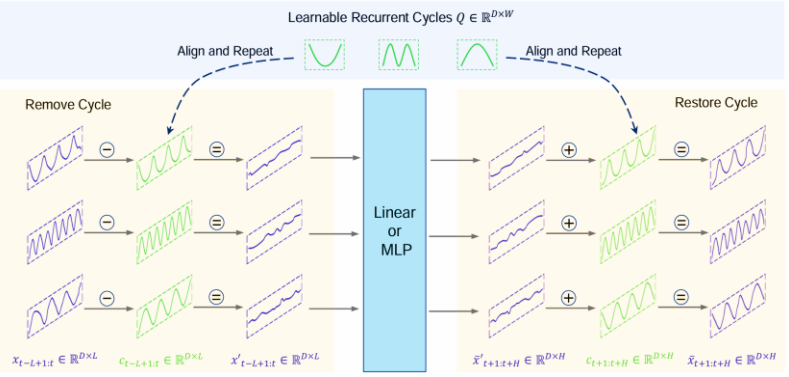
CycleNet: Enhancing Time Series Forecasting through Modeling Periodic Patterns

Shengsheng Lin[1], **Weiwei Lin**[1,2,*], **Xinyi Hu**[3], **Wentai Wu**[4], **Ruichao Mo**[1], **Haocheng Zhong**[1]
[1]School of Computer Science and Engineering, South China University of Technology, China
[2]Pengcheng Laboratory, China
[3]Department of Computer Science and Engineering, The Chinese University of Hong Kong
[4]College of Information Science and Technology, Jinan University, China
cslinshengsheng@mail.scut.edu.cn, linww@scut.edu.cn, xyhu@cse.cuhk.edu.hk, wentaiwu@jnu.edu.cn, {cs_moruichao, cshczhong}@mail.scut.edu.cn



Dataset	ETTh1		ETTh2		ETTm1		ETTm2		Electricity		Solar-Energy		Traffic		Weather	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Autoformer 2021	0.496	0.487	0.450	0.459	0.588	0.517	0.327	0.371	0.227	0.338	0.885	0.711	0.628	0.379	0.338	0.382
FEDformer 2022	0.440	0.460	0.437	0.449	0.448	0.452	0.305	0.349	0.214	0.327	0.291	0.381	0.610	0.376	0.309	0.360
SCINet 2022	0.747	0.647	0.954	0.723	0.485	0.481	0.571	0.537	0.268	0.365	0.282	0.375	0.804	0.509	0.292	0.363
DLinear 2023	0.456	0.452	0.559	0.515	0.403	0.407	0.350	0.401	0.212	0.300	0.330	0.401	0.625	0.383	0.265	0.317
TimesNet 2023	0.458	0.450	0.414	0.427	0.400	0.406	0.291	0.333	0.192	0.295	0.301	0.319	0.620	0.336	0.259	0.287
TiDE 2023	0.541	0.507	0.611	0.550	0.419	0.419	0.358	0.404	0.251	0.344	0.347	0.417	0.760	0.473	0.271	0.320
Crossformer 2023	0.529	0.522	0.942	0.684	0.513	0.496	0.757	0.610	0.244	0.334	0.641	0.639	0.550	0.304	0.259	0.315
PatchTST 2023	0.469	0.454	0.387	0.407	0.387	0.400	0.281	0.326	0.205	0.290	0.270	0.307	0.481	0.304	0.259	0.281
TimeMixer 2024	0.447	0.440	**0.364**	**0.395**	0.381	0.395	0.275	0.323	0.182	0.272	0.216	0.280	0.484	0.297	**0.240**	0.271
iTransformer 2024	0.454	0.447	0.383	0.407	0.407	0.410	0.288	0.332	0.178	0.270	0.233	0.262	**0.428**	**0.282**	0.258	0.278
CycleNet/Linear	**0.432**	**0.427**	0.383	0.404	0.386	**0.395**	0.272	0.315	0.170	0.260	0.235	0.270	0.485	0.313	0.254	0.279
CycleNet/MLP	0.457	0.441	0.388	0.409	**0.379**	0.396	**0.266**	**0.314**	**0.168**	**0.259**	**0.210**	**0.261**	0.472	0.301	0.243	**0.271**

CycleNet

通过建模周期模式增强时间序列预测

NIPS 2024

2024 NeurIPS（CCF-A）

CycleNet

Enhancing Time Series Forecasting through Modeling Periodic Patterns

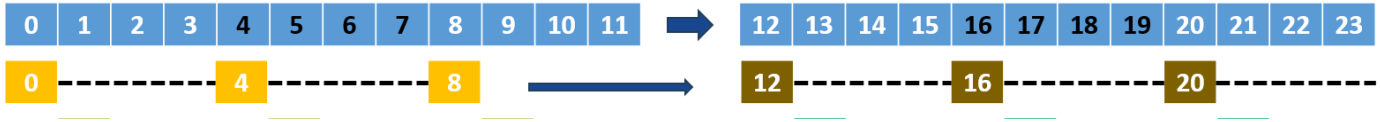
24.10.31

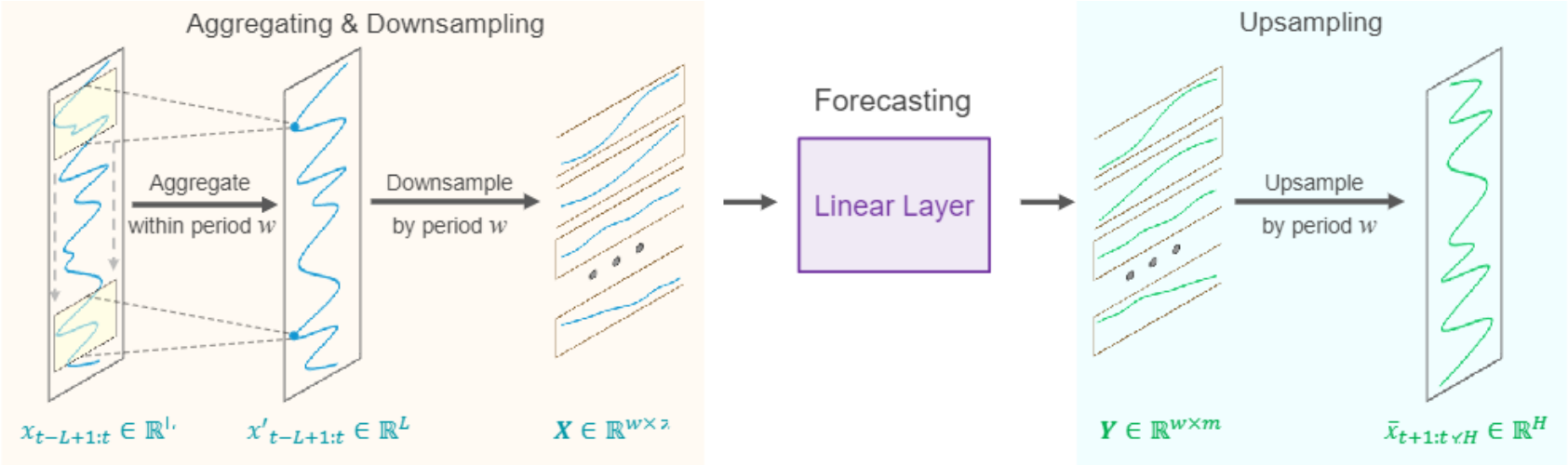
Presented by Yyyq

00 SparseTSF回顾（ICML2024）

Modeling Long-term Time Series Forecasting with 1k Parameters

- **滑动聚合**
  - $x^{(i)}_{t-L+1:t} = x^{(i)}_{t-L+1:t} + \text{Conv1D}(x^{(i)}_{t-L+1:t})$ kernel size of $2 \times \lfloor \frac{w}{2} \rfloor + 1$.
- **下采样：** $\mathrm{L} \rightarrow w \times n$
- **稀疏滑动预测：** $w \times n \rightarrow w \times m$
- **上采样：** $w \times m \rightarrow H$



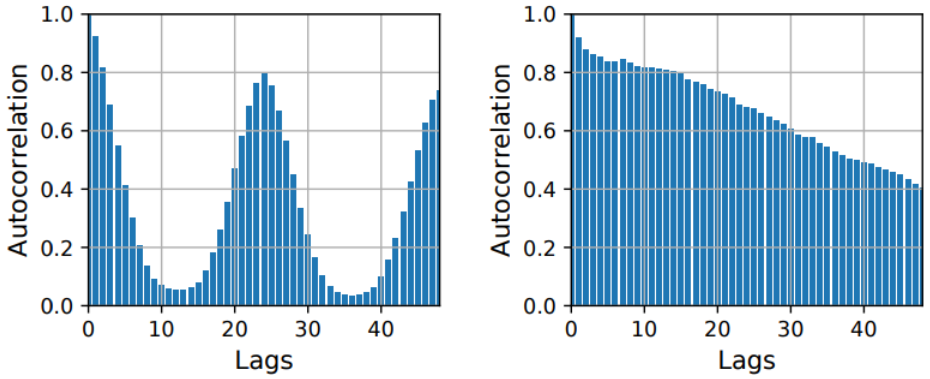


00 SparseTSF回顾（ICML2024）

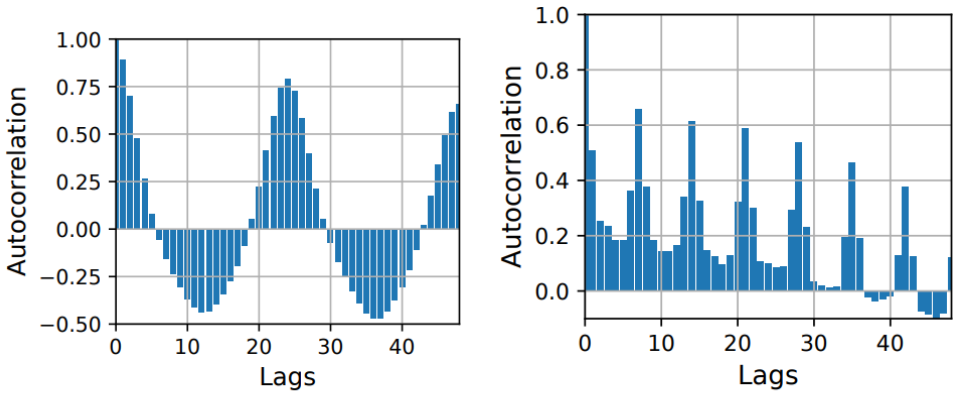
- **超长周期问题**
- **多周期问题**

回望窗口长度L=720

Datasets	ETTh1 & ETTh2	Electricity	Traffic
Channels	7	321	862
Frequency	hourly	hourly	hourly
Timesteps	17,420	26,304	17,544



(a) Original (b) Downsampled

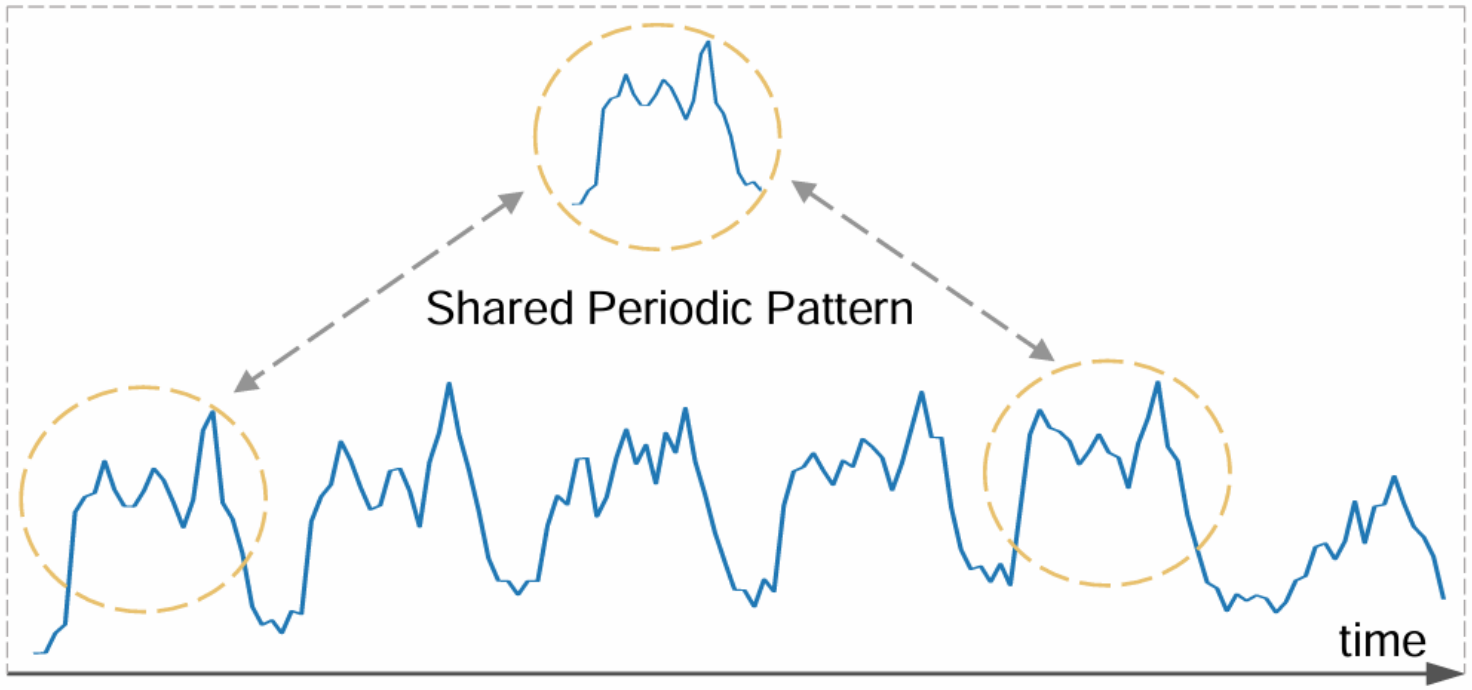


(a) Original (b) Downsampled

01 问题描述

- **时序数据内部潜在的稳定周期性：**长时预测中的共享周期模式

提取周期性特征 → **建模周期模式**

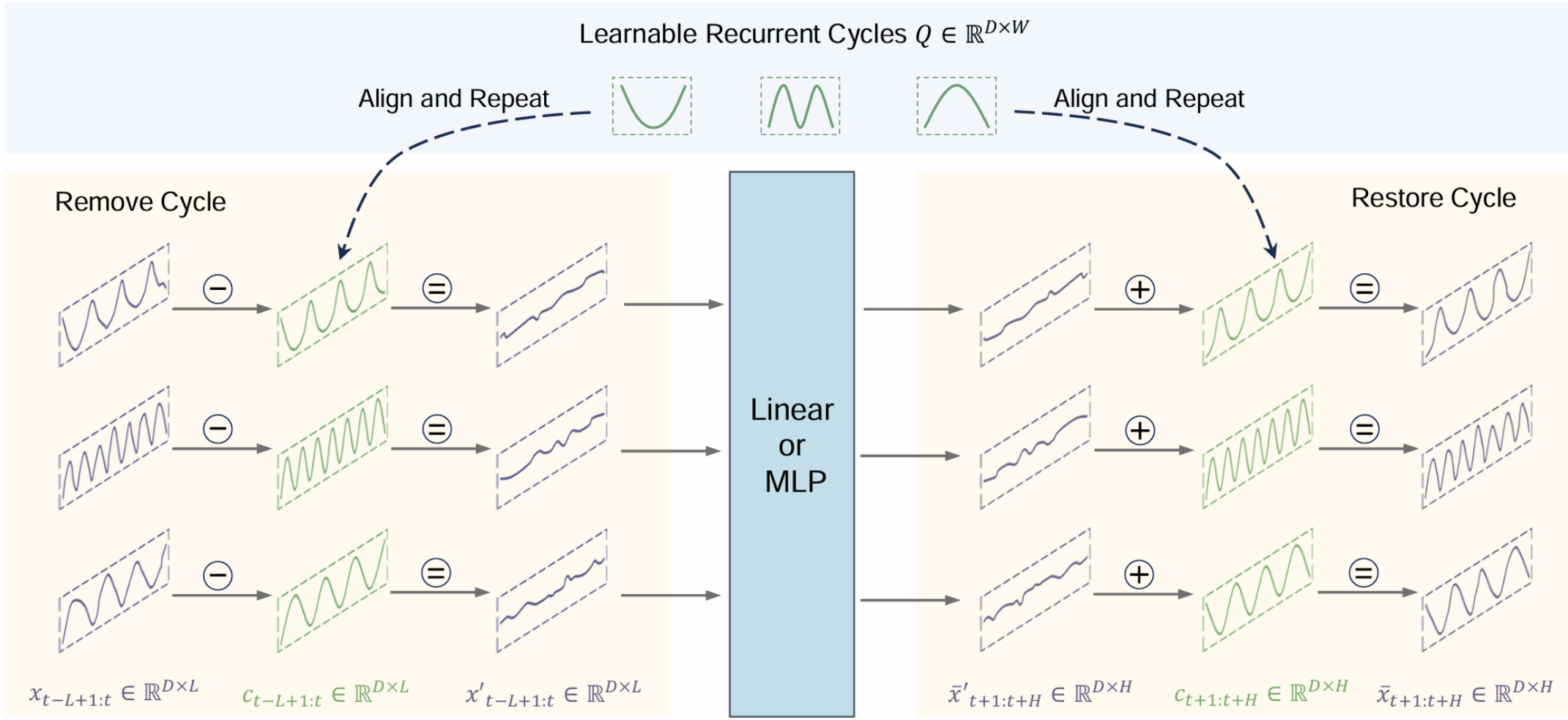


02 创新点

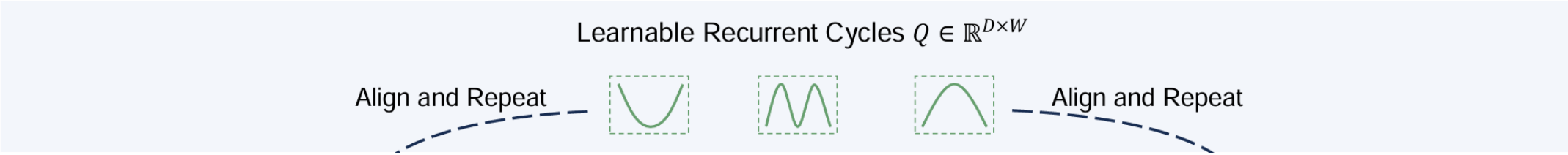
- **对周期模式进行显示建模**
- **RCF（Residual Cycle Forecasting）技术**
- **CycleNet**
  - RCF结合Linear/MLP，SOTA的同时，减少90%的参数量

概念简单，计算效率高

03 整体结构


Learnable Recurrent Cycles $Q \in \mathbb{R}^{D \times W}$
Align and Repeat
Align and Repeat
Remove Cycle
Restore Cycle
Linear or MLP
$x_{t-L+1:t} \in \mathbb{R}^{D \times L}$
$c_{t-L+1:t} \in \mathbb{R}^{D \times L}$
$x'_{t-L+1:t} \in \mathbb{R}^{D \times L}$
$\bar{x}'_{t+1:t+H} \in \mathbb{R}^{D \times H}$
$c_{t+1:t+H} \in \mathbb{R}^{D \times H}$
$\bar{x}_{t+1:t+H} \in \mathbb{R}^{D \times H}$

04 算法设计：周期模式建模



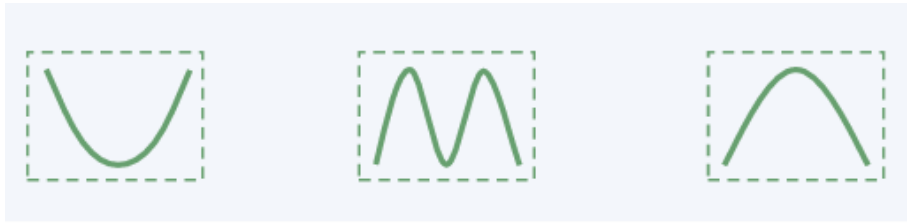
- **初始化：**可学习的循环周期Q，初始化为全0

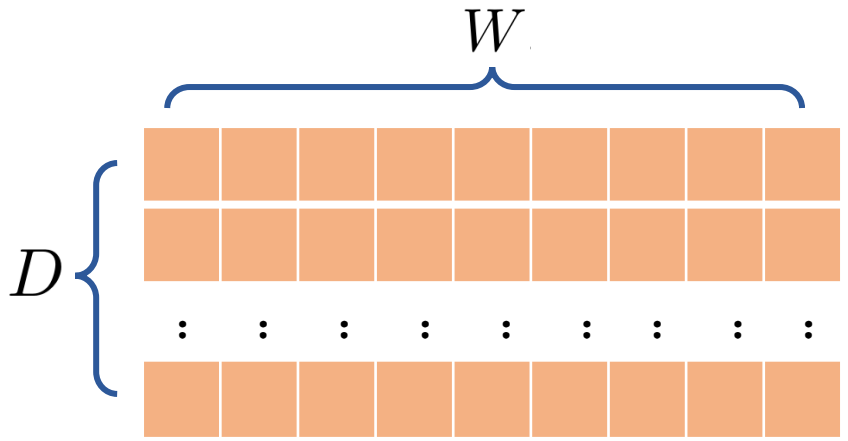
$$Q \in \mathbb{R}^{W \times D}$$

W：先验周期长度（最大稳定周期）

D：通道数/变量数

- **训练和更新：**全局共享，随主干一起反向传播-梯度更新



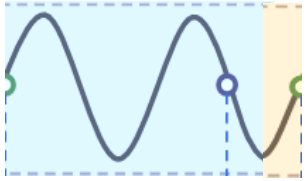


04 算法设计：RCF （Residual cycle forecasting）

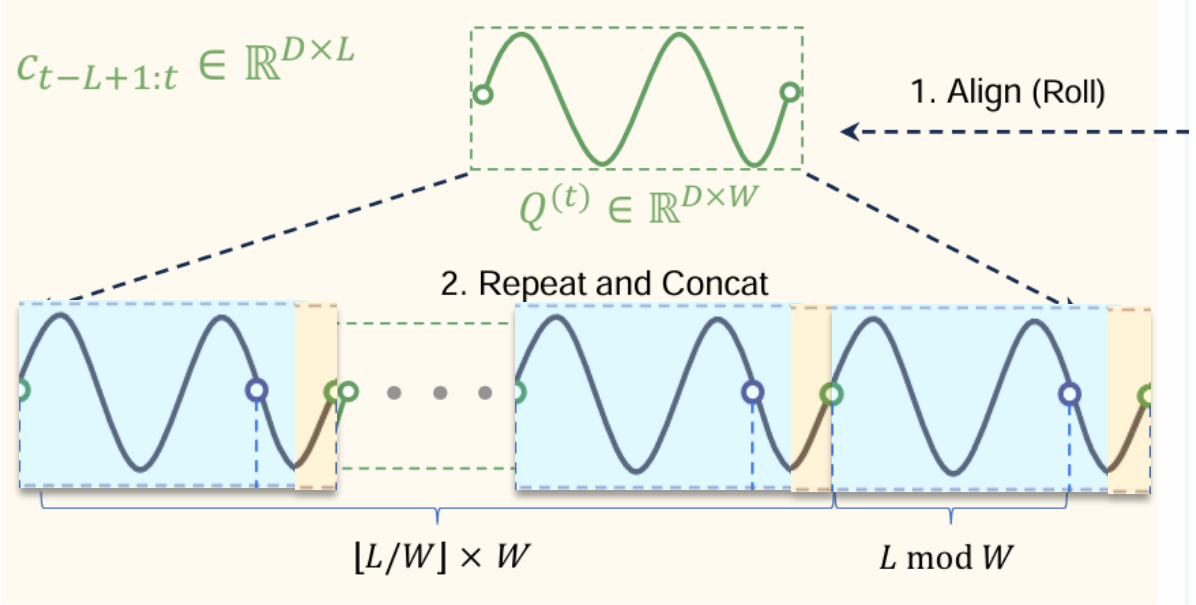
- **周期复制：**循环周期的排列和重复，匹配预测序列长度

L预测H

$$Q \in \mathbb{R}^{W \times D}$$



$$c_{t-L+1:t} \in \mathbb{R}^{D \times L}$$

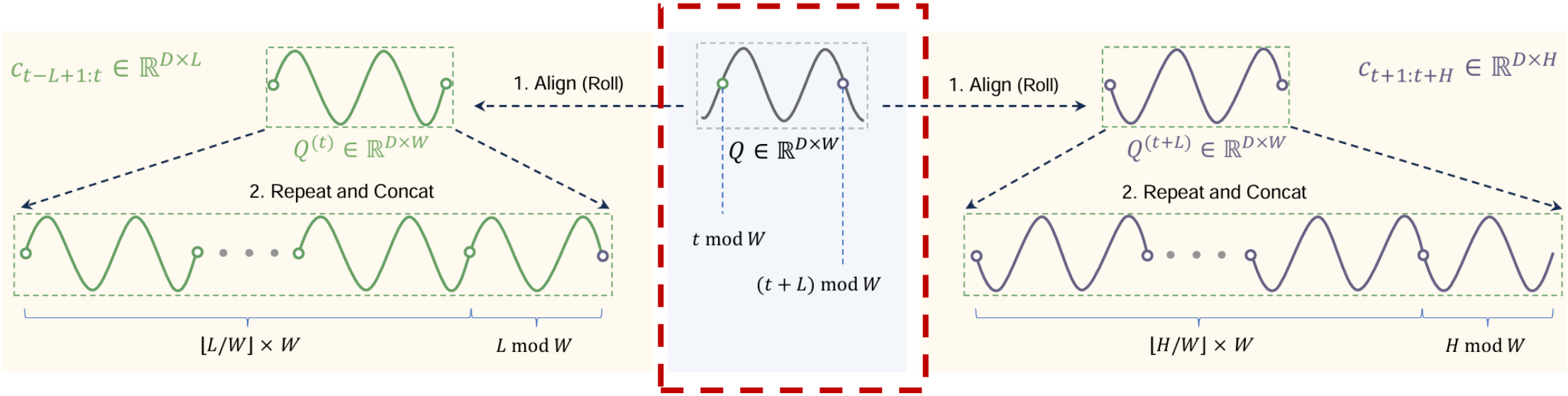


$$c_{t-L+1:t} = [\underbrace{Q^{(t)}, \cdots, Q^{(t)}}_{\lfloor L/W \rfloor}, Q^{(t)}_{0:L \bmod W}],$$

04 算法设计：RCF （Residual cycle forecasting）

- **周期复制：**循环周期的排列和重复，匹配预测序列长度

L预测H

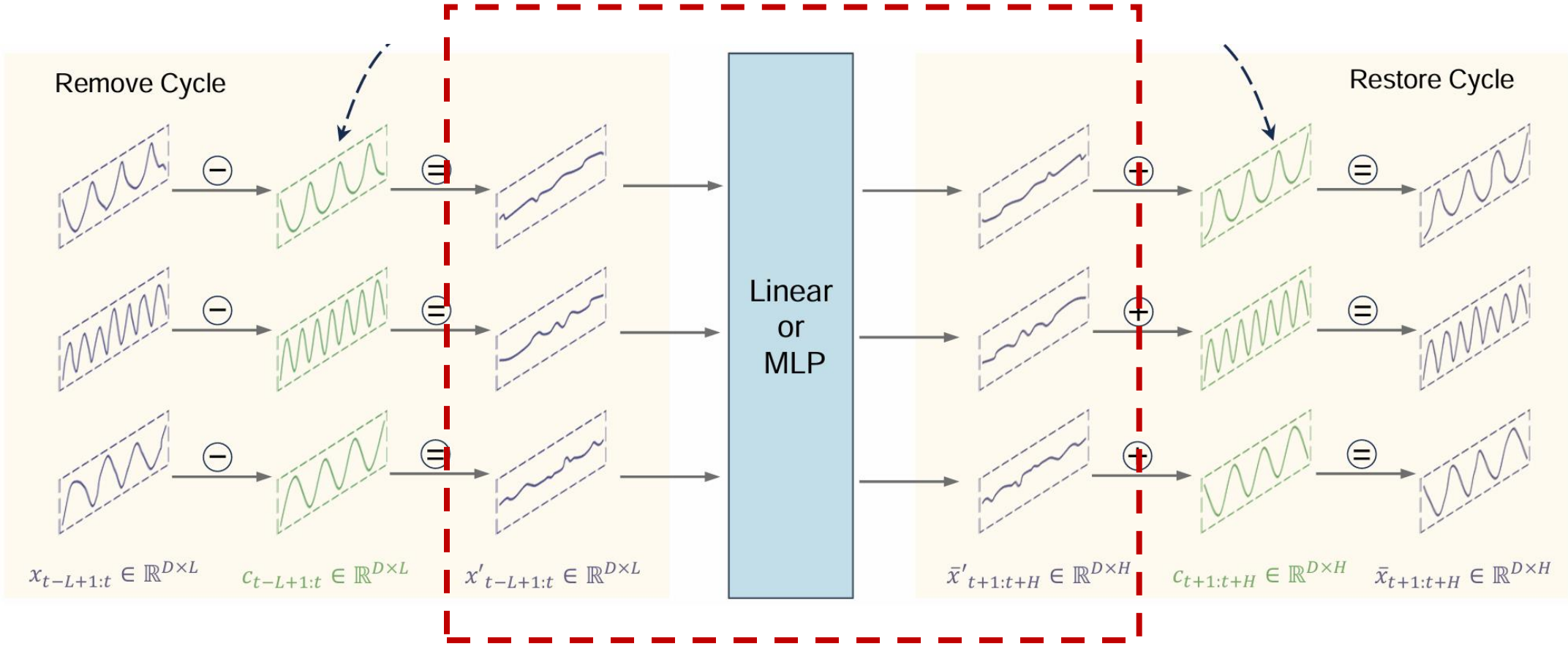


$$c_{t-L+1:t} = [\underbrace{Q^{(t)}, \cdots, Q^{(t)}}_{\lfloor L/W \rfloor}, Q^{(t)}_{0:L \bmod W}],$$

$$c_{t+1:t+H} = [\underbrace{Q^{(t+L)}, \cdots, Q^{(t+L)}}_{\lfloor H/W \rfloor}, Q^{(t+L)}_{0:H \bmod W}].$$

04 算法设计：RCF （Residual cycle forecasting）

- **剩余量预测**



04 算法设计：CycleNet

Linear or MLP

- **BackBone**

  通道独立：

  - 单层Linear（CycleNet/Linear）
  - 双层MLP（CycleNet/MLP）

- **归一化策略：RevIN**

$$x_{t-L+1:t} = \frac{x_{t-L+1:t} - \mu}{\sqrt{\sigma + \epsilon}},$$

$$\bar{x}_{t+1:t+H} = \bar{x}_{t+1:t+H} \times \sqrt{\sigma + \epsilon} + \mu,$$

05 实验：数据集

Dataset	ETTh1 & ETTh2	ETTm1 & ETTm2	Electricity	Solar-Energy	Traffic	Weather
Timesteps	17,420	69,680	26,304	52,560	17,544	52,696
Channels	7	7	321	137	862	21
Frequency	1 hour	15 mins	1 hour	10 mins	1 hour	10 mins
Cyclic Patterns	Daily	Daily	Daily & Weekly	Daily	Daily & Weekly	Daily
Cycle Length	**24**	**96**	**168**	**144**	**168**	**144**

05 实验：对比实验

L = 96

Dataset	ETTh1		ETTh2		ETTm1		ETTm2		Electricity		Solar-Energy		Traffic		Weather	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Autoformer [2021]	0.496	0.487	0.450	0.459	0.588	0.517	0.327	0.371	0.227	0.338	0.885	0.711	0.628	0.379	0.338	0.382
FEDformer [2022]	0.440	0.460	0.437	0.449	0.448	0.452	0.305	0.349	0.214	0.327	0.291	0.381	0.610	0.376	0.309	0.360
SCINet [2022]	0.747	0.647	0.954	0.723	0.485	0.481	0.571	0.537	0.268	0.365	0.282	0.375	0.804	0.509	0.292	0.363
DLinear [2023]	0.456	0.452	0.559	0.515	0.403	0.407	0.350	0.401	0.212	0.300	0.330	0.401	0.625	0.383	0.265	0.317
TimesNet [2023]	0.458	0.450	0.414	0.427	0.400	0.406	0.291	0.333	0.192	0.295	0.301	0.319	0.620	0.336	0.259	0.287
TiDE [2023]	0.541	0.507	0.611	0.550	0.419	0.419	0.358	0.404	0.251	0.344	0.347	0.417	0.760	0.473	0.271	0.320
Crossformer [2023]	0.529	0.522	0.942	0.684	0.513	0.496	0.757	0.610	0.244	0.334	0.641	0.639	0.550	0.304	0.259	0.315
PatchTST [2023]	0.469	0.454	0.387	0.407	0.387	0.400	0.281	0.326	0.205	0.290	0.270	0.307	0.481	0.304	0.259	0.281
TimeMixer [2024]	0.447	0.440	**0.364**	**0.395**	0.381	0.395	0.275	0.323	0.182	0.272	0.216	0.280	0.484	0.297	**0.240**	0.271
iTransformer [2024]	0.454	0.447	0.383	0.407	0.407	0.410	0.288	0.332	0.178	0.270	0.233	0.262	**0.428**	**0.282**	0.258	0.278
CycleNet/Linear	**0.432**	**0.427**	0.383	0.404	0.386	**0.395**	0.272	0.315	0.170	0.260	0.235	0.270	0.485	0.313	0.254	0.279
CycleNet/MLP	0.457	0.441	0.388	0.409	**0.379**	0.396	**0.266**	**0.314**	**0.168**	**0.259**	**0.210**	**0.261**	0.472	0.301	0.243	**0.271**

05 实验：对比实验——交通场景下的进一步分析

Dataset	Traffic	
Metric	MSE	MAE
Autoformer [2021]	0.628	0.379
FEDformer [2022]	0.610	0.376
SCINet [2022]	0.804	0.509
DLinear [2023]	0.625	0.383
TimesNet [2023]	0.620	0.336
TiDE [2023]	0.760	0.473
Crossformer [2023]	0.550	0.304
PatchTST [2023]	0.481	0.304
TimeMixer [2024]	0.484	0.297
iTransformer [2024]	**0.428**	**0.282**
CycleNet/Linear	0.485	0.313
CycleNet/MLP	0.472	0.301

MSE差距约10%

Model		CycleNet /MLP		CycleNet /Linear		RLinear [2023]		iTransformer [2024]		PatchTST [2023]		Crossformer [2023]		DLinear [2023]	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
PEMS03	12	**0.066**	0.172	0.080	0.192	0.126	0.236	0.071	0.174	0.099	0.216	0.090	0.203	0.122	0.243
	24	0.089	0.201	0.120	0.237	0.246	0.334	0.093	0.201	0.142	0.259	0.121	0.240	0.201	0.317
	48	0.136	0.247	0.156	0.258	0.551	0.529	**0.125**	**0.236**	0.211	0.319	0.202	0.317	0.333	0.425
	96	0.182	0.282	0.199	0.292	1.057	0.787	**0.164**	**0.275**	0.269	0.370	0.262	0.367	0.457	0.515
PEMS04	12	0.078	0.186	0.089	0.201	0.138	0.252	0.078	0.183	0.105	0.224	0.098	0.218	0.148	0.272
	24	0.099	0.212	0.127	0.245	0.258	0.348	0.095	0.205	0.153	0.275	0.131	0.256	0.224	0.340
	48	0.133	0.248	0.169	0.286	0.572	0.544	0.120	0.233	0.229	0.339	0.205	0.326	0.355	0.437
	96	0.167	0.281	0.189	0.293	1.137	0.820	0.150	0.262	0.291	0.389	0.402	0.457	0.452	0.504
PEMS07	12	**0.062**	**0.162**	0.075	0.183	0.118	0.235	0.067	0.165	0.095	0.207	0.094	0.200	0.115	0.242
	24	**0.086**	0.192	0.113	0.225	0.242	0.341	0.088	**0.190**	0.150	0.262	0.139	0.247	0.210	0.329
	48	0.128	0.234	0.157	0.254	0.562	0.541	**0.110**	**0.215**	0.253	0.340	0.311	0.369	0.398	0.458
	96	0.176	0.268	0.207	0.291	1.096	0.795	**0.139**	0.245	0.346	0.404	0.396	0.442	0.594	0.553
PEMS08	12	0.082	0.185	0.091	0.201	0.133	0.247	**0.079**	**0.182**	0.168	0.232	0.165	0.214	0.154	0.276
	24	0.117	0.226	0.140	0.251	0.249	0.343	**0.115**	**0.219**	0.224	0.281	0.215	0.260	0.248	0.353
	48	**0.169**	0.268	0.200	0.291	0.569	0.544	0.186	**0.235**	0.321	0.354	0.315	0.355	0.440	0.470
	96	0.233	0.306	0.272	0.328	1.166	0.814	**0.221**	**0.267**	0.408	0.417	0.377	0.397	0.674	0.565
Avg.		0.125	0.229	0.149	0.252	0.514	0.482	**0.119**	**0.218**	0.217	0.306	0.220	0.304	0.320	0.394

MSE差距约5%

05 实验：对比实验——交通场景下的进一步分析

	Traffic	Electricity	Solar-Energy	ETTh1	PEMS03	PEMS04	PEMS07	PEMS08
Avg. Extreme Points	**23.8**	1.4	0	0	0.9	0.1	3.5	4.8
Avg. Max Extreme	**9.27**	4.14	2.92	4.08	2.87	2.66	2.61	2.77
Cosine Similarity	0.56	0.46	**0.92**	0.21	0.84	0.77	0.80	0.78

- 每个通道内极值点（ Z-Score > 6 ）的平均数量
- 每个通道的平均最大极值
- 通道之间的余弦相似度

- **对RCF的影响：** RCF 中学习到的平均循环向显著的异常值倾斜，循环中某个点的平均值被夸大
- **时空关系建模的重要性：** 交通场景中合理的时空关系建模(或多元关系建模)是必不可少的

05 实验：效率实验

Model	Parameters	MACs	Training Time(s)
Informer [2021]	12.53M	3.97G	70.1
Autoformer [2021]	12.22M	4.41G	107.7
FEDformer [2022]	17.98M	4.41G	238.7
DLinear [2023]	139.6K	44.91M	18.1
PatchTST [2023]	10.74M	25.87G	129.5
iTransformer [2024]	5.15M	1.65G	35.1
CycleNet/MLP	472.9K	134.84M	30.8
CycleNet/Linear	123.7K	22.42M	29.6
RCF part	53.9K	0	12.8

05 实验：模型分析——RCF的有效性

Dataset	Electricity								Traffic							
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
Linear	0.197	0.274	0.197	0.277	0.212	0.292	0.253	0.324	0.645	0.383	0.598	0.361	0.605	0.362	0.643	0.381
+ RCF	0.141	0.234	0.155	0.247	0.172	0.264	0.210	0.296	0.480	0.314	0.482	0.313	0.476	0.303	0.503	0.320
Improve	**28.6%**	**14.6%**	**21.4%**	**10.8%**	**18.8%**	**9.5%**	**17.1%**	**8.7%**	**25.6%**	**18.0%**	**19.5%**	**13.2%**	**21.3%**	**16.2%**	**21.8%**	**16.1%**
MLP	0.175	0.259	0.181	0.265	0.197	0.282	0.240	0.317	0.500	0.325	0.496	0.321	0.509	0.325	0.542	0.342
+ RCF	0.136	0.229	0.152	0.244	0.170	0.264	0.212	0.299	0.458	0.296	0.457	0.294	0.470	0.299	0.502	0.314
Improve	**22.2%**	**11.6%**	**15.9%**	**8.0%**	**13.6%**	**6.3%**	**11.6%**	**5.7%**	**8.5%**	**8.9%**	**7.9%**	**8.3%**	**7.7%**	**8.0%**	**7.3%**	**8.1%**
DLinear	0.195	0.278	0.194	0.281	0.207	0.297	0.243	0.331	0.649	0.398	0.599	0.372	0.606	0.375	0.646	0.396
+ RCF	0.143	0.240	0.156	0.253	0.171	0.270	0.204	0.302	0.506	0.317	0.499	0.317	0.512	0.325	0.545	0.343
Improve	**26.6%**	**13.6%**	**19.7%**	**10.0%**	**17.4%**	**8.9%**	**16.3%**	**8.8%**	**22.1%**	**20.4%**	**16.6%**	**14.6%**	**15.4%**	**13.3%**	**15.6%**	**13.5%**
PatchTST	0.168	0.260	0.176	0.266	0.193	0.282	0.233	0.317	0.436	0.281	0.449	0.285	0.464	0.293	0.499	0.310
+ RCF	0.136	0.231	0.153	0.246	0.170	0.264	0.211	0.299	0.438	0.264	0.457	0.270	0.469	0.275	0.509	0.292
Improve	**19.0%**	**11.0%**	**13.0%**	**7.6%**	**11.7%**	**6.6%**	**9.4%**	**5.7%**	-0.5%	**6.1%**	-1.8%	**5.5%**	-1.0%	**6.3%**	-2.0%	**6.1%**
iTransformer	0.148	0.240	0.162	0.253	0.178	0.269	0.225	0.317	0.395	0.268	0.417	0.276	0.433	0.283	0.467	0.302
+ RCF	0.136	0.231	0.153	0.247	0.168	0.263	0.194	0.287	0.415	0.263	0.440	0.271	0.456	0.278	0.491	0.294
Improve	**8.1%**	**3.7%**	**5.6%**	**2.4%**	**5.8%**	**2.2%**	**13.8%**	**9.5%**	-5.1%	**1.9%**	-5.5%	**1.8%**	-5.3%	**1.8%**	-5.1%	**2.6%**

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad \mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

05 实验：模型分析——和其他STD方法的比较

- 使用纯线性模型作为主干

Setup	CLinear (RCF+Linear)		LDLinear (LD+Linear)		DLinear (MOV+Linear)		SLinear (Sparse+Linear)		Linear	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	**0.418**	**0.434**	0.427	0.439	0.425	0.437	0.424	0.436	0.427	0.439
ETTh2	**0.451**	**0.456**	0.455	0.457	0.471	0.467	0.460	0.460	0.460	0.462
ETTm1	**0.349**	**0.382**	0.365	0.387	0.367	0.390	0.362	0.383	0.362	0.384
ETTm2	**0.266**	**0.330**	0.273	0.336	0.280	0.341	0.290	0.352	0.269	0.331
Electricity	**0.157**	**0.255**	0.167	0.264	0.167	0.264	0.172	0.268	0.167	0.265
Solar-Energy	**0.220**	**0.259**	0.253	0.316	0.254	0.318	0.255	0.315	0.253	0.318
Traffic	**0.423**	**0.289**	0.434	0.296	0.434	0.296	0.435	0.292	0.434	0.296
Weather	0.245	0.300	0.244	0.297	**0.244**	**0.296**	0.246	0.298	0.245	0.297

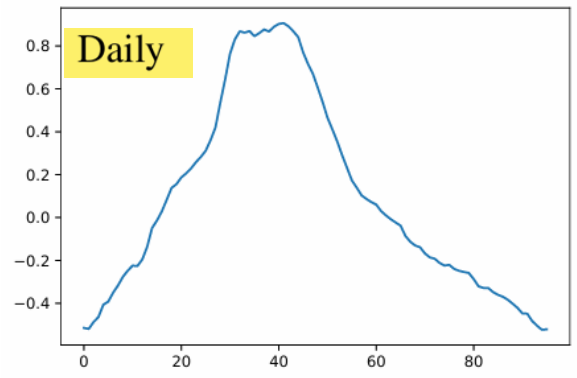
05 实验：模型分析——超参数W的影响

	Setup	RCF/W=168		RCF/W=144		RCF/W=96		RCF/W=24		W/o. RCF	
	Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Daily & Weekly	Electricity	**0.142**	**0.234**	0.196	0.275	0.196	0.274	0.195	0.274	0.197	0.274
Daily & Weekly	Trafiic	**0.480**	**0.314**	0.617	0.386	0.617	0.385	0.618	0.385	0.645	0.383
10 mins Daily	Solar-Energy	0.289	0.376	**0.208**	**0.256**	0.276	0.365	0.287	0.375	0.286	0.375
15 mins Daily	ETTm1	0.350	0.369	0.340	0.366	**0.325**	**0.363**	0.348	0.367	0.351	0.372
Daily	ETTh1	0.395	0.402	0.384	0.395	0.383	0.393	**0.377**	**0.391**	0.384	0.392

- **W正确设置**：RCF发挥重要作用，产生较大的性能差距
- **W错误设置**：RCF 也不会带来显着的负面影响

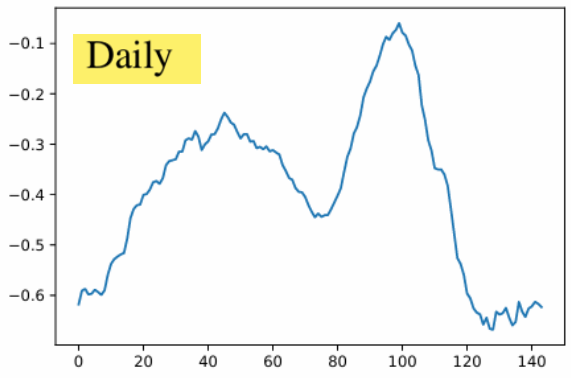
05 实验：模型分析——学习周期模式的可视化

RCF/W=96



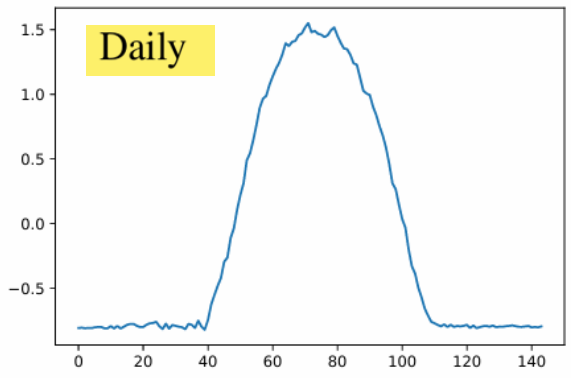
(a) ETTm1, 7th

RCF/W=144



(b) Weather, 7th

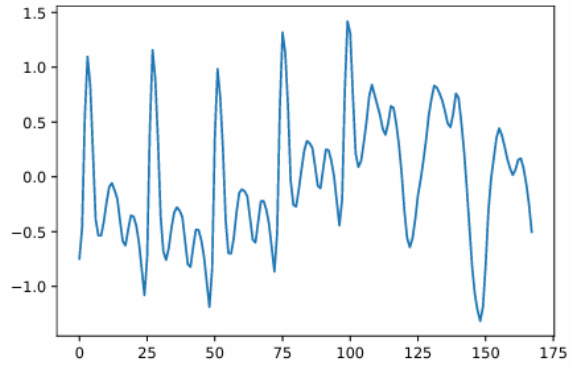
RCF/W=144



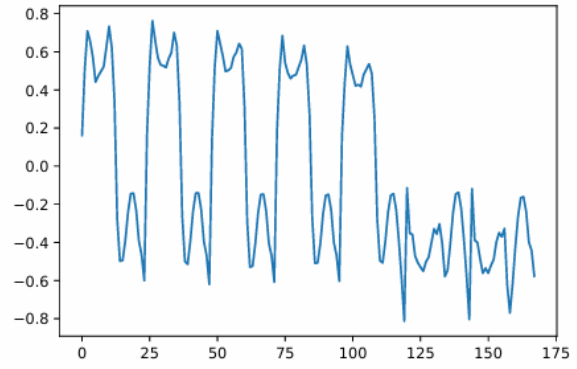
(c) Solar-Energy, 137th

Daily & Weekly

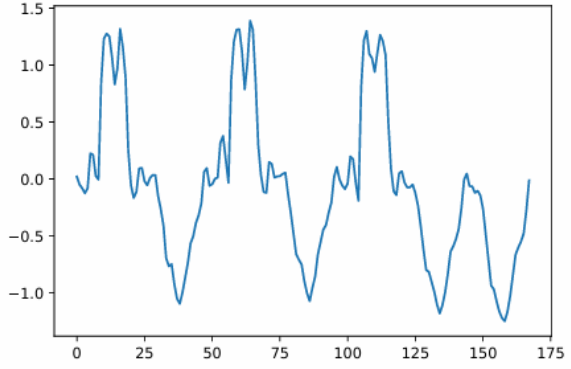
RCF/W=168



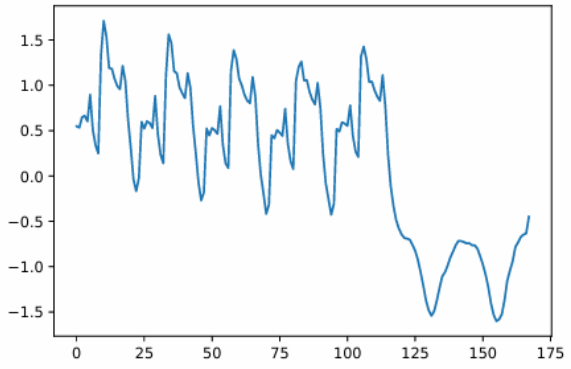
(d) Traffic, 607th



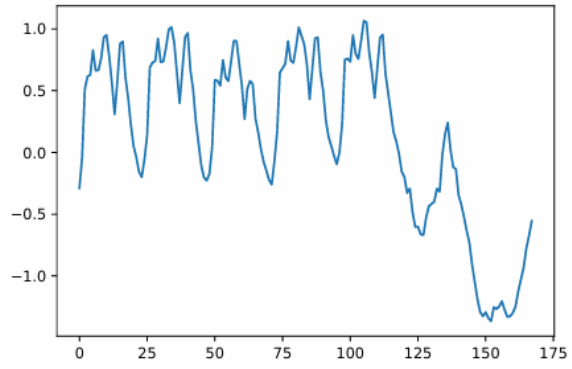
(e) Electricity, 311st



(f) Electricity, 318th



(g) Electricity, 320th

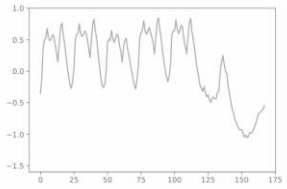


(h) Electricity, 321st

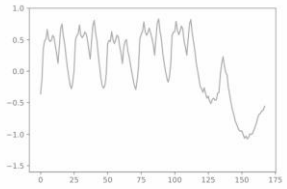
RCF/W=168

Daily & Weekly

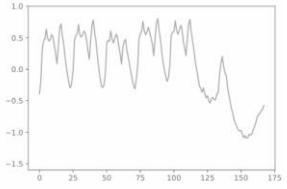
05 实验：模型分析——学习周期模式的可视化（不同配置下）



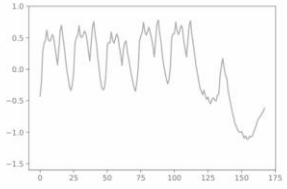
(a) Horizon-96



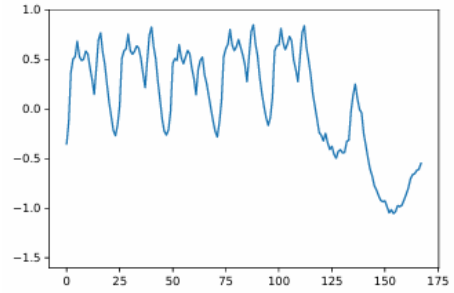
(b) Horizon-192



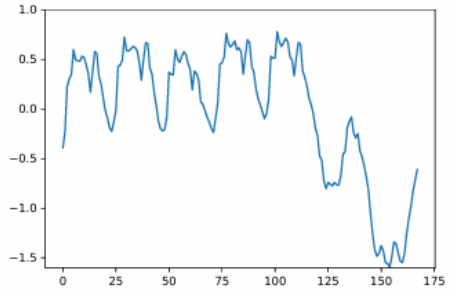
(c) Horizon-336



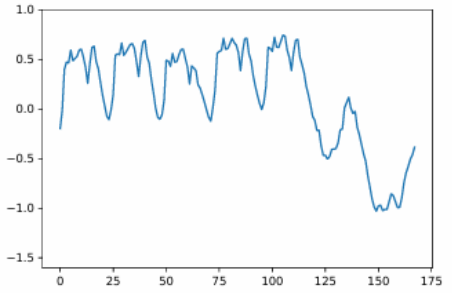
(d) Horizon-720



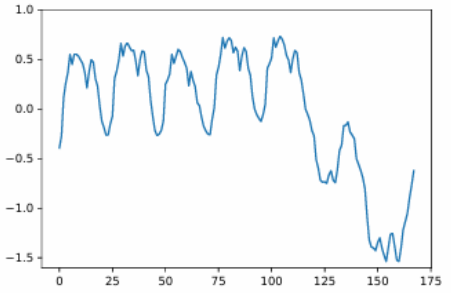
(e) Lookback-96



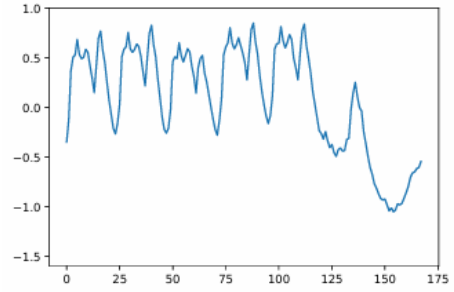
(f) Lookback-192



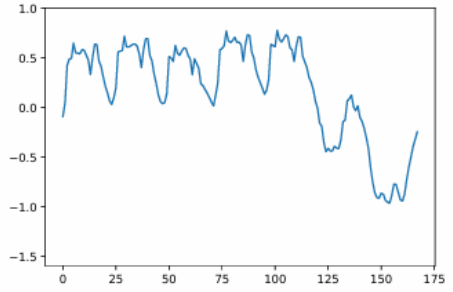
(g) Lookback-336



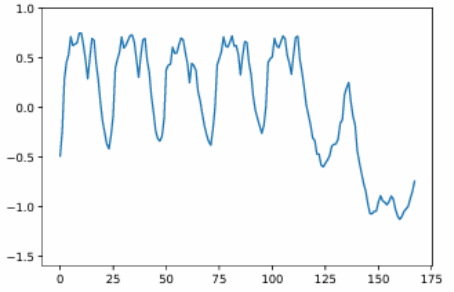
(h) Lookback-720



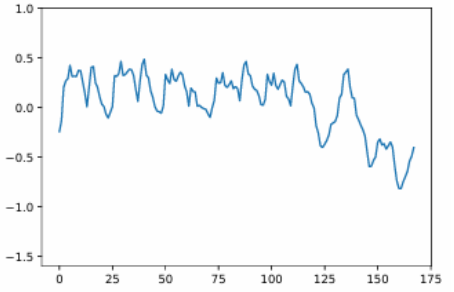
(i) Backbone-Linear



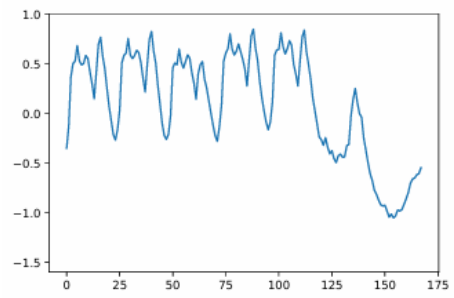
(j) Backbone-DLinear



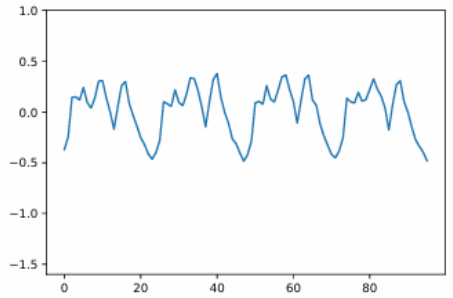
(k) Backbone-PatchTST



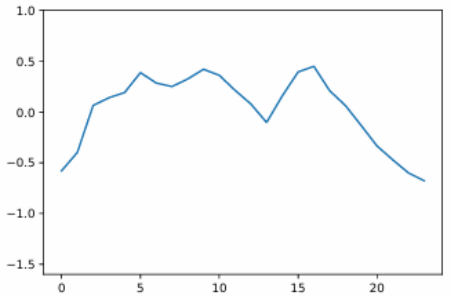
(l) Backbone-iTransformer



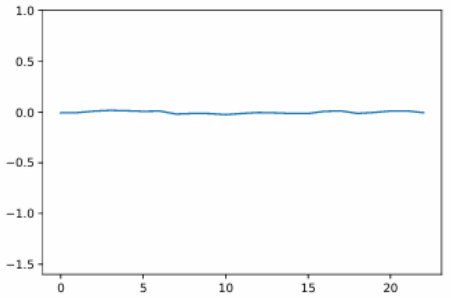
(m) W-168



(n) W-96



(o) W-24



(p) W-23

06 总结与讨论

- SparseTSF
  - 超长周期问题
  - 多周期问题
- **CycleNet/RCF**
  - **只能学习固定长度的周期**：不适合周期长度随时间变化的数据集，如心电图(ECG)数据
  - **不同通道表现出不同长度的循环**
  - **数据集存在显著的异常值**
  - **更长的依赖关系（例如年度周期）具有挑战**

谢谢观看

MANY THANKS !

24.10.31