西北師範大學
NORTHWEST NORMAL UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

PhysDiff: A Physically-Guided Diffusion Model for Multivariate Time Series Anomaly Detection

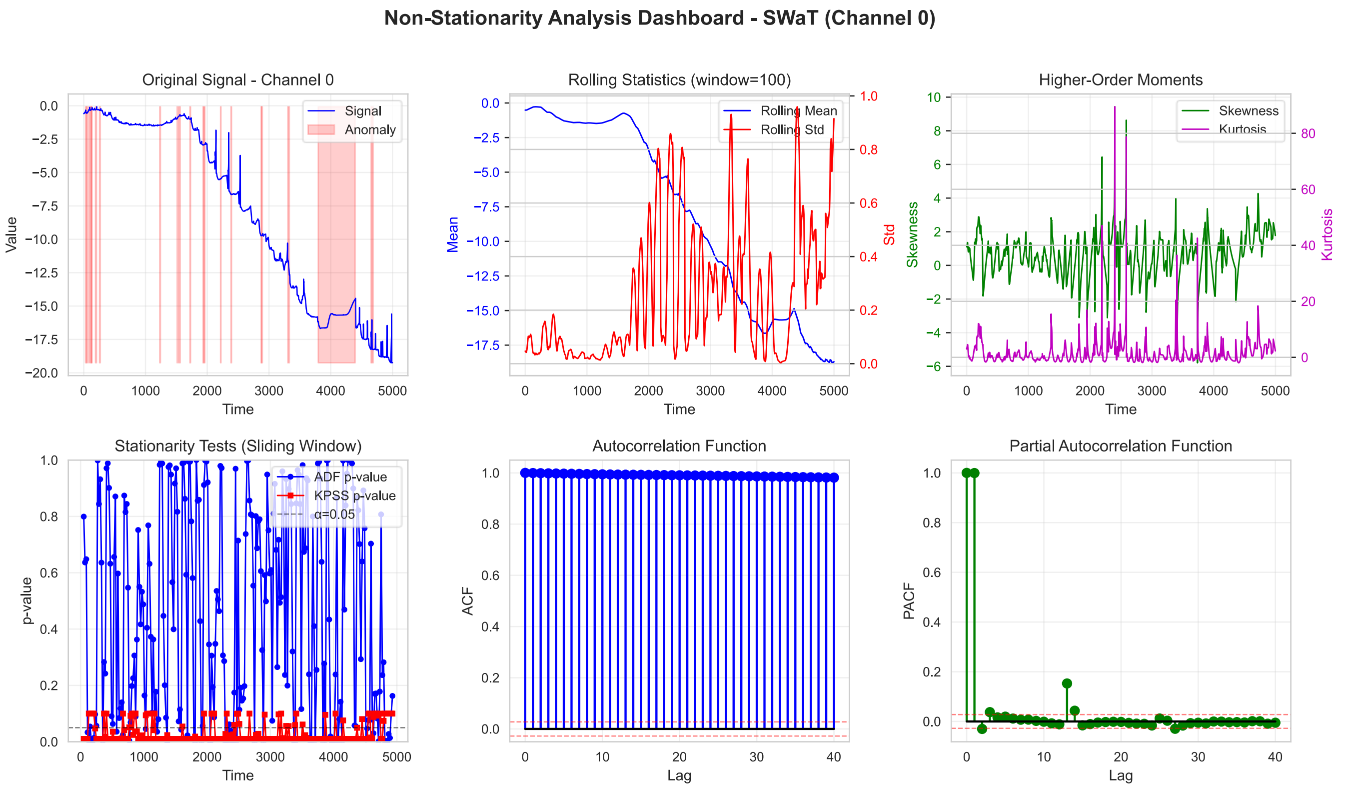
Long Li *, Wencheng Zhang*, Shi Yuan*, Hongle Guo and Wanghu Chen†

*These authors contributed equally to this work.
†Corresponding author.
Li L, Chen W, et al. PhysDiff: A Physically-Guided Diffusion Model for Multivariate Time Series Anomaly Detection[J]. Advances in Neural Information Processing Systems(NeurIPS), 2025.

1 Background & Motivations

The Challenge

Multivariate time series anomaly detection faces critical challenges in non-stationary environments with high false-positive rates and limited interpretability.



Key Findings: Non-Stationarity Analysis

1. Strong Non-Stationarity Confirmed

Significant downward trend (0 → −20)

ADF (p≈1.0) and KPSS (p≈0) both reject stationarity

ACF shows no decay → requires differencing

2. Time-Varying Statistics

Variance fluctuates dramatically (peaks during anomalies)

Extreme kurtosis spikes (>80) indicate heavy tails

Statistical properties unstable across time

- Pattern anomalies are subtle and remain within normal ranges
- Non-stationary data exhibits overlapping dynamics
- Methods struggle when point anomalies are obscured

2 Analysis & Solutions

西北師範大學 NORTHWEST NORMAL UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

Two Key Challenges

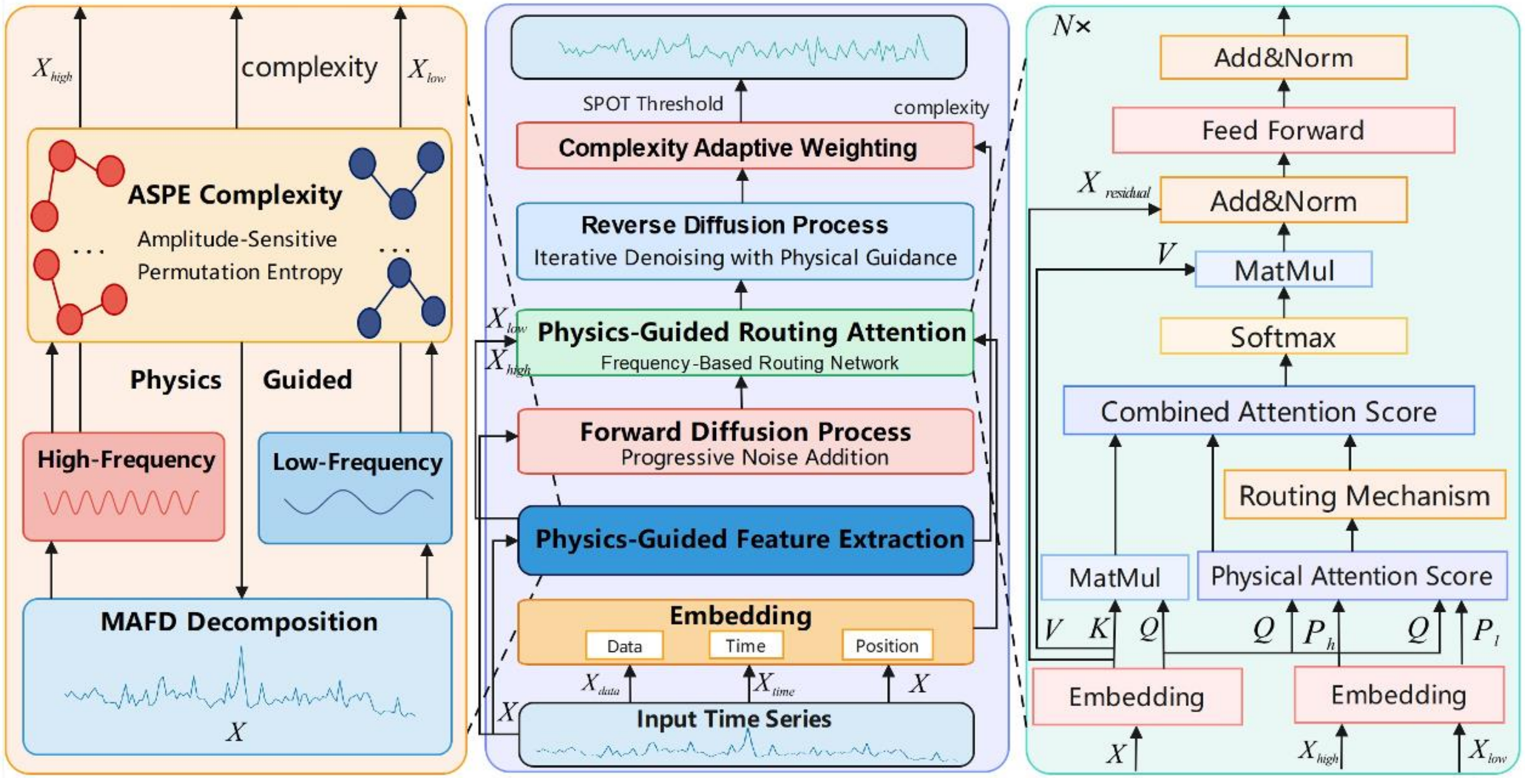
Adaptive Decomposition: Existing methods predominantly employ fixed decomposition patterns, making it difficult to handle complex data.

Frequency Integration: During signal reconstruction, how to strike a well-founded balance between detail and stability remains a challenge.

Solutions

Proposing the PhysDiff Framework

- **Physically Guided Feature Extraction:** Separates high-frequency oscillations from low-frequency trends to reduce noise and provide physical prior information;
- **Physically Informed Diffusion Model:** Performs reconstruction via diffusion proc-esses conditioned on physical prior inform ation to highlight anomalous deviations;
- **Anomaly Detection Scoring:** Fuses reco -nstruction errors with physical plausibility to generate anomaly scores.



2 Analysis & Solutions

NORTHWEST NORMAL UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

(1) Physics-Guided Feature Extraction

Through the combination of Multi-channel Adaptive Fourier Decomposition (MAFD) and Amplitude-Sensitive Permutation Entropy (ASPE), providing interpretable inputs for subsequent diffusion models.

(2) Physically-Informed Diffusion Model

Through the conditional diffusion process, focusing on solving the problem of the lack of physical constraints in the reconstruction of non-stationary data in traditional diffusion models.

Frequency-Based Routing Attention

$$\mathrm{Attention}(Q,K,V,\mathbf{P}_h,\mathbf{P}_l)=\mathrm{softmax}\left(\frac{QK^T+g_h\cdot Q\mathbf{P}_h^T+g_l\cdot Q\mathbf{P}_l^T}{\sqrt{d_k}}\right)V$$

Physical Consistency through Energy Guidance

$$\mathcal{E}(\mathbf{x}_t)=\|\nabla_{\mathbf{x}}\Phi(\mathbf{x}_t)\|^2 \qquad \mathbf{x}_{t-1}\leftarrow\mathbf{x}_t-\lambda\nabla_{\mathbf{x}_t}\mathcal{E}(\mathbf{x}_t)+\sqrt{2\lambda}\mathbf{n}$$

(3) Physics-Driven Anomaly Detection :

For each sliding window, we define an anomaly score that fuses reconstruction fidelity with physical plausibility

$$\mathrm{Score}(t)=\alpha\cdot\underbrace{\|\mathbf{X}(t)-\mathbf{X}(t)\|_2}_{\text{Reconstruction Error}}+(1-\alpha)\cdot\underbrace{D_{\mathrm{KL}}(P_{\mathrm{TTFD}}\|P_{\mathrm{prior}})}_{\text{Time-Frequency Distribution Divergence}}$$

3 Experiments

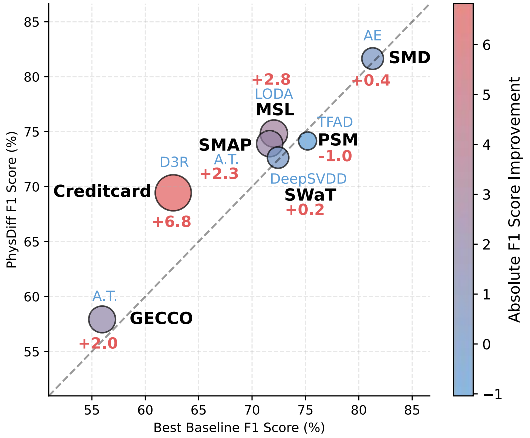
NORTHWEST NORMAL UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

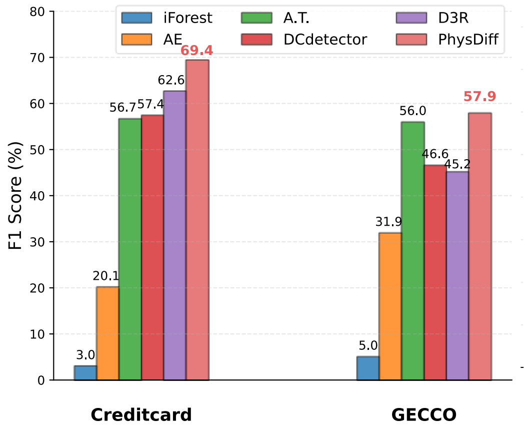
Comparative Study

Our comprehensive comparative analysis evaluates **PhysDiff** over 18 state-of-the-art anomaly detection methods across real-world benchmark dataset.

Dataset	SMD			MSL			SMAP			SWaT			PSM		
Metric	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**
OCSVM	66.98	62.03	73.75	50.26	99.86	66.87	41.05	69.37	51.58	56.08	98.72	72.11	57.51	58.11	57.81
PCA	64.92	40.19	54.34	52.69	98.33	68.61	50.62	98.48	66.87	62.32	82.96	71.18	77.44	37.71	53.53
HBOS	56.28	63.11	62.17	59.25	83.32	69.25	41.54	66.17	51.04	57.71	29.82	43.21	100.00	6.54	12.28
LOF	57.69	99.10	72.92	49.89	72.18	59.00	47.92	82.86	60.72	53.20	96.73	68.65	53.90	99.91	70.02
IForset	100.00	9.37	17.13	53.87	94.58	68.65	41.12	68.91	51.51	53.03	62.80	62.03	100.00	3.35	6.48
LODA	59.02	66.18	62.40	57.79	95.65	72.05	51.51	100.00	68.00	56.30	70.14	62.54	62.22	40.17	56.05
AE	69.22	98.48	81.30	55.75	96.66	70.72	39.42	70.31	50.52	54.92	98.20	70.45	60.67	98.24	75.01
DAGMM	63.57	70.83	67.00	54.07	92.11	68.14	50.75	96.38	66.49	59.42	92.36	72.32	68.22	70.50	69.34
LSTM	60.12	84.77	70.35	58.82	14.68	23.49	55.25	27.70	36.90	49.99	82.11	62.15	57.06	95.92	71.55
BeatGAN	74.11	81.64	77.69	55.74	98.94	71.30	54.04	98.30	69.74	61.89	83.46	71.08	58.81	99.08	73.81
Omni	79.09	75.77	77.40	51.23	99.40	67.61	52.74	98.51	68.70	62.76	82.82	71.41	69.20	80.79	74.55
A.T.	100.00	3.19	6.19	51.04	95.36	66.49	56.91	96.69	71.65	53.63	59.94	57.59	52.01	82.18	64.55
DCdetector	50.93	95.57	66.45	55.94	95.53	70.56	53.12	98.37	68.99	53.25	98.12	69.03	54.72	86.36	66.99
SensitiveHUE	60.34	90.13	72.29	55.92	98.95	71.46	53.63	98.37	69.42	58.91	91.71	71.74	56.15	98.75	71.59
DeepSVDD	64.98	64.77	64.88	10.53	100.00	19.06	29.73	7.09	11.45	59.11	93.53	72.44	74.05	50.64	60.15
MTAD-GAT	85.90	67.69	75.71	54.96	94.93	69.81	39.05	93.99	55.08	65.90	77.51	71.23	79.90	60.14	68.63
TFAD	56.32	97.83	71.49	54.96	94.93	69.81	39.05	93.99	55.08	60.38	81.96	69.53	79.14	71.63	**75.20**
D3R	64.87	97.93	78.04	56.45	95.55	71.81	51.08	94.46	66.30	64.25	77.50	70.25	53.17	100.00	69.43
PhysDiff	71.03	96.00	**81.65**	62.75	92.66	**74.83**	64.36	86.81	**73.91**	60.00	92.04	**72.64**	66.09	84.47	74.16

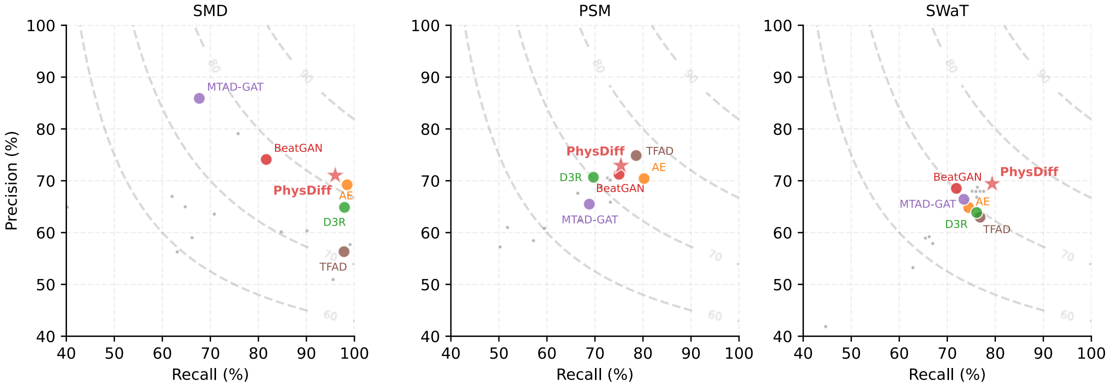


Performance comparison between PhysDiff and baseline methods across seven datasets.



Performance comparison on NeurIPS-TS datasets (Creditcard and GECCO).

Precision-Recall analysis across SMD, PSM, and SWaT datasets. PhysDiff (marked with a star) achieves a better balance between precision and recall compared to competing methods, maintaining high scores on both metrics. Contour lines represent F1 score values.

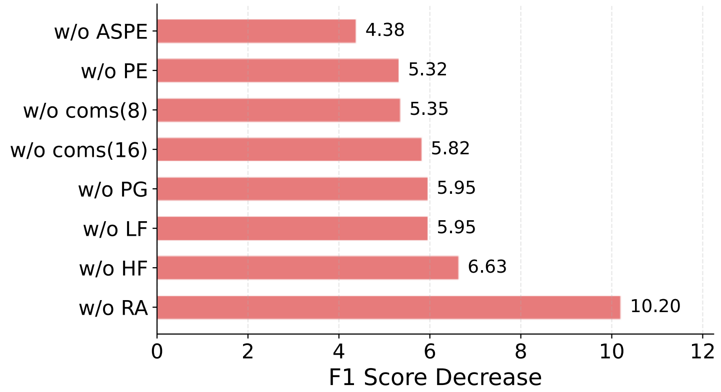


3 Experiments

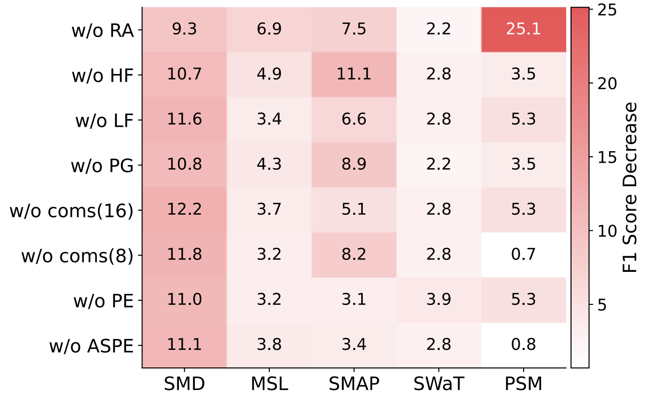
西北師範大學 NORTHWEST NORMAL UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

Ablation Study: Each Component's Contribution

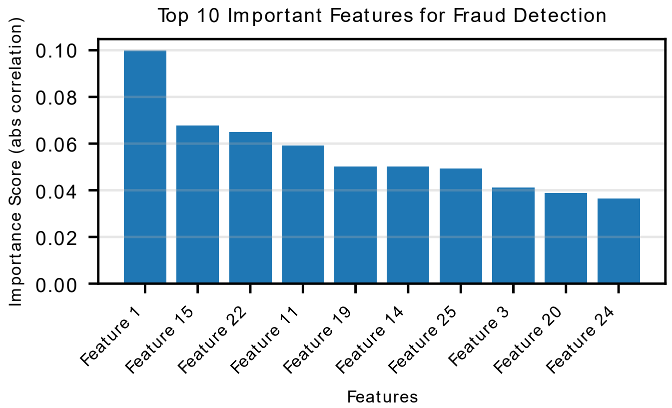


Impact of removing individual components from PhysDiff, measured by F1 score decrease averaged across all datasets.

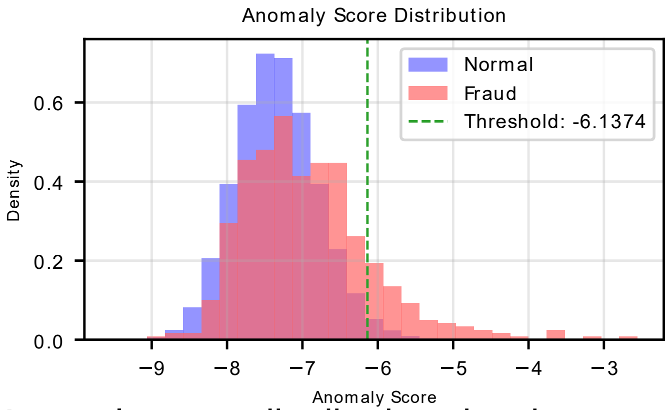


The dataset-specific impact of removing each component. Darker colors indicate larger F1 score decreases.

Case Study: Fraud Detection in Financial Transactions



Top 10 features for fraud detection, with Feature 1 (0.0998) contributing most significantly.



Anomaly score distribution showing separation between normal transactions and fraudulent ones.

西北師範大學
NORTHWEST NORMAL UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

Thank You for Listening

References

Ze W, ChiMan W, et al. Adaptive fourier decomposition for multi-channel signal analysis. IEEE Transactions on Signal Processing, 70:903–918,Jan 2022.

Chaoli Z, Tian Z, et al. Tfad: A decomposition time series anomaly detection architecture with time-frequency analysis. In Proceedings of the 31st ACM international conference, pages 2497–2507, 2022.

Contact

Long Li	2023222197@nwnu.edu.cn
Wencheng Zhang	202421162217@nwnu.edu.cn
Shi Yuan	2023222147@nwnu.edu.cn
Hongle Guo	guohongleself@nwnu.edu.cn
Wanghu Chen	chenwh@nwnu.edu.cn