



PAMod: Modeling Cyclical Shifts via Phase-Amplitude Modulation for Non-stationary Time Series Forecasting

Yingbo Zhou¹, Yutong Ye², Shuhao Li¹, Rui Qian¹, Qiang Huang¹,

Lemao Liu¹, Li Sun³, Dejing Dou^{1,4*}

¹Fudan University,

²Zhejiang Sci-Tech University,

³Beijing University of Posts and Telecommunications,

⁴BEDI Cloud, *Corresponding author

Motivation

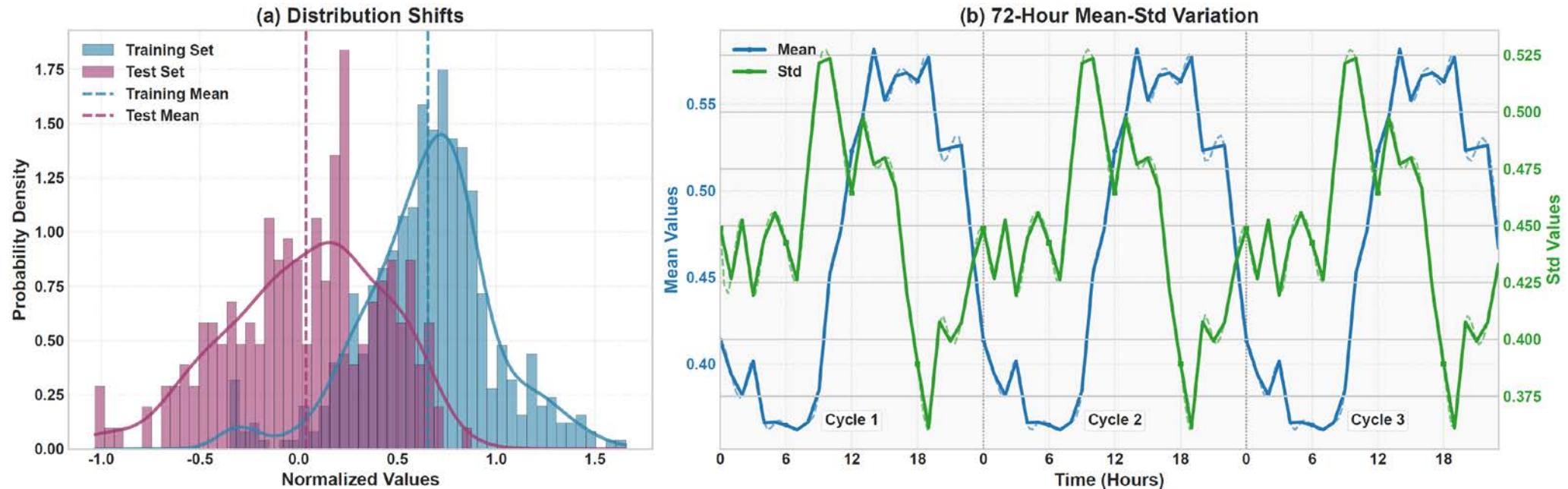


Figure 1: Cyclical distribution shifts exhibit stable mean-variance variations.

- **The Problem:** Real-world time series are non-stationary (shifting mean & variance).
- **The Insight:** Shifts are structured, not random. They follow **cyclical patterns** (e.g., 24h periodicity).
- **Our Goal:** Explicitly model these predictable cyclical shifts instead of just removing them.

Methodology

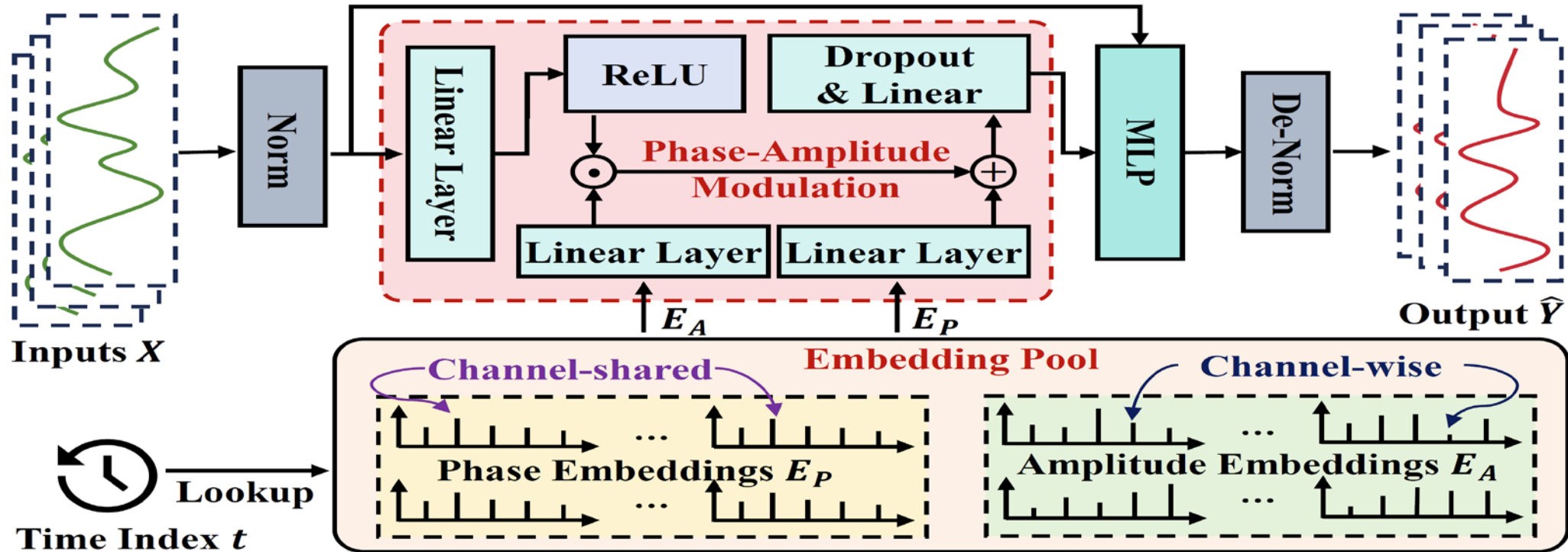


Figure 2: Overview of PAMod. Modulating in normalized space equals dynamic denormalization.

- **PAMod**: A lightweight, plug-and-play framework.
- **How it works**: Uses learnable cyclical embeddings to modulate normalized features, where **phase (additive)** for **mean shifts** and **amplitude (multiplicative)** for **variance changes**.

Main Results

Model	PAMod		TQNet		TimeEmb		FilterTS		Amplifier		TimeXer		CycleNet		TimeMixer		iTransformer		PatchTST	
	(Ours)		<u>[2025]</u>		<u>[2025]</u>		<u>[2025]</u>		<u>[2025]</u>		<u>[2024]</u>		<u>[2024]</u>		<u>[2024]</u>		<u>[2024]</u>		<u>[2023]</u>	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	0.363	0.378	0.377	0.393	<u>0.368</u>	<u>0.385</u>	0.385	0.396	0.382	0.395	0.382	0.397	0.379	0.396	0.381	0.395	0.407	0.410	0.387	0.400
ETTm2	0.263	0.307	0.277	0.323	<u>0.265</u>	<u>0.308</u>	0.277	0.322	0.280	0.326	0.274	0.322	0.266	0.314	0.275	0.323	0.288	0.332	0.281	0.326
ETTh1	0.413	0.420	0.441	0.434	<u>0.425</u>	<u>0.425</u>	0.434	0.430	0.430	0.428	0.437	0.437	0.457	0.441	0.447	0.440	0.454	0.448	0.469	0.455
ETTh2	0.360	0.387	0.378	0.402	<u>0.362</u>	<u>0.390</u>	0.375	0.398	0.381	0.405	0.368	0.396	0.388	0.409	0.364	0.395	0.383	0.407	0.387	0.407
ECL	<u>0.165</u>	0.254	0.164	<u>0.259</u>	0.168	0.261	0.180	0.272	0.172	0.266	0.171	0.270	0.168	0.259	0.182	0.272	0.178	0.270	0.205	0.290
Traffic	<u>0.434</u>	0.272	0.445	<u>0.276</u>	0.453	0.295	0.470	0.315	0.483	0.317	0.466	0.287	0.472	0.314	0.484	0.297	0.428	0.282	0.481	0.300
Weather	<u>0.239</u>	0.261	0.242	0.269	0.237	<u>0.262</u>	0.245	0.274	0.253	0.275	0.241	0.271	0.243	0.271	0.240	0.271	0.258	0.278	0.259	0.273
Solar	<u>0.206</u>	0.226	0.198	<u>0.256</u>	0.248	0.270	0.215	0.277	0.241	0.270	0.237	0.302	0.210	0.261	0.216	0.280	0.233	0.262	0.270	0.307
PEMS03	0.091	0.189	<u>0.097</u>	<u>0.203</u>	0.104	0.207	0.134	0.246	0.131	0.239	0.112	0.214	0.118	0.226	0.167	0.267	0.113	0.222	0.180	0.291
PEMS04	0.086	0.185	<u>0.091</u>	<u>0.197</u>	0.096	0.200	0.125	0.241	0.135	0.249	0.105	0.209	0.119	0.232	0.185	0.287	0.111	0.221	0.195	0.307
PEMS07	<u>0.080</u>	0.168	0.075	<u>0.171</u>	0.097	0.188	0.120	0.220	0.122	0.226	0.085	0.182	0.113	0.214	0.181	0.271	0.101	0.204	0.211	0.303
PEMS08	0.119	0.190	0.142	0.229	<u>0.132</u>	<u>0.221</u>	0.180	0.266	0.183	0.271	0.175	0.250	0.150	0.246	0.226	0.299	0.150	0.226	0.280	0.321
1 st Count	7	12	<u>3</u>	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

Table 1: Multivariate forecasting performance.

➤ **SOTA Accuracy:** 4%-16% MSE reduction on 12 benchmarks.

Main Results

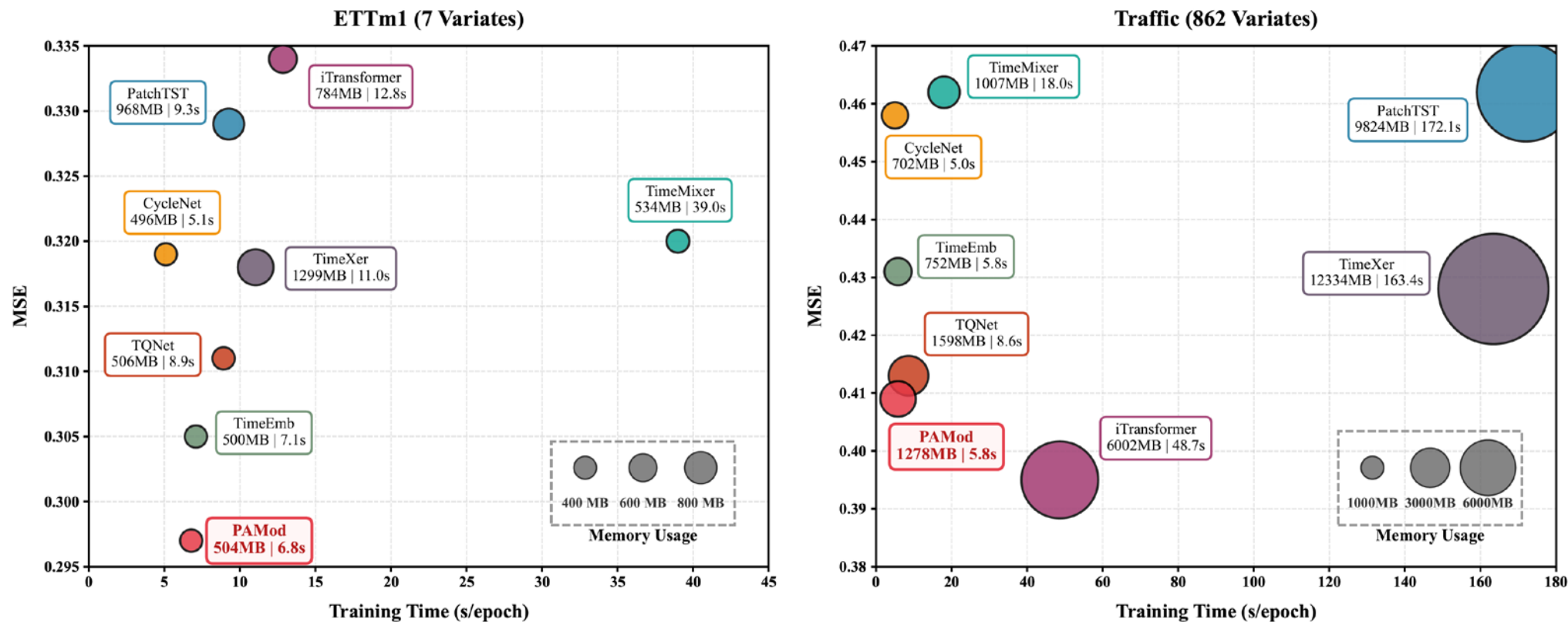


Figure 3: Computational efficiency under the input-96-predict-96 setting.

- **High Efficiency:** 5-10x fewer parameters; drastically lower memory usage.

Main Results

Dataset	ETTh2								ETTm1							
Horizon	96		192		336		720		96		192		336		720	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
DLinear [2023] +PAMod	0.333	0.387	0.477	0.476	0.594	0.541	0.831	0.657	0.345	0.372	0.380	0.389	0.413	0.413	0.474	0.453
	0.312	0.342	0.426	0.433	0.445	0.472	0.484	0.489	0.317	0.348	0.355	0.372	0.395	0.392	0.448	0.432
Imp.	6.3%	11.6%	10.7%	9.0%	25.1%	12.8%	41.8%	25.6%	8.1%	6.5%	6.6%	4.4%	4.4%	5.1%	5.5%	4.6%
PatchTST [2023] +PAMod	0.302	0.348	0.388	0.400	0.426	0.433	0.431	0.466	0.329	0.367	0.367	0.385	0.399	0.410	0.545	0.439
	0.293	0.334	0.365	0.388	0.401	0.418	0.419	0.438	0.312	0.345	0.352	0.370	0.380	0.394	0.478	0.428
Imp.	3.0%	4.0%	5.9%	3.0%	5.9%	3.5%	2.8%	6.0%	5.2%	6.0%	4.1%	3.9%	4.8%	3.9%	12.3%	2.5%
iTransformer [2024] +PAMod	0.297	0.349	0.380	0.400	0.428	0.432	0.427	0.445	0.334	0.368	0.377	0.391	0.426	0.420	0.491	0.459
	0.284	0.332	0.363	0.385	0.406	0.414	0.410	0.431	0.312	0.349	0.354	0.376	0.393	0.399	0.446	0.434
Imp.	4.4%	4.9%	4.5%	3.8%	5.1%	4.2%	4.0%	3.1%	6.6%	5.2%	6.1%	3.8%	7.7%	5.0%	9.2%	5.4%
TQNet [2025] +PAMod	0.295	0.343	0.367	0.393	0.417	0.427	0.433	0.446	0.311	0.353	0.356	0.378	0.390	0.401	0.452	0.440
	0.280	0.329	0.358	0.378	0.399	0.413	0.412	0.433	0.299	0.337	0.347	0.369	0.375	0.388	0.435	0.425
Imp.	5.1%	4.1%	2.5%	3.8%	4.3%	3.3%	4.8%	2.9%	3.9%	4.5%	2.5%	2.4%	3.8%	3.2%	3.8%	3.4%

Table 3: Improves existing backbones with simple integration.

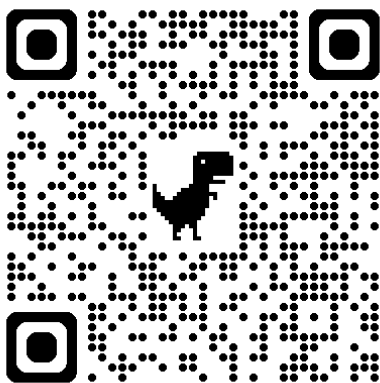
- **Architecture-Agnostic:** Seamlessly integrates into both MLP-based and Transformer-based backbones with better forecasting performance.

Conclusion

- ✓ **Modulation-based Paradigm:** We introduce Phase-Amplitude Modulation as a powerful alternative to traditional normalization, unifying dynamic denormalization with representation learning.
- ✓ **Structure over Complexity:** By explicitly capturing predictable cyclical shifts through decoupled mean-variance variations, we demonstrate that lightweight designs can significantly outperform massive architectures.
- ✓ **Practical Breakthrough:** Achieving 4%-16% MSE reduction with 5-10x fewer parameters establishes a new benchmark for highly accurate, interpretable, and resource-efficient forecasting.



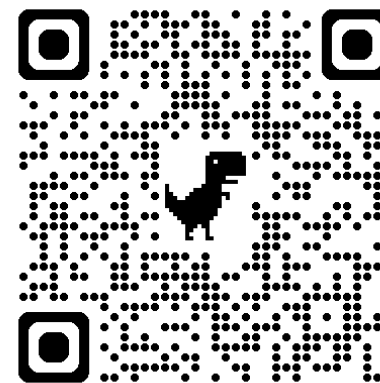
Thank You!



Paper



Code



Our Team