^{(s)}$. Independent retrieval at the three scales is followed by additive reconstruction:

$$R_{T,b,:,c} = \sum_{s \in \mathcal{S}} \tilde{r}_{b,c}^{(s)}. \tag{6}$$

The predictive encoder of TimeGeo is trained with predictive supervision from the target variable's own future. After training, its encoder and memory are fixed, and online queries retrieve predictive evidence with similar temporal shapes from the target variable's history.

3.3 RoleGeo: Stable System-Role Retrieval

RoleGeo retrieves the dynamic role of a target variable within the multivariate system. Similar local temporal segments can occur in distinct system states. RoleGeo constructs role-predictive supervision from stable cross-variate relations in the training period, so that its retrieval keys capture both target future behavior and system role.

For a training series $V \in \mathbb{R}^{T \times C}$, RoleGeo first computes first-order differences $D_{t,c} = V_{t+1,c} - V_{t,c}$ and partitions the time axis into two contiguous regions. Let $R^{(1)}$ and $R^{(2)}$ be relation matrices from standardized differences in the two regions. StableRel retains signed relations supported by both regions:

$$S_{ij} = \sqrt{\max(R_{ij}^{(1)} R_{ij}^{(2)}, 0)}, \qquad Q_{ij} = \operatorname{sign}(R_{ij}^{(1)} + R_{ij}^{(2)}) S_{ij}. \tag{7}$$

Removing the diagonal and normalizing yield relation weights W and a variable-level relation consistency $c_i \in [0, 1]$. Using variable-wise normalized first differences from the complete training period, the time-aligned role response of variable i is

$$E_{t,i} = \bar{D}_{t,i} - \sum_{j=1}^{C} W_{ij} \bar{D}_{t,j}. \tag{8}$$

The response captures target dynamics after accounting for stable system relations. For a target future $f_{t,i}$ and its same-start role-response future $e_{t,i}$, RoleGeo normalizes both and uses the following role-predictive signature to define historical analogies:

$$z_{t,i} = \operatorname{concat}\left(\sqrt{1 - c_i}\, \overline{f}_{t,i}, \sqrt{c_i}\, \overline{e}_{t,i}\right). \tag{9}$$

The online input to RoleEncoder is target history $h_{t,i}$, and contrastive positives are determined by $z_{t,i}$. Its key space is organized by future behavior and system role, while RoleGeo memory stores normalized future values of the target variable itself. The retrieved values therefore provide target-local predictive evidence.

3.4 Geometry-Specific Readout and FixedMean CoMerge

The two retrieval geometries have distinct reliability conditions, so CoReG assigns them independent GeoRead modules. For path $v \in \{T, R\}$, GeoRead maps its adjusted Top-K scores S_v to two coefficients for the retrieval forecast and the Anchor forecast. These coefficients are normalized only within the corresponding GeoRead; no cross-path normalization is performed between TimeGeo and RoleGeo:

$$\beta_v[b, c, :] = \operatorname{Softmax}\left(W_2^{(v)} \operatorname{GELU}(W_1^{(v)} S_v[b, c, :] + b_1^{(v)}) + b_2^{(v)}\right). \tag{10}$$

The coefficients for a given (b, c) are shared across all forecast steps, yielding the path view

$$Y_v[b, h, c] = \beta_v[b, c, 0] R_v[b, h, c] + \beta_v[b, c, 1] B_{\text{anc}}[b, h, c]. \tag{11}$$

Each GeoRead determines how strongly evidence from its path enhances the Anchor. Since both path views already include the Anchor forecast, FixedMean CoMerge applies a fixed symmetric average:

$$\widehat{Y} = \frac{1}{2}(Y_T + Y_R). \tag{12}$$

This merger preserves symmetric roles for the two retrieval geometries and forms the final prediction from their complementary complete evidence views.

3.5 Training and Information Boundary

CoReG directly optimizes the final forecast loss during fusion training. Let ℓ denote the loss used to train the paired Backbone. The training objective is

$$\mathcal{L}_{\text{fuse}} = \ell(\widehat{Y}, Y). \tag{13}$$

Training first constructs or loads the paired Anchor, then trains the predictive encoders of TimeGeo and RoleGeo and builds their two memories. Fusion training fixes the Anchor, both predictive encoders, and all memory tensors, and updates the non-encoder readout parameters in both retrieval paths together with the fusion parameters. At inference, the Anchor first produces B_{anc}, the target history independently queries the two memories, the GeoRead modules produce Y_T and Y_R, and FixedMean CoMerge returns $\widehat{Y}$.

Online forecasting uses the current history, frozen Anchor, and training-period memories. Inference reads no test future or test label and performs no parameter update. StableRel is formed during training and provides stable system-role evidence through the RoleGeo key space.

4 Experiments

This section evaluates CoReG from three perspectives: forecasting effectiveness, the complementarity of TimeGeo and RoleGeo, and its ability to enhance different forecasting backbones. Experiments span multiple datasets, forecasting horizons, and model architectures; experimental settings, per-horizon results, complete ablations, and additional comparisons are provided in Appendix C, Appendix D, and Appendix E.

4.1 Experimental Settings

Datasets. We use ten datasets (ETTh1, ETTh2, ETTm1, ETTm2, Weather, Electricity, Exchange, PEMS04, PEMS08, and Solar), covering energy, weather, traffic, and financial multivariate settings. These datasets vary substantially in sequence length, number of variables, periodic structure, and cross-variate correlation, allowing evaluation under different temporal structures and system relations.

Baselines. CoReG is paired with DLinear (Zeng et al., 2023), MICN (Wang et al., 2023), iTransformer (Liu et al., 2024b), PatchMLP (Tang & Zhang, 2025), xPatch (Stitsyuk & Choi, 2025), and Amplifier (Fei et al., 2025). These six Backbones cover linear, convolutional multi-scale, attention-based, patch-based MLP, and hybrid architectures, allowing us to examine whether retrieval enhancement adapts to different forecasting inductive biases.

Implementation Details. The input length is fixed to 720, and we evaluate horizons 96, 192, 336, and 720 using MSE and MAE. Backbones are trained for up to 10 epochs with learning rate 0.0001, batch size 64, and early-stopping patience 3; the CoReG forecast encoder is trained for 80 epochs, followed by at most 10 fusion-training epochs. Unless otherwise stated, comparisons use the same data splits, preprocessing, input lengths, and forecast windows.

4.2 Main Results

This subsection examines whether CoReG consistently enhances forecasters with different inductive biases. For each Dataset–Backbone pair, the Backbone checkpoint and evaluation protocol are held fixed; CoReG is added to the original forecaster and the resulting error change is measured. Table 2 summarizes the paired results across four forecasting horizons to evaluate the applicability of complementary retrieval across architectures and data conditions.

Table 2 shows that CoReG reduces MSE in 58 of 60 paired settings and MAE in all 60 settings, with both metrics improving jointly in 58 settings and mean reductions of 14.94% and 10.52%, respectively. Gains are most pronounced on Exchange, Electricity, and the PEMS datasets, while improvements on ETT and Weather are more moderate. This consistency across architectures shows that complementary retrieval can adapt to forecasters with different inductive biases.

Table 2: Paired CoReG results across six forecasting Backbones, averaged over four horizons. +CoReG also reports relative changes; green and red denote decreased and non-decreased errors.

Dataset	Variant	DLinear (2023)		MICN (2023)		iTransformer (2024)		PatchMLP (2025)		xPatch (2025)		Amplifier (2025)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	Original	0.5123	0.5142	0.5337	0.5441	0.5388	0.5338	0.5761	0.5541	0.5194	0.5099	0.6129	0.5721
	+CoReG	0.5020 ↓2.01%	0.5085 ↓1.11%	0.5182 ↓2.90%	0.5320 ↓2.22%	0.5143 ↓4.55%	0.5193 ↓2.72%	0.5319 ↓7.67%	0.5274 ↓4.82%	0.5114 ↓1.54%	0.5059 ↓0.78%	0.5218 ↓14.86%	0.5222 ↓8.72%
ETTh2	Original	0.3231	0.3989	0.3348	0.4101	0.3278	0.3943	0.3918	0.4289	0.3033	0.3678	0.3621	0.4182
	+CoReG	0.2932 ↓9.25%	0.3772 ↓5.44%	0.2934 ↓12.37%	0.3791 ↓7.56%	0.3119 ↓4.85%	0.3827 ↓2.94%	0.3589 ↓8.40%	0.4090 ↓4.64%	0.2964 ↓2.27%	0.3640 ↓1.03%	0.3156 ↓12.84%	0.3847 ↓8.01%
ETTm1	Original	0.3964	0.4193	0.4019	0.4255	0.4083	0.4363	0.4163	0.4421	0.3955	0.4164	0.4020	0.4277
	+CoReG	0.3810 ↓3.88%	0.4155 ↓0.91%	0.3833 ↓4.63%	0.4185 ↓1.65%	0.3914 ↓4.14%	0.4267 ↓2.20%	0.3965 ↓4.76%	0.4291 ↓2.94%	0.3895 ↓1.52%	0.4134 ↓0.72%	0.3900 ↓2.99%	0.4218 ↓1.38%
ETTm2	Original	0.2010	0.2983	0.2103	0.3079	0.2198	0.3130	0.2275	0.3157	0.1961	0.2884	0.2054	0.3018
	+CoReG	0.2010 ↑0.00%	0.2957 ↓0.87%	0.2037 ↓3.14%	0.2986 ↓3.02%	0.2098 ↓4.55%	0.3047 ↓2.65%	0.2161 ↓5.01%	0.3070 ↓2.76%	0.1943 ↓0.92%	0.2862 ↓0.76%	0.1999 ↓2.68%	0.2968 ↓1.66%
Weather	Original	0.2380	0.2891	0.2421	0.2938	0.2389	0.2788	0.2297	0.2711	0.2212	0.2508	0.2388	0.2790
	+CoReG	0.2282 ↓4.12%	0.2747 ↓4.98%	0.2283 ↓5.70%	0.2747 ↓6.50%	0.2349 ↓1.67%	0.2753 ↓1.26%	0.2290 ↓0.30%	0.2705 ↓0.22%	0.2208 ↓0.18%	0.2505 ↓0.12%	0.2329 ↓2.47%	0.2734 ↓2.01%
Exchange	Original	0.5786	0.5423	0.9735	0.7638	0.5018	0.4888	0.6953	0.5509	0.4706	0.4562	0.4574	0.4721
	+CoReG	0.2929 ↓49.38%	0.4012 ↓26.02%	0.3376 ↓65.32%	0.4364 ↓42.86%	0.4504 ↓10.24%	0.4617 ↓5.54%	0.5581 ↓19.73%	0.5048 ↓8.37%	0.4380 ↓6.93%	0.4453 ↓2.39%	0.4525 ↓1.07%	0.4626 ↓2.01%
Electricity	Original	0.3065	0.3729	0.3913	0.4305	0.1698	0.2686	0.4161	0.4123	0.3145	0.3697	0.4345	0.4222
	+CoReG	0.1595 ↓47.96%	0.2556 ↓31.46%	0.1700 ↓56.56%	0.2754 ↓36.03%	0.1576 ↓7.18%	0.2526 ↓5.96%	0.1617 ↓61.14%	0.2586 ↓37.28%	0.1626 ↓48.30%	0.2550 ↓31.03%	0.1610 ↓62.95%	0.2554 ↓39.51%
PEMS04	Original	0.2553	0.3453	0.2609	0.3502	0.1759	0.2870	0.1663	0.2831	0.1931	0.2815	0.3207	0.4078
	+CoReG	0.1460 ↓42.81%	0.2549 ↓26.18%	0.1464 ↓43.89%	0.2551 ↓27.16%	0.1418 ↓19.39%	0.2507 ↓12.65%	0.1392 ↓16.30%	0.2498 ↓11.76%	0.1477 ↓23.51%	0.2484 ↓11.76%	0.1487 ↓53.63%	0.2561 ↓37.20%
PEMS08	Original	0.4172	0.3905	0.2757	0.2826	0.3099	0.3062	0.3247	0.3034	0.3766	0.3287	0.4829	0.4235
	+CoReG	0.2902 ↓30.44%	0.2837 ↓27.35%	0.2494 ↓9.54%	0.2620 ↓7.29%	0.2937 ↓5.23%	0.2754 ↓10.06%	0.2987 ↓8.01%	0.2740 ↓9.69%	0.3274 ↓13.06%	0.2835 ↓13.75%	0.3150 ↓34.77%	0.2818 ↓33.46%
Solar	Original	0.2127	0.2747	0.2054	0.2599	0.1985	0.2651	0.1915	0.2505	0.1924	0.2213	0.2040	0.2620
	+CoReG	0.1913 ↓10.06%	0.2345 ↓14.63%	0.1994 ↓2.92%	0.2515 ↓3.23%	0.1880 ↓5.29%	0.2339 ↓11.77%	0.1877 ↓1.98%	0.2416 ↓3.55%	0.1994 ↑3.64%	0.2120 ↓4.20%	0.1906 ↓6.57%	0.2344 ↓10.53%

Table 3: CoReG ablation with DLinear. Entries average four horizons; the last two columns report mean changes from FULL across the four datasets. Positive values indicate increased error; green and red denote values below and not below FULL.

Variant	ETTm1		Weather		Exchange		PEMS08		Avg. Δ	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	ΔMSE	ΔMAE
FULL	0.3810	0.4155	0.2282	0.2747	0.2929	0.4012	0.2902	0.2837	—	—
Raw Backbone	0.3964	0.4193	0.2380	0.2891	0.5786	0.5423	0.4172	0.3905	+37.40%	+19.73%
Retrieval-only CoReG	0.4693	0.4874	0.3072	0.3360	0.9592	0.7350	0.6331	0.5349	+100.86%	+52.84%
Same-channel Retrieval	0.3829	0.4169	0.2308	0.2769	0.3039	0.4077	0.2931	0.2858	+1.59%	+0.88%
TimeGeo Only	0.3840	0.4172	0.2297	0.2760	0.2947	0.4028	0.2968	0.2863	+1.08%	+0.54%
RoleGeo Only	0.3830	0.4168	0.2305	0.2775	0.3007	0.4043	0.2917	0.2867	+1.18%	+0.78%
w/ Shifted Peer Responses	0.3807	0.4154	0.2441	0.2868	0.2942	0.4024	0.2899	0.2836	+1.80%	+1.15%
w/o Stable Relations	0.3816	0.4158	0.2288	0.2752	0.2993	0.4056	0.2938	0.2851	+0.96%	+0.47%
w/ Shared Readout	0.3815	0.4158	0.2287	0.2755	0.2933	0.4015	0.2909	0.2841	+0.18%	+0.14%

4.3 Ablation Studies

Table 3 compares the complete model with eight DLinear variants across ETTm1, Weather, Exchange, and PEMS08, covering the interaction between retrieved evidence and Backbone forecasts, TimeGeo–RoleGeo complementarity, RoleGeo relation construction, temporal alignment, and independent path readout.

Raw Backbone, Retrieval-only CoReG, and Same-channel Retrieval show that retrieved evidence forms stable gains when combined with a parametric Backbone. TimeGeo Only and RoleGeo Only both retain effective forecasting ability, while FULL provides more consistent low error across the four datasets, reflecting the complementarity of the two geometries. Removing Stable Relations or temporal alignment increases error, and Shared Readout is also higher than independent readout, showing that relation construction, temporal alignment, and path-specific readout jointly affect prediction quality.

4.4 How Do the Two Geometries Form Complementary Predictive Evidence?

This subsection analyzes how the complementarity of TimeGeo and RoleGeo emerges during inference. Figure 3 follows three levels: GeoRead behavior for the two evidence sources, coordinated variable-pair improvements produced by RoleGeo, and the distinct trajectory corrections and joint result within a concrete forecasting window.

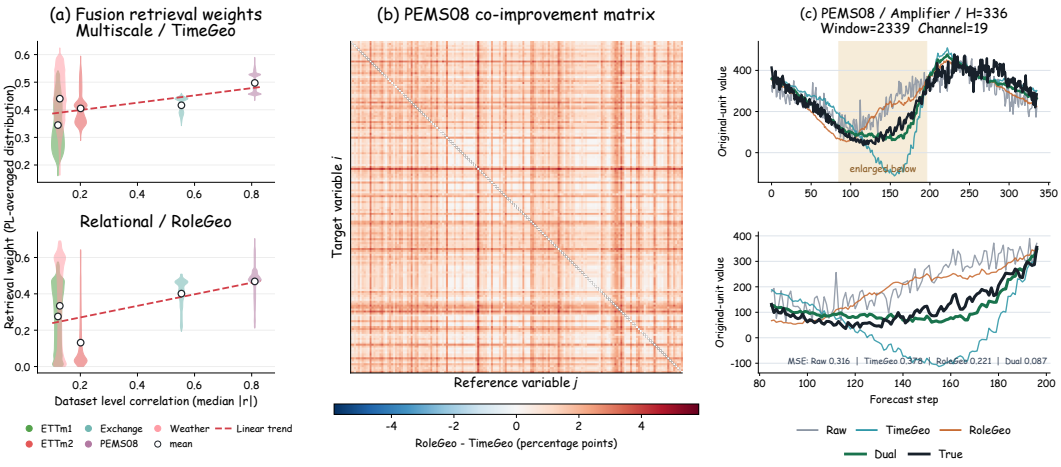


Figure 3: Complementary predictive evidence from TimeGeo and RoleGeo. (a) GeoRead coefficients and dataset correlation structure. (b) RoleGeo-minus-TimeGeo variable-pair co-improvement rates on PEMS08. (c) Distinct corrections and joint prediction in a representative PEMS08/Amplifier/Horizon-336 window.

Retrieval evidence is read differently across dataset structures. Figure 3(a) shows substantial cross-dataset variation in both GeoRead retrieval coefficients. PEMS08 and Exchange retain relatively high average coefficients for both retrieved forecasts, whereas ETTm2 has a markedly lower RoleGeo coefficient; both paths exhibit positive cross-dataset linear trends. This provides a descriptive link between stronger cross-variate structure and greater use of retrieved forecasts. The remaining differences among datasets with similar correlation levels further indicate that the realized readout also reflects dataset-specific dynamics.

RoleGeo produces coordinated gains across related variables. Figure 3(b) moves the analysis to variable pairs. The PEMS08 co-improvement matrix contains broad positive regions, indicating that RoleGeo more frequently reduces window-level MSE for both variables in a pair than TimeGeo. Its stripe structure concentrates this advantage in groups of variables, showing how the system roles learned by RoleGeo appear as coordinated cross-variate improvements.

Complementary corrections translate into a better forecast. Figure 3(c) shows how the two evidence sources interact within one window. This case connects the conditioned readout and cross-variate improvement patterns to the final forecast: the two geometries provide distinct error corrections whose combination produces a more accurate trajectory.

4.5 Plug-and-Play Comparison

This subsection evaluates CoReG as a plug-in enhancement module in terms of forecasting accuracy and computational resources. We compare CoReG with SARAF (Zhou et al., 2026), GTR (Cao et al., 2026), the online δ-Adapter protocol (Liang et al., 2026), and GLAFF (Wang et al., 2024). Table 4 evaluates forecasting accuracy with MICN, while Figure 4 examines the corresponding latency and memory costs through a representative case and all-setting distributions.

Table 4 shows that CoReG reduces both MSE and MAE for MICN on all five datasets and attains the lowest error in $17/20$ settings for each metric; SARAF is particularly effective on Exchange, while the δ-Adapter is competitive on selected Weather and PEMS08 metrics. Figure 4 further shows that CoReG achieves a clear improvement in the representative iTransformer case at latency close to Raw and with modest additional memory; across the DLinear and MICN settings, it also has the highest median MSE improvement at median latency around one to two times Raw. SARAF, δ-Adapter, and GLAFF have median latencies several to tens of times Raw, whereas GTR has lower latency but a median improvement near zero. Together, the accuracy and resource results show a more favorable accuracy–latency trade-off for CoReG across Backbones.

Table 4: Average plug-and-play comparison on MICN. Dataset rows report MSE and MAE averaged over four horizons; summary rows report mean relative gain, improved cases, and best cases across 20 settings. Green and red indicate errors below and not below Raw, respectively; best and second-best results are bold and underlined.

Dataset	Horizon	Raw		CoReG		SARAF (2026)		GTR (2026)		δ-Adapter (2026)		GLAFF (2024)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	Avg.	0.5337	0.5441	**0.5182** ↓2.90%	**0.5320** ↓2.22%	0.5433 ↑1.80%	0.5408 ↓0.61%	0.5412 ↑1.41%	0.5512 ↑1.30%	0.5427 ↑1.69%	0.5436 ↓0.09%	0.5690 ↑6.61%	0.5630 ↑3.47%
ETTh2	Avg.	0.3348	0.4101	**0.2934** ↓12.37%	**0.3791** ↓7.56%	0.3208 ↓4.18%	0.3905 ↓4.78%	0.3373 ↑0.75%	0.4173 ↑1.76%	0.3417 ↑2.06%	0.4114 ↑0.32%	0.4270 ↑27.54%	0.4644 ↑13.24%
Exchange	Avg.	0.9735	0.7638	**0.3376** ↓65.32%	**0.4364** ↓42.86%	0.4574 ↓53.01%	0.4492 ↓41.19%	1.0303 ↑5.83%	0.7048 ↓7.72%	0.8743 ↓10.19%	0.7116 ↓6.83%	1.2927 ↑32.79%	0.8645 ↑13.18%
Weather	Avg.	0.2421	0.2938	**0.2283** ↓5.70%	**0.2747** ↓6.50%	0.2491 ↑2.89%	0.3000 ↑2.11%	0.2435 ↑0.58%	0.2946 ↑0.27%	0.2326 ↓3.92%	0.2874 ↓2.18%	0.2327 ↓3.88%	0.2843 ↓3.23%
PEMS08	Avg.	0.2757	0.2826	**0.2494** ↓9.54%	0.2620 ↓7.29%	0.2605 ↓5.51%	0.2750 ↓2.69%	0.2617 ↓5.08%	0.2725 ↓3.57%	0.2540 ↓7.87%	**0.2615** ↓7.47%	0.2596 ↓5.84%	0.2725 ↓3.57%
Sum.	Mean relative gain	—	—	**+17.37%**	**+12.32%**	+10.34%	+8.59%	-0.80%	+0.93%	+3.17%	+2.88%	-22.88%	-7.11%
	Improved cases	—	—	**20/20**	**20/20**	9/20	12/20	9/20	9/20	13/20	14/20	12/20	10/20
	Best cases	0/20	0/20	**17/20**	**17/20**	2/20	2/20	0/20	0/20	1/20	1/20	0/20	0/20

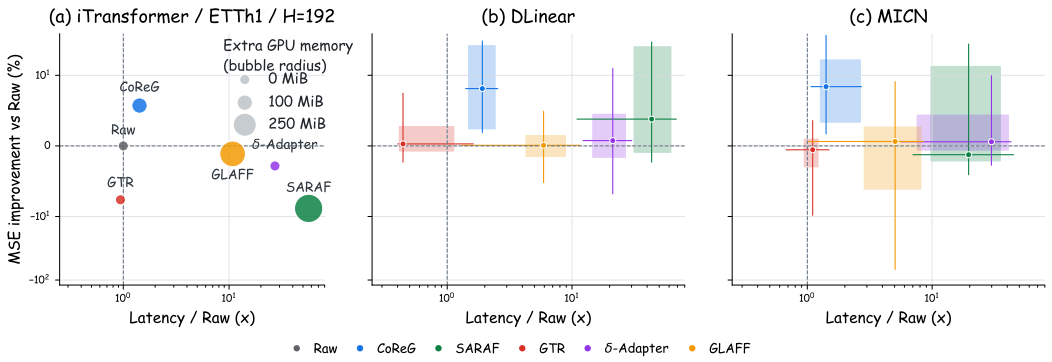


Figure 4: Accuracy–resource trade-offs of plug-in modules. (a) MSE improvement, latency, and additional memory for iTransformer/ETTh1/Horizon 192. (b–c) MSE improvement and latency across 20 matched DLinear and MICN settings. Dots, boxes, and whiskers denote the median, 25th–75th, and 10th–90th percentiles.

5 Conclusion

This paper presents CoReG, which enhances multivariate time-series forecasters through complementary retrieval geometries. CoReG preserves the semantics of temporal-shape and system-role similarities in independent paths, allowing distinct historical evidence to contribute separately to prediction. Experiments across datasets and forecasting backbones show that both geometries carry independent predictive value, their relative advantages vary with forecasting conditions, and their combination yields the most consistent improvements. These findings frame retrieval-augmented forecasting as a problem of organizing predictive analogies: the value of historical evidence depends on the predictive semantics expressed by similarity. This perspective moves retrieval-augmented forecasting from constructing a single retrieval space toward composing multiple semantically distinct and complementary views of historical evidence.

References

Fanpu Cao, Lu Dai, Jindong Han, and Hui Xiong. Enhancing multivariate time series forecasting with global temporal retrieval. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/7a3284ce53f620d244fca0c79d94c478-Abstract-Conference.html.

Dazhao Du, Tao Han, and Song Guo. Predicting the future by retrieving the past. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 20896–20904, 2026. doi: 10.1609/aaai.v40i25.39230. URL https://ojs.aaai.org/index.php/AAAI/article/view/39230.

Jingru Fei, Kun Yi, Wei Fan, Qi Zhang, and Zhendong Niu. Amplifier: Bringing attention to neglected low-energy components in time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 11645–11653, 2025. doi: 10.1609/aaai.v39i11.33267.

Sungwon Han, Seungeon Lee, Meeyoung Cha, Sercan O. Arik, and Jinsung Yoon. Retrieval augmented time series forecasting. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 21774–21797. PMLR, 2025. URL https://proceedings.mlr.press/v267/han25d.html.

Jinkwan Jang, Hyungjin Park, Jinmyeong Choi, and Taesup Kim. Towards robust real-world multivariate time series forecasting: A unified framework for dependency, asynchrony, and missingness. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/f924dab7a06d0463fe238904678acaee-Abstract-Conference.html.

Baoyu Jing, Si Zhang, Yada Zhu, Bin Peng, Kaiyu Guan, Andrew Margenot, and Hanghang Tong. Retrieval based time series forecasting. *arXiv preprint arXiv:2209.13525*, 2022. URL https://arxiv.org/abs/2209.13525.

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.

Jingwei Liu, Ling Yang, Hongyan Li, and Shenda Hong. Retrieval-augmented diffusion models for time series forecasting. In *Advances in Neural Information Processing Systems*, volume 37, 2024a. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/053ee34c0971568bfa5c773015c10502-Paper-Conference.pdf.

Kun Liu, Renjun Jia, Ruifeng Yang, Xirui Zeng, Yuqi Liang, and Cen Chen. Are global dependencies necessary? scalable time series forecasting via local cross-variate modeling. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/a832b65da26c4867a8581cb13d46ad6d-Abstract-Conference.html.

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 *The Twelfth International Conference on Learning Representations*, 2024b. URL https://openreview.net/forum?id=JePfAI8fah.

Jiaming Ma, Qihe Huang, Haofeng Ma, Guanjun Wang, Sheng Huang, Zhengyang Zhou, Pengkun Wang, Xu Wang, Binwu Wang, and Yang Wang. PHAT: Modeling period heterogeneity for multivariate time series forecasting. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/dbb5180957513805ebeea787b8c66ac9-Abstract-Conference.html.

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

Xiangfei Qiu, Jilin Hu, Lekui Zhou, Xingjian Wu, Junyang Du, Buang Zhang, Chenjuan Guo, Aoying Zhou, Christian S. Jensen, Zhenli Sheng, and Bin Yang. TFB: Towards comprehensive and fair benchmarking of time series forecasting methods. *Proceedings of the VLDB Endowment*, 17 (9):2363–2377, 2024. doi: 10.14778/3665844.3665863.

Chao Song, Youfang Lin, Shengnan Guo, and Huaiyu Wan. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 914–921, 2020. doi: 10.1609/aaai.v34i01.5438.

Artyom Stitsyuk and Jaesik Choi. xPatch: Dual-stream time series forecasting with exponential seasonal-trend decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 20601–20609, 2025. doi: 10.1609/aaai.v39i19.34270.

Peiwang Tang and Weitai Zhang. Unlocking the power of patch: Patch-based MLP for long-term time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 12640–12648, 2025. doi: 10.1609/aaai.v39i12.33378.

Changyuan Tian, Zhicong Lu, Zequn Zhang, Heming Yang, Wei Cao, Zhi Guo, Xian Sun, and Li Jin. Hypermixer: Specializable hypergraph channel mixing for long-term multivariate time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 20885–20893, 2025. doi: 10.1609/aaai.v39i19.34302.

Chengsen Wang, Qi Qi, Jingyu Wang, Haifeng Sun, Zirui Zhuang, Jinming Wu, and Jianxin Liao. Rethinking the power of timestamps for robust time series forecasting: A global-local fusion perspective. In *Advances in Neural Information Processing Systems*, 2024. URL https://arxiv.org/abs/2409.18696.

Haoxin Wang, Yipeng Mo, Kunlan Xiang, Nan Yin, Honghe Dai, Bixiong Li, Songhai Fan, and Site Mo. CSformer: Combining channel independence and mixing for robust multivariate time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 21090–21098, 2025. doi: 10.1609/aaai.v39i20.35406.

Huiqiang Wang, Jian Peng, Feihu Huang, Jince Wang, Junhui Chen, and Yifei Xiao. MICN: Multi-scale local and global context modeling for long-term series forecasting. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=zt53IDUR1U.

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, 2021.

Jiyuan Xu, Wenyu Zhang, Xin Jing, Jiahao Nie, Shuai Chen, and Shuai Zhang. CPiRi: Channel permutation-invariant relational interaction for multivariate time series forecasting. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/ebba182cb97864368fdb6ae00773a5e4-Abstract-Conference.html.

Silin Yang, Dong Wang, Haoqi Zheng, and Ruochun Jin. TimeRAG: Boosting LLM time series forecasting via retrieval-augmented generation. In *ICASSP 2025 – 2025 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1–5. IEEE, 2025. doi: 10.1109/ICASSP49660.2025.10889933. URL https://ieeexplore.ieee.org/document/10889933.

Sitan Yang, Carson Eisenach, and Dhruv Madeka. MQ-ReTCNN: Multi-horizon time series forecasting with retrieval-augmentation. In *Proceedings of the 8th KDD Workshop on Mining and Learning from Time Series (MileTS'22)*, 2022.

Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 11121–11128, 2023. doi: 10.1609/aaai.v37i9.26317.

Huanyu Zhang, Chang Xu, Yi-Fan Zhang, Zhang Zhang, Liang Wang, Jiang Bian, and Tieniu Tan. TimeRAF: Retrieval-augmented foundation model for zero-shot time series forecasting. *IEEE Transactions on Knowledge and Data Engineering*, 37(9):5654–5665, 2025. doi: 10.1109/TKDE.2025.3579137. URL https://doi.org/10.1109/TKDE.2025.3579137.

Yunhao Zhang and Junchi Yan. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=vSVLM2j9eie.

Lifan Zhao and Yanyan Shen. Rethinking channel dependence for multivariate time series forecasting: Learning from leading indicators. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=QcSxEwOQwu.

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 *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 11106–11115, 2021. doi: 10.1609/aaai.v35i12.17325.

Shiqiao Zhou, Holger Schöner, Zipeng Wu, Edouard Fouché, IAG Wilson, and Shuo Wang. Stationarity-aware retrieval-augmented time series forecasting, 2026. URL https://arxiv.org/abs/2606.04135.

APPENDIX CONTENTS

A Extended Related Works **14**

B Complete Methodology and Implementation Details **16**

C Experimental Settings **21**

D Supplementary Experiments **23**

D.1 RQ1: How Do TimeGeo and RoleGeo Provide Complementary Retrieval Evidence? 23

D.2 RQ2: Does the Fusion Readout Adapt Its Retrieval Evidence to Dataset Structure and Forecast Horizon? 27

D.3 RQ3: Do TimeGeo and RoleGeo Provide Differentiated Window-Level Gains Across Datasets and Forecast States? 29

D.4 RQ4: Does Cross-Variate Relation Information Provide Stable Predictive Gains for RoleGeo? 34

D.5 RQ5: Is a Simple Static Relation Residual Sufficient for RoleGeo's Predictive Gains? 35

D.6 RQ6: How Sensitive Is CoReG to Its Key Retrieval and Geometry Hyperparameters? 36

D.7 RQ7: Does CoReG Remain Stable Across Random Seeds? 39

D.8 RQ8: How Does CoReG Respond to Reduced Training Data Across Dataset Structures and Backbones? 41

D.9 RQ9: Does Final-View Merging Require Complex Learnable Weights? 44

D.10 RQ10: Which Online Paths Account for CoReG's Computation? 45

E Additional Experimental Results **47**

E.1 Complete Paired Results 47

E.2 Complete Ablation Results 50

E.3 Complete Plug-in Comparisons 50

E.4 Complete CoMerge Tables 55

E.5 Complete Multi-Seed Tables 57

E.6 Complete Data-Sparsity Tables 58

F LLM Usage **62**

A Extended Related Works

A.1 Multivariate Time Series Forecasting

Multivariate time series forecasting builds parameterized predictors around temporal and variable dependencies. DLinear shows that simple linear mappings can form strong forecasting baselines (Zeng et al., 2023); PatchTST extends the effective history through patching and channel-independent modeling (Nie et al., 2023); iTransformer treats variables as tokens and learns their dependencies with attention (Liu et al., 2024b); and Crossformer jointly captures temporal and cross-variate dependencies (Zhang & Yan, 2023). Recent work continues these directions: PatchMLP combines cross-variate mixing with channel-independent processing (Tang & Zhang, 2025), xPatch uses a dual-stream architecture to fuse seasonal–trend information (Stitsyuk & Choi, 2025), and CSformer coordinates channel-independent representations with channel mixing (Wang et al., 2025). PHAT further models heterogeneous periodicity across variables (Ma et al., 2026), while ChannelTokenFormer jointly addresses channel dependencies, asynchronous sampling, and missing observations (Jang et al., 2026). Together, these methods establish a lineage from temporal representation and variable interaction to adaptation under complex data conditions.

A.2 Retrieval-Augmented and Similarity-Based Forecasting

Retrieval-augmented forecasting explicitly supplies a predictor with reference evidence from a dataset or case repository. MQ-ReTCNN selects relevant entities from an offline entity-representation database to approximate large-scale cross-entity attention (Yang et al., 2022); ReTime retrieves reference series through relational structure before content synthesis (Jing et al., 2022); and RATD finds reference series by distances in a pre-trained representation space and uses them to guide diffusion generation (Liu et al., 2024a). Retrieval has also been coupled with specific forecasters: TimeRAG retrieves DTW-similar reference segments from a knowledge base and supplies them as text prompts to an LLM (Yang et al., 2025), RAFT directly retrieves pattern-similar historical candidates and their subsequent values from the training set (Han et al., 2025), and TimeRAF augments a frozen time-series foundation model through a learnable retriever and channel prompting (Zhang et al., 2025). More recent methods enlarge the retrieval scope: PFRP builds a fixed-size global memory bank with predictive contrastive learning and fuses its retrieved global forecast with a local forecaster (Du et al., 2026), GTR maintains global temporal representations of complete cycles and retrieves relevant global segments to extend a host forecaster's effective context (Cao et al., 2026), and SARAF forms a candidate pool through time-aligned temporal similarity before using dataset-level stationarity to regulate candidate diversity and future-sequence aggregation (Zhou et al., 2026). This lineage spans cross-entity, relational, representational, historical-pattern, global-periodic, and stationarity-aware retrieval, with individual methods generally selecting evidence through one principal criterion.

A.3 Cross-Variate Dependency Modeling

Cross-variate dependency modeling incorporates structural information from other variables into a target forecast. Crossformer uses cross-dimensional attention to capture joint temporal and variable dependencies (Zhang & Yan, 2023), iTransformer models multivariate correlations through variable tokens (Liu et al., 2024b), and LIFT explicitly estimates locally stationary lead–lag relations and supplies leading-variable information to the forecaster (Zhao & Shen, 2024). Recent studies consider richer relation forms: HyperMixer uses specializable hypergraphs for higher-order, time-varying channel relations (Tian et al., 2025); CSformer coordinates channel independence and channel mixing (Wang et al., 2025); VPNet enables scalable local cross-variate interactions over a variable–patch field (Liu et al., 2026); and CPiRi learns relational structures invariant to channel permutations (Xu et al., 2026). These methods mainly use variable relations for joint representation and prediction of the current input.

A.4 Detailed Comparison with Retrieval-Augmented Forecasting Methods

Table A.1 summarizes representative retrieval designs spanning cross-entity, relational, representational, historical-pattern, global-periodic, and stationarity-aware evidence. TimeRAF and TimeRAG

couple retrieval with a time-series foundation model and an LLM, respectively. PFRP uses a fixed-size global memory bank to provide retrieval forecasts for arbitrary local univariate forecasters, GTR establishes retrieval as an add-on for multivariate forecasting backbones, and SARAF improves retrieval robustness under non-stationarity through time alignment, candidate diversity, and stationarity-aware aggregation. CoReG combines temporal-shape and system-role retrieval geometries, then reads future values from a fixed-size, target-variable-specific memory to form a retrieval-augmented path for multivariate forecasting.

Table A.1: Detailed comparison between CoReG and representative retrieval-augmented forecasting methods. Methods are ordered by publication year. The symbol △ denotes that the method is closely related to the property but does not fully report it under the channel-wise multivariate setting or a general module interface used here. Cross-variate retrieval evidence indicates that retrieval explicitly uses structural or role information across variables. A bounded memory bank has a controlled candidate-pool size after construction.

Method	Retrieval criterion	Multivariate Forecasting	Plug-and-Play	Cross-variate evidence	Future values as values	Bounded memory bank
MQ-ReTCNN (2022)	Cross-entity encoded context	✓	△	×	×	×
ReTime (2022)	Relational structure	△	×	×	×	×
RATD (2024)	Pre-trained representation distance	✓	×	×	✓	×
TimeRAG (2025)	DTW distance	×	×	×	×	×
RAFT (2025)	Historical-pattern similarity	✓	×	×	✓	×
TimeRAF (2025)	Learnable feature similarity	✓	×	×	×	×
PFRP (2026)	Representation-feature similarity	×	✓	×	✓	✓
GTR (2026)	Global temporal representation	✓	✓	×	×	×
SARAF (2026)	Time-aligned similarity + stationarity-controlled diversity	✓	✓	×	✓	×
CoReG (Ours)	Temporal shape + system role	✓	✓	✓	✓	✓

B Complete Methodology and Implementation Details

This appendix gives the complete computational definition of CoReG. CoReG combines a frozen paired Anchor Forecaster with two independent retrieval paths, TimeGeo and RoleGeo: the offline stage constructs their frozen geometry assets, and the online stage reads these assets to form and merge two complete forecast views.

B.1 Offline Construction of Frozen Geometry Assets

CoReG constructs a paired Anchor, a TimeGeo memory, and a RoleGeo memory from the training series. As shown in Figure 5, the Anchor Forecaster is trained with the same data split, preprocessing, input length, horizon, and random seed as CoReG. Fusion training and inference use the same frozen checkpoint. The two retrieval paths construct their encoders and memories independently on the training split, and fusion training does not update these geometry assets.

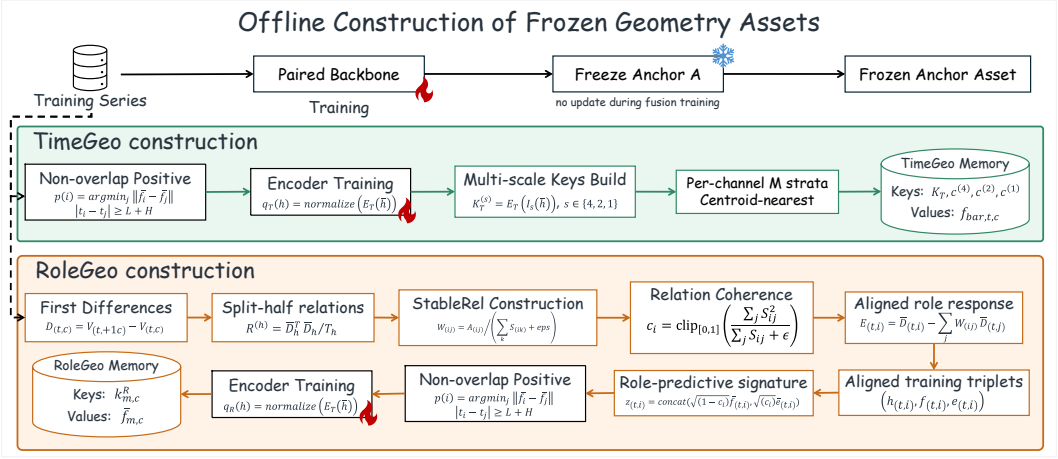


Figure 5: Offline construction of CoReG's frozen geometry assets. The training series first produces a paired Anchor checkpoint. TimeGeo constructs its memory through non-overlapping predictive contrastive learning, multi-scale history keys, and per-variable time-stratified representatives. RoleGeo constructs StableRel, relation coherence, and aligned role responses from training-period first differences, then trains a RoleEncoder with role-predictive signatures before constructing its memory. Both memories store target-owned, training-period normalized future values.

For a training-series start index t and variable c, define the history–future pair

$$h_{t,c} = V_{t:t+L-1,c}, \qquad f_{t,c} = V_{t+L:t+L+H-1,c}. \tag{14}$$

Every one-dimensional window u is standardized within the window:

$$\mu(u) = \frac{1}{|u|}\sum_i u_i, \qquad \sigma(u) = \mathrm{Std}(u) \vee 10^{-6}, \qquad \bar{u} = \frac{u - \mu(u)}{\sigma(u)}. \tag{15}$$

Each geometry has an independent two-layer predictive encoder,

$$e(u) = W_2\mathrm{ReLU}(W_1\bar{u} + b_1) + b_2 \in \mathbb{R}^d, \qquad q(u) = \frac{e(u)}{\|e(u)\|_2}. \tag{16}$$

Contrastive candidates must not overlap temporally:

$$\mathcal{J}(i) = \{j : |t_i - t_j| \geq L + H\}, \qquad p(i) = \arg\min_{j\in\mathcal{J}(i)} \|z_i - z_j\|_2, \tag{17}$$

$$\mathcal{L}_{\mathrm{PCL}} = -\frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} \log \frac{\exp(q_i^\top q_{p(i)}/\tau)}{\sum_{j\in\mathcal{J}(i)} \exp(q_i^\top q_j/\tau)}. \tag{18}$$

TimeGeo uses the target variable's normalized future as its supervision signature, whereas RoleGeo uses the role-predictive signature defined in Section B.5. Both encoders are frozen after training.

B.2 Online Forecasting and Two-view Merging

Given a multivariate history $X \in \mathbb{R}^{N\times L\times C}$ and future label $Y \in \mathbb{R}^{N\times H\times C}$, online forecasting uses only the current history, the frozen Anchor, and the two training-period memories. Figure 6 shows that the Anchor produces a shared forecast B_{anc}, while TimeGeo and RoleGeo independently retrieve target-local future memories to form retrieval forecasts R_T, R_R and adjusted Top-K scores S_T, S_R. Two independent GeoRead modules then produce complete views Y_T, Y_R, which Fixed CoMerge averages.

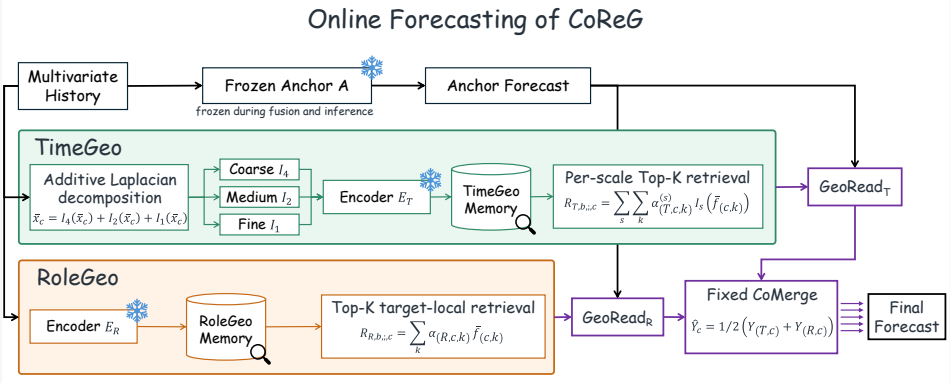


Figure 6: Online forecasting with CoReG. The multivariate history enters the frozen Anchor, TimeGeo, and RoleGeo. TimeGeo retrieves separately from three additive Laplacian components, while RoleGeo retrieves with the role-predictive geometry formed offline; both read target-owned training-period future values. GeoRead$_T$ and GeoRead$_R$ separately fuse their retrieval forecast with the Anchor forecast. Fixed CoMerge takes the arithmetic mean of the two complete views.

$$B_{\text{anc}} = A(X) \in \mathbb{R}^{N\times H\times C}, \qquad R_T, R_R \in \mathbb{R}^{N\times H\times C}, \qquad S_T, S_R \in \mathbb{R}^{N\times C\times K}, \tag{19}$$

$$Y_T = G_T(R_T, B_{\text{anc}}, S_T), \qquad Y_R = G_R(R_R, B_{\text{anc}}, S_R), \qquad \widehat{Y} = \frac{1}{2}(Y_T + Y_R). \tag{20}$$

At inference, the RoleGeo query encodes only the target variable's history. StableRel and role responses shape its key geometry during the offline stage; they are not recomputed from the current query. Both paths retrieve target-owned values, and no test label, test future, memory update, or parameter update is used.

B.3 Shared Predictive Memory and Retrieval Protocol

For each variable, each path stores a normalized key, target-owned normalized future value, and window start:

$$\mathcal{M}_c = \{(k_{m,c}, \tilde{f}_{m,c}, t_m)\}_{m=1}^{M}. \tag{21}$$

All training windows of each target variable are partitioned into M contiguous time strata. The representative nearest the predictive-embedding centroid is retained within each stratum, preserving temporal coverage and predictive representation coverage. Training queries mask memory windows that overlap their history or future. The two memories, encoders, and retrieval computations are independent, while both preserve target-local future-value ownership.

B.4 TimeGeo: Additive Multi-scale Temporal Retrieval

Let $P_s(u)$ denote non-overlapping block averaging with block length s, repeated to the original length. With $\mathcal{S} = (4, 2, 1)$,

$$\ell_4(u) = P_4(u), \qquad \ell_2(u) = P_2(u) - P_4(u), \qquad \ell_1(u) = u - P_2(u), \tag{22}$$

$$u = \ell_4(u) + \ell_2(u) + \ell_1(u). \tag{23}$$

For query $x_{b,c}$ and scale s, TimeGeo encodes normalized $\ell_s(x_{b,c})$ and retrieves Top-K entries from variable c's TimeGeo memory by dot-product similarity. It applies the same component $\ell_s(\tilde{f}_m)$ to each candidate future and computes a confidence score from the current history and candidate future:

$$g^{(s)}_{b,c,m} = \sigma\left(w_2^\top \mathrm{GELU}(W_x \bar{x}^{(s)}_{b,c} + W_f \ell_s(\tilde{f}_m) + b)\right). \tag{24}$$

$$r^{(s)}_{b,c,m} = a^{(s)}_{b,c,m} g^{(s)}_{b,c,m}, \qquad \alpha^{(s)}_{b,c,:} = \mathrm{Softmax}(r^{(s)}_{b,c,:}), \tag{25}$$

$$\tilde{r}^{(s)}_{b,c} = \sum_{m=1}^{K} \alpha^{(s)}_{b,c,m} \ell_s(\tilde{f}_m), \qquad \tilde{r}^{T}_{b,c} = \sum_{s\in\mathcal{S}} \tilde{r}^{(s)}_{b,c}. \tag{26}$$

An output affine gate and the query statistics restore the retrieved future:

$$R_{T,b,:,c} = \sigma(x_{b,c}) \left[a_o(\bar{x}_{b,c}) \odot \tilde{r}^{T}_{b,c} + b_o(\bar{x}_{b,c})\right] + \mu(x_{b,c}). \tag{27}$$

The three adjusted Top-K score vectors are averaged to form $S_T[b, c, :]$. TimeGeo uses a common predictive encoder for the three Laplacian views and produces its temporal geometry through independent retrieval followed by additive reconstruction.

B.5 RoleGeo: Stable-relation-guided Role-predictive Geometry

RoleGeo extracts stable cross-variate dynamics during training to shape target-variable predictive similarity. For $V \in \mathbb{R}^{T\times C}$, let

$$D_{t,c} = V_{t+1,c} - V_{t,c}. \tag{28}$$

After splitting D into two contiguous halves and centering and RMS-standardizing each variable within each half,

$$R^{(h)} = \frac{(\bar{D}^{(h)})^\top \bar{D}^{(h)}}{T_h}, \qquad h \in \{1, 2\}. \tag{29}$$

StableRel retains relations supported by both halves:

$$S_{ij} = \sqrt{\max(R^{(1)}_{ij} R^{(2)}_{ij}, 0)}, \qquad Q_{ij} = \mathrm{sign}(R^{(1)}_{ij} + R^{(2)}_{ij}) S_{ij}, \tag{30}$$

$$W_{ii} = 0, \qquad W_{ij} = \frac{Q_{ij}}{\sum_k S_{ik} + \epsilon}, \qquad c_i = \mathrm{clip}_{[0,1]}\left(\frac{\sum_j S^2_{ij}}{\sum_j S_{ij} + \epsilon}\right). \tag{31}$$

For the full training period, variable-wise normalized first differences yield the aligned role response

$$E_{t,i} = \bar{D}_{t,i} - \sum_{j=1}^{C} W_{ij} \bar{D}_{t,j}. \tag{32}$$

An all-zero row is prepended to E for alignment with the original time indices. The target future $f_{t,i}$ and role-response future $e_{t,i}$ yield the role-predictive signature

$$z_{t,i} = \mathrm{concat}\left(\sqrt{1 - c_i}\, \overline{f}_{t,i}, \sqrt{c_i}\, \overline{e}_{t,i}\right). \tag{33}$$

The RoleEncoder receives target history $h_{t,i}$ and uses $z_{t,i}$ for predictive contrastive supervision. Its key space is organized by future behavior and system role, while its memory stores target-owned normalized future values.

B.6 GeoRead, Fixed CoMerge, and Fusion Training

The paths use independent GeoRead modules. For $v \in \{T, R\}$, GeoRead maps its adjusted Top-K scores to two coefficients for the retrieval forecast and the Anchor forecast. These coefficients are normalized only within the corresponding GeoRead; no cross-path normalization is performed between TimeGeo and RoleGeo:

$$\beta_v[b, c, :] = \mathrm{Softmax}\left(W^{(v)}_2 \mathrm{GELU}(W^{(v)}_1 S_v[b, c, :] + b^{(v)}_1) + b^{(v)}_2\right) \in \mathbb{R}^2, \tag{34}$$

$$Y_v[b, h, c] = \beta_v[b, c, 0] R_v[b, h, c] + \beta_v[b, c, 1] B_{\mathrm{anc}}[b, h, c]. \tag{35}$$

One coefficient pair is shared across all forecast steps for the same (b, c). Fixed CoMerge averages the two complete views. The retrieval coefficients from the two GeoRead modules act on R_T and R_R, respectively, and jointly enter the final prediction after the $1/2$ FixedMean scaling; each coefficient pair describes the retrieval readout within its path:

$$\mathcal{L}_{\mathrm{fuse}} = \ell(\widehat{Y}, Y). \tag{36}$$

Algorithm 1 Training and memory construction for CoReG

Require: Training series V, Backbone A, forecast horizon H
Ensure: Frozen Anchor, TimeGeo memory, RoleGeo memory, and GeoRead parameters
1: Train A on the training split and freeze A
2: Train the TimeGeo encoder; select time-stratified representatives and store multi-scale keys with target-owned futures
3: Estimate StableRel W and coherence c; construct the aligned role response E
4: Train the RoleEncoder; select representatives and store role keys with target-owned futures
5: Freeze both predictive encoders and all memory tensors
6: **for** each fusion batch (X, Y, index) **do**
7: $\quad B_{\text{anc}} \leftarrow A(X)$; $(R_T, S_T) \leftarrow \text{TimeGeo.retrieve}(X, \text{index})$
8: $\quad (R_R, S_R) \leftarrow \text{RoleGeo.retrieve}(X, \text{index})$
9: $\quad Y_T \leftarrow G_T(R_T, B_{\text{anc}}, S_T)$; $Y_R \leftarrow G_R(R_R, B_{\text{anc}}, S_R)$
10: $\quad \widehat{Y} \leftarrow 0.5(Y_T + Y_R)$; optimize $\mathcal{L}_{\text{fuse}}$
11: **end for**

Algorithm 2 Inference with CoReG

Require: Current multivariate history X
Ensure: Forecast $\widehat{Y}$
1: $B_{\text{anc}} \leftarrow A(X)$; $(R_T, S_T) \leftarrow \text{TimeGeo.retrieve}(X, \text{test mode})$
2: $(R_R, S_R) \leftarrow \text{RoleGeo.retrieve}(X, \text{test mode})$
3: $Y_T \leftarrow G_T(R_T, B_{\text{anc}}, S_T)$; $Y_R \leftarrow G_R(R_R, B_{\text{anc}}, S_R)$
4: **return** $\widehat{Y} \leftarrow 0.5(Y_T + Y_R)$

B.7 Staged Training and Inference

Table B.2 summarizes the parameter states of offline construction and fusion training.

Table B.2: CoReG stages, frozen parameters, and resulting assets.

Stage	Data	Trainable parameters	Frozen parameters	Artifact
Anchor	Train/validation split	Backbone	Retrieval paths and GeoRead	Anchor checkpoint
TimeGeo asset	Training split	TimeGeo encoder	Anchor	TimeGeo memory
RoleGeo asset	Training split	RoleEncoder	Anchor	RoleGeo memory
Fusion training	Train/validation split	Non-encoder path readouts and two GeoRead modules	Anchor, encoders, and memories	CoReG checkpoint

Before fusion training, both predictive encoders and all memory tensors are frozen. The following algorithms state the training and inference procedures.

B.8 Information Boundary, Cache Identity, and Complexity

Table B.3 states the information boundary of the two paths. Test labels and futures are unavailable, and inference does not update memory.

Cache identity binds the preprocessed training-data hash, L, H, variable count, embedding dimension, memory capacity, retrieval range, representative-selection strategy, random seed, and RoleGeo relation configuration. A cache or checkpoint identity mismatch causes an explicit failure.

Table B.3: Information available to TimeGeo and RoleGeo during construction and inference.

Information	TimeGeo	RoleGeo	Availability at inference
Current target history	Query history	Query history	Available
Current peer history	Not used	Outside the online query	Unavailable
Training target future	PCL and memory value	PCL and memory value	Memory asset
Training peer history	Not used	StableRel and role response	Offline asset
Test future	Unavailable	Unavailable	Unavailable

Let M be the per-variable memory capacity, d the embedding dimension, and $S = 3$ the number of TimeGeo scales. The main retrieval cost per batch and main memory storage are

$$O\left(BC\left[S(Md + KH) + (Md + KH)\right]\right), \tag{37}$$

$$O\left(CM[(S + 1)d + 2H]\right), \tag{38}$$

respectively. The offline StableRel matrix costs $O(C^2)$, and the implementation streams candidate futures by query to control peak activation memory.

C Experimental Settings

C.1 Datasets, TFB Framework, and Evaluation Protocol

We implement and evaluate CoReG within the Time Series Forecasting Benchmark (TFB) framework (Qiu et al., 2024). TFB provides unified data loading, experimental configuration, and evaluation pipelines. Our evaluation covers ten public multivariate datasets from transformer monitoring, weather, electricity consumption, exchange rates, traffic flow, and photovoltaic generation. Table C.1 summarizes their statistics. Unless otherwise specified, all methods use the same fixed data splits, preprocessing, input length, prediction horizons, and rolling evaluation windows. The input length is 720, the prediction horizons are 96, 192, 336, and 720, and performance is evaluated using mean squared error (MSE) and mean absolute error (MAE).

Table C.1: Statistics of the datasets used for multivariate time-series forecasting evaluation.

Dataset	Domain	Variables	Time points	Sampling frequency
ETTh	Transformer temperature and load	7	17,420	1 h
ETTm	Transformer temperature and load	7	69,680	15 min
Weather	Meteorological observations	21	52,696	10 min
Electricity	Customer electricity load	321	26,304	1 h
Exchange	Exchange rates	8	7,588	1 day
PEMS04	Traffic flow	307	16,992	5 min
PEMS08	Traffic flow	170	17,856	5 min
Solar	Photovoltaic generation	137	52,560	10 min

ETT. The Electricity Transformer Temperature (ETT) dataset (Zhou et al., 2021) contains oil-temperature and load-related measurements from electricity transformers in two regions of China. It comprises ETTm1, ETTm2, ETTh1, and ETTh2: the ETTm subsets are sampled every 15 minutes and the ETTh subsets hourly. Each subset contains oil temperature and six load-related variables, enabling evaluation across sampling frequencies.

Weather. The Weather dataset (Wu et al., 2021) records 21 meteorological variables in Germany throughout 2020 at 10-minute intervals, including temperature, humidity, pressure, wind speed, and visibility. Its continuous high-frequency measurements assess multivariate weather dynamics and local fluctuations.

Electricity. The Electricity dataset (Lai et al., 2018) contains electricity-consumption records for 321 clients from 2012 to 2014. We use the commonly adopted hourly version. Its periodic yet heterogeneous client-level patterns assess cross-variable modeling in high-dimensional load series.

Exchange. The Exchange dataset (Lai et al., 2018) contains daily exchange-rate series for eight countries from 1990 to 2016. Its non-stationary and irregular fluctuations evaluate generalization in a low-dimensional but dynamically varying setting.

PEMS04 and PEMS08. PEMS04 and PEMS08 (Song et al., 2020) are 5-minute traffic-flow datasets from the California traffic-monitoring system, with 307 and 170 road sensors, respectively. Their temporal periodicity, spatial correlation, and local congestion changes assess high-frequency traffic-network forecasting.

Solar. The Solar dataset (Lai et al., 2018) contains 10-minute solar-power generation series from 137 photovoltaic plants in 2006. Daily cycles and weather disturbances make it suitable for evaluating high-frequency, strongly non-stationary renewable-energy series.

C.2 Implementation and Training Configuration

We pair CoReG with DLinear (Zeng et al., 2023), PatchMLP (Tang & Zhang, 2025), iTransformer (Liu et al., 2024b), xPatch (Stitsyuk & Choi, 2025), MICN (Wang et al., 2023), and Amplifier (Fei et al., 2025). All backbones are trained for up to 10 epochs, with patience 3, learning rate

0.0001, and batch size 64; their remaining model parameters follow the official TFB scripts. Each paired experiment uses the same trained checkpoint as its corresponding backbone baseline as the Anchor, and the checkpoint remains frozen during retrieval augmentation.

Table C.2: Implementation and training configuration of CoReG.

Item	Configuration
TimeGeo scales	$(4, 2, 1)$
Retrieval scope	Variable-specific future values
Top-K	50
Memory entries per variable	1,000
Forecast encoding dimension	16
Forecast encoder epochs	80
Fusion-training epochs	10
Early-stopping patience	3
Learning rate	0.0001
Batch size	64

As shown in Table C.2, CoReG uses the same retrieval and fusion-training configuration in all paired experiments. During fusion training, the Anchor, forecast encoder, and memory remain frozen, while only the retrieval readout and within-path fusion parameters are optimized.

C.3 Computing Environment

Table C.3: Computing environment for CoReG experiments.

Item	Configuration
GPU	NVIDIA GeForce RTX 4070 Ti SUPER (16 GB)
CPU	Intel Core i7-14650HX (16 cores / 24 threads)
System memory	32 GB
Operating system	Windows 11
Python	3.10.20
PyTorch	2.4.1+cu121
CUDA runtime	12.1
cuDNN	9.1

All CoReG experiments run on the same single-machine workstation. Table C.3 reports the hardware and software versions required to reproduce the experiments.

C.4 Reproducibility Assets and Result Traceability

Each reported result is associated with its dataset, prediction horizon, paired backbone, random seed 2027, TFB configuration file, CoReG configuration, backbone-checkpoint identifier, and the construction identities of the TimeGeo and RoleGeo memories. Experimental records retain the launch configuration, execution log, window-level forecasts, and aggregation script, so that every number in a main table can be traced to its prediction outputs and experimental configuration.

D SUPPLEMENTARY EXPERIMENTS

D.1 RQ1: HOW DO TIMEGEO AND ROLEGEO PROVIDE COMPLEMENTARY RETRIEVAL EVIDENCE?

This section presents case studies of the complementary roles of TimeGeo and RoleGeo along concrete forecast trajectories. The cases cover ETTh2, ETTm1, ETTm2, Electricity, PEMS04, PEMS08, and Solar, together with DLinear, iTransformer, PatchMLP, Amplifier, and MICN, at horizons 96, 192, 336, and 720. Each composite figure contains two independent cases: the upper row shows the complete forecasting window, and the lower row enlarges the key interval marked by shading above. The curves show the original Backbone (Raw), TimeGeo, RoleGeo, the dual-retrieval fusion (Dual), and the ground truth (True), with the case MSE reported at the bottom.

Figures D.1 and D.2 focus on periodic and long–short-term changes in the ETT datasets. Across ETTh2, ETTm1, and ETTm2, TimeGeo and RoleGeo track different historical patterns in different intervals: one path follows local periodic or trend structure, while the other supplies a distinct trajectory reference around fluctuations and turning points. After combining the two sources, Dual stays closer to the ground-truth trajectory in the enlarged intervals, together with lower case-level MSE.

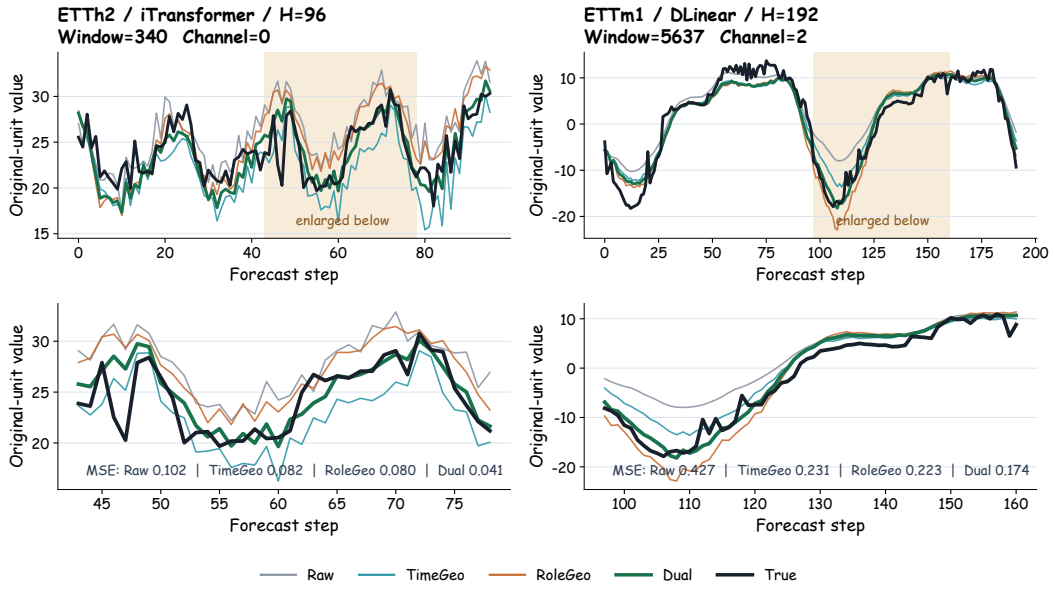


Figure D.1: Dual-retrieval cases on ETTh2/iTransformer/Horizon 96 and ETTm1/DLinear/Horizon 192. For each case, the upper panel shows the complete window and the lower panel enlarges the shaded interval. Curves represent Raw, TimeGeo, RoleGeo, Dual, and True; the case MSE is reported at the bottom.

Figures D.3 and D.4 show Electricity and PEMS04 cases with pronounced periodic states or rapid state transitions. Around periodic plateaus, transition boundaries, and trough recovery, a single retrieval path can exhibit phase displacement, amplitude shrinkage, or local overshoot. The other geometry supplies a different historical reference, allowing Dual to preserve both state boundaries and waveform amplitude. The enlarged intervals make this correction directly visible.

Figures D.5 and D.6 extend the analysis to PEMS08, MICN, and Solar, covering high-dimensional traffic signals and strongly periodic solar signals. In the PEMS08 cases, the system-state reference supplied by RoleGeo complements the local-shape reference supplied by TimeGeo. In the Solar case, the two paths jointly preserve the rising and falling boundaries of periodic segments. Across these Backbones and data regimes, dual retrieval converts complementary historical evidence into a more stable forecast trajectory.

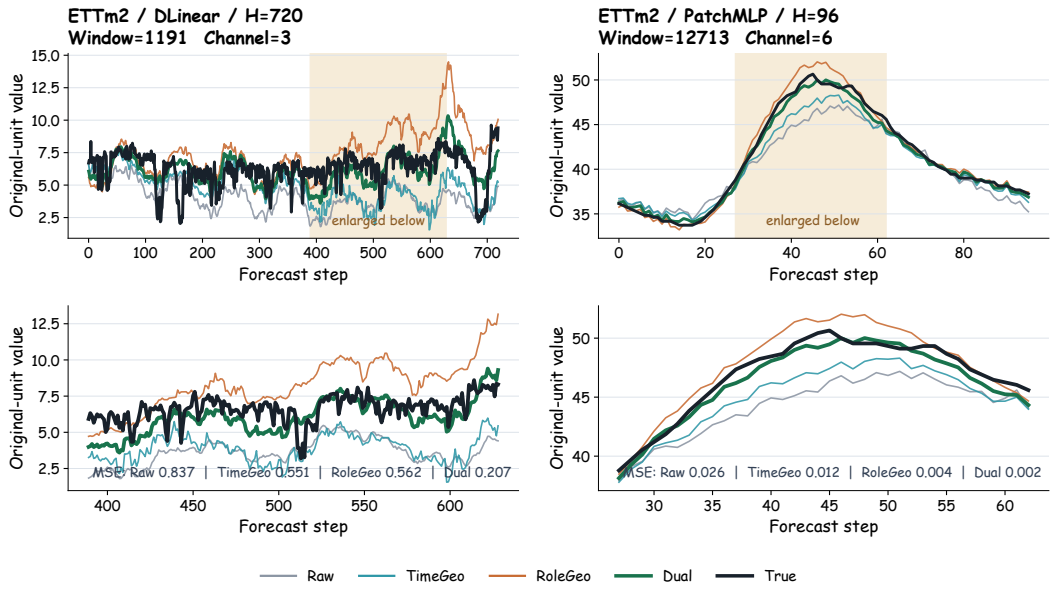


Figure D.2: Dual-retrieval cases on ETTm2/DLinear/Horizon 720 and ETTm2/PatchMLP/Horizon 96. The left case highlights persistent bias correction at a long horizon, while the right case shows complementary retrieval around a short-horizon periodic peak.

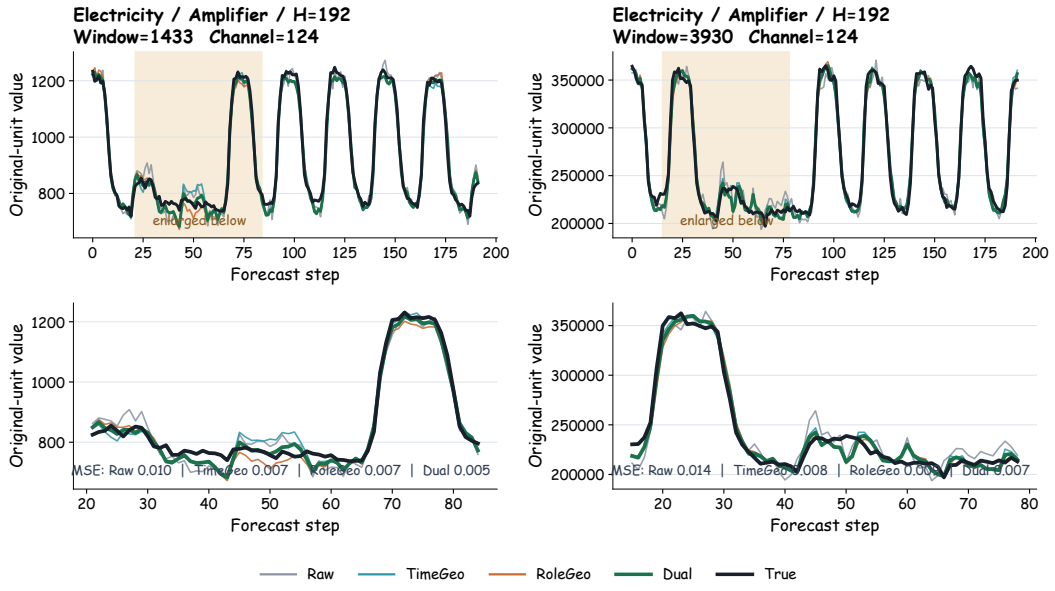


Figure D.3: Two Electricity/Amplifier/Horizon 192 cases with periodic state transitions. The upper panels show complete cycles and the lower panels enlarge low-amplitude fluctuation intervals where Dual combines the two retrieval sources.

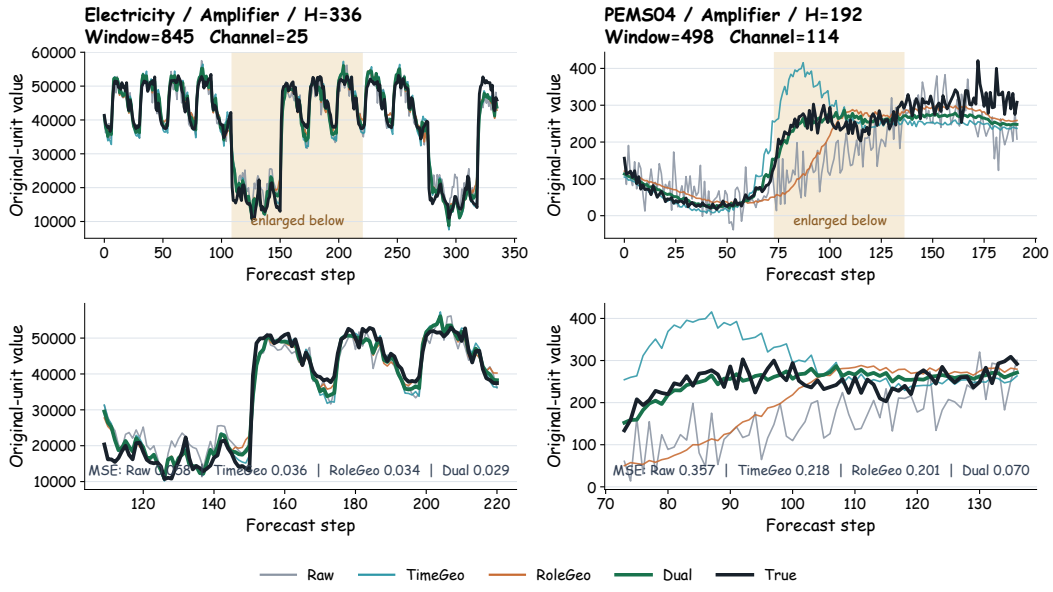


Figure D.4: Electricity/Amplifier/Horizon 336 and PEMS04/Amplifier/Horizon 192 cases. The examples cover high-frequency periodic behavior and nonstationary state transitions; the enlarged intervals show joint phase and amplitude correction by Dual.

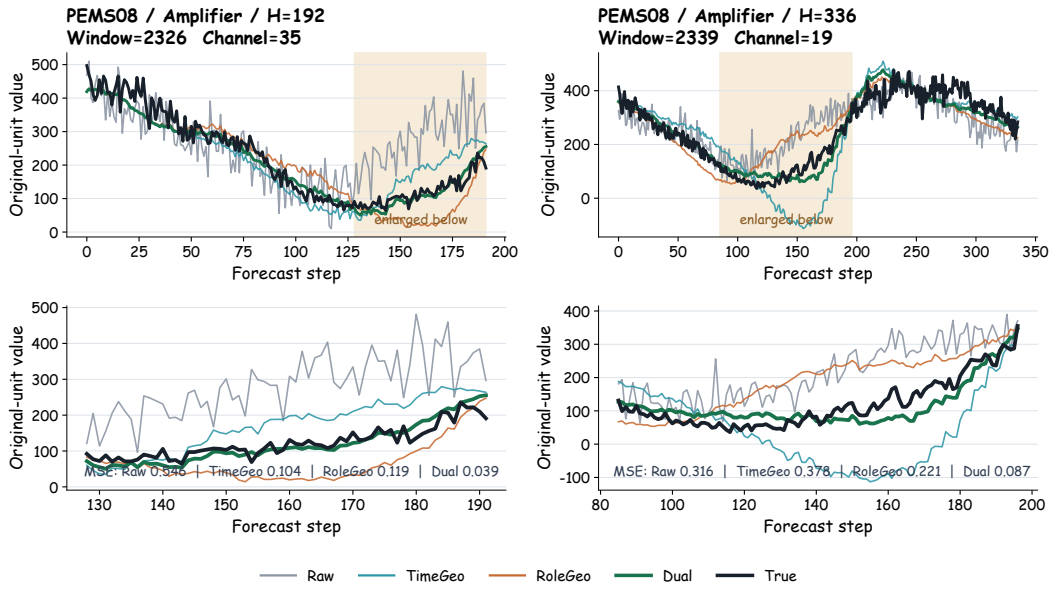


Figure D.5: PEMS08/Amplifier cases at horizons 192 and 336. The examples highlight local decline–recovery and periodic peak intervals; Dual maintains dynamics closer to True in the enlarged regions.

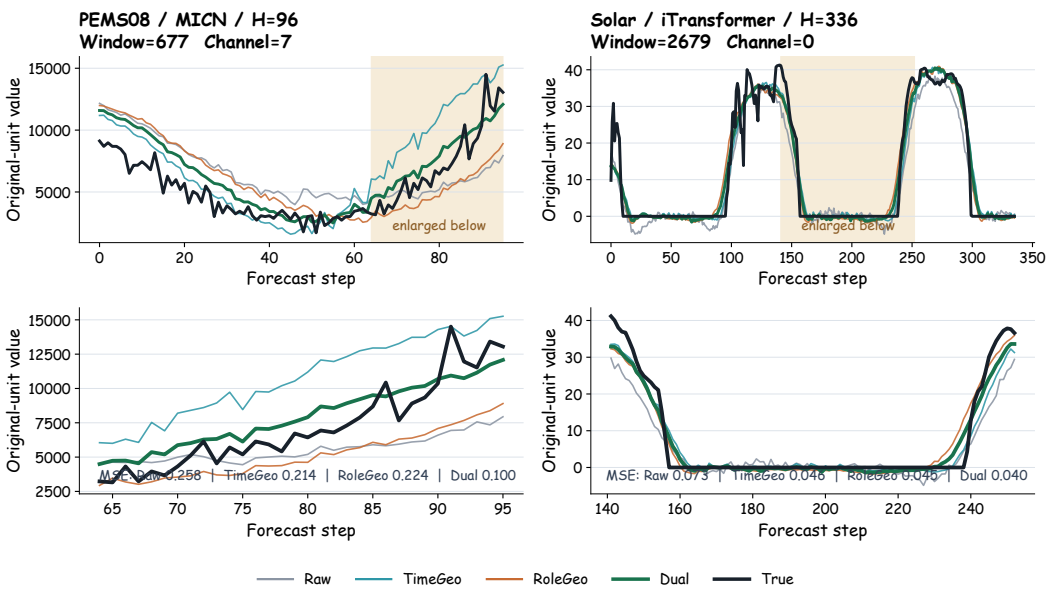


Figure D.6: PEMS08/MICN/Horizon 96 and Solar/iTransformer/Horizon 336 cases. The left case shows a nonstationary high-dimensional traffic sequence, while the right case shows strong solar periodicity; Dual combines the complementary TimeGeo and RoleGeo references in both trajectories.

These cases expose the central mechanism of CoReG at the trajectory level. TimeGeo supplies temporal-shape evidence from the target variable, and RoleGeo supplies historical analogies under stable system roles. Their evidence is complementary around local fluctuations, state transitions, and long-horizon error accumulation. Dual therefore improves trajectory shape, phase alignment, and amplitude tracking across the presented datasets and Backbones, as reflected by the case-level MSE annotations.

D.2 RQ2: Does the Fusion Readout Adapt Its Retrieval Evidence to Dataset Structure and Forecast Horizon?

This section analyzes whether the GeoRead modules adjust their reliance on retrieval forecasts according to dataset structure and forecast horizon. The analysis uses DLinear-paired experiments on ETTh1, ETTh2, ETTm1, ETTm2, Exchange, PEMS08, and Weather at horizons 96, 192, 336, and 720, covering 28 tasks. Each trace contains target-variable-level coefficients from evaluation batches. The TimeGeo and RoleGeo retrieval coefficients denote the local retrieval-prediction coefficient within the corresponding "retrieval–Anchor" pair; their Anchor coefficients are the complements. They are not cross-path probabilities. FixedMean scales both local retrieval contributions by $1/2$ before they enter the final prediction.

For target variable c, the local readouts and their final aggregation are

$$Y_T = \beta_T^R R_T + \beta_T^B B, \qquad Y_R = \beta_R^R R_R + \beta_R^B B, \tag{39}$$

$$\widehat{Y} = \frac{1}{2}\beta_T^R R_T + \frac{1}{2}\beta_R^R R_R + \frac{1}{2}(\beta_T^B + \beta_R^B)B. \tag{40}$$

Thus, the coefficients in the following table and figures describe local GeoRead behavior; their sum after FixedMean scaling gives the aggregate retrieval contribution.

Table D.1: GeoRead retrieval coefficients across datasets and forecast horizons. The parenthetical value after each dataset gives its variable count, and Level correlation $|r|$ is the median absolute Pearson correlation among raw variables. The first header row groups columns by prediction length (PL), while the second row lists the TimeGeo and RoleGeo coefficients. The final two columns average each coefficient over the four PLs. These coefficients are independently normalized within their GeoRead and are not cross-path probabilities.

Dataset	Level corr. $\lvert r\rvert$	PL 96		PL 192		PL 336		PL 720		Average	
		TimeGeo	RoleGeo	TimeGeo	RoleGeo	TimeGeo	RoleGeo	TimeGeo	RoleGeo	TimeGeo	RoleGeo
ETTh1 (7)	0.119	0.337	0.293	0.268	0.217	0.308	0.278	0.218	0.203	0.283	0.248
ETTh2 (7)	0.200	0.476	0.252	0.435	0.261	0.325	0.298	0.244	0.225	0.370	0.259
ETTm1 (7)	0.122	0.454	0.280	0.321	0.250	0.317	0.296	0.287	0.276	0.345	0.276
ETTm2 (7)	0.201	0.358	0.100	0.370	0.119	0.418	0.108	0.474	0.200	0.405	0.132
Exchange (8)	0.554	0.441	0.424	0.392	0.333	0.440	0.465	0.392	0.382	0.416	0.401
PEMS08 (170)	0.811	0.470	0.387	0.527	0.455	0.535	0.552	0.458	0.480	0.497	0.468
Weather (21)	0.128	0.530	0.413	0.407	0.306	0.457	0.366	0.367	0.251	0.440	0.334

Table D.1 shows substantial task-conditioned variation in the local retrieval coefficients. PEMS08 and Exchange generally retain relatively high coefficients for both geometries, whereas ETTm2 has a lower RoleGeo coefficient. Weather assigns a higher TimeGeo coefficient at the shortest horizon and reduces it as the horizon grows. The within-dataset changes across horizons further show that GeoRead adapts its local retrieval reading to the forecast range.

Figure D.7 shows that datasets with stronger overall level correlation often exhibit more pronounced GeoRead retrieval-coefficient patterns, especially PEMS08 and Exchange. Distribution location and spread still vary across horizons, so overall correlation explains only part of the cross-dataset differences; the realized coefficients also depend on the forecast horizon, target variable, and local dynamics.

Figure D.8 shows that larger systems are associated with different local coefficient patterns, while variable count alone does not determine the reading degree of either geometry. PEMS08 retains relatively high coefficients for both retrieval geometries despite its much larger variable count, whereas low-dimensional datasets also differ substantially. System scale, dataset dynamics, and forecast horizon jointly shape the evidence readout.

Together, Table D.1 and Figures D.7–D.8 show conditional evidence reading by the two GeoRead modules. Their local retrieval coefficients vary with dataset structure and forecast horizon, and the two geometries follow different patterns. FixedMean then applies the symmetric $1/2$ scaling to combine their contributions in the final forecast.

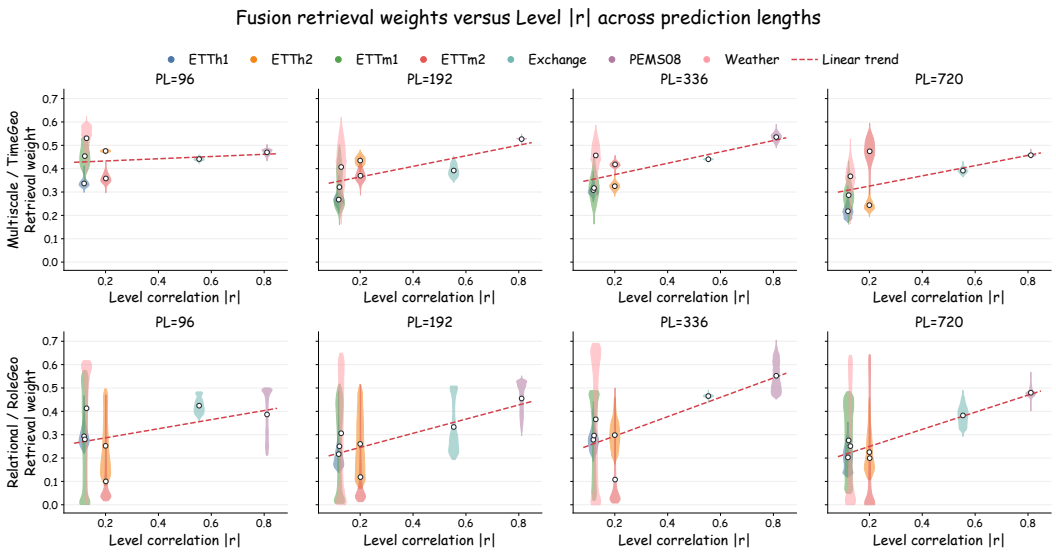


Figure D.7: GeoRead retrieval coefficients versus dataset-level correlation across forecast horizons. The upper row shows TimeGeo and the lower row RoleGeo; columns correspond to horizons 96, 192, 336, and 720. The horizontal axis is the median absolute Pearson correlation among raw variables, and the vertical axis is the corresponding local retrieval coefficient. Violins show evaluation-batch coefficient distributions, white circles show dataset means, and colors identify datasets.

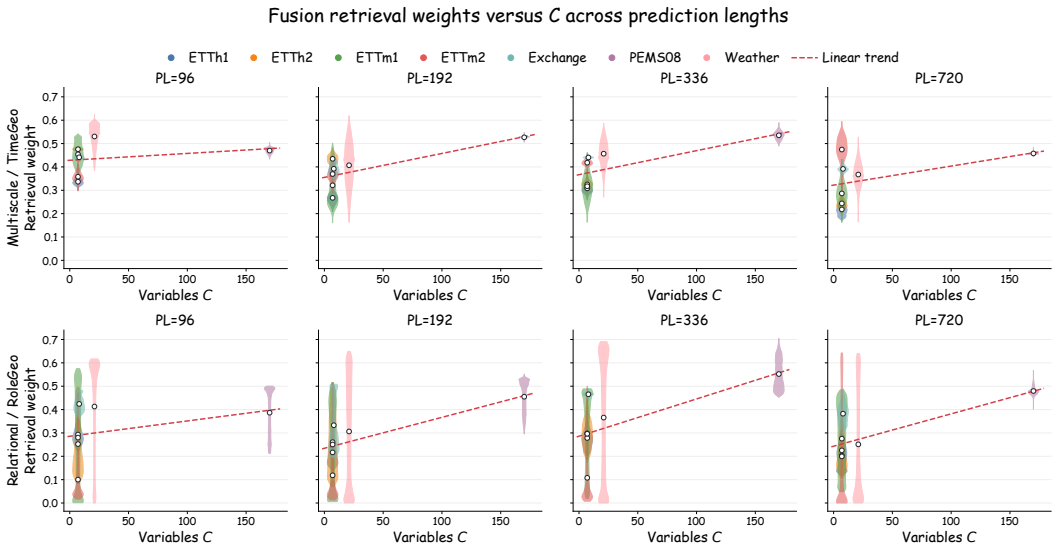


Figure D.8: GeoRead retrieval coefficients versus the number of variables across forecast horizons. The upper row shows TimeGeo and the lower row RoleGeo; columns correspond to horizons 96, 192, 336, and 720. The horizontal axis is the dataset variable count C, and the vertical axis is the corresponding local retrieval coefficient. Violins, mean markers, and dataset colors follow Figure D.7.

D.3 RQ3: Do TimeGeo and RoleGeo Provide Differentiated Window-Level Gains Across Datasets and Forecast States?

This section compares the gain patterns of TimeGeo and RoleGeo at the window and variable-pair levels. We use DLinear on ETTm1, ETTm2, Exchange, PEMS08, and Weather at horizons 96, 192, 336, and 720; all results use seed 2027. For test window w and geometry $G \in \{\text{TimeGeo}, \text{RoleGeo}\}$, we define the branch gain relative to Raw DLinear by

$$g_G(w) = \frac{\text{MSE}_{\text{Raw}}(w) - \text{MSE}_G(w)}{\text{MSE}_{\text{Raw}}(w)}. \tag{41}$$

Here, $\text{MSE}_G(w)$ is the MSE of the branch-fused prediction retaining geometry G on window w. A larger $g_G(w)$ indicates a larger error reduction by that branch. The window winner is the branch with the larger of $g_{\text{TimeGeo}}(w)$ and $g_{\text{RoleGeo}}(w)$.

Figure D.9 gives the ordered window-level difference $g_{\text{RoleGeo}}(w) - g_{\text{TimeGeo}}(w)$ for four datasets. Positive values favor RoleGeo and negative values favor TimeGeo. The traces repeatedly cross zero across horizons and local test-window intervals, showing that branch advantage changes with forecasting state. The smoothed traces are more positive for Exchange and PEMS08, while ETTm1 and ETTm2 remain closer to balanced competition.

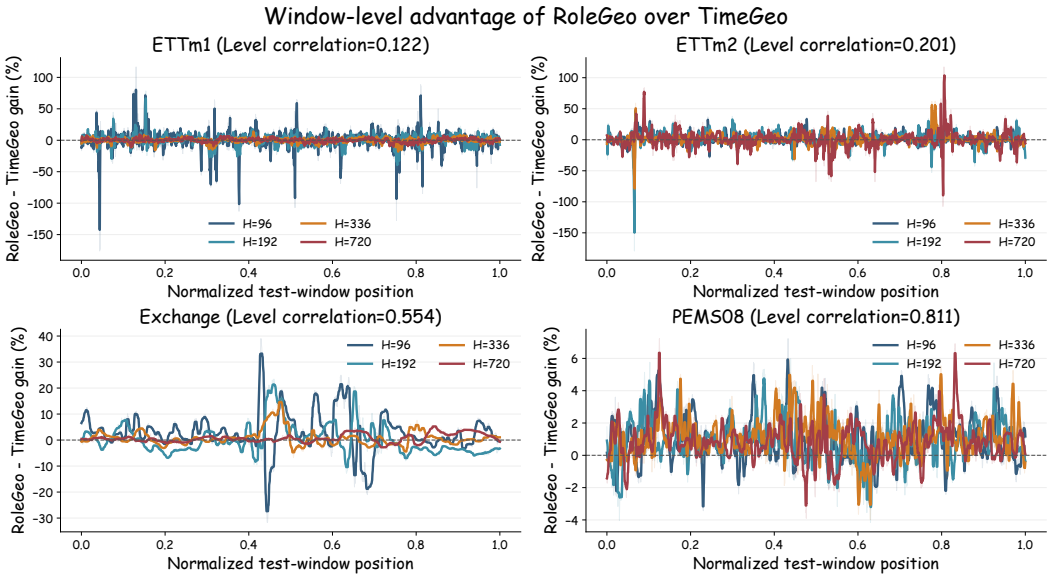


Figure D.9: Ordered window-level gain differences $g_{\text{RoleGeo}} - g_{\text{TimeGeo}}$ across ETTm1, ETTm2, Exchange, and PEMS08. Faint curves show per-window differences and thick curves show rolling medians; colors identify the four forecast horizons.

Figures D.10–D.14 provide dataset-specific curves and embedded winner strips. RoleGeo wins 60.4%, 58.0%, 54.4%, and 54.0% of ETTm1 windows across the four horizons, and 55.1%, 63.1%, 51.6%, and 55.5% on ETTm2. Exchange shows stronger horizon dependence, with RoleGeo winning 72.9%, 40.6%, 57.5%, and 65.0% of windows. PEMS08 maintains a RoleGeo winner fraction of 77.5%–84.4%. Weather favors TimeGeo at horizons 96, 336, and 720, with TimeGeo winning 57.5%, 58.2%, and 62.9% of windows, while the two branches are nearly balanced at horizon 192. The "Any gain" annotations show that these winner comparisons occur in windows where at least one branch improves over Raw DLinear.

The window-level result compares temporal gain patterns, while the variable-pair analysis compares cross-variable improvement patterns. For variable pair (i, j), the co-improvement rate is the fraction of real test windows in which both variables reduce MSE relative to Raw DLinear. Each cell in Figure D.15 gives the difference between the RoleGeo and TimeGeo co-improvement rates, averaged



Figure D.10: Window-level branch gains and embedded winner strips on ETTm1. Each panel corresponds to one forecast horizon; the strip marks the branch with the larger gain and the annotation reports gain and winner fractions.

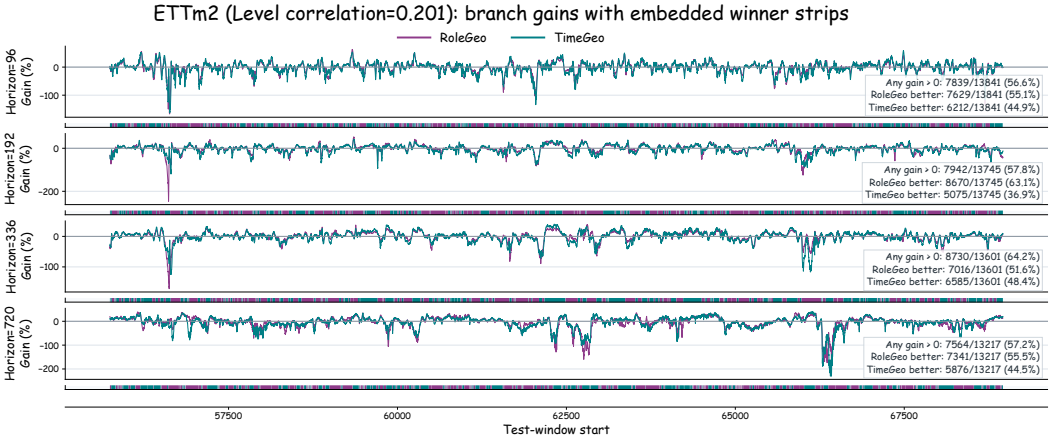


Figure D.11: Window-level branch gains and embedded winner strips on ETTm2. RoleGeo has its largest winner fraction at horizon 192, while the other horizons show closer branch competition.

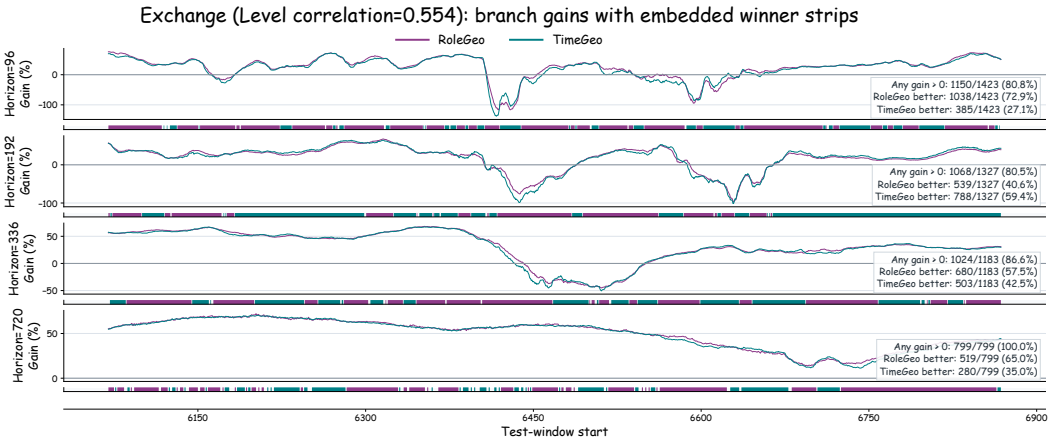


Figure D.12: Window-level branch gains and embedded winner strips on Exchange. RoleGeo leads at horizons 96, 336, and 720, while TimeGeo leads at horizon 192.

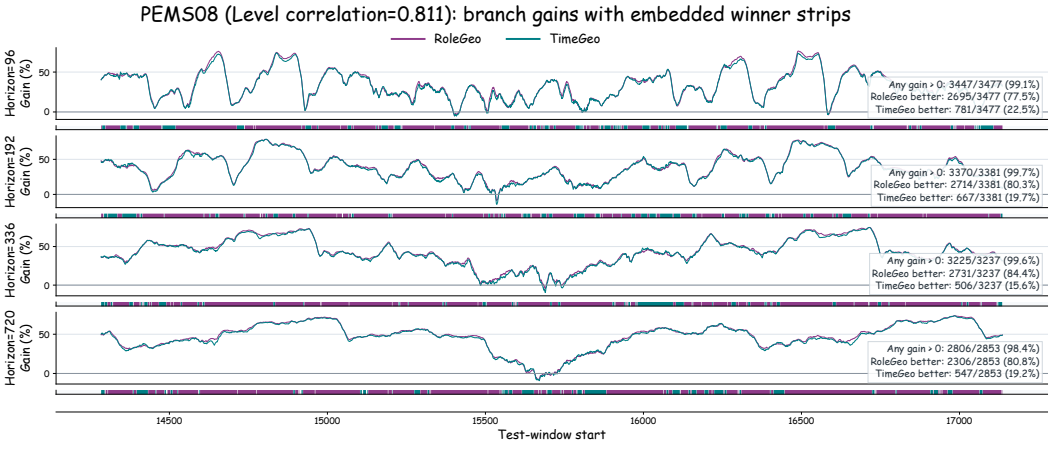


Figure D.13: Window-level branch gains and embedded winner strips on PEMS08. RoleGeo maintains a higher winner fraction at every horizon, indicating sustained value from system-state relations in this high-dimensional traffic dataset.



Figure D.14: Window-level branch gains and embedded winner strips on Weather. TimeGeo leads at horizons 96, 336, and 720, while the branches are nearly balanced at horizon 192.

over the four horizons. A positive value means that RoleGeo more frequently produces simultaneous improvements for that variable pair, and a negative value means the same for TimeGeo.

Figure D.15 contains broad positive regions for ETTm2, Exchange, and PEMS08, where RoleGeo more frequently produces simultaneous improvements for variable pairs. Weather contains distinct positive and negative blocks, with TimeGeo attaining higher co-improvement rates for some pairs. Together with the window-level comparison, these results establish distinct gain profiles: TimeGeo is more effective in particular local windows and variable pairs, whereas RoleGeo has broader window-level advantages and stronger grouped-variable improvements on datasets such as PEMS08. The two branches thus provide differentiated retrieval forecasts.

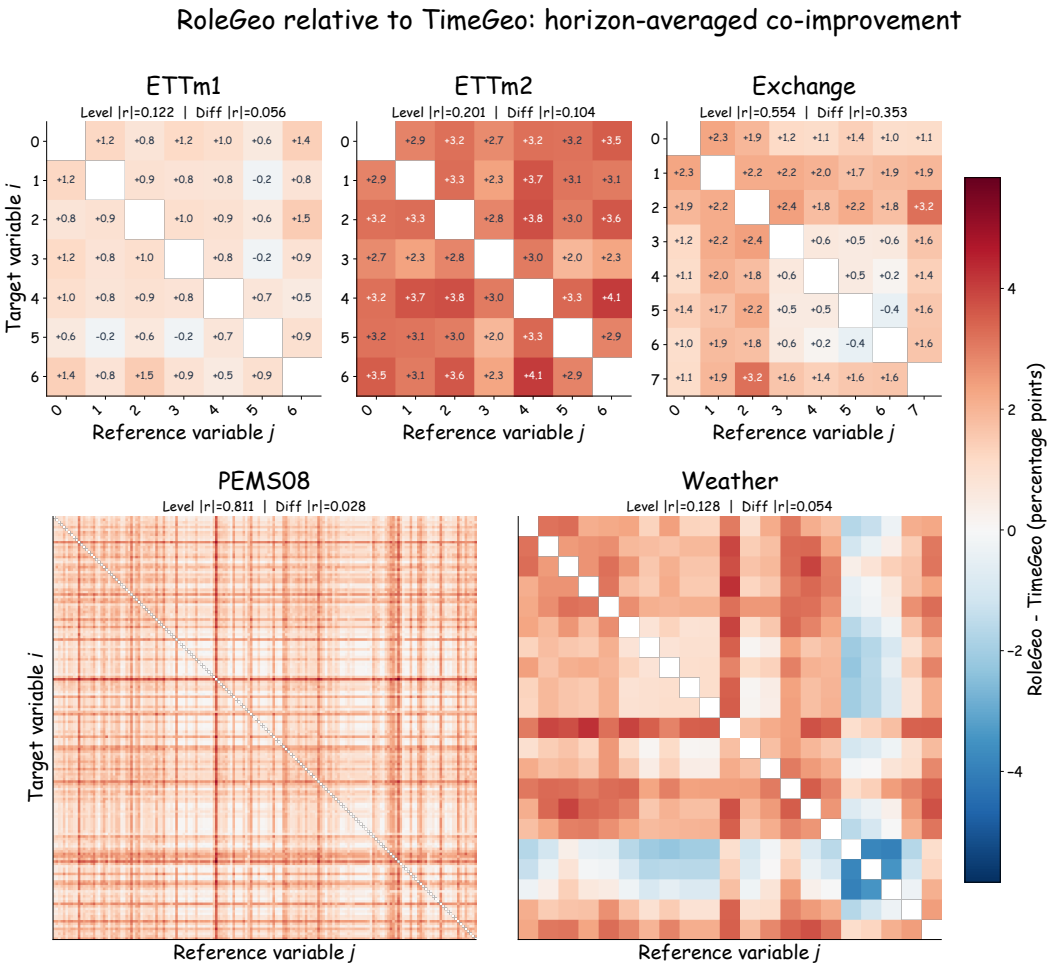


Figure D.15: Difference in variable-pair co-improvement rates between TimeGeo and RoleGeo. Rows and columns index variables i and j, and diagonal entries are hidden. Each cell gives the horizon-averaged RoleGeo co-improvement rate minus the TimeGeo co-improvement rate in percentage points. Red indicates more frequent co-improvement under RoleGeo, while blue indicates more frequent co-improvement under TimeGeo.

D.4 RQ4: Does Cross-Variate Relation Information Provide Stable Predictive Gains for RoleGeo?

This section treats the relation matrix W as the only intervened modeling factor. For each perturbed W, the RoleEncoder is retrained and the RoleGeo memory keys and memory bank are rebuilt from the resulting relational representations; relation residuals are recomputed using the current W, while the consistency coefficients c_i are fixed to their Full StableRel values. The Anchor, Backbone, RoleGeo architecture, retrieval capacity, training protocol, fusion module, and random seed remain identical across variants, and the TimeGeo branch and its cache are not perturbed. The analysis covers ETTh1, ETTh2, ETTm1, ETTm2, Weather, Exchange, and PEMS08, averaged over horizons 96, 192, 336, and 720. Full StableRel uses the relation matrix estimated from the complete training period. Zero relation removes relation information; Single-half relation uses one training interval; Random relation preserves edge density and weight scale while randomly rewiring connections; Permuted relation reassigns variable identities; and Sign-inverted relation preserves connectivity and magnitude while reversing relation signs.

Table D.2 reports the four-horizon average error changes relative to Full StableRel. Positive values indicate higher error and negative values indicate lower error; each variant reports MSE and MAE separately.

Table D.2: Relation-matrix variants relative to Full StableRel. Values are four-horizon average relative error changes (%). Positive values indicate error increases and are shown in red; negative values indicate error decreases and are shown in green.

Dataset	Zero relation		Single-half relation		Random relation		Permuted relation		Sign-inverted relation	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	+0.029	+0.004	-0.031	+0.010	+0.095	+0.047	+0.003	+0.021	+0.040	+0.014
ETTh2	+0.482	+0.210	+0.022	+0.017	-0.010	-0.007	+0.120	+0.083	+0.051	+0.033
ETTm1	+0.066	+0.057	+0.028	+0.011	+0.095	+0.041	+0.061	+0.074	+0.053	+0.034
ETTm2	+0.480	+0.116	+0.063	+0.040	+0.035	+0.025	+0.141	+0.092	-0.038	+0.006
Weather	+0.241	+0.194	+0.307	+0.215	+0.191	+0.154	+0.194	+0.154	+0.148	+0.127
Exchange	+5.947	+2.180	+0.078	+0.082	+0.608	+0.222	+0.661	+0.359	+0.000	+0.000
PEMS08	+0.330	+0.007	+0.316	-0.038	+0.095	+0.018	+0.096	+0.148	+0.357	+0.101
Avg.	+1.082	+0.395	+0.112	+0.048	+0.158	+0.071	+0.182	+0.133	+0.087	+0.045

Zero relation produces the clearest degradation. Across 28 dataset–horizon settings, its average MSE and MAE increase by 1.08% and 0.395%, with increases in 21/28 MSE and 20/28 MAE entries. The largest effect appears on Exchange, where MSE and MAE rise by 5.947% and 2.180%; ETTh2 and ETTm2 also show MSE increases of approximately 0.48%. These results indicate that cross-variate relation information contributes directly to RoleGeo's predictive geometry.

The other relation constructions remain close to Full StableRel overall. Their average MSE changes are +0.112% for Single-half, +0.158% for Random, +0.182% for Permuted, and +0.087% for Sign-inverted. RoleGeo retains stable behavior under changes to the relation partition, topology, variable ordering, and sign convention, while the magnitude of relation-conditioned gains varies with dataset dynamics.

D.5 RQ5: Is a Simple Static Relation Residual Sufficient for RoleGeo's Predictive Gains?

RoleGeo builds its relational representation from first differences, a split-half StableRel matrix, and a linear relation residual. This section tests whether this static residual is sufficient for forecasting, and whether window-level relations, current peer histories, or a wider memory candidate pool provide consistent additional gains.

All variants keep DLinear, input and prediction lengths, retrieval capacity, fusion modules, and the random seed fixed, changing one relational factor at a time. The analysis covers ETTh1, ETTh2, Exchange, ETTm1, ETTm2, and Weather at horizons 96, 192, 336, and 720. PEMS08 has Full and Backbone results only and is therefore excluded from the variant averages.

Table D.3 reports four-horizon average error changes relative to Full. Negative values indicate lower error and positive values indicate higher error. The overall columns average the six datasets with complete variant results; Exchange, Weather, and ETTm1 illustrate distinct dataset-level patterns.

Table D.3: RoleGeo mechanism variants relative to Full. Values are four-horizon average relative error changes (%). Positive values indicate error increases and are shown in red; negative values indicate error decreases and are shown in green. Dataset groups follow the appendix order, with the six-dataset average in the final group.

Variant	Changed factor	ETTh1		ETTh2		Exchange		ETTm1		ETTm2		Weather		Overall	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
First-half relation	relation interval	-0.02	+0.01	+0.01	+0.01	+0.11	+0.09	-0.05	-0.02	+0.08	-0.03	+0.31	+0.21	+0.07	+0.05
Second-half relation	relation interval	-0.01	+0.02	+0.07	+0.03	-0.02	-0.15	-0.06	-0.05	+0.04	+0.02	+0.17	+0.07	+0.03	-0.01
Global relation	remove split-half	+0.00	+0.00	+0.06	+0.04	+1.87	+0.49	-0.01	-0.02	+0.24	+0.03	-0.04	+0.01	+0.36	+0.09
Spearman relation	correlation estimator	+0.03	+0.02	+0.06	+0.04	+0.57	-0.03	-0.06	-0.05	+0.12	-0.03	+0.34	+0.22	+0.18	+0.03
Dynamic-W target-only	window relation	+0.02	+0.01	+0.03	+0.01	-0.27	-0.05	+0.03	-0.01	-0.04	-0.02	+0.10	+0.02	-0.02	-0.01
Static peer query	current peer histories	+0.61	+0.52	-1.99	-0.92	+6.77	+4.14	+0.86	+0.29	-1.93	-0.85	+1.38	+1.71	+0.95	+0.81
Dynamic peer query	dynamic relation and peers	+0.67	+0.50	-2.32	-1.03	+6.28	+3.45	+0.74	+0.49	-1.98	-0.88	+1.30	+1.72	+0.78	+0.71
Level role signal	relation signal	+0.04	+0.03	+0.26	+0.13	-0.01	-0.13	-0.11	-0.02	+0.02	-0.02	+0.33	+0.20	+0.09	+0.03
All-channel memory	candidate scope	-0.01	-0.25	+0.12	+0.07	+2.51	+0.90	+0.01	-0.36	-0.09	-0.13	+0.50	+0.25	+0.51	+0.08
Relation-peer memory	candidate scope	+0.04	-0.24	+0.24	+0.15	+0.68	+0.18	-0.07	-0.49	-0.00	-0.09	+0.36	+0.22	+0.21	-0.04

Replacing the relation estimator produces only small overall changes. First-half, second-half, global, and Spearman relation yield average MSE changes of +0.07%, +0.03%, +0.36%, and +0.18%, respectively. Exchange is more sensitive to the global relation, whereas Weather and ETTm1 remain close to zero. The predictive gain therefore does not depend on one particular correlation estimator; the static StableRel constraint remains effective across the tested datasets.

Query formation produces the clearest separation. Dynamic-W target-only nearly matches Full, with an overall MSE change of -0.02%. Static peer query and dynamic peer query increase overall MSE by 0.95% and 0.78%, respectively. On Exchange, their MSE increases reach 6.77% and 6.28%; Weather and ETTm1 show the same positive direction. Injecting current peer histories into the query reduces RoleGeo retrieval stability, while window-level relation updates provide no consistent additional gain.

Signal form and candidate scope are secondary factors. Level role signal changes overall MSE by +0.09%; all-channel and relation-peer memory change it by +0.51% and +0.21%. Neither a wider candidate set nor a hard peer restriction yields a consistent cross-dataset improvement. The relation matrix primarily guides retrieval and constructs the relational residual.

Together, Table D.3 shows that static StableRel and the target-history relation residual can support RoleGeo's predictive gains across six datasets and four horizons. Window-level relation updates add no average benefit, and current peer histories can substantially reduce retrieval quality. The evidence supports describing RoleGeo as a relation-guided representation and retrieval mechanism. The experiment establishes predictive sufficiency for the tested relation representation; lead–lag, nonlinear, and finer-grained locally time-varying dependencies require dedicated modeling and experiments.

D.6 RQ6: How Sensitive Is CoReG to Its Key Retrieval and Geometry Hyperparameters?

CoReG shows stronger performance sensitivity to retrieval-capacity parameters and flatter responses to TimeGeo scales and StableRel peers. The analysis sweeps five component groups on DLinear and xPatch across ETTh1, ETTh2, ETTm1, ETTm2, and Exchange at horizons 96, 192, 336, and 720. Each sweep changes one component; MSE and MAE are reported as relative changes from the reference value within the same Dataset–Horizon setting. The Top-K reference is $K = 50$ with a candidate pool of 100, while the remaining reference values appear in Table D.4. Negative and positive changes indicate lower and higher error, respectively. All results use seed 2027.

Table D.4: Candidate values and reference values for the hyperparameter sensitivity analysis.

Parameter group	Candidate values	Reference value
Top-K	10, 25, 50, 100	50
Memory size	250, 500, 1000, 1500, 2000	1000
Embedding dimension	8, 16, 32	16
TimeGeo scales	1; 2–1; 4–2–1; 8–4–2–1	4–2–1
StableRel peers	1, 2, 4, all	all

Figure D.16 shows that Top-K, memory size, and embedding dimension have broader MSE and MAE change distributions, with especially pronounced tails for DLinear. The distributions for TimeGeo scales and StableRel peers concentrate closer to zero. This complete-setting summary associates retrieval candidate count, memory size, and matching representation capacity with larger error-response ranges, while the scale composition of the geometry views and the StableRel peer range have flatter responses over the tested candidates.

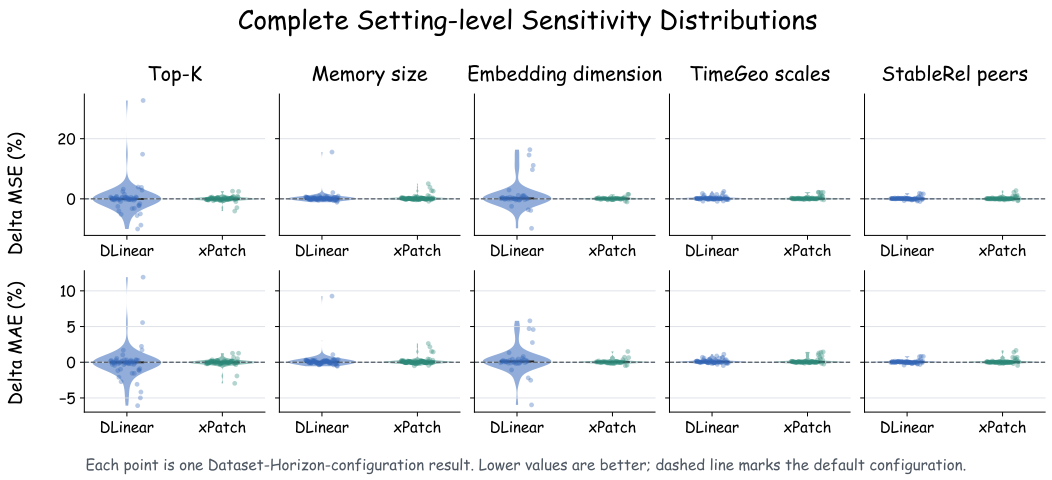


Figure D.16: Complete setting-level hyperparameter sensitivity distributions. Columns correspond to Top-K, memory size, embedding dimension, TimeGeo scales, and StableRel peers; rows report relative MSE and MAE changes from the reference value, respectively. Each point is one Dataset–Horizon–configuration result. The dashed line denotes zero change, and lower values are better.

Figures D.17–D.20 expand the horizon-wise trajectories on ETTh1 and ETTh2 and compare the same parameter groups across DLinear and xPatch. The ETTh1 curves remain close to zero across the five groups. On ETTh2, DLinear shows larger horizon-dependent responses to Top-K and embedding dimension: small Top-K values yield pronounced negative changes in some long-horizon settings, whereas larger Top-K values produce positive changes in other settings. The corresponding xPatch trajectories are more compact. Together, the two Dataset–Backbone comparisons show that hyperparameter effects depend jointly on component type, Dataset, Backbone, and horizon.

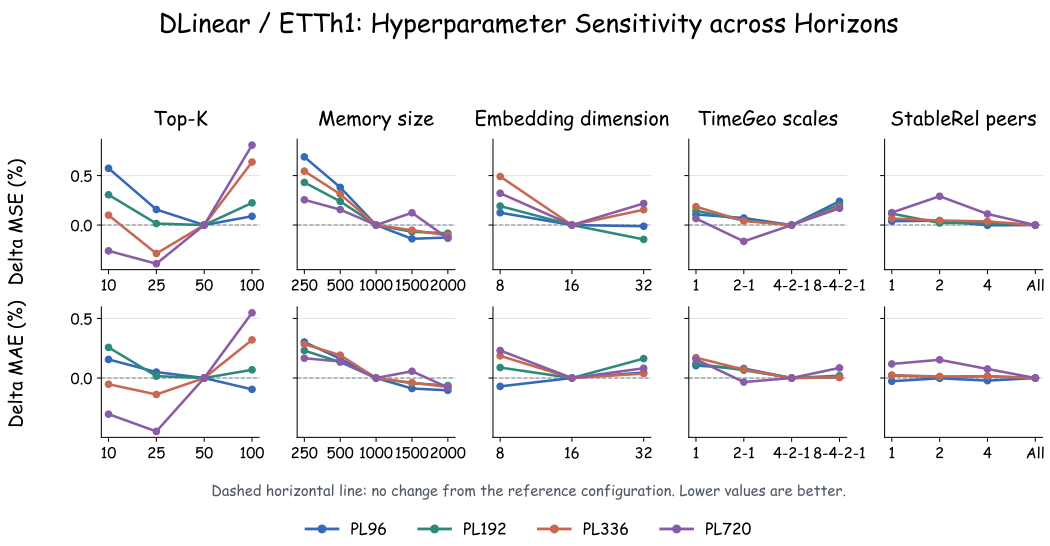


Figure D.17: Horizon-wise hyperparameter sensitivity of DLinear on ETTh1. Columns correspond to the five parameter-group sweeps; rows report relative MSE and MAE changes from the reference value. Colors identify the four horizons, and the dashed line denotes zero change.

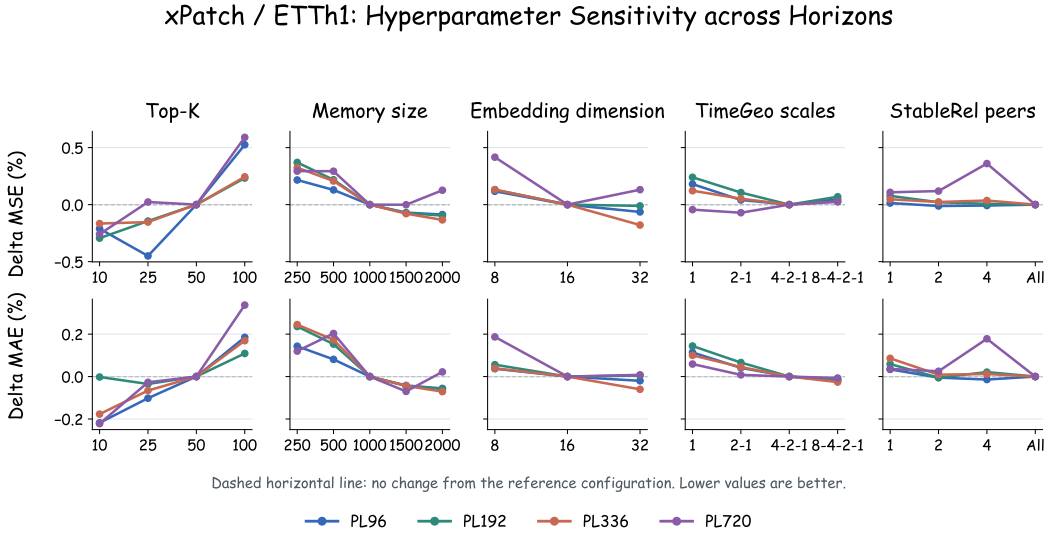


Figure D.18: Horizon-wise hyperparameter sensitivity of xPatch on ETTh1. The fields, colors, and reference line follow Figure D.17.

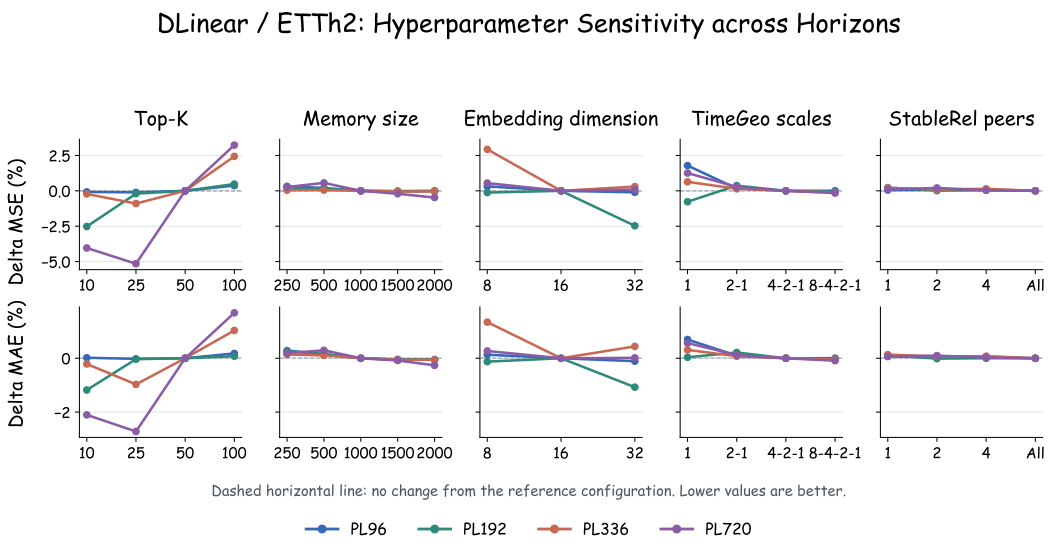


Figure D.19: Horizon-wise hyperparameter sensitivity of DLinear on ETTh2. Columns correspond to the five parameter-group sweeps; rows report relative MSE and MAE changes from the reference value. Colors identify the four horizons, and the dashed line denotes zero change.

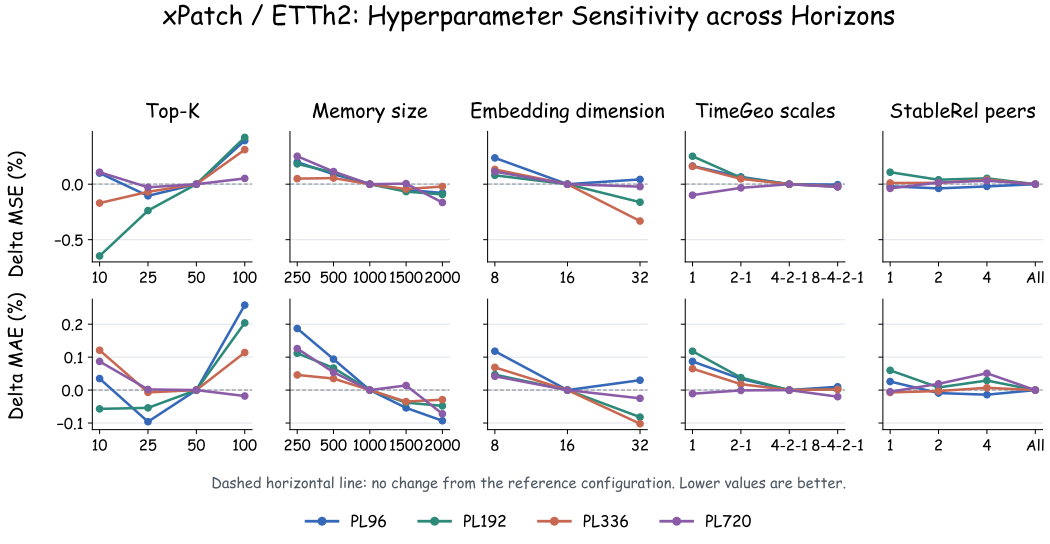


Figure D.20: Horizon-wise hyperparameter sensitivity of xPatch on ETTh2. The fields, colors, and reference line follow Figure D.19.

D.7 RQ7: Does CoReG Remain Stable Across Random Seeds?

CoReG maintains lower forecasting errors than Raw across five random seeds. The multi-seed study covers ETTh1 and ETTh2 with DLinear, iTransformer, and MICN at horizons 96, 192, 336, and 720. Each Dataset–Backbone–Horizon configuration includes seeds 2025–2029. We retain the five seeds as independent paired observations and conduct two-sided paired Wilcoxon and paired t tests on 120 configuration–seed observations, with Holm correction for multiple comparisons.

Table D.5 reports the seed-level paired effects and Holm-corrected significance tests. CoReG is lower than Raw in all 120 paired observations for both metrics, while comparisons with other plug-in modules quantify their relative differences against both references.

Table D.5: Seed-level paired effects and Holm-corrected significance tests. Raw and CoReG are the top-level references. Each comparison uses 120 unaggregated Dataset–Backbone–Horizon–seed observations. Δ, Rel., and Wins denote mean error difference, relative change, and lower-error observations. Each Holm p cell lists the paired Wilcoxon (W) and paired t-test (t) results.

Method	Metric	Reference Raw				Reference CoReG			
		Δ	Rel.	Wins	Holm p	Δ	Rel.	Wins	Holm p
CoReG	MSE	-0.022589	-5.67%	120/120	$W: 9.86\text{e}{-21}$ $t: 1.86\text{e}{-26}$			–	
	MAE	-0.016043	-3.59%	120/120	$W: 9.86\text{e}{-21}$ $t: 9.38\text{e}{-31}$			–	
SARAF	MSE	-0.000406	-0.46%	62/120	$W: 0.895$ $t: 0.796$	+0.022183	+5.61%	5/120	$W: 2.29\text{e}{-20}$ $t: 7.23\text{e}{-27}$
	MAE	-0.005448	-1.32%	77/120	$W: 2.38\text{e}{-4}$ $t: 5.08\text{e}{-5}$	+0.010595	+2.37%	12/120	$W: 8.82\text{e}{-19}$ $t: 4.75\text{e}{-24}$
GTR	MSE	+0.068077	+19.47%	34/120	$W: 7.76\text{e}{-7}$ $t: 8.02\text{e}{-4}$	+0.090666	+26.68%	5/120	$W: 1.61\text{e}{-20}$ $t: 1.73\text{e}{-6}$
	MAE	+0.039114	+9.41%	36/120	$W: 1.77\text{e}{-6}$ $t: 5.08\text{e}{-5}$	+0.055157	+13.53%	5/120	$W: 3.56\text{e}{-20}$ $t: 2.43\text{e}{-9}$
δ-Adapter	MSE	+0.007739	+2.35%	23/120	$W: 6.30\text{e}{-10}$ $t: 3.26\text{e}{-9}$	+0.030328	+8.74%	6/120	$W: 3.66\text{e}{-20}$ $t: 3.19\text{e}{-30}$
	MAE	+0.002916	+0.79%	41/120	$W: 2.38\text{e}{-4}$ $t: 8.15\text{e}{-4}$	+0.018959	+4.62%	6/120	$W: 1.31\text{e}{-19}$ $t: 9.67\text{e}{-29}$
GLAFF	MSE	+0.024834	+7.87%	35/120	$W: 4.33\text{e}{-6}$ $t: 9.56\text{e}{-4}$	+0.047423	+15.08%	2/120	$W: 1.27\text{e}{-20}$ $t: 4.84\text{e}{-9}$
	MAE	+0.013480	+3.43%	47/120	$W: 1.48\text{e}{-3}$ $t: 8.15\text{e}{-4}$	+0.029523	+7.50%	3/120	$W: 1.15\text{e}{-20}$ $t: 8.93\text{e}{-10}$

Figure D.21 visualizes the distribution of these seed-level effects across all modules and configurations.

Complete per-configuration five-seed means and sample standard deviations are listed in Tables E.11–E.13.

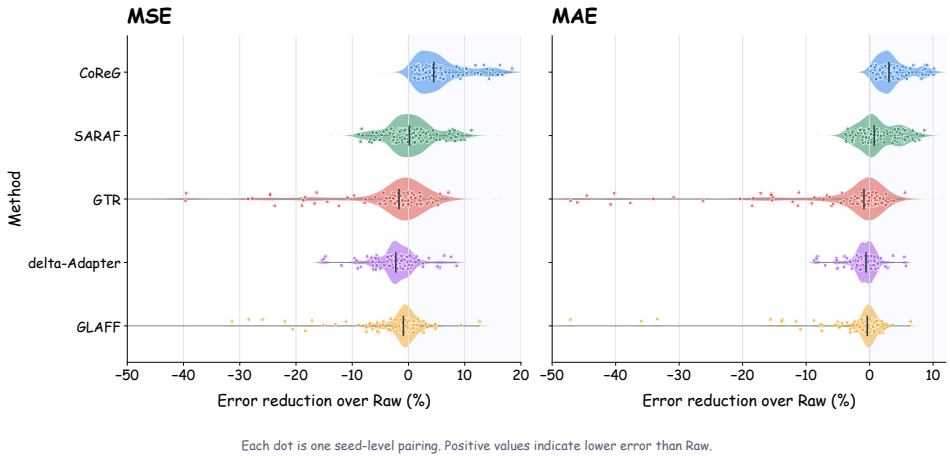


Figure D.21: Distribution of seed-level paired effects across all modules and forecasting configurations. Each point corresponds to one Dataset–Backbone–Horizon–seed observation; negative values indicate lower error for the compared method than its reference.

D.8 RQ8: How Does CoReG Respond to Reduced Training Data Across Dataset Structures and Backbones?

This section analyzes how CoReG responds as the available training data is progressively reduced, and how the response varies with dataset structure and host backbone. We evaluate ETTm1, ETTm2, Exchange, and PEMS08 with DLinear, iTransformer, PatchMLP, and Amplifier at horizons 96, 192, 336, and 720. For each dataset, nested training-window subsets retain 100%, 50%, 25%, and 10% of the training data, while validation and test windows, input length, forecasting horizon, CoReG configuration, and random seed remain fixed. All changes are measured relative to the corresponding 100% setting for the same dataset, backbone, and horizon.

Table D.6 summarizes the dataset characteristics used to contextualize the comparison. ETTm1 and ETTm2 have the same time-point count, variable count, and dominant period, but differ in seasonality, trend strength, and cross-variable correlation. Exchange has the fewest time points and the longest dominant period, whereas PEMS08 has the largest variable count and strongest level correlation together with low differenced correlation and effective-rank ratio. The table reports the definitions of all eight features used by the plots.

Table D.6: Time-scale, variable-structure, and correlation characteristics of the four datasets used in the data-sparsity experiment.

Dataset	T	C	Period	Seasonality	Trend	Level Corr.	Diff Corr.	Corr. Eff. Rank
ETTm1	69,680	7	96	0.924709	0.024729	0.121568	0.055520	0.649875
ETTm2	69,680	7	96	0.955785	0.229625	0.201076	0.104373	0.618859
Exchange	7,588	8	690	0.950949	0.232419	0.553930	0.353110	0.406098
PEMS08	17,856	170	288	0.914021	0.000451	0.810796	0.028456	0.020390

T is the number of observed time points and C is the number of jointly observed variables. Period is the median dominant period identified from detrended FFT spectra. Seasonality is the median detrended lag correlation at the corresponding dominant period. Trend is the median normalized low-frequency trend strength. Level Corr. and Diff Corr. are the median absolute Pearson correlations for level and first-differenced variables, respectively. Corr. Eff. Rank is the effective-rank ratio of the cross-variable correlation structure; lower values indicate a more concentrated low-rank structure.

Figures D.22–D.25 show the MAE response at 100%, 50%, 25%, and 10% training data. Each figure contains eight feature panels, with datasets reordered from low to high according to the feature shown in that panel. The open square marks the 100% baseline, and the three filled bars show changes from that baseline at the reduced training ratios. Avg. reports the arithmetic mean over the four horizons. The panels provide alternative dataset-ordering views of the same sparsity response.

Across the four backbone groups, the MAE response shows clear dataset and architecture differences as training data is reduced. The overall change tends to grow at lower retained ratios, while individual dataset–horizon combinations can fluctuate around this trend. Exchange and PEMS08 usually show larger responses across several ratios and horizons, whereas ETTm1 and ETTm2 follow different response trajectories despite their similar time-point and variable counts. Dataset orderings also vary across feature panels, backbones, and ratios. These observations support a joint view in which data coverage, dataset structure, forecasting horizon, and the host backbone shape CoReG's sample efficiency.

Tables E.14–E.17 provide the complete Dataset–Horizon–Ratio results. Raw columns give the original backbone reference, and the remaining columns report CoReG at 100%, 50%, 25%, and 10% training ratios.

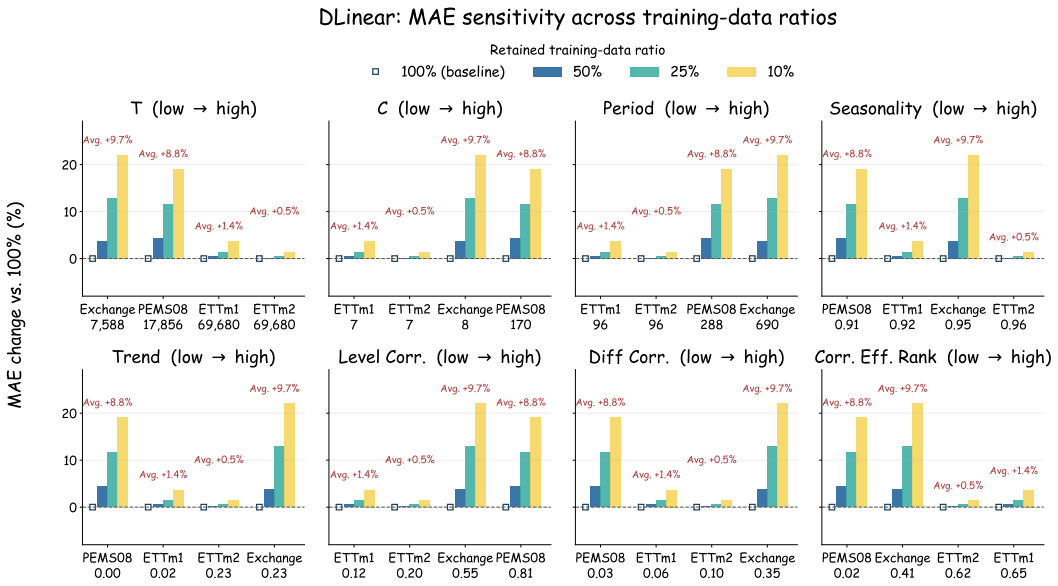


Figure D.22: DLinear MAE response across retained training-data ratios. The eight panels sort datasets by T, C, Period, seasonality, trend, Level Corr., Diff Corr., and Corr. Eff. Rank; bars show changes relative to the corresponding 100% setting, with Avg. denoting the four-horizon mean.

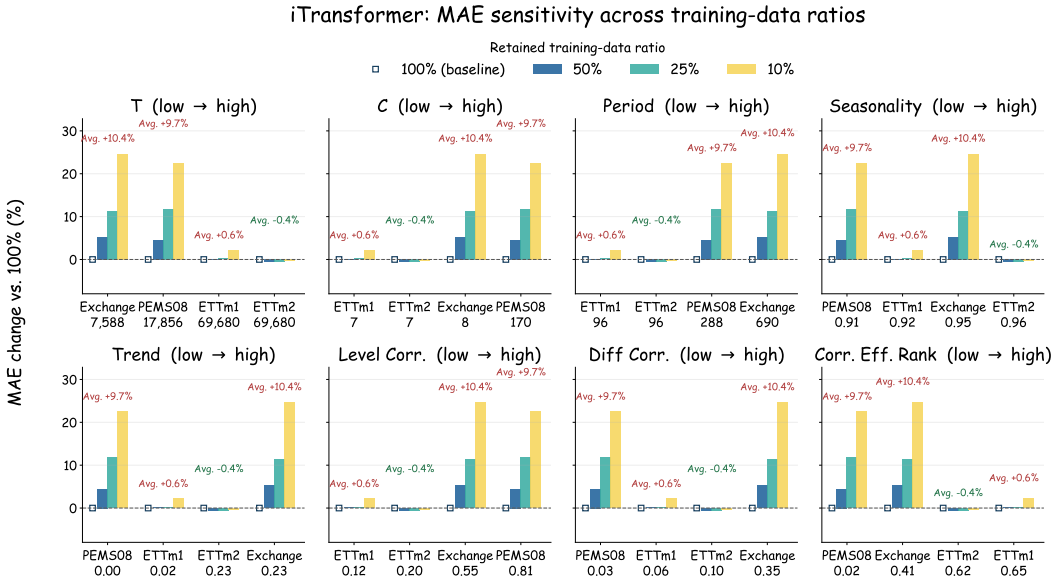


Figure D.23: iTransformer MAE response across retained training-data ratios. The eight panels sort datasets by T, C, Period, seasonality, trend, Level Corr., Diff Corr., and Corr. Eff. Rank; bars show changes relative to the corresponding 100% setting, with Avg. denoting the four-horizon mean.

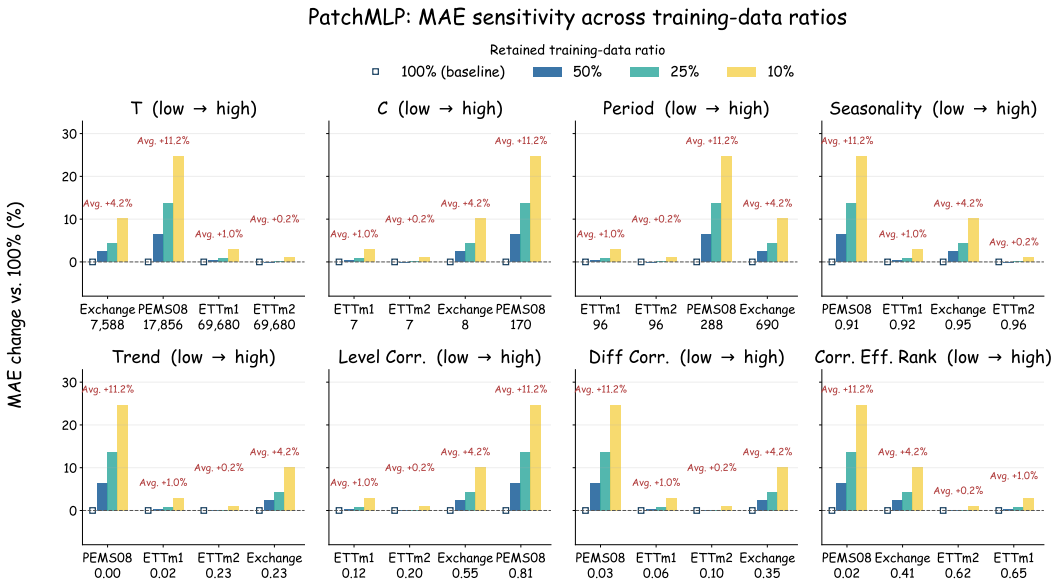


Figure D.24: PatchMLP MAE response across retained training-data ratios. The eight panels sort datasets by T, C, Period, seasonality, trend, Level Corr., Diff Corr., and Corr. Eff. Rank; bars show changes relative to the corresponding 100% setting, with Avg. denoting the four-horizon mean.

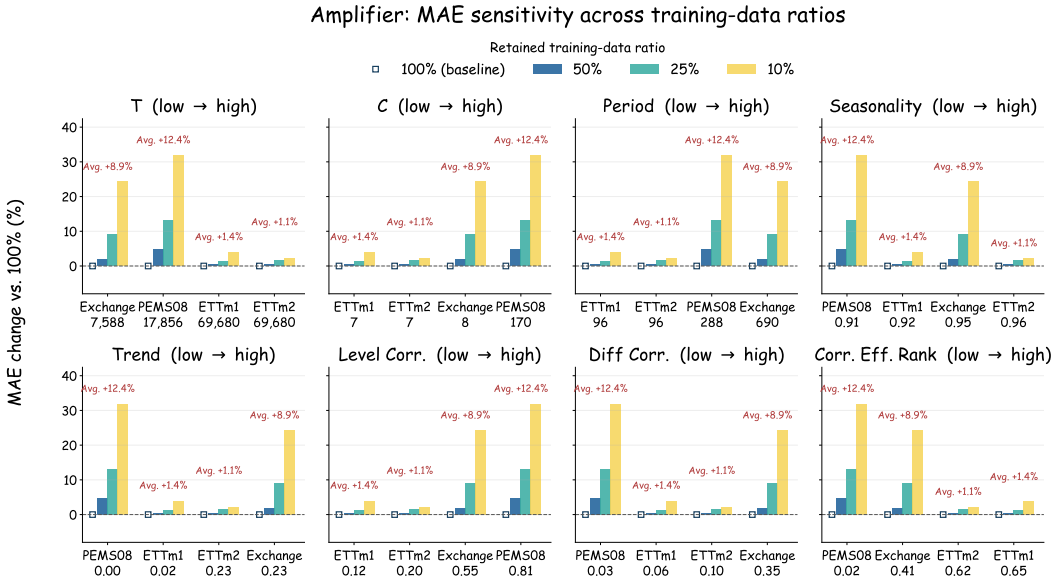


Figure D.25: Amplifier MAE response across retained training-data ratios. The eight panels sort datasets by T, C, Period, seasonality, trend, Level Corr., Diff Corr., and Corr. Eff. Rank; bars show changes relative to the corresponding 100% setting, with Avg. denoting the four-horizon mean.

D.9 RQ9: Does Final-View Merging Require Complex Learnable Weights?

FixedMean provides a reasonable, lightweight, and stable default rule for merging the two prediction views. CoReG produces two views through the independent GeoRead modules of TimeGeo and RoleGeo, and combines them as $\hat{Y} = 0.5Y_T + 0.5Y_R$. Single-geometry ablations show that TimeGeo Only and RoleGeo Only each retain an effective retrieval prediction path, while FULL achieves the most consistent low errors by combining both paths. We therefore test whether input-dependent or fine-grained weights deliver consistent additional gains after the two geometries have independently read their retrieval evidence.

To isolate the final merging rule, we use the Post-hoc protocol: the Backbone, retrieval representations, and two trained GeoRead modules are fixed, while only the parameterization of the mixing coefficient η changes. FixedMean uses $\eta = 0.5$ throughout. GlobalScalar learns one task-level scalar, ChannelGate learns one scalar per variable channel, HorizonGate learns one scalar per forecast position, and EvidenceGate dynamically produces one weight per sample and variable channel from retrieval scores and disagreement features. All learnable variants initialize at $\eta = 0.5$. We evaluate ETTh1, ETTh2, ETTm1, ETTm2, Weather, and Exchange with DLinear, PatchMLP, iTransformer, and MICN at horizons 96 and 720.

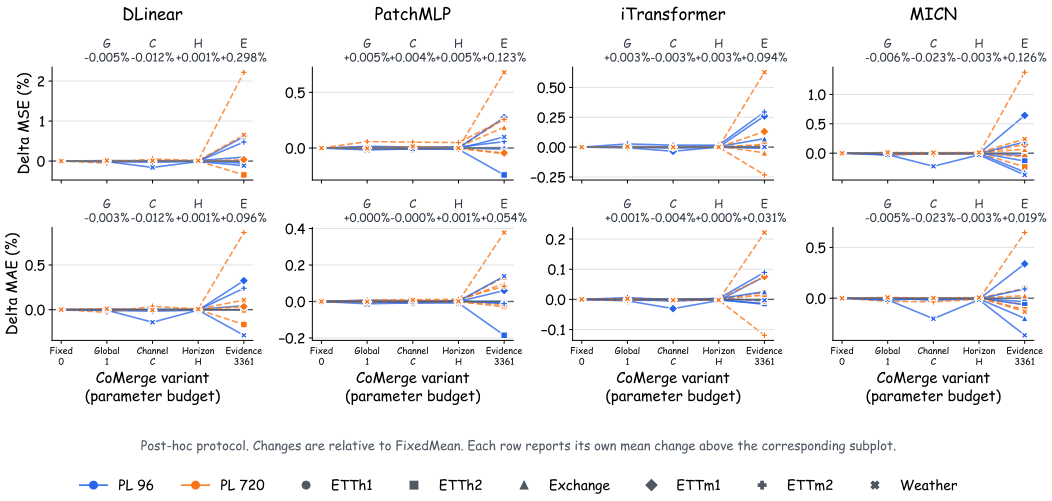


Figure D.26: Post-hoc error changes of final-view merging rules relative to FixedMean. Columns correspond to DLinear, PatchMLP, iTransformer, and MICN; rows report MSE and MAE, respectively. The horizontal axis orders FixedMean, GlobalScalar, ChannelGate, HorizonGate, and EvidenceGate and gives their CoMerge parameter budgets. Color and line style distinguish horizons 96 and 720, and marker shapes distinguish datasets. The vertical axis reports the error change for the same setting relative to FixedMean; the zero line denotes parity. The annotations above each subplot give the average change for the corresponding rule and Backbone.

Figure D.26 shows that the learnable rules induce only near-zero changes across Backbones, with directions that vary by setting. EvidenceGate produces no additional average error reduction on any of the four Backbones. Complex final weighting therefore yields no consistent cross-architecture gain. FixedMean directly integrates the two read prediction views with no additional gate parameters, providing a reasonable, lightweight, and stable default merging rule for CoReG. Complete per-configuration results for the merging rules are listed in Tables E.7–E.10.

D.10 RQ10: Which Online Paths Account for CoReG's Computation?

This section uses isolated FP32 CUDA-forward timing to attribute CoReG's online computation across its execution paths. We select three configurations with clearly visible incremental costs to expose the time composition of Shared, TimeGeo, RoleGeo, and Merge/output: Exchange/DLinear/96, Weather/iTransformer/336, and PEMS08/MICN/720. The three configurations span low, intermediate, and high absolute CUDA workloads. Each case uses paired Raw inputs, the matched Backbone checkpoint, retrieval caches, batch size 64, and seed 2027, with 5 warmup and 20 measurement iterations.

Table D.7: Online CUDA-time attribution in three representative configurations. Values report isolated-forward CUDA time in ms per batch. Shared includes input preparation, H2D transfer, and Backbone forward computation; Merge/output includes FixedMean merging, test diagnostic synchronization, and output handling. RoleGeo relation construction is completed offline and excluded from the online paths.

Case	Raw	CoReG	Shared	TimeGeo	RoleGeo	Merge/output
Exchange / DLinear / 96	0.518	5.438	0.442	3.381	1.310	0.305
Weather / iTransformer / 336	1.886	13.563	1.924	8.715	2.452	0.471
PEMS08 / MICN / 720	82.757	235.205	79.472	123.761	27.173	4.800

The CoReG-to-Raw CUDA-forward time ratios are $10.50\times$, $7.19\times$, and $2.84\times$ for the three configurations, respectively. These ratios characterize path-level increments in the selected diagnostic cases and locate their computational sources. Table D.7 and Figure D.27 show that TimeGeo is the largest online path in all three cases, accounting for 62.2%, 64.3%, and 52.6% of total CoReG CUDA time. RoleGeo accounts for 24.1%, 18.1%, and 11.6%, respectively. Final merging and output handling contribute only 2.0%–5.6% across the settings, so FixedMean remains a small online component.

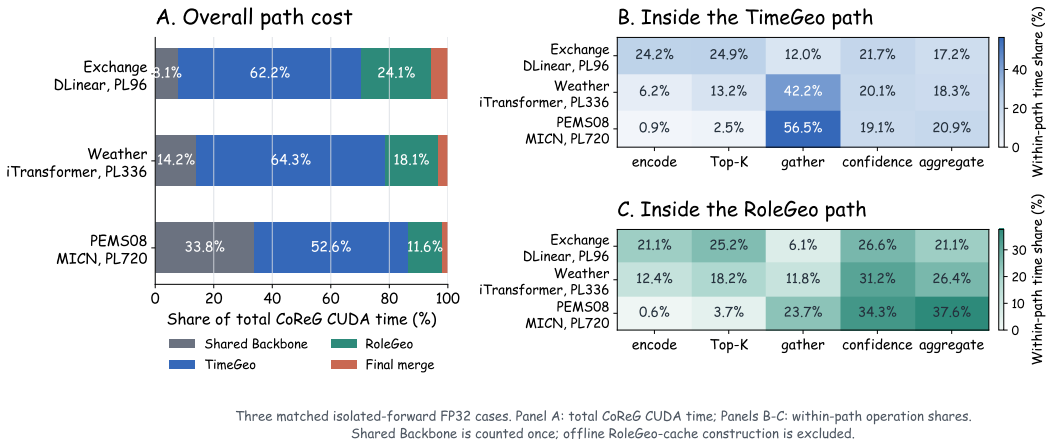


Figure D.27: Hierarchical online CUDA-time attribution for three representative CoReG cases. Panel A reports the shares of Shared Backbone, TimeGeo, RoleGeo, and Final merge in total CoReG CUDA time. Panels B and C aggregate the within-path shares of encode, Top-K, gather, confidence, and aggregate operations for TimeGeo and RoleGeo, respectively. Shared Backbone is counted once, and offline RoleGeo-cache construction is excluded.

Figure D.27 further identifies the within-path hotspot structure. TimeGeo cost on Exchange is distributed across encoding, Top-K, confidence, and aggregation, whereas future gather accounts for 42.2% and 56.5% of the TimeGeo path on Weather and PEMS08, respectively. RoleGeo cost concentrates on confidence and aggregation, which account for 34.3% and 37.6% of the RoleGeo path in the PEMS08/MICN/720 case. Multiscale candidate access and the subsequent evidence readout provide the principal optimization targets for CoReG's online implementation.

Figure D.28 decomposes TimeGeo by scale. Scale 2 has the highest measured CUDA time in each profiled case, with the largest separation from adjacent scales in the heavy case. This pattern identifies Scale 2 candidate access and readout as a priority for further implementation optimization.

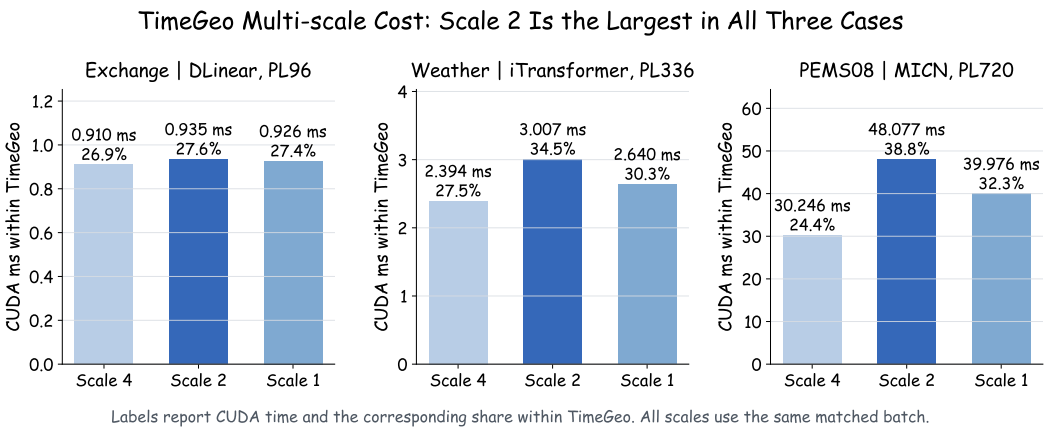


Figure D.28: Online CUDA-time decomposition of TimeGeo's multiscale path. The panels correspond to Exchange/DLinear/96, Weather/iTransformer/336, and PEMS08/MICN/720. Each bar reports CUDA time for one scale under the same paired batch, along with its share within the TimeGeo path. Scale 2 is the highest-time scale in all profiled cases.

This path-level decomposition complements the end-to-end latency distribution over the complete matched-setting collection: the end-to-end measurement describes overall latency across datasets and horizons, while the isolated CUDA profiler exposes how incremental computation is allocated across online paths. The three appendix cases therefore provide computation attribution at selected diagnostic settings. CoReG's main online costs are TimeGeo multiscale retrieval, particularly Scale 2 candidate access and future gather, together with RoleGeo confidence computation and neighbor aggregation. FixedMean combines the two prediction views with fixed weights and contributes a small online time share.

E Additional Experimental Results

E.1 Complete Paired Results

CoReG consistently reduces forecasting errors across modeling families, datasets, and forecasting horizons. To answer this question, Tables E.1 and E.2 report complete paired results on ten multivariate forecasting datasets using MSE and MAE, respectively. Each Backbone and its CoReG-enhanced counterpart use the same full training run and checkpoint, and all metrics are shown to four decimal places.

Table E.1: Complete paired MSE results of CoReG across forecasting backbones and datasets. Lower is better. Green indicates a decrease after adding CoReG, while red indicates an unchanged or increased error. Average is computed over four forecasting horizons, with the relative change shown below each CoReG average.

Dataset	Horizon	DLinear (2023)		MICN (2023)		iTransformer (2024)		PatchMLP (2025)		xPatch (2025)		Amplifier (2025)	
		Original	+CoReG	Original	+CoReG	Original	+CoReG	Original	+CoReG	Original	+CoReG	Original	+CoReG
ETTh1	96	0.4302	0.4223	0.4895	0.4606	0.4421	0.4229	0.4798	0.4356	0.4227	0.4159	0.5720	0.4285
	192	0.4711	0.4628	0.5041	0.4869	0.4995	0.4710	0.5270	0.4841	0.4731	0.4653	0.5855	0.4783
	336	0.5202	0.5067	0.5273	0.5183	0.5392	0.5146	0.5854	0.5341	0.5262	0.5147	0.6133	0.5213
	720	0.6278	0.6163	0.6141	0.6072	0.6745	0.6488	0.7119	0.6739	0.6555	0.6496	0.6808	0.6591
	Avg.	0.5123	0.5020 ↓2.01%	0.5337	0.5182 ↓2.90%	0.5388	0.5143 ↓4.55%	0.5761	0.5319 ↓7.67%	0.5194	0.5114 ↓1.54%	0.6129	0.5218 ↓14.86%
ETTh2	96	0.2275	0.2232	0.2597	0.2249	0.2426	0.2287	0.2772	0.2464	0.2167	0.2138	0.3066	0.2304
	192	0.2867	0.2723	0.3020	0.2701	0.2943	0.2855	0.3366	0.2983	0.2738	0.2696	0.3312	0.2828
	336	0.3261	0.2969	0.3361	0.3015	0.3560	0.3256	0.4141	0.3907	0.3163	0.3101	0.3648	0.3278
	720	0.4522	0.3804	0.4414	0.3772	0.4183	0.4077	0.5395	0.5000	0.4063	0.3922	0.4459	0.4215
	Avg.	0.3231	0.2932 ↓9.25%	0.3348	0.2934 ↓12.37%	0.3278	0.3119 ↓4.85%	0.3918	0.3589 ↓8.40%	0.3033	0.2964 ↓2.27%	0.3621	0.3156 ↓12.84%
ETTm1	96	0.3274	0.3132	0.3395	0.3164	0.3343	0.3193	0.3414	0.3234	0.3189	0.3114	0.3308	0.3188
	192	0.3760	0.3563	0.3807	0.3614	0.3894	0.3677	0.3968	0.3784	0.3728	0.3671	0.3775	0.3665
	336	0.4181	0.4027	0.4202	0.4026	0.4318	0.4120	0.4432	0.4195	0.4196	0.4112	0.4218	0.4095
	720	0.4639	0.4518	0.4670	0.4528	0.4779	0.4667	0.4838	0.4647	0.4707	0.4684	0.4777	0.4652
	Avg.	0.3964	0.3810 ↓3.88%	0.4019	0.3833 ↓4.63%	0.4083	0.3914 ↓4.14%	0.4163	0.3965 ↓4.76%	0.3955	0.3895 ↓1.52%	0.4020	0.3900 ↓2.99%
ETTm2	96	0.1453	0.1462	0.1518	0.1487	0.1619	0.1522	0.1627	0.1571	0.1421	0.1403	0.1511	0.1462
	192	0.1785	0.1781	0.1789	0.1810	0.1970	0.1877	0.2054	0.1901	0.1743	0.1728	0.1902	0.1814
	336	0.2133	0.2101	0.2346	0.2163	0.2237	0.2163	0.2379	0.2326	0.2060	0.2050	0.2150	0.2093
	720	0.2669	0.2698	0.2758	0.2688	0.2965	0.2830	0.3039	0.2846	0.2621	0.2591	0.2653	0.2629
	Avg.	0.2010	0.2010 ↑0.00%	0.2103	0.2037 ↓3.14%	0.2198	0.2098 ↓4.55%	0.2275	0.2161 ↓5.01%	0.1961	0.1943 ↓0.92%	0.2054	0.1999 ↓2.68%
Weather	96	0.1688	0.1565	0.1741	0.1542	0.1613	0.1586	0.1536	0.1524	0.1434	0.1424	0.1702	0.1596
	192	0.2116	0.2007	0.2139	0.2000	0.2093	0.2055	0.2061	0.2038	0.1863	0.1861	0.2114	0.2047
	336	0.2563	0.2468	0.2585	0.2476	0.2615	0.2564	0.2436	0.2439	0.2378	0.2379	0.2554	0.2524
	720	0.3152	0.3089	0.3220	0.3114	0.3236	0.3189	0.3153	0.3160	0.3174	0.3167	0.3180	0.3149
	Avg.	0.2380	0.2282 ↓4.12%	0.2421	0.2283 ↓5.70%	0.2389	0.2349 ↓1.67%	0.2297	0.2290 ↓0.30%	0.2212	0.2208 ↓0.18%	0.2388	0.2329 ↓2.47%
Exchange	96	0.1559	0.1051	1.3625	0.1450	0.1284	0.1212	0.1470	0.1367	0.0931	0.0963	0.1809	0.1370
	192	0.2700	0.1937	0.3236	0.2057	0.2488	0.2325	0.3369	0.2707	0.1911	0.1917	0.2157	0.2201
	336	0.5119	0.2931	0.5493	0.3217	0.4118	0.3795	0.6113	0.5226	0.3754	0.3622	0.3663	0.3775
	720	1.3767	0.5797	1.6588	0.6779	1.2182	1.0685	1.6860	1.3026	1.2230	1.1019	1.0668	1.0753
	Avg.	0.5786	0.2929 ↓49.38%	0.9735	0.3376 ↓65.32%	0.5018	0.4504 ↓10.24%	0.6953	0.5581 ↓19.73%	0.4706	0.4380 ↓6.93%	0.4574	0.4525 ↓1.07%
Electricity	96	0.1438	0.1314	0.2090	0.1546	0.1383	0.1312	0.1434	0.1317	0.1473	0.1333	0.1560	0.1322
	192	0.1574	0.1470	0.1828	0.1562	0.1565	0.1467	0.1622	0.1488	0.1591	0.1491	0.1596	0.1478
	336	0.1721	0.1616	0.1933	0.1685	0.1728	0.1601	0.1817	0.1656	0.1718	0.1650	0.1785	0.1631
	720	0.7528	0.1980	0.9798	0.2006	0.2116	0.1925	1.1770	0.2007	0.7797	0.2032	1.2438	0.2009
	Avg.	0.3065	0.1595 ↓47.96%	0.3913	0.1700 ↓56.56%	0.1698	0.1576 ↓7.18%	0.4161	0.1617 ↓61.14%	0.3145	0.1626 ↓48.30%	0.4345	0.1610 ↓62.95%
PEMS04	96	0.2203	0.1438	0.2329	0.1447	0.1384	0.1299	0.1399	0.1255	0.1577	0.1363	0.3224	0.1434
	192	0.2304	0.1382	0.2370	0.1385	0.1587	0.1332	0.1569	0.1330	0.2085	0.1456	0.3079	0.1405
	336	0.2512	0.1418	0.2500	0.1414	0.1785	0.1416	0.1634	0.1381	0.2203	0.1484	0.3138	0.1450
	720	0.3192	0.1603	0.3236	0.1609	0.2280	0.1625	0.2050	0.1604	0.1859	0.1606	0.3386	0.1659
	Avg.	0.2553	0.1460 ↓42.81%	0.2609	0.1464 ↓43.89%	0.1759	0.1418 ↓19.39%	0.1663	0.1392 ↓16.30%	0.1931	0.1477 ↓23.51%	0.3207	0.1487 ↓53.63%
PEMS08	96	0.3656	0.2399	0.3198	0.2366	0.2230	0.2194	0.2292	0.2183	0.3255	0.2684	0.4914	0.2502
	192	0.3967	0.2861	0.2498	0.2365	0.2921	0.2908	0.3038	0.2903	0.3751	0.3138	0.4704	0.3161
	336	0.4219	0.3018	0.2595	0.2531	0.3217	0.3069	0.3573	0.3218	0.3987	0.3374	0.4671	0.3265
	720	0.4844	0.3331	0.2737	0.2713	0.4029	0.3575	0.4082	0.3644	0.4071	0.3900	0.5026	0.3671
	Avg.	0.4172	0.2902 ↓30.44%	0.2757	0.2494 ↓9.54%	0.3099	0.2937 ↓5.23%	0.3247	0.2987 ↓8.01%	0.3766	0.3274 ↓13.06%	0.4829	0.3150 ↓34.77%
Solar	96	0.1884	0.1747	0.1866	0.1723	0.1661	0.1678	0.1700	0.1672	0.1736	0.1817	0.1896	0.1742
	192	0.2076	0.1868	0.2083	0.2070	0.1926	0.1846	0.1848	0.1834	0.1940	0.2017	0.2135	0.1869
	336	0.2233	0.1994	0.2207	0.2191	0.2129	0.1958	0.1968	0.1943	0.2084	0.2097	0.2042	0.1970
	720	0.2314	0.2043	0.2059	0.1994	0.2223	0.2037	0.2143	0.2059	0.1937	0.2045	0.2086	0.2043
	Avg.	0.2127	0.1913 ↓10.06%	0.2054	0.1994 ↓2.92%	0.1985	0.1880 ↓5.29%	0.1915	0.1877 ↓1.98%	0.1924	0.1994 ↑3.64%	0.2040	0.1906 ↓6.57%

Table E.2: Complete paired MAE results of CoReG across forecasting backbones and datasets. Lower is better. Green indicates a decrease after adding CoReG, while red indicates an unchanged or increased error. Average is computed over four forecasting horizons, with the relative change shown below each CoReG average.

Dataset	Horizon	DLinear (2023)		MICN (2023)		iTransformer (2024)		PatchMLP (2025)		xPatch (2025)		Amplifier (2025)	
		Original	+CoReG	Original	+CoReG	Original	+CoReG	Original	+CoReG	Original	+CoReG	Original	+CoReG
ETTh1	96	0.4528	0.4518	0.5142	0.4922	0.4710	0.4582	0.4935	0.4651	0.4414	0.4388	0.5436	0.4582
	192	0.4851	0.4809	0.5225	0.5099	0.5091	0.4930	0.5278	0.4996	0.4809	0.4769	0.5549	0.4969
	336	0.5229	0.5149	0.5396	0.5321	0.5356	0.5221	0.5598	0.5285	0.5183	0.5114	0.5735	0.5265
	720	0.5959	0.5864	0.6001	0.5939	0.6194	0.6039	0.6353	0.6164	0.5990	0.5966	0.6164	0.6070
	Avg.	0.5142	0.5085 ↓ 1.11%	0.5441	0.5320 ↓ 2.22%	0.5338	0.5193 ↓ 2.72%	0.5541	0.5274 ↓ 4.82%	0.5099	0.5059 ↓ 0.78%	0.5721	0.5222 ↓ 8.72%
ETTh2	96	0.3273	0.3205	0.3548	0.3237	0.3386	0.3283	0.3616	0.3386	0.3121	0.3090	0.3875	0.3284
	192	0.3742	0.3602	0.3867	0.3608	0.3761	0.3686	0.3974	0.3734	0.3494	0.3469	0.4008	0.3654
	336	0.4072	0.3839	0.4154	0.3888	0.4130	0.3916	0.4477	0.4333	0.3780	0.3739	0.4193	0.3936
	720	0.4867	0.4442	0.4834	0.4432	0.4493	0.4424	0.5088	0.4907	0.4315	0.4263	0.4654	0.4515
	Avg.	0.3989	0.3772 ↓ 5.44%	0.4101	0.3791 ↓ 7.56%	0.3943	0.3827 ↓ 2.94%	0.4289	0.4090 ↓ 4.64%	0.3678	0.3640 ↓ 1.03%	0.4182	0.3847 ↓ 8.01%
ETTm1	96	0.3732	0.3693	0.3860	0.3722	0.3875	0.3783	0.3929	0.3807	0.3655	0.3617	0.3811	0.3731
	192	0.4034	0.3981	0.4085	0.4036	0.4229	0.4115	0.4314	0.4163	0.3989	0.3958	0.4102	0.4053
	336	0.4317	0.4292	0.4351	0.4301	0.4494	0.4388	0.4586	0.4425	0.4318	0.4272	0.4395	0.4344
	720	0.4689	0.4655	0.4725	0.4679	0.4855	0.4780	0.4855	0.4770	0.4693	0.4687	0.4799	0.4744
	Avg.	0.4193	0.4155 ↓ 0.91%	0.4255	0.4185 ↓ 1.65%	0.4363	0.4267 ↓ 2.20%	0.4421	0.4291 ↓ 2.94%	0.4164	0.4134 ↓ 0.72%	0.4277	0.4218 ↓ 1.38%
ETTm2	96	0.2508	0.2503	0.2597	0.2542	0.2686	0.2581	0.2665	0.2606	0.2438	0.2418	0.2582	0.2523
	192	0.2805	0.2785	0.2819	0.2812	0.2972	0.2896	0.3016	0.2874	0.2729	0.2704	0.2893	0.2818
	336	0.3100	0.3050	0.3303	0.3101	0.3178	0.3109	0.3246	0.3204	0.2978	0.2961	0.3110	0.3058
	720	0.3521	0.3491	0.3598	0.3491	0.3684	0.3600	0.3703	0.3596	0.3391	0.3367	0.3487	0.3475
	Avg.	0.2983	0.2957 ↓ 0.87%	0.3079	0.2986 ↓ 3.02%	0.3130	0.3047 ↓ 2.65%	0.3157	0.3070 ↓ 2.76%	0.2884	0.2862 ↓ 0.76%	0.3018	0.2968 ↓ 1.66%
Weather	96	0.2287	0.2112	0.2360	0.2082	0.2171	0.2128	0.2074	0.2056	0.1842	0.1831	0.2258	0.2129
	192	0.2683	0.2545	0.2708	0.2538	0.2563	0.2543	0.2564	0.2545	0.2261	0.2263	0.2598	0.2541
	336	0.3060	0.2916	0.3088	0.2924	0.2969	0.2940	0.2849	0.2858	0.2678	0.2683	0.2927	0.2913
	720	0.3534	0.3416	0.3596	0.3443	0.3448	0.3400	0.3356	0.3361	0.3251	0.3244	0.3377	0.3352
	Avg.	0.2891	0.2747 ↓ 4.98%	0.2938	0.2747 ↓ 6.50%	0.2788	0.2753 ↓ 1.26%	0.2711	0.2705 ↓ 0.22%	0.2508	0.2505 ↓ 0.12%	0.2790	0.2734 ↓ 2.01%
Exchange	96	0.3066	0.2421	0.9857	0.2800	0.2627	0.2509	0.2831	0.2715	0.2177	0.2215	0.3194	0.2731
	192	0.4025	0.3318	0.4594	0.3508	0.3740	0.3546	0.4314	0.3837	0.3158	0.3168	0.3390	0.3421
	336	0.5580	0.4187	0.5888	0.4488	0.4824	0.4570	0.5618	0.5240	0.4497	0.4411	0.4440	0.4519
	720	0.9020	0.6122	1.0215	0.6662	0.8364	0.7842	0.9271	0.8398	0.8417	0.8017	0.7860	0.7831
	Avg.	0.5423	0.4012 ↓ 26.02%	0.7638	0.4364 ↓ 42.86%	0.4888	0.4617 ↓ 5.54%	0.5509	0.5048 ↓ 8.37%	0.4562	0.4453 ↓ 2.39%	0.4721	0.4626 ↓ 2.01%
Electricity	96	0.2470	0.2284	0.3234	0.2665	0.2381	0.2283	0.2472	0.2305	0.2460	0.2280	0.2633	0.2287
	192	0.2600	0.2432	0.2966	0.2641	0.2558	0.2418	0.2662	0.2470	0.2556	0.2424	0.2600	0.2428
	336	0.2758	0.2598	0.3092	0.2799	0.2732	0.2561	0.2870	0.2657	0.2672	0.2588	0.2798	0.2588
	720	0.7087	0.2908	0.7926	0.2912	0.3073	0.2844	0.8486	0.2911	0.7099	0.2907	0.8856	0.2914
	Avg.	0.3729	0.2556 ↓ 31.46%	0.4305	0.2754 ↓ 36.03%	0.2686	0.2526 ↓ 5.96%	0.4123	0.2586 ↓ 37.28%	0.3697	0.2550 ↓ 31.03%	0.4222	0.2554 ↓ 39.51%
PEMS04	96	0.3197	0.2546	0.3277	0.2545	0.2505	0.2402	0.2593	0.2404	0.2522	0.2360	0.4129	0.2524
	192	0.3269	0.2469	0.3334	0.2473	0.2708	0.2427	0.2728	0.2432	0.2937	0.2481	0.4008	0.2482
	336	0.3427	0.2495	0.3434	0.2495	0.2899	0.2509	0.2797	0.2479	0.3011	0.2512	0.4033	0.2520
	720	0.3918	0.2685	0.3964	0.2692	0.3367	0.2689	0.3206	0.2678	0.2788	0.2582	0.4143	0.2719
	Avg.	0.3453	0.2549 ↓ 26.18%	0.3502	0.2551 ↓ 27.16%	0.2870	0.2507 ↓ 12.65%	0.2831	0.2498 ↓ 11.76%	0.2815	0.2484 ↓ 11.76%	0.4078	0.2561 ↓ 37.20%
PEMS08	96	0.3579	0.2700	0.3330	0.2650	0.2607	0.2513	0.2650	0.2512	0.3054	0.2690	0.4312	0.2615
	192	0.3724	0.2759	0.2544	0.2475	0.2869	0.2653	0.2772	0.2613	0.3221	0.2761	0.4005	0.2752
	336	0.3915	0.2811	0.2604	0.2566	0.3115	0.2840	0.3109	0.2792	0.3341	0.2817	0.4063	0.2823
	720	0.4401	0.3078	0.2824	0.2791	0.3655	0.3009	0.3605	0.3042	0.3533	0.3070	0.4561	0.3080
	Avg.	0.3905	0.2837 ↓ 27.35%	0.2826	0.2620 ↓ 7.29%	0.3062	0.2754 ↓ 10.06%	0.3034	0.2740 ↓ 9.69%	0.3287	0.2835 ↓ 13.75%	0.4235	0.2818 ↓ 33.46%
Solar	96	0.2550	0.2220	0.2601	0.2374	0.2360	0.2203	0.2281	0.2206	0.2131	0.2002	0.2557	0.2198
	192	0.2704	0.2298	0.2451	0.2443	0.2607	0.2310	0.2475	0.2383	0.2229	0.2100	0.2714	0.2307
	336	0.2836	0.2397	0.2652	0.2635	0.2762	0.2381	0.2681	0.2595	0.2274	0.2161	0.2576	0.2402
	720	0.2899	0.2464	0.2691	0.2606	0.2875	0.2464	0.2581	0.2481	0.2217	0.2215	0.2632	0.2469
	Avg.	0.2747	0.2345 ↓ 14.63%	0.2599	0.2515 ↓ 3.23%	0.2651	0.2339 ↓ 11.77%	0.2505	0.2416 ↓ 3.55%	0.2213	0.2120 ↓ 4.20%	0.2620	0.2344 ↓ 10.53%

E.2 Complete Ablation Results

This section evaluates the independent capability and complementarity of the two retrieval geometries, the construction of RoleGeo, temporal alignment, and path-specific readout on DLinear. The experiments cover ETTm1, Weather, Exchange, and PEMS08. Each Dataset $\times$ Horizon cell uses the same frozen Anchor checkpoint for all variants; all variants share data splits, forecasting horizons, random seed 2027, history length 720, candidate-retrieval constraints, and evaluation implementation.

Backbone, single-view retrieval, and retrieval-only prediction. Raw Backbone provides the original predictor reference. Retrieval-only CoReG averages the two retrieval predictions before readout. Same-channel Retrieval retrieves from the target variable's own historical memory and fuses the retrieved prediction with the Backbone output.

TimeGeo and RoleGeo. TimeGeo Only uses the multi-scale temporal geometry and one GeoRead, while RoleGeo Only uses the stable relations and time-aligned peer responses defining the role geometry. Together with FULL, these variants present the individual capability and complementarity of the two geometries.

Temporal alignment and stable relations. w/ Shifted Peer Responses replaces time-aligned peer responses with shifted responses, while w/o Stable Relations sets the relation matrix to zero. These variants correspond to the temporal-alignment and stable-relation designs.

Independent readouts. w/ Shared Readout shares readout parameters across the two views, enabling comparison with path-specific readouts.

E.3 Complete Plug-in Comparisons

CoReG provides competitive error reductions in comparisons with external enhancement modules across three Backbones, five datasets, and four forecasting horizons, while the relative strengths of the modules vary with the Backbone, dataset, and horizon. We compare CoReG with four external enhancement modules: the historical-segment retrieval module SARAF (Zhou et al., 2026), the global temporal retrieval module GTR (Cao et al., 2026), an online residual adapter, and the timestamp global-local fusion module GLAFF (Wang et al., 2024). The comparison covers DLinear, iTransformer, and MICN on ETTh1, ETTh2, Exchange, Weather, and PEMS08, with horizons 96, 192, 336, and 720 under seed 2027.

Each module follows its original valid protocol. SARAF constructs and uses retrieval memory once per query; GTR is trained end to end through the host forecasting loss; the online residual adapter follows the official online X+Y additive protocol, producing a forecast before receiving delayed labels to update its corrector; and the joint training of GLAFF, SARAF, and GTR starts from independent initialization. Tables E.4, E.5, and E.6 provide complete results for the three backbones and summarize each module's average relative change from Raw, count of error reductions, and count of lowest-error results.

CoReG maintains the lowest error in most paired settings for DLinear and MICN, and shows stable advantages on ETTh1 and ETTh2 with iTransformer. SARAF, GTR, the online residual adapter, and GLAFF attain lower individual metrics for a small number of Backbone–Dataset combinations, reflecting the distinct operating conditions of external enhancement mechanisms.

Figure E.1 summarizes the accuracy–latency distribution of each plug-in across all matched settings. It compresses the setting-level comparisons in the complete tables into Backbone-specific quantile statistics, allowing typical gains and variation across datasets and horizons to be examined jointly.

Figure E.1 shows that CoReG achieves positive median MSE reductions on DLinear, iTransformer, and MICN while maintaining a stable accuracy–latency trade-off across all three Backbones. The gain and latency distributions of the other plug-ins vary substantially with the host architecture; CoReG therefore exhibits more consistent enhancement across matched settings.

Table E.3: Complete ablation MSE and MAE results for DLinear on ETTm1, Weather, Exchange, and PEMS08. Green indicates an error lower than FULL; red indicates an error equal to or higher than FULL. Avg. is the arithmetic mean across four horizons. For each horizon, the rightmost columns average relative changes from FULL across the four datasets; the Avg. row averages the four dataset-level Avg. changes. Percentages use the unrounded source values.

Variant	Horizon	ETTm1		Weather		Exchange		PEMS08		Avg. Relative Change	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	ΔMSE	ΔMAE
FULL	96	0.3132	0.3693	0.1565	0.2112	0.1051	0.2421	0.2399	0.2700	–	–
	192	0.3563	0.3981	0.2007	0.2545	0.1937	0.3318	0.2861	0.2759	–	–
	336	0.4027	0.4292	0.2468	0.2916	0.2931	0.4187	0.3018	0.2811	–	–
	720	0.4518	0.4655	0.3089	0.3416	0.5797	0.6122	0.3331	0.3078	–	–
	Avg.	0.3810	0.4155	0.2282	0.2747	0.2929	0.4012	0.2902	0.2837	–	–
Raw Backbone	96	0.3274	0.3732	0.1688	0.2287	0.1559	0.3066	0.3656	0.3579	↑ 28.30%	↑ 17.14%
	192	0.3760	0.4034	0.2116	0.2683	0.2700	0.4025	0.3967	0.3724	↑ 22.26%	↑ 15.75%
	336	0.4181	0.4317	0.2563	0.3060	0.5119	0.5580	0.4219	0.3915	↑ 30.54%	↑ 19.51%
	720	0.4639	0.4689	0.3152	0.3534	1.3767	0.9020	0.4844	0.4401	↑ 46.91%	↑ 23.63%
	Avg.	0.3964	0.4193	0.2380	0.2891	0.5786	0.5423	0.4172	0.3905	↑ 37.40%	↑ 19.73%
Retrieval-only CoReG	96	0.4317	0.4634	0.2752	0.3167	0.6092	0.6008	1.1129	0.8234	↑ 239.35%	↑ 107.14%
	192	0.4527	0.4761	0.2907	0.3253	0.7302	0.6617	0.5848	0.5323	↑ 113.33%	↑ 59.94%
	336	0.4798	0.4934	0.3122	0.3378	0.9278	0.7482	0.4046	0.3816	↑ 74.07%	↑ 36.31%
	720	0.5131	0.5169	0.3508	0.3642	1.5694	0.9294	0.4302	0.4023	↑ 56.76%	↑ 25.05%
	Avg.	0.4693	0.4874	0.3072	0.3360	0.9592	0.7350	0.6331	0.5349	↑ 100.86%	↑ 52.84%
Same-channel Retrieval	96	0.3147	0.3699	0.1584	0.2133	0.1142	0.2529	0.2419	0.2719	↑ 2.79%	↑ 1.58%
	192	0.3582	0.4002	0.2054	0.2590	0.1976	0.3337	0.2867	0.2786	↑ 1.26%	↑ 0.96%
	336	0.4045	0.4309	0.2489	0.2930	0.2921	0.4181	0.3070	0.2831	↑ 0.67%	↑ 0.37%
	720	0.4542	0.4668	0.3108	0.3424	0.6116	0.6262	0.3368	0.3096	↑ 1.94%	↑ 0.84%
	Avg.	0.3829	0.4169	0.2308	0.2769	0.3039	0.4077	0.2931	0.2858	↑ 1.59%	↑ 0.88%
TimeGeo Only	96	0.3185	0.3733	0.1608	0.2151	0.1095	0.2467	0.2473	0.2730	↑ 2.92%	↑ 1.48%
	192	0.3607	0.4003	0.2031	0.2568	0.1907	0.3296	0.2939	0.2778	↑ 0.90%	↑ 0.37%
	336	0.4031	0.4289	0.2479	0.2922	0.2963	0.4220	0.3088	0.2839	↑ 0.99%	↑ 0.48%
	720	0.4536	0.4663	0.3069	0.3397	0.5825	0.6130	0.3372	0.3105	↑ 0.36%	↑ 0.16%
	Avg.	0.3840	0.4172	0.2297	0.2760	0.2947	0.4028	0.2968	0.2863	↑ 1.08%	↑ 0.54%
RoleGeo Only	96	0.3153	0.3703	0.1582	0.2134	0.1031	0.2400	0.2388	0.2723	↓ 0.15%	↑ 0.33%
	192	0.3581	0.3997	0.2040	0.2581	0.1996	0.3366	0.2854	0.2793	↑ 1.25%	↑ 1.12%
	336	0.4051	0.4309	0.2505	0.2946	0.3183	0.4277	0.3032	0.2838	↑ 2.79%	↑ 1.13%
	720	0.4535	0.4664	0.3092	0.3440	0.5819	0.6129	0.3394	0.3115	↑ 0.69%	↑ 0.56%
	Avg.	0.3830	0.4168	0.2305	0.2775	0.3007	0.4043	0.2917	0.2867	↑ 1.18%	↑ 0.79%
w/ Shifted Peer Responses	96	0.3125	0.3690	0.1576	0.2128	0.1071	0.2440	0.2413	0.2700	↑ 0.75%	↑ 0.37%
	192	0.3559	0.3978	0.2007	0.2544	0.1933	0.3318	0.2817	0.2753	↓ 0.47%	↓ 0.09%
	336	0.4029	0.4293	0.3097	0.3392	0.2946	0.4200	0.3041	0.2813	↑ 6.70%	↑ 4.18%
	720	0.4516	0.4654	0.3084	0.3409	0.5816	0.6138	0.3326	0.3077	↓ 0.01%	↑ 0.00%
	Avg.	0.3807	0.4154	0.2441	0.2868	0.2942	0.4024	0.2899	0.2836	↑ 1.80%	↑ 1.15%
w/o Stable Relations	96	0.3136	0.3695	0.1583	0.2132	0.1133	0.2518	0.2430	0.2709	↑ 2.59%	↑ 1.34%
	192	0.3569	0.3984	0.2018	0.2554	0.1939	0.3314	0.2889	0.2773	↑ 0.45%	↑ 0.20%
	336	0.4031	0.4295	0.2466	0.2915	0.2941	0.4200	0.3072	0.2829	↑ 0.53%	↑ 0.24%
	720	0.4528	0.4659	0.3085	0.3409	0.5961	0.6193	0.3363	0.3095	↑ 0.97%	↑ 0.39%
	Avg.	0.3816	0.4158	0.2288	0.2752	0.2993	0.4056	0.2938	0.2851	↑ 0.96%	↑ 0.47%
w/ Shared Readout	96	0.3141	0.3699	0.1579	0.2128	0.1051	0.2421	0.2406	0.2709	↑ 0.37%	↑ 0.32%
	192	0.3570	0.3985	0.2016	0.2556	0.1942	0.3325	0.2877	0.2765	↑ 0.36%	↑ 0.23%
	336	0.4029	0.4294	0.2473	0.2922	0.2932	0.4188	0.3022	0.2811	↑ 0.11%	↑ 0.06%
	720	0.4519	0.4655	0.3081	0.3414	0.5806	0.6126	0.3332	0.3079	↓ 0.01%	↑ 0.01%
	Avg.	0.3815	0.4158	0.2287	0.2755	0.2933	0.4015	0.2909	0.2841	↑ 0.18%	↑ 0.14%

Table E.4: Detailed plug-and-play comparison on DLinear. Each entry reports MSE and MAE at the corresponding forecasting horizon. Avg. is the arithmetic mean across four horizons, with the relative change from Raw shown below each plug-in value. The summary rows average relative gains from Raw, count settings with lower error than Raw, and count the lowest-error results among all methods; each count is out of 20 dataset–horizon settings and is computed separately for MSE and MAE using the original precision. Green indicates lower error than Raw, while red indicates error equal to or higher than Raw. In the detailed rows, the lowest and second-lowest errors are bold and underlined; in the summary rows, the largest and second-largest statistics are bold and underlined.

Dataset	Horizon	Raw		CoReG		SARAF (2026)		GTR (2026)		δ-Adapter (2026)		GLAFF (2024)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4302	0.4528	**0.4223**	0.4518	0.4417	0.4650	0.4307	0.4534	0.4400	0.4584	0.4302	**0.4515**
	192	0.4711	0.4851	**0.4628**	**0.4809**	0.4787	0.4924	0.4833	0.4935	0.4828	0.4900	0.4795	0.4892
	336	0.5202	0.5229	**0.5067**	**0.5149**	0.5208	0.5217	0.5335	0.5296	0.5213	0.5179	0.5197	0.5191
	720	0.6278	0.5959	**0.6163**	**0.5864**	0.6298	0.5962	0.6258	0.5907	0.6244	0.5888	0.6298	0.5973
	Avg.	0.5123	0.5142	**0.5020** ↓2.01%	**0.5085** ↓1.11%	0.5177 ↑1.05%	0.5188 ↑0.89%	0.5183 ↑1.17%	0.5168 ↑0.51%	0.5171 ↑0.94%	0.5138 ↓0.08%	0.5148 ↑0.49%	0.5143 ↑0.02%
ETTh2	96	0.2275	0.3273	0.2232	**0.3205**	0.2246	0.3216	**0.2226**	0.3232	0.2306	0.3300	0.2389	0.3322
	192	0.2867	0.3742	**0.2723**	0.3602	0.2737	**0.3556**	0.2805	0.3696	0.2895	0.3708	0.2909	0.3756
	336	0.3261	0.4072	**0.2969**	**0.3839**	0.3161	0.3856	0.3277	0.4087	0.3476	0.4140	0.3309	0.4095
	720	0.4522	0.4867	**0.3804**	**0.4442**	0.4123	0.4504	0.4323	0.4763	0.4522	0.4860	0.4304	0.4738
	Avg.	0.3231	0.3989	**0.2932** ↓9.25%	**0.3772** ↓5.44%	0.3067 ↓5.08%	0.3783 ↓5.16%	0.3158 ↓2.26%	0.3944 ↓1.13%	0.3300 ↑2.14%	0.4002 ↑0.33%	0.3228 ↓0.09%	0.3978 ↓0.28%
Exchange	96	0.1559	0.3066	0.1051	0.2421	**0.0872**	**0.2052**	0.1306	0.2817	0.1750	0.3204	0.1673	0.3217
	192	0.2700	0.4025	0.1937	0.3318	**0.1895**	**0.3098**	0.2473	0.3898	0.2946	0.4182	0.3037	0.4272
	336	0.5119	0.5580	**0.2931**	**0.4187**	0.3396	0.4223	0.4740	0.5557	0.4896	0.5423	0.5071	0.5393
	720	1.3767	0.9020	**0.5797**	**0.6122**	0.8711	0.6950	1.2948	0.8697	1.3102	0.8635	1.0110	0.7420
	Avg.	0.5786	0.5423	**0.2929** ↓49.38%	**0.4012** ↓26.02%	0.3719 ↓35.72%	0.4081 ↓24.75%	0.5367 ↓7.24%	0.5242 ↓3.34%	0.5673 ↓1.95%	0.5361 ↓1.14%	0.4973 ↓14.05%	0.5075 ↓6.42%
Weather	96	0.1688	0.2287	**0.1565**	**0.2112**	0.1702	0.2333	0.1706	0.2330	0.1652	0.2298	0.1583	0.2115
	192	0.2116	0.2683	**0.2007**	**0.2545**	0.2139	0.2709	0.2124	0.2689	0.2096	0.2702	0.2034	0.2562
	336	0.2563	0.3060	0.2468	**0.2916**	0.2622	0.3098	0.2578	0.3071	**0.2461**	0.2998	0.2486	0.2936
	720	0.3152	0.3534	0.3089	**0.3416**	0.3294	0.3605	0.3226	0.3636	**0.3059**	0.3450	0.3124	0.3473
	Avg.	0.2380	0.2891	**0.2282** ↓4.12%	**0.2747** ↓4.98%	0.2439 ↑2.48%	0.2936 ↑1.56%	0.2409 ↑1.22%	0.2931 ↑1.38%	0.2317 ↓2.65%	0.2862 ↓1.00%	0.2307 ↓3.07%	0.2772 ↓4.12%
PEMS08	96	0.3656	0.3579	**0.2399**	**0.2700**	0.2688	0.3087	0.3597	0.3519	0.3014	0.2954	0.3629	0.3553
	192	0.3967	0.3724	**0.2861**	**0.2759**	0.3020	0.3171	0.3932	0.3653	0.3371	0.3167	0.3983	0.3709
	336	0.4219	0.3915	**0.3018**	**0.2811**	0.3061	0.3120	0.4284	0.4008	0.3679	0.3432	0.4194	0.3935
	720	0.4844	0.4401	0.3331	**0.3078**	**0.3232**	0.3159	0.4832	0.4364	0.4255	0.3895	0.4841	0.4406
	Avg.	0.4172	0.3905	**0.2902** ↓30.44%	**0.2837** ↓27.35%	0.3000 ↓28.09%	0.3134 ↓19.74%	0.4161 ↓0.26%	0.3886 ↓0.49%	0.3579 ↓14.21%	0.3362 ↓13.91%	0.4162 ↓0.24%	0.3901 ↓0.10%
Sum.	Mean relative gain	—	—	**+17.07%**	**+12.49%**	+12.98%	+9.49%	+1.90%	+0.70%	+2.10%	+2.83%	+0.88%	+1.42%
	Improved cases	—	—	**20/20**	**20/20**	12/20	13/20	11/20	11/20	11/20	12/20	11/20	11/20
	Best cases	0/20	0/20	**14/20**	**16/20**	3/20	3/20	1/20	0/20	2/20	0/20	0/20	1/20

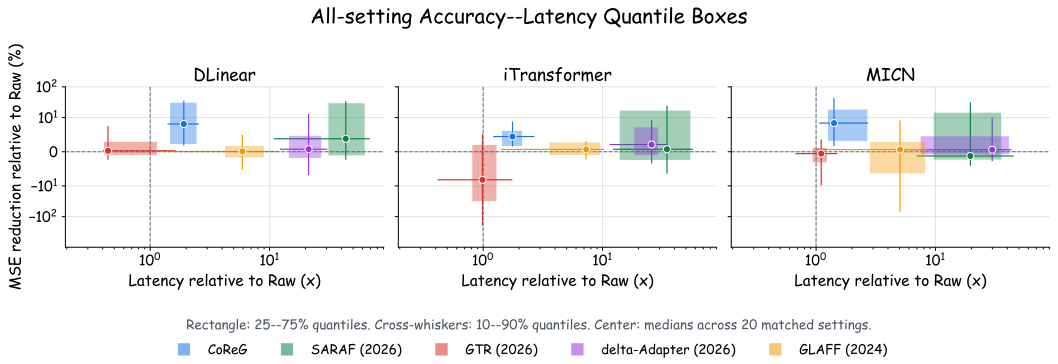


Figure E.1: All-setting accuracy–latency quantile boxes for plug-in enhancement modules. The panels correspond to DLinear, iTransformer, and MICN. The horizontal axis is the end-to-end inference latency relative to the matched Raw Backbone, and the vertical axis is the MSE reduction relative to Raw. Each rectangle spans the 25th–75th percentiles across 20 matched dataset–horizon settings, cross-whiskers span the 10th–90th percentiles, and the center dot is the median. The vertical and horizontal zero lines denote the Raw MSE and latency reference, respectively.

Table E.5: Detailed plug-and-play comparison on iTransformer. Each entry reports MSE and MAE at the corresponding forecasting horizon. Avg. is the arithmetic mean across four horizons, with the relative change from Raw shown below each plug-in value. The summary rows average relative gains from Raw, count settings with lower error than Raw, and count the lowest-error results among all methods; each count is out of 20 dataset–horizon settings and is computed separately for MSE and MAE using the original precision. Green indicates lower error than Raw, while red indicates error equal to or higher than Raw. In the detailed rows, the lowest and second-lowest errors are bold and underlined; in the summary rows, the largest and second-largest statistics are bold and underlined.

Dataset	Horizon	Raw		CoReG		SARAF (2026)		GTR (2026)		δ-Adapter (2026)		GLAFF (2024)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4421	0.4710	**0.4229**	**0.4582**	0.4720	0.4884	0.5234	0.5429	0.4526	0.4785	0.4618	0.4821
	192	0.4995	0.5091	**0.4710**	**0.4930**	0.5438	0.5284	0.5377	0.5461	0.5137	0.5190	0.5052	0.5126
	336	0.5392	0.5356	**0.5146**	**0.5221**	0.5364	0.5333	0.6059	0.5675	0.5472	0.5421	0.5422	0.5372
	720	0.6745	0.6194	**0.6488**	**0.6039**	0.6728	0.6143	0.6922	0.6314	0.6572	0.6114	0.6838	0.6209
	Avg.	0.5388	0.5338	**0.5143** ↓4.55%	**0.5193** ↓2.72%	0.5563 ↑3.25%	0.5411 ↑1.37%	0.5898 ↑9.47%	0.5720 ↑7.16%	0.5427 ↑0.72%	0.5377 ↑0.73%	0.5483 ↑1.76%	0.5382 ↑0.82%
ETTh2	96	0.2426	0.3386	**0.2287**	**0.3283**	0.2581	0.3478	0.6341	0.5890	0.2785	0.3664	0.2442	0.3395
	192	0.2943	0.3761	**0.2855**	**0.3686**	0.2933	0.3767	0.4790	0.5043	0.3216	0.3956	0.3078	0.3834
	336	0.3560	0.4130	**0.3256**	**0.3916**	0.3298	0.4017	0.4972	0.5214	0.3510	0.4099	0.3393	0.4040
	720	0.4183	0.4493	0.4077	0.4424	0.4144	0.4527	0.4975	0.5001	**0.3949**	**0.4367**	0.4141	0.4500
	Avg.	0.3278	0.3943	**0.3119** ↓4.85%	**0.3827** ↓2.94%	0.3239 ↓1.19%	0.3947 ↑0.10%	0.5270 ↑60.77%	0.5287 ↑34.09%	0.3365 ↑2.65%	0.4021 ↑1.98%	0.3263 ↓0.46%	0.3942 ↓0.03%
Exchange	96	0.1284	0.2627	0.1212	0.2509	**0.0869**	**0.2074**	1.3031	0.9727	0.1254	0.2625	0.1244	0.2580
	192	0.2488	0.3740	0.2325	0.3546	**0.1844**	**0.3070**	0.2498	0.3717	0.2155	0.3526	0.2449	0.3691
	336	0.4118	0.4824	0.3795	0.4570	**0.3375**	**0.4225**	2.3541	1.3170	0.4142	0.4798	0.4013	0.4746
	720	1.2182	0.8364	**1.0685**	**0.7842**	1.1426	0.7973	1.3251	0.8777	1.1366	0.8067	1.2052	0.8323
	Avg.	0.5018	0.4888	0.4504 ↓10.24%	0.4617 ↓5.54%	**0.4379** ↓12.73%	**0.4335** ↓11.31%	1.3080 ↑160.66%	0.8848 ↑81.01%	0.4729 ↓5.76%	0.4754 ↓2.74%	0.4940 ↓1.55%	0.4835 ↓1.08%
Weather	96	0.1613	0.2171	0.1586	**0.2128**	0.1709	0.2320	**0.1531**	0.2169	0.1611	0.2227	0.1610	0.2133
	192	0.2093	0.2563	0.2055	0.2543	0.2125	0.2702	**0.1979**	0.2568	0.2053	0.2636	0.2030	**0.2497**
	336	0.2615	0.2969	0.2564	0.2940	0.2649	0.3146	0.2765	0.3225	0.2545	0.3013	**0.2530**	**0.2908**
	720	0.3236	0.3448	0.3189	0.3400	0.3374	0.3666	**0.3015**	**0.3376**	0.3177	0.3480	0.3300	0.3408
	Avg.	0.2389	0.2788	0.2349 ↓1.67%	0.2753 ↓1.26%	0.2464 ↑3.14%	0.2959 ↑6.13%	**0.2323** ↓2.76%	0.2834 ↑1.65%	0.2346 ↓1.80%	0.2839 ↑1.83%	0.2367 ↓0.92%	**0.2736** ↓1.87%
PEMS08	96	0.2230	0.2607	0.2194	0.2513	**0.1886**	0.2372	0.2190	0.2596	0.2022	**0.2371**	0.2230	0.2609
	192	0.2921	0.2869	0.2908	0.2653	**0.2394**	**0.2551**	0.3693	0.3210	0.2695	0.2605	0.2906	0.2854
	336	0.3217	0.3115	0.3069	0.2840	**0.2718**	**0.2721**	0.3059	0.3079	0.2967	0.2855	0.3137	0.3071
	720	0.4029	0.3655	0.3575	0.3009	**0.3038**	**0.2935**	0.3946	0.3649	0.3645	0.3326	0.3994	0.3633
	Avg.	0.3099	0.3062	0.2937 ↓5.23%	0.2754 ↓10.06%	**0.2509** ↓19.04%	**0.2645** ↓13.62%	0.3222 ↑3.97%	0.3133 ↑2.32%	0.2832 ↓8.62%	0.2789 ↓8.92%	0.3067 ↓1.03%	0.3042 ↓0.65%
Sum.	Mean relative gain	—	—	+4.77%	**+4.33%**	**+6.52%**	+3.84%	-86.28%	-32.02%	+2.08%	+1.22%	+0.48%	+0.57%
	Improved cases	—	—	**20/20**	**20/20**	13/20	11/20	6/20	6/20	14/20	11/20	12/20	12/20
	Best cases	0/20	0/20	**8/20**	**9/20**	7/20	6/20	3/20	1/20	1/20	2/20	1/20	2/20

Table E.6: Detailed plug-and-play comparison on MICN. Each entry reports MSE and MAE at the corresponding forecasting horizon. Avg. is the arithmetic mean across four horizons, with the relative change from Raw shown below each plug-in value. The summary rows average relative gains from Raw, count settings with lower error than Raw, and count the lowest-error results among all methods; each count is out of 20 dataset–horizon settings and is computed separately for MSE and MAE using the original precision. Green indicates lower error than Raw, while red indicates error equal to or higher than Raw. In the detailed rows, the lowest and second-lowest errors are bold and underlined; in the summary rows, the largest and second-largest statistics are bold and underlined.

Dataset	Horizon	Raw		CoReG		SARAF (2026)		GTR (2026)		δ-Adapter (2026)		GLAFF (2024)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4895	0.5142	**0.4606**	**0.4922**	0.4957	0.5074	0.5028	0.5294	0.5010	0.5150	0.6285	0.5945
	192	0.5041	0.5225	**0.4869**	**0.5099**	0.5185	0.5226	0.5128	0.5295	0.5178	0.5258	0.4942	0.5126
	336	0.5273	0.5396	**0.5183**	**0.5321**	0.5360	0.5380	0.5348	0.5451	0.5370	0.5397	0.5224	0.5359
	720	0.6141	0.6001	**0.6072**	**0.5939**	0.6232	0.5953	0.6146	0.6010	0.6148	0.5940	0.6309	0.6090
	Avg.	0.5337	0.5441	**0.5182** ↓ 2.90%	**0.5320** ↓ 2.22%	0.5433 ↑ 1.80%	0.5408 ↓ 0.61%	0.5412 ↑ 1.41%	0.5512 ↑ 1.30%	0.5427 ↑ 1.69%	0.5436 ↓ 0.09%	0.5690 ↑ 6.61%	0.5630 ↑ 3.47%
ETTh2	96	0.2597	0.3548	**0.2249**	**0.3237**	0.2373	0.3347	0.2680	0.3694	0.2601	0.3535	0.6019	0.5485
	192	0.3020	0.3867	**0.2701**	**0.3608**	0.2902	0.3713	0.3172	0.4089	0.3123	0.3889	0.3413	0.4164
	336	0.3361	0.4154	**0.3015**	**0.3888**	0.3281	0.3957	0.3320	0.4120	0.3539	0.4214	0.3270	0.4087
	720	0.4414	0.4834	**0.3772**	**0.4432**	0.4277	0.4603	0.4320	0.4789	0.4404	0.4820	0.4379	0.4839
	Avg.	0.3348	0.4101	**0.2934** ↓ 12.37%	**0.3791** ↓ 7.56%	0.3208 ↓ 4.18%	0.3905 ↓ 4.78%	0.3373 ↑ 0.75%	0.4173 ↑ 1.76%	0.3417 ↑ 2.06%	0.4114 ↑ 0.32%	0.4270 ↑ 27.54%	0.4644 ↑ 13.24%
Exchange	96	1.3625	0.9857	0.1450	0.2800	**0.0888**	**0.2069**	0.3491	0.4428	1.1571	0.8871	2.2098	1.1999
	192	0.3236	0.4594	**0.2057**	0.3508	0.2067	**0.3220**	0.3514	0.4575	0.3098	0.4402	1.2756	0.8891
	336	0.5493	0.5888	**0.3217**	**0.4488**	0.3891	0.4510	0.6668	0.6471	0.4975	0.5636	0.5475	0.5776
	720	1.6588	1.0215	**0.6779**	**0.6662**	1.1452	0.8169	2.7541	1.2719	1.5326	0.9554	1.1378	0.7914
	Avg.	0.9735	0.7638	**0.3376** ↓ 65.32%	**0.4364** ↓ 42.86%	0.4574 ↓ 53.01%	0.4492 ↓ 41.19%	1.0303 ↑ 5.83%	0.7048 ↓ 7.72%	0.8743 ↓ 10.19%	0.7116 ↓ 6.83%	1.2927 ↑ 32.79%	0.8645 ↑ 13.18%
Weather	96	0.1741	0.2360	**0.1542**	**0.2082**	0.1762	0.2410	0.1725	0.2335	0.1680	0.2336	0.1611	0.2228
	192	0.2139	0.2708	**0.2000**	**0.2538**	0.2168	0.2744	0.2128	0.2688	0.2075	0.2684	0.2034	0.2610
	336	0.2585	0.3088	**0.2476**	**0.2924**	0.2633	0.3111	0.2577	0.3066	0.2479	0.3008	0.2517	0.3037
	720	0.3220	0.3596	0.3114	**0.3443**	0.3403	0.3734	0.3312	0.3693	**0.3070**	0.3467	0.3147	0.3497
	Avg.	0.2421	0.2938	**0.2283** ↓ 5.70%	**0.2747** ↓ 6.50%	0.2491 ↑ 2.89%	0.3000 ↑ 2.11%	0.2435 ↑ 0.58%	0.2946 ↑ 0.27%	0.2326 ↓ 3.92%	0.2874 ↓ 2.18%	0.2327 ↓ 3.88%	0.2843 ↓ 3.23%
PEMS08	96	0.3198	0.3330	0.2366	0.2650	**0.2244**	0.2644	0.2648	0.2848	0.2357	**0.2528**	0.2411	0.2712
	192	0.2498	0.2544	**0.2365**	**0.2475**	0.2642	0.2780	0.2523	0.2626	0.2474	0.2502	0.2486	0.2625
	336	0.2595	0.2604	**0.2531**	**0.2566**	0.2689	0.2709	0.2563	0.2608	0.2591	0.2601	0.2694	0.2710
	720	0.2737	0.2824	**0.2713**	**0.2791**	0.2846	0.2865	0.2734	0.2816	0.2736	0.2828	0.2794	0.2851
	Avg.	0.2757	0.2826	**0.2494** ↓ 9.54%	0.2620 ↓ 7.29%	0.2605 ↓ 5.51%	0.2750 ↓ 2.69%	0.2617 ↓ 5.08%	0.2725 ↓ 3.57%	0.2540 ↓ 7.87%	**0.2615** ↓ 7.47%	0.2596 ↓ 5.84%	0.2725 ↓ 3.57%
Sum.	Mean relative gain	—	—	**+17.37%**	**+12.32%**	+10.34%	+8.59%	-0.80%	+0.93%	+3.17%	+2.88%	-22.88%	-7.11%
	Improved cases	—	—	**20/20**	**20/20**	9/20	12/20	9/20	9/20	13/20	14/20	12/20	10/20
	Best cases	0/20	0/20	**17/20**	**17/20**	2/20	2/20	0/20	0/20	1/20	1/20	0/20	0/20

Figure E.2 further presents the relationship among accuracy, inference latency, and additional GPU memory in three representative configurations. Each case retains the measured results of every module, showing their resource–accuracy locations for the same forecasting task.

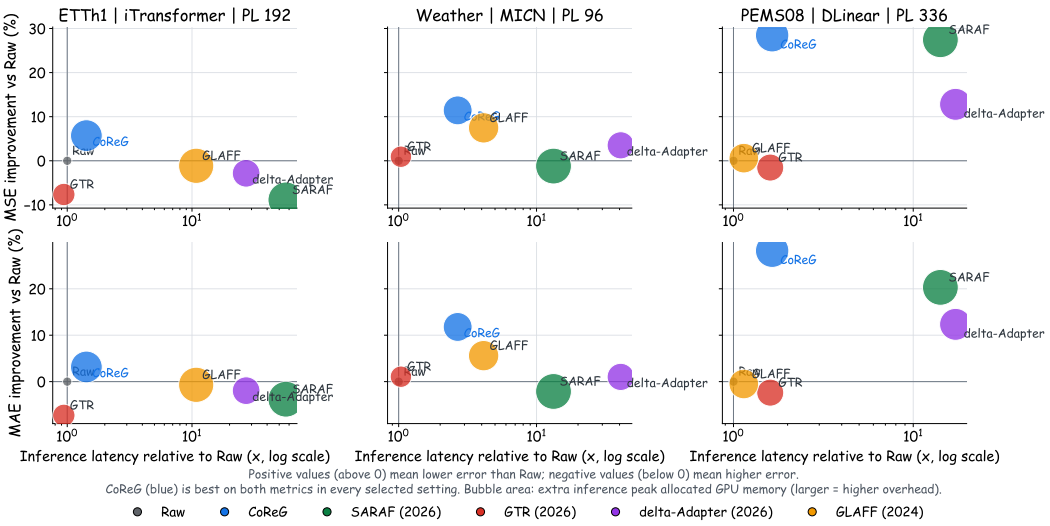


Figure E.2: Accuracy–resource comparison in three representative forecasting settings. Columns correspond to iTransformer on ETTh1, MICN on Weather, and DLinear on PEMS08; rows report MSE and MAE improvement relative to Raw. The horizontal axis is the end-to-end inference latency relative to Raw on a logarithmic scale, and the vertical axis is the relative error improvement, where positive values indicate lower error. Bubble area denotes the additional inference peak allocated GPU memory relative to the matched Raw setting, and color identifies the plug-in module.

In these cases, CoReG achieves the largest improvements in both MSE and MAE while maintaining smaller additional inference-memory bubbles. This figure complements the distributional evidence in Figure E.1 by showing the concrete accuracy, latency, and memory trade-offs within individual forecasting tasks.

E.4 Complete CoMerge Tables

Table E.7: Post-hoc sensitivity of final-view merging rules, DLinear panel. Datasets form the primary row groups and each group covers horizons 96 and 720. FixedMean reports raw MSE and MAE. The remaining entries report the relative error change (%) from FixedMean for the same dataset–horizon setting; negative and positive values indicate lower and higher errors, respectively. Metrics retain four decimal places to expose the small differences in this sensitivity analysis.

Dataset	Horizon	FixedMean		GlobalScalar Δ (%)		ChannelGate Δ (%)		HorizonGate Δ (%)		EvidenceGate Δ (%)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4223	0.4518	-0.0002	-0.0004	-0.0007	-0.0007	-0.0002	-0.0002	+0.0038	-0.0024
	720	0.6163	0.5864	+0.0002	+0.0002	-0.0016	-0.0014	+0.0000	+0.0002	+0.0018	+0.0022
ETTh2	96	0.2232	0.3205	+0.0013	+0.0006	-0.0004	-0.0016	+0.0031	+0.0012	-0.0515	-0.0072
	720	0.3804	0.4442	-0.0050	-0.0018	-0.0084	-0.0032	+0.0011	+0.0005	-0.3460	-0.1661
ETTm1	96	0.3132	0.3693	-0.0112	-0.0092	-0.0160	-0.0133	-0.0067	-0.0065	+0.6115	+0.3246
	720	0.4518	0.4655	+0.0000	+0.0000	-0.0071	-0.0045	-0.0002	+0.0004	+0.0334	+0.0350
ETTm2	96	0.1462	0.2503	-0.0068	-0.0040	-0.0212	-0.0128	-0.0055	-0.0040	+0.4748	+0.2397
	720	0.2698	0.3491	-0.0508	-0.0307	+0.0534	+0.0375	+0.0122	+0.0077	+2.2210	+0.8626
Weather	96	0.1565	0.2112	-0.0115	-0.0062	-0.1597	-0.1406	-0.0121	-0.0047	-0.1156	-0.2865
	720	0.3089	0.3416	+0.0110	+0.0091	+0.0084	-0.0044	+0.0087	+0.0070	+0.6532	+0.1074
Exchange	96	0.1051	0.2421	+0.0114	+0.0050	+0.0086	+0.0033	+0.0114	+0.0050	+0.0990	+0.0438
	720	0.5797	0.6122	+0.0002	+0.0000	-0.0033	-0.0018	+0.0047	+0.0018	-0.0040	-0.0025

Table E.8: Post-hoc sensitivity of final-view merging rules, PatchMLP panel. Datasets form the primary row groups and each group covers horizons 96 and 720. FixedMean reports raw MSE and MAE. The remaining entries report the relative error change (%) from FixedMean for the same dataset–horizon setting; negative and positive values indicate lower and higher errors, respectively.

Dataset	Horizon	FixedMean		GlobalScalar Δ (%)		ChannelGate Δ (%)		HorizonGate Δ (%)		EvidenceGate Δ (%)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4356	0.4651	-0.0021	-0.0009	-0.0023	-0.0015	+0.0000	+0.0002	-0.0103	-0.0054
	720	0.6739	0.6164	-0.0030	-0.0018	-0.0105	-0.0067	-0.0018	-0.0011	-0.0513	-0.0292
ETTh2	96	0.2464	0.3386	-0.0162	-0.0127	-0.0081	-0.0077	-0.0114	-0.0083	-0.2411	-0.1852
	720	0.5000	0.4907	+0.0074	+0.0037	+0.0026	+0.0016	+0.0054	+0.0024	+0.2618	+0.1329
ETTm1	96	0.3234	0.3807	+0.0003	-0.0003	-0.0022	-0.0076	+0.0012	+0.0000	+0.2733	+0.0599
	720	0.4647	0.4770	-0.0047	-0.0027	-0.0073	-0.0046	-0.0028	-0.0019	-0.0439	-0.0168
ETTm2	96	0.1571	0.2606	+0.0153	+0.0073	+0.0096	+0.0042	+0.0096	+0.0038	+0.0573	-0.0115
	720	0.2846	0.3596	+0.0583	+0.0072	+0.0534	+0.0100	+0.0492	+0.0117	+0.2576	+0.0834
Weather	96	0.1524	0.2056	-0.0066	+0.0005	-0.0079	-0.0029	-0.0053	+0.0015	+0.0998	+0.1386
	720	0.3160	0.3361	-0.0025	-0.0015	+0.0152	+0.0074	-0.0006	+0.0012	+0.6807	+0.3779
Exchange	96	0.1367	0.2715	+0.0007	+0.0004	+0.0007	+0.0000	+0.0015	+0.0007	+0.0044	+0.0015
	720	1.3026	0.8398	+0.0097	+0.0051	+0.0090	+0.0038	+0.0093	+0.0045	+0.1874	+0.0992

Table E.9: Post-hoc sensitivity of final-view merging rules, iTransformer panel. Datasets form the primary row groups and each group covers horizons 96 and 720. FixedMean reports raw MSE and MAE. The remaining entries report the relative error change (%) from FixedMean for the same dataset–horizon setting; negative and positive values indicate lower and higher errors, respectively.

Dataset	Horizon	FixedMean		GlobalScalar Δ (%)		ChannelGate Δ (%)		HorizonGate Δ (%)		EvidenceGate Δ (%)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4229	0.4582	-0.0005	-0.0009	-0.0026	-0.0017	-0.0002	-0.0004	-0.0104	-0.0164
	720	0.6488	0.6039	+0.0025	+0.0031	+0.0000	+0.0003	+0.0008	+0.0012	+0.0148	+0.0230
ETTh2	96	0.2287	0.3283	-0.0013	-0.0027	-0.0061	-0.0034	-0.0013	-0.0021	-0.0009	-0.0134
	720	0.4077	0.4424	+0.0012	+0.0007	+0.0005	-0.0002	+0.0022	+0.0014	+0.0213	+0.0142
ETTm1	96	0.3193	0.3783	-0.0050	-0.0048	-0.0348	-0.0307	+0.0009	-0.0037	+0.2602	+0.0756
	720	0.4667	0.4780	-0.0004	+0.0000	-0.0004	-0.0015	-0.0004	+0.0000	+0.1307	+0.0759
ETTm2	96	0.1522	0.2581	+0.0269	+0.0058	+0.0164	-0.0012	+0.0164	+0.0031	+0.2943	+0.0895
	720	0.2830	0.3600	-0.0025	-0.0025	-0.0032	-0.0014	-0.0004	-0.0019	-0.2325	-0.1189
Weather	96	0.1586	0.2128	+0.0032	+0.0014	-0.0038	-0.0056	+0.0025	+0.0005	+0.0013	-0.0028
	720	0.3189	0.3400	+0.0044	+0.0024	+0.0000	-0.0026	+0.0009	-0.0021	+0.6271	+0.2215
Exchange	96	0.1212	0.2509	+0.0058	+0.0024	+0.0025	+0.0012	+0.0050	+0.0024	+0.0710	+0.0259
	720	1.0685	0.7842	-0.0001	+0.0013	+0.0004	+0.0003	+0.0073	+0.0031	-0.0474	-0.0057

Table E.10: Post-hoc sensitivity of final-view merging rules, MICN panel. Datasets form the primary row groups and each group covers horizons 96 and 720. FixedMean reports raw MSE and MAE. The remaining entries report the relative error change (%) from FixedMean for the same dataset–horizon setting; negative and positive values indicate lower and higher errors, respectively.

Dataset	Horizon	FixedMean		GlobalScalar Δ (%)		ChannelGate Δ (%)		HorizonGate Δ (%)		EvidenceGate Δ (%)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4606	0.4922	+0.0000	+0.0000	-0.0020	-0.0010	-0.0002	-0.0002	-0.0382	-0.0341
	720	0.6072	0.5939	-0.0023	-0.0022	+0.0013	+0.0002	-0.0018	-0.0010	-0.0601	-0.0623
ETTh2	96	0.2249	0.3237	+0.0031	+0.0022	-0.0151	-0.0188	+0.0040	+0.0019	-0.1334	-0.0646
	720	0.3772	0.4432	-0.0005	+0.0009	-0.0005	+0.0014	+0.0003	+0.0002	-0.2336	-0.1126
ETTm1	96	0.3164	0.3722	-0.0013	-0.0011	-0.0149	-0.0134	-0.0082	-0.0089	+0.6423	+0.3380
	720	0.4528	0.4679	+0.0002	+0.0002	-0.0038	-0.0024	-0.0013	+0.0009	+0.1495	+0.0964
ETTm2	96	0.1487	0.2542	-0.0013	-0.0008	-0.0034	-0.0043	-0.0047	-0.0055	+0.1903	+0.0905
	720	0.2688	0.3491	-0.0268	-0.0304	-0.0156	-0.0349	-0.0097	-0.0097	+1.3774	+0.6456
Weather	96	0.1542	0.2082	-0.0246	-0.0159	-0.2224	-0.2013	-0.0292	-0.0144	-0.3696	-0.3641
	720	0.3114	0.3443	+0.0119	+0.0081	+0.0090	+0.0020	+0.0119	+0.0076	+0.2415	-0.1322
Exchange	96	0.1450	0.2800	-0.0324	-0.0200	+0.0021	+0.0004	-0.0062	-0.0064	-0.3201	-0.1975
	720	0.6779	0.6662	+0.0040	+0.0012	-0.0055	-0.0020	+0.0091	+0.0032	+0.0674	+0.0248

E.5 Complete Multi-Seed Tables

Table E.11: DLinear results across five random seeds. Each cell reports the five-seed mean on the first line and sample standard deviation on the second line; MSE and MAE are separated.

Dataset	Horizon	Raw		CoReG		SARAF (2026)		GTR (2026)		δ-Adapter (2026)		GLAFF (2024)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4302 ±0.0002	0.4528 ±0.0004	0.4223 ±0.0005	0.4512 ±0.0005	0.4414 ±0.0003	0.4649 ±0.0002	0.4345 ±0.0033	0.4547 ±0.0028	0.4404 ±0.0008	0.4588 ±0.0006	0.4327 ±0.0030	0.4531 ±0.0019
	192	0.4716 ±0.0008	0.4856 ±0.0007	0.4638 ±0.0006	0.4820 ±0.0007	0.4793 ±0.0005	0.4929 ±0.0003	0.4775 ±0.0077	0.4884 ±0.0054	0.4839 ±0.0010	0.4908 ±0.0007	0.4740 ±0.0036	0.4860 ±0.0027
	336	0.5174 ±0.0022	0.5201 ±0.0021	0.5073 ±0.0012	0.5142 ±0.0010	0.5210 ±0.0003	0.5219 ±0.0002	0.5222 ±0.0070	0.5213 ±0.0050	0.5230 ±0.0019	0.5191 ±0.0016	0.5218 ±0.0060	0.5208 ±0.0027
	720	0.6280 ±0.0010	0.5963 ±0.0012	0.6178 ±0.0028	0.5875 ±0.0015	0.6298 ±0.0001	0.5962 ±0.0001	0.6336 ±0.0101	0.5944 ±0.0046	0.6250 ±0.0011	0.5887 ±0.0008	0.6307 ±0.0029	0.5964 ±0.0024
ETTh2	96	0.2286 ±0.0015	0.3280 ±0.0014	0.2245 ±0.0015	0.3211 ±0.0008	0.2242 ±0.0002	0.3212 ±0.0003	0.2256 ±0.0023	0.3260 ±0.0024	0.2318 ±0.0015	0.3307 ±0.0012	0.2313 ±0.0047	0.3276 ±0.0028
	192	0.2830 ±0.0025	0.3712 ±0.0022	0.2712 ±0.0020	0.3587 ±0.0013	0.2737 ±0.0001	0.3557 ±0.0001	0.2802 ±0.0027	0.3699 ±0.0021	0.2901 ±0.0011	0.3714 ±0.0010	0.2898 ±0.0046	0.3749 ±0.0035
	336	0.3285 ±0.0074	0.4082 ±0.0050	0.2979 ±0.0009	0.3854 ±0.0011	0.3165 ±0.0006	0.3858 ±0.0005	0.3247 ±0.0019	0.4064 ±0.0015	0.3485 ±0.0078	0.4145 ±0.0056	0.3277 ±0.0032	0.4070 ±0.0018
	720	0.4532 ±0.0081	0.4871 ±0.0042	0.3799 ±0.0005	0.4436 ±0.0008	0.4131 ±0.0006	0.4506 ±0.0002	0.4396 ±0.0064	0.4808 ±0.0037	0.4537 ±0.0101	0.4869 ±0.0053	0.4737 ±0.0354	0.4972 ±0.0189

Table E.12: iTransformer (2024) results across five random seeds. Each cell reports the five-seed mean on the first line and sample standard deviation on the second line; MSE and MAE are separated.

Dataset	Horizon	Raw		CoReG		SARAF (2026)		GTR (2026)		δ-Adapter (2026)		GLAFF (2024)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4511 ±0.0053	0.4762 ±0.0030	0.4272 ±0.0027	0.4617 ±0.0021	0.4714 ±0.0044	0.4873 ±0.0030	0.5130 ±0.0540	0.5279 ±0.0411	0.4644 ±0.0068	0.4862 ±0.0044	0.4569 ±0.0037	0.4780 ±0.0034
	192	0.5037 ±0.0030	0.5121 ±0.0020	0.4728 ±0.0011	0.4938 ±0.0006	0.5447 ±0.0016	0.5293 ±0.0013	0.6430 ±0.1633	0.6002 ±0.0895	0.5192 ±0.0038	0.5231 ±0.0025	0.5037 ±0.0034	0.5113 ±0.0030
	336	0.5447 ±0.0033	0.5387 ±0.0019	0.5168 ±0.0028	0.5232 ±0.0012	0.5362 ±0.0008	0.5336 ±0.0003	0.7116 ±0.1847	0.6297 ±0.0945	0.5524 ±0.0031	0.5456 ±0.0022	0.5502 ±0.0064	0.5421 ±0.0036
	720	0.6699 ±0.0071	0.6170 ±0.0037	0.6447 ±0.0045	0.6016 ±0.0026	0.6710 ±0.0042	0.6137 ±0.0022	0.8292 ±0.2671	0.6946 ±0.1175	0.6541 ±0.0102	0.6112 ±0.0056	0.6758 ±0.0058	0.6160 ±0.0036
ETTh2	96	0.2413 ±0.0013	0.3382 ±0.0008	0.2278 ±0.0012	0.3266 ±0.0017	0.2532 ±0.0039	0.3453 ±0.0023	0.5083 ±0.1119	0.5227 ±0.0650	0.2735 ±0.0070	0.3636 ±0.0055	0.2453 ±0.0026	0.3411 ±0.0013
	192	0.2973 ±0.0020	0.3800 ±0.0025	0.2853 ±0.0015	0.3685 ±0.0012	0.2939 ±0.0011	0.3777 ±0.0010	0.6800 ±0.4265	0.5978 ±0.1892	0.3214 ±0.0036	0.3961 ±0.0030	0.3067 ±0.0086	0.3871 ±0.0068
	336	0.3424 ±0.0118	0.4063 ±0.0062	0.3245 ±0.0033	0.3929 ±0.0026	0.3292 ±0.0010	0.4013 ±0.0004	0.7325 ±0.6396	0.5864 ±0.2438	0.3522 ±0.0026	0.4135 ±0.0025	0.3409 ±0.0134	0.4046 ±0.0055
	720	0.4275 ±0.0155	0.4629 ±0.0146	0.4120 ±0.0115	0.4519 ±0.0114	0.4151 ±0.0015	0.4531 ±0.0017	0.4750 ±0.0717	0.4841 ±0.0383	0.3989 ±0.0086	0.4441 ±0.0074	0.4328 ±0.0160	0.4669 ±0.0155

Table E.13: MICN (2023) results across five random seeds. Each cell reports the five-seed mean on the first line and sample standard deviation on the second line; MSE and MAE are separated.

Dataset	Horizon	Raw		CoReG		SARAF (2026)		GTR (2026)		δ-Adapter (2026)		GLAFF (2024)	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.4782 ±0.0077	0.5089 ±0.0051	0.4513 ±0.0055	0.4864 ±0.0037	0.4960 ±0.0052	0.5080 ±0.0031	0.4749 ±0.0224	0.5091 ±0.0172	0.4910 ±0.0077	0.5095 ±0.0050	0.6325 ±0.1246	0.5894 ±0.0644
	192	0.5026 ±0.0035	0.5225 ±0.0023	0.4868 ±0.0024	0.5098 ±0.0015	0.5172 ±0.0015	0.5224 ±0.0006	0.5039 ±0.0116	0.5233 ±0.0090	0.5160 ±0.0035	0.5251 ±0.0024	0.4923 ±0.0031	0.5120 ±0.0010
	336	0.5284 ±0.0046	0.5391 ±0.0023	0.5195 ±0.0017	0.5325 ±0.0011	0.5378 ±0.0014	0.5392 ±0.0010	0.5278 ±0.0042	0.5388 ±0.0037	0.5366 ±0.0039	0.5388 ±0.0026	0.5254 ±0.0091	0.5360 ±0.0054
	720	0.6160 ±0.0039	0.5986 ±0.0009	0.6094 ±0.0033	0.5928 ±0.0011	0.6218 ±0.0039	0.5937 ±0.0019	0.6210 ±0.0099	0.6015 ±0.0050	0.6174 ±0.0032	0.5938 ±0.0007	0.6301 ±0.0087	0.6055 ±0.0035
ETTh2	96	0.2595 ±0.0028	0.3550 ±0.0027	0.2249 ±0.0018	0.3236 ±0.0006	0.2372 ±0.0006	0.3346 ±0.0004	0.2786 ±0.0143	0.3803 ±0.0127	0.2596 ±0.0030	0.3530 ±0.0024	0.5682 ±0.1172	0.5374 ±0.0472
	192	0.2974 ±0.0072	0.3838 ±0.0054	0.2688 ±0.0018	0.3600 ±0.0009	0.2902 ±0.0003	0.3712 ±0.0001	0.3031 ±0.0100	0.3915 ±0.0109	0.3074 ±0.0072	0.3860 ±0.0056	0.3486 ±0.0274	0.4167 ±0.0166
	336	0.3340 ±0.0024	0.4133 ±0.0020	0.3037 ±0.0025	0.3894 ±0.0008	0.3283 ±0.0003	0.3958 ±0.0002	0.3366 ±0.0075	0.4157 ±0.0056	0.3530 ±0.0021	0.4193 ±0.0019	0.3353 ±0.0107	0.4148 ±0.0083
	720	0.4455 ±0.0056	0.4846 ±0.0032	0.3778 ±0.0028	0.4432 ±0.0025	0.4277 ±0.0004	0.4605 ±0.0001	0.4375 ±0.0132	0.4801 ±0.0065	0.4522 ±0.0145	0.4867 ±0.0068	0.4499 ±0.0525	0.4883 ±0.0294

E.6 Complete Data-Sparsity Tables

Table E.14: DLinear results at different training ratios in the data-sparsity experiment. Raw denotes the original Backbone result; 100%, 50%, 25%, and 10% report CoReG at the corresponding training ratios. Non-Raw cells report the value; green indicates a decrease and red indicates an unchanged or increased error relative to the corresponding 100% setting. Relative changes from the 100% setting are reported only in the Avg. row. Each Dataset block ends with an Avg. row averaged over the four horizons.

Dataset	Horizon	Raw		100%		50%		25%		10%	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	96	0.3274	0.3732	0.3132	0.3693	0.3141	0.3717	0.3188	0.3763	0.3318	0.3874
	192	0.3760	0.4034	0.3563	0.3981	0.3612	0.4028	0.3639	0.4046	0.3778	0.4146
	336	0.4181	0.4317	0.4027	0.4292	0.4069	0.4299	0.4070	0.4329	0.4182	0.4412
	720	0.4639	0.4689	0.4518	0.4655	0.4559	0.4671	0.4603	0.4709	0.4675	0.4787
	Avg.	0.3964	0.4193	0.3810 ↓ 0.00%	0.4155 ↓ 0.00%	0.3846 ↑ 0.94%	0.4179 ↑ 0.58%	0.3875 ↑ 1.71%	0.4212 ↑ 1.37%	0.3988 ↑ 4.67%	0.4305 ↑ 3.61%
ETTm2	96	0.1453	0.2508	0.1462	0.2503	0.1470	0.2534	0.1485	0.2551	0.1506	0.2585
	192	0.1785	0.2805	0.1781	0.2785	0.1759	0.2775	0.1816	0.2828	0.1811	0.2839
	336	0.2133	0.3100	0.2101	0.3050	0.2114	0.3064	0.2094	0.3058	0.2120	0.3074
	720	0.2669	0.3521	0.2698	0.3491	0.2600	0.3455	0.2571	0.3429	0.2616	0.3467
	Avg.	0.2010	0.2983	0.2010 ↓ 0.00%	0.2957 ↓ 0.00%	0.1986 ↓ 1.19%	0.2957 ↓ 0.00%	0.1992 ↓ 0.90%	0.2966 ↑ 0.30%	0.2013 ↑ 0.15%	0.2991 ↑ 1.15%
Exchange	96	0.1559	0.3066	0.1051	0.2421	0.1175	0.2604	0.1592	0.3122	0.1942	0.3475
	192	0.2700	0.4025	0.1937	0.3318	0.2053	0.3455	0.2364	0.3815	0.2759	0.4178
	336	0.5119	0.5580	0.2931	0.4187	0.3167	0.4408	0.3376	0.4590	0.3700	0.4705
	720	1.3767	0.9020	0.5797	0.6122	0.5474	0.5985	0.5511	0.6009	0.7102	0.6498
	Avg.	0.5786	0.5423	0.2929 ↓ 0.00%	0.4012 ↓ 0.00%	0.2967 ↑ 1.30%	0.4113 ↑ 2.52%	0.3211 ↑ 9.63%	0.4384 ↑ 9.27%	0.3876 ↑ 32.33%	0.4714 ↑ 17.50%
PEMS08	96	0.3656	0.3579	0.2399	0.2700	0.2649	0.2882	0.3025	0.3137	0.3455	0.3460
	192	0.3967	0.3724	0.2861	0.2759	0.3004	0.2880	0.3346	0.3280	0.3413	0.3247
	336	0.4219	0.3915	0.3018	0.2811	0.3105	0.2914	0.3308	0.2992	0.3582	0.3390
	720	0.4844	0.4401	0.3331	0.3078	0.3379	0.3163	0.3512	0.3234	0.3619	0.3379
	Avg.	0.4172	0.3905	0.2902 ↓ 0.00%	0.2837 ↓ 0.00%	0.3034 ↑ 4.55%	0.2960 ↑ 4.34%	0.3298 ↑ 13.65%	0.3161 ↑ 11.42%	0.3517 ↑ 21.19%	0.3369 ↑ 18.75%

Table E.15: iTransformer results at different training ratios in the data-sparsity experiment. Raw denotes the original Backbone result; 100%, 50%, 25%, and 10% report CoReG at the corresponding training ratios. Non-Raw cells report the value; green indicates a decrease and red indicates an unchanged or increased error relative to the corresponding 100% setting. Relative changes from the 100% setting are reported only in the Avg. row. Each Dataset block ends with an Avg. row averaged over the four horizons.

Dataset	Horizon	Raw		100%		50%		25%		10%	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	96	0.3343	0.3875	0.3193	0.3783	0.3189	0.3796	0.3219	0.3814	0.3318	0.3893
	192	0.3894	0.4229	0.3677	0.4115	0.3709	0.4136	0.3673	0.4104	0.3785	0.4180
	336	0.4318	0.4494	0.4120	0.4388	0.4093	0.4383	0.4122	0.4399	0.4229	0.4480
	720	0.4779	0.4855	0.4667	0.4780	0.4663	0.4786	0.4673	0.4788	0.4784	0.4875
	Avg.	0.4083	0.4363	0.3914 ↓0.00%	0.4267 ↓0.00%	0.3913 ↓0.03%	0.4275 ↑0.19%	0.3922 ↑0.20%	0.4276 ↑0.21%	0.4029 ↑2.94%	0.4357 ↑2.11%
ETTm2	96	0.1619	0.2686	0.1522	0.2581	0.1544	0.2590	0.1530	0.2605	0.1533	0.2622
	192	0.1970	0.2972	0.1877	0.2896	0.1847	0.2872	0.1851	0.2878	0.1850	0.2893
	336	0.2237	0.3178	0.2163	0.3109	0.2172	0.3120	0.2155	0.3109	0.2172	0.3121
	720	0.2965	0.3684	0.2830	0.3600	0.2728	0.3529	0.2697	0.3500	0.2696	0.3492
	Avg.	0.2198	0.3130	0.2098 ↓0.00%	0.3047 ↓0.00%	0.2073 ↓1.19%	0.3028 ↓0.62%	0.2058 ↓1.91%	0.3023 ↓0.79%	0.2063 ↓1.67%	0.3032 ↓0.49%
Exchange	96	0.1284	0.2627	0.1212	0.2509	0.1222	0.2543	0.1575	0.2948	0.2333	0.3615
	192	0.2488	0.3740	0.2325	0.3546	0.2731	0.3875	0.2977	0.4087	0.3781	0.4618
	336	0.4118	0.4824	0.3795	0.4570	0.4419	0.4971	0.4503	0.5014	0.5255	0.5463
	720	1.2182	0.8364	1.0685	0.7842	1.0968	0.7995	1.1136	0.8088	1.1789	0.8228
	Avg.	0.5018	0.4888	0.4504 ↓0.00%	0.4617 ↓0.00%	0.4835 ↑7.35%	0.4846 ↑4.96%	0.5048 ↑12.08%	0.5034 ↑9.03%	0.5790 ↑28.55%	0.5481 ↑18.71%
PEMS08	96	0.2230	0.2607	0.2194	0.2513	0.2464	0.2683	0.2976	0.2979	0.3705	0.3436
	192	0.2921	0.2869	0.2908	0.2653	0.3064	0.2789	0.3452	0.3008	0.3748	0.3294
	336	0.3217	0.3115	0.3069	0.2840	0.3244	0.2868	0.3499	0.3017	0.3797	0.3285
	720	0.4029	0.3655	0.3575	0.3009	0.3699	0.3156	0.3779	0.3276	0.3817	0.3412
	Avg.	0.3099	0.3062	0.2937 ↓0.00%	0.2754 ↓0.00%	0.3118 ↑6.16%	0.2874 ↑4.36%	0.3426 ↑16.65%	0.3070 ↑11.47%	0.3767 ↑28.26%	0.3357 ↑21.90%

Table E.16: PatchMLP results at different training ratios in the data-sparsity experiment. Raw denotes the original Backbone result; 100%, 50%, 25%, and 10% report CoReG at the corresponding training ratios. Non-Raw cells report the value; green indicates a decrease and red indicates an unchanged or increased error relative to the corresponding 100% setting. Relative changes from the 100% setting are reported only in the Avg. row. Each Dataset block ends with an Avg. row averaged over the four horizons.

Dataset	Horizon	Raw		100%		50%		25%		10%	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	96	0.3414	0.3929	0.3234	0.3807	0.3221	0.3805	0.3275	0.3840	0.3456	0.3980
	192	0.3968	0.4315	0.3784	0.4163	0.3870	0.4247	0.3799	0.4200	0.3898	0.4271
	336	0.4432	0.4586	0.4195	0.4425	0.4108	0.4406	0.4196	0.4461	0.4287	0.4523
	720	0.4838	0.4855	0.4647	0.4770	0.4635	0.4766	0.4670	0.4792	0.4809	0.4886
	Avg.	0.4163	0.4421	0.3965 ↓0.00%	0.4291 ↓0.00%	0.3958 ↓0.18%	0.4306 ↑0.35%	0.3985 ↑0.50%	0.4323 ↑0.75%	0.4112 ↑3.71%	0.4415 ↑2.89%
ETTm2	96	0.1627	0.2665	0.1571	0.2606	0.1548	0.2597	0.1568	0.2628	0.1593	0.2674
	192	0.2054	0.3015	0.1901	0.2874	0.1913	0.2892	0.1917	0.2912	0.1952	0.2961
	336	0.2379	0.3246	0.2326	0.3204	0.2293	0.3181	0.2265	0.3176	0.2305	0.3216
	720	0.3039	0.3703	0.2846	0.3596	0.2840	0.3586	0.2796	0.3556	0.2762	0.3533
	Avg.	0.2275	0.3157	0.2161 ↓0.00%	0.3070 ↓0.00%	0.2148 ↓0.60%	0.3064 ↓0.20%	0.2136 ↓1.16%	0.3068 ↓0.07%	0.2153 ↓0.37%	0.3096 ↑0.85%
Exchange	96	0.1470	0.2831	0.1367	0.2715	0.1565	0.2923	0.1749	0.3090	0.2103	0.3399
	192	0.3369	0.4314	0.2707	0.3837	0.3021	0.4071	0.3129	0.4150	0.3585	0.4479
	336	0.6113	0.5618	0.5226	0.5240	0.5043	0.5232	0.4856	0.5152	0.5093	0.5314
	720	1.6860	0.9271	1.3026	0.8398	1.1269	0.8097	1.1391	0.8150	1.1508	0.8155
	Avg.	0.6953	0.5509	0.5581 ↓0.00%	0.5048 ↓0.00%	0.5225 ↓6.38%	0.5081 ↑0.65%	0.5281 ↓5.38%	0.5135 ↑1.72%	0.5572 ↓0.16%	0.5337 ↑5.73%
PEMS08	96	0.2292	0.2650	0.2183	0.2512	0.2584	0.2761	0.3096	0.3072	0.3691	0.3460
	192	0.3038	0.2772	0.2903	0.2613	0.3047	0.2811	0.3367	0.3015	0.3763	0.3339
	336	0.3573	0.3109	0.3218	0.2792	0.3365	0.2909	0.3536	0.3030	0.3760	0.3233
	720	0.4082	0.3605	0.3644	0.3042	0.3747	0.3167	0.3823	0.3300	0.3952	0.3580
	Avg.	0.3247	0.3034	0.2987 ↓0.00%	0.2740 ↓0.00%	0.3186 ↑6.66%	0.2912 ↑6.28%	0.3455 ↑15.67%	0.3104 ↑13.28%	0.3791 ↑26.92%	0.3403 ↑24.20%

Table E.17: Amplifier results at different training ratios in the data-sparsity experiment. Raw denotes the original Backbone result; 100%, 50%, 25%, and 10% report CoReG at the corresponding training ratios. Non-Raw cells report the value; green indicates a decrease and red indicates an unchanged or increased error relative to the corresponding 100% setting. Relative changes from the 100% setting are reported only in the Avg. row. Each Dataset block ends with an Avg. row averaged over the four horizons.

Dataset	Horizon	Raw		100%		50%		25%		10%	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	96	0.3308	0.3811	0.3188	0.3731	0.3211	0.3770	0.3212	0.3797	0.3363	0.3931
	192	0.3775	0.4102	0.3665	0.4053	0.3649	0.4062	0.3706	0.4119	0.3843	0.4227
	336	0.4218	0.4395	0.4095	0.4344	0.4047	0.4346	0.4127	0.4394	0.4246	0.4494
	720	0.4777	0.4799	0.4652	0.4744	0.4630	0.4751	0.4671	0.4786	0.4808	0.4884
	Avg.	0.4020	0.4277	0.3900 ↓0.00%	0.4218 ↓0.00%	0.3884 ↓0.41%	0.4232 ↑0.33%	0.3929 ↑0.74%	0.4274 ↑1.33%	0.4065 ↑4.23%	0.4384 ↑3.94%
ETTm2	96	0.1511	0.2582	0.1462	0.2523	0.1529	0.2577	0.1525	0.2597	0.1518	0.2606
	192	0.1902	0.2893	0.1814	0.2818	0.1786	0.2810	0.1822	0.2849	0.1888	0.2920
	336	0.2150	0.3110	0.2093	0.3058	0.2126	0.3080	0.2140	0.3107	0.2169	0.3117
	720	0.2653	0.3487	0.2629	0.3475	0.2640	0.3462	0.2690	0.3488	0.2670	0.3478
	Avg.	0.2054	0.3018	0.1999 ↓0.00%	0.2968 ↓0.00%	0.2020 ↑1.05%	0.2982 ↑0.47%	0.2044 ↑2.25%	0.3010 ↑1.42%	0.2061 ↑3.10%	0.3030 ↑2.09%
Exchange	96	0.1809	0.3194	0.1370	0.2731	0.1167	0.2458	0.1563	0.2930	0.2648	0.3854
	192	0.2157	0.3390	0.2201	0.3421	0.2524	0.3703	0.2810	0.3947	0.3734	0.4577
	336	0.3663	0.4440	0.3775	0.4519	0.4234	0.4838	0.4598	0.5028	0.5074	0.5330
	720	1.0668	0.7860	1.0753	0.7831	1.1137	0.8045	1.1117	0.8062	1.1580	0.8156
	Avg.	0.4574	0.4721	0.4525 ↓0.00%	0.4626 ↓0.00%	0.4766 ↑5.33%	0.4761 ↑2.92%	0.5022 ↑10.98%	0.4992 ↑7.91%	0.5759 ↑27.27%	0.5479 ↑18.44%
PEMS08	96	0.4914	0.4312	0.2502	0.2615	0.2724	0.2803	0.3211	0.3153	0.4884	0.4254
	192	0.4704	0.4004	0.3161	0.2752	0.3242	0.2870	0.3492	0.3025	0.4114	0.3674
	336	0.4671	0.4062	0.3265	0.2823	0.3371	0.2930	0.3690	0.3279	0.3818	0.3386
	720	0.5026	0.4561	0.3671	0.3080	0.3797	0.3200	0.3844	0.3254	0.3839	0.3434
	Avg.	0.4829	0.4235	0.3150 ↓0.00%	0.2818 ↓0.00%	0.3284 ↑4.25%	0.2951 ↑4.72%	0.3559 ↑12.98%	0.3178 ↑12.78%	0.4164 ↑32.19%	0.3687 ↑30.84%

F LLM USAGE

We used GPT-5.6 as an auxiliary language tool during the manuscript preparation process. Its use was limited to polishing the language, organizing sentences, and improving readability.

GPT-5.6 had no role in defining the research questions, designing the study, conceiving the methodology, implementing or conducting the experiments, analyzing the results, or formulating the conclusions. The authors conceived, executed, and independently verified all methodological contributions, experimental implementation and execution, analyses, and conclusions. The authors take full responsibility for the technical content, reported results, and final wording of this paper.