STABLE-STRUCTURE-GUIDED SPECTRAL CORRUPTION FOR DIFFUSION-BASED MULTIVARIATE TIME SERIES ANOMALY DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Diffusion-based reconstruction has shown promising results in multivariate time series anomaly detection, using discrepancies between observations and their reconstructions to identify anomalies. However, isotropic corruption may affect the input information available to the denoiser, raising the question of how to design anisotropic corruption based on structural characteristics. Existing frequency-aware diffusion methods typically allocate corruption under prior assumptions that equate absolute frequency location or spectral energy with information value; these assumptions may overlook normal high-frequency transients and low-energy frequency components that encode reproducible, stable structures. To address this issue, we propose a stable-structure-guided diffusion framework. The framework constructs adaptive soft frequency bands from local window features and combines local spectral prototype support, cross-window reproducibility, and cross-variable relational consistency provided by a stable reference bank to form band-level structural evidence. This evidence guides the allocation of corruption strengths to preserve components supported by stable patterns during denoising. Across five multivariate benchmark datasets, the framework improves mean F1 by 33.10% over the state-of-the-art diffusion model and by 24.86% over the state-of-the-art model among all evaluated baselines, which comprise 41 representative anomaly detection methods.

1 INTRODUCTION

Multivariate time series anomaly detection (MTSAD) aims to learn the structure of normal system operation and identify observations that deviate from normal system behavior. It is widely applied to financial data monitoring (Crépey et al., 2022), industrial control (Li et al., 2019), spacecraft telemetry (Hundman et al., 2018), and network intrusion detection (Mirsky et al., 2018). However, real-world systems often exhibit long-term trends, periodic oscillations, short-lived transients, and cross-variable coupling simultaneously (Zhong et al., 2025; Wu et al., 2025). These components of normal operation are superimposed on noise and disturbances in real-world systems, making it more difficult to distinguish normal variations, background interference, and genuine anomalies (Wang et al., 2023; Zhong et al., 2025).

Diffusion models identify anomalies by recovering normal patterns and have achieved competitive performance in multivariate time series anomaly detection (Xiao et al., 2023; Chen et al., 2023). Standard DDPMs apply isotropic corruption, using the same signal attenuation coefficient and noise intensity across frequency components (Ho et al., 2020). However, this treatment causes low-frequency information to decay rapidly and affects training stability (Wang et al., 2025), suggesting that the allocation of corruption across frequency components influences both the information available for reconstruction and the model's learning process.

Existing work primarily addresses this issue using two types of prior assumptions, taking either absolute frequency location or spectral energy as the basis for allocating corruption. The first prior associates information value with absolute frequency location, assuming that low-frequency components contain more useful information (Wu et al., 2021), and preferentially corrupts high-frequency components through progressive smoothing (Wang et al., 2025). However, this assumption may

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

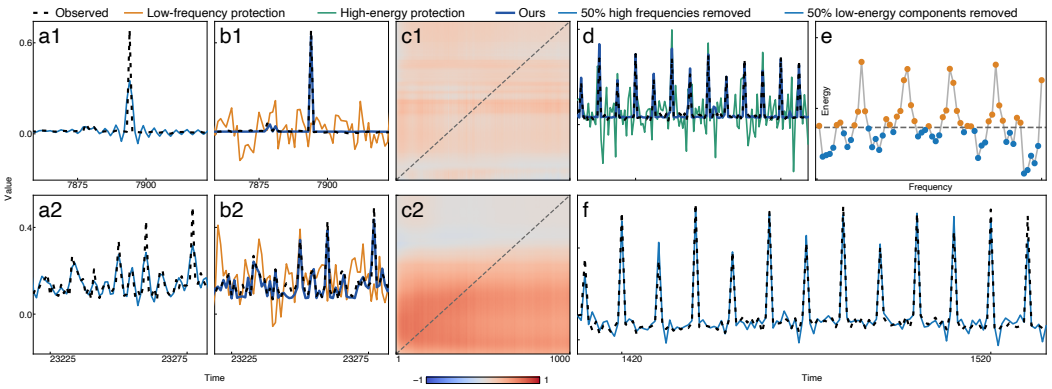

Figure 1: Analysis of two protection priors on the SMD dataset. (a1, a2) Loss of detail caused by removing high-frequency components; (b1, b2) structural protection reconstructs transient changes better than low-frequency protection; (c1, c2) gradient cosine similarity over 1000 training steps, reflecting differences in the update directions of the two strategies; (d) insufficient reconstruction of low-energy details by energy protection; (e, f) removal of low-energy frequency components and its effect on the time-domain series.

overlook transient information carried by high-frequency components. Across the 22 datasets examined (Qiu et al., 2025a), a criterion based on absolute frequency location covers only 38.82% of anomaly occurrences under the statistical definition used in this study.[1] Figure 1(a1, a2) shows that removing high-frequency components in the low-frequency protection test loses information characterizing short-lived transient changes; in Figure 1(b1, b2), the low-frequency-protection strategy insufficiently reconstructs some transient spikes and introduces additional fluctuations, whereas the structural-protection strategy more closely matches the original series. The second prior assesses information value using spectral power, assuming that high-power components contain more useful information and should therefore receive stronger protection (Shen et al., 2026). However, the association between spectral energy and anomaly occurrence is approximately 45.1% on these datasets.[1] Figure 1(e, f) shows that after low-energy frequency components are removed, the major spikes remain well preserved, but additional fluctuations appear in the originally relatively smooth intervals between spikes. The energy-protection strategy in Figure 1(d) exhibits a similar phenomenon: the major spikes are reconstructed well, whereas reconstruction deviations are more pronounced in smooth intervals. Therefore, a key question remains: **how should corruption be allocated across frequency components to preserve valuable information and support the reconstruction of normal patterns?**

These observations motivate assessing the value of frequency-domain information through system structure. Multivariate time series can be viewed as multivariate observations generated by latent system dynamics (Karl et al., 2017), whose structural meaning arises from evolving system states and inter-variable coupling (Kipf et al., 2018). Assessing a frequency component therefore requires considering not only its frequency location and spectral power, but also whether it consistently represents system states across time and captures coordinated changes across variables. Real-world time series often lack explicit physical interpretations; in such cases, structural relationships estimated from data provide evidence for information selection. Moreover, normal patterns observed over long periods may evolve as the operating environment changes (Zhu et al., 2023), with phenomena such as concept drift or equipment aging. Drawing on the modeling approach of Hallac et al. (2017), which characterizes time series through stable dependencies within each state, we consider that the system has its own relatively stable normal patterns at different operating stages, whose stability is reflected in temporal structures and cross-variable dependencies within each stage.

Building on this perspective, we propose a frequency-aware diffusion framework guided by evidence of stable structure, where stability refers to the recurrence of patterns and relationships rather than slow signal variation. The core idea is to allocate corruption strengths according to the stable-structure evidence of different frequency bands, providing stronger protection to bands with stronger structural evidence. The framework first organizes local spectral information using

[1] See Appendix A for the statistical definition and methodology.

window-dependent Gaussian soft frequency bands. It then constructs a reference bank based on stable operating states to jointly assess support from local spectral prototypes, cross-window reproducibility, and consistency of cross-variable relationships. Local structural information thus contributes to structural assessment rather than being directly equated with information value. The resulting structural evidence guides forward corruption by reallocating the corruption budget across frequency bands while leaving the global DDPM diffusion schedule unchanged. This design decouples the overall progression of diffusion from noise allocation across bands. Bands with stronger structural evidence receive weaker corruption, with the aim of preserving reproducible stable structures and reducing interference from irrelevant components during reconstruction. Our method, guided by evidence of stable structure, provides a basis for allocating corruption strengths across frequency components and can be extended to univariate anomaly detection by removing cross-variable relationships. Across five univariate benchmarks, our method improves AUROC by 10.99%.

Our contributions are summarized as follows:

- We propose normal-reference consistency as the basis for allocating corruption across frequency components, assessing structural evidence by measuring the consistency between the current window and stable operating reference states rather than relying solely on absolute frequency location or spectral energy.
- We propose a spectral corruption allocation framework guided by evidence of normal-reference consistency. The framework combines adaptive Gaussian soft frequency bands with a frozen normal reference bank and integrates local spectral prototype support, cross-window reproducibility, and cross-variable relational consistency. While keeping the global scalar diffusion schedule and the total noise budget unchanged, it reallocates noise variance across frequency components so that bands with stronger structural evidence undergo weaker corruption.
- Across five multivariate benchmark datasets, our method improves mean F1 by 33.10% over the strongest diffusion baseline; compared with representative methods using isotropic noise, frequency-position-based protection, and spectral-energy-based protection, it achieves improvements of 33.10%, 88.78%, and 150.53%, respectively.

2 Related Work

2.1 Diffusion Models for Time Series Anomaly Detection

Diffusion models have increasingly been applied to time series analysis. TimeGrad (Rasul et al., 2021) and CSDI (Tashiro et al., 2021) use diffusion models for multivariate probabilistic forecasting and conditional imputation, respectively. For anomaly detection, DiffAD (Xiao et al., 2023) and ImDiffusion (Chen et al., 2023) adopt masked imputation or conditional reconstruction to identify anomalies from discrepancies between observations and reconstructions. Their forward processes retain standard isotropic corruption without explicitly distinguishing perturbation strengths across frequency components. Subsequent studies have incorporated time series characteristics into diffusion and reconstruction. For example, PhysDiff (Li et al., 2025) extracts high- and low-frequency components through adaptive frequency decomposition, uses them as physical priors, and dynamically guides reverse denoising through a dual-path conditioning mechanism and frequency-routing attention. However, this frequency-structure information primarily serves as additional conditioning for reconstruction; the forward process still uses isotropic corruption without adapting perturbation strengths to the structural characteristics of different frequency components.

2.2 Frequency-Domain Modeling vs Frequency-Aware Diffusion

Modeling with frequency-domain representations. These methods complement time-domain information with frequency-domain representations to characterize different types of anomalies. TFAD (Zhang et al., 2022) decomposes time series into trend and residual components, learns representations through time- and frequency-domain branches, and combines information from both domains for anomaly detection. CATCH (Wu et al., 2025) divides the spectrum into frequency patches and uses patch-level masks and attention mechanisms to learn variable correlations within different frequency bands, capturing fine-grained spectral features.

Designing diffusion corruption in the frequency domain. These methods adjust corruption strengths according to frequency characteristics, replacing isotropic corruption with anisotropic corruption. MATSD (Wang et al., 2025) uses moving-average kernels to progressively smooth signals, preferentially corrupting high-frequency components. FSDI (Shen et al., 2026) estimates frequency importance from the spectral power of the data and allocates noise variance to each frequency in proportion to the reciprocal of its importance, while preserving the total noise energy. The two methods organize corruption allocation according to frequency location and spectral power, respectively.

3 PRELIMINARIES

3.1 PROBLEM DEFINITION

Let $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_T]^\top \in \mathbb{R}^{T \times C}$ denote a multivariate time series, where T is the sequence length, C is the number of variables, and $\mathbf{x}_t \in \mathbb{R}^C$ is the observation at time t. Multivariate time series anomaly detection aims to identify anomalous observations at individual time points (Chen et al., 2023). In our setting, we extract sliding windows of length L: $\mathbf{X}_i = [\mathbf{x}_i, \ldots, \mathbf{x}_{i+L-1}]^\top \in \mathbb{R}^{L \times C}$, where $1 \le i \le T - L + 1$. The detector assigns an anomaly score $a_t \in \mathbb{R}_{\ge 0}$ to each time point, with larger scores indicating stronger evidence of an anomaly.

3.2 DENOISING DIFFUSION PROBABILISTIC MODELS

Denoising diffusion probabilistic models (DDPMs; Ho et al., 2020) progressively add Gaussian noise through a Markov chain of length N. We use n to denote the diffusion step. The forward transition distribution is determined by the noise variance β_n and the signal retention coefficient α_n:

$$q(\mathbf{X}^n \mid \mathbf{X}^{n-1}) = \mathcal{N}(\sqrt{\alpha_n}\mathbf{X}^{n-1}, \beta_n \mathbf{I}), \qquad \alpha_n = 1 - \beta_n \tag{1}$$

Let $\bar{\alpha}_n = \prod_{j=1}^{n} \alpha_j$. A sample at any diffusion step can then be constructed directly from the original sample and standard Gaussian noise, without sequentially applying all forward transitions:

$$\mathbf{X}^n = \sqrt{\bar{\alpha}_n}\mathbf{X}^0 + \sqrt{1 - \bar{\alpha}_n}\boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}) \tag{2}$$

The reverse process gradually recovers the data through parameterized Gaussian distributions, whose means are predicted by the model and whose covariances are denoted by Σ_n:

$$p_\theta(\mathbf{X}^{n-1} \mid \mathbf{X}^n) = \mathcal{N}(\mu_\theta(\mathbf{X}^n, n), \Sigma_n) \tag{3}$$

A standard DDPM is typically trained to predict the noise ϵ at each diffusion step.

4 METHOD

4.1 WINDOW-CONDITIONED ADAPTIVE SPECTRAL PARTITIONING

Nearby time points typically exhibit stronger correlations than distant ones (Park & Kang, 2026). Meanwhile, local statistics and inter-variable relationships change across operating states, and fixed frequency-band boundaries shared by all windows cannot effectively represent local spectral structures. We therefore adopt adaptive spectral partitioning conditioned on the input window. To accommodate potential coupling across frequency locations, we use a soft partition that allows each frequency component to contribute to multiple bands with different weights.

We first apply the real fast Fourier transform along the temporal dimension to obtain a complex spectral representation of the input window, $\mathbf{U} = \mathrm{rFFT}(\mathbf{X}) \in \mathbb{C}^{K \times C}$, where $K = \lfloor L/2 \rfloor + 1$. The spectrum of each variable is then normalized separately to remove differences in units and amplitude. The normalized spectrum $\widetilde{\mathbf{U}}$ is used only to construct features for mask prediction. To jointly represent spectral amplitude and phase, inter-variable dependencies, and frequency location, we augment the real and imaginary parts of the normalized spectrum with log-magnitude and frequency coordinates to form the features in Equation (4).

$$\mathbf{q}_k = \left[\mathrm{Re}\,\widetilde{\mathbf{U}}_k; \mathrm{Im}\,\widetilde{\mathbf{U}}_k; \log\left(1 + \left|\widetilde{\mathbf{U}}_k\right|\right); f_k\right], \qquad f_k = \frac{\log(1+k)}{\log K} \tag{4}$$

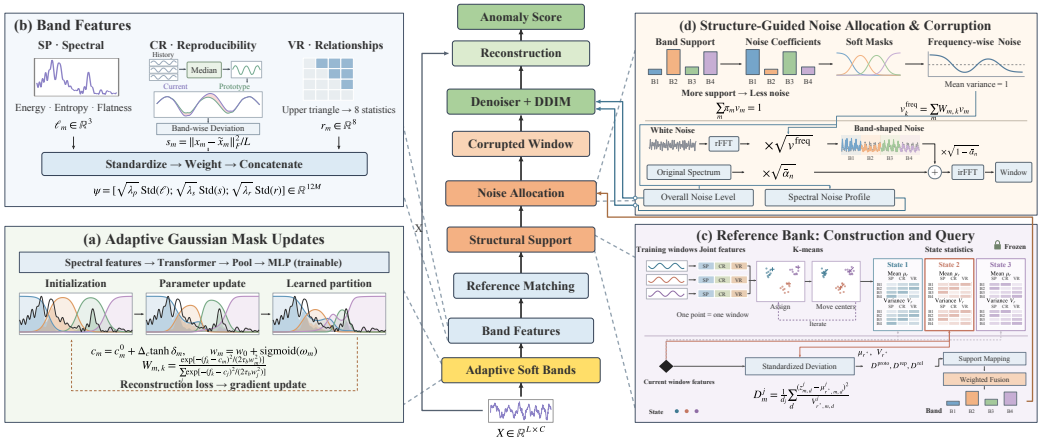

Figure 2: Overview of the proposed framework, which extracts band-level structural evidence to guide diffusion noise allocation, reconstruction, and anomaly scoring. (a) The frequency-band mask generator (FBMG) learns Gaussian soft bands. (b) Band features describe spectral properties, reproducibility, and variable relationships. (c) Pretraining alternates FBMG and reference-bank updates; main training reinitializes FBMG and freezes the bank. Current windows query reference statistics to compute structural support. (d) Support guides budget-conserving spectral noise allocation, followed by conditional reconstruction and residual-based anomaly scoring.

The real and imaginary parts retain the complex spectral information, while the log-magnitude compresses its dynamic range to reduce the dominance of a few high-energy components.

The joint representation is then encoded along the frequency axis, allowing each frequency location to aggregate contextual spectral information. The band generator uses the encoded representations to predict center offsets δ_m and scale parameters ω_m for M bands in the current window(5).

$$\{(\delta_m, \omega_m)\}_{m=1}^{M} = \mathrm{MLP}(\mathrm{Pool}(\mathrm{Transformer}(\mathbf{q}))) \tag{5}$$

To preserve band ordering and prevent scale collapse or uncontrolled expansion, we prescribe band centers c_m^0 and region boundaries w^0, and use bounded mappings to determine the adjusted band centers c_m and scales w_m(6).

$$c_m = c_m^0 + \Delta_c \tanh(\delta_m), \qquad w_m = w_0 + \mathrm{sigmoid}(\omega_m). \tag{6}$$

These parameters define Gaussian soft masks, which are normalized across bands(7), with τ_b controlling the partition temperature.

$$G(f; c, w) = \exp\left(-\frac{(f-c)^2}{2\tau_b w^2}\right), \quad W_{m,k} = \frac{G(f_k; c_m, w_m)}{\sum_{j=1}^{M} G(f_k; c_j, w_j) + \varepsilon}, \tag{7}$$
$$\mathbf{W} = [W_{m,k}] \in \mathbb{R}^{M \times K}$$

Here, ε prevents a zero denominator, and $K = \lfloor L/2 \rfloor + 1$. The resulting matrix $\mathbf{W} \in [0,1]^{M \times K}$ satisfies $\sum_{m=1}^{M} W_{m,k} = 1$ at every frequency location. Each band representation is obtained by weighting the original spectrum with its corresponding mask, $\mathbf{U}_m = \mathbf{W}_m \odot \mathbf{U}$, and the original window can be recovered through the inverse transform, $\mathbf{X} = \mathrm{irFFT}(\mathbf{U})$. Detailed computations and procedures are provided in Appendix B.

4.2 Frozen Joint Stable-Structure Reference

The importance of a frequency band is related to its absolute frequency location and spectral energy, but depends more on its ability to represent recurrent stable structures. We therefore seek to characterize structure through three complementary aspects: local spectral prototype support, cross-window reproducibility, and cross-variable relational consistency. We also construct a structural reference bank and use clustering to obtain reference prototypes.

Local spectral prototype support characterizes the spectral shape of a band. We use mean energy E, spectral entropy H, and spectral flatness F to describe overall band intensity, the concentration or dispersion of energy, and whether the spectrum is approximately flat or contains pronounced peaks, respectively. Stacking the three-dimensional local spectral features across all M bands gives

$$\ell = [\log(\mathbf{E} + \epsilon), \mathbf{H}, \mathbf{F}] \in \mathbb{R}^{M\times 3}, \qquad \mathbf{E}, \mathbf{H}, \mathbf{F} \in \mathbb{R}^{M}. \tag{8}$$

Cross-window reproducibility is assessed through deviation from a training-window prototype. Inspired by the use of median aggregation to obtain prototypes from images of the same class, we stack the training windows and take their element-wise median as a robust prototype. We then apply the same frequency mask to the current window and the prototype and measure their deviation within each band.

$$\begin{aligned} s_m &= \frac{1}{L}\left\|\mathbf{x}_m - \mathrm{irFFT}\left(W_m \odot \mathrm{rFFT}(\widetilde{\mathbf{x}});\, n = L\right)\right\|_F^2 \in \mathbb{R}, \\ \widetilde{\mathbf{x}} &= \mathrm{median}_{n\in\mathcal{I}_{\mathrm{train}}}\, \mathbf{x}_n, \qquad \mathbf{s} = [s_1, \ldots, s_M]^\top \in \mathbb{R}^M. \end{aligned} \tag{9}$$

Cross-variable relational consistency captures coordinated changes among variables. Multivariate anomalies may alter individual variables or disrupt their existing relationships. We therefore compute pairwise cross-variable spectral relationships within each band, extract the upper-triangular entries excluding the diagonal, and compress them into eight relational statistics. The resulting band-level representations are then concatenated across all M bands.

$$\begin{aligned} \mathbf{A}_m &= \mathrm{Rel}(\mathbf{U}, W_m) \in \mathbb{R}^{C\times C}, \quad \mathbf{q}_m = \mathrm{vec}_{\mathrm{upper}}(\mathbf{A}_m) \in \mathbb{R}^{C(C-1)/2}, \\ \mathbf{r}_m &= \mathrm{Stats}(\mathbf{q}_m) \in \mathbb{R}^8, \qquad\qquad \mathbf{r} = \mathrm{Concat}_{m=1}^{M}(\mathbf{r}_m) \in \mathbb{R}^{8M}. \end{aligned} \tag{10}$$

These three views assess whether within-window frequency-domain features occur stably, whether they recur across windows, and whether inter-variable dependencies remain stable. Together, they provide complementary descriptions of stable patterns through local structure, cross-window consistency, and variable coupling. We standardize the three feature groups separately and combine them into a $12M$-dimensional window-level joint representation.

$$\psi = \left[\sqrt{\lambda_{\mathrm{proto}}}\,\mathrm{Std}(\ell); \sqrt{\lambda_{\mathrm{rep}}}\,\mathrm{Std}(s); \sqrt{\lambda_{\mathrm{rel}}}\,\mathrm{Std}(r)\right] \in \mathbb{R}^{12M} \qquad \sum \lambda = 1 \tag{11}$$

Stable patterns may comprise multiple operating states with distinct reference characteristics. We cluster window-level joint representations ψ to identify these states and construct a reference bank of stable window-level structures. The bank is established before the main training stage and remains frozen thereafter.

For each input window, we select the nearest reference cluster $r^\star$ based on the standardized deviations of the three feature groups and use this cluster as a proxy for the latent operating state. We then use the means $\mu^j_{r^\star,m,d}$ and variances $V^j_{r^\star,m,d}$ of the three feature groups in each band within cluster $r^\star$ as stable band statistics to compute the local spectral prototype deviation D^{proto}, cross-window reproducibility deviation D^{rep}, and cross-variable relational deviation D^{rel} for each band of the current input window. The three deviations are mapped to support scores in $[0, 1]$ and combined by weighted fusion to obtain band-level normal-structure evidence e_m:

$$\begin{aligned} D^j_{r^\star,m} &= \frac{1}{\dim_j}\sum_{d=1}^{\dim_j}\frac{\left(z^j_{m,d} - \mu^j_{r^\star,m,d}\right)^2}{V^j_{r^\star,m,d}}, \qquad j \in \{\mathrm{proto}, \mathrm{rep}, \mathrm{rel}\}, \\ e_m &= \lambda_{\mathrm{proto}}\,\sigma\left(\frac{\tau_p - D^{\mathrm{proto}}_{r^\star,m}}{T_p}\right) + \lambda_{\mathrm{rep}}\,\exp\left(-\frac{D^{\mathrm{rep}}_{r^\star,m}}{2}\right) + \lambda_{\mathrm{rel}}\,\exp\left(-\frac{D^{\mathrm{rel}}_{r^\star,m}}{2}\right). \end{aligned} \tag{12}$$

Here, $\sigma(x) = 1/(1 + \exp(-x))$ denotes the sigmoid function. Here, m indexes frequency bands, $z^j_{m,d}$ is the corresponding feature component of the current window, and $\dim_j$ is the dimensionality of feature group j. The fusion weights satisfy $\lambda_j \geq 0$ and $\sum_j \lambda_j = 1$. The parameters τ_p and T_p control the deviation threshold and the smoothness of the mapping for prototype support, respectively. A larger e_m indicates stronger consistency between the band and the selected reference state; this score is used for subsequent noise allocation. Detailed computations are provided in Appendix C.

4.3 BUDGET-CONSERVING STRUCTURE-GUIDED DIFFUSION

Standard diffusion models use isotropic noise corruption. Our method reallocates corruption across frequency bands using structural evidence while preserving the global diffusion schedule and total noise budget. The diffusion step determines the overall corruption level, whereas the spectral structure of the current window determines its allocation across bands. For an input window, let $\mathbf{e} = [e_1, \ldots, e_M]$ denote band-level normal-structure evidence and $\boldsymbol{\pi} = [\pi_1, \ldots, \pi_M]$ denote the spectral masses of the bands. The budget mapping $\mathcal{B}$ converts structural evidence into band-level corruption coefficients[(13)].

$$\mathbf{v} = \mathcal{B}(-\mathbf{e}; \boldsymbol{\pi}) \in \mathbb{R}^M \tag{13}$$

Bands with stronger normal-structure evidence receive weaker corruption, guiding the model toward informative patterns and reducing interference from irrelevant components. The soft band masks map the corruption coefficients determined by band-level structural evidence to individual frequency locations, $v_k^{\text{freq}} = \sum_{m=1}^{M} W_{m,k} v_m$. Let $\bar{\alpha}_n$ denote the cumulative signal-retention factor at diffusion step n. We obtain frequency-domain noise by applying the rFFT to standard Gaussian noise sampled in the time domain, $\Xi = \mathrm{rFFT}(\epsilon)$. In the structure-guided forward process[(14)], the global schedule $\bar{\alpha}_n$ controls the overall corruption level, while $\sqrt{v_k^{\text{freq}}}$ determines how corruption is distributed across frequencies.

$$\mathbf{U}^n = \sqrt{\bar{\alpha}_n}\,\mathbf{U}^0 + \sqrt{1-\bar{\alpha}_n}\sqrt{v_k^{\text{freq}}}\,\Xi, \qquad \mathbf{X}^n = \mathrm{irFFT}(\mathbf{U}^n) \tag{14}$$

Band-selective corruption allows the same diffusion step to produce different frequency-wise corruption patterns. To make these differences explicit to the denoiser, we construct two complementary conditions from the effective corruption strength $p_k^n = (1-\bar{\alpha}_n)v_k^{\text{freq}}$. The condition C_{mean}^n represents the overall corruption level, whereas C_{spec}^n represents frequency-wise deviations from the average level. The denoising network follows the CSDI architecture (Tashiro et al., 2021) and predicts the uncorrupted window from the corrupted input and these two conditions[(15)].

$$\widehat{\mathbf{X}}^{(0)} = f_\theta(\mathbf{X}^n, C_{\text{mean}}^n, C_{\text{spec}}^n) \tag{15}$$

The conditions inform the denoiser of the overall recovery strength and the differences required across frequencies. During training, we sample diffusion steps uniformly, draw Gaussian noise, and optimize the model using an L_1 reconstruction loss[(16)].

$$\mathcal{L} = \mathbb{E}_i\left[\left\|\widehat{\mathbf{X}}_i - \mathbf{X}_i\right\|_1\right] \tag{16}$$

4.4 DETERMINISTIC RECONSTRUCTION AND ANOMALY SCORING

At inference, the model determines the adaptive frequency-band partition and structural corruption coefficients from the input window and holds them fixed throughout reconstruction. After the input is corrupted to diffusion step 50, the denoiser predicts the clean window under the guidance of the step-dependent conditions C_{mean}^n and C_{spec}^n. Inference uses 50-step deterministic DDIM sampling (Song et al., 2021) with $\eta_{\text{DDIM}} = 0$, introducing no additional random noise during the reverse process. We then compute element-wise absolute reconstruction errors and standardize them using per-variable statistics of training calibration residuals. The standardized errors are averaged first over overlapping windows and then over variables to obtain a point-wise anomaly score.

All five datasets use the TAB anomaly-ratio grid $[0.1, 0.5, 1, 2, 3, 5, 10, 15, 20, 25]\%$ (Qiu et al., 2025a). We also aim to avoid substantial distortions in method comparison introduced by point adjustment (Kim et al., 2022).

5 EXPERIMENTS

5.1 EXPERIMENTAL SETUP

Datasets and Evaluation Metrics. We evaluate our method on five real-world datasets: Mars Science Laboratory rover (MSL; Hundman et al., 2018), Soil Moisture Active Passive satellite (SMAP;

Hundman et al., 2018), Server Machine Dataset (SMD; Su et al., 2019), Pooled Server Metrics (PSM; Abdulaal et al., 2021), and Secure Water Treatment (SWaT; Mathur & Tippenhauer, 2016). We report F1, AUPRC, and AUROC. Results for 13 evaluation metrics are provided in Appendix D.

Baseline Models. We compare the reported baseline results with the full model (Full) using the TAB grid evaluation described above (Qiu et al., 2025a).

Table 1: Results on five benchmark datasets (%). Double and single underlines indicate the best and second-best results, respectively. The two upper panels jointly rank non-diffusion models and report F1; the lower panel separately ranks diffusion models and reports F1, AUPRC, and AUROC. See Appendix D.2 for univariate datasets. MA-TSD and FSDI are extended to anomaly detection based on their core code.

Model	MSL	SMAP	SMD	PSM	SWaT
ML					
IForest	2.50	0.40	3.92	0.71	0.41
LOF	19.30	27.80	9.50	43.51	21.74
OCSVM	12.62	5.50	15.80	38.72	23.00
CBLOF	12.71	5.42	15.53	42.60	49.80
EIF	12.83	22.52	18.24	46.41	68.80
HBOS	12.74	6.00	14.84	38.92	41.73
KMeans	18.44	6.30	16.90	57.60	30.74
KNN	25.60	28.30	11.20	43.64	21.70
LODA	17.34	0.40	15.50	38.30	67.80
PCA	13.33	5.91	17.61	42.40	26.10
DL					
AE	2.44	0.11	0.41	17.80	20.00
VAE	2.70	0.20	2.10	1.11	32.40
OmniAnomaly	17.76	13.67	17.08	41.79	49.72
USAD	20.23	12.03	18.21	38.56	7.68
GANF	15.93	13.12	17.52	50.53	69.64
DAGMM	4.34	0.11	0.00	0.70	22.91
LSTMED	2.71	0.23	2.54	16.20	5.64
Ours	**30.45**	25.51	**45.10**	40.93	68.78

Model	MSL	SMAP	SMD	PSM	SWaT
DL (cont.)					
Torsk	16.14	21.04	10.20	28.64	14.90
TranAD	9.63	0.24	16.33	40.54	32.30
AnomTrans	5.60	4.60	2.30	1.92	1.94
DCdetector	4.90	8.04	6.62	8.40	7.12
ContraAD	2.03	0.11	0.00	2.12	0.43
DLinear	11.31	11.00	17.44	4.70	5.13
NLinear	11.71	11.01	17.72	4.60	5.94
DUET	11.00	10.94	16.00	9.44	6.64
iTransformer	11.61	10.50	18.93	10.90	3.91
PatchTST	11.13	11.00	18.70	11.11	6.60
TimesNet	14.21	11.50	19.10	11.70	8.20
ModernTCN	11.40	10.73	15.13	12.42	9.24
Frequency					
TFAD	0.98	0.98	0.94	1.00	0.99
DualTF	1.91	16.30	13.54	50.60	21.20
TFMAE	10.89	21.30	4.62	15.78	19.09
CATCH	14.43	10.24	22.23	8.30	6.41
Ours	**30.45**	25.51	**45.10**	40.93	68.78

Diff-Model	MSL			SMAP			SMD			PSM			SWaT		
DiffAD	24.63	18.4	65.5	17.31	13.0	51.6	8.54	5.17	54.0	30.04	30.9	54.0	8.94	11.7	20.4
ImDiffusion	17.03	13.8	54.9	15.32	12.0	49.5	18.91	11.64	70.1	27.90	29.6	51.8	6.51	11.7	50.7
D3R	24.01	15.8	66.3	12.81	10.7	41.8	31.34	23.82	74.4	25.37	26.7	47.9	13.41	14.3	59.7
AnomFilter	21.83	14.5	59.5	18.71	13.5	53.4	12.96	9.65	58.9	33.40	35.9	56.5	16.55	12.8	56.3
MATSD	16.57	11.6	47.8	11.77	12.6	40.4	27.41	20.13	78.6	42.70	43.7	64.5	13.20	10.7	36.0
ICDiffAD	18.62	15.6	60.7	13.23	11.9	50.5	17.48	9.86	71.7	36.04	40.8	66.7	8.14	13.0	54.4
FSDI	15.49	10.8	51.1	16.42	12.5	49.1	8.40	4.72	53.7	26.83	28.0	50.6	16.99	12.5	51.3
PhysDiff	17.61	15.3	53.5	15.81	12.3	47.7	16.15	9.39	61.4	39.57	40.8	59.8	69.22	70.6	81.5
Ours	**30.45**	**25.52**	**71.03**	**25.51**	**16.90**	**62.01**	**45.10**	**42.83**	**80.70**	40.93	**46.45**	62.40	68.78	69.66	79.50

5.2 Baseline Comparison

Table 1 compares our method with existing anomaly detection methods. Compared with diffusion baselines whose performance varies considerably across datasets, our method remains highly competitive. Further analysis shows that our method not only retains some anomalies already detected by the baselines but also identifies anomalies they miss. On MSL and SMAP, compared with DiffAD, ImDiffusion, D3R, AnomFilter, and ICDiffAD, the additional detection proportion is 29.36%–54.14%, while the false positive rate decreases by 5.02%–10.85% in relative terms. On SWaT, our method additionally detects 69.60% of the anomalous points jointly missed by MA-TSD, which uses low-frequency protection, and FSDI, which uses high-energy protection, while reducing the false positive rate by approximately 89% relative to each. These results demonstrate combined improvements in anomaly detection and false positive suppression.

5.3 Case Study

Figure 3 shows reconstruction differences under different criteria for corruption allocation. Frequency position and spectral energy do not directly correspond to normality, as anomalous changes

may also contain low-frequency or high-energy components. In Figure 3(a), the reconstructions of MA-TSD and FSDI follow the anomalous fluctuations more closely, whereas our method is closer to the normal constant pattern, retaining larger anomaly residuals. In Figure 3(b), MA-TSD exhibits oscillations near state-transition boundaries, and both baselines also follow the input more closely within the anomalous interval, while our method maintains a relatively stable reconstruction plateau. In Figure 3(c), FSDI exhibits pronounced fluctuations when reconstructing the near-constant segment shown, whereas the reconstructions of MA-TSD and our method are closer to the original series. These observations are consistent with our motivation to guide corruption allocation using normal reference structure, suggesting that structural consistency can provide an allocation criterion distinct from frequency position and spectral energy.

5.4 Ablation Studies

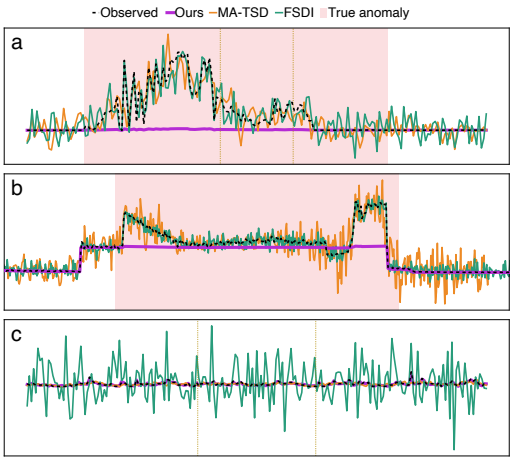

Figure 3: Reconstruction comparison of Ours, MA-TSD, and FSDI on the SMD dataset: shaded regions indicate the time intervals of labeled anomalous events.

Table 2: Module and evidence ablations. S: soft bands; E: evidence-guided corruption; B: budget conservation; CR: cross-window reproducibility; VR: cross-variable relationships; SP: spectral prototype; w/o: removal. Fixed bands are equal-width. The starred row denotes DDIM.

(a) Module ablations

S/E/B	MSL	SMAP	SMD	PSM	SWaT
× × ✓*	16.27	15.32	19.43	25.90	50.04
× ✓ ×	28.75	20.14	41.19	31.03	61.81
× ✓ ✓	29.38	21.03	43.18	31.42	63.72
✓ ✓ ×	29.53	25.17	44.53	39.18	66.00

(b) Evidence ablations

Variant	MSL	SMAP	SMD	PSM	SWaT
w/o CR	29.92	22.69	44.23	40.16	67.26
w/o VR	29.56	20.94	44.44	40.74	67.84
w/o SP	29.91	21.11	44.87	40.89	65.90
Full	**30.45**	**25.51**	**45.10**	**40.93**	**68.78**

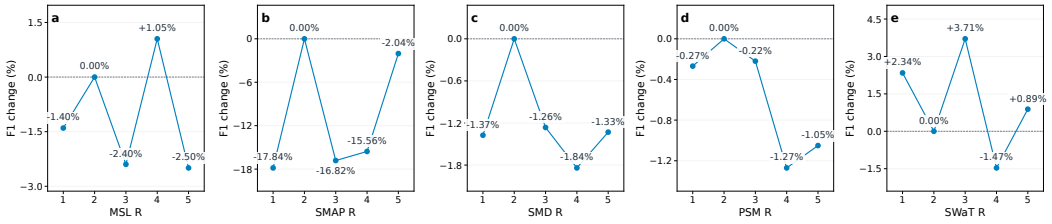

Figure 4: Relative changes in F1 as the number of clusters R in the reference bank varies.

6 Conclusion

This paper proposes a frequency-aware diffusion model that identifies stable operating states and uses deviations from these states to guide which frequency components are corrupted or protected. This design avoids relying solely on potentially restrictive priors based on absolute frequency location or spectral energy and supports the reconstruction of stable time-series patterns. Experiments demonstrate improved anomaly detection performance compared with a range of representative methods, including diffusion-based baselines.

AI USE STATEMENT

Generative AI tools were used to assist with the following tasks:

Translation and typesetting. AI was used to translate the authors' Chinese manuscript into English, including the translation of technical concepts and terminology and the contextual disambiguation of terms with multiple meanings. AI also assisted with PDF rendering, layout adjustments, and resolving typesetting issues.

Programming assistance. AI provided assistance with environment setup, code reproduction, and troubleshooting. The design and implementation of the core algorithms in this work did not rely on AI.

Literature retrieval. AI was used to help find references as an alternative to searching with Google Scholar. The papers related to the core research ideas of this work were not retrieved using AI.

The authors take full responsibility for the final content, technical claims, and experimental results of this paper.

REFERENCES

Ahmed Abdulaal, Zhuanghua Liu, and Tomer Lancewicki. Practical approach to asynchronous multivariate time series anomaly detection and localization. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 2485–2494, 2021. doi: 10.1145/3447548.3467174. URL https://doi.org/10.1145/3447548.3467174.

Julien Audibert, Pietro Michiardi, Frédéric Guyard, Sébastien Marti, and Maria A. Zuluaga. USAD: UnSupervised anomaly detection on multivariate time series. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 3395–3404, 2020. doi: 10.1145/3394486.3403392. URL https://doi.org/10.1145/3394486.3403392.

Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. LOF: Identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, pp. 93–104, 2000. doi: 10.1145/342009.335388. URL https://doi.org/10.1145/342009.335388.

Yuhang Chen, Chaoyun Zhang, Minghua Ma, Yudong Liu, Ruomeng Ding, Bowen Li, Shilin He, Saravan Rajmohan, Qingwei Lin, and Dongmei Zhang. ImDiffusion: Imputed diffusion models for multivariate time series anomaly detection. *Proceedings of the VLDB Endowment*, 17(3):359–372, 2023. doi: 10.14778/3632093.3632101. URL https://doi.org/10.14778/3632093.3632101.

Stéphane Crépey, Noureddine Lehdili, Nisrine Madhar, and Maud Thomas. Anomaly detection in financial time series by principal component analysis and neural networks. *Algorithms*, 15(10):385, 2022. doi: 10.3390/a15100385. URL https://doi.org/10.3390/a15100385.

Enyan Dai and Jie Chen. Graph-augmented normalizing flows for anomaly detection of multiple time series. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2202.07857.

Yuchen Fang, Jiandong Xie, Yan Zhao, Lu Chen, Yunjun Gao, and Kai Zheng. Temporal-frequency masked autoencoders for time series anomaly detection. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*, pp. 1228–1241, 2024. doi: 10.1109/ICDE60146.2024.00099. URL https://doi.org/10.1109/ICDE60146.2024.00099.

Markus Goldstein and Andreas Dengel. Histogram-based outlier score (HBOS): A fast unsupervised anomaly detection algorithm. In *KI-2012: Poster and Demo Track*, 2012. URL https://www.dfki.de/fileadmin/user_upload/import/6431_HBOS-poster.pdf.

David Hallac, Sagar Vare, Stephen Boyd, and Jure Leskovec. Toeplitz inverse covariance-based clustering of multivariate time series data. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 215–223, 2017. doi: 10.1145/3097983.3098060. URL https://arxiv.org/abs/1706.03161.

Sahand Hariri, Matias Carrasco Kind, and Robert J. Brunner. Extended isolation forest. *IEEE Transactions on Knowledge and Data Engineering*, 33(4):1479–1489, 2021. doi: 10.1109/TKDE.2019.2947676. URL https://doi.org/10.1109/TKDE.2019.2947676.

Zengyou He, Xiaofei Xu, and Shengchun Deng. Discovering cluster-based local outliers. *Pattern Recognition Letters*, 24(9–10):1641–1650, 2003. doi: 10.1016/S0167-8655(03)00003-5. URL https://doi.org/10.1016/S0167-8655(03)00003-5.

Niklas Heim and James E. Avery. Adaptive anomaly detection in chaotic time series with a spatially aware echo state network. *arXiv preprint arXiv:1909.01709*, 2019. URL https://arxiv.org/abs/1909.01709.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851, 2020. URL https://papers.nips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

Kyle Hundman, Valentino Constantinou, Christopher Laporte, Ian Colwell, and Tom Soderstrom. Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 387–395, 2018. doi: 10.1145/3219819.3219845. URL https://doi.org/10.1145/3219819.3219845.

Maximilian Karl, Maximilian Soelch, Justin Bayer, and Patrick van der Smagt. Deep variational bayes filters: Unsupervised learning of state space models from raw data. In *International Conference on Learning Representations*, 2017. URL https://arxiv.org/abs/1605.06432.

Siwon Kim, Kukjin Choi, Hyun-Soo Choi, Byunghan Lee, and Sungroh Yoon. Towards a rigorous evaluation of time-series anomaly detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 7194–7201, 2022. doi: 10.1609/aaai.v36i7.20680. URL https://ojs.aaai.org/index.php/AAAI/article/view/20680.

Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. URL https://arxiv.org/abs/1312.6114.

Thomas Kipf, Ethan Fetaya, Kuan-Chieh Wang, Max Welling, and Richard Zemel. Neural relational inference for interacting systems. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2688–2697. PMLR, 2018. URL https://proceedings.mlr.press/v80/kipf18a.html.

Dan Li, Dacheng Chen, Baihong Jin, Lei Shi, Jonathan Goh, and See-Kiong Ng. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. In *Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series*, pp. 703–716. Springer, 2019. doi: 10.1007/978-3-030-30490-4_56. URL https://doi.org/10.1007/978-3-030-30490-4_56.

Long Li, Wencheng Zhang, Shi Yuan, Hongle Guo, and Wanghu Chen. PhysDiff: A physically-guided diffusion model for multivariate time series anomaly detection. In *Advances in Neural Information Processing Systems*, volume 38, pp. 105720–105750, 2025. doi: 10.52202/085713-3531. URL https://papers.nips.cc/paper_files/paper/2025/hash/980ea04d23d1f6908964eba2a74afe45-Abstract-Conference.html.

Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422, 2008. doi: 10.1109/ICDM.2008.17. URL https://doi.org/10.1109/ICDM.2008.17.

Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. iTransformer: Inverted transformers are effective for time series forecasting. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/2ea18fdc667e0ef2ad82b2b4d65147ad-Abstract-Conference.html.

Donghao Luo and Xue Wang. ModernTCN: A modern pure convolution structure for general time series analysis. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/86b1437c1e4c3b3c4debff98234a67e7-Abstract-Conference.html.

J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1. University of California Press, 1967. URL https://projecteuclid.org/euclid.bsmsp/1200512992.

Pankaj Malhotra, Anusha Ramakrishnan, Gaurangi Anand, Lovekesh Vig, Puneet Agarwal, and Gautam Shroff. LSTM-based encoder-decoder for multi-sensor anomaly detection. *arXiv preprint arXiv:1607.00148*, 2016. URL https://arxiv.org/abs/1607.00148.

Aditya P. Mathur and Nils Ole Tippenhauer. SWaT: A water treatment testbed for research and training on ICS security. In *2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater)*, pp. 31–36, 2016. doi: 10.1109/CySWater.2016.7469060. URL https://doi.org/10.1109/CySWater.2016.7469060.

Yisroel Mirsky, Tomer Doitshman, Yuval Elovici, and Asaf Shabtai. Kitsune: An ensemble of autoencoders for online network intrusion detection. In *Proceedings of the Network and Distributed System Security Symposium*, 2018. doi: 10.14722/ndss.2018.23204. URL https://doi.org/10.14722/ndss.2018.23204.

Youngeun Nam, Susik Yoon, Yooju Shin, Minyoung Bae, Hwanjun Song, Jae-Gil Lee, and Byung Suk Lee. Breaking the time-frequency granularity discrepancy in time-series anomaly detection. In *Proceedings of the ACM Web Conference 2024*, pp. 4204–4215, 2024. doi: 10.1145/3589334.3645556. URL https://doi.org/10.1145/3589334.3645556.

Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2211.14730.

Kohei Obata, Zheng Chen, Yasuko Matsubara, Lingwei Zhu, and Yasushi Sakurai. AnomalyFilter: Selective denoising diffusion model for time series anomaly detection. In *Advances in Knowledge Discovery and Data Mining*, Lecture Notes in Computer Science, pp. 429–441. Springer, 2026. doi: 10.1007/978-981-92-1300-9_34. URL https://doi.org/10.1007/978-981-92-1300-9_34.

Jinju Park and Seokho Kang. PaAno: Patch-based representation learning for time-series anomaly detection. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/5991785875cf5eee9c29e2ed7d26d1c0-Abstract-Conference.html.

Tomáš Pevný. Loda: Lightweight on-line detector of anomalies. *Machine Learning*, 102(2):275–304, 2016. doi: 10.1007/s10994-015-5521-0. URL https://doi.org/10.1007/s10994-015-5521-0.

Xiangfei Qiu, Zhe Li, Wanghui Qiu, Shiyan Hu, Lekui Zhou, Xingjian Wu, Zhengyu Li, Chenjuan Guo, Aoying Zhou, Zhenli Sheng, Jilin Hu, Christian S. Jensen, and Bin Yang. TAB: Unified benchmarking of time series anomaly detection methods. *Proceedings of the VLDB Endowment*, 18(9):2775–2789, 2025a. doi: 10.14778/3746405.3746407. URL https://doi.org/10.14778/3746405.3746407.

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, 2025b. doi: 10.1145/3690624.3709325. URL https://doi.org/10.1145/3690624.3709325.

Sridhar Ramaswamy, Rajeev Rastogi, and Kyuseok Shim. Efficient algorithms for mining outliers from large data sets. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, pp. 427–438, 2000. doi: 10.1145/342009.335437. URL https://doi.org/10.1145/342009.335437.

Kashif Rasul, Calvin Seward, Ingmar Schuster, and Roland Vollgraf. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8857–8868. PMLR, 2021. URL https://proceedings.mlr.press/v139/rasul21a.html.

Mayu Sakurada and Takehisa Yairi. Anomaly detection using autoencoders with nonlinear dimensionality reduction. In *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, pp. 4–11, 2014. doi: 10.1145/2689746.2689747. URL https://doi.org/10.1145/2689746.2689747.

Bernhard Schölkopf, Robert C. Williamson, Alex Smola, John Shawe-Taylor, and John Platt. Support vector method for novelty detection. In *Advances in Neural Information Processing Systems*, volume 12, 1999. URL https://papers.nips.cc/paper/1723-support-vector-method-for-novelty-detection.

Wangmeng Shen, Hongfan Gao, Qingsong Zhong, Dingli Xu, and Jilin Hu. FSDI: Frequency-shaped diffusion for time-series imputation. In *Proceedings of the ACM Web Conference 2026*, pp. 5310–5321, 2026. doi: 10.1145/3774904.3792330. URL https://doi.org/10.1145/3774904.3792330.

Mei-Ling Shyu, Shu-Ching Chen, Kanoksri Sarinnapakorn, and LiWu Chang. A novel anomaly detection scheme based on principal component classifier. In *Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop*, pp. 172–179, 2003.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=St1giarCHLP.

Ya Su, Youjian Zhao, Chenhao Niu, Rong Liu, Wei Sun, and Dan Pei. Robust anomaly detection for multivariate time series through stochastic recurrent neural network. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2828–2837, 2019. doi: 10.1145/3292500.3330672. URL https://doi.org/10.1145/3292500.3330672.

Yusuke Tashiro, Jiaming Song, Yang Song, and Stefano Ermon. CSDI: Conditional score-based diffusion models for probabilistic time series imputation. In *Advances in Neural Information Processing Systems*, volume 34, pp. 24804–24816, 2021. URL https://papers.nips.cc/paper_files/paper/2021/hash/cfe8504bda37b575c70ee1a8276f3486-Abstract.html.

Shreshth Tuli, Giuliano Casale, and Nicholas R. Jennings. TranAD: Deep transformer networks for anomaly detection in multivariate time series data. *Proceedings of the VLDB Endowment*, 15(6): 1201–1214, 2022. doi: 10.14778/3514061.3514067. URL https://doi.org/10.14778/3514061.3514067.

Chengsen Wang, Zirui Zhuang, Qi Qi, Jingyu Wang, Xingyu Wang, Haifeng Sun, and Jianxin Liao. Drift doesn't matter: Dynamic decomposition with diffusion reconstruction for unstable multivariate time series anomaly detection. In *Advances in Neural Information Processing Systems*, volume 36, pp. 10758–10774, 2023. doi: 10.52202/075280-0473. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/22f5d8e689d2a011cd8ead552ed59052-Abstract-Conference.html.

Chenxi Wang, Linxiao Yang, Zhixian Wang, Liang Sun, and Yi Wang. A non-isotropic time series diffusion model with moving average transitions. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 65144–65166. PMLR, 2025. URL https://proceedings.mlr.press/v267/wang25dy.html.

Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Autoformer: Decomposition transformers with Auto-Correlation for long-term series forecasting. In *Advances in Neural Information Processing Systems*, volume 34, pp. 22419–22430, 2021. URL https://papers.nips.cc/paper_files/paper/2021/hash/bcc0d400288793e8bdcd7c19a8ac0c2b-Abstract.html.

Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. TimesNet: Temporal 2D-variation modeling for general time series analysis. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2210.02186.

Xingjian Wu, Xiangfei Qiu, Zhengyu Li, Yihang Wang, Jilin Hu, Chenjuan Guo, Hui Xiong, and Bin Yang. CATCH: Channel-aware multivariate time series anomaly detection via frequency patching. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/2b25c39788e5cf11d3541de433ebf4c0-Abstract-Conference.html.

Chunjing Xiao, Zehua Gou, Wenxin Tai, Kunpeng Zhang, and Fan Zhou. Imputation-based time-series anomaly detection with conditional weight-incremental diffusion models. In *Proceedings*

of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 2742–2751, 2023. doi: 10.1145/3580305.3599391. URL https://doi.org/10.1145/3580305.3599391.

Jiehui Xu, Haixu Wu, Jianmin Wang, and Mingsheng Long. Anomaly transformer: Time series anomaly detection with association discrepancy. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2110.02642.

Yiyuan Yang, Chaoli Zhang, Tian Zhou, Qingsong Wen, and Liang Sun. DCdetector: Dual attention contrastive representation learning for time series anomaly detection. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3033–3045, 2023. doi: 10.1145/3580305.3599295. URL https://doi.org/10.1145/3580305.3599295.

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

Chaoli Zhang, Tian Zhou, Qingsong Wen, and Liang Sun. TFAD: A decomposition time series anomaly detection architecture with time-frequency analysis. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management*, pp. 2497–2507, 2022. doi: 10.1145/3511808.3557470. URL https://doi.org/10.1145/3511808.3557470.

Fan Zhang, Sinchee Chin, Jing-Hao Xue, and Wenming Yang. ICDiffAD: Implicit conditioning diffusion model for time series anomaly detection. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/59a9cc95f046e9125d8816ef971873e7-Abstract-Conference.html.

Guojin Zhong, Pan Wang, Jin Yuan, Zhiyong Li, and Long Chen. Multi-resolution decomposable diffusion model for non-stationary time series anomaly detection. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/9f7f063144103bf6debb09a3f15e00fb-Abstract-Conference.html.

Jiaqi Zhu, Shaofeng Cai, Fang Deng, Beng Chin Ooi, and Wenqiao Zhang. METER: A dynamic concept adaptation framework for online anomaly detection. *Proceedings of the VLDB Endowment*, 17(4):794–807, 2023. doi: 10.14778/3636218.3636233. URL https://www.vldb.org/pvldb/vol17/p794-zhu.pdf.

Zhihao Zhuang, Yingying Zhang, Kai Zhao, Chenjuan Guo, Bin Yang, Qingsong Wen, and Lunting Fan. Noise matters: Cross contrastive learning for Flink anomaly detection. *Proceedings of the VLDB Endowment*, 18(4):1159–1168, 2024. doi: 10.14778/3717755.3717773. URL https://www.vldb.org/pvldb/vol18/p1159-zhuang.pdf.

Bo Zong, Qi Song, Martin Renqiang Min, Wei Cheng, Cristian Lumezanu, Daeki Cho, and Haifeng Chen. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=BJJLHbb0-.

A Definitions of Statistical Criteria

We use overlapping windows of length 128 with a stride of 32 to compute a frequency score and a spectral-energy score for each variable. Reference distributions are constructed from the training data, with their 95th percentiles serving as the thresholds for the corresponding criteria. Test data are used exclusively for evaluation and do not contribute to threshold selection. The two scores quantify departures in frequency distribution and spectral energy from normal patterns while accounting for normal variability; a test window is assigned a high score when it exceeds the corresponding threshold. For each window, if the corresponding score of any variable exceeds the threshold, a response is recorded for the current interval of 32 consecutive timestamps. We then define the association rate between each criterion and ground-truth anomalies as the proportion of ground-truth anomalous points (`label = 1`) covered by that criterion.

Figure A illustrates the computation and two representative cases. Red shading marks ground-truth anomalous intervals, and dashed lines indicate the 95th-percentile thresholds determined from the training data. Orange and blue curves represent the spectral-energy and frequency scores, respectively. In the left panel, the spectral-energy score exceeds its threshold during the anomaly, whereas the frequency distribution shows no substantial deviation. In the right panel, neither score exceeds its corresponding threshold despite the presence of labeled anomalies. These examples show that frequency- and spectral-energy-based criteria capture some anomalies but miss others.

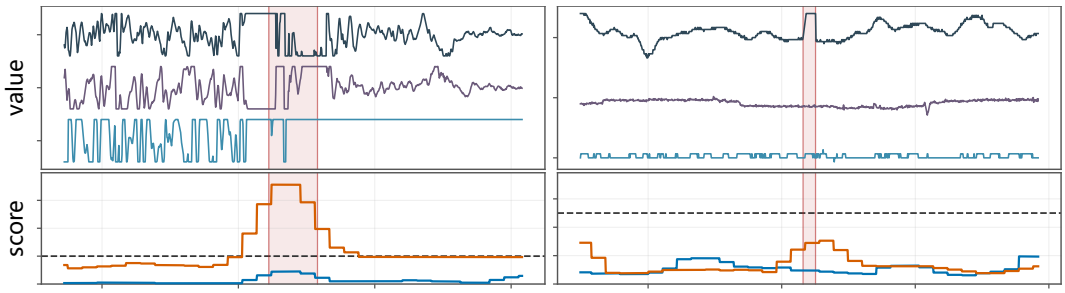

Figure A: Frequency- and spectral-energy-based criteria and two representative anomaly cases. Red shading marks labeled anomalies; dashed lines denote the 95th-percentile thresholds estimated from training data. Orange and blue curves show the spectral-energy and frequency scores, respectively.

B PRINCIPLES OF THE FREQUENCY-BAND MASK GENERATOR

B.1 SOFT-MASK VISUALIZATION

To illustrate how soft masks are generated and applied, we selected a time series segment containing both normal observations and an anomalous event (`label=1`). Figure B.1(a) shows the selected segment, with the shaded region marking the anomalous interval. Figure B.1(b) presents the spectra before, during, and after the event. Figure B.1(c) shows the initial Gaussian soft frequency-band masks generated for the anomalous interval. In this example, $M = 4$. The preset band centers are determined by uniform partitioning based on absolute frequency positions. The frequency-band generator then adaptively adjusts the center and scale of each band. Figure B.1(d) presents the spectra obtained by applying the soft bands, together with their respective spectral energy proportions.

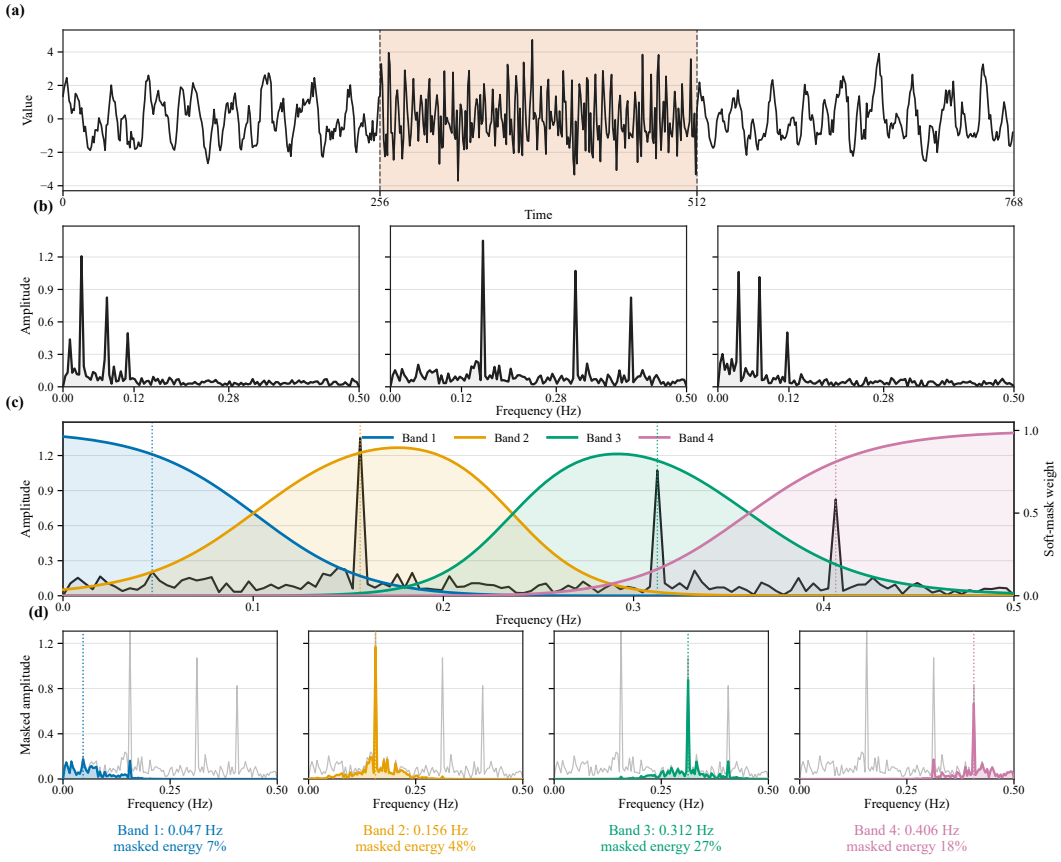

Figure B.1: Initialization of the frequency-band mask generator on the PSM dataset.

B.2 MASK-GENERATOR LEARNING DYNAMICS

To illustrate the learning process of the frequency-band generator, we selected the time series shown in Figure B.2(a). Figure B.2(b) presents the amplitude-normalized spectrum and the soft bands initialized from a fixed frequency partition. Figure B.2(c) shows how the band centers shift and their scales change during learning. The lower- and higher-frequency regions are ultimately both included in Band2, giving the band a two-peaked structure. This observation indicates coupling between different frequency positions and points to a potential limitation of applying corruption solely on the basis of absolute frequency position. Figure B.2(d) shows the L_1 reconstruction loss during training of the frequency-band generator. The loss decreases overall and gradually converges as the generator is updated.

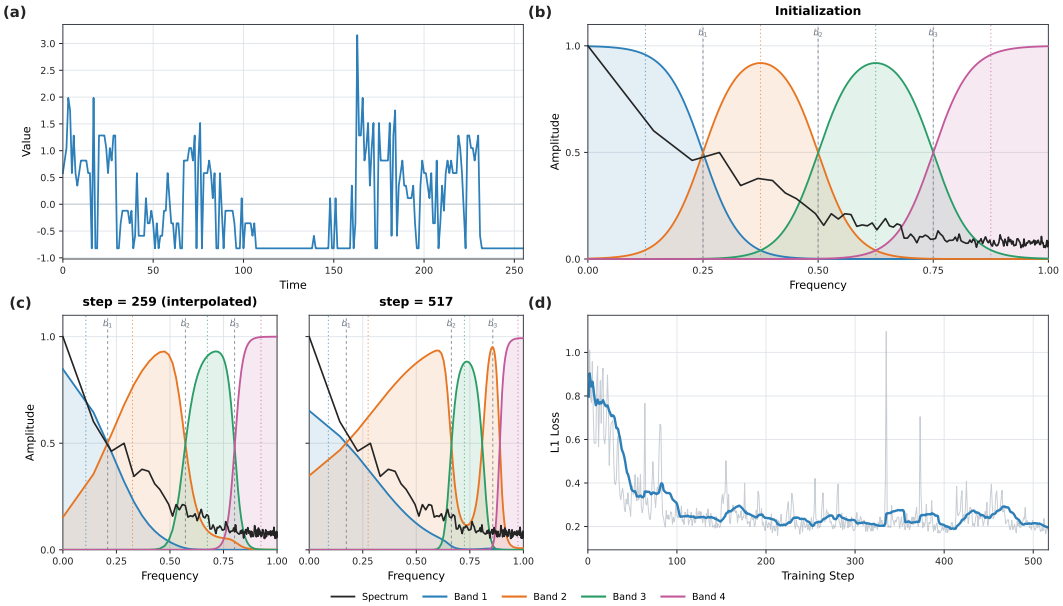

Figure B.2: Update process of the frequency-band mask generator on the PSM dataset.

C Reference Bank and Corruption Budget

C.1 Spectral Reference Features

Applying rFFT along the time dimension of the input window yields the complex spectrum U. We use the mean spectral power across variables to describe the intensity of frequency components and construct band statistics using the soft masks $W_{m,k}$:

$$P_k = \frac{1}{C}\sum_{c=1}^{C}|U_{k,c}|^2, \qquad D_m = \sum_k W_{m,k}, \qquad \omega_{m,k} = \frac{W_{m,k}P_k}{\sum_j W_{m,j}P_j + \varepsilon}. \tag{17}$$

Here, C is the number of variables; k, j range over the retained frequency bins; D_m is the total band weight; $\omega_{m,k}$ is the normalized power weight within a band; and $\varepsilon > 0$ is a numerical stabilizer. Mean band energy, spectral entropy, and spectral flatness are defined as

$$\begin{aligned} E_m &= \frac{\sum_k W_{m,k}P_k}{D_m + \varepsilon}, \\ H_m &= -\frac{\sum_k \omega_{m,k}\log(\omega_{m,k}+\varepsilon)}{\log D_m}, \\ F_m &= \frac{\exp\left[D_m^{-1}\sum_k W_{m,k}\log(P_k+\varepsilon)\right]}{E_m + \varepsilon}. \end{aligned} \tag{18}$$

These quantities describe overall band intensity, dispersion of the power distribution, and flatness, respectively, and form the local spectral features:

$$\ell_m = [\log(E_m + \varepsilon),\ H_m,\ F_m] \in \mathbb{R}^3. \tag{19}$$

For a pair of variables c-d, we define their normalized cross-spectral relationship within band m as

$$G_m^{c,d} = \frac{\mathrm{Re}\Big(\sum_{k=0}^{K-1} W_{m,k}\overline{U}_{k,c}U_{k,d}\Big)}{\sqrt{\sum_{k=0}^{K-1} W_{m,k}|U_{k,c}|^2}\sqrt{\sum_{k=0}^{K-1} W_{m,k}|U_{k,d}|^2} + \varepsilon}. \tag{20}$$

Here, $\overline{U}_{k,c}$ denotes complex conjugation, and $\mathrm{Re}(\cdot)$ extracts the real part. The relationship matrix is symmetric, so $G_m^{c,d} = G_m^{d,c}$. Its diagonal entries equal 1 when the numerical stabilizer is omitted and the corresponding band energy is nonzero. We therefore retain only the $\binom{C}{2}$ unordered variable pairs in the strictly upper-triangular part:

$$q_m = \left[q^1, q^2, \ldots, q^{\binom{C}{2}}\right]. \tag{21}$$

These pairwise relationships are compressed into an eight-dimensional statistical representation:

$$\mathbf{r}_m = [\mu, \sigma, \mu_{\mathrm{abs}}, q_{\max}, q_{\min}, \mu_+, \mu_-, \mu_w] \in \mathbb{R}^8. \tag{22}$$

The local spectral features ℓ, cross-window reproducibility representation s, and cross-variable relationship representation $\mathbf{r}$ are standardized separately, weighted, and concatenated into the joint window-level representation ψ:

$$\psi = \left[\sqrt{\lambda_{\mathrm{proto}}}\,\mathrm{Std}(\ell); \sqrt{\lambda_{\mathrm{rep}}}\,\mathrm{Std}(s); \sqrt{\lambda_{\mathrm{rel}}}\,\mathrm{Std}(\mathbf{r})\right]. \tag{23}$$

K-means clustering is applied to the joint training-window representations $\{\psi_i\}$ to construct a reference bank containing R normal operating states.

C.2 Support and Corruption Budget

For the reference state $r^\star$ selected in the main text[(12)], monotonically decreasing functions convert the three structural deviations into support scores:

$$\begin{aligned} \mathcal{Q}_{\mathrm{rep}}(D) = \mathcal{Q}_{\mathrm{rel}}(D) &= \exp\left(-\frac{D}{2}\right), \\ \mathcal{Q}_{\mathrm{proto}}(D) &= \sigma\left(\frac{\tau_p - D}{T_p}\right). \end{aligned} \tag{24}$$

Here, $D \geq 0$, σ is the sigmoid function, and τ_p and $T_p > 0$ control the threshold and temperature of prototype support, respectively. The three support scores are combined by weighted fusion into band-level structural evidence:

$$e_m = \sum_{j\in\{\mathrm{proto,rep,rel}\}} \lambda_j \mathcal{Q}_j\Big(D^j_{r^\star,m}\Big), \qquad \lambda_j \geq 0, \quad \sum_j \lambda_j = 1. \tag{25}$$

When all three evidence components are enabled, each weight is $1/3$. A larger e_m indicates closer agreement between the band and the selected reference state.

The band budget weights are defined by averaging the soft masks over the frequency dimension:

$$\pi_m = \frac{1}{K}\sum_{k=0}^{K-1} W_{m,k}, \qquad \sum_m \pi_m = 1. \tag{26}$$

This requires $\sum_m W_{m,k} = 1$ at each frequency position. The weight π_m describes the band's soft coverage proportion over frequency bins and does not depend on the power of the observed signal.

The budget mapping allocates corruption according to the difference between each band's support and the weighted average:

$$\bar{e} = \sum_m \pi_m e_m, \qquad h_m = \tanh(\bar{e} - e_m), \qquad \bar{h} = \sum_m \pi_m h_m. \tag{27}$$

$$v_m = 1 + \gamma(h_m - \bar{h}), \qquad v_k^{\text{freq}} = \sum_m W_{m,k} v_m, \qquad \gamma \in [0, 1]. \tag{28}$$

Here, γ controls the magnitude of corruption redistribution. Within a window, higher support leads to a smaller allocated corruption variance coefficient. Subtracting $\bar{h}$ ensures

$$\sum_m \pi_m v_m = \frac{1}{K} \sum_{k=0}^{K-1} v_k^{\text{freq}} = 1,$$

so the mapping changes the allocation of corruption across frequencies while preserving the mean corruption variance coefficient over the one-sided frequency bins.

C.3 Reference-Bank Visualization

The reference bank and frequency-band generator depend on each other during initialization. Evidence from the reference bank guides diffusion corruption, and the resulting reconstruction error provides the training signal for the band generator. However, the statistics and clustering used to construct the bank also depend on the generator's band partition. To resolve this initialization dependency, we introduce a pretraining stage before formal training. A temporary reference and the band generator are updated alternately at the batch level, after which the pretrained generator is used to construct and freeze the formal reference bank. This design is inspired by TranAD (Tuli et al., 2022), which uses deviations from an initial reconstruction to guide a second reconstruction. We adopt the underlying idea of using preliminary feedback to support subsequent modeling.

For the first training batch, the initialized band generator partitions each window into soft frequency bands, with band centers and scales close to their preset values. We then extract the band-level features required for reference construction and compute their means and variances across windows to initialize a temporary reference. This reference evaluates deviations within the current batch itself. These deviations are combined with local spectral structure to form corruption evidence. Evidence-guided structured corruption is then applied, and the reconstruction error is backpropagated to update both the band generator and the denoiser.

For each subsequent batch, we first extract band-level features using the generator updated on the preceding batch. The current batch statistics are then incorporated into the temporary reference through exponential smoothing. The updated reference includes information from the current batch and is used to evaluate that batch, construct corruption evidence, and complete the model update. Each batch therefore follows the sequence of band partitioning, feature computation, temporary-reference updating, evidence-guided corruption, and model updating.

Let $\widehat{\boldsymbol{\mu}}_b, \widehat{\boldsymbol{v}}_b \in \mathbb{R}^{M \times d}$ denote the feature means and variances computed across windows in training batch b, where d is the feature dimension. The corresponding temporary-reference statistics are denoted by $\boldsymbol{\mu}_b$ and $\boldsymbol{v}_b$. The first batch initializes these statistics:

$$\boldsymbol{\mu}_1 = \widehat{\boldsymbol{\mu}}_1, \qquad \boldsymbol{v}_1 = \widehat{\boldsymbol{v}}_1. \tag{29}$$

For subsequent batches, the updates are

$$\boldsymbol{\mu}_b = (1 - \rho)\boldsymbol{\mu}_{b-1} + \rho\widehat{\boldsymbol{\mu}}_b, \tag{30}$$

$$\boldsymbol{v}_b = (1 - \rho)\boldsymbol{v}_{b-1} + \rho\widehat{\boldsymbol{v}}_b + \rho(1 - \rho)\left(\widehat{\boldsymbol{\mu}}_b - \boldsymbol{\mu}_{b-1}\right)^{\odot 2}. \tag{31}$$

The exponential smoothing coefficient ρ is set to 0.02. After pretraining, only the band-generator parameters are retained and frozen. The pretrained denoiser and temporary reference are not reused.

The frozen band generator then processes all training windows, producing a soft frequency-band partition for each window. We compute local spectral features, cross-window reproducibility representations, and cross-variable relationship features. These band-level features are compressed and concatenated into a joint vector for each window. K-means clustering then constructs the formal reference bank with R normal operating regimes. Figure C.3 shows a PCA projection of the joint representations for $R = 3$.

C.4 Operating-Regime Visualization

Figure C.5 presents the operating-regime assignments of 25 variables in the PSM dataset over the displayed interval for $R = 3$. Using one variable as an example, it also compares time windows from different clusters and their corresponding rFFT amplitude spectra. Differences in temporal patterns and spectral characteristics across intervals suggest that the same variable may exhibit different stable operating regimes over time.

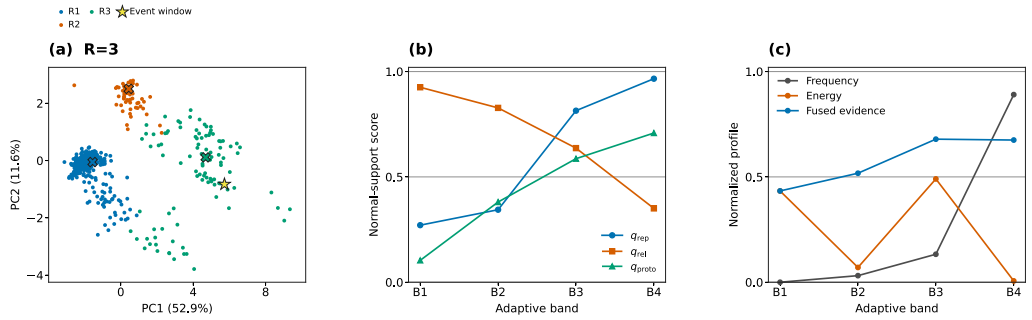

Figure C.3: (a) Two-dimensional PCA projection of the joint features of a subset of normal training windows. (b) Local prototype support, cross-window reproducibility, and cross-variable relationship consistency across the four frequency bands of the current window. Band B1 exhibits high cross-variable relationship consistency but low cross-window reproducibility and local prototype support. In the higher-frequency bands B3 and B4, cross-variable relationship consistency decreases, whereas cross-window reproducibility and local prototype support increase. (c) Comparison of frequency position, spectral energy proportion, and fused evidence. Band B4 receives the strongest support from the fused evidence, whereas B1 receives the weakest support despite being a low-frequency band with high spectral energy.

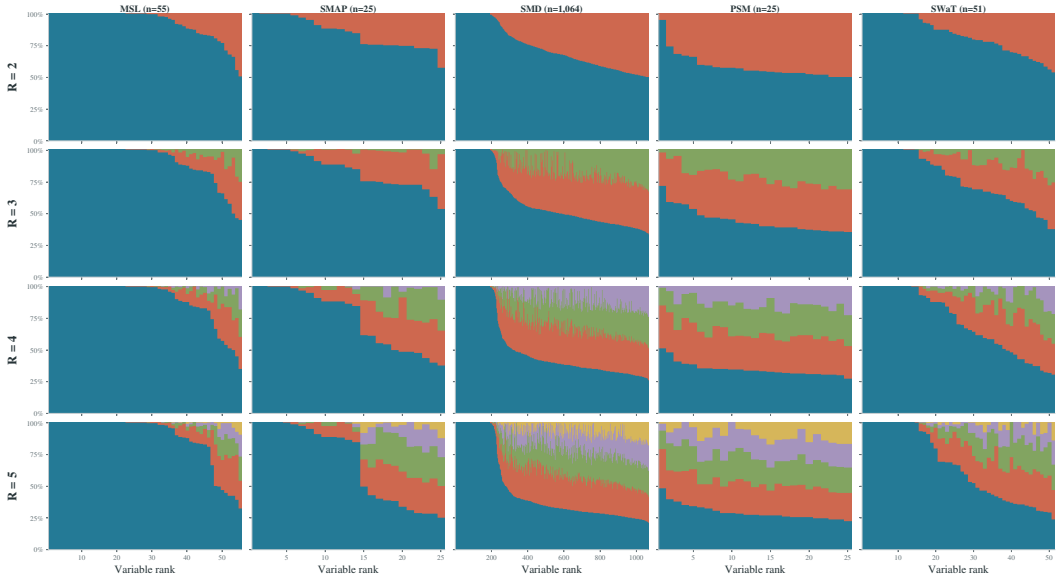

Figure C.4: Per-variable clustering results and state window proportions across five datasets with preset numbers of states $R = 2, 3, 4, 5$. Each column corresponds to one dataset, and each row corresponds to one value of R. Each vertical stacked bar shows the proportions of windows of a variable assigned to the respective cluster states, with variables sorted in descending order of their largest window proportion. This is an independent per-variable diagnostic analysis and does not participate in the main model's reference-state selection for multivariate windows; states with the same index across different variables do not represent the same operating state.



Figure C.5: Variable-clustering timeline, example windows, and rFFT spectra for PSM with $R = 3$.

D Evaluation Metrics Based on TAB

D.1 Baseline Models

Traditional Methods. IForest (Liu et al., 2008), LOF (Breunig et al., 2000), OCSVM (Schölkopf et al., 1999), CBLOF (He et al., 2003), EIF (Hariri et al., 2021), HBOS (Goldstein & Dengel, 2012), KMeans (MacQueen, 1967), KNN (Ramaswamy et al., 2000), LODA (Pevný, 2016), and PCA (Shyu et al., 2003).

Time Series Models. AE (Sakurada & Yairi, 2014), VAE (Kingma & Welling, 2013), OmniAnomaly (Su et al., 2019), USAD (Audibert et al., 2020), GANF (Dai & Chen, 2022), DAGMM (Zong et al., 2018), LSTMED (Malhotra et al., 2016), Torsk (Heim & Avery, 2019), TranAD (Tuli et al., 2022), AnomTrans (Xu et al., 2022), DCdetector (Yang et al., 2023), ContraAD (Zhuang et al., 2024), DLinear and NLinear (Zeng et al., 2023), DUET (Qiu et al., 2025b), iTransformer (Liu et al., 2024), PatchTST (Nie et al., 2023), TimesNet (Wu et al., 2023), and ModernTCN (Luo & Wang, 2024).

Time-Frequency Models. TFAD (Zhang et al., 2022), DualTF (Nam et al., 2024), TFMAE (Fang et al., 2024), and CATCH (Wu et al., 2025).

Diffusion Models. DiffAD (Xiao et al., 2023), ImDiffusion (Chen et al., 2023), D3R (Wang et al., 2023), AnomFilter (Obata et al., 2026), MATSD (Wang et al., 2025), ICDiffAD (Zhang et al., 2026), FSDI (Shen et al., 2026), and PhysDiff (Li et al., 2025).

D.2 Univariate Time Series Results

Table D.2.1: Average Metrics (%) on five univariate time series: UCR-DISTORTEDECG2, NAB-KnownCause-5, Yahoo-A2synthetic-30, E-1, and Periodic-7491.

Method	AUPRC	AUROC	Range-AUPRC	Range-AUROC	VUS-PR	VUS-ROC
TranAD	42.04	64.85	36.96	73.44	34.67	70.66
DCdetector	5.61	54.73	9.87	57.74	9.77	57.49
Ours	**42.61**	**71.98**	**38.04**	**75.32**	**35.79**	**72.33**

D.3 Cumulative-Score Ranking

Drawing on the evaluation protocol of TAB (Qiu et al., 2025a), we adopt a cumulative point-based ranking scheme. For each metric on each dataset, the first-, second-, and third-ranked models receive 3, 2, and 1 points, respectively. The calculation covers the 13 appendix metrics; tied ranks are defined as one plus the number of methods with strictly higher values. We sum these points across datasets and metrics and rank the models by their cumulative scores, yielding the leaderboard below.

Rank	Model	Gold	Silver	Bronze	Total
1	**Ours**	**21**	**4**	**2**	**73**
2	KNN	6	5	8	36
3	AE	2	5	10	26
4	EIF	4	4	2	22
4	LSTMED	6	1	2	22
6	VAE	4	2	3	19
7	LOF	3	3	3	18
8	KMeans	4	1	3	17
8	PhysDiff	3	3	2	17
10	GANF	1	5	1	14
11	CATCH	0	5	3	13
12	FSDI	4	0	0	12
13	TimesNet	1	3	2	11
14	IForest	1	3	1	10
15	TranAD	1	3	0	9
16	D3R	0	4	0	8
16	DAGMM	0	4	0	8
18	CBLOF	1	1	1	6
18	DiffAD	0	2	2	6
18	HBOS	1	1	1	6
18	iTransformer	2	0	0	6
22	DualTF	1	1	0	5
22	ImDiffusion	1	1	0	5
22	TFMAE	0	0	5	5
25	MATSD	1	0	1	4
25	OmniAnomaly	0	2	0	4
27	AnomFilter	0	1	1	3
27	LODA	0	1	1	3
27	ModernTCN	0	0	3	3
27	NLinear	1	0	0	3
27	TFAD	0	1	1	3
27	Torsk	1	0	0	3
33	OCSVM	0	0	2	2
34	PatchTST	0	0	1	1
34	USAD	0	0	1	1
36	AnomTrans	0	0	0	0
36	ContraAD	0	0	0	0
36	DCdetector	0	0	0	0
36	DLinear	0	0	0	0
36	DUET	0	0	0	0
36	ICDiffAD	0	0	0	0
36	PCA	0	0	0	0

D.4 Additional Evaluation Metrics

Point-wise F1 is reported in Table 1; 13 additional metrics are listed below.

MSL

Table D.4.1: Additional TAB evaluation metrics (%) on MSL.

Model	ACC	AFF-F1	AFF-P	AFF-R	R-F1	R-P	R-R	A-P	A-R	R-A-P	R-A-R	V-PR	V-ROC
Traditional methods													
IForest	89.9	57.9	50.5	69.6	13.7	17.9	11.2	11.4	52.5	17.5	57.3	17.2	57.4
LOF	75.4	69.8	54.9	97.9	16.3	13.8	20.6	11.9	55.6	18.3	62.8	18.2	61.8
OCSVM	82.3	63.9	49.5	90.8	13.2	9.9	20.0	15.4	52.1	18.4	59.2	18.6	59.3
CBLOF	80.9	67.4	51.2	98.6	13.5	9.7	22.7	18.1	62.0	22.8	69.3	22.4	68.1
EIF	86.4	69.4	57.1	88.5	16.2	14.7	18.0	18.0	62.6	21.5	69.5	21.3	68.8
HBOS	82.3	67.8	51.7	98.5	15.0	12.0	20.0	13.3	57.9	18.9	63.9	19.0	63.2
KMeans	70.1	58.0	48.9	71.2	22.7	15.0	46.4	17.6	60.6	22.1	61.6	21.5	61.0
KNN	79.2	69.1	56.2	92.6	17.7	15.5	20.5	20.1	62.1	24.6	64.5	24.3	64.8
LODA	88.8	61.4	49.3	81.3	13.9	17.2	11.7	14.0	49.2	17.8	55.5	17.8	54.2
PCA	80.8	67.8	53.7	92.0	15.4	11.4	23.5	15.7	55.0	20.3	63.5	19.9	62.4
Time-series models													
AE	89.0	61.2	51.5	75.1	12.8	25.4	8.6	13.8	51.6	19.2	60.6	19.2	60.5
VAE	88.5	64.4	53.7	80.4	14.7	21.9	11.0	14.0	52.7	19.0	61.1	18.9	60.3
OmniAnomaly	70.8	67.8	51.6	98.9	14.1	10.2	22.8	13.2	52.3	18.8	61.3	18.6	60.5
USAD	71.7	67.2	53.2	91.3	20.1	15.6	28.3	17.3	58.1	20.8	63.0	20.7	62.3
GANF	70.1	62.3	53.3	74.9	17.2	11.2	36.9	17.7	60.1	20.4	62.1	20.4	61.4
DAGMM	88.2	72.0	58.0	94.9	18.8	25.5	14.9	13.7	57.0	20.2	63.5	20.0	62.6
LSTMED	88.7	62.1	52.1	78.8	14.1	22.0	10.4	13.7	53.0	19.3	61.9	19.0	61.1
Torsk	69.9	54.0	44.5	68.7	16.7	11.8	29.2	10.7	50.9	15.7	47.1	15.7	46.7
TranAD	86.8	72.8	58.5	96.3	16.0	14.2	18.7	12.7	47.6	17.9	56.6	17.7	55.2
AnomTrans	86.6	68.6	53.0	97.1	14.0	11.0	19.2	10.7	50.7	16.1	51.8	15.9	51.3
DCdetector	86.8	67.3	51.9	95.6	13.4	11.2	17.4	10.5	49.5	15.4	53.2	15.6	52.8
ContraAD	88.8	67.0	51.5	95.7	13.3	11.4	16.1	11.0	54.5	16.5	64.3	16.6	64.0
DLinear	86.0	71.6	57.5	97.9	15.5	12.7	20.0	13.9	60.1	20.9	67.8	20.5	67.0
NLinear	85.7	72.6	58.0	97.4	15.3	12.8	19.0	14.0	60.1	21.9	68.4	21.6	67.9
DUET	85.6	71.7	56.8	98.2	15.1	12.1	20.0	14.6	62.6	22.1	69.8	21.8	69.1
iTransformer	86.0	70.0	56.5	94.8	14.8	12.6	18.0	14.2	58.6	21.9	67.2	21.8	66.4
PatchTST	86.1	71.2	56.6	96.1	14.5	11.9	18.9	14.8	62.6	22.7	70.4	22.2	70.1
TimesNet	80.6	72.5	57.2	98.9	16.1	11.6	26.6	14.7	61.5	22.9	70.5	22.6	69.9
ModernTCN	85.4	72.7	58.1	98.4	15.2	12.5	19.3	14.4	62.0	22.5	69.8	22.3	68.6
Time-frequency models													
TFAD	89.1	66.5	51.5	93.6	11.9	10.8	13.2	10.5	50.0	17.1	51.5	16.7	51.5
DualTF	88.9	25.8	54.7	17.0	1.2	63.2	0.6	11.4	49.8	18.1	60.4	17.1	58.9
TFMAE	68.3	66.6	50.3	98.7	14.8	10.7	24.2	9.6	45.2	13.6	50.4	13.6	50.1
CATCH	82.0	69.3	55.2	97.2	16.4	13.5	20.9	15.6	63.0	23.4	71.4	23.1	70.6
Diffusion models													
DiffAD	73.2	69.6	54.6	96.1	19.4	16.7	23.1	18.4	65.5	21.7	68.8	21.6	68.3
ImDiffusion	70.5	69.1	52.8	99.8	14.8	11.1	22.3	13.8	54.9	19.0	63.2	18.9	62.5
D3R	73.0	67.8	54.3	90.4	19.2	15.2	26.2	15.8	66.3	21.5	69.1	21.3	67.2
AnomFilter	72.2	70.8	54.9	99.7	14.8	10.0	28.7	14.5	59.5	21.1	68.5	20.9	67.7
MATSD	70.4	61.9	48.0	87.2	13.0	8.1	31.9	11.6	47.8	18.0	53.3	17.7	52.3
ICDiffAD	71.1	67.7	51.4	98.9	14.8	10.7	24.0	15.6	60.7	20.1	66.7	20.0	66.1
FSDI	70.0	67.8	51.3	99.9	14.5	10.9	21.7	10.8	51.1	15.9	60.3	15.9	59.6
PhysDiff	70.7	69.0	52.7	99.8	15.1	10.5	26.5	15.3	53.5	21.0	63.2	20.8	62.5
Ours	**75.29**	**75.43**	**62.18**	**95.88**	**24.62**	**18.67**	**36.13**	**25.52**	**71.03**	**29.82**	**75.33**	**29.41**	**74.61**

SMAP

Table D.4.2: Additional TAB evaluation metrics (%) on SMAP.

Model	ACC	AFF-F1	AFF-P	AFF-R	R-F1	R-P	R-R	A-P	A-R	R-A-P	R-A-R	V-PR	V-ROC
Traditional methods													
IForest	86.7	51.1	42.4	64.2	9.2	9.3	9.2	12.2	48.8	13.5	49.8	13.6	50.3
LOF	63.6	64.4	50.2	90.0	16.3	13.0	21.7	17.5	62.6	19.2	58.4	19.3	58.4
OCSVM	81.2	50.0	38.9	69.8	9.9	7.9	13.5	10.2	39.3	11.7	43.1	11.7	42.7
CBLOF	81.3	49.7	39.3	70.6	9.8	7.8	13.0	15.1	57.2	16.8	56.8	16.6	56.6
EIF	70.3	61.3	47.3	89.6	13.1	9.2	22.2	16.4	64.5	17.9	61.5	18.0	62.3
HBOS	81.5	50.6	39.8	70.6	9.4	7.5	12.7	14.7	58.8	16.6	56.3	16.4	56.4
KMeans	86.1	51.3	49.6	53.1	11.9	11.7	12.1	12.7	40.1	13.2	38.7	13.2	38.3
KNN	66.2	62.5	48.6	89.2	16.5	14.6	19.5	16.6	63.2	17.9	56.9	18.0	56.9
LODA	87.4	50.0	39.8	67.6	8.3	8.3	8.3	9.5	36.1	10.3	41.2	10.4	40.6
PCA	81.5	50.2	39.7	68.3	10.5	8.5	13.8	10.4	39.9	11.8	42.0	11.7	41.5
Time-series models													
AE	87.2	43.0	46.2	40.2	7.6	14.7	5.1	10.1	39.8	11.7	42.8	11.7	43.1
VAE	86.4	48.4	40.8	59.5	7.1	9.6	5.7	10.4	41.5	12.2	45.1	12.2	45.1
OmniAnomaly	67.4	65.6	49.2	98.3	12.9	8.5	26.9	11.5	46.6	13.0	49.6	13.1	49.5
USAD	66.8	58.2	44.9	82.6	10.9	6.8	28.3	12.0	50.0	13.6	50.9	13.6	50.9
GANF	67.2	52.1	41.9	69.0	12.8	8.4	27.5	11.6	36.3	12.3	36.2	12.3	36.1
DAGMM	16.4	39.4	39.2	39.7	6.1	7.5	5.1	14.0	56.7	16.8	57.5	17.0	57.8
LSTMED	86.8	48.5	40.8	60.3	7.2	9.6	5.7	10.8	43.8	12.4	46.0	12.3	46.0
Torsk	69.7	64.7	52.0	85.6	14.7	9.2	36.2	15.1	51.7	17.2	51.9	17.3	52.5
TranAD	87.9	53.8	50.1	58.1	7.6	10.2	6.0	9.6	37.0	11.2	37.5	11.2	37.7
AnomTrans	86.1	69.3	53.5	98.4	14.9	12.2	19.3	12.8	50.3	14.8	50.4	14.8	50.8
DCdetector	83.0	68.2	52.3	98.2	15.8	13.5	19.0	12.8	50.0	14.9	50.4	14.8	50.7
ContraAD	16.4	39.3	39.1	39.6	6.1	7.5	5.1	9.6	36.4	10.3	37.7	10.4	37.9
DLinear	68.5	60.6	44.4	95.4	13.0	9.0	23.9	10.3	39.0	11.9	42.8	11.9	42.7
NLinear	69.0	60.5	45.4	94.8	12.4	8.3	25.6	11.0	43.3	12.7	47.4	12.6	47.2
DUET	68.4	61.7	45.9	93.7	12.0	8.0	24.3	11.0	43.3	12.5	47.4	12.5	46.6
iTransformer	68.4	57.4	43.5	85.6	11.3	7.8	20.3	11.3	40.9	12.5	43.9	12.6	44.1
PatchTST	72.0	62.4	46.2	96.0	12.2	8.2	25.1	11.0	44.2	12.9	48.2	13.0	47.7
TimesNet	68.4	63.9	47.6	97.9	14.4	9.3	32.3	12.0	43.9	13.5	48.3	13.6	48.4
ModernTCN	68.6	61.9	46.0	95.3	12.6	8.3	26.4	11.1	43.5	12.5	47.0	12.4	47.1
Time-frequency models													
TFAD	86.8	67.5	51.6	97.8	12.3	12.5	12.1	12.8	50.0	14.2	50.5	14.2	50.5
DualTF	65.9	66.9	51.4	99.6	14.8	11.0	22.7	12.3	46.4	14.1	50.9	13.8	49.9
TFMAE	70.3	70.4	54.4	99.6	17.3	14.2	22.1	15.4	58.5	17.4	61.3	17.4	61.1
CATCH	73.1	57.9	43.9	84.8	11.2	7.8	21.0	11.5	42.9	13.0	46.3	12.9	45.5
Diffusion models													
DiffAD	68.8	70.7	54.7	99.8	16.5	13.4	21.3	13.0	51.6	15.1	55.9	15.1	55.7
ImDiffusion	68.0	66.8	50.2	99.9	16.1	12.7	21.8	12.0	49.5	13.6	52.7	13.7	52.7
D3R	67.1	54.8	45.5	69.0	10.0	8.9	11.4	10.7	41.8	12.0	40.7	12.0	40.5
AnomFilter	69.3	68.4	52.1	99.4	18.2	14.1	25.5	13.5	53.4	15.3	56.6	15.2	56.5
MATSD	66.7	53.1	41.4	74.2	11.3	7.9	19.5	12.6	40.4	13.3	42.9	13.3	42.8
ICDiffAD	67.2	55.7	42.4	80.9	15.7	12.5	21.0	11.9	50.5	13.7	51.8	13.7	51.8
FSDI	68.4	67.8	51.3	99.9	15.7	12.5	21.1	12.5	49.1	14.2	52.7	14.2	52.6
PhysDiff	68.2	67.5	51.0	99.9	15.5	11.9	22.3	12.3	47.7	14.2	51.5	14.2	51.4
Ours	**71.85**	**71.66**	**55.96**	**99.57**	**19.86**	**14.70**	**30.62**	**16.90**	**62.01**	**19.15**	**65.26**	**19.14**	**65.09**

SMD

Table D.4.3: Additional TAB evaluation metrics (%) on SMD.

Model	ACC	AFF-F1	AFF-P	AFF-R	R-F1	R-P	R-R	A-P	A-R	R-A-P	R-A-R	V-PR	V-ROC
Traditional methods													
IForest	95.9	62.5	79.6	51.4	19.3	45.7	12.3	12.1	65.9	9.8	68.2	9.9	68.3
LOF	31.2	67.9	51.5	99.6	6.2	3.2	71.4	6.7	64.0	6.9	62.0	6.9	62.1
OCSVM	86.1	74.2	65.0	86.4	11.2	6.7	32.3	10.4	60.7	7.9	57.9	8.1	57.7
CBLOF	84.9	73.2	62.0	89.2	13.1	8.0	37.4	12.2	66.4	9.9	65.7	10.0	65.6
EIF	88.1	72.2	63.1	84.7	12.0	7.4	31.9	16.5	64.7	10.9	63.3	11.0	62.5
HBOS	84.6	62.6	55.2	72.4	10.3	6.2	30.0	14.6	62.6	8.8	59.2	8.7	60.1
KMeans	69.0	69.0	62.5	77.0	16.2	9.3	63.3	20.1	82.0	14.1	69.0	14.3	68.6
KNN	45.6	69.2	53.8	99.8	7.4	3.9	76.3	19.0	71.3	13.4	71.2	13.4	71.5
LODA	86.9	73.6	64.5	85.6	12.2	7.6	31.5	11.2	62.6	8.6	60.7	8.8	60.9
PCA	86.5	73.2	67.4	80.1	18.9	12.4	40.4	12.9	68.5	10.9	66.0	10.9	64.9
Time-series models													
AE	95.2	11.1	88.0	5.9	2.5	91.1	1.3	20.9	79.3	20.3	79.9	20.1	79.4
VAE	96.4	44.9	87.8	30.2	14.7	53.5	8.5	11.5	64.6	9.7	64.2	9.8	64.3
OmniAnomaly	84.1	80.3	72.1	90.7	21.5	14.2	44.3	9.4	69.5	8.6	67.7	8.6	67.7
USAD	88.4	42.5	79.8	29.0	24.7	36.8	18.6	10.2	68.0	7.2	51.5	7.2	51.5
GANF	88.3	40.6	78.3	27.4	25.4	28.9	22.6	12.0	71.3	7.8	53.3	7.8	53.3
DAGMM	95.1	2.3	48.6	1.2	0.0	0.0	0.0	4.5	52.3	28.8	53.2	28.7	53.4
LSTMED	94.9	52.0	88.2	36.9	13.9	19.7	10.7	12.4	67.8	11.5	69.1	11.5	68.3
Torsk	74.0	59.4	52.6	69.4	7.9	4.5	32.6	6.3	57.7	5.6	43.8	5.6	43.6
TranAD	95.0	78.5	77.6	79.4	21.4	16.8	29.3	11.2	66.4	10.4	66.6	10.4	67.5
AnomTrans	94.5	69.3	55.6	92.1	11.7	8.8	17.7	4.5	50.9	5.8	50.0	5.8	50.3
DCdetector	77.4	67.1	51.0	99.4	7.0	4.1	25.9	4.2	49.5	4.6	49.2	4.6	49.3
ContraAD	4.2	2.3	48.2	1.2	0.0	0.0	0.0	4.2	49.2	4.3	43.7	4.3	44.0
DLinear	94.0	82.9	72.8	96.3	18.8	11.9	44.0	13.3	71.5	13.8	73.6	13.8	73.8
NLinear	94.0	83.8	74.0	96.6	19.3	12.4	42.5	13.8	72.3	14.3	74.7	14.2	74.6
DUET	93.4	82.5	75.0	91.6	13.8	8.5	36.1	13.6	72.5	13.5	73.1	13.4	73.6
iTransformer	87.2	81.6	70.5	98.9	14.2	8.7	42.3	14.4	73.9	14.1	75.4	14.3	75.9
PatchTST	91.9	82.5	73.2	97.4	16.3	10.1	41.5	14.7	73.2	14.5	75.6	14.4	76.1
TimesNet	95.2	81.0	77.4	86.6	20.7	14.8	35.9	16.0	75.7	15.9	78.0	15.9	78.5
ModernTCN	93.5	83.1	76.8	93.0	14.3	8.9	36.8	13.2	71.6	13.0	74.3	13.1	73.5
Time-frequency models													
TFAD	95.4	65.5	51.2	91.0	4.7	4.2	5.4	4.2	50.0	4.8	50.0	4.8	50.0
DualTF	85.9	67.5	52.4	95.4	8.5	5.5	18.6	17.4	69.6	10.4	60.2	10.6	59.6
TFMAE	86.5	65.5	49.1	98.3	7.1	4.3	20.4	3.5	39.9	3.7	38.8	3.7	38.6
CATCH	91.3	83.2	74.5	96.2	19.0	12.2	42.6	16.3	77.6	16.8	80.9	16.6	80.2
Diffusion models													
DiffAD	73.3	68.3	52.1	99.2	7.6	4.5	25.6	5.17	54.0	5.3	55.0	5.3	54.8
ImDiffusion	88.5	77.0	63.7	97.1	11.8	6.9	39.5	11.64	70.1	11.4	71.6	11.4	71.5
D3R	95.1	75.0	81.0	69.8	26.0	23.5	29.0	23.82	74.4	19.4	72.2	19.4	72.1
AnomFilter	92.0	75.5	63.5	92.9	12.9	8.4	28.4	9.65	58.9	8.3	60.1	8.3	60.0
MATSD	93.4	76.5	72.4	81.2	17.8	11.6	38.5	20.13	78.6	19.3	79.1	19.3	79.0
ICDiffAD	92.4	52.0	64.2	43.8	9.0	6.9	13.1	9.86	71.7	9.0	65.8	9.0	65.7
FSDI	73.3	67.8	51.3	99.8	7.9	4.6	29.1	4.72	53.7	5.1	55.5	5.1	55.3
PhysDiff	92.3	75.9	61.6	98.9	14.2	9.0	34.4	9.39	61.4	9.2	63.2	9.3	63.1
Ours	**94.55**	**80.71**	**76.35**	**89.94**	**28.43**	**29.61**	**39.18**	**42.83**	**80.70**	**40.25**	**82.48**	**40.17**	**82.19**

PSM

Table D.4.4: Additional TAB evaluation metrics (%) on PSM.

Model	ACC	AFF-F1	AFF-P	AFF-R	R-F1	R-P	R-R	A-P	A-R	R-A-P	R-A-R	V-PR	V-ROC
Traditional methods													
IForest	71.9	61.5	90.4	47.5	23.6	79.7	13.9	33.6	54.0	33.7	54.8	33.2	54.1
LOF	28.1	69.2	53.4	99.6	8.8	4.6	98.6	46.6	73.1	42.3	63.8	42.1	63.5
OCSVM	73.8	52.7	65.1	45.0	19.6	38.4	13.4	41.6	62.0	36.8	53.5	36.9	53.6
CBLOF	69.1	78.5	74.2	83.5	44.6	48.3	42.0	53.8	71.8	43.9	63.0	44.4	63.8
EIF	71.6	77.4	69.0	90.1	40.4	38.9	42.1	49.1	68.8	45.8	63.3	46.1	63.2
HBOS	74.1	65.2	62.6	70.1	33.8	46.4	26.7	48.3	70.8	43.9	63.9	44.1	65.1
KMeans	65.4	73.5	66.0	83.0	48.6	37.4	70.8	54.7	76.8	45.2	62.5	45.1	62.7
KNN	28.4	69.2	53.0	99.8	0.2	0.1	99.8	54.8	75.1	45.1	67.0	45.7	66.8
LODA	70.7	72.2	65.5	80.5	33.7	34.9	32.5	42.1	59.6	38.6	53.6	38.7	53.8
PCA	67.9	70.1	70.7	69.5	48.1	45.9	50.5	47.7	66.9	40.9	57.2	41.4	57.7
Time-series models													
AE	70.8	70.8	76.7	67.1	31.0	44.0	24.2	54.8	74.8	48.2	69.2	47.8	68.7
VAE	71.7	29.7	99.3	17.4	11.8	99.6	6.3	41.6	64.1	39.0	58.4	39.1	58.3
OmniAnomaly	69.3	74.3	68.2	81.6	45.1	48.8	41.8	44.4	67.4	40.9	61.3	40.9	61.2
USAD	67.6	52.9	71.5	42.0	30.6	29.1	32.4	47.7	59.9	34.0	43.1	34.2	43.8
GANF	73.9	71.2	69.9	72.5	44.7	37.7	55.1	52.8	73.2	48.2	65.2	48.2	65.0
DAGMM	70.8	46.0	50.4	42.7	5.3	11.0	3.5	41.5	63.7	40.7	61.1	40.4	61.3
LSTMED	73.5	70.6	80.9	63.5	37.2	69.0	25.5	58.8	76.7	52.8	70.5	51.9	70.1
Torsk	62.6	67.6	61.9	74.5	35.8	28.6	48.5	28.1	51.0	28.5	43.2	28.8	43.1
TranAD	72.5	74.5	76.2	73.6	46.0	56.4	39.2	49.9	63.5	44.1	57.7	43.8	57.5
AnomTrans	71.9	63.1	53.5	77.9	13.3	28.9	8.7	27.8	50.3	27.8	48.8	27.6	49.1
DCdetector	70.7	67.0	54.7	86.4	11.7	27.4	7.4	27.5	50.3	27.3	49.3	27.4	48.6
ContraAD	71.7	65.1	54.1	81.7	14.2	28.7	9.5	30.0	53.2	29.6	43.1	29.5	43.6
DLinear	73.2	82.8	77.4	89.0	47.4	58.5	40.4	35.5	56.0	35.1	55.7	35.4	56.2
NLinear	72.8	82.6	78.0	89.0	44.3	57.2	36.4	36.2	56.6	36.1	56.9	35.9	56.6
DUET	73.4	81.2	75.1	88.4	42.1	60.5	32.8	38.4	58.1	38.9	58.2	38.4	57.6
iTransformer	73.2	85.2	77.5	96.9	50.4	54.0	47.2	37.2	58.1	37.2	58.0	37.6	57.7
PatchTST	73.0	83.3	75.8	94.1	45.0	55.1	38.4	37.7	58.3	37.4	58.5	38.1	58.6
TimesNet	72.1	83.7	73.8	96.8	48.8	46.8	50.9	35.6	55.4	35.6	55.6	35.7	55.8
ModernTCN	72.5	81.5	72.0	93.9	44.5	52.5	38.6	39.0	58.3	39.0	58.0	38.9	58.0
Time-frequency models													
TFAD	72.0	62.2	53.1	75.1	9.1	27.1	5.4	27.8	50.0	29.4	49.9	29.3	50.0
DualTF	69.6	71.9	61.6	86.3	42.9	45.5	41.2	30.9	54.5	30.8	49.8	30.9	50.0
TFMAE	55.6	65.6	51.1	91.7	21.5	23.5	19.8	21.6	33.3	22.1	27.4	22.2	27.0
CATCH	73.4	83.6	75.9	94.1	47.9	54.5	43.0	38.6	59.7	38.9	59.4	38.6	60.1
Diffusion models													
DiffAD	63.1	68.7	53.5	95.9	29.4	30.3	28.6	30.9	54.0	30.7	52.0	30.8	51.7
ImDiffusion	62.0	69.7	54.3	97.3	26.3	27.4	25.3	29.6	51.8	29.5	49.2	29.5	49.0
D3R	60.6	66.7	51.1	96.2	27.7	28.6	26.9	26.7	47.9	27.0	45.2	27.1	45.0
AnomFilter	64.9	72.9	57.5	99.6	38.2	30.3	51.8	35.9	56.5	35.9	56.1	35.9	56.0
MATSD	69.8	79.7	66.4	99.8	43.5	30.6	75.1	43.7	64.5	44.1	64.3	44.2	64.3
ICDiffAD	72.7	24.2	69.4	14.6	17.9	46.0	11.1	40.8	66.7	38.3	59.2	38.4	59.3
FSDI	61.4	69.5	53.7	98.3	29.0	28.1	30.0	28.0	50.6	28.2	49.5	28.3	49.0
PhysDiff	68.1	75.5	60.8	99.4	33.3	26.5	45.0	40.8	59.8	40.2	59.1	40.1	58.6
Ours	**68.84**	**80.39**	**67.88**	**98.55**	**43.29**	**35.27**	**56.02**	**46.45**	**62.40**	**42.55**	**60.46**	**42.59**	**60.54**

SWaT

Table D.4.5: Additional TAB evaluation metrics (%) on SWaT.

Model	ACC	AFF-F1	AFF-P	AFF-R	R-F1	R-P	R-R	A-P	A-R	R-A-P	R-A-R	V-PR	V-ROC
Traditional methods													
IForest	87.0	68.0	57.3	83.5	12.9	12.9	13.3	9.3	34.9	13.3	43.0	12.9	42.1
LOF	12.1	68.7	52.6	99.1	0.2	0.1	99.1	73.0	84.1	48.5	71.7	46.1	70.3
OCSVM	19.5	69.7	53.8	98.9	2.7	1.4	93.3	16.9	65.6	53.6	63.3	52.8	61.6
CBLOF	80.2	73.1	62.7	87.6	14.9	8.7	52.9	76.5	88.5	55.2	80.7	53.5	78.3
EIF	90.6	76.5	75.8	77.4	13.0	7.7	40.5	50.2	88.0	51.5	84.0	49.5	81.5
HBOS	74.9	65.1	52.9	84.3	11.6	6.7	43.8	73.6	82.3	58.0	81.9	56.1	78.1
KMeans	55.5	60.2	63.5	57.1	32.1	23.2	52.1	72.3	82.7	42.1	60.6	40.8	58.5
KNN	12.0	68.7	52.5	99.3	0.7	0.3	99.9	77.9	86.7	52.3	73.9	51.1	71.7
LODA	93.1	53.0	62.7	45.9	7.3	9.0	6.1	69.3	79.4	40.8	69.8	40.4	68.9
PCA	37.5	71.1	57.3	93.4	12.8	7.0	72.0	73.0	81.8	53.7	72.2	50.7	68.8
Time-series models													
AE	71.1	72.8	60.0	92.4	15.7	15.1	16.8	70.6	81.8	45.8	64.5	44.2	62.9
VAE	89.1	70.8	63.9	79.9	19.8	56.2	12.0	74.9	83.3	60.4	77.9	57.0	74.0
OmniAnomaly	83.8	62.0	65.5	58.8	26.3	28.9	24.2	29.9	75.7	30.2	69.6	29.7	67.6
USAD	65.7	46.6	57.8	39.1	6.3	3.5	31.6	12.7	55.0	17.2	47.8	17.0	46.9
GANF	93.3	55.2	72.4	44.6	11.5	7.9	21.4	72.0	84.6	48.5	74.9	46.5	72.1
DAGMM	75.6	73.2	62.5	89.6	17.2	18.6	16.4	12.7	53.0	57.6	53.3	57.7	52.9
LSTMED	76.2	66.2	54.4	86.6	8.1	5.5	15.1	10.7	23.5	12.1	29.5	11.7	28.4
Torsk	69.0	66.4	53.8	87.1	22.4	14.7	46.4	13.5	58.8	18.6	65.1	17.9	63.2
TranAD	60.6	71.9	64.5	82.1	26.3	17.8	50.4	73.2	82.4	52.5	71.2	49.7	69.3
AnomTrans	88.0	69.3	54.2	96.1	15.5	13.3	18.5	12.9	49.9	15.0	50.7	14.7	51.3
DCdetector	85.8	69.0	53.5	97.6	20.4	20.4	20.3	12.2	50.1	15.3	50.7	15.0	51.3
ContraAD	12.1	16.3	53.1	9.6	0.2	0.1	2.8	8.1	22.3	11.0	29.9	10.8	28.6
DLinear	83.5	71.3	56.9	96.8	11.3	7.9	19.9	8.6	30.6	11.6	39.9	11.4	39.0
NLinear	83.5	71.9	57.6	95.6	12.6	9.5	19.9	8.1	23.7	10.9	33.9	10.9	33.4
DUET	83.5	71.2	59.7	89.3	11.0	7.6	19.4	8.6	24.3	12.5	35.8	12.0	34.6
iTransformer	86.8	71.9	58.5	93.3	12.8	10.5	16.9	8.3	23.5	12.0	34.9	11.5	33.4
PatchTST	83.0	71.4	58.4	91.8	11.4	8.3	18.4	8.6	23.9	12.4	35.6	12.2	34.5
TimesNet	87.1	79.6	70.7	91.1	16.4	13.8	20.2	11.0	28.9	18.4	40.0	17.4	39.4
ModernTCN	79.6	71.8	58.0	94.2	10.1	6.6	22.0	9.2	23.8	13.5	35.2	12.9	34.3
Time-frequency models													
TFAD	87.5	68.6	53.1	96.6	14.9	12.1	19.6	12.2	50.0	16.2	50.9	16.0	50.8
DualTF	62.9	69.0	53.4	99.8	15.7	12.1	22.0	14.3	56.8	17.2	60.7	17.2	60.4
TFMAE	69.9	68.2	52.0	99.0	16.5	13.6	21.0	14.0	56.0	16.5	59.4	16.4	59.3
CATCH	83.2	71.4	57.9	93.2	11.1	8.1	18.0	8.2	23.1	12.0	34.7	11.6	33.9
Diffusion models													
DiffAD	79.8	59.2	57.6	60.9	8.8	5.7	19.5	11.7	20.4	11.3	28.0	11.1	27.2
ImDiffusion	84.0	31.4	54.7	22.0	4.2	7.2	3.0	11.7	50.7	15.8	44.2	15.7	43.6
D3R	67.8	72.2	56.5	99.8	16.7	12.9	23.7	14.3	59.7	19.1	65.4	18.6	64.8
AnomFilter	69.0	72.9	58.7	96.0	17.3	14.1	22.5	12.8	56.3	15.7	60.0	15.6	59.3
MATSD	67.8	73.4	58.6	97.9	13.1	8.2	32.2	10.7	36.0	15.1	46.9	14.6	45.8
ICDiffAD	84.3	23.2	61.0	14.3	0.1	0.0	14.3	13.0	54.4	17.1	49.0	17.0	47.9
FSDI	69.2	68.9	52.6	99.9	15.5	12.5	20.3	12.5	51.3	15.2	56.3	15.1	56.0
PhysDiff	93.2	73.7	65.8	83.8	10.6	7.4	18.8	70.6	81.5	65.6	81.7	64.0	81.1
Ours	**93.09**	**45.39**	**59.36**	**36.74**	**8.04**	**5.32**	**16.42**	**69.66**	**79.50**	**38.56**	**61.22**	**37.74**	**60.00**