BETTER DEPENDENCIES, BETTER FORECASTS: DEPENDENCY RECTIFICATION FOR ROBUST TIME SERIES FORECASTING

Anonymous authors
Paper under double-blind review

ABSTRACT

Time series forecasting typically assumes that predictive dependencies can be reliably learned from observed histories. However, real-world observations are often imperfect: noise, missing values, and localized distortions not only perturb individual values but also alter the apparent temporal dependencies that models rely on. As a result, directly fitting observed correlations may cause forecasting models to absorb spurious and non-generalizable relations, degrading robustness under corrupted inputs. To address this issue, we propose FenDiM, a Frequency-Enhanced Diffusion-Mamba forecasting architecture for prediction-oriented dependency rectification. FenDiM progressively suppresses corruption-induced dependencies through diffusion-based rectification, uses frequency-domain evidence to compensate for unreliable time-domain patterns, and leverages Mamba for efficient long-range temporal modeling. Across six benchmarks, FenDiM achieves the best horizon-averaged MSE on five datasets under standard observations, reducing MSE by 20.7% on average relative to the strongest competing baseline across these five datasets. At a prediction length of 192, it ranks within the top two in both MAE and MSE across all tested dataset–missing-rate settings and remains competitive under Gaussian noise. These results demonstrate the effectiveness of prediction-oriented dependency rectification for robust time series forecasting. The code and datasets are available at `https://anonymous.4open.science/r/FenDim-5035`.

1 INTRODUCTION

Time series forecasting is a fundamental task with broad applications in energy management, traffic control He et al. (2025); Zhu et al. (2024); Wen et al. (2025), economic planning Zhang et al. (2025b); Zhu et al. (2025); You et al. (2026), weather prediction Zheng et al. (2025); Zhao et al. (2024); Zhang et al. (2025c), and ocean monitoring Xu et al. (2025); Huang et al. (2025a); El Aouni et al. (2025). Modern forecasting models have achieved strong performance by learning temporal patterns from historical observations and extrapolating them into the future. Despite their architectural diversity, most existing methods share a common premise: the dependencies fitted from observed histories provide reliable predictive evidence for future dynamics.

However, the premise can be fragile in real-world settings. Historical observations are often incomplete, noisy, or locally unstable due to missing values Khayati et al. (2025), noise Lu et al. (2026), abrupt changes Liu et al. (2026), and local fluctuations Moebus et al. (2025). In forecasting, dependencies refer to the predictive relations that connect historical evidence to future outcomes, such as which past segments, variables, or temporal patterns a model should rely on for prediction. Imperfect observations not only corrupt individual values, but can also bias these learned dependencies: transient segments, corrupted variables, or short-term fluctuations may appear predictive despite being irrelevant to future evolution. Consequently, directly fitting observed associations may cause forecasting models to internalize spurious and non-generalizable temporal relations, especially under long prediction horizons or distributional perturbations.

Although recent methods improve forecasting through stronger temporal architectures Qiu et al. (2025); Wu et al. (2025), robust modeling strategies Wang et al. (2025b); An et al. (2026), or

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.

reconstruction-oriented treatments Sun et al. (2025); Huang et al. (2025b) of imperfect inputs, most of them still build predictions from dependencies extracted from the observed sequence. When imperfect observations bias these dependencies, forecasting models may optimize over relations that are statistically salient in the history but irrelevant or harmful to future dynamics. Therefore, the central challenge is not to repair corrupted values as a preprocessing step, but to rectify dependency formation under the guidance of the prediction objective. This motivates **prediction-oriented dependency rectification**: suppressing spurious relations while preserving forecasting-relevant dependencies.

To this end, we propose **FenDiM**, a **F**requency-**En**hanced **Di**ffusion-**M**amba forecasting framework for prediction-oriented dependency rectification. FenDiM does not treat imperfect observations as a preprocessing problem; instead, it rectifies misleading dependencies within the forecasting process. **Progressive Dependency Rectification** (PDR) performs diffusion-based dependency rectification to suppress spurious temporal relations induced by imperfect observations. **Frequency-Compensated Temporal Learning** (FCTL) compensates the rectified representation with frequency-domain structure when local time-domain evidence is unreliable. A Mamba-based temporal architecture reduces the sequential modeling cost of diffusion-based rectification, enabling efficient long-horizon forecasting. A **Prediction-Oriented Joint Calibration** objective further ensures that all components serve forecasting utility rather than value reconstruction alone.

Experiments on six real-world benchmarks show that FenDiM achieves the best horizon-averaged MSE on five datasets under standard observations, with maximum MSE and MAE reductions of 38.5% and 11.5% relative to the strongest competing baseline for each dataset and metric. At prediction horizon $T = 192$, the largest MAE reductions under missing and Gaussian-noisy inputs are 53.4% and 21.1%, respectively. These results support prediction-oriented dependency rectification across standard and corrupted observations, while the gains vary by dataset and error metric.

We summarize our main contributions as follows:

- We highlight a dependency-level perspective on robust time series forecasting, showing that imperfect observations can degrade prediction not only by corrupting values, but also by making unreliable temporal relations appear predictive. This motivates prediction-oriented dependency rectification, which learns forecasting-relevant relations under the guidance of the prediction objective.
- We propose FenDiM, a Frequency-Enhanced Diffusion-Mamba framework that rectifies misleading dependencies for forecasting. FenDiM couples diffusion-based rectification, frequency-domain compensation, and Mamba-based temporal modeling to learn reliable long-range temporal relations from imperfect histories.
- Extensive experiments under standard, missing-observation, noisy-observation, and compound missing-noise settings show that FenDiM achieves the best dataset-average MSE on five of six benchmarks and consistently maintains strong robustness across diverse observation perturbations.

2 Related Work

2.1 Forecasting with Imperfect Observations

Time-series forecasting often assumes reliable historical observations, yet missingness, noise, and local distortions can corrupt both values and inferred dependencies. A common strategy is imputation-then-forecast, using models such as BRITS Cao et al. (2018), SAITS Du et al. (2023), and CSDI Tashiro et al. (2021); however, reconstruction-oriented imputation may propagate errors into prediction. Recent methods instead model imperfect observations directly: BiTGraph Chen et al. (2023) accounts for missingness bias, ProFITi Yalavarthi et al. (2025) models irregular observations probabilistically, and Merlin Yu et al. (2025) improves robustness through distillation and contrastive learning. Despite improved robustness, these methods still condition on perturbation-biased dependencies, allowing spurious relations to be treated as predictive evidence.

2.2 Robust Forecasting under Noisy Evidence

Robust forecasting methods seek to prevent models from fitting unreliable temporal evidence. DecompNet Luo & Wang (2025) injects implicit trend-seasonal decomposition to mitigate noisy high-frequency variations. Selective Learning Fu et al. (2025) masks uncertain and anomalous timesteps to avoid optimizing non-generalizable targets. Time-o1 Wang et al. (2025a) transforms correlated labels into decorrelated components to reduce label-side bias. AMRC Liang et al. (2025) adaptively masks redundant historical segments, while MEW Zhang et al. (2025a) identifies the maximum effective look-back window and filters redundant subsequences through an information bottleneck. PIR Liu et al. (2025) revises biased forecasting instances after prediction, and AliO Park et al. (2025) improves consistency by aligning overlapping forecasts across lagged windows. These studies reveal that not all temporal evidence should be optimized or trusted equally. However, they mainly treat unreliability as an evidence-selection or error-correction problem, rather than a dependency-formation bias.

2.3 Diffusion Models for Temporal Recovery

Diffusion models have been introduced into time-series modeling to handle uncertainty in temporal prediction and recovery. TimeGrad Rasul et al. (2021) adapts diffusion to probabilistic forecasting through autoregressive denoising, but does not target incomplete observations. CSDI Tashiro et al. (2021) extends conditional diffusion to missing-value imputation, while PriSTI Liu et al. (2023a) further incorporates spatiotemporal priors for structured recovery; however, both are primarily optimized for reconstruction fidelity. Recent methods such as OVC-Diff Wang et al. (2023) and TSDE Senane et al. (2024) improve observed-value consistency and broaden diffusion-based learning to imputation, interpolation, and forecasting-oriented masking. Nevertheless, this line of work largely prioritizes recovery over forecasting utility, which may allow perturbation-induced dependencies to persist and harm prediction.

3 Problem Definition

We consider multivariate forecasting from complete or imperfect histories. Given a clean historical sequence $\mathbf{X} \in \mathbb{R}^{L\times C}$, the goal is to predict the clean future $\mathbf{Y} \in \mathbb{R}^{L_p\times C_{\text{tar}}}$. We model the accessible history as

$$\widetilde{\mathbf{X}} = \mathbf{\Omega} \odot (\mathbf{X} + \mathbf{E}) + (1 - \mathbf{\Omega}) \odot \mathbf{\Phi}, \tag{1}$$

where $\mathbf{\Omega} \in \{0, 1\}^{L\times C}$ is the observation mask, $\mathbf{\Phi}$ is the placeholder for unavailable values, and $E_{\ell,c} \sim \mathcal{N}(0, \sigma^2 s_c^2)$, with s_c denoting the training-set standard deviation of variable c. This formulation covers complete ($\mathbf{\Omega} = \mathbf{1}, \mathbf{E} = \mathbf{0}$), missing ($\mathbf{E} = \mathbf{0}$), noisy ($\mathbf{\Omega} = \mathbf{1}$), and compound missing-noise observations. The forecasting task is

$$\widehat{\mathbf{Y}} = f_\theta(\widetilde{\mathbf{X}}, \mathbf{\Omega}), \tag{2}$$

where the target $\mathbf{Y}$ remains clean.

4 Methodology

4.1 Framework Overview

We present **FenDiM**, an end-to-end framework for robust multivariate time series forecasting, as illustrated in Figure 1. Rather than treating perturbations as an isolated preprocessing issue, FenDiM addresses a more fundamental problem: dependencies directly extracted from observed histories may be biased by missingness, noise, or localized distortions, and therefore become unreliable for prediction.

FenDiM formulates forecasting as *prediction-oriented dependency rectification*. It consists of three coupled components: Progressive Dependency Rectification (PDR), which corrects biased dependency representations; Frequency-Compensated Temporal Learning (FCTL), which uses frequency-domain structural cues to recalibrate unreliable temporal evidence; and a Mamba-based forecasting backbone, which models the calibrated representation for long-range prediction. All components are optimized jointly, so that dependency rectification and frequency compensation are aligned with forecasting utility rather than independent reconstruction or denoising goals.

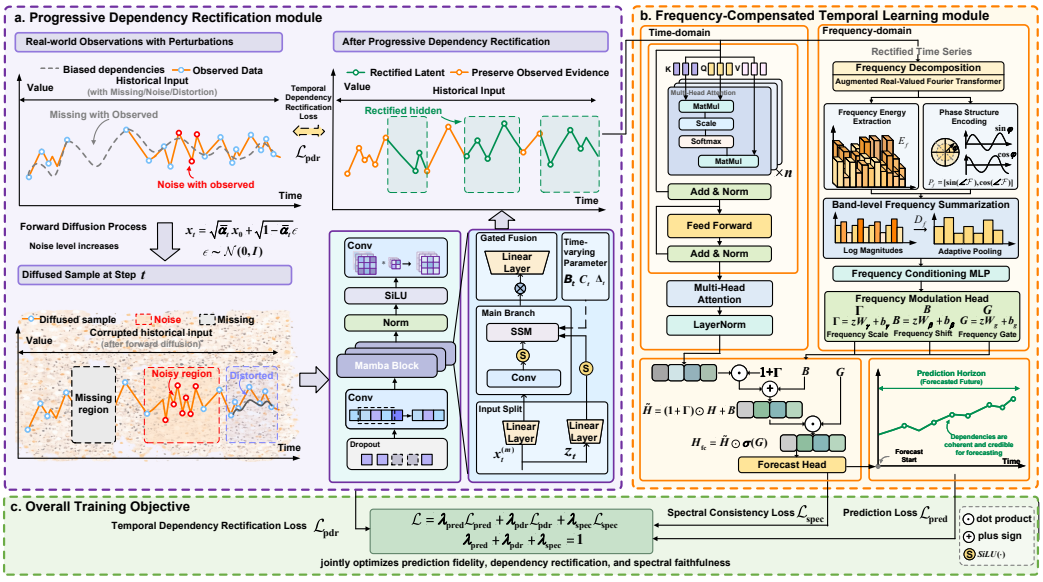

Figure 1: **Framework overview of FenDiM.** FenDiM performs prediction-oriented dependency rectification for robust forecasting. PDR suppresses corruption-induced spurious dependencies, FCTL injects frequency-domain structural cues to compensate unreliable temporal evidence, and a Mamba-based forecasting backbone produces long-range predictions from the calibrated representation under a joint training objective.

4.2 Progressive Dependency Rectification via Mamba-based Conditional Diffusion

PDR serves as the first stage of FenDiM and performs conditional diffusion in the latent dependency space. Instead of correcting corrupted values directly, it transforms the initially fitted representation into a rectified forecasting state, providing the downstream forecaster with more stable and prediction-relevant temporal evidence.

Given the imperfect input $\widetilde{X} \in \mathbb{R}^{B\times L\times C}$ and the observation mask $\Omega \in \{0,1\}^{B\times L\times C}$, we first encode an initial dependency representation:

$$Z_0 = g_{\theta_g}(\widetilde{X}, \Omega), \qquad Z_0 \in \mathbb{R}^{B\times L_z\times d}, \tag{3}$$

where g_{θ_g} is a temporal encoder. For purely noisy observations, Ω is set to an all-one mask. Although Z_0 summarizes dependencies extracted from available observations, it may also contain misleading relations amplified by corruptions.

To make rectification condition-aware, we construct a partial evidence context. During training, available entries are randomly subsampled:

$$\Omega_{\text{cond}} = \Omega \odot M_{\text{cond}}, \qquad \widetilde{X}_{\text{cond}} = \widetilde{X} \odot \Omega_{\text{cond}}, \tag{4}$$

where M_{cond} is a random conditioning mask. The retained observations are encoded as

$$Z_{\text{cond}} = c_\psi(\widetilde{X}_{\text{cond}}, \Omega_{\text{cond}}), \tag{5}$$

where c_ψ is a conditional evidence encoder. At inference, we use all available observations as the conditioning context by setting $\Omega_{\text{cond}} = \Omega$. This conditional design prevents the rectifier from performing unconditional denoising and encourages it to recover stable predictive dependencies from available evidence.

PDR performs conditional diffusion in the representation space. We perturb Z_0 as

$$Z_t = \sqrt{\bar{\alpha}_t} Z_0 + \sqrt{1-\bar{\alpha}_t}\epsilon, \qquad \epsilon \sim \mathcal{N}(0, I), \tag{6}$$

where t is the diffusion step and $\bar{\alpha}_t$ is the cumulative signal-preservation coefficient. A Mamba-based conditional denoiser predicts the injected perturbation:

$$\hat{\epsilon}_{\theta_d} = \mathcal{D}_{\theta_d}^{\mathrm{Mamba}} \left(Z_t + e(t), Z_{\mathrm{cond}}, \Omega_{\mathrm{cond}} \right), \tag{7}$$

where $e(t)$ is the timestep embedding. Mamba is adopted because its selective state-space updates capture long-range temporal dependencies with linear-time complexity, making iterative representation rectification more practical than attention-based denoising for long-horizon forecasting.

The corresponding denoising objective is introduced in Eq. 17, where it is jointly optimized with the forecasting and spectral consistency losses.

After training, the learned reverse process maps the initially fitted representation to a rectified forecasting state:

$$Z_{\mathrm{rect}} = \mathcal{R}_{\theta_d} \left(Z_0, Z_{\mathrm{cond}}, \Omega_{\mathrm{cond}} \right). \tag{8}$$

By operating on dependency representations rather than raw values, PDR corrects misleading temporal relations before prediction, enabling the downstream forecaster to extrapolate from more prediction-relevant evidence.

4.3 Frequency-Compensated Temporal Learning via Frequency Evidence

FCTL complements PDR by compensating the rectified temporal representation with frequency-domain evidence. It injects sequence-level spectral regularities into temporal modeling, recalibrating unstable local fluctuations without introducing an independent frequency forecasting branch.

Let $Z_{\mathrm{rect}} \in \mathbb{R}^{B \times L_z \times d}$ denote the rectified representation produced by PDR. We first project Z_{rect} into a temporal carrier:

$$H = \mathcal{E}_{\mathrm{T}} \left(Z_{\mathrm{rect}} W_{\mathrm{in}} + b_{\mathrm{in}} \right), \qquad H \in \mathbb{R}^{B \times L_z \times D}, \tag{9}$$

where $\mathcal{E}_{\mathrm{T}}(\cdot)$ is a temporal encoder, and W_{in}, b_{in} project the rectified representation into the hidden dimension D.

We then extract complementary frequency evidence from Z_{rect} using the one-sided Fourier projection along the temporal dimension:

$$\mathcal{F}_{b,k,:} = \sum_{\ell=0}^{L_z-1} Z_{\mathrm{rect}}[b, \ell, :] \exp\left(-\mathrm{i} \frac{2\pi k \ell}{L_z} \right), \qquad k = 0, \ldots, F_z - 1, \tag{10}$$

where $F_z = \lfloor L_z/2 \rfloor + 1$.

We summarize the spectrum into a compact descriptor:

$$U_f = \mathrm{Pool}_K \left(\mathrm{Concat} \left[\log(1 + |\mathcal{F}|), \sin(\angle \mathcal{F}), \cos(\angle \mathcal{F}) \right] \right), \tag{11}$$

where $\mathrm{Pool}_K(\cdot)$ adaptively pools the frequency axis into K bands. The descriptor is then projected into a frequency code:

$$z_f = \psi_f \left(\mathrm{vec}(U_f) \right), \qquad z_f \in \mathbb{R}^{B \times D}, \tag{12}$$

where $\psi_f(\cdot)$ is a lightweight projection network.

The frequency code generates channel-wise compensation parameters:

$$\Gamma, \Delta, G_f = \phi_f(z_f), \qquad \Gamma, \Delta, G_f \in \mathbb{R}^{B \times 1 \times D}, \tag{13}$$

where $\phi_f(\cdot)$ maps the frequency code into channel-wise modulation parameters. The temporal carrier is recalibrated by residual frequency modulation:

$$H_{\mathrm{fc}} = H + \sigma(G_f) \odot (\Gamma \odot H + \Delta). \tag{14}$$

This residual form preserves the original temporal carrier while allowing frequency evidence to selectively compensate unreliable temporal activations.

Finally, the calibrated representation is fed into the Mamba-based forecasting backbone:

$$\widehat{Y} = f_{\omega}^{\mathrm{Mamba}} \left(H_{\mathrm{fc}} \right). \tag{15}$$

In this way, FCTL realizes frequency-to-time compensation: temporal modeling captures local sequential dependencies, while frequency evidence provides global structural guidance for stable forecasting.

4.4 Loss Function

FenDiM is optimized with a prediction-oriented joint objective that couples forecasting accuracy, dependency rectification, and frequency consistency. Given the prediction $\widehat{Y} \in \mathbb{R}^{B\times L_p\times C_{\text{tar}}}$ and the ground truth $Y \in \mathbb{R}^{B\times L_p\times C_{\text{tar}}}$, the point-wise forecasting loss is

$$\mathcal{L}_{\text{pred}} = \frac{1}{BL_pC_{\text{tar}}} \sum_{b=1}^{B}\sum_{\tau=1}^{L_p}\sum_{c=1}^{C_{\text{tar}}} \left(\widehat{Y}_{b,\tau,c} - Y_{b,\tau,c}\right)^2. \tag{16}$$

For dependency rectification, PDR is supervised in the latent dependency space by predicting the injected perturbation:

$$\mathcal{L}_{\text{pdr}} = \mathbb{E}_{t,\epsilon}\left[\|\epsilon - \hat{\epsilon}_{\theta_d}\|_2^2\right], \tag{17}$$

where ϵ is the injected Gaussian noise and $\hat{\epsilon}_{\theta_d}$ is predicted by the Mamba-based conditional denoiser in Eq. 7.

To preserve future-relevant spectral structure, we define a frequency consistency loss on the log-magnitude spectrum:

$$\mathcal{L}_{\text{spec}} = \frac{1}{BF_pC_{\text{tar}}} \left\|\log\left(1 + \left|\mathcal{F}^+(\widehat{Y})\right|\right) - \log\left(1 + \left|\mathcal{F}^+(Y)\right|\right)\right\|_1, \tag{18}$$

where $\mathcal{F}^+(\cdot)$ denotes the one-sided Fourier projection along the prediction horizon and $F_p = \lfloor L_p/2\rfloor + 1$.

The overall objective is

$$\mathcal{L} = \lambda_{\text{pred}}\mathcal{L}_{\text{pred}} + \lambda_{\text{pdr}}\mathcal{L}_{\text{pdr}} + \lambda_{\text{spec}}\mathcal{L}_{\text{spec}}, \tag{19}$$

where

$$\lambda_{\text{pred}} + \lambda_{\text{pdr}} + \lambda_{\text{spec}} = 1, \qquad \lambda_{\text{pred}}, \lambda_{\text{pdr}}, \lambda_{\text{spec}} \geq 0. \tag{20}$$

This joint objective keeps PDR and FCTL aligned with forecasting utility, rather than optimizing independent reconstruction or frequency-matching goals.

5 Experiments

To comprehensively evaluate FenDiM, we organize our experiments along six complementary perspectives: general forecasting accuracy under standard observations, robustness to incomplete observations, robustness to compound corruptions involving both missingness and noise, component-wise effectiveness, and hyperparameter sensitivity. Specifically, our experiments are designed to answer the following research questions:

- **RQ1:** How does FenDiM perform compared with strong state-of-the-art forecasting methods under standard observation settings?
- **RQ2:** Can FenDiM maintain robust forecasting performance when observations are incomplete, especially under high missing ratios?
- **RQ3:** How robust is FenDiM when observations are corrupted by Gaussian noise?
- **RQ4:** How does FenDiM perform under compound corruptions where missingness and noise occur simultaneously?
- **RQ5:** How does each key component of FenDiM contribute to forecasting accuracy?
- **RQ6:** Is FenDiM stable to hyperparameter choices and practical in computational efficiency?

5.1 Experimental Settings

Datasets. We conduct extensive experiments on six real-world multivariate time series datasets, including Electricity Transformer Temperature (ETTh1, ETTh2, ETTm1, ETTm2) Zhou et al. (2021), Weather, and Exchange Rate. Detailed dataset statistics and preprocessing protocols are provided in Appendix A.

Baselines. We compare FenDiM with recent state-of-the-art forecasting models, including TimeMosaic Ding et al. (2026), SparseTSF Lin et al. (2025), SimpleTM Chen et al. (2025), DPEM Hou et al. (2025), xPatch Stitsyuk & Choi (2025), AMD Hu et al. (2025), DeepEDM Majeedi et al. (2025), PDF Dai et al. (2024), iTransformer Liu et al. (2023b), BiTGraph Chen et al. (2023), SSSD Alcaraz & Strodthoff (2022), and Autoformer Wu et al. (2021). Detailed descriptions are provided in Appendix B.

Implementation Details. Following prior work, we report Mean Squared Error (MSE) and Mean Absolute Error (MAE) as evaluation metrics. All experiments are implemented in PyTorch Paszke et al. (2019) and conducted on a single NVIDIA RTX 5090 GPU. Detailed implementation settings and the unified experimental protocol are provided in Appendices B.2 and B.3.

Table 1: Performance comparison on multivariate forecasting benchmarks with historical length $L = 96$ and prediction lengths $T \in \{96, 192, 336, 720\}$. Lower MSE/MAE indicates better forecasting accuracy. The best and second-best results in each row are highlighted in red and blue, respectively.

Model		FenDiM (Ours)		TimeMosaic (2026)		SparseTSF (2025)		SimpleTM (2025)		DPEM (2025)		xPatch (2025)		AMD (2025)		DeepEDM (2025)		PDF (2024)		iTransformer (2023)		BiTGraph (2023)		SSSD (2023)		Autoformer (2021)	
Metric		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
ETTh1	96	0.089	0.187	0.135	0.531	0.159	0.714	0.143	0.574	0.138	0.512	0.117	0.361	0.108	0.346	0.106	0.293	0.149	0.664	0.205	1.279	0.164	0.451	0.098	0.271	0.202	1.228
	192	0.099	0.224	0.132	0.452	0.165	0.749	0.153	0.651	0.142	0.552	0.121	0.364	0.117	0.395	0.107	0.295	0.158	0.708	0.204	1.183	0.186	0.579	0.098	0.274	0.208	1.235
	336	0.101	0.244	0.138	0.471	0.172	0.791	0.162	0.724	0.151	0.617	0.125	0.378	0.125	0.441	0.107	0.298	0.125	0.441	0.214	1.331	0.208	0.724	0.100	0.282	0.227	1.507
	720	0.111	0.304	0.144	0.490	0.189	0.911	0.171	0.753	0.158	0.660	0.107	0.254	0.125	0.426	0.109	0.304	0.185	0.867	0.223	1.338	0.226	0.846	0.100	0.287	0.225	1.464
	Avg	0.100	0.240	0.137	0.486	0.171	0.791	0.157	0.675	0.147	0.585	0.118	0.339	0.119	0.402	0.107	0.297	0.154	0.670	0.212	1.283	0.196	0.650	0.099	0.279	0.216	1.359
ETTh2	96	0.147	0.450	0.232	1.472	0.278	2.363	0.235	1.517	0.227	1.164	0.214	1.320	0.198	1.045	0.165	0.718	0.246	1.808	0.328	3.450	0.282	1.338	0.202	0.953	0.289	2.115
	192	0.160	0.538	0.258	1.877	0.309	2.941	0.153	0.651	0.257	1.539	0.246	1.794	0.225	1.343	0.167	0.817	0.281	2.481	0.347	3.718	0.325	1.765	0.213	1.039	0.294	2.216
	336	0.166	0.593	0.286	2.413	0.340	3.637	0.293	2.618	0.279	1.882	0.278	2.423	0.260	1.791	0.170	0.985	0.313	3.243	0.401	4.917	0.348	2.001	0.212	1.040	0.321	2.787
	720	0.180	0.697	0.333	3.259	0.407	5.138	0.342	3.562	0.308	2.210	0.231	1.734	0.253	1.704	0.177	1.190	1.704	4.545	0.541	7.224	0.358	2.106	0.213	1.044	0.370	3.652
	Avg	0.163	0.570	0.277	2.255	0.334	3.520	0.256	2.087	0.268	1.699	0.242	1.818	0.234	1.471	0.170	0.927	0.636	3.019	0.404	4.827	0.328	1.803	0.210	1.019	0.318	2.692
ETTm1	96	0.054	0.074	0.111	0.371	0.133	0.551	0.112	0.373	0.098	0.231	0.081	0.187	0.077	0.164	0.060	0.092	0.119	0.463	0.178	1.015	0.115	0.236	0.064	0.100	0.220	1.614
	192	0.059	0.086	0.118	0.420	0.142	0.632	0.127	0.483	0.111	0.319	0.087	0.202	0.087	0.205	0.060	0.094	0.133	0.569	0.171	0.852	0.140	0.343	0.064	0.100	0.229	1.728
	336	0.062	0.095	0.108	0.277	0.151	0.699	0.139	0.572	0.119	0.373	0.091	0.213	0.094	0.238	0.061	0.096	0.145	0.659	0.199	1.147	0.169	0.494	0.064	0.100	0.208	1.411
	720	0.068	0.112	0.106	0.260	0.162	0.778	0.151	0.645	0.134	0.493	0.231	1.734	0.091	0.213	0.062	0.098	0.158	0.741	0.195	1.074	0.182	0.565	0.064	0.100	0.209	1.387
	Avg	0.061	0.092	0.111	0.332	0.147	0.665	0.132	0.518	0.115	0.354	0.122	0.584	0.087	0.205	0.061	0.095	0.139	0.608	0.186	1.022	0.151	0.409	0.064	0.100	0.216	1.535
ETTm2	96	0.092	0.179	0.165	0.714	0.204	1.149	0.169	0.752	0.166	0.623	0.116	0.376	0.132	0.432	0.100	0.252	0.181	0.979	0.257	2.011	0.182	0.598	0.104	0.267	0.247	1.604
	192	0.102	0.219	0.192	0.977	0.228	1.520	0.197	1.057	0.188	0.792	0.125	0.456	0.152	0.604	0.102	0.296	0.211	1.332	0.266	2.181	0.229	0.917	0.104	0.269	0.239	1.329
	336	0.104	0.229	0.224	1.391	0.252	1.930	0.223	1.342	0.200	0.906	0.134	0.526	0.166	0.866	0.105	0.362	0.237	1.745	0.297	2.704	0.273	1.269	0.105	0.270	0.243	1.501
	720	0.069	0.113	0.216	1.269	0.287	2.569	0.250	1.745	0.264	1.780	0.134	0.526	0.181	1.035	0.111	0.541	0.274	2.379	0.327	3.418	0.301	1.514	0.105	0.270	0.260	1.786
	Avg	0.092	0.185	0.200	1.088	0.243	1.792	0.210	1.224	0.205	1.025	0.127	0.471	0.158	0.734	0.104	0.363	0.226	1.609	0.287	2.579	0.246	1.074	0.104	0.269	0.247	1.555
Weather	96	0.253	0.933	0.354	0.225	0.465	0.303	0.253	0.137	0.254	0.146	0.266	0.133	0.303	0.270	0.307	0.159	0.423	0.289	0.618	0.937	0.338	0.592	0.379	0.224	0.432	0.303
	192	0.252	0.957	0.355	0.229	0.472	0.317	0.257	0.143	0.258	0.153	0.263	0.136	0.304	0.204	0.309	0.160	0.437	0.307	0.817	0.875	0.261	0.523	0.377	0.229	0.453	0.308
	336	0.253	0.956	0.356	0.230	0.481	0.326	0.266	0.146	0.264	0.156	0.270	0.137	0.315	0.164	0.313	0.161	0.452	0.318	0.791	0.888	0.274	0.791	0.377	0.232	0.472	0.309
	720	0.253	0.945	0.362	0.237	0.492	0.336	0.288	0.146	0.271	0.160	0.266	0.135	0.325	0.130	0.313	0.163	0.469	0.329	0.851	0.906	0.293	1.087	0.377	0.235	0.481	0.310
	Avg	0.253	0.948	0.357	0.230	0.478	0.321	0.266	0.143	0.262	0.154	0.266	0.135	0.312	0.192	0.311	0.161	0.445	0.311	0.769	0.902	0.292	0.748	0.378	0.230	0.460	0.308
Exchange	96	0.061	0.082	0.198	0.114	0.218	0.137	0.152	0.137	0.180	0.144	0.062	0.094	0.151	0.477	0.088	0.185	0.193	0.109	0.248	0.257	0.233	0.253	0.068	0.116	0.728	0.709
	192	0.063	0.090	0.278	0.205	0.286	0.235	0.195	0.137	0.232	0.153	0.070	0.127	0.180	0.696	0.092	0.194	0.270	0.199	0.324	0.343	0.290	0.269	0.071	0.122	0.727	0.708
	336	0.066	0.107	0.356	0.367	0.380	0.397	0.247	0.129	0.261	0.175	0.080	0.169	0.209	0.984	0.095	0.206	0.357	0.343	0.422	0.484	0.378	0.274	0.072	0.123	0.726	0.705
	720	0.073	0.123	0.553	0.871	0.559	0.762	0.300	0.126	0.292	0.185	0.103	0.272	0.259	1.543	0.101	0.220	0.519	0.664	0.621	0.465	0.339	0.301	0.072	0.121	0.732	0.710
	Avg	0.066	0.101	0.346	0.389	0.361	0.383	0.224	0.132	0.241	0.164	0.079	0.166	0.200	0.925	0.094	0.201	0.335	0.329	0.404	0.387	0.310	0.274	0.071	0.121	0.728	0.708

5.2 Main Results.

FenDiM achieves the best dataset-averaged MSE on five of the six benchmarks and the best or tied-best dataset-averaged MAE on five benchmarks (Table 1). Relative to the strongest competing baseline for each dataset and metric, FenDiM reduces averaged MSE by 38.5% on ETTh2, 31.2% on ETTm2, and 14.0% on ETTh1, while reducing averaged MAE by 11.5% on ETTm2. On Weather, it achieves the best averaged MAE, improving over DPEM by 3.4%, but does not improve MSE. On Exchange, it reduces averaged MAE by 7.0% over the strongest baseline. These results show broad forecasting gains, with an MAE–MSE trade-off on Weather.

We further compare FenDiM with representative baselines to clarify the source of its gains. Compared with observation-driven forecasters, FenDiM shows that stronger dependency extraction from observed histories is not sufficient when the observed relations are biased or unstable. Compared with reconstruction-oriented methods, its advantage suggests that recovering latent sequence structures does not necessarily yield forecasting-useful dependencies. These results support the effectiveness of prediction-oriented dependency rectification, through which FenDiM suppresses biased temporal relations and aligns the corrected representation with forecasting utility.

Table 2: Robustness comparison under four corrupted settings: missing rates 30%, 50%, 70%, and Gaussian noise, with $L = 96$ and $T = 192$. The best and second-best results are highlighted in red and blue.

Model		FenDiM (Ours)		TimeMosaic (2026)		SparseTSF (2025)		SimpleTM (2025)		DPEM (2025)		xPatch (2025)		AMD (2025)		DeepEDM (2025)		PDF (2024)		iTransformer (2023)		BiTGraph (2023)		SSSD (2023)		Autoformer (2021)	
Dataset	Robustness Setting	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
ETTh1	Missing 30%	0.114	0.311	0.190	0.990	0.193	1.003	0.193	0.959	0.187	0.880	0.197	1.038	0.159	0.678	0.267	1.438	0.189	0.963	0.219	1.417	0.267	1.438	0.114	0.383	0.258	1.740
	Missing 50%	0.133	0.424	0.230	1.472	0.225	1.393	0.203	1.130	0.219	1.238	0.231	1.461	0.214	1.267	0.224	1.374	0.224	1.374	0.241	1.686	0.264	1.482	0.139	0.316	0.278	2.029
	Missing 70%	0.190	0.861	0.265	1.956	0.265	1.957	0.220	1.464	0.259	1.773	0.231	1.461	0.276	2.157	0.263	1.844	0.266	1.974	0.270	2.079	0.281	1.735	0.202	2.845	0.302	2.426
	Gaussian Noise	0.124	0.255	0.136	0.423	0.167	0.755	0.206	0.975	0.157	0.657	0.166	0.775	0.950	25.999	0.109	0.247	0.162	0.708	0.205	1.180	0.340	3.151	0.135	0.336	0.229	1.434
ETTh2	Missing 30%	0.228	1.078	0.385	3.856	0.368	3.857	0.421	5.578	0.348	2.914	0.438	6.593	0.308	2.371	0.397	5.148	0.350	3.551	0.469	6.484	0.461	5.205	0.242	1.140	0.492	6.477
	Missing 50%	0.263	1.434	0.478	7.083	0.427	5.084	0.400	4.868	0.416	4.289	0.438	6.593	0.376	3.603	0.465	7.450	0.416	4.887	0.487	7.171	0.465	5.485	0.212	1.056	0.537	7.585
	Missing 70%	0.354	2.757	0.545	9.251	0.496	6.836	0.386	4.293	0.496	6.326	0.569	9.544	0.559	8.328	0.496	6.820	0.489	6.664	0.517	7.961	0.494	6.262	0.369	1.265	0.590	8.963
	Gaussian Noise	0.238	1.714	0.274	1.951	0.311	2.953	0.470	7.045	0.286	2.008	0.360	4.542	4.330	38.028	0.187	0.993	0.288	2.593	0.457	5.511	0.655	10.423	0.413	5.321	0.386	3.909
ETTm1	Missing 30%	0.071	0.120	0.172	0.923	0.283	2.201	0.328	3.597	0.182	0.891	0.197	1.240	0.142	0.585	0.185	1.089	0.170	0.871	0.197	1.180	0.248	1.541	0.299	2.907	0.255	2.032
	Missing 50%	0.090	0.193	0.211	1.417	0.312	2.689	0.321	3.621	0.216	1.280	0.235	1.824	0.193	1.154	0.206	1.314	0.204	1.279	0.217	1.451	0.269	1.759	0.299	2.904	0.272	2.338
	Missing 70%	0.149	0.552	0.251	2.054	0.352	3.449	0.310	3.535	0.251	1.826	0.264	2.320	0.248	2.011	0.242	1.853	0.244	1.879	0.246	1.890	0.246	1.890	0.298	2.904	0.309	3.088
	Gaussian Noise	0.107	0.637	0.113	0.661	0.147	0.640	0.352	3.591	0.133	0.490	0.143	0.682	2.259	96.975	0.101	0.219	0.139	0.586	0.181	0.972	0.314	2.779	0.475	4.483	0.284	1.996
ETTm2	Missing 30%	0.119	0.295	0.237	1.507	0.261	1.940	0.382	3.827	0.263	1.747	0.253	1.830	0.219	1.214	0.280	2.269	0.251	1.830	0.303	2.782	0.365	2.640	0.124	0.299	0.389	4.119
	Missing 50%	0.145	0.440	0.285	2.216	0.299	2.528	0.367	3.515	0.306	2.462	0.303	2.636	0.278	2.070	0.310	2.762	0.295	2.468	0.331	3.221	0.370	2.770	0.186	0.474	0.395	3.994
	Missing 70%	0.233	1.238	0.340	3.215	0.346	3.361	0.349	3.212	0.353	3.383	0.350	3.480	0.346	3.312	0.353	3.383	0.348	3.426	0.364	3.805	0.398	3.213	0.292	1.412	0.501	6.862
	Gaussian Noise	0.158	0.510	0.189	0.882	0.229	1.527	0.419	4.478	0.211	1.035	0.203	1.084	1.241	41.493	0.167	0.435	0.213	1.373	0.274	2.341	0.484	5.620	0.269	2.339	0.328	2.704
Weather	Missing 30%	0.252	0.961	0.604	3.846	0.608	3.982	0.280	1.660	0.314	1.866	0.333	2.099	0.341	16.000	0.709	5.061	0.588	3.927	1.195	8.781	0.722	4.341	0.377	2.287	0.499	3.230
	Missing 50%	0.259	0.970	0.775	5.490	0.735	5.321	0.335	2.099	0.374	2.461	0.405	2.863	0.398	14.310	0.898	6.253	0.722	5.285	1.180	8.834	0.672	4.346	0.428	2.409	0.519	3.659
	Missing 70%	0.263	0.987	0.945	7.707	0.882	7.301	0.400	2.824	0.445	3.374	0.473	3.845	0.471	7.039	1.023	7.300	0.880	7.303	1.155	8.931	0.642	4.634	0.775	7.087	0.549	4.281
	Gaussian Noise	0.257	0.952	0.355	2.293	0.472	3.175	0.264	1.427	0.259	1.526	0.275	1.717	0.980	3.937	0.309	1.467	0.438	3.074	1.217	8.750	0.724	5.449	0.260	7.565	0.492	3.066
Exchange	Missing 30%	0.063	0.090	0.429	0.431	0.409	0.409	0.711	1.989	0.352	0.270	0.493	0.658	0.411	0.383	0.327	0.284	0.406	0.392	0.433	0.416	0.616	0.695	0.077	0.115	0.728	0.791
	Missing 50%	0.063	0.089	0.616	0.949	0.519	0.625	0.637	0.881	0.479	0.635	0.590	0.906	0.574	0.755	0.482	0.567	0.532	0.633	0.524	0.577	0.604	0.630	0.070	0.121	0.827	0.831
	Missing 70%	0.095	0.170	0.901	2.218	0.647	0.901	0.543	0.749	0.625	0.827	0.672	1.095	0.711	1.112	0.619	0.869	0.665	0.943	0.646	0.845	0.687	0.768	0.077	0.237	1.302	3.294
	Gaussian Noise	0.060	0.082	1.520	3.690	0.280	0.131	1.406	2.773	0.193	0.548	0.270	0.134	1.383	7.631	0.678	0.695	0.309	0.150	0.143	0.318	0.803	0.683	0.076	0.099	0.757	1.061

5.3 Robustness Analysis

We evaluate missing observations, Gaussian noise, and compound corruptions using $L = 96$ and $T \in \{96, 192, 336, 720\}$. Table 2 summarizes results at $T = 192$; full missing and Gaussian-noise results appear in Appendix C, and compound-corruption results in Appendix D. Unless otherwise stated, the comparisons below use Table 2. Relative reductions are computed as $(e_{\text{base}} - e_{\text{FenDiM}})/e_{\text{base}} \times 100\%$, where e_{base} is the lowest error among competing baselines for the same dataset, setting, and metric.

Robustness under Missing Observations. Across the 18 dataset–missing-rate settings, FenDiM ranks within the top two in both MAE and MSE, attaining the best or tied-best MAE in 16 cases and the best MSE in 15 cases. At 30% missingness, it achieves the best or tied-best MAE and the best MSE on all six datasets, with the MAE on ETTh1 tied with SSSD. At 50% missingness, it achieves the best MAE on five datasets and the best MSE on four; at 70%, it achieves the best MAE and MSE on five datasets each. Representative reductions include 5.8% in MAE and 5.4% in MSE on ETTh2 at 30% missingness; the largest MAE reduction is 53.4% on ETTm1 at 50% missingness. These results support robust forecasting from incomplete observations across the tested missing rates.

Robustness under Noisy Observations. With additive Gaussian noise, FenDiM achieves the best MAE on three of six datasets: ETTm2, Weather, and Exchange, with reductions of 5.4%, 0.8%, and 21.1%, respectively, over the strongest competing baseline. It ranks within the top two in MAE on all six datasets and in MSE on five, achieving the best MSE on Weather and Exchange. FenDiM ranks second in both metrics on ETTh1 and ETTh2; on ETTm1, it ranks second in MAE but fourth in MSE. Thus, its noise robustness varies across datasets and error metrics, with the clearest joint MAE/MSE gains on Weather and Exchange.

5.4 Ablation Study.

We evaluate FenDiM's key components on ETTh1, ETTh2, Exchange, and Weather using three variants: *w/o PDR*, *w/o frequency-temporal anchoring*, and *w/o frequency backbone*. Besides MAE/MSE trajectories, we report iAUC, computed by trapezoidal integration over the error–epoch curve, where lower values indicate smaller accumulated errors under the same training budget.

As shown in Figure 2, FenDiM consistently obtains lower error trajectories and smaller iAUC values than its variants. The degradation of *w/o PDR* confirms the importance of rectifying biased observed dependencies before forecasting. The drop caused by removing frequency-temporal anchoring, especially on ETTh2, Exchange, and Weather, shows that frequency-domain compensation helps

preserve global temporal structure and reduce error accumulation. Finally, *w/o frequency backbone* verifies the necessity of temporal encoding for modeling sequential evolution. Together, these results demonstrate that FenDiM's gains come from the coordinated design of dependency rectification, frequency-temporal anchoring, and temporal representation learning.

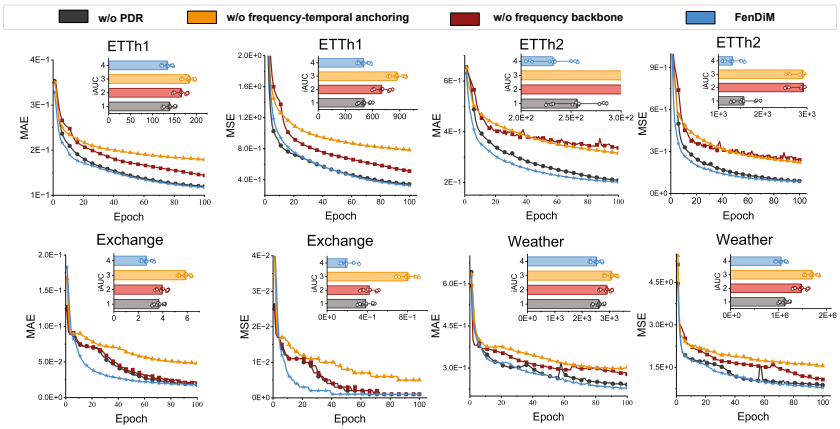

Figure 2: Ablation study of FenDiM across ETTh1, ETTh2, Exchange, and Weather, with epoch-wise MAE/MSE trajectories and corresponding iAUC summaries.

5.5 Sensitivity and Efficiency Analysis

We further evaluate FenDiM from the perspectives of hyperparameter sensitivity and computational efficiency, with full results provided in Appendix D.1 and Appendix E. FenDiM remains stable across a broad range of λ_{pdr} and λ_{spec}, suggesting that its gains do not rely on aggressive tuning. Efficiency results further show that the Mamba-based denoiser alleviates the overhead of diffusion-based rectification and keeps FenDiM practical for long-horizon forecasting.

6 Conclusion

We present FenDiM, a Frequency-Enhanced Diffusion-Mamba architecture for reliable time series forecasting under imperfect observations. FenDiM moves beyond observation-driven dependency extraction by formulating forecasting as prediction-oriented dependency rectification, where corruption-induced temporal relations are progressively suppressed, and forecasting-relevant dependencies are preserved. This formulation is instantiated through diffusion-based progressive rectification, frequency-domain structural compensation, and efficient Mamba-based temporal modeling. Across six benchmarks, FenDiM reduces horizon-averaged MSE and MAE by up to 38.5% and 11.5% under standard observations, and shows competitive robustness under missing, noisy, and compound-corrupted inputs. These results highlight dependency rectification as a general principle for reliable forecasting under both standard and imperfect observations.

AI Use Statement

Generative AI tools were used for language polishing and improving the readability of the manuscript. All AI-assisted edits were reviewed by the authors.

Reproducibility Statement

We provide implementation details, preprocessing and corruption protocols, and hyperparameter settings in the Appendix. Anonymous code and datasets are provided to facilitate reproducibility.

REFERENCES

Juan Miguel Lopez Alcaraz and Nils Strodthoff. Diffusion-based time series imputation and forecasting with structured state space models. *arXiv preprint arXiv:2208.09399*, 2022.

Zhongde An, Jinhong You, Jiyanglin Li, Yiming Tang, Wen Li, Heming Du, and Shouguo Du. Fredn: Spectral disentanglement for time series forecasting via learnable frequency decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 19623–19631, 2026.

Wei Cao, Dong Wang, Jian Li, Hao Zhou, Lei Li, and Yitan Li. Brits: Bidirectional recurrent imputation for time series. *Advances in neural information processing systems*, 31, 2018.

Hui Chen, Viet Luong, Lopamudra Mukherjee, and Vikas Singh. Simpletm: A simple baseline for multivariate time series forecasting. In *The Thirteenth International Conference on Learning Representations*, 2025.

Xiaodan Chen, Xiucheng Li, Bo Liu, and Zhijun Li. Biased temporal convolution graph network for time series forecasting with missing values. In *The Twelfth International Conference on Learning Representations*, 2023.

Tao Dai, Beiliang Wu, Peiyuan Liu, Naiqi Li, Jigang Bao, Yong Jiang, and Shu-Tao Xia. Periodicity decoupling framework for long-term series forecasting. In *The twelfth international conference on learning representations*, 2024.

Kuiye Ding, Fanda Fan, Chunyi Hou, Zheya Wang, Lei Wang, Zhengxin Yang, and Jianfeng Zhan. Timemosaic: Temporal heterogeneity guided time series forecasting via adaptive granularity patch and segment-wise decoding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 20790–20798, 2026.

Wenjie Du, David Côté, and Yan Liu. Saits: Self-attention-based imputation for time series. *Expert Systems with Applications*, 219:119619, 2023.

Anass El Aouni, Quentin Gaudel, Juan Emmanuel Johnson, REGNIER Charly, Julien Le Sommer, Ronan Fablet, Marie Drevillon, Yann Drillet, Pierre Yves Le Traon, et al. Oceanbench: A benchmark for data-driven global ocean forecasting systems. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2025.

Yisong Fu, Zezhi Shao, Chengqing Yu, Yujie Li, Zhulin An, Qi Wang, Yongjun Xu, and Fei Wang. Selective learning for deep time series forecasting. *arXiv preprint arXiv:2510.25207*, 2025.

Sicheng He, Junzhong Ji, and Minglong Lei. Decomposed spatio-temporal mamba for long-term traffic prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 11772–11780, 2025.

Rui Hou, Qiao Liu, Peng He, Yao Liu, Yanwen Huang, Jun Xie, Yang Xie, and Tingting Dai. Dpem: Dual-perspective enhanced mamba for multivariate time series forecasting. *Information Fusion*, 123:103250, 2025.

Yifan Hu, Peiyuan Liu, Peng Zhu, Dawei Cheng, and Tao Dai. Adaptive multi-scale decomposition framework for time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 17359–17367, 2025.

Qiusheng Huang, Yuan Niu, Xiaohui Zhong, Anboyu Guo, Lei Chen, Dianjun Zhang, Xuefeng Zhang, and Hao Li. Fuxi-ocean: A global ocean forecasting system with sub-daily resolution. *arXiv preprint arXiv:2506.03210*, 2025a.

Songtao Huang, Zhen Zhao, Can Li, and Lei Bai. Timekan: Kan-based frequency decomposition learning architecture for long-term time series forecasting. *arXiv preprint arXiv:2502.06910*, 2025b.

Mourad Khayati, Guillaume Chacun, Zakhar Tymchenko, and Philippe Cudré-Mauroux. A-darts: stable model selection for data repair in time series. In *2025 IEEE 41st International Conference on Data Engineering (ICDE)*, pp. 2009–2023. IEEE, 2025.

Renzhao Liang, Sizhe Xu, Chenggang Xie, Jingru Chen, Feiyang Ren, Shu Yang, and Takahiro Yabe. Abstain mask retain core: Time series prediction by adaptive masking loss with representation consistency. *arXiv preprint arXiv:2510.19980*, 2025.

Shengsheng Lin, Weiwei Lin, Wentai Wu, Haojun Chen, and CL Philip Chen. Sparsetsf: Lightweight and robust time series forecasting via sparse modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

Mingzhe Liu, Han Huang, Hao Feng, Leilei Sun, Bowen Du, and Yanjie Fu. Pristi: A conditional diffusion framework for spatiotemporal imputation. In *2023 IEEE 39th international conference on data engineering (ICDE)*, pp. 1927–1939. IEEE, 2023a.

Xvyuan Liu, Xiangfei Qiu, Xingjian Wu, Zhengyu Li, Chenjuan Guo, Jilin Hu, and Bin Yang. Rethinking irregular time series forecasting: A simple yet effective baseline. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 23873–23881, 2026.

Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. itransformer: Inverted transformers are effective for time series forecasting. *arXiv preprint arXiv:2310.06625*, 2023b.

Zhiding Liu, Mingyue Cheng, Guanhao Zhao, Jiqian Yang, Qi Liu, and Enhong Chen. Improving time series forecasting via instance-aware post-hoc revision. *arXiv preprint arXiv:2505.23583*, 2025.

Junkai Lu, Peng Chen, Chenjuan Guo, Yang Shu, Meng Wang, and Bin Yang. Towards non-stationary time series forecasting with temporal stabilization and frequency differencing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 24070–24078, 2026.

Donghao Luo and Xue Wang. Decompnet: Enhancing time series forecasting models with implicit decomposition. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Abrar Majeedi, Viswanatha Reddy Gajjala, Satya Sai Srinath Namburi Gnvv, Nada Magdi Elkordi, and Yin Li. Lets forecast: Learning embedology for time series forecasting. *arXiv preprint arXiv:2506.06454*, 2025.

Max Moebus, Björn Braun, and Christian Holz. Contimask: Explaining irregular time series via perturbations in continuous time. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Kwangryeol Park, Jaeho Kim, and Seulki Lee. Alio: Output alignment matters in long-term time series forecasting. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

Xiangfei Qiu, Xingjian Wu, Yan Lin, Chenjuan Guo, Jilin Hu, and Bin Yang. Duet: Dual clustering enhanced multivariate time series forecasting. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pp. 1185–1196, 2025.

Kashif Rasul, Calvin Seward, Ingmar Schuster, and Roland Vollgraf. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting. In *International conference on machine learning*, pp. 8857–8868. PMLR, 2021.

Zineb Senane, Lele Cao, Valentin Leonhard Buchner, Yusuke Tashiro, Lei You, Pawel Andrzej Herman, Mats Nordahl, Ruibo Tu, and Vilhelm Von Ehrenheim. Self-supervised learning of time series representation via diffusion process and imputation-interpolation-forecasting mask. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*, pp. 2560–2571, 2024.

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.

Yu Sun, Xinyu Yang, Shaoxu Song, Ying Zhang, and Xiaojie Yuan. Collaborative imputation for multivariate time series with convergence guarantee. In *2025 IEEE 41st International Conference on Data Engineering (ICDE)*, pp. 1111–1124. IEEE, 2025.

Yusuke Tashiro, Jiaming Song, Yang Song, and Stefano Ermon. Csdi: Conditional score-based diffusion models for probabilistic time series imputation. *Advances in neural information processing systems*, 34:24804–24816, 2021.

Hao Wang, Licheng Pan, Zhichao Chen, Xu Chen, Qingyang Dai, Lei Wang, Haoxuan Li, and Zhouchen Lin. Time-o1: Time-series forecasting needs transformed label alignment. *arXiv preprint arXiv:2505.17847*, 2025a.

Xu Wang, Hongbo Zhang, Pengkun Wang, Yudong Zhang, Binwu Wang, Zhengyang Zhou, and Yang Wang. An observed value consistent diffusion model for imputing missing values in multivariate time series. In *Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining*, pp. 2409–2418, 2023.

Yihang Wang, Yuying Qiu, Peng Chen, Yang Shu, Zhongwen Rao, Lujia Pan, Bin Yang, and Chenjuan Guo. Lightgts: A lightweight general time series forecasting model. *arXiv preprint arXiv:2506.06005*, 2025b.

Di Wen, Yu Wang, Zhigang Wu, Zhaocheng He, Zhe Wu, and Zheng Qingfang. A driving-style-adaptive framework for vehicle trajectory prediction. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in neural information processing systems*, 34:22419–22430, 2021.

Xingjian Wu, Xiangfei Qiu, Hanyin Cheng, Zhengyu Li, Jilin Hu, Chenjuan Guo, and Bin Yang. Enhancing time series forecasting through selective representation spaces: A patch perspective. *arXiv preprint arXiv:2510.14510*, 2025.

Jingyi Xu, Shengnan Wang, Weidong Yang, Keyi Liu, Yeqi Luo, Ben Fei, and Lei Bai. Sifusion: A unified fusion framework for multi-granularity arctic sea ice forecasting. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Vijaya Krishna Yalavarthi, Randolf Scholz, Stefan Born, and Lars Schmidt-Thieme. Probabilistic forecasting of irregularly sampled time series with missing values via conditional normalizing flows. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 21877–21885, 2025.

Mingjie You, Dawei Cheng, Meilin Zhang, Peng Zhu, and Yuqi Liang. Delay-aware graph neural stochastic differential equations for financial time series modeling and forecasting. In *Proceedings of the ACM Web Conference 2026*, pp. 7967–7978, 2026.

Chengqing Yu, Fei Wang, Chuanguang Yang, Zezhi Shao, Tao Sun, Tangwen Qian, Wei Wei, Zhulin An, and Yongjun Xu. Merlin: Multi-view representation learning for robust multivariate time series forecasting with unfixed missing rates. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pp. 3633–3644, 2025.

Jiahui Zhang, Zhengyang Zhou, Wenjie Du, and Yang Wang. Enhancing the maximum effective window for long-term time series forecasting. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025a.

Xu Zhang, Zhengang Huang, Yunzhi Wu, Xun Lu, Erpeng Qi, Yunkai Chen, Zhongya Xue, Qitong Wang, Peng Wang, and Wei Wang. Multi-period learning for financial time series forecasting. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pp. 2848–2859, 2025b.

Xukai Zhang, Shuliang Wang, Guangyin Jin, Ziqiang Yuan, Hanning Yuan, and Sijie Ruan. Eulerian neural network informed by chemical transport for air quality forecasting. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025c.

Yongning Zhao, Shiji Pan, Yanxu Chen, Haohan Liao, Yingying Zheng, and Lin Ye. Intraday wind power forecasting by ensemble of overlapping historical numerical weather predictions. *IEEE Transactions on Sustainable Energy*, 16(2):1315–1328, 2024.

Zinan Zheng, Yang Liu, and Jia Li. Mesh interpolation graph network for dynamic and spatially irregular global weather forecasting. *arXiv preprint arXiv:2509.20911*, 2025.

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.

Yuanshao Zhu, James Jianqiao Yu, Xiangyu Zhao, Xun Zhou, Liang Han, Xuetao Wei, and Yuxuan Liang. Unitraj: Learning a universal trajectory foundation model from billion-scale worldwide traces. *arXiv preprint arXiv:2411.03859*, 2024.

Zhuohang Zhu, Haodong Chen, Qiang Qu, and Vera Chung. Fincast: A foundation model for financial time-series forecasting. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*, pp. 4539–4549, 2025.

Table 3: Summary of datasets used in the experiments.

Datasets	Dimension	Forecasting Horizons	Dataset Size	Frequency	Domain
ETTm1	7	{96,192,336,720}	(34465,11521,11521)	15mins	Electricity
ETTm2	7	{96,192,336,720}	(34465,11521,11521)	15mins	Electricity
ETTh1	7	{96,192,336,720}	(8545,2881,2881)	1 hour	Electricity
ETTh2	7	{96,192,336,720}	(8545,2881,2881)	1 hour	Electricity
Weather	21	{96,192,336,720}	(36792, 5271, 10540)	10mins	Weather
Exchange	8	{96,192,336,720}	(5120,665,1422)	Daily	Exchange rate

A DATASET

We evaluate our method on six widely used long-term time-series forecasting benchmarks, as summarized in Table 3. The ETT datasets contain electricity transformer records, including load and oil temperature; ETTh1 and ETTh2 are recorded hourly, while ETTm1 and ETTm2 are recorded every 15 minutes, covering the period from July 2016 to July 2018. The Weather dataset contains 21 meteorological indicators, such as air temperature and humidity, recorded every 10 minutes throughout 2020. The Exchange dataset records daily exchange rates from eight countries between 1990 and 2016. Following the standard long-term forecasting protocol, all datasets are split chronologically into training, validation, and test sets. The ETT datasets use the commonly adopted fixed chronological partitions, while Weather and Exchange follow the 7:1:2 chronological split used in prior benchmark settings. The dataset sizes reported in Table 3 denote the commonly used training/validation/test sequence counts under the standard look-back setting.

B BASELINES

We compare FenDiM with recent state-of-the-art time-series forecasting models from three dependency-learning paradigms.

Observation-driven temporal dependency models.

- **TimeMosaic** Ding et al. (2026) uses adaptive patch mosaics to capture heterogeneous temporal patterns.
- **SparseTSF** Lin et al. (2025) exploits cross-period sparsity for lightweight periodic dependency modeling.
- **SimpleTM** Chen et al. (2025) uses simple token mixing to model temporal and cross-variate dependencies.
- **xPatch** Stitsyuk & Choi (2025) combines exponential decomposition with patch-wise temporal modeling.
- **AMD** Hu et al. (2025) adaptively decomposes and mixes multi-scale temporal components.
- **PDF** Dai et al. (2024) reshapes sequences by periods to model intra- and inter-period dependencies.
- **Autoformer** Wu et al. (2021) couples series decomposition with auto-correlation for periodic pattern discovery.

Correlation-driven multivariate dependency models.

- **iTransformer** Liu et al. (2023b) treats variables as tokens to model inter-variate correlations.
- **DPEM** Hou et al. (2025) uses dual temporal-spatial encoders to capture coupled multivariate dependencies.

Imperfect-observation-aware forecasting models.

- **BiTGraph** Chen et al. (2023) incorporates missingness bias into temporal and graph propagation.

Table 4: Code repositories and dependency-acquisition categories of baseline models.

Model Type	Model	Code Repository
Temporal dependency	TimeMosaic	https://github.com/BenchCouncil/TimeMosaic
	SparseTSF	https://github.com/lss-1138/SparseTSF
	SimpleTM	https://github.com/vsingh-group/SimpleTM
	xPatch	https://github.com/stitsyuk/xPatch
	AMD	https://github.com/TROUBADOUR000/AMD
	PDF	https://github.com/Hank0626/PDF
	Autoformer	https://github.com/thuml/Autoformer
Cross-variate dependency	iTransformer	https://github.com/thuml/iTransformer
	DPEM	https://github.com/HRFarsighted/DPEM
Imperfect-aware forecasting	BiTGraph	https://github.com/chenxiaodanhit/BiTGraph
	DeepEDM	https://github.com/abrarmajeedi/DeepEDM
	SSSD	https://github.com/AI4HealthUOL/SSSD

- **DeepEDM** Majeedi et al. (2025) learns latent delay embeddings for noise-robust dynamical forecasting.
- **SSSD** Alcaraz & Strodthoff (2022) uses conditional diffusion and structured state-space modeling for imputation and forecasting.

The code repositories and dependency-learning categories of these baselines are summarized in Table 4.

B.1 Evaluation Metrics

Following standard long-term forecasting practice, we evaluate prediction accuracy using Mean Absolute Error (MAE) and Mean Squared Error (MSE). For a test set with N forecasting windows, prediction horizon H, and C variables, let $\mathbf{Y}^{(n)} \in \mathbb{R}^{H\times C}$ and $\widehat{\mathbf{Y}}^{(n)} \in \mathbb{R}^{H\times C}$ denote the ground truth and prediction of the n-th sample, respectively. MAE and MSE are defined as

$$\mathrm{MAE} = \frac{1}{NHC}\sum_{n=1}^{N}\sum_{t=1}^{H}\sum_{c=1}^{C}\left|Y_{t,c}^{(n)} - \widehat{Y}_{t,c}^{(n)}\right|. \tag{21}$$

$$\mathrm{MSE} = \frac{1}{NHC}\sum_{n=1}^{N}\sum_{t=1}^{H}\sum_{c=1}^{C}\left(Y_{t,c}^{(n)} - \widehat{Y}_{t,c}^{(n)}\right)^2. \tag{22}$$

MAE reflects the average absolute forecasting error, whereas MSE emphasizes large deviations. Lower values indicate better performance. For each horizon, the metrics are computed over all test windows and all variables. The reported average across horizons is the arithmetic mean over $\{96, 192, 336, 720\}$. For repeated runs, each metric is first computed on the complete test set for each run and then averaged across runs.

B.2 Implementation Details.

FenDiM is implemented in PyTorch, and all experiments are conducted on a single NVIDIA GeForce RTX 5090 GPU. We train the model using the Adam optimizer with an initial learning rate of 1×10^{-4}. The point-wise prediction term $\mathcal{L}_{\mathrm{pred}}$ is implemented as the mean squared error, and the full training objective follows Eq. 19, which jointly optimizes forecasting accuracy, dependency rectification, and spectral consistency. Unless otherwise specified, we set $\lambda_{\mathrm{pdr}} = 0.005$, $\lambda_{\mathrm{spec}} = 0.1$, and $\lambda_{\mathrm{pred}} = 1 - \lambda_{\mathrm{pdr}} - \lambda_{\mathrm{spec}}$. The batch size is set to 64 and the maximum number of training epochs is 100. For all datasets, the look-back window is fixed to 96, and the prediction horizon is selected from $96, 192, 336, 720$. All observed variables in each dataset are used for evaluation. We additionally employ a learnable sequence-index embedding with dimension 6, resulting in an internal feature dimension of $C + 6$, where C denotes the number of observed variables.

For the progressive dependency rectification module, we adopt a mask-based denoising strategy implemented with a Block-Mamba denoiser. During training, random masks are applied to the input sequence with a missing rate of 0.2. The diffusion scheduler contains 1000 diffusion steps, where β linearly increases from 10^{-4} to 0.02. To reduce inference cost, only 2 denoising steps are used during prediction. For the frequency structural anchoring module, we use a dual-branch frequency-domain backbone with hidden dimension 256, 8 attention heads, 4 layers, 16 frequency bands, and a dropout rate of 0.1. Only frequency magnitudes are used in the frequency-domain constraint, encouraging the model to preserve stable frequency structures without over-constraining phase variations.

For fair comparison, all baselines are implemented with their official codebases and recommended configurations when available. We retrain all methods from scratch using the same chronological splits, look-back length, prediction horizons, corruption protocols, and evaluation metrics. Under missing and noisy settings, all methods receive the same corrupted inputs, without additional information beyond their official input interfaces.

B.3 Unified Experimental Settings.

To ensure a fair comparison, all experiments are conducted under unified forecasting protocols. We use a look-back length of $L = 96$ and prediction lengths of $T \in \{96, 192, 336, 720\}$ for all datasets and compared methods. We evaluate four settings: standard observations, missing observations, Gaussian-noisy observations, and compound corruptions. For missing-observation experiments, the input sequence is randomly masked with missing ratios of 30%, 50%, and 70%. For Gaussian-noisy experiments, additive Gaussian noise is injected into the input sequence to simulate measurement perturbations. For Gaussian-noisy experiments, we sample additive noise as $\epsilon_{\ell,c} \sim \mathcal{N}(0, \sigma^2 s_c^2)$, where s_c denotes the standard deviation of variable c computed from the training set. We set $\sigma = 0.1$ for all datasets and apply the same noise-generation protocol to all compared methods. For compound-corruption experiments, additive Gaussian noise is further injected into the masked input sequence, resulting in observations that are both incomplete and distorted. All models are evaluated using the same data splits, corruption protocols, and metrics. We report MAE and MSE as the primary evaluation metrics, with standard-setting results reported in Table 1 and robustness results reported in Tables 5–9.

C Robustness Analysis

Tables 5–7 report the full missing-observation results, with Avg denoting the arithmetic mean over four forecasting horizons. At 30% missingness, FenDiM achieves the best average MAE on five datasets and the best average MSE on all six; on ETTh2, the reductions over the strongest baseline are 4.6% and 4.3%, respectively. At 50% missingness, it achieves the best average MAE on five datasets and the best average MSE on four, including a 5.3% MSE reduction on ETTm2 and a 7.0% MAE reduction on Exchange. At 70% missingness, it achieves the best average MAE and MSE on five datasets each; the corresponding reductions include 5.9% in MAE on ETTh1 and 10.2% in MSE on ETTm2. FenDiM remains within the top two in both horizon-averaged metrics across all 18 dataset–missing-rate settings.

Table 8 reports Gaussian-noise results across all four horizons. FenDiM achieves the best average MAE on ETTm2, Weather, and Exchange, with reductions of 5.9%, 1.5%, and 21.3% over the strongest competing baseline, respectively, and ranks second on the remaining datasets. It also achieves the best average MSE on ETTm2, Weather, and Exchange and ranks second on ETTh1 and ETTh2, but ranks fifth on ETTm1. These results indicate competitive noise robustness while identifying ETTm1 MSE as a remaining weakness.

D Robustness under compound corruptions.

Real-world observation errors rarely appear in isolation. We therefore evaluate a more challenging setting where 30% of the input observations are masked and Gaussian noise is simultaneously added to the remaining inputs. As shown in Table 9, FenDiM remains within the top two in all dataset-level average MAE and MSE comparisons, achieving the best average MAE on five of the six datasets

Table 5: Performance comparison under 30% missing observations with $L = 96$ and $T \in \{96, 192, 336, 720\}$ on multivariate forecasting benchmarks. The best and second-best results are highlighted in red and blue.

Model		FenDiM (Ours)		TimeMosaic (2026)		SparseTSF (2025)		SimpleTM (2025)		DPEM (2025)		xPatch (2025)		AMD (2025)		DeepEDM (2025)		PDF (2024)		iTransformer (2023)		BiTGraph (2023)		SSSD (2023)		Autoformer (2021)	
Metric		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
ETTh1	96	0.114	0.308	0.184	0.943	0.188	0.968	0.195	0.956	0.186	0.877	0.186	0.950	0.155	0.656	0.259	1.378	0.182	0.913	0.214	1.379	0.259	1.378	0.113	0.379	0.255	1.696
	192	0.114	0.311	0.190	0.990	0.193	1.003	0.193	0.959	0.187	0.880	0.197	1.038	0.159	0.678	0.267	1.438	0.189	0.963	0.219	1.417	0.267	1.438	0.114	0.383	0.258	1.740
	336	0.115	0.312	0.196	1.044	0.199	1.051	0.192	0.963	0.188	0.891	0.207	1.114	0.164	0.706	0.278	1.525	0.196	1.015	0.226	1.472	0.278	1.525	0.115	0.387	0.263	1.797
	720	0.116	0.319	0.213	1.201	0.213	1.175	0.189	0.951	0.194	0.935	0.229	1.316	0.174	0.768	0.313	1.832	0.212	1.144	0.240	1.600	0.313	1.832	0.115	0.392	0.275	1.931
	Avg	0.115	0.312	0.196	1.044	0.198	1.049	0.192	0.957	0.189	0.896	0.205	1.104	0.163	0.702	0.279	1.543	0.195	1.009	0.225	1.467	0.279	1.543	0.114	0.385	0.263	1.791
ETTh2	96	0.229	1.082	0.349	3.398	0.342	3.314	0.424	5.612	0.338	2.754	0.366	3.838	0.290	2.076	0.351	3.568	0.322	2.993	0.444	5.956	0.431	4.660	0.232	1.095	0.497	6.600
	192	0.228	1.078	0.385	3.856	0.368	3.857	0.421	5.578	0.348	2.914	0.438	6.593	0.308	2.371	0.397	5.148	0.350	3.551	0.469	6.484	0.461	5.205	0.242	1.140	0.492	6.477
	336	0.228	1.076	0.407	4.749	0.395	4.506	0.415	5.507	0.357	3.050	0.537	12.960	0.323	2.617	0.444	7.275	0.380	4.221	0.492	7.049	0.502	5.982	0.243	1.141	0.501	6.768
	720	0.230	1.088	0.464	6.004	0.455	5.857	0.394	5.191	0.378	3.344	0.708	26.380	0.360	3.243	0.589	19.420	0.444	5.611	0.539	7.994	0.605	8.124	0.243	1.145	0.540	7.771
	Avg	0.229	1.081	0.401	4.502	0.390	4.383	0.414	5.472	0.355	3.016	0.512	12.443	0.320	2.577	0.445	8.852	0.374	4.094	0.486	6.871	0.500	5.993	0.240	1.130	0.508	6.904
ETTm1	96	0.071	0.121	0.163	0.837	0.266	1.881	0.330	3.603	0.172	0.825	0.177	0.985	0.137	0.559	0.171	0.914	0.159	0.789	0.191	1.128	0.217	1.325	0.298	2.903	0.269	2.290
	192	0.071	0.120	0.172	0.923	0.283	2.201	0.328	3.597	0.182	0.891	0.197	1.240	0.142	0.585	0.185	1.089	0.170	0.871	0.197	1.180	0.248	1.541	0.299	2.907	0.255	2.032
	336	0.071	0.120	0.183	1.028	0.303	2.586	0.327	3.626	0.190	0.948	0.220	1.608	0.145	0.602	0.201	1.319	0.178	0.940	0.203	1.230	0.284	1.872	0.300	2.914	0.249	1.921
	720	0.071	0.120	0.197	1.153	0.334	3.191	0.326	3.629	0.194	0.974	0.245	1.920	0.150	0.624	0.221	1.621	0.188	1.015	0.210	1.287	0.303	1.993	0.302	2.908	0.248	1.861
	Avg	0.071	0.120	0.179	0.985	0.296	2.465	0.328	3.614	0.184	0.910	0.210	1.438	0.144	0.592	0.195	1.236	0.174	0.904	0.200	1.206	0.263	1.683	0.300	2.908	0.255	2.026
ETTm2	96	0.120	0.296	0.217	1.221	0.243	1.616	0.384	3.813	0.246	1.526	0.234	1.518	0.207	1.068	0.255	1.865	0.228	1.462	0.287	2.493	0.315	1.981	0.124	0.297	0.397	4.279
	192	0.119	0.295	0.237	1.507	0.261	1.940	0.382	3.827	0.263	1.747	0.253	1.830	0.219	1.214	0.280	2.269	0.251	1.830	0.303	2.782	0.365	2.640	0.124	0.299	0.389	4.119
	336	0.119	0.296	0.257	1.846	0.280	2.305	0.380	3.869	0.281	1.976	0.278	2.288	0.232	1.390	0.303	2.731	0.273	2.221	0.320	3.115	0.419	3.471	0.125	0.300	0.384	3.995
	720	0.120	0.298	0.289	2.371	0.310	2.901	0.372	3.738	0.310	2.359	0.328	3.308	0.255	1.746	0.341	3.549	0.308	2.854	0.347	3.699	0.480	4.555	0.125	0.310	0.403	4.512
	Avg	0.119	0.296	0.250	1.736	0.273	2.191	0.380	3.812	0.275	1.902	0.273	2.236	0.228	1.355	0.295	2.604	0.265	2.092	0.314	3.022	0.395	3.162	0.124	0.302	0.393	4.226
Weather	96	0.253	0.935	0.628	3.932	0.637	4.018	0.281	1.547	0.326	1.876	0.342	2.153	0.353	21.621	0.709	4.721	0.611	3.947	1.285	9.410	0.759	4.652	0.379	2.243	0.524	3.445
	192	0.252	0.961	0.604	3.846	0.608	3.982	0.280	1.660	0.314	1.866	0.333	2.099	0.341	16.000	0.709	5.061	0.588	3.927	1.195	8.781	0.722	4.341	0.377	2.287	0.499	3.230
	336	0.253	0.960	0.620	3.957	0.619	4.076	0.282	1.717	0.320	1.903	0.342	2.124	0.353	12.642	0.725	5.250	0.603	4.037	1.210	8.912	0.720	4.446	0.377	2.317	0.500	3.266
	720	0.253	0.946	0.639	4.065	0.632	4.178	0.284	1.744	0.327	1.944	0.352	2.117	0.367	9.884	0.741	5.732	0.619	4.147	1.228	9.092	0.724	4.574	0.377	2.353	0.506	3.325
	Avg	0.253	0.950	0.623	3.950	0.624	4.064	0.281	1.667	0.322	1.897	0.342	2.123	0.354	15.037	0.721	5.191	0.605	4.015	1.230	9.049	0.731	4.503	0.378	2.300	0.507	3.316
Exchange	96	0.061	0.083	0.364	0.325	0.348	0.348	0.698	1.043	0.335	0.243	0.417	0.482	0.365	0.305	0.299	0.236	0.341	0.283	0.373	0.319	0.577	0.594	0.055	0.110	0.736	0.797
	192	0.063	0.090	0.429	0.431	0.409	0.409	0.711	1.989	0.352	0.270	0.493	0.658	0.411	0.383	0.327	0.284	0.406	0.392	0.433	0.416	0.616	0.695	0.077	0.115	0.728	0.791
	336	0.066	0.099	0.510	0.619	0.493	0.493	0.716	1.149	0.377	0.314	0.653	1.258	0.464	0.493	0.364	0.353	0.490	0.551	0.513	0.565	0.681	0.872	0.083	0.117	0.730	0.793
	720	0.072	0.123	0.700	1.202	0.653	0.653	0.722	1.245	0.410	0.379	0.735	1.793	0.569	0.761	0.436	0.510	0.644	0.887	0.674	0.916	0.881	1.483	0.090	0.117	0.782	0.721
	Avg	0.066	0.098	0.501	0.644	0.476	0.476	0.712	1.357	0.369	0.302	0.575	1.048	0.452	0.486	0.357	0.346	0.470	0.528	0.498	0.554	0.689	0.911	0.076	0.115	0.744	0.776

Table 6: Performance comparison under 50% missing observations with $L = 96$ and $T \in \{96, 192, 336, 720\}$ on multivariate forecasting benchmarks. The best and second-best results are highlighted in red and blue.

Model		FenDiM (Ours)		TimeMosaic (2026)		SparseTSF (2025)		SimpleTM (2025)		DPEM (2025)		xPatch (2025)		AMD (2025)		DeepEDM (2025)		PDF (2024)		iTransformer (2023)		BiTGraph (2023)		SSSD (2023)		Autoformer (2021)	
Metric		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
ETTh1	96	0.132	0.421	0.224	1.413	0.221	1.356	0.201	1.104	0.217	1.228	0.220	1.370	0.209	1.226	0.218	1.326	0.218	1.326	0.236	1.647	0.258	1.440	0.139	0.313	0.276	2.009
	192	0.133	0.424	0.230	1.472	0.225	1.393	0.203	1.130	0.219	1.238	0.231	1.461	0.214	1.267	0.224	1.374	0.224	1.374	0.241	1.686	0.264	1.482	0.139	0.316	0.278	2.029
	336	0.134	0.427	0.237	1.538	0.231	1.446	0.204	1.150	0.222	1.259	0.239	1.537	0.219	1.310	0.230	1.431	0.230	1.431	0.246	1.743	0.275	1.559	0.140	0.324	0.281	2.076
	720	0.136	0.442	0.253	1.716	0.243	1.576	0.206	1.179	0.230	1.325	0.258	1.717	0.232	1.432	0.244	1.569	0.244	1.569	0.259	1.875	0.300	1.775	0.141	0.329	0.292	2.212
	Avg	0.134	0.428	0.236	1.535	0.230	1.442	0.204	1.140	0.222	1.262	0.237	1.521	0.219	1.309	0.229	1.425	0.229	1.425	0.246	1.738	0.274	1.564	0.140	0.321	0.282	2.081
ETTh2	96	0.262	1.421	0.442	5.887	0.405	4.576	0.393	4.662	0.404	4.075	0.366	3.838	0.376	3.603	0.401	4.527	0.392	4.371	0.464	6.664	0.436	4.983	0.204	0.974	0.547	7.797
	192	0.263	1.434	0.478	7.083	0.427	5.084	0.400	4.868	0.416	4.289	0.438	6.593	0.376	3.603	0.465	7.450	0.416	4.887	0.487	7.171	0.465	5.485	0.212	1.056	0.537	7.585
	336	0.265	1.468	0.497	7.448	0.450	5.687	0.404	5.062	0.427	4.520	0.537	12.960	0.432	5.036	0.512	10.150	0.441	5.520	0.510	7.718	0.504	6.211	0.212	1.063	0.544	7.839
	720	0.272	1.550	0.550	8.834	0.504	6.906	0.403	5.242	0.456	4.982	0.708	26.380	0.491	6.970	0.673	23.270	0.495	6.729	0.556	8.630	0.604	8.192	0.215	1.071	0.580	8.758
	Avg	0.266	1.468	0.492	7.313	0.447	5.563	0.400	4.959	0.426	4.467	0.512	12.443	0.419	4.803	0.513	11.349	0.436	5.377	0.504	7.546	0.502	6.218	0.211	1.041	0.552	7.995
ETTm1	96	0.089	0.189	0.204	1.344	0.299	2.428	0.319	3.581	0.207	1.209	0.214	1.476	0.188	1.113	0.199	1.238	0.196	1.206	0.212	1.405	0.233	1.478	0.297	2.900	0.288	2.648
	192	0.090	0.193	0.211	1.417	0.312	2.689	0.321	3.621	0.216	1.280	0.235	1.824	0.193	1.154	0.206	1.314	0.204	1.279	0.217	1.451	0.269	1.759	0.299	2.904	0.272	2.338
	336	0.091	0.199	0.221	1.548	0.328	3.014	0.323	3.696	0.223	1.337	0.265	2.422	0.197	1.187	0.218	1.482	0.211	1.344	0.222	1.496	0.312	2.159	0.300	2.911	0.266	2.199
	720	0.093	0.204	0.235	1.734	0.355	3.580	0.326	3.744	0.227	1.354	0.288	2.806	0.203	1.212	0.232	1.706	0.219	1.411	0.228	1.544	0.327	2.227	0.302	2.906	0.264	2.121
	Avg	0.091	0.196	0.218	1.511	0.324	2.928	0.322	3.660	0.218	1.295	0.251	2.132	0.195	1.166	0.214	1.435	0.207	1.310	0.220	1.474	0.285	1.906	0.299	2.905	0.272	2.327
ETTm2	96	0.144	0.432	0.268	1.928	0.286	2.256	0.362	3.356	0.292	2.240	0.286	2.305	0.266	1.877	0.289	2.369	0.276	2.133	0.318	2.968	0.322	2.150	0.186	0.472	0.401	4.071
	192	0.145	0.440	0.285	2.216	0.299	2.528	0.367	3.515	0.306	2.462	0.303	2.636	0.278	2.070	0.310	2.762	0.295	2.468	0.331	3.221	0.370	2.770	0.186	0.474	0.395	3.994
	336	0.147	0.450	0.303	2.585	0.314	2.847	0.372	3.708	0.323	2.729	0.329	3.193	0.293	2.310	0.329	3.191	0.313	2.838	0.345	3.521	0.424	3.577	0.186	0.474	0.394	4.012
	720	0.150	0.476	0.332	3.132	0.339	3.402	0.373	3.762	0.352	3.194	0.387	4.512	0.321	2.798	0.364	3.967	0.345	3.488	0.369	4.071	0.497	4.823	0.187	0.476	0.410	4.472
	Avg	0.147	0.449	0.297	2.465	0.310	2.758	0.369	3.585	0.318	2.656	0.326	3.162	0.290	2.264	0.323	3.072	0.307	2.732	0.341	3.445	0.403	3.330	0.186	0.474	0.400	4.137
Weather	96	0.253	0.937	0.822	5.906	0.793	5.640	0.339	2.003	0.399	2.589	0.432	3.074	0.423	4.232	0.879	6.069	0.778	5.630	1.268	9.481	0.709	4.743	0.434	2.384	0.565	4.075
	192	0.259	0.970	0.775	5.490	0.735	5.321	0.335	2.099	0.374	2.461	0.405	2.863	0.398	14.310	0.898	6.253	0.722	5.285	1.180	8.834	0.672	4.346	0.428	2.409	0.519	3.659
	336	0.259	0.969	0.795	5.619	0.748	5.438	0.342	2.182	0.380	2.508	0.411	2.866	0.410	11.220	0.926	6.625	0.737	5.396	1.194	8.967	0.677	4.466	0.427	2.436	0.519	3.688
	720	0.259	0.958	0.817	5.723	0.763	5.579	0.349	2.235	0.388	2.565	0.420	2.864	0.427	8.808	0.947	7.797	0.753	5.523	1.213	9.151	0.686	4.611	0.426	2.469	0.527	3.767
	Avg	0.258	0.959	0.802	5.685	0.760	5.494	0.341	2.130	0.385	2.531	0.417	2.917	0.414	9.642	0.913	6.686	0.747	5.459	1.214	9.108	0.686	4.542	0.429	2.425	0.532	3.797
Exchange	96	0.062	0.083	0.569	0.868	0.469	0.529	0.608	0.839	0.457	0.615	0.520	0.712	0.518	0.614	0.444	0.486	0.478	0.525	0.472	0.482	0.574	0.559	0.068	0.115	0.829	0.732
	192	0.063	0.089	0.616	0.949	0.519	0.625	0.637	0.881	0.479	0.635	0.590	0.906	0.574	0.755	0.482	0.567	0.532	0.633	0.524	0.577	0.604	0.630	0.070	0.121	0.827	0.831
	336	0.066	0.099	0.681	1.178	0.591	0.772	0.662	0.936	0.512	0.710	0.782	1.769	0.626	0.946	0.530	0.679	0.604	0.788	0.594	0.721	0.665	0.781	0.071	0.123	0.827	0.831
	720	0.072	0.123	0.848	1.940	0.734	1.109	0.697	1.017	0.748	0.817	0.857	1.833	0.784	1.433	0.638	0.948	0.742	1.121	0.740	1.065	0.863	1.349	0.074	0.122	0.931	0.934
	Avg	0.066	0.099	0.679	1.234	0.578	0.759	0.651	0.918	0.549	0.694	0.687	1.305	0.626	0.937	0.524	0.670	0.589	0.767	0.583	0.711	0.677	0.830	0.071	0.120	0.854	0.832

and the best or tied-best average MSE on four datasets. This indicates that FenDiM remains reliable when temporal evidence is simultaneously incomplete and noisy.

Table 7: Performance comparison under 70% missing observations with $L = 96$ and $T \in \{96, 192, 336, 720\}$ on multivariate forecasting benchmarks. The best and second-best results are highlighted in red and blue.

Model		FenDiM (Ours)		TimeMosaic (2026)		SparseTSF (2025)		SimpleTM (2025)		DPEM (2025)		xPatch (2025)		AMD (2025)		DeepEDM (2025)		PDF (2024)		iTransformer (2023)		BiTGraph (2023)		SSSD (2023)		Autoformer (2021)	
Metric		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
ETTh1	96	0.186	0.824	0.261	1.908	0.261	1.919	0.215	1.411	0.256	1.750	0.220	1.370	0.270	2.095	0.257	1.816	0.261	1.925	0.266	2.040	0.278	1.714	0.202	2.817	0.301	2.414
	192	0.190	0.861	0.265	1.956	0.265	1.957	0.220	1.464	0.259	1.773	0.231	1.461	0.276	2.157	0.263	1.844	0.266	1.974	0.270	2.079	0.281	1.735	0.202	2.845	0.302	2.426
	336	0.191	0.874	0.271	2.017	0.270	2.016	0.225	1.504	0.262	1.812	0.239	1.537	0.281	2.217	0.270	1.904	0.272	2.036	0.276	2.140	0.290	1.797	0.203	2.918	0.304	2.470
	720	0.197	0.924	0.284	2.168	0.281	2.154	0.231	1.578	0.272	1.911	0.258	1.717	0.295	2.401	0.298	2.241	0.284	2.183	0.287	2.277	0.305	1.935	0.204	2.972	0.314	2.611
	Avg	0.191	0.870	0.270	2.012	0.269	2.012	0.223	1.489	0.262	1.811	0.237	1.521	0.280	2.218	0.272	1.951	0.271	2.029	0.275	2.134	0.288	1.795	0.203	2.888	0.305	2.480
ETTh2	96	0.353	2.719	0.520	8.375	0.478	6.377	0.367	3.862	0.482	6.034	0.504	7.265	0.497	6.412	0.453	5.393	0.470	6.214	0.496	7.486	0.463	5.755	0.360	1.148	0.599	9.179
	192	0.354	2.757	0.545	9.251	0.496	6.836	0.386	4.293	0.496	6.326	0.569	9.544	0.559	8.328	0.496	6.820	0.489	6.664	0.517	7.961	0.494	6.262	0.369	1.265	0.590	8.963
	336	0.361	2.890	0.563	9.500	0.516	7.384	0.400	4.725	0.512	6.663	0.652	13.770	0.645	12.567	0.535	8.335	0.509	7.233	0.538	8.483	0.536	6.997	0.373	1.307	0.596	9.197
	720	0.382	3.274	0.615	11.080	0.562	8.443	0.423	5.538	0.547	7.286	0.786	20.550	0.777	20.206	0.657	14.070	0.552	8.183	0.582	9.366	0.647	9.063	0.384	1.382	0.627	9.993
	Avg	0.363	2.910	0.561	9.551	0.513	7.260	0.394	4.605	0.509	6.577	0.628	12.782	0.619	11.878	0.535	8.655	0.505	7.074	0.533	8.324	0.535	7.019	0.371	1.275	0.603	9.333
ETTm1	96	0.145	0.516	0.247	1.992	0.343	3.253	0.304	3.453	0.244	1.763	0.245	1.971	0.245	1.974	0.238	1.812	0.239	1.818	0.243	1.854	0.243	1.854	0.297	2.899	0.324	3.419
	192	0.149	0.552	0.251	2.054	0.352	3.449	0.310	3.535	0.251	1.826	0.264	2.320	0.248	2.011	0.242	1.853	0.244	1.879	0.246	1.890	0.246	1.890	0.298	2.904	0.309	3.088
	336	0.154	0.593	0.259	2.193	0.364	3.712	0.316	3.648	0.257	1.881	0.293	2.948	0.252	2.045	0.246	1.898	0.250	1.937	0.250	1.926	0.250	1.926	0.300	2.910	0.301	2.869
	720	0.158	0.622	0.272	2.394	0.386	4.229	0.323	3.740	0.260	1.903	0.313	3.321	0.256	2.052	0.252	1.944	0.256	1.991	0.255	1.959	0.255	1.959	0.302	2.905	0.298	2.781
	Avg	0.151	0.571	0.257	2.158	0.361	3.661	0.313	3.594	0.253	1.843	0.279	2.640	0.250	2.021	0.245	1.877	0.247	1.906	0.248	1.907	0.248	1.907	0.299	2.905	0.308	3.039
ETTm2	96	0.227	1.163	0.326	2.927	0.337	3.154	0.336	2.890	0.341	3.176	0.338	3.198	0.334	3.078	0.341	3.176	0.335	3.145	0.355	3.603	0.353	2.644	0.292	1.411	0.518	7.318
	192	0.233	1.238	0.340	3.215	0.346	3.361	0.349	3.212	0.353	3.383	0.350	3.480	0.346	3.312	0.353	3.383	0.348	3.426	0.364	3.805	0.398	3.213	0.292	1.412	0.501	6.862
	336	0.241	1.346	0.355	3.560	0.356	3.624	0.362	3.565	0.367	3.663	0.371	4.018	0.362	3.620	0.367	3.663	0.362	3.744	0.375	4.061	0.453	4.033	0.293	1.526	0.479	6.289
	720	0.256	1.558	0.379	4.098	0.377	4.130	0.375	3.821	0.394	4.181	0.423	5.317	0.396	4.291	0.394	4.181	0.389	4.359	0.395	4.567	0.564	5.872	0.294	1.558	0.504	7.080
	Avg	0.239	1.326	0.350	3.450	0.354	3.567	0.355	3.372	0.364	3.601	0.371	4.003	0.360	3.575	0.364	3.601	0.359	3.669	0.373	4.009	0.442	3.941	0.293	1.477	0.501	6.887
Weather	96	0.259	0.953	1.022	8.608	0.971	8.021	0.413	2.831	0.484	3.681	0.515	4.230	0.512	7.308	1.038	7.471	0.967	8.075	1.241	9.612	0.693	5.174	0.800	7.307	0.613	4.881
	192	0.263	0.987	0.945	7.707	0.882	7.301	0.400	2.824	0.445	3.374	0.473	3.845	0.471	7.039	1.023	7.300	0.880	7.303	1.155	8.931	0.642	4.634	0.775	7.087	0.549	4.281
	336	0.263	0.985	0.970	7.894	0.897	7.456	0.411	2.934	0.462	3.523	0.480	3.871	0.485	6.480	1.070	7.880	0.895	7.433	1.170	9.068	0.647	4.744	0.790	7.346	0.547	4.311
	720	0.261	0.977	0.995	8.041	0.916	7.657	0.424	3.019	0.462	3.523	0.490	3.910	0.504	5.831	1.107	8.952	0.914	7.596	1.188	9.256	0.656	4.888	0.791	7.500	0.556	4.407
	Avg	0.262	0.976	0.983	8.062	0.916	7.609	0.412	2.902	0.463	3.525	0.489	3.964	0.493	6.665	1.059	7.901	0.914	7.602	1.189	9.216	0.660	4.860	0.789	7.310	0.566	4.470
Exchange	96	0.094	0.164	0.849	2.056	0.607	0.839	0.491	0.586	0.599	0.755	0.622	0.951	0.659	0.959	0.590	0.784	0.622	0.849	0.605	0.757	0.670	0.795	0.075	0.229	1.311	3.373
	192	0.095	0.170	0.901	2.218	0.647	0.901	0.543	0.749	0.625	0.827	0.672	1.095	0.711	1.112	0.619	0.869	0.665	0.943	0.646	0.845	0.687	0.768	0.077	0.237	1.302	3.294
	336	0.103	0.185	0.929	2.280	0.705	0.976	0.593	0.916	0.665	0.944	0.851	1.920	0.777	1.332	0.671	0.993	0.723	1.079	0.704	0.977	0.728	0.780	0.078	0.240	1.306	3.294
	720	0.112	0.233	1.050	2.946	0.830	1.131	0.674	1.226	0.762	1.269	0.929	1.976	0.919	1.895	0.788	1.309	0.842	1.393	0.830	1.304	0.893	1.429	0.079	0.243	1.359	3.567
	Avg	0.101	0.188	0.932	2.375	0.697	0.962	0.575	0.869	0.663	0.949	0.769	1.486	0.767	1.325	0.667	0.989	0.713	1.066	0.696	0.971	0.745	0.943	0.077	0.237	1.320	3.382

Table 8: Performance comparison with Gaussian-corrupted observations on multivariate forecasting benchmarks, with historical length $L = 96$ and prediction lengths $T \in \{96, 192, 336, 720\}$. The best and second-best results in each row are highlighted in red and blue, respectively. Avg is the arithmetic mean of the four displayed horizon values, rounded to three decimals. Ties share the same highlight.

Model		FenDiM (Ours)		TimeMosaic (2026)		SparseTSF (2025)		SimpleTM (2025)		DPEM (2025)		xPatch (2025)		AMD (2025)		DeepEDM (2025)		PDF (2024)		iTransformer (2023)		BiTGraph (2023)		SSSD (2023)		Autoformer (2021)	
Metric		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
ETTh1	96	0.122	0.243	0.132	0.407	0.162	0.720	0.213	1.017	0.157	0.657	0.157	0.711	0.892	24.401	0.104	0.233	0.154	0.655	0.199	1.141	0.333	3.072	0.135	0.333	0.223	1.346
	192	0.124	0.255	0.136	0.423	0.167	0.755	0.206	0.975	0.157	0.657	0.166	0.775	0.950	25.999	0.109	0.247	0.162	0.708	0.205	1.180	0.340	3.151	0.135	0.336	0.229	1.434
	336	0.128	0.284	0.141	0.443	0.174	0.797	0.201	0.949	0.157	0.656	0.173	0.820	1.027	29.082	0.113	0.263	0.170	0.757	0.211	1.233	0.342	3.186	0.136	0.343	0.234	1.503
	720	0.134	0.292	0.154	0.519	0.191	0.918	0.190	0.877	0.160	0.672	0.189	0.946	1.110	34.360	0.121	0.294	0.188	0.877	0.227	1.361	0.351	3.275	0.138	0.355	0.246	1.619
	Avg	0.127	0.269	0.141	0.448	0.174	0.798	0.203	0.955	0.158	0.661	0.171	0.813	0.995	28.461	0.112	0.259	0.169	0.749	0.211	1.229	0.342	3.171	0.136	0.342	0.233	1.476
ETTh2	96	0.237	1.667	0.247	1.554	0.281	2.374	0.485	7.405	0.276	1.878	0.298	2.617	3.966	31.966	0.181	0.833	0.255	2.018	0.431	4.941	0.636	9.970	0.399	5.008	0.380	3.751
	192	0.238	1.714	0.274	1.951	0.311	2.953	0.470	7.045	0.286	2.008	0.360	4.542	4.330	38.028	0.187	0.993	0.288	2.593	0.457	5.511	0.655	10.423	0.413	5.321	0.386	3.909
	336	0.234	1.825	0.297	2.388	0.342	3.650	0.451	6.604	0.292	2.080	0.440	8.837	5.210	57.072	0.195	1.246	0.320	3.262	0.481	6.114	0.678	10.999	0.434	5.795	0.398	4.229
	720	0.240	2.212	0.360	3.660	0.409	5.152	0.405	5.527	0.307	2.227	0.575	17.570	7.254	120.589	0.216	1.671	0.392	4.717	0.527	7.128	0.731	12.007	0.484	6.859	0.435	5.093
	Avg	0.237	1.855	0.295	2.388	0.336	3.532	0.453	6.645	0.290	2.048	0.418	8.392	5.190	61.914	0.195	1.186	0.314	3.148	0.474	5.924	0.675	10.850	0.433	5.746	0.400	4.246
ETTm1	96	0.103	0.598	0.106	0.635	0.138	0.561	0.360	3.681	0.130	0.477	0.134	0.569	2.043	81.399	0.094	0.189	0.127	0.490	0.174	0.916	0.310	2.782	0.438	3.826	0.297	2.236
	192	0.107	0.637	0.113	0.661	0.147	0.640	0.352	3.591	0.133	0.490	0.143	0.682	2.259	96.975	0.101	0.219	0.139	0.586	0.181	0.972	0.314	2.779	0.475	4.483	0.284	1.996
	336	0.112	0.686	0.120	0.690	0.154	0.706	0.345	3.536	0.134	0.494	0.155	0.837	2.411	113.150	0.108	0.249	0.149	0.659	0.188	1.023	0.319	2.764	0.510	5.285	0.280	1.912
	720	0.112	0.766	0.131	0.637	0.165	0.784	0.336	3.458	0.137	0.512	0.173	1.010	2.570	132.960	0.122	0.306	0.161	0.747	0.196	1.088	0.327	2.814	0.603	7.088	0.281	1.877
	Avg	0.109	0.672	0.118	0.656	0.151	0.673	0.348	3.567	0.134	0.493	0.151	0.775	2.321	106.121	0.106	0.241	0.144	0.621	0.185	1.000	0.318	2.785	0.507	5.171	0.286	2.005
ETTm2	96	0.158	0.504	0.171	0.695	0.205	1.157	0.432	4.692	0.196	0.908	0.187	0.903	1.175	40.068	0.160	0.358	0.185	0.990	0.255	2.013	0.481	5.568	0.249	2.028	0.323	2.575
	192	0.158	0.510	0.189	0.882	0.229	1.527	0.419	4.478	0.211	1.035	0.203	1.084	1.241	41.493	0.167	0.435	0.213	1.373	0.274	2.341	0.484	5.620	0.269	2.339	0.328	2.704
	336	0.159	0.521	0.204	1.063	0.253	1.935	0.407	4.290	0.221	1.111	0.219	1.286	1.318	45.408	0.176	0.532	0.239	1.773	0.293	2.705	0.495	5.812	0.288	2.643	0.333	2.832
	720	0.164	0.557	0.230	1.420	0.288	2.574	0.383	3.854	0.236	1.259	0.250	1.766	1.582	70.994	0.177	0.801	0.276	2.384	0.324	3.320	0.546	6.803	0.319	3.229	0.352	3.260
	Avg	0.160	0.523	0.199	1.015	0.244	1.798	0.410	4.329	0.216	1.078	0.215	1.260	1.329	49.491	0.170	0.532	0.228	1.630	0.287	2.595	0.502	5.951	0.281	2.560	0.334	2.843
Weather	96	0.259	0.912	0.355	2.246	0.465	3.033	0.268	1.370	0.254	1.463	0.267	1.678	1.288	5.998	0.308	1.445	0.423	2.899	1.309	9.366	0.786	6.156	0.261	7.639	0.488	3.002
	192	0.257	1.052	0.355	2.293	0.472	3.175	0.264	1.427	0.259	1.526	0.275	1.717	0.980	3.937	0.309	1.467	0.438	3.074	1.217	8.750	0.724	5.449	0.260	7.565	0.492	3.066
	336	0.258	1.056	0.357	2.308	0.481	3.265	0.269	1.452	0.264	1.563	0.282	1.744	1.011	4.052	0.313	1.496	0.452	3.184	1.231	8.880	0.733	5.572	0.261	7.612	0.493	3.078
	720	0.259	1.342	0.363	2.366	0.492	3.355	0.264	1.451	0.271	1.599	0.292	1.741	1.079	4.490	0.314	1.494	0.469	3.289	1.250	9.056	0.748	5.745	0.264	7.776	0.496	3.102
	Avg	0.258	1.091	0.358	2.303	0.478	3.207	0.266	1.425	0.262	1.538	0.279	1.720	1.090	4.619	0.311	1.476	0.446	3.112	1.252	9.013	0.748	5.731	0.262	7.648	0.492	3.062
Exchange	96	0.058	0.076	1.548	3.827	0.283	0.133	1.268	2.260	0.196	0.560	0.244	0.112	1.151	5.897	0.589	0.538	0.299	0.140	0.143	0.316	0.840	0.744	0.073	0.090	0.794	1.442
	192	0.060	0.082	1.520	3.690	0.280	0.131	1.406	2.773	0.193	0.548	0.270	0.134	1.383	7.631	0.678	0.695	0.309	0.150	0.143	0.318	0.803	0.683	0.076	0.099	0.757	1.061
	336	0.063	0.092	1.563	4.044	0.278	0.130	1.563	3.440	0.188	0.529	0.326	0.215	1.666	9.629	0.700	0.736	0.314	0.156	0.143	0.320	0.775	0.640	0.081	0.112	0.734	0.834
	720	0.069	0.114	1.935	7.190	0.271	0.126	1.778	4.477	0.185	0.528	0.646	1.106	2.165	13.423	0.692	0.721	0.329	0.174	0.147	0.339	0.754	0.610	0.091	0.144	0.735	0.784
	Avg	0.063	0.091	1.642	4.688	0.278	0.130	1.504	3.238	0.191	0.541	0.372	0.392	1.591	9.145	0.665	0.673	0.313	0.155	0.144	0.323	0.793	0.669	0.080	0.111	0.755	1.030

The results reveal different aspects of FenDiM's robustness under compound perturbations. On ETTh1, where long-horizon dependencies are vulnerable to error propagation once observations are corrupted, FenDiM reduces average MAE by 3.4% over the second-best method while ranking

Table 9: Performance comparison under simultaneous 30% missingness and Gaussian noise on multivariate forecasting benchmarks, with historical length $L = 96$ and prediction lengths $T \in \{96, 192, 336, 720\}$. The best and second-best results in each row are highlighted in red and blue, respectively.

Model		FenDiM (Ours)		TimeMosaic (2026)		SparseTSF (2025)		SimpleTM (2025)		DPEM (2025)		xPatch (2025)		AMD (2025)		DeepEDM (2025)		PDF (2024)		iTransformer (2023)		BiTGraph (2023)		SSSD (2023)		Autoformer (2021)	
Metric		MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
ETTh1	96	0.114	0.307	0.184	0.952	0.187	0.966	0.196	0.963	0.186	0.878	0.183	0.916	0.705	17.15	0.185	0.95	0.181	0.911	0.215	1.382	0.258	1.375	0.118	0.291	0.255	1.701
	192	0.114	0.311	0.19	0.996	0.193	1.003	0.194	0.963	0.187	0.882	0.195	1.009	0.744	18.53	0.194	1.028	0.189	0.961	0.220	1.422	0.266	1.435	0.118	0.294	0.258	1.744
	336	0.115	0.312	0.197	1.051	0.199	1.05	0.193	0.967	0.188	0.892	0.204	1.087	0.783	20.25	0.202	1.093	0.196	1.015	0.226	1.475	0.278	1.521	0.12	0.302	0.263	1.801
	720	0.116	0.319	0.213	1.203	0.213	1.175	0.190	0.956	0.194	0.936	0.225	1.280	0.845	23.68	0.224	1.295	0.212	1.142	0.24	1.605	0.312	1.826	0.12	0.307	0.275	1.932
	Avg	0.115	0.312	0.196	1.051	0.198	1.049	0.193	0.962	0.189	0.897	0.202	1.073	0.769	19.903	0.201	1.092	0.195	1.007	0.225	1.471	0.279	1.539	0.119	0.299	0.263	1.794
ETTh2	96	0.229	1.078	0.348	3.353	0.342	3.319	0.425	5.655	0.338	2.751	0.350	3.555	1.444	92.01	0.353	3.654	0.322	2.989	0.443	5.917	0.430	4.661	0.322	1.077	1.496	6.534
	192	0.229	1.079	0.384	4.186	0.368	3.86	0.422	5.601	0.348	2.916	0.418	6.039	1.668	120.2	0.399	5.238	0.350	3.552	0.468	6.447	0.461	5.21	0.313	1.08	1.491	6.433
	336	0.228	1.077	0.406	4.677	0.395	4.507	0.416	5.538	0.357	3.047	0.512	11.68	1.997	180.6	0.446	7.323	0.380	4.222	0.492	7.01	0.502	5.987	0.312	1.081	1.501	6.736
	720	0.23	1.088	0.464	5.974	0.455	5.855	0.395	5.177	0.378	3.326	0.669	23.66	2.891	428.5	0.591	18.21	0.444	5.612	0.538	7.955	0.604	8.127	0.313	1.084	1.540	7.758
	Avg	0.229	1.081	0.401	4.548	0.390	4.385	0.415	5.493	0.355	3.010	0.487	11.234	2.000	205.32	0.447	8.606	0.374	4.094	0.485	6.832	0.499	5.996	0.315	1.081	1.507	6.865
ETTm1	96	0.071	0.121	0.163	0.839	0.163	0.815	0.33	3.606	0.172	0.819	0.175	0.962	0.867	29.71	0.171	0.914	0.160	0.790	0.191	1.126	0.217	1.324	0.298	2.903	0.268	2.263
	192	0.071	0.12	0.173	0.926	0.17	0.875	0.328	3.597	0.181	0.887	0.195	1.219	0.973	38.35	0.186	1.091	0.17	0.872	0.197	1.178	0.248	1.539	0.299	2.907	0.254	2.012
	336	0.071	0.12	0.183	1.031	0.177	0.927	0.328	3.624	0.189	0.945	0.218	1.584	1.026	43.22	0.201	1.318	0.178	0.939	0.203	1.228	0.284	1.869	0.300	2.914	0.249	1.907
	720	0.071	0.12	0.198	1.158	0.185	0.984	0.326	3.625	0.194	0.974	0.243	1.904	1.083	49.04	0.222	1.621	0.188	1.015	0.210	1.285	0.303	1.991	0.302	2.908	0.248	1.851
	Avg	0.071	0.12	0.179	0.989	0.174	0.901	0.328	3.613	0.184	0.906	0.208	1.417	0.987	40.08	0.195	1.236	0.174	0.904	0.201	1.204	0.263	1.681	0.300	2.908	0.255	2.008
ETTm2	96	0.12	0.296	0.217	1.215	0.243	1.615	0.385	3.812	0.246	1.523	0.234	1.527	0.722	19.39	0.255	1.87	0.228	1.468	0.288	2.507	0.315	1.976	0.124	0.307	0.395	4.211
	192	0.119	0.296	0.237	1.499	0.261	1.938	0.383	3.825	0.263	1.744	0.252	1.833	0.800	24.53	0.28	2.273	0.252	1.832	0.303	2.798	0.365	2.637	0.124	0.309	0.388	4.054
	336	0.119	0.296	0.257	1.837	0.280	2.306	0.381	3.871	0.281	1.969	0.276	2.296	0.876	30.15	0.303	2.737	0.273	2.222	0.32	3.129	0.419	3.468	0.125	0.310	0.382	3.932
	720	0.12	0.298	0.288	2.369	0.310	2.902	0.372	3.743	0.310	2.357	0.326	3.352	1.071	48.97	0.341	3.552	0.308	2.853	0.347	3.709	0.480	4.553	0.125	0.310	0.402	4.468
	Avg	0.12	0.297	0.250	1.73	0.273	2.190	0.38	3.813	0.275	1.898	0.272	2.252	0.867	30.76	0.295	2.608	0.265	2.094	0.315	3.036	0.395	3.159	0.124	0.309	0.392	4.166
Weather	96	0.253	0.937	0.628	3.922	0.637	4.005	0.279	1.542	0.326	1.873	0.343	2.177	1.358	27.72	0.709	4.660	0.612	3.950	1.286	9.41	0.759	4.648	0.379	2.242	0.525	3.465
	192	0.252	0.961	0.603	3.842	0.608	3.969	0.279	1.656	0.314	1.863	0.333	2.111	1.296	21.72	0.707	5.049	0.588	3.933	1.196	8.782	0.722	4.339	0.377	2.287	0.498	3.241
	336	0.253	0.961	0.620	3.951	0.618	4.064	0.28	1.71	0.320	1.901	0.342	2.130	1.323	18.92	0.723	5.058	0.603	4.037	1.211	8.914	0.720	4.445	0.377	2.317	0.500	3.273
	720	0.253	0.948	0.639	4.06	0.631	4.167	0.283	1.734	0.327	1.943	0.352	2.122	1.363	17.07	0.738	5.539	0.619	4.143	1.229	9.094	0.724	4.574	0.377	2.353	0.506	3.331
	Avg	0.253	0.952	0.623	3.944	0.624	4.051	0.280	1.661	0.322	1.895	0.343	2.135	1.335	21.36	0.719	5.077	0.606	4.016	1.231	9.050	0.731	4.502	0.378	2.300	0.507	3.328
Exchange	96	0.095	0.134	1.339	3.067	0.538	0.51	1.628	3.98	0.391	0.301	0.903	1.336	0.888	3.907	0.623	0.651	0.451	0.376	0.441	0.376	0.558	0.569	0.112	0.261	1.229	3.258
	192	0.096	0.196	1.395	3.329	0.582	0.609	1.690	4.278	0.408	0.330	0.903	1.956	1.048	4.857	0.733	0.899	0.517	0.503	0.517	0.503	0.598	0.669	0.114	0.277	1.17	2.867
	336	0.101	0.441	1.517	4.093	0.645	0.761	1.748	4.568	0.431	0.377	1.077	2.441	1.251	6.207	0.81	1.121	0.599	0.678	0.599	0.678	0.663	0.838	0.115	0.289	1.09	2.437
	720	0.113	0.461	1.989	4.805	0.773	1.092	1.795	4.884	0.466	0.459	1.475	2.462	1.592	7.502	0.948	1.594	0.752	1.046	0.752	1.046	0.861	1.407	0.117	0.303	1.152	2.631
	Avg	0.101	0.308	1.560	3.824	0.635	0.743	1.715	4.428	0.424	0.367	1.09	2.049	1.195	5.618	0.779	1.066	0.580	0.651	0.577	0.651	0.670	0.871	0.115	0.283	1.16	2.798

second in average MSE. On ETTm2, where high-resolution sampling makes local dependencies more sensitive to missing-and-noisy inputs, FenDiM reduces average MAE and MSE by 3.2% and 3.9%, respectively. On Weather, where short-term fluctuations and longer-range structures coexist, FenDiM reduces average MAE by 9.6% and achieves the best average MSE, showing stronger control of unstable dependency evidence. Even on Exchange, whose non-stationarity makes corrupted observations harder to interpret, FenDiM remains within the top two for both average MAE and MSE. These moderate but consistent improvements suggest that FenDiM is not simply exploiting clean temporal patterns, but can retain predictive dependencies when different corruption sources interact.

This robustness stems from prediction-oriented dependency rectification. Missing values break temporal continuity, while Gaussian noise perturbs the remaining observed values, together inducing biased dependency estimates. Methods that directly extract temporal patterns, estimate cross-variable correlations, or reconstruct sequence structures may still inherit such corrupted dependencies. FenDiM instead calibrates dependency formation toward forecasting utility, suppressing missingness- and noise-induced spurious relations while preserving stable predictive evidence. This explains its consistent top-tier performance under compound corruptions.

D.1 Hyperparameter Sensitivity

FenDiM introduces two key hyperparameters: the diffusion loss weight λ_{pdr} and the frequency loss weight λ_{spec}. λ_{pdr} determines the strength of prediction-oriented dependency rectification, while λ_{spec} controls the degree of frequency-domain structural alignment. We evaluate their impact by varying one hyperparameter at a time under forecasting horizons $T \in \{96, 192, 336, 720\}$, as shown in Figure 3.

Diffusion loss weight. We evaluate $\lambda_{\text{pdr}} \in \{0, 0.005, 0.01, 0.03, 0.05, 0.1\}$ with λ_{spec} fixed. The results show that a lightweight diffusion loss consistently improves forecasting performance, peaking at $\lambda_{\text{pdr}} = 0.005$. Compared with removing the diffusion loss, it reduces the average MAE and MSE by 1.14% and 1.12%, respectively. Notably, the performance remains stable across a broad range of small weights, indicating that the rectification loss is robust and does not require horizon-

specific tuning. However, a larger weight such as $\lambda_{\text{pdr}} = 0.03$ degrades performance, suggesting that excessive dependency rectification may suppress forecasting-relevant variations.

Frequency loss weight. We evaluate $\lambda_{\text{spec}} \in \{0.01, 0.03, 0.05, 0.1, 0.2, 0.3\}$ with λ_{pdr} fixed. The results demonstrate that moderate frequency alignment is beneficial, with the best performance achieved at $\lambda_{\text{spec}} = 0.1$. Compared with the weakest frequency setting $\lambda_{\text{spec}} = 0.01$, it improves the average MAE and MSE by 4.76% and 9.18%, respectively. The gains are consistent across horizons, indicating that frequency alignment helps preserve global frequency-domain structures. In contrast, overly large weights, e.g., $\lambda_{\text{spec}} = 0.2$ or 0.3, lead to diminished performance, implying that excessive frequency constraints may over-emphasize frequency reconstruction and weaken time-domain prediction fidelity.

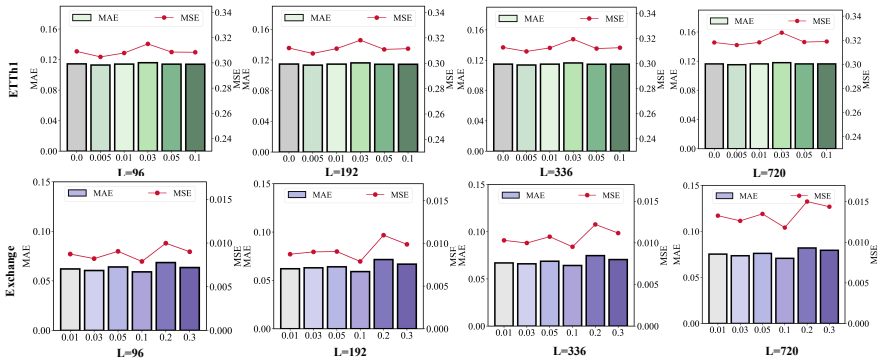

Figure 3: Sensitivity analysis of FenDiM across forecasting horizons, where λ_{pdr} is evaluated on ETTh1 and λ_{spec} is evaluated on Exchange.

E Efficiency Analysis

Diffusion-based rectification improves the reliability of the learned forecasting representation, but a direct implementation with attention-based denoising may introduce non-negligible computational overhead. To examine whether the proposed Mamba-based denoiser makes FenDiM computationally practical, we evaluate both component-level runtime and end-to-end training efficiency.

We report two types of runtime measurements. First, the component-level runtime in Figure 4 is measured at the batch level, where each number denotes the summed forward-pass time of the corresponding model components within a single mini-batch. This measurement isolates the computational cost of diffusion rectification and its related modules. Second, the training time in Table 10 is measured at the epoch level, where each number denotes the end-to-end wall-clock time for one complete training epoch under the same hardware, batch size, look-back length, prediction horizon, and training protocol.

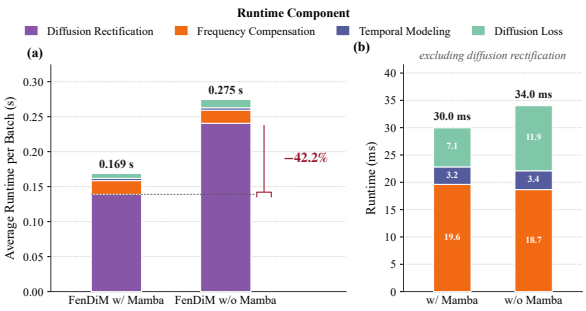

Figure 4: Component-level per-batch forward runtime of FenDiM with and without the Mamba-based denoiser. Each value denotes the summed forward-pass time of the corresponding component within a single mini-batch. Mamba mainly reduces the cost of diffusion inference, leading to a 42.2% reduction in diffusion inference time and a 38.5% reduction in total per-batch forward runtime.

Component-level runtime. Figure 4 compares the runtime decomposition of FenDiM with and without the Mamba-based denoiser. Replacing the denoising backbone with Mamba reduces the diffusion inference time from 0.241 to 0.139, corresponding to a 42.2% reduction. The

Table 10: Training efficiency comparison across six benchmarks. We report training time and the relative speedup of FenDiM over the average training time of compared baselines.

Dataset	FenDiM Time (s)	Avg. Baseline Time (s)	Speedup
ETTh1	97.1	335.3	3.45×
ETTh2	93.8	214.0	2.28×
ETTm1	132.2	741.7	5.61×
ETTm2	278.2	932.4	3.35×
Weather	37.7	114.7	3.04×
Exchange	149.4	1030.0	6.89×
Average	131.4	561.4	4.27×

Table 11: PDR backbone comparison on ETTh1 and Exchange. The best result in each column is shown in bold. Ties are both bolded.

Dataset	**ETTh1**								**Exchange**							
I/O	96/96		96/192		96/336		96/720		96/96		96/192		96/336		96/720	
Metrics	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**
LCR	0.116	0.311	0.116	0.314	0.117	0.317	0.118	0.322	0.366	0.326	0.387	0.379	0.417	0.442	0.461	0.564
PSSR	**0.114**	0.309	0.115	0.312	**0.115**	0.313	**0.116**	**0.319**	0.377	0.339	0.396	0.385	0.424	0.452	0.470	0.563
HCSR	**0.114**	**0.308**	**0.114**	**0.311**	**0.115**	**0.312**	**0.116**	**0.319**	**0.181**	**0.179**	**0.179**	**0.175**	**0.181**	**0.195**	**0.186**	**0.231**

total inference time is reduced from 0.275 to 0.169, yielding a 38.5% reduction. In contrast, the FFT and spatiotemporal modules contribute only a small fraction of the total runtime. These results indicate that the main computational bottleneck lies in diffusion-based rectification, and that Mamba effectively alleviates this bottleneck through efficient long-range sequence modeling.

End-to-end training efficiency. Table 10 reports the training time of FenDiM and representative forecasting baselines across six benchmarks. Although FenDiM introduces diffusion-based rectification, it remains efficient in full training. Compared with the average training time of the baselines, FenDiM achieves an average speedup of $4.27\times$, with dataset-wise speedups ranging from $2.28\times$ on ETTh2 to $6.89\times$ on Exchange. This suggests that the Mamba-based rectification design makes diffusion-guided dependency correction practical for long-horizon forecasting.

E.1 Model Analysis

PDR Backbone Analysis. We evaluate three backbone choices for PDR while keeping all other components fixed: *LCR*, *HCSR*, and *PSSR*. As shown in Table 11, *HCSR* achieves the best or tied-best results in all ETTh1 columns and shows a clear advantage on Exchange. Averaged over the four horizons, *HCSR* reduces MAE and MSE on Exchange by 55.4% and 54.4%, respectively, compared with the strongest alternative backbone. On ETTh1, where the competing backbones already perform closely, *HCSR* remains competitive and slightly improves the averaged MAE and MSE by 0.2% and 0.2%, respectively. These results indicate that PDR benefits from introducing *Mamba* as the core rectification operator, while retaining local convolutional projections remains important for stable dependency correction.

LCR captures local continuity well through convolutional residual blocks, yet remains limited in repairing long-range dependency distortions. *PSSR*, by contrast, strengthens long-range propagation through a fully Mamba-based architecture, but sacrifices local structural anchoring after removing convolutional projection stages that explicitly organize observation and mask information. *HCSR* strikes the best balance: convolutional projections preserve local evidence fidelity, while the Mamba backbone enables efficient propagation of corrected dependency information over long ranges. Since PDR aims at progressive correction of perturbation-distorted dependencies rather than simple denoising, this hybrid Mamba-centered design is better aligned with the role of the module.

Table 12: Comparison of FCTL variants on ETTh1 and Exchange. The best result in each column is shown in bold.

Dataset	ETTh1								Exchange							
I/O	96/96		96/192		96/336		96/720		96/96		96/192		96/336		96/720	
Metrics	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
FCTL-Mag	0.108	**0.178**	0.108	0.296	0.108	0.299	0.112	**0.303**	0.063	0.091	0.066	0.093	0.069	**0.097**	0.077	0.136
FCTL-PESA	**0.089**	0.187	**0.099**	**0.224**	**0.101**	**0.244**	**0.111**	0.304	**0.061**	**0.082**	**0.063**	**0.090**	**0.066**	0.107	**0.073**	**0.123**

FCTL Phase-Encoding Analysis. To further examine the role of phase information in FCTL, we provide an illustration of phase-aware spectral anchoring in Figure 5 and compare two variants while keeping all other components fixed: **FCTL-PESA**, the full model with phase-encoded frequency anchoring, and **FCTL-Mag**, which removes phase encoding and retains only magnitude-based frequency statistics. As shown in Table 12, **FCTL-PESA** achieves the best MAE in all eight horizon-dataset settings and the best MSE in five of the eight settings. Averaged over ETTh1 and Exchange, **FCTL-PESA** reduces MAE and MSE by 6.8% and 8.8%, respectively, compared with **FCTL-Mag**. These results indicate that phase-aware frequency structure provides more informative spectral anchoring than magnitude-only conditioning, although magnitude-only conditioning remains competitive in several MSE cases.

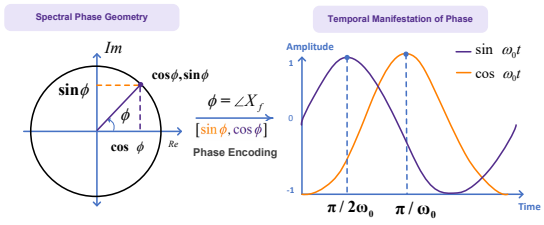

Figure 5: **Phase-aware spectral anchoring in FCTL.** Magnitude captures the distribution of spectral energy across frequency bands, whereas phase preserves temporal alignment, lag structure, and oscillatory coordination. FCTL-PESA encodes phase with sinusoidal descriptors and aggregates phase-aware frequency cues at the band level, providing a more discriminative structural anchor for forecasting under corrupted or incomplete observations.

The advantage of **FCTL-PESA** lies in making the frequency anchor more informative for temporal modeling. Magnitude characterizes how frequency energy is distributed across frequency bands, but it does not specify how oscillatory components are aligned in time. As illustrated in Figure 5, phase encoding complements magnitude information by preserving temporal alignment, lag structure, and oscillatory coordination. Under corrupted or incomplete observations, different temporal patterns may exhibit similar magnitude spectra while corresponding to substantially different temporal organizations. As a result, magnitude-only frequency conditioning remains insufficient for disambiguating biased temporal evidence. By additionally encoding phase through sinusoidal descriptors and aggregating it at the band level, **FCTL-PESA** provides a stable and discriminative structural anchor while avoiding angular discontinuity and local phase instability.

This effect is tightly connected to the role of FCTL in our framework. FCTL is introduced not as an auxiliary frequency branch, but as a structural conditioner that stabilizes temporal learning when local observations are noisy or unreliable. The comparison with **FCTL-Mag** shows that incorporating phase information improves overall forecasting accuracy, especially in MAE, while the remaining competitive MSE cases of magnitude-only conditioning suggest that magnitude statistics still provide useful spectral regularization. These results validate that the effectiveness of FCTL primarily stems from injecting phase-aware frequency structure into the forecasting process.

F Theoretical Insight into Dependency Rectification

We provide a theoretical view of why FenDiM improves forecasting robustness under imperfect observations. The purpose of this analysis is not to establish a distribution-free guarantee, but to clarify the mechanism behind prediction-oriented dependency rectification. Our key claim is that observational perturbations do not merely corrupt individual values; more fundamentally, they bias the dependency evidence extracted from the observation space. FenDiM addresses this issue through

two complementary mechanisms: progressive dependency rectification, which moves perturbed representations toward a cleaner dependency-expression space, and frequency structural anchoring, which provides a lower-sensitivity global reference when local temporal evidence is unreliable.

Let $X \in \mathbb{R}^{L\times C}$ denote the clean historical sequence and let the accessible observation be

$$\widetilde{X} = \Omega \odot X + (1-\Omega)\odot \Phi + E + \Delta, \tag{23}$$

where $\Omega \in \{0,1\}^{L\times C}$ is the observation mask, Φ denotes the placeholder used for unavailable values, E denotes stochastic noise, and Δ denotes a generic localized distortion term. In the main experiments, this perturbation model is primarily instantiated by missingness, Gaussian noise, and compound missing-noise corruption; Δ is retained to represent localized deviations or abrupt distortions that may occur in real-world observations.

Following the methodology, FenDiM first maps the observed history into a latent dependency representation. Let

$$Z_0 = G(\widetilde{X}, \Omega), \qquad Z^\star = G(X, \mathbf{1}), \tag{24}$$

where G denotes the temporal encoder used to construct the initial dependency representation, Z_0 is the representation extracted from the imperfect observation, and $Z^\star$ is the oracle representation that would be obtained from the clean history. To formalize dependency formation, let ϕ be a relational feature map and define the dependency evidence induced by a latent representation Z as

$$\mathcal{A}(Z) = \phi(Z)\phi(Z)^\top. \tag{25}$$

Here, $\mathcal{A}(Z)$ serves as a generic relational substrate upon which the downstream predictor relies.

Perturbation-induced dependency bias. Assume that G and ϕ are Lipschitz continuous with constants L_G and L_ϕ, respectively, and that $\|\phi(Z)\|_F \le C_H$ for all representations encountered during training. Then,

$$\|\mathcal{A}(Z_0) - \mathcal{A}(Z^\star)\|_F \le 2C_H L_\phi \|Z_0 - Z^\star\|_F. \tag{26}$$

Using the Lipschitz continuity of G, we further obtain

$$\|\mathcal{A}(Z_0) - \mathcal{A}(Z^\star)\|_F \le 2C_H L_\phi L_G \|\widetilde{X} - X\|_F. \tag{27}$$

Substituting Eq. (23) gives

$$\|\mathcal{A}(Z_0) - \mathcal{A}(Z^\star)\|_F \le 2C_H L_\phi L_G \Big(\|(1-\Omega)\odot(X-\Phi)\|_F + \|E\|_F + \|\Delta\|_F \Big). \tag{28}$$

Equation (28) shows that missing entries, stochastic noise, and localized distortions can all bias the dependency evidence extracted from observed histories. Thus, forecasting degradation under imperfect observations originates not only from point-wise value corruption, but also from biased dependency formation.

Effect of progressive dependency rectification. Let the output of the PDR module be

$$Z_{\mathrm{rect}} = \mathcal{R}_\theta(Z_0, Z_{\mathrm{cond}}, \Omega_{\mathrm{cond}}), \tag{29}$$

where Z_{cond} and Ω_{cond} denote the conditional evidence used by the rectifier. We assume that the learned rectification process reduces the expected distance from the clean dependency representation:

$$\mathbb{E}\big[\|Z_{\mathrm{rect}} - Z^\star\|_F^2\big] \le \kappa^2 \mathbb{E}\big[\|Z_0 - Z^\star\|_F^2\big], \qquad 0 < \kappa < 1. \tag{30}$$

This assumption abstracts the intended effect encouraged by the conditional diffusion objective and the joint forecasting loss: the rectifier is expected to suppress perturbation-induced components in the latent dependency space rather than perform unconditional value denoising.

Combining Eq. (26) with Eq. (30) yields

$$\mathbb{E}\big[\|\mathcal{A}(Z_{\mathrm{rect}}) - \mathcal{A}(Z^\star)\|_F^2\big] \le 4C_H^2 L_\phi^2 \kappa^2 \mathbb{E}\big[\|Z_0 - Z^\star\|_F^2\big]. \tag{31}$$

This result justifies interpreting PDR as dependency rectification: if the rectified representation becomes closer to the clean dependency representation, the induced relational structure also becomes more faithful to the underlying forecasting evidence.

Stability of frequency structural anchoring. PDR may still leave residual local unreliability in the temporal representation. FenDiM therefore introduces a frequency structural branch to provide global temporal evidence. To analyze its stabilizing effect, consider an idealized band-averaged spectral descriptor

$$\mathcal{Q}(Z) = \mathcal{P}(\log(1 + |\mathcal{F}Z|)), \tag{32}$$

where $\mathcal{F}$ denotes the unitary discrete Fourier transform along the temporal dimension and $\mathcal{P}$ denotes average pooling over disjoint frequency bands of size b. Since the Fourier transform is unitary, the magnitude and $\log(1 + \cdot)$ operations are non-expansive, and average pooling reduces sensitivity over each band, $\mathcal{Q}$ is Lipschitz with constant $L_Q \le 1/\sqrt{b}$:

$$\|\mathcal{Q}(U) - \mathcal{Q}(V)\|_F \le \frac{1}{\sqrt{b}}\|U - V\|_F. \tag{33}$$

Equation (33) formalizes the role of frequency structural anchoring: after band aggregation, the global spectral descriptor is less sensitive to localized perturbations. The implemented FCTL follows this principle by using log-magnitude and phase-aware frequency encodings for modulation, while the spectral consistency loss constrains the forecasted log-magnitude spectrum to preserve stable future-relevant structure.

Forecasting risk under joint calibration. Let the temporal branch be denoted by $\mathcal{T}$, the fusion operator by Ψ, and the prediction head by g. The final prediction is

$$\widehat{Y} = g\big(\Psi\big(\mathcal{T}(Z_{\mathrm{rect}}), \mathcal{Q}(Z_{\mathrm{rect}})\big)\big). \tag{34}$$

The corresponding oracle clean prediction is

$$\widehat{Y}^\star = g\big(\Psi\big(\mathcal{T}(Z^\star), \mathcal{Q}(Z^\star)\big)\big), \tag{35}$$

and the oracle forecasting risk is

$$\mathcal{R}_{\mathrm{oracle}} = \mathbb{E}\Big[\|\widehat{Y}^\star - Y\|_2^2\Big]. \tag{36}$$

Assume that $\mathcal{T}$, Ψ, and g are Lipschitz continuous with constants $L_{\mathcal{T}}$, L_Ψ, and L_g, respectively. Then,

$$\mathbb{E}\Big[\|\widehat{Y} - Y\|_2^2\Big] \le 2\mathcal{R}_{\mathrm{oracle}} + 2L_g^2 L_\Psi^2 \left(L_{\mathcal{T}} + L_Q\right)^2 \mathbb{E}\big[\|Z_{\mathrm{rect}} - Z^\star\|_F^2\big]. \tag{37}$$

Combining Eq. (37) with Eq. (30) gives

$$\mathbb{E}\Big[\|\widehat{Y} - Y\|_2^2\Big] \le 2\mathcal{R}_{\mathrm{oracle}} + 2L_g^2 L_\Psi^2 \left(L_{\mathcal{T}} + L_Q\right)^2 \kappa^2 \mathbb{E}\big[\|Z_0 - Z^\star\|_F^2\big]. \tag{38}$$

Together with Eq. (28), this indicates that the excess forecasting risk is controlled by the magnitude of observation-induced perturbations, the rectification fidelity of PDR, and the stability of the frequency structural anchor.

Summary. This analysis reveals a coherent division of labor in FenDiM. Imperfect observations first bias dependency formation in the latent observation space. PDR counteracts this effect by moving the representation toward a cleaner dependency-expression space. FCTL supplies a lower-sensitivity global spectral reference when local temporal evidence remains unreliable. The joint objective then aligns prediction accuracy, dependency rectification, and spectral consistency with forecasting utility, rather than optimizing independent reconstruction or frequency-matching goals.