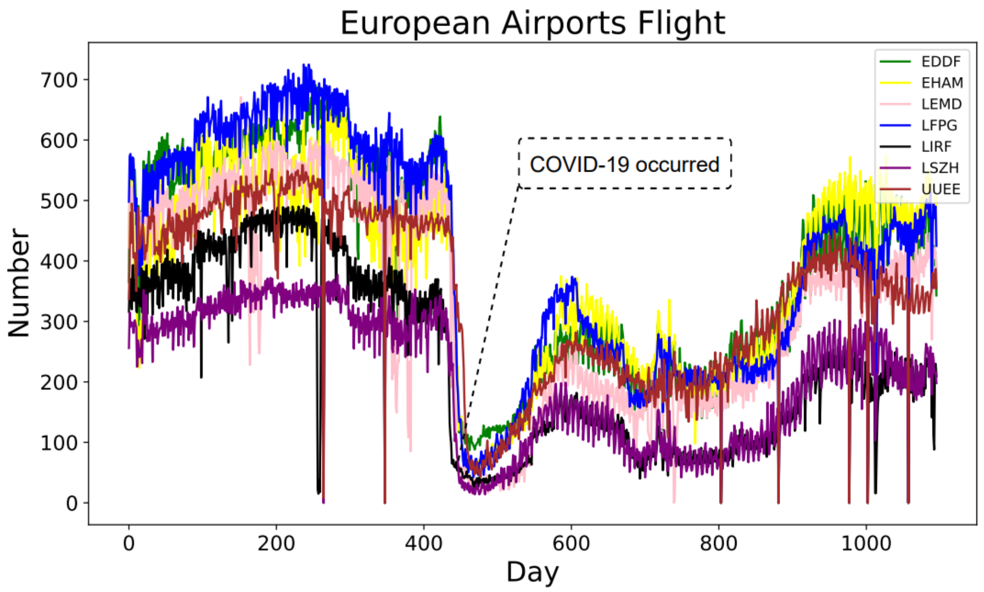
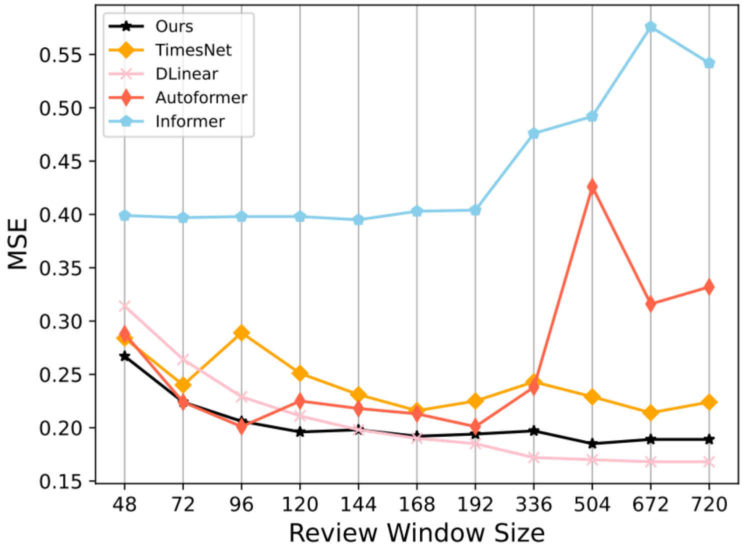
2024 AAAI CCF-A

MSGNet

Learning Multi-Scale Inter-Series Correlations for Multivariate Time Series Forecasting





2024 AAAI CCF-A

MSGNet

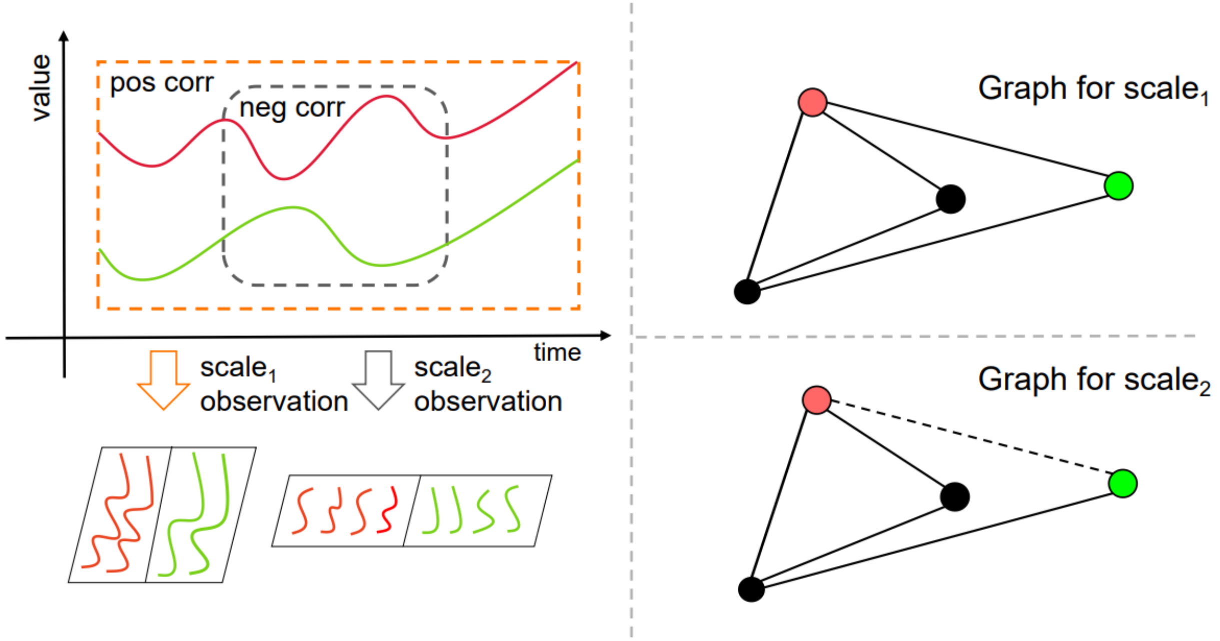
Learning Multi-Scale Inter-Series Correlations for Multivariate Time Series Forecasting

24.3. 5

Presented by Yyyq

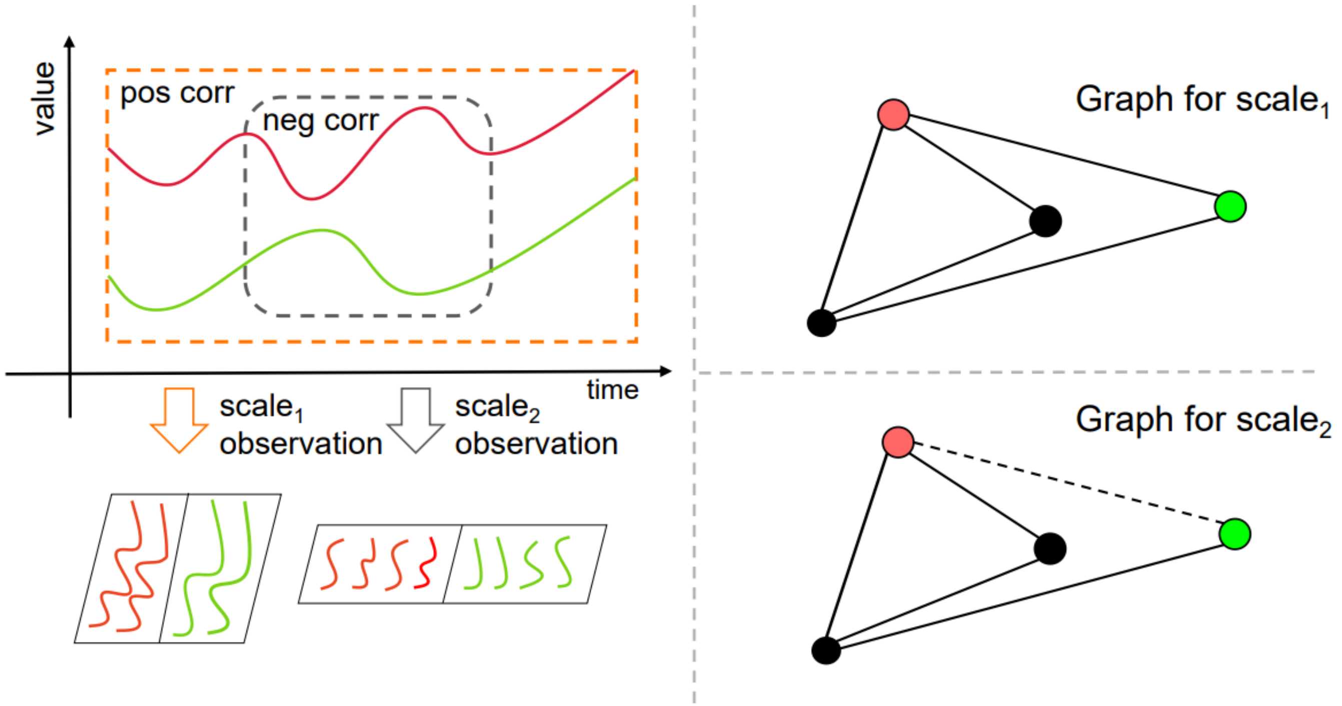
01 问题描述

- **多变量时序：**序列内(intra)和序列间(inter)相关性
- **不同时序、不同尺度之间的不同序列间关系**
  - 静态GNN不能捕获复杂变化的序列相关性
  - 动态时变图结构忽略了和时间尺度的相关性
- 应用于**分布外样本**时，
  也表现出很强的泛化能力



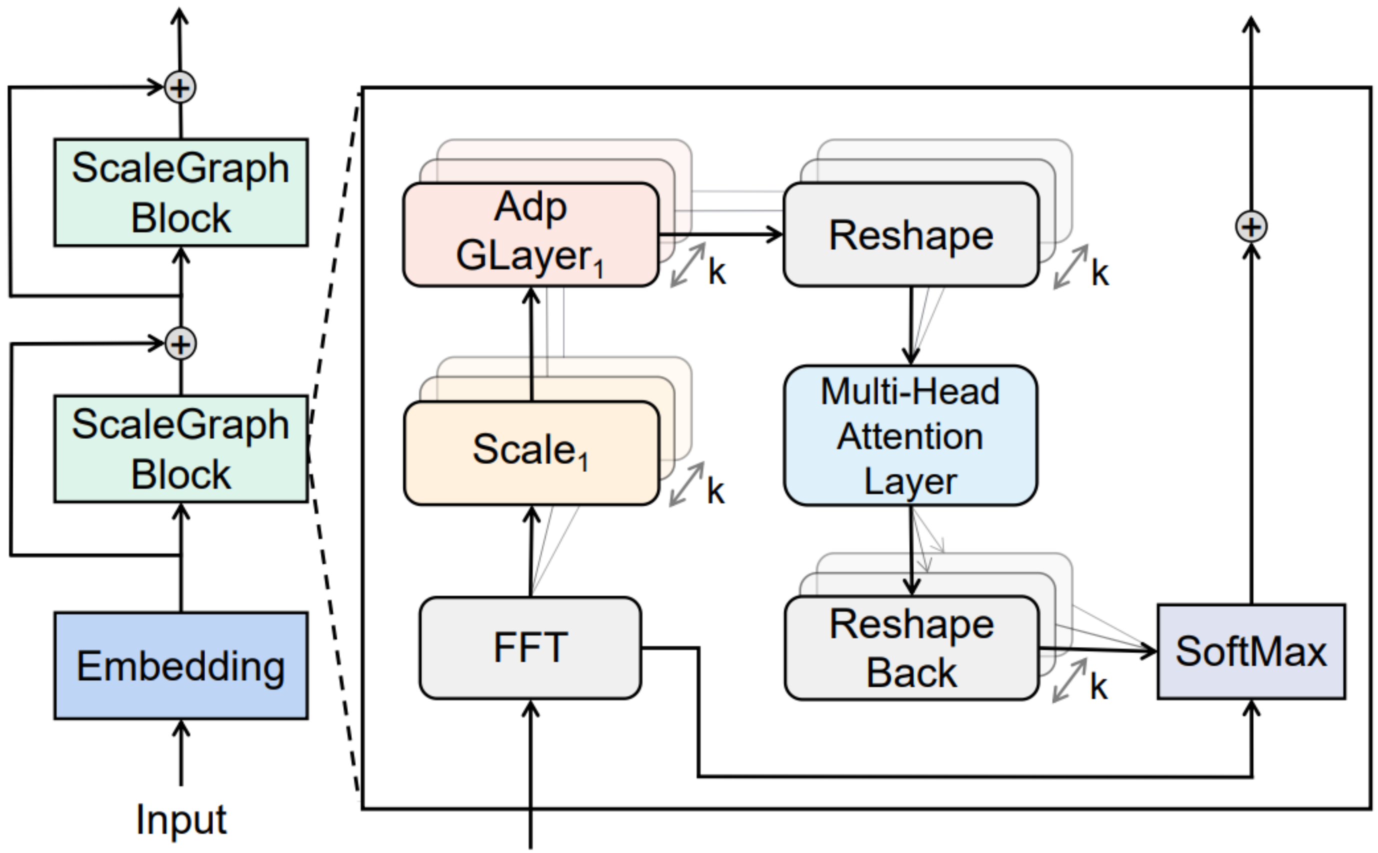
01 **问题描述**

- 不同时序、不同尺度之间的不同序列间关系
  - 频域分析：提取周期模式
  - 自注意力机制：序列内关系
  - **多尺度自适应图卷积**



02

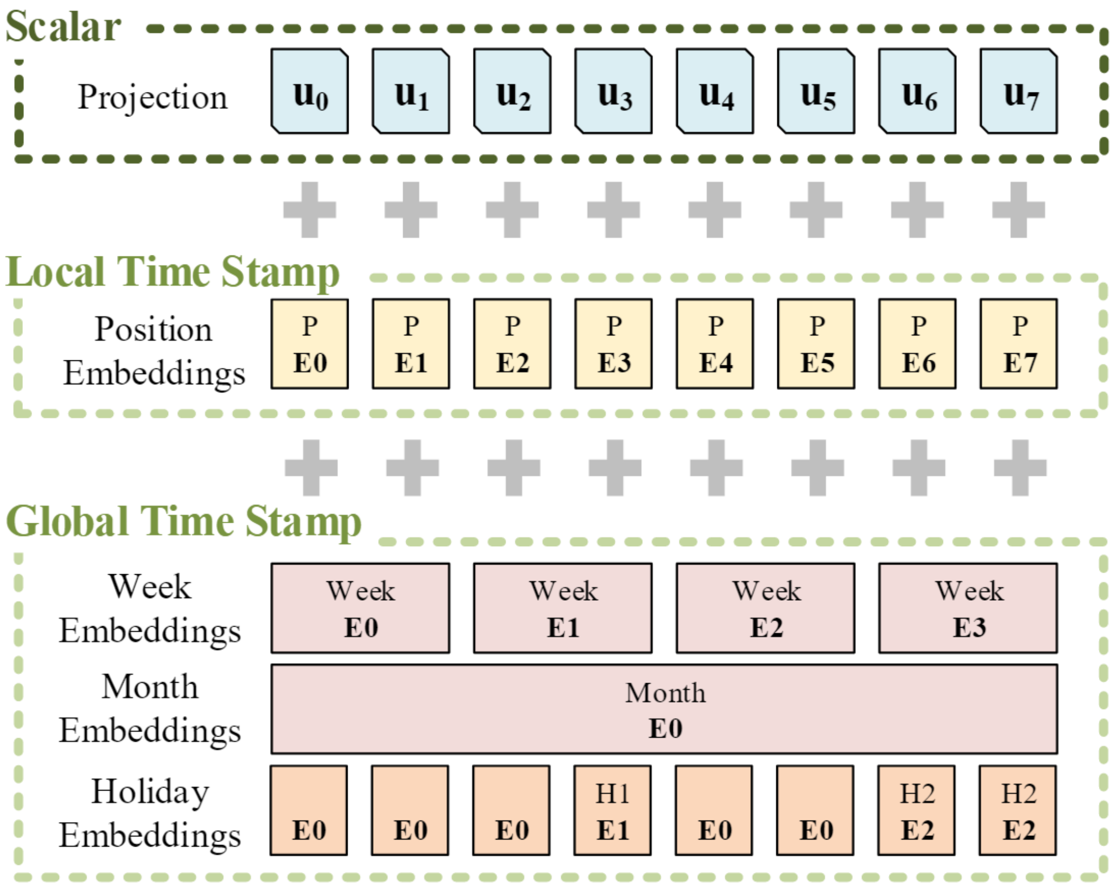
整体结构：MSGNet



03 算法描述：输入嵌入

➢ 维度嵌入（Informer）

$$\mathbf{X}_{\text{emb}} = \alpha \text{Conv1D}(\hat{\mathbf{X}}_{t-L:t}) + \mathbf{PE} + \sum_{p=1}^{P} \mathbf{SE}_p.$$



03 算法描述：尺度识别

➢ **FFT检测显著周期性（TimesNet）**

- k个不同的尺度：<u>周期性是在输入历史数据的范围内提取的</u>

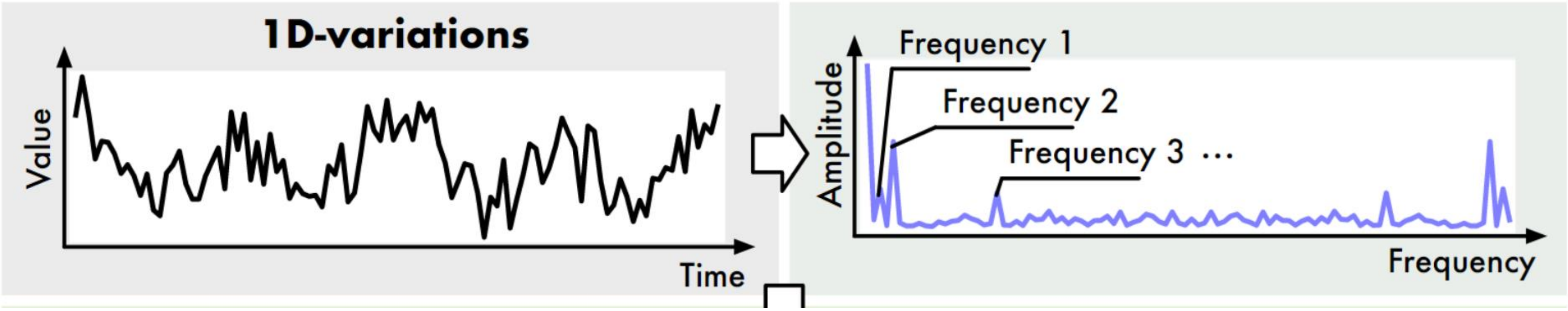
$$\mathbf{F} = \mathrm{Avg}\left(\mathrm{Amp}\left(\mathrm{FFT}(\mathbf{X}_{\mathrm{emb}})\right)\right),$$

$$f_1, \cdots, f_k = \underset{f_* \in \{1, \cdots, \frac{L}{2}\}}{\mathrm{argTopk}} (\mathbf{F}), s_i = \frac{L}{f_i}.$$

$s_i \times f_i = L$

（L=96，根据数据集，步数大小：10/15分钟，小时，日，

那么提取的周期范围就是：16小时内的周期性、4天内的周期性、3个月内的周期性）



03 算法描述：多尺度自适应图卷积

- 针对每个尺度，捕获不同的序列间关系

$$\mathcal{X}^i = \mathrm{Reshape}_{s_i,f_i}(\mathrm{Padding}(\mathbf{X}_{\mathrm{in}})), \quad i \in \{1,\ldots,k\}$$

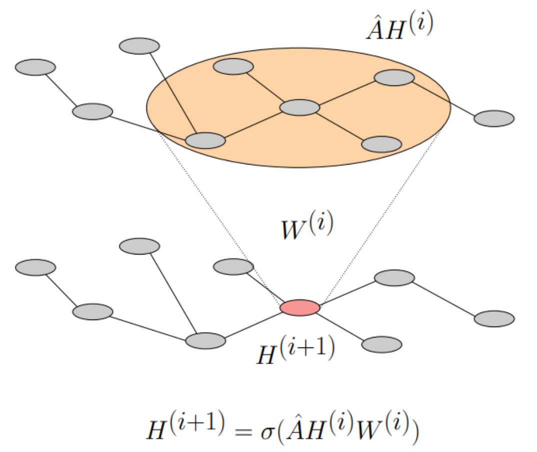
$$\mathcal{H}^i = \mathbf{W}^i \mathcal{X}^i. \quad \mathcal{H}^i \in \mathbb{R}^{N \times s_i \times f_i}$$

- **Mixhop图卷积：** 重复混合不同距离的邻居的特征表示

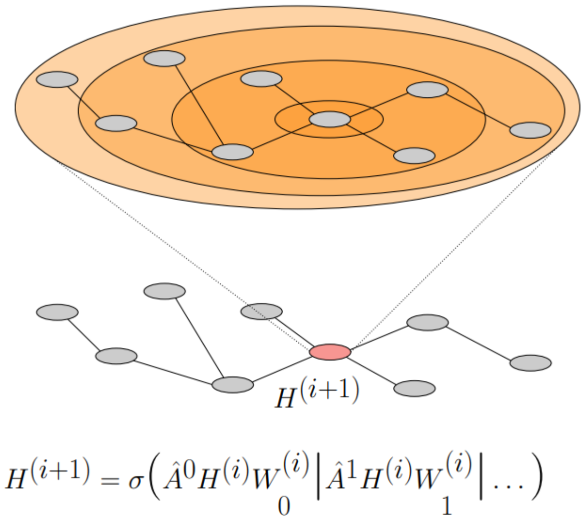
$$\mathbf{A}^i = \mathrm{SoftMax}(\mathrm{ReLu}(\mathbf{E}_1^i(\mathbf{E}_2^i)^T)).$$

```
self.nodevec1 = nn.Parameter(torch.randn(c_out, node_dim), requires_grad=True)
self.nodevec2 = nn.Parameter(torch.randn(node_dim, c_out), requires_grad=True)
```

$$\mathcal{H}_{\mathrm{out}}^i = \sigma\left(\Big\Vert_{j\in\mathcal{P}} (\mathbf{A}^i)^j \mathcal{H}^i\right),$$



(a) Traditional graph convolution.



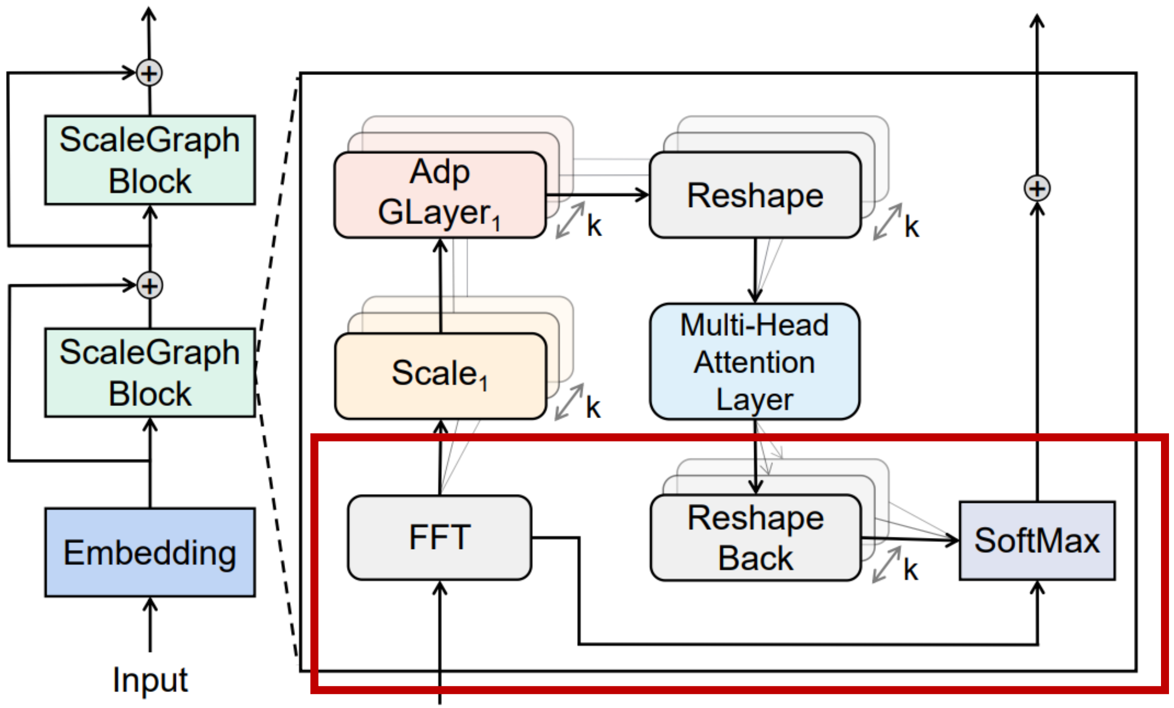
(b) Our mixed feature model.

03 算法描述：多头注意和尺度融合

- **多头注意力MHA：**捕获序列内相关性
- **尺度融合：**按振幅，根据各自的幅度强调来自不同尺度的信息

$$\hat{a}_1, \cdots, \hat{a}_k = \mathrm{SoftMax}(\mathbf{F}_{f_1}, \cdots, \mathbf{F}_{f_k}),$$

$$\hat{\mathbf{X}}_{\mathrm{out}} = \sum_{i=1}^{k} \hat{a}_i \hat{\mathbf{X}}^i_{\mathrm{out}}.$$



04 实验：数据集

Datasets	Nodes	Input Length	Output Length	Train / test / valid Size	Frequency
Flight	7	96	{96, 192, 336, 720}	(18317, 2633, 5261)	Hourly
Weather	21	96	{96, 192, 336, 720}	(36792, 5271, 10540)	10 minutes
ETTm1	7	96	{96, 192, 336, 720}	(34465, 11521, 11521)	15 minutes
ETTm2	7	96	{96, 192, 336, 720}	(34465, 11521, 11521)	15 minutes
ETTh1	7	96	{96, 192, 336, 720}	(8545, 2881, 2881)	Hourly
ETTh2	7	96	{96, 192, 336, 720}	(8545, 2881, 2881)	Hourly
Electricity	321	96	{96, 192, 336, 720}	(18317, 2633, 5261)	Hourly
Exchange	8	96	{96, 192, 336, 720}	(5120, 665, 1422)	Daily

Flight： 飞行信息（出发和目的地机场、出发时间、着陆时间等），**包括与 COVID-19 特别相关的飞行数据（2020 年之后）**

04 实验1：对比实验

Models		Ours		TimesNet		DLinear		NLinear		MTGnn		Autoformer		Informer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Flight	96	**0.183**	**0.301**	0.237	0.350	0.221	0.337	0.270	0.379	0.196	0.316	0.204	0.319	0.333	0.405
	192	**0.189**	**0.306**	0.224	0.337	0.220	0.336	0.272	0.380	0.272	0.379	0.200	0.314	0.358	0.421
	336	0.206	0.320	0.289	0.394	0.229	0.342	0.280	0.385	0.260	0.369	**0.201**	**0.318**	0.398	0.446
	720	**0.253**	**0.358**	0.310	0.408	0.263	0.366	0.316	0.409	0.390	0.449	0.345	0.426	0.476	0.484
Weather	96	**0.163**	**0.212**	0.172	0.220	0.196	0.255	0.196	0.235	0.171	0.231	0.266	0.336	0.300	0.384
	192	**0.212**	**0.254**	0.219	0.261	0.237	0.296	0.241	0.271	0.215	0.274	0.307	0.367	0.598	0.544
	336	0.272	**0.299**	0.280	0.306	0.283	0.335	0.293	0.308	**0.266**	0.313	0.359	0.395	0.578	0.523
	720	0.350	**0.348**	0.365	0.359	0.345	0.381	0.366	0.356	**0.344**	0.375	0.419	0.428	1.059	0.741
ETTm1	96	**0.319**	**0.366**	0.338	0.375	0.345	0.372	0.350	0.371	0.381	0.415	0.505	0.475	0.672	0.571
	192	0.376	0.397	**0.374**	**0.387**	0.380	0.389	0.389	0.390	0.442	0.451	0.553	0.496	0.795	0.669
	336	0.417	0.422	**0.410**	**0.411**	0.413	0.413	0.422	0.412	0.475	0.475	0.621	0.537	1.212	0.871
	720	0.481	0.458	0.478	0.450	**0.474**	0.453	0.482	**0.446**	0.531	0.507	0.671	0.561	1.166	0.823
ETTm2	96	**0.177**	**0.262**	0.187	0.267	0.193	0.292	0.188	0.272	0.240	0.343	0.255	0.339	0.365	0.453
	192	**0.247**	**0.307**	0.249	0.309	0.284	0.362	0.253	0.312	0.398	0.454	0.281	0.340	0.533	0.563
	336	**0.312**	**0.346**	0.321	0.351	0.369	0.427	0.314	0.350	0.568	0.555	0.339	0.372	1.363	0.887
	720	0.414	**0.403**	**0.408**	0.403	0.554	0.522	0.414	0.405	1.072	0.767	0.433	0.432	3.379	1.338
ETTh1	96	0.390	0.411	**0.384**	0.402	0.386	**0.400**	0.393	0.400	0.440	0.450	0.449	0.459	0.865	0.713
	192	0.442	0.442	**0.436**	**0.429**	0.437	0.432	0.449	0.433	0.449	0.433	0.500	0.482	1.008	0.792
	336	**0.480**	0.468	0.491	0.469	0.481	0.459	0.485	**0.448**	0.598	0.554	0.521	0.496	1.107	0.809
	720	0.494	0.488	0.521	0.500	0.519	0.516	**0.469**	**0.461**	0.685	0.620	0.514	0.512	1.181	0.865
ETTh2	96	0.328	0.371	0.340	0.374	0.333	0.387	**0.322**	**0.369**	0.496	0.509	0.346	0.388	3.755	1.525
	192	**0.402**	**0.414**	0.402	0.414	0.477	0.476	0.410	0.419	0.716	0.616	0.456	0.452	5.602	1.931
	336	**0.435**	**0.443**	0.452	0.452	0.594	0.541	0.444	0.449	0.718	0.614	0.482	0.486	4.721	1.835
	720	**0.417**	**0.441**	0.462	0.468	0.831	0.657	0.450	0.462	1.161	0.791	0.515	0.511	3.647	1.625
Electricity	96	**0.165**	0.274	0.168	**0.272**	0.197	0.282	0.198	0.274	0.211	0.305	0.201	0.317	0.274	0.368
	192	**0.184**	0.292	0.184	0.289	0.196	0.285	0.197	**0.277**	0.225	0.319	0.222	0.334	0.296	0.386
	336	**0.195**	0.302	0.198	0.300	0.209	0.301	0.211	**0.292**	0.247	0.340	0.231	0.338	0.300	0.394
	720	0.231	0.332	**0.220**	**0.320**	0.245	0.333	0.253	0.325	0.287	0.373	0.254	0.361	0.373	0.439
Exchange	96	0.102	0.230	0.107	0.234	0.088	0.218	**0.088**	**0.205**	0.267	0.378	0.197	0.323	0.847	0.752
	192	0.195	0.317	0.226	0.344	**0.176**	0.315	0.177	**0.297**	0.590	0.578	0.300	0.369	1.204	0.895
	336	0.359	0.436	0.367	0.448	**0.313**	0.427	0.323	**0.409**	0.939	0.749	0.509	0.524	1.672	1.036
	720	0.940	0.738	0.964	0.746	**0.839**	**0.695**	0.923	0.725	1.107	0.834	1.447	0.941	2.478	1.310
Avg Rank		**1.813**		2.750		3.563		2.813		5.313		4.750		7.000	

04

实验1：对比实验——Flight 数据集

Models		Ours		TimesNet		DLinear		NLinear		MTGnn		Autoformer		Informer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Flight	96	**0.183**	**0.301**	0.237	0.350	0.221	0.337	0.270	0.379	0.196	0.316	0.204	0.319	0.333	0.405
	192	**0.189**	**0.306**	0.224	0.337	0.220	0.336	0.272	0.380	0.272	0.379	0.200	0.314	0.358	0.421
	336	0.206	0.320	0.289	0.394	0.229	0.342	0.280	0.385	0.260	0.369	**0.201**	**0.318**	0.398	0.446
	720	**0.253**	**0.358**	0.310	0.408	0.263	0.366	0.316	0.409	0.390	0.449	0.345	0.426	0.476	0.484

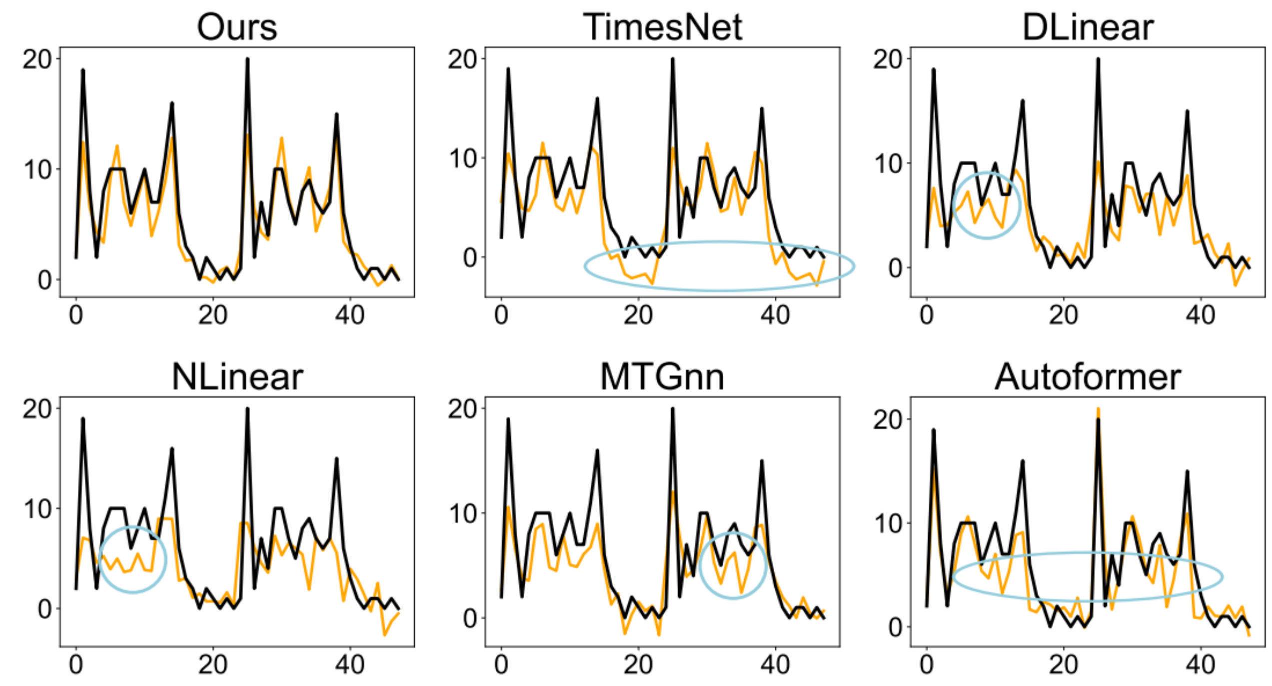


Figure 3: Visualization of Flight prediction results: black

04 实验2：学习序列间相关性的可视化

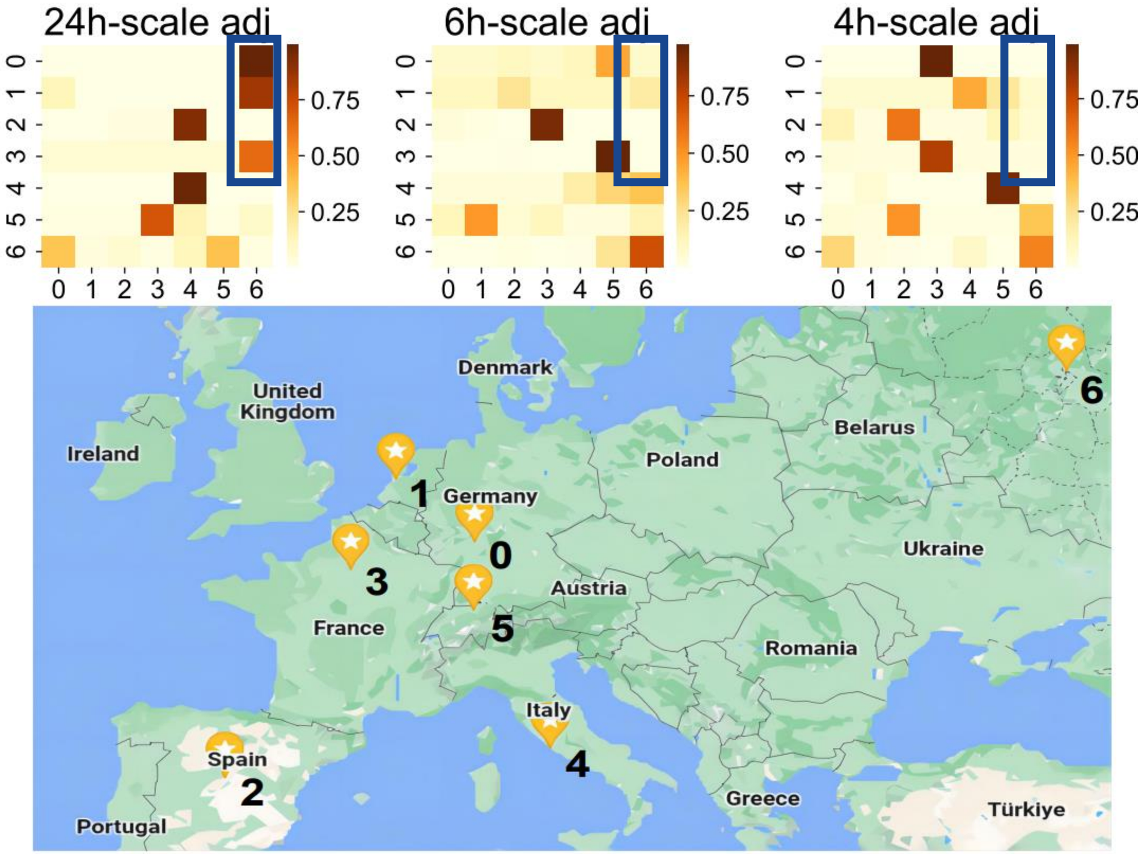


Figure 4: Learned adjacency matrices (24h, 6h, and 4h of the first layer) and airport map for Flight dataset.

04

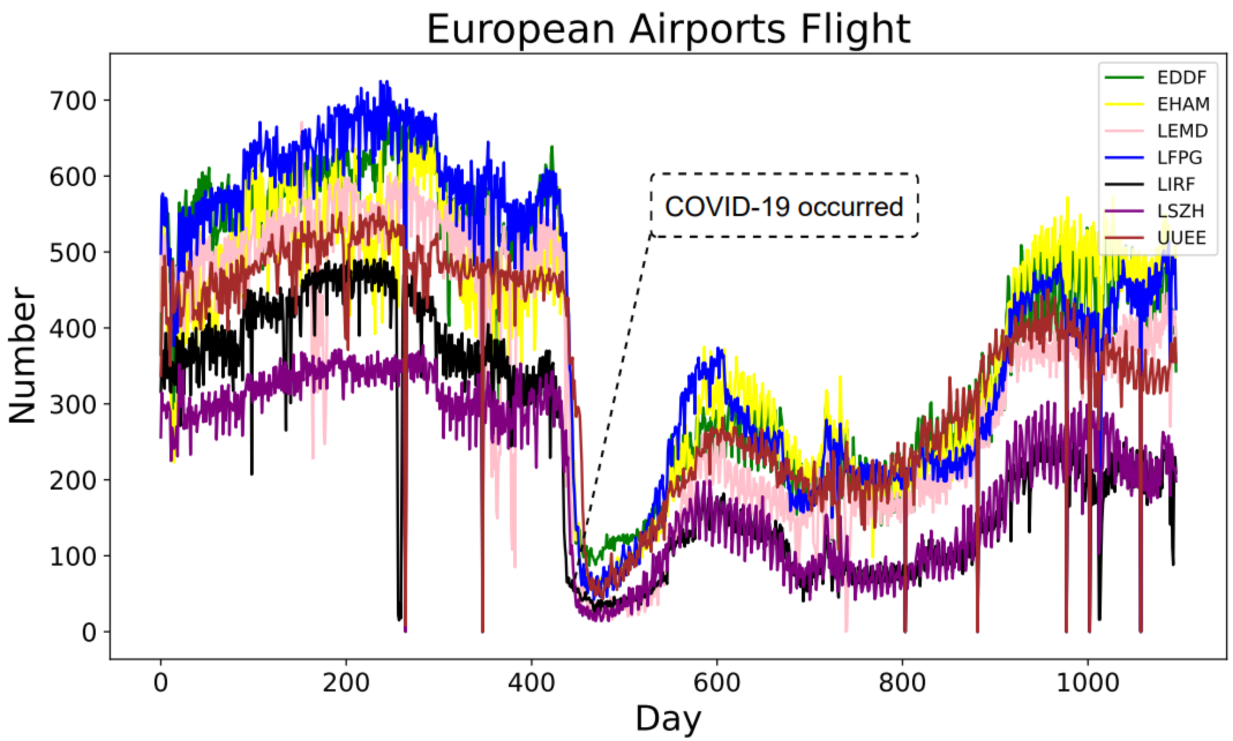
实验3：消融实验

Dataset	Flight		Weather		ETTm2	
Metric	MSE	MAE	MSE	MAE	MSE	MAE
MSGNet	**0.195**	**0.311**	**0.218**	**0.255**	**0.245**	**0.304**
w/o-AdapG	0.302	0.401	0.232	0.270	0.253	0.313
w/o-MG	0.213	0.331	0.226	0.261	0.250	0.307
w/o-A	0.198	0.314	0.224	0.259	0.247	0.306
w/o-Mix	0.202	0.318	0.224	0.260	0.247	0.304
TimesNet	0.263	0.372	0.226	0.263	0.254	0.309

04

实验4：泛化能力

Models	**Ours**		TimesNet		DLinear		NLinear		MTGnn		Autoformer	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Flight(7:1:2)	**0.208**	**0.321**	0.265	0.372	0.233	0.345	0.285	0.388	0.280	0.378	0.238	0.344
Flight(4:4:2)	**0.252**	**0.366**	0.335	0.426	0.332	0.448	0.365	0.447	0.407	0.501	0.307	0.424
Decrease(%)	**21.29**	**13.80**	26.47	14.32	42.29	29.87	28.19	15.17	45.74	32.52	29.17	23.09



04

实验5：较长输入序列下的性能

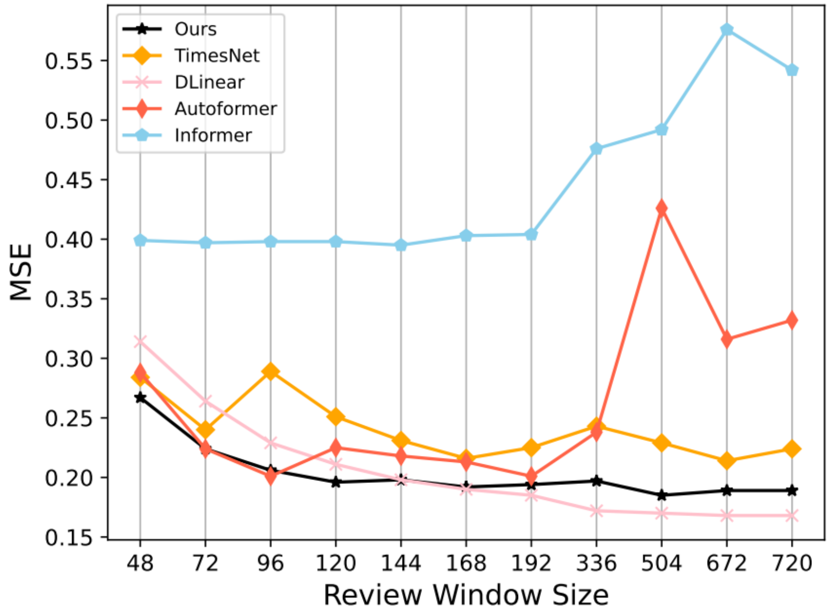


Figure 9: Flight dataset predictions for 336 time steps with different review windows. We use four other models for comparison.

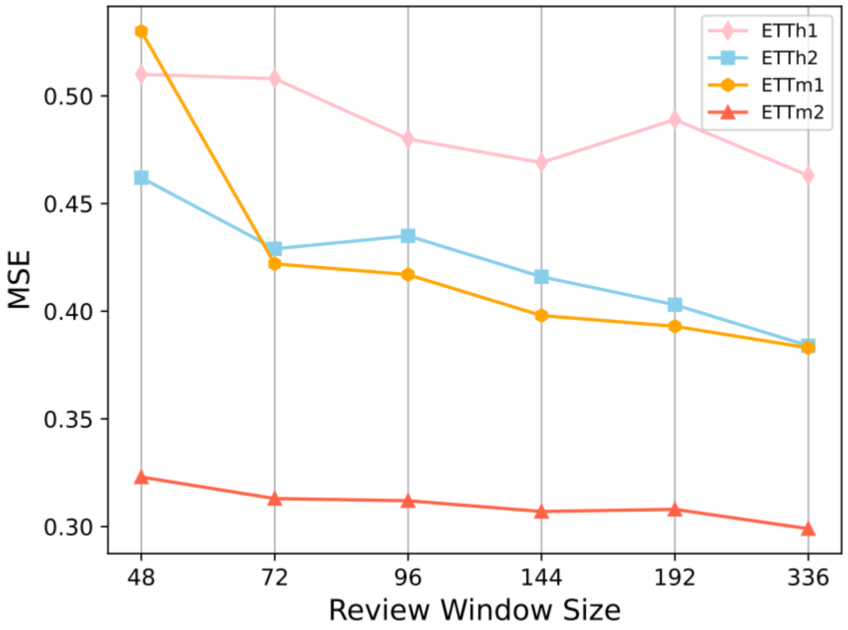


Figure 10: MSGnet's ETT dataset prediction performance for 336 time steps with different review windows.

谢谢观看

MANY THANKS !

24.3.5