NEURAL INFORMATION PROCESSING SYSTEMS

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

# Personalized Adapter for Large Meteorology Model on Devices: Towards Weather Foundation Models


**<u>Shengchao Chen</u>**[1], Guodong Long[1], Jing Jiang[1], Chengqi Zhang[2]

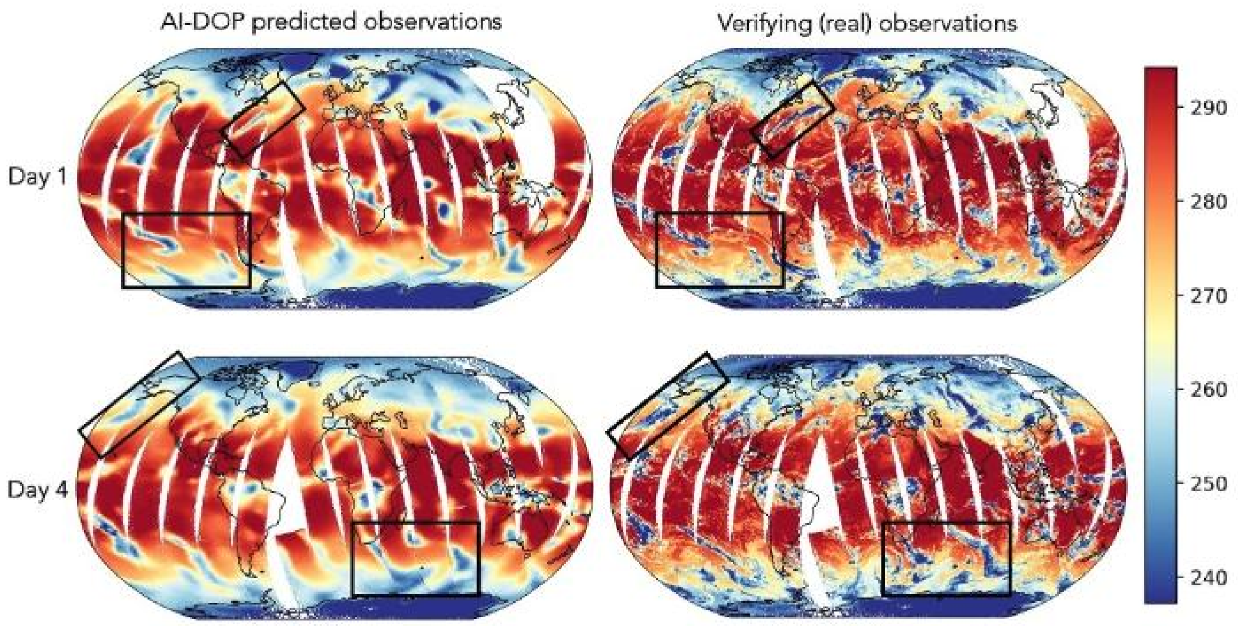
1 Australian AI Institute, FEIT, University of Technology Sydney
2 Department of Data Science and AI, The Hong Kong Polytechnic University

Accurate meteorology variables modeling is crucial in addressing the global threat of climate change.

Is it possible to *develop weather foundation models (WFMs)* that can effectively model meteorological variables across various regions?

❑ From on-device meteorology variables modeling toward WFMs:

- ❖ Meteorology varibales across regions interact significantly (e.g., spatial relations), allowing for ***mutually beneficial modeling***.
- ❖ Mainstream methods typically use ***task/data-specific deep models***, the intuition for achieving excellent performance is that fuses ***large-scale cross-region datasets to centralised traning.***




❑ **Current advancements about WFMs**

- ❖ **Traning datasets:** Large reanalysis (simulation) datasets (e.g., ERA5)
- ❖ **Models parameters:** Large-scale model parameters
- ❖ **Traning strategy:** centralized training

***Not realistic for low-resource weather device in practice***

1. Motivations

2. Contributions

3. Framework

4. Datasets

5. Results

## ❑ Generic framework:

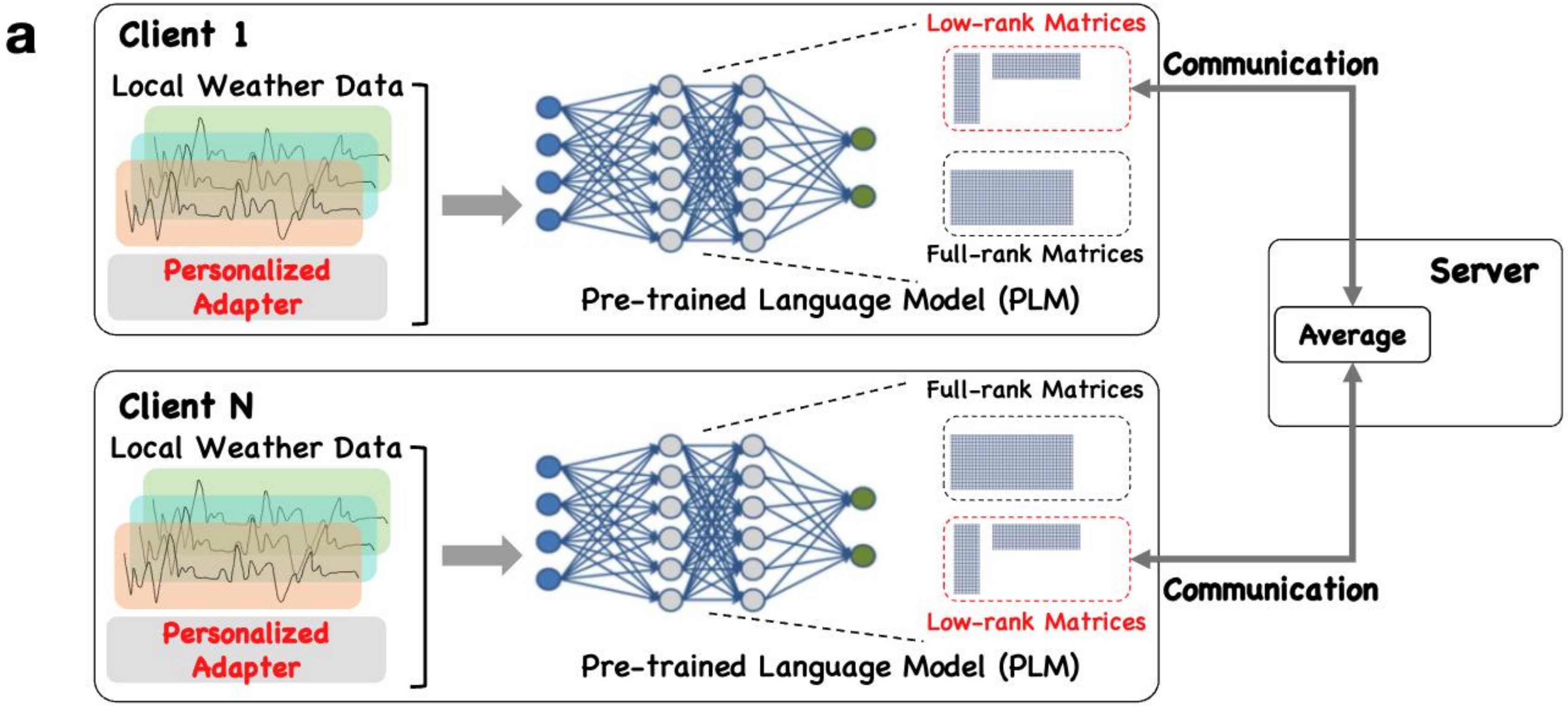
- ❖ We propose *LM-Weather*, a generic FL framework that transforms Pretrained Language Models into customized models for on-device meteorological variable modeling. LM-Weather is *parameter-, communication-, and data-efficient.*

## ❑ Adaption and communication mechansim:

- ❖ We propose *personalized adapter* for local PLMs with LoRA, to facilitate knowledge transfer from text to weather sequences. In addition, we introduce low-rank communication to reduce overhead while maintaining performance.

## ❑ Real-world datasets:

- ❖ We compile *real-world, real-observation datasets* for on-device meteorological variable modeling, which pioneers in the field of on-device meteorological variable modeling.




- **Distribued Architecture:** Using federated learning to handle the Non-IID data among devices, while ensuring privacy and computional & communicatin burden from centralized training.

- **Personliazed Adapter on PLM:** Taming the local PLM via achieving knowledge transfering between text and weather sequence.

- **Low-rank Communication:** A minimal number of parameters are both computed and communicated.

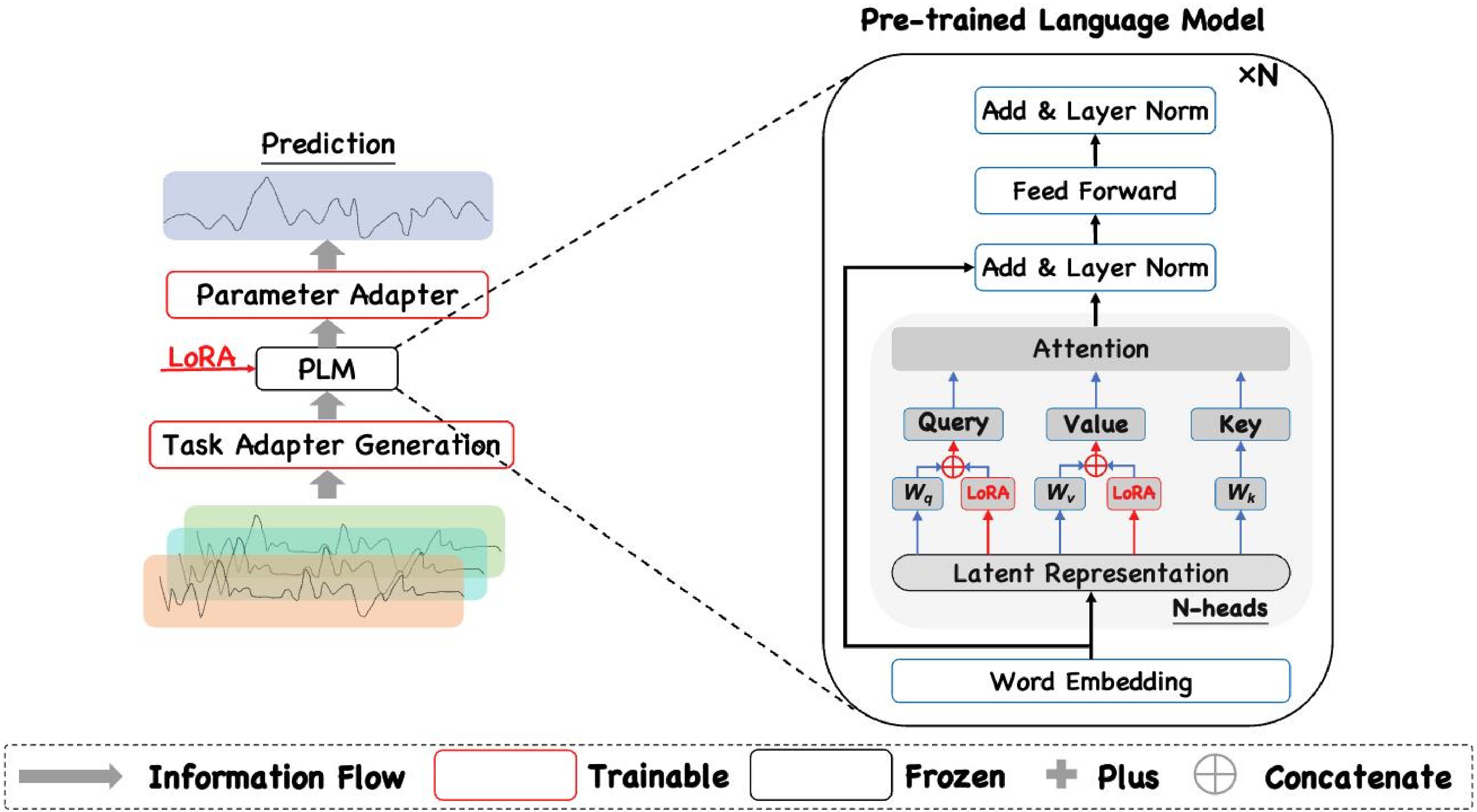
1. Motivations

2. Contributions

**3. Framework**

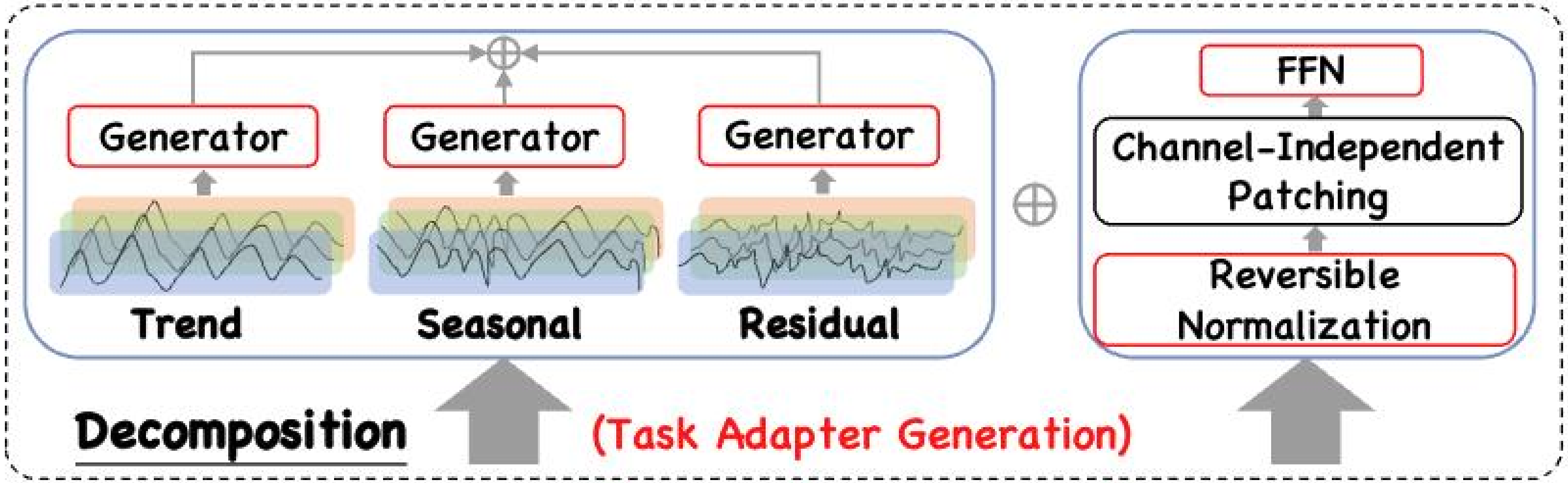
4. Datasets

5. Results




Personalized adapter consists of Task Adapter and Parameter Adapter, the low-rank adaption (LoRA) works on PLM backbone.

c

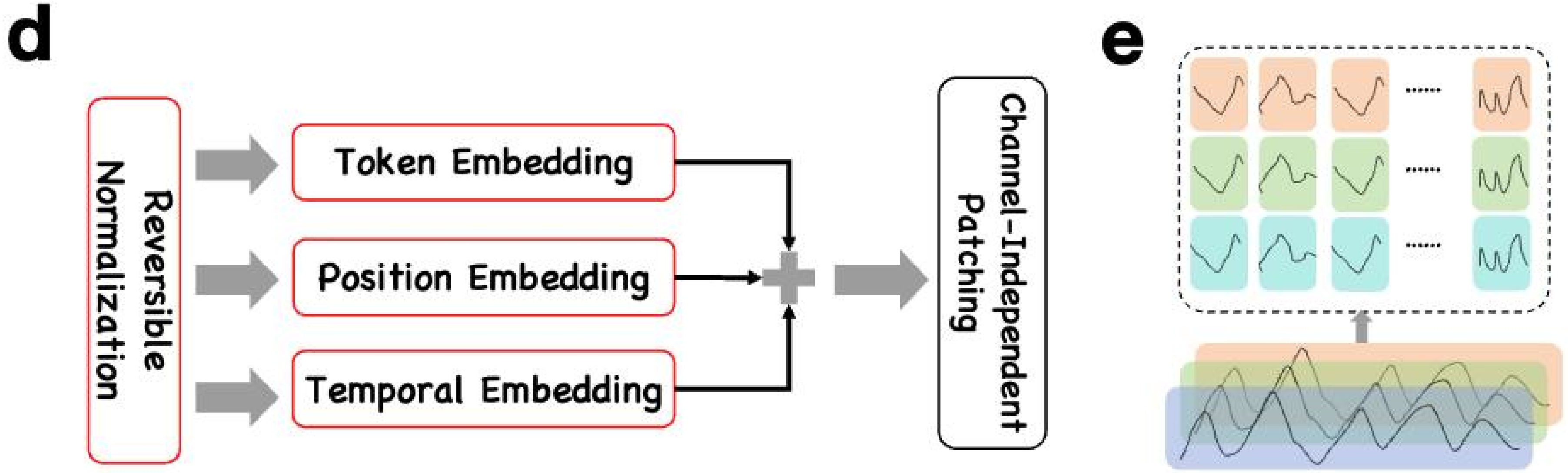



❑ **Task Adapter have three independent generators, but firstly:**

Sequence Decomposition: $\mathcal{X}_{\text{Trend}}^{k} + \mathcal{X}_{\text{Seasonal}}^{k} + \mathcal{X}_{\text{Residual}}^{k} = \texttt{Decomp}(\mathcal{X}^{k})$

Reversible Normalization: $\mathcal{X}_{\text{Trend}}' = \gamma_T \left( \mathcal{X}_{\text{Trend}} - \frac{\mathbb{E}[\mathcal{X}_{\text{Trend}}]}{\sqrt{\text{Var}\left[\mathcal{X}_{\text{Trend}}\right] + \epsilon_T}} \right) + \beta_T$

$$\boldsymbol{P}_d = \boldsymbol{P}_{\text{TO}}^{d} + \boldsymbol{P}_{\text{PO}}^{d} + \boldsymbol{P}_{\text{TE}}^{d}, \text{ where } d \in \{\text{Trend}, \text{Seasonal}, \text{Residual}\}$$




❑ **For each generator:**

❖ Token Embedding, using 1DCNN to tokenize each sample:

$$P_{\mathrm{TO}}^{k} = \textsc{Conv1D}(\mathcal{X}^{k}), \quad P_{\mathrm{TO}} = \textsc{Conv1D}(\mathcal{X})$$

❖ Position Embedding, using a trainable lookup table to mapping each point's position:

$$P_{\mathrm{PO}} = E(\textsc{Index}(\mathcal{X}))$$

❖ Temporal Embedding, encoding different time attributes to samples:

$$P_{\mathrm{TE}} = \sum_{\alpha \in \{\mathrm{mins},\mathrm{hours},\mathrm{days},\mathrm{weeks},\mathrm{months}\}} E_{\alpha}(\mathcal{X})$$

# Contains 20 variables

## ❑ Four real-world datasets:

- **ODW1 Series:** ODW1T & ODW1V
- **ODW2 Series:** ODW2T & ODW2V

**T:** has a heterogeneous time span, meaning the data collection start and end times vary by location.

**V:** extends T-version by adding variability in the observed variables.

| Abbreviation | Full name | Unit |
|---|---|---|
| ap | Air Pressure | $hpa$ |
| t | Air Temperature | $^{\circ}C$ |
| mxt | Maximum Temperature | $^{\circ}C$ |
| mnt | Minimum Temperature | $^{\circ}C$ |
| dt | Dewpoint Temperature | $^{\circ}C$ |
| rh | Relative Humidity | $\%$ |
| wvp | Water Vapor Pressure | $hpa$ |
| p1 | Precipitation in 1h | $mm$ |
| p2 | Precipitation in 3h | $mm$ |
| p3 | Precipitation in 6h | $mm$ |
| p4 | Precipitation in 12h | $mm$ |
| p5 | Precipitation in 24h | $mm$ |
| wd | Wind Direction | $^{\circ}C$ |
| ws | Wind Speed | $ms^{-1}$ |
| mwd | Maximum Wind Direction | $^{\circ}$ |
| st | Land Surface Temperature | $^{\circ}C$ |
| hv1 | Horizontal Visibility in 1 min | $m$ |
| hv2 | Horizontal Visibility in 10 min | $m$ |
| vv | Vertical Visibility | $m$ |

1. Motivations

2. Contributions

3. Framework

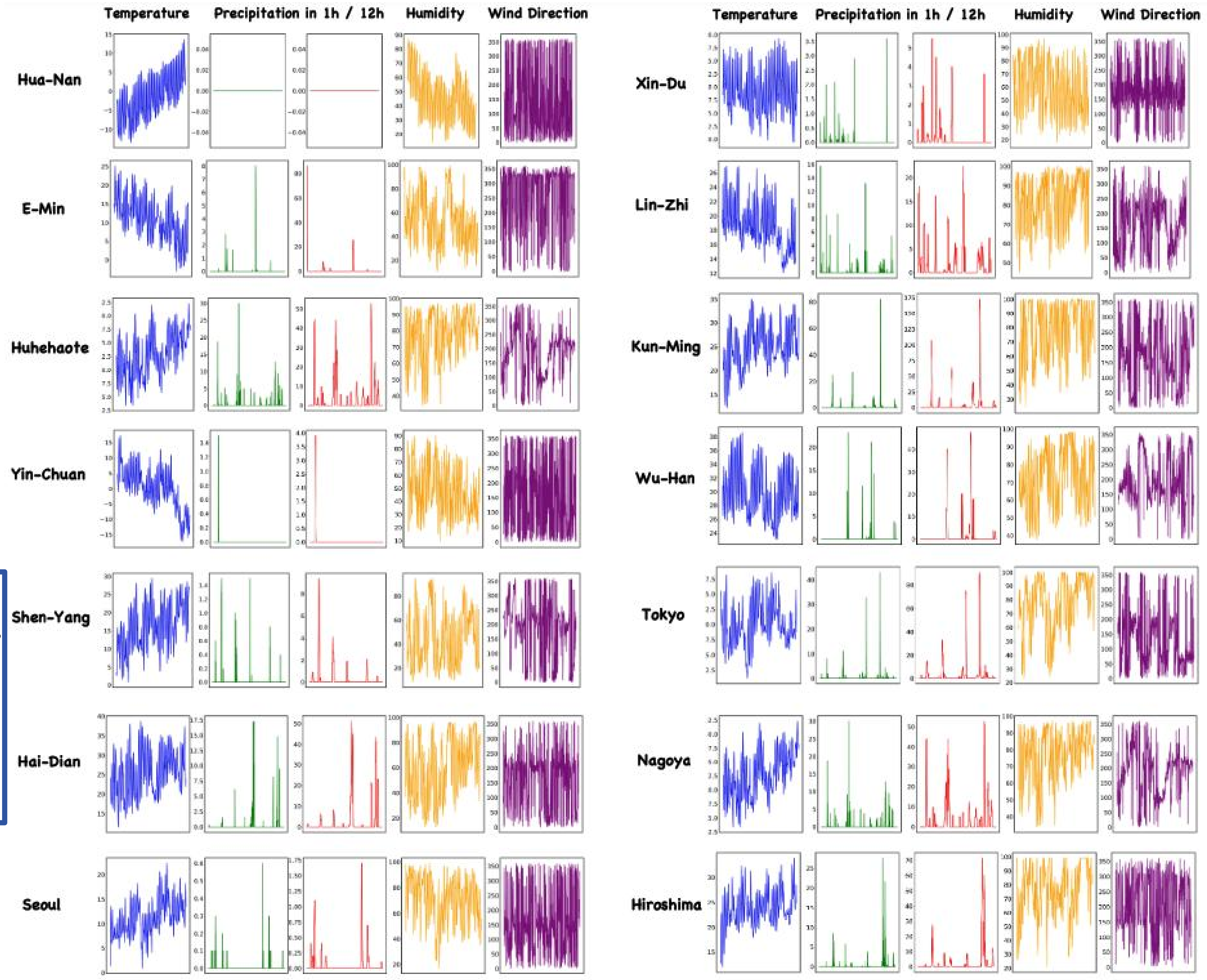
**4. Datasets**

5. Results

# ODW1 Dataset:

Partial visualization

15 stations in China, Japan, and South Korea



1. Motivations

2. Contributions

3. Framework

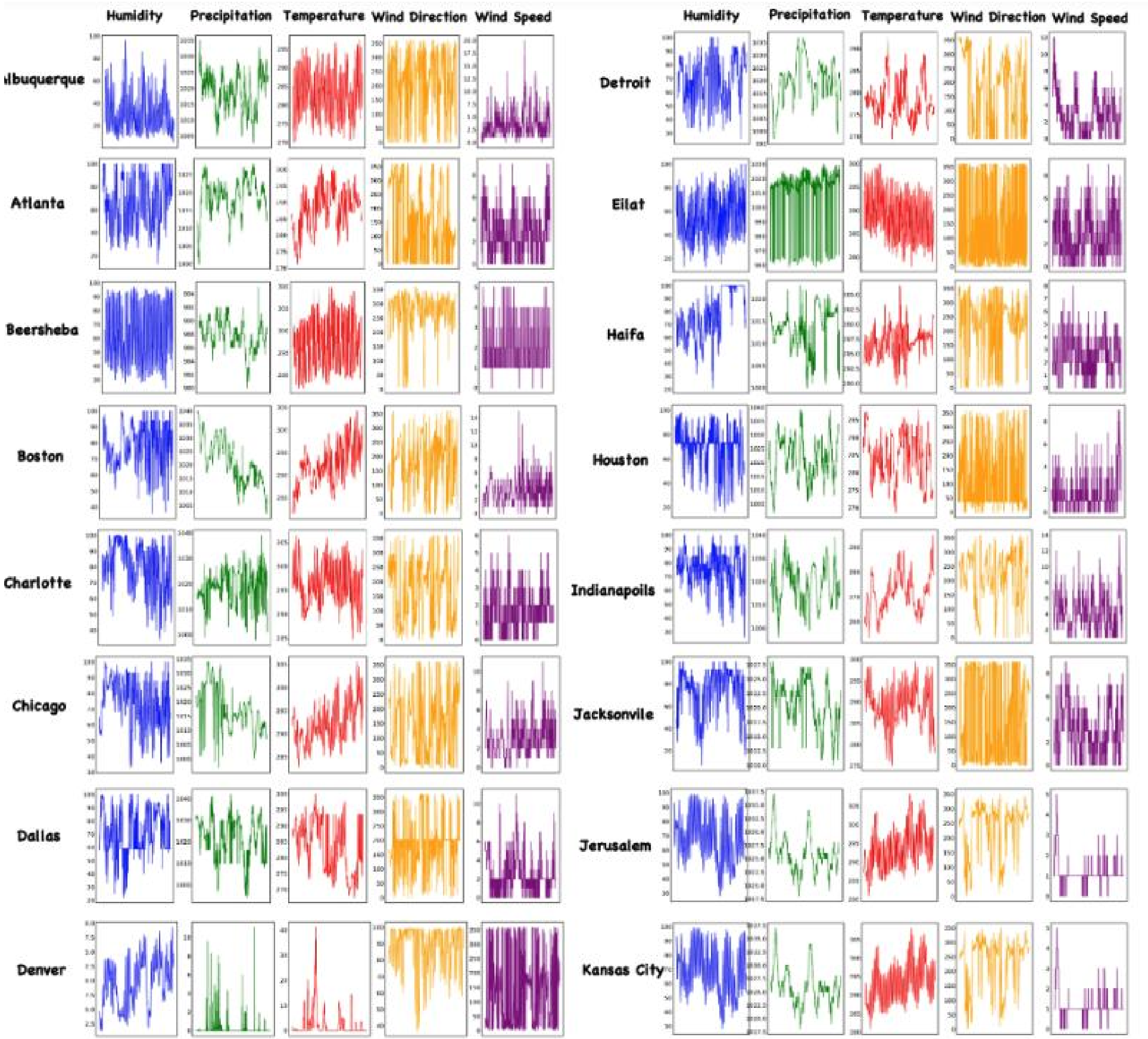
**4. Datasets**

5. Results

# ODW2 Dataset:

Partial visualization

36 stations in US, Canada, and Israel



# ❑ Main Results (Multivariate Forecasting)

| Method | | LM-WEATHER-AVE | | LM-WEATHER | | FL-GPT4TS | | FL-Reformer | | FL-Pyraformer | | FL-DLinear | | FL-PatchTST | | FL-iTransformer | | FL-LightTS | | FL-Transformer | | FL-Informer | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dataset | Length | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| ODW1T | 96 | 44.1 | 74.8 | **42.3** | **71.1** | 46.3 | 78.5 | 70.7 | 92.9 | 67.2 | 86.1 | 49.7 | 78.6 | 45.0 | 77.0 | 48.4 | 80.2 | 54.8 | 85.6 | 50.7 | 82.1 | 51.9 | 83.2 |
| | 192 | 46.3 | 77.5 | **44.4** | **73.6** | 48.6 | 81.3 | 75.1 | 98.3 | 70.0 | 90.9 | 52.3 | 81.8 | 47.3 | 79.8 | 51.8 | 84.3 | 59.5 | 90.6 | 52.1 | 84.0 | 52.9 | 84.6 |
| | 336 | 47.9 | 79.3 | **45.8** | **75.2** | 50.3 | 83.2 | 79.8 | 100.5 | 74.1 | 92.8 | 53.9 | 83.7 | 49.0 | 81.7 | 54.5 | 87.3 | 64.0 | 94.6 | 52.9 | 85.2 | 53.5 | 85.6 |
| | 720 | 51.8 | 83.0 | **49.2** | **78.5** | 54.4 | 87.2 | 87.1 | 102.9 | 80.5 | 95.2 | 57.2 | 87.3 | 53.3 | 85.6 | 60.1 | 93.1 | 72.4 | 102.7 | 55.4 | 87.6 | 55.3 | 87.4 |
| | Avg. | 47.5 | 78.7 | **45.4** | **74.6** | 49.9 | 82.5 | 78.2 | 98.7 | 73.0 | 91.3 | 53.3 | 82.8 | 48.6 | 81.0 | 53.7 | 63.7 | 62.7 | 93.4 | 52.8 | 84.7 | 53.4 | 85.2 |
| ODW1V | 96 | 42.7 | 69.5 | **42.3** | 69.6 | 44.0 | 71.4 | 42.9 | 67.8 | 57.7 | **67.2** | 46.4 | 73.3 | 44.3 | 69.6 | 56.8 | 76.8 | 48.0 | 75.1 | 67.0 | 89.4 | 59.0 | 80.3 |
| | 192 | 45.5 | 72.6 | **44.4** | 71.7 | 47.0 | 75.8 | 48.4 | 75.4 | 59.2 | **69.4** | 47.9 | 75.1 | 46.8 | 72.1 | 55.0 | 75.0 | 49.1 | 79.2 | 69.9 | 93.0 | 61.2 | 82.8 |
| | 336 | 47.2 | 74.3 | **46.0** | **72.4** | 48.8 | 77.7 | 51.0 | 77.0 | 63.4 | 73.3 | 49.1 | 76.9 | 48.5 | 74.8 | 62.4 | 83.7 | 50.8 | 77.9 | 71.4 | 94.8 | 63.7 | 85.8 |
| | 720 | 51.2 | 78.2 | **49.7** | **74.0** | 53.3 | 81.7 | 54.5 | 82.3 | 67.3 | 76.1 | 52.5 | 80.3 | 54.3 | 79.1 | 72.1 | 96.2 | 54.7 | 82.7 | 76.2 | 87.3 | 68.4 | 91.8 |
| | Avg. | 46.6 | 73.6 | **45.6** | 71.9 | 48.3 | 76.7 | 49.2 | 75.6 | 61.9 | **71.5** | 49.0 | 76.4 | 48.5 | 73.9 | 58.1 | 85.0 | 50.7 | 78.7 | 71.1 | 91.1 | 63.1 | 85.2 |
| ODW2T | 96 | 64.3 | 88.2 | 62.8 | 85.5 | 66.8 | 91.7 | 100.3 | 126.3 | 95.0 | 120.3 | 67.9 | 84.7 | 70.2 | 88.1 | 68.6 | 86.5 | 68.4 | 85.4 | 85.0 | 103.0 | 84.7 | 102.7 |
| | 192 | 67.7 | 91.5 | 66.2 | 89.1 | 71.1 | 96.1 | 102.1 | 130.3 | 99.9 | 125.8 | 71.4 | 88.1 | 72.2 | 90.7 | 71.1 | 88.9 | 71.9 | 88.9 | 85.0 | 103.0 | 84.9 | 102.8 |
| | 336 | 69.5 | 93.7 | 67.9 | 91.1 | 72.9 | 98.4 | 104.2 | 130.0 | 102.0 | 128.5 | 73.0 | 89.5 | 73.0 | 91.9 | 71.8 | 89.6 | 73.7 | 90.5 | 82.6 | 100.5 | 84.8 | 102.9 |
| | 720 | 72.6 | 97.3 | 70.7 | 94.6 | 76.2 | 101.2 | 107.3 | 134.2 | 104.2 | 131.4 | 76.1 | 92.9 | 75.1 | 93.3 | 72.9 | 91.0 | 76.7 | 93.7 | 84.1 | 105.1 | 85.4 | 103.8 |
| | Avg. | 68.5 | 92.7 | **66.9** | **90.1** | 71.8 | 96.9 | 103.5 | 130.2 | 100.3 | 126.5 | 72.1 | 88.8 | 72.6 | 91.0 | 71.1 | 89.0 | 72.7 | 89.6 | 84.2 | 102.9 | 84.9 | 103.1 |
| ODW2V | 96 | 76.8 | 99.7 | **65.1** | **88.4** | 78.5 | 102.7 | 89.6 | 112.7 | 89.1 | 112.5 | 74.8 | 96.8 | 76.3 | 99.9 | 73.5 | 97.7 | 92.2 | 117.7 | 77.0 | 100.1 | 77.4 | 100.4 |
| | 192 | 77.9 | 100.8 | **68.3** | **91.4** | 79.7 | 103.8 | 90.5 | 114.2 | 96.4 | 120.1 | 76.6 | 98.9 | 79.9 | 103.3 | 78.8 | 103.6 | 100.5 | 128.1 | 78.3 | 101.8 | 78.0 | 101.1 |
| | 336 | 78.5 | 101.5 | **69.9** | **93.0** | 80.3 | 104.5 | 94.2 | 119.3 | 98.4 | 122.2 | 77.6 | 100.2 | 81.8 | 105.3 | 82.1 | 107.5 | 105.5 | 134.4 | 79.4 | 103.3 | 78.7 | 102.0 |
| | 720 | 79.9 | 103.6 | **72.9** | **96.5** | 82.0 | 106.7 | 97.4 | 120.4 | 100.5 | 125.0 | 79.6 | 103.0 | 86.2 | 100.2 | 86.2 | 112.7 | 111.0 | 141.3 | 86.1 | 112.3 | 81.3 | 105.6 |
| | Avg. | 78.3 | 101.4 | **69.0** | **92.3** | 80.1 | 104.4 | 92.9 | 116.6 | 96.1 | 120.0 | 77.2 | 99.7 | 81.1 | 102.2 | 80.2 | 105.4 | 102.3 | 130.4 | 80.2 | 104.4 | 78.8 | 102.2 |
| $1^{st}$ Count | | 0 | | **29** | | 0 | | 0 | | 4 | | 0 | | 0 | | 0 | | 0 | | 0 | | 0 | |

# ❖ Few-shot Experiments (5% training data)

| Method | | LM-WEATHER-AVE | | LM-WEATHER | | FL-GPT4TS | | FL-Reformer | | FL-Pyraformer | | FL-Dlinear | | FL-PatchTST | | FL-iTransformer | | FL-Lights | | FL-Transformer | | FL-Informer | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Metrics | Length | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| ODW1T | 96 | 88.1 | 95.1 | 87.3 | **93.9** | 91.6 | 100.9 | 166.9 | 296.0 | 173.6 | 299.2 | 92.4 | 187.5 | **85.1** | 182.7 | 103.3 | 204.8 | 185.8 | 328.3 | 93.7 | 193.5 | 91.1 | 190.0 |
| | 192 | 90.2 | 98.4 | **89.6** | **96.5** | 95.8 | 104.6 | 166.9 | 297.3 | 176.0 | 303.0 | 94.4 | 192.4 | 90.7 | 191.6 | 106.7 | 210.7 | 188.1 | 336.0 | 96.8 | 200.1 | 93.5 | 195.9 |
| | 336 | 94.2 | 101.7 | **92.2** | **99.7** | 100.0 | 108.2 | 168.9 | 297.5 | 177.6 | 303.0 | 95.9 | 193.2 | 96.5 | 197.4 | 108.7 | 211.8 | 188.4 | 334.6 | 100.2 | 203.3 | 99.3 | 201.0 |
| | 720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 90.8 | **98.4** | 89.7 | 96.7 | 95.8 | 104.6 | 167.7 | 296.9 | 175.7 | 301.7 | 94.2 | 191.0 | **90.7** | 190.6 | 106.3 | 209.1 | 187.5 | 333.0 | 96.9 | 199.0 | 94.6 | 195.6 |
| ODW1V | 96 | 79.6 | 104.3 | **75.7** | **98.1** | 82.0 | 108.5 | 101.5 | 130.2 | 81.6 | 107.5 | 98.8 | 127.4 | 327.6 | 392.4 | 135.0 | 168.3 | 111.0 | 141.6 | 116.6 | 155.8 | 111.5 | 144.9 |
| | 192 | 87.8 | 115.5 | **82.5** | **108.4** | 90.8 | 120.3 | 107.0 | 136.7 | 90.2 | 118.9 | 110.4 | 141.6 | 334.4 | 403.4 | 145.4 | 180.2 | 117.5 | 149.2 | 123.4 | 164.1 | 116.0 | 152.4 |
| | 336 | 103.9 | 133.4 | **98.7** | **125.4** | 107.0 | 138.7 | 113.2 | 142.4 | 106.1 | 137.5 | 120.0 | 153.2 | 341.6 | 413.7 | 122.1 | 153.5 | 126.3 | 159.7 | 133.6 | 161.3 | 123.2 | 167.4 |
| | 720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 90.4 | 117.7 | **85.6** | **110.6** | 93.3 | 122.5 | 107.2 | 136.4 | 92.6 | 121.3 | 109.7 | 140.7 | 334.5 | 403.2 | 134.2 | 167.3 | 118.3 | 150.2 | 124.5 | 160.4 | 116.9 | 154.9 |
| ODW2T | 96 | 111.0 | 159.4 | **99.0** | **135.5** | 127.9 | 178.2 | 158.3 | 241.2 | 173.3 | 247.1 | 107.1 | 152.8 | 101.2 | 147.9 | 115.9 | 166.3 | 183.6 | 273.3 | 142.3 | 199.6 | 158.8 | 201.3 |
| | 192/336/720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 111.0 | 159.4 | **99.0** | **135.5** | 127.9 | 178.2 | 158.3 | 241.2 | 173.3 | 247.1 | 107.1 | 152.8 | 101.2 | 147.9 | 115.9 | 166.3 | 183.6 | 273.3 | 142.3 | 199.6 | 158.8 | 201.3 |
| ODW2V | 96 | 105.3 | 135.7 | **96.8** | **122.1** | 110.2 | 155.4 | 151.5 | 190.7 | 150.5 | 189.3 | 112.2 | 141.2 | 115.5 | 145.8 | 110.2 | 143.4 | 162.1 | 212.5 | 106.4 | 136.8 | 149.6 | 188.2 |
| | 192/336/720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 105.3 | 135.7 | **96.8** | **122.1** | 110.2 | 155.4 | 151.5 | 190.7 | 150.5 | 189.3 | 112.2 | 141.2 | 115.5 | 145.8 | 110.2 | 143.4 | 162.1 | 212.5 | 106.4 | 136.8 | 149.6 | 188.2 |
| $1^{st}$ Count | | 1 | | **16** | | 0 | | 0 | | 0 | | 0 | | 2 | | 0 | | 0 | | 0 | | 0 | |

1. Motivations

2. Contributions

3. Framework

4. Datasets

5. Results

## ❑ Main Results (Imputation, 50% masking ratio)

| Dataset | Length | LM-WEATHER-AVE MAE | LM-WEATHER-AVE RMSE | LM-WEATHER MAE | LM-WEATHER RMSE | FL-GPT4TS MAE | FL-GPT4TS RMSE | FL-Reformer MAE | FL-Reformer RMSE | FL-Pyraformer MAE | FL-Pyraformer RMSE | FL-DLinear MAE | FL-DLinear RMSE | FL-PatchTST MAE | FL-PatchTST RMSE | FL-iTransformer MAE | FL-iTransformer RMSE | FL-LightTS MAE | FL-LightTS RMSE | FL-Transformer MAE | FL-Transformer RMSE | FL-Informer MAE | FL-Informer RMSE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ODW1T | 96 | 22.4 | 43.5 | **21.7** | **41.8** | 23.3 | 45.2 | 63.7 | 88.4 | 62.2 | 85.9 | 29.2 | 50.8 | 28.9 | 54.6 | 22.8 | 44.5 | 24.4 | 43.7 | 58.3 | 82.8 | 70.8 | 99.6 |
| | 192 | 23.4 | 43.7 | **22.6** | **42.0** | 24.6 | 45.9 | 67.2 | 91.2 | 65.5 | 88.5 | 28.7 | 50.2 | 47.5 | 77.3 | 23.8 | 44.1 | 25.7 | 45.3 | 57.3 | 82.4 | 66.3 | 92.1 |
| | 336 | 24.1 | 44.1 | **23.2** | **42.4** | 25.3 | 46.3 | 70.4 | 93.4 | 68.5 | 90.6 | 28.3 | 49.4 | 48.6 | 77.0 | 27.2 | 47.7 | 26.9 | 46.6 | 58.4 | 83.5 | 36.9 | 55.3 |
| | 720 | 26.0 | 45.1 | **24.9** | **43.3** | 27.3 | 47.4 | 77.9 | 96.8 | 75.8 | 93.9 | 28.0 | 49.0 | 56.6 | 85.1 | 36.5 | 56.2 | 27.2 | 47.4 | 56.6 | 80.4 | 71.7 | 96.7 |
| | Avg. | 24.0 | 44.1 | **23.1** | **42.4** | 25.1 | 46.2 | 69.8 | 92.5 | 68.0 | 89.7 | 28.5 | 49.9 | 45.4 | 73.5 | 27.6 | 48.2 | 26.1 | 45.7 | 57.6 | 82.3 | 61.4 | 85.9 |
| ODW1V | 96 | 42.1 | 62.0 | 41.1 | 60.4 | 42.9 | 63.8 | 43.8 | 64.9 | 42.3 | **53.0** | 43.0 | 63.0 | 53.6 | 77.1 | **38.7** | 58.2 | 41.5 | 61.5 | 37.8 | 56.9 | 41.1 | 59.2 |
| | 192 | 43.9 | 64.5 | 42.8 | 62.8 | 45.6 | 66.9 | 45.8 | 67.6 | 44.7 | 56.2 | 49.3 | 71.2 | 57.5 | 81.5 | 49.3 | 68.9 | **41.9** | 62.0 | 44.1 | **57.4** | 48.8 | 66.8 |
| | 336 | 45.7 | 66.6 | **44.6** | **64.9** | 47.5 | 69.2 | 47.6 | 69.8 | 54.6 | 65.7 | 53.4 | 76.6 | 60.7 | 85.0 | 60.0 | 79.8 | 47.3 | 64.6 | 48.5 | 68.0 | 50.2 | 67.1 |
| | 720 | 47.5 | 68.7 | **46.3** | **66.9** | 49.4 | 71.4 | 49.6 | 72.0 | 59.2 | 73.5 | 56.8 | 80.7 | 63.3 | 87.4 | 61.6 | 80.4 | 52.5 | 72.9 | 52.7 | 70.1 | 60.3 | 77.2 |
| | Avg. | 44.8 | 65.5 | **43.7** | 63.8 | 46.4 | 67.8 | 46.7 | 68.6 | 50.2 | 62.1 | 50.6 | 72.9 | 58.8 | 82.7 | 52.4 | 71.8 | 45.8 | 65.3 | 45.8 | **63.1** | 50.1 | 67.6 |
| ODW2T | 96 | 38.0 | 56.6 | **36.9** | **54.9** | 39.1 | 58.3 | 50.3 | 70.3 | 95.4 | 120.8 | 40.8 | 60.0 | 38.4 | 58.6 | 39.1 | 58.3 | 38.8 | 57.8 | 65.5 | 86.6 | 51.7 | 72.0 |
| | 192 | 38.3 | 56.6 | **37.2** | **54.9** | 39.8 | 58.9 | 52.1 | 74.2 | 96.2 | 122.3 | 42.9 | 62.7 | 66.7 | 87.8 | 39.4 | 58.3 | 39.5 | 58.4 | 71.4 | 92.8 | 55.0 | 75.7 |
| | 336 | 43.5 | 65.5 | **42.2** | **63.5** | 44.8 | 68.1 | 56.6 | 78.9 | 97.8 | 125.5 | 46.0 | 67.7 | 68.7 | 90.1 | 44.8 | 67.5 | 47.8 | 65.3 | 66.8 | 88.8 | 51.5 | 72.8 |
| | 720 | 47.9 | 68.8 | **46.5** | **66.7** | 49.8 | 71.5 | 64.3 | 87.7 | 99.1 | 129.9 | 52.8 | 76.1 | 70.4 | 93.5 | 49.3 | 71.0 | 48.0 | 68.0 | 67.4 | 89.2 | 51.5 | 73.0 |
| | Avg. | 41.9 | 61.9 | **38.8** | **61.7** | 43.4 | 64.2 | 55.8 | 77.8 | 97.1 | 124.6 | 45.6 | 66.6 | 61.1 | 82.5 | 43.2 | 63.8 | 43.5 | 62.4 | 67.8 | 89.4 | 52.4 | 73.4 |
| ODW2V | 96 | 28.1 | 45.3 | **27.5** | **44.0** | 28.4 | 45.8 | 50.3 | 70.3 | 53.2 | 72.4 | 72.1 | 92.0 | 39.8 | 58.4 | 72.7 | 94.7 | 96.4 | 123.5 | 52.7 | 73.2 | 54.8 | 76.9 |
| | 192 | 28.6 | 45.3 | **28.0** | **44.0** | 29.2 | 46.1 | 51.0 | 71.1 | 46.1 | 65.2 | 75.7 | 95.9 | 44.9 | 63.7 | 79.1 | 102.0 | 98.6 | 125.8 | 53.9 | 74.7 | 56.2 | 78.8 |
| | 336 | 33.7 | 49.8 | **32.7** | **48.4** | 34.9 | 51.8 | 54.2 | 76.6 | 74.2 | 97.3 | 77.3 | 97.8 | 50.9 | 70.1 | 82.6 | 106.1 | 101.2 | 128.8 | 54.4 | 75.4 | 56.8 | 79.7 |
| | 720 | 37.1 | 53.1 | **36.0** | **51.5** | 39.3 | 56.3 | 59.4 | 81.7 | 82.4 | 100.9 | 77.1 | 97.3 | 59.2 | 79.3 | 83.0 | 106.0 | 98.5 | 124.3 | 55.4 | 77.5 | 56.4 | 78.6 |
| | Avg. | 31.9 | 48.4 | **31.1** | **47.0** | 33.0 | 50.0 | 53.7 | 74.9 | 64.0 | 84.0 | 75.5 | 95.8 | 48.7 | 67.9 | 79.4 | 102.2 | 98.7 | 125.6 | 54.1 | 75.2 | 56.0 | 78.5 |
| $1^{st}$ Count | | 0 | | **30** | | 0 | | 0 | | 0 | | 0 | | 0 | | 0 | | 0 | | 0 | | 0 | |

## ❖ Few-shot Experiments (5% training data, 50% masking ratio)

| Ratio | Length | LM-WEATHER-AVE MAE | LM-WEATHER-AVE RMSE | LM-WEATHER MAE | LM-WEATHER RMSE | FL-GPT4TS MAE | FL-GPT4TS RMSE | FL-Reformer MAE | FL-Reformer RMSE | FL-Pyraformer MAE | FL-Pyraformer RMSE | FL-DLinear MAE | FL-DLinear RMSE | FL-PatchTST MAE | FL-PatchTST RMSE | FL-iTransformer MAE | FL-iTransformer RMSE | FL-LightTS MAE | FL-LightTS RMSE | FL-Transformer MAE | FL-Transformer RMSE | FL-Informer MAE | FL-Informer RMSE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ODW1T | 96 | 61.2 | 121.2 | **59.9** | **120.8** | 62.4 | 138.6 | 147.4 | 261.3 | 149.5 | 256.4 | 110.0 | 209.1 | 64.2 | 147.0 | 119.0 | 228.5 | 173.0 | 310.8 | 140.8 | 260.0 | 143.9 | 264.7 |
| | 192 | 69.1 | 130.2 | **64.7** | **127.7** | 67.3 | 145.2 | 151.3 | 267.8 | 152.0 | 258.1 | 110.1 | 203.4 | 74.1 | 155.1 | 120.9 | 223.0 | 172.2 | 301.4 | 149.1 | 262.4 | 150.3 | 264.2 |
| | 336/720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 65.2 | 125.7 | **62.3** | **124.4** | 64.9 | 141.9 | 149.4 | 264.6 | 150.8 | 257.3 | 110.0 | 206.2 | 69.2 | 151.0 | 120.0 | 225.7 | 172.6 | 306.1 | 145.0 | 261.2 | 147.1 | 264.4 |
| ODW1V | 96 | 62.2 | 134.1 | 62.8 | 135.5 | 67.3 | **131.2** | 103.9 | 189.8 | **61.5** | 132.6 | 112.2 | 208.4 | 161.1 | 281.5 | 117.6 | 219.5 | 119.6 | 223.5 | 98.0 | 198.5 | 94.2 | 188.1 |
| | 192 | 71.4 | 140.5 | 72.2 | 142.1 | 74.6 | 152.7 | 103.3 | 182.4 | **70.6** | **138.8** | 113.3 | 200.6 | 160.5 | 272.4 | 124.5 | 218.9 | 122.7 | 217.8 | 101.8 | 191.3 | 96.7 | 181.5 |
| | 336/720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 66.8 | 137.3 | 67.5 | 138.8 | 71.0 | 142.0 | 103.6 | 186.1 | **66.1** | **135.7** | 112.7 | 204.5 | 160.8 | 276.9 | 121.0 | 219.2 | 121.2 | 220.7 | 101.8 | 194.9 | 95.5 | 184.8 |
| ODW2T | 96 | 102.5 | 156.3 | **99.4** | **151.6** | 112.0 | 157.2 | 116.2 | 161.3 | 124.9 | 165.6 | 123.7 | 178.3 | 173.0 | 256.7 | 127.3 | 190.6 | 133.8 | 200.3 | 124.3 | 187.5 | 105.7 | 161.1 |
| | 192/336/720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 102.5 | 156.3 | **99.4** | **151.6** | 112.0 | 157.2 | 116.2 | 161.3 | 124.9 | 165.6 | 123.7 | 178.3 | 173 | 256.7 | 127.3 | 190.6 | 133.8 | 200 | 124.3 | 188 | 105.7 | 161.0 |
| ODW2V | 96 | 42.4 | 62.9 | **35.7** | 112.1 | 56.4 | 77.3 | 106.8 | 135.5 | 70.8 | **95.5** | 113.3 | 148.0 | 153.8 | 199.5 | 101.8 | 136.4 | 106.1 | 142.2 | 100.1 | 134.6 | 89.8 | 119.0 |
| | 192/336/720 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | Avg. | 42.4 | 62.9 | **35.7** | 112.1 | 56.4 | 77.3 | 106.8 | 135.5 | 70.8 | **95.5** | 113.3 | 148.0 | 153.8 | 199.5 | 101.8 | 136.4 | 106.1 | 142.2 | 100.1 | 134.6 | 89.8 | 119.0 |
| $1^{st}$ Count | | 0 | | **12** | | 1 | | 0 | | 6 | | 0 | | 0 | | 0 | | 0 | | 0 | | 0 | |

Personalized Adapter for Large Meteorology Model on Devices: Towards Weather Foundation Models (NerulPS 2024)

1. Motivations

2. Contributions

3. Framework

4. Datasets

5. Results

## ❑ Ablation Study

| Method | Task | | Ablation Perspective | | Ave. Variations | | Params.# | |
|---|---|---|---|---|---|---|---|---|
| | Forecasting | Imputation | Model Component | Personalized Method | Forecasting | Imputation | Train.# | Comm.# |
| LM-WEATHER | 45.4/74.6 | 23.1/40.0 | Original | Original | - | - | 10.38 M | 0.38 M |
| LM-WEATHER-A | 50.8/87.6 | 26.0/47.7 | *wo* Decomposition | Original | ↓11.8% | ↓12.6% | 10.38 M | 0.38 M |
| LM-WEATHER-B | 50.9/85.6 | 25.4/47.1 | *wo* Trend Component | Original | ↓12.1% | ↓10.0% | 10.37 M | 0.38 M |
| LM-WEATHER-C | 50.1/83.6 | 25.0/46.1 | *wo* Seasonal Component | Original | ↓10.3% | ↓8.2% | 10.37 M | 0.38 M |
| LM-WEATHER-D | 49.3/81.7 | 24.4/45.6 | *wo* Residual Component | Original | ↓8.6% | ↓5.6% | 10.37 M | 0.38 M |
| LM-WEATHER-E | 53.8/95.6 | 25.5/47.0 | *wo* Prompt Generator | Original | ↓18.5% | ↓10.4% | 10.36 M | 0.38 M |
| LM-WEATHER-F | 49.4/82.3 | 28.1/52.0 | Original | *w* LoRA, Local: Low-Rank Matrix, Global: the rest of trainable param. | ↓8.8% | ↓21.6% | 10.38 M | 10.00 M |
| LM-WEATHER-G | 43.2/71.4 | 22.4/39.1 | Original | *wo* LoRA, Local: Attention Param. Global: Attention Param | ↑5.1% | ↑3.1% | 52.01 M | 41.99 M |
| LM-WEATHER-H | **42.7/71.2** | **22.2/39.3** | Original | *wo* LoRA, Local: Attention Param. Global: the rest of trainable param. | ↑6.3% | ↑4.1% | 52.01 M | 10.00 M |

❖ The personalized adapter *effectively strength PLM's capabilities for weather data modeling*, providing a balance between performance and efficiency.

❖ Compared to other tuning strategies, our LoRA-based local tuning and communication method *significantly improves the computional and communication efficiency* of our framework.

# ❑ Framework Analysis

(compared to FL baselines that prioritize communication efficiency)

| Method | Forecasting | Imputation | Train. | Comm. Params. | Comm. |
|---|---|---|---|---|---|
| FL-Pyraformer | 73.0/91.3 | 68.0/89.7 | 153.32 M | 153.32 M | 0.07× |
| FL-PatchTST | 48.6/81.0 | 45.4/73.5 | 74.74 M | 74.74 M | 0.14× |
| FL-LightTS | 62.7/93.4 | 26.1/45.7 | 1.68 M | 1.68 M | 6.2× |
| FL-DLinear | 63.3/82.8 | 28.5/49.9 | 1.06 M | 1.06 M | 9.8× |
| LM-WEATHER-Ave | 47.5/78.7 | 24.0/44.1 | 10.38 M | 10.38 M | 1× |
| LM-WEATHER (Ours) | **45.4/74.6** | **23.1/42.4** | 10.38 M | **0.38 M** | **27.3×** |
| LM-WEATHER (*w* FedKD) | 49.6/76.2 | 27.5/43.6 | 10.38 M | 1.68 M | 6.2× |
| LM-WEATHER (*w* FedPer) | 52.1/79.0 | 25.1/44.3 | 10.38 M | 8.46 M | 1.2× |
| LM-WEATHER (*w* FedBF) | 46.2/78.1 | 23.7/44.0 | 10.49 M | 10.49 M | 0.9× |
| LM-WEATHER (*w* FedAP) | 47.4/79.2 | 24.3/44.7 | 10.38 M | 9.6 M | 1.1× |
| LM-WEATHER (*w* PromptFL) | 46.0/78.4 | 23.8/45.1 | 10.38 M | 8.4 M | 1.2× |

❖ Our method significantly outperforms the communication-efficient FL baselines in terms of communication efficiency.

# ❑ Framework Analysis
(Robustness to Number of Devices)

| Dataset | Rate / Devices | Normal Forecasting | Normal Imputation | Few-Shot (15%) Forecasting | Few-Shot (15%) Imputation |
|---|---|---|---|---|---|
| **ODW1T** | 0.1 (2/round) | 44.4/73.6 | 22.6/42.0 | 64.7/100.4 | 40.2/68.2 |
| | 0.3 (5/round) | 43.7/72.5 (↑ 1.55) | 24.2/43.7 (↓ 5.55) | 63.4/99.7 (↑ 1.40) | 41.4/68.7 (↓ 1.85) |
| | 0.5 (8/round) | 42.9/72.0 (↑ 2.85) | 21.0/42.1 (↑ 3.90) | 63.7/99.2 (↑ 1.40) | 42.3/68.5 (↓ 2.8) |
| | 0.7 (11/round) | 43.9/74.1 (↑ 0.25) | 21.8/41.2 (↑ 2.80) | 64.5/101.0 (↑ 0.10) | 39.5/66.7 (↑ 2.00) |
| | 1.0 (16/round) | 44.2/74.0 (0 -) | 21.3/41.6 (↑ 3.10) | 63.6/100.2 (↑ 0.95) | 40.4/68.0 (↓ 0.1) |
| **ODW2T** | 0.1 (4/round) | 66.2/89.1 | 37.2/54.9 | 89.7/131.8 | 77.2/112.6 |
| | 0.3 (11/round) | 68.2/89.7 (↓ 1.85) | 36.5/53.1 (↑ 2.65) | 90.2/132.5 (↓ 0.55) | 75.4/110.3 (↑ 2.25) |
| | 0.5 (18/round) | 65.4/89.2 (↑ 0.55) | 36.7/53.4 (↑ 2.05) | 89.1/131.4 (↑ 0.50) | 76.5/111.2 (↑ 1.10) |
| | 0.7 (25/round) | 65.7/88.8 (↑ 0.90) | 36.1/53.9 (↑ 2.45) | 88.9/130.9 (↑ 0.80) | 76.9/112.3 (↑ 0.35) |
| | 1.0 (36/round) | 65.9/89.0 (↑ 0.25) | 36.9/55.0 (↑ 0.30) | 89.1/130.7 (↑ 0.75) | 76.7/112.1 (↑ 0.50) |

- ❖ Increasing device count during training slightly boosts performance in both regular and few-shot settings.
- ❖ Data distribution imbalances can degrade performance when adding devices, showing non-linear gains.
- ❖ More deivces incrsase communication costs, potentially outweighting minor performance gains.

NEURAL INFORMATION PROCESSING SYSTEMS

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

THE HONG KONG POLYTECHNIC UNIVERSITY 香港理工大學

# Thank you!

[Paper] [Code] [Our Survey]