



Interpretable weather forecasting for worldwide stations with a unified deep model

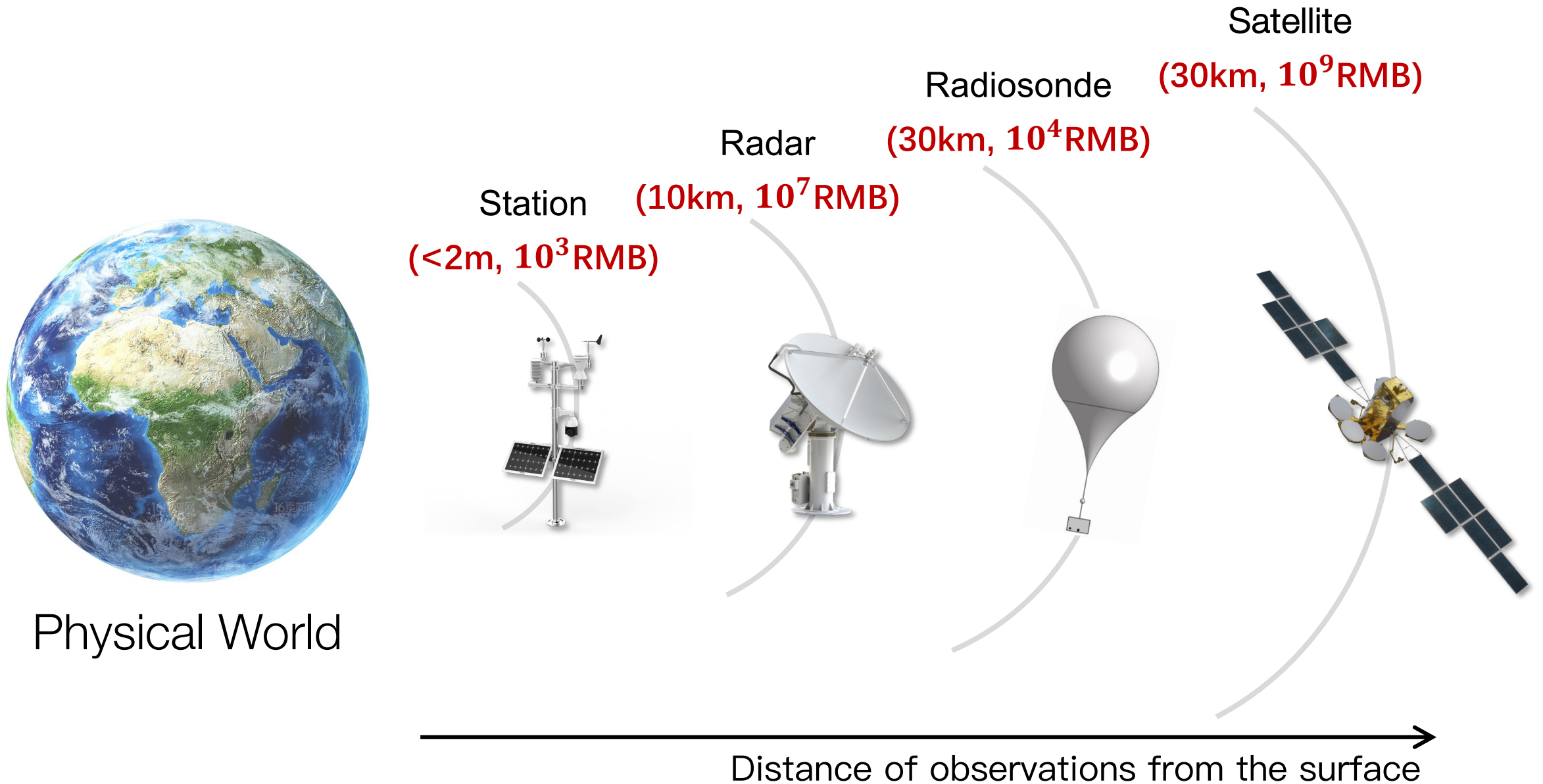
Haixu Wu, Hang Zhou, Mingsheng Long, Jianmin Wang

School of Software, Tsinghua University

Nature Machine Intelligence 2023

<https://www.nature.com/articles/s42256-023-00667-9>

How to Observe the Weather?



Automatic Weather Station

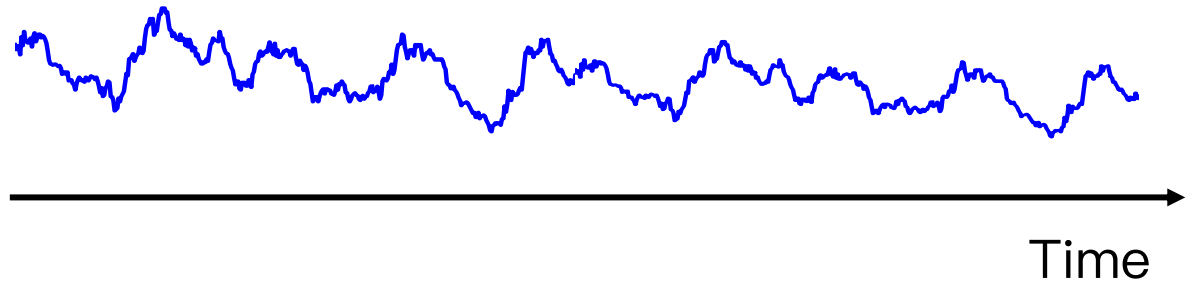


Physical World

Station
($<2\text{m}$, 10^3RMB)



**the most accurate and reliable
near-surface weather observation**



Automatic Weather Station

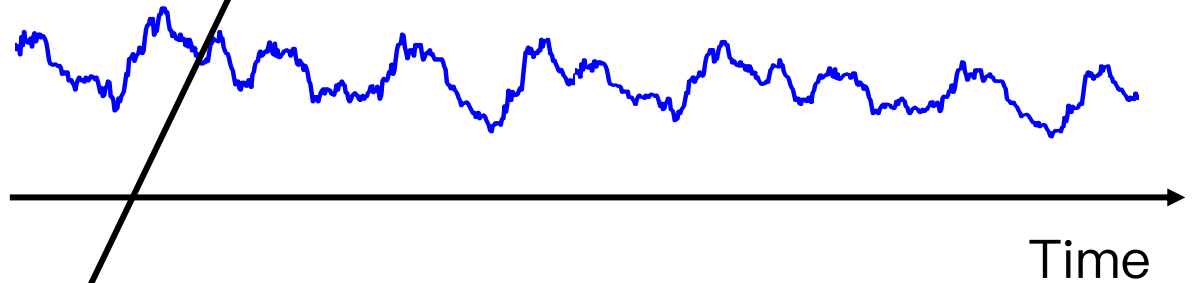


Physical World

Station
($<2\text{m}$, 10^3RMB)



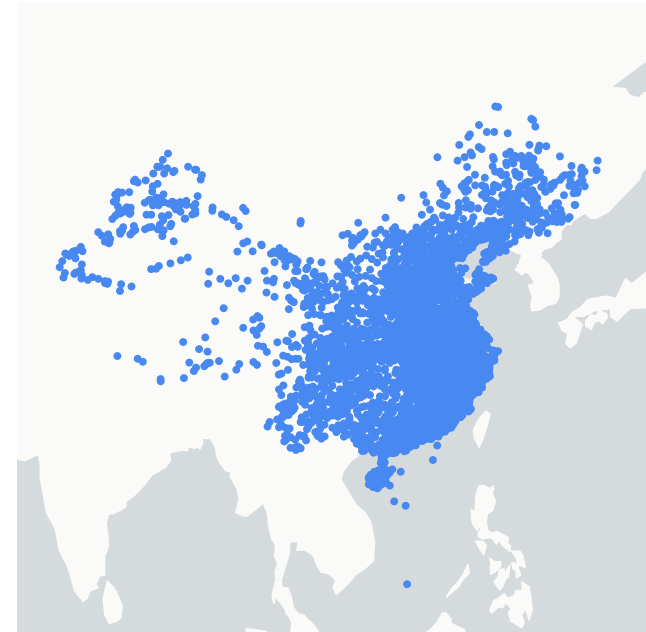
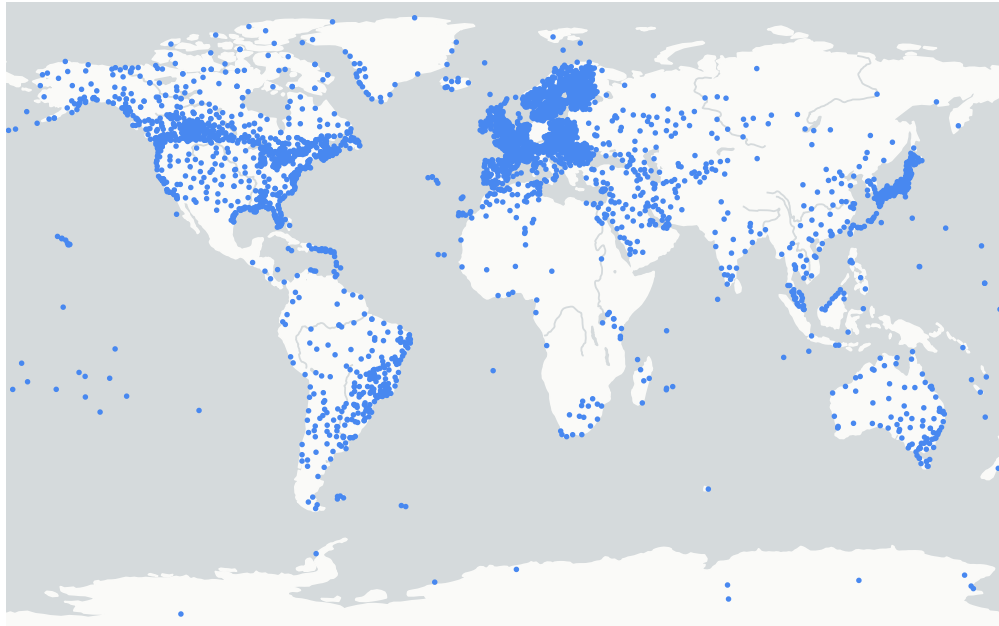
**the most accurate and reliable
near-surface weather observation**



Directly Affect Human Life

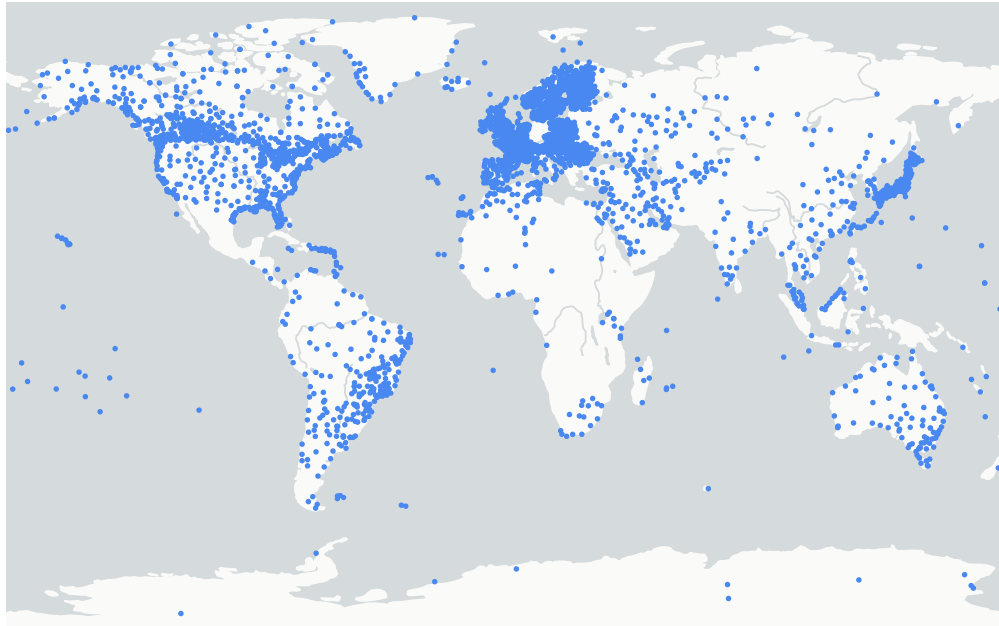
Objective

Tens of thousands of weather stations are **scattered** around the world and record the near-surface weather **every minute, even every second.**



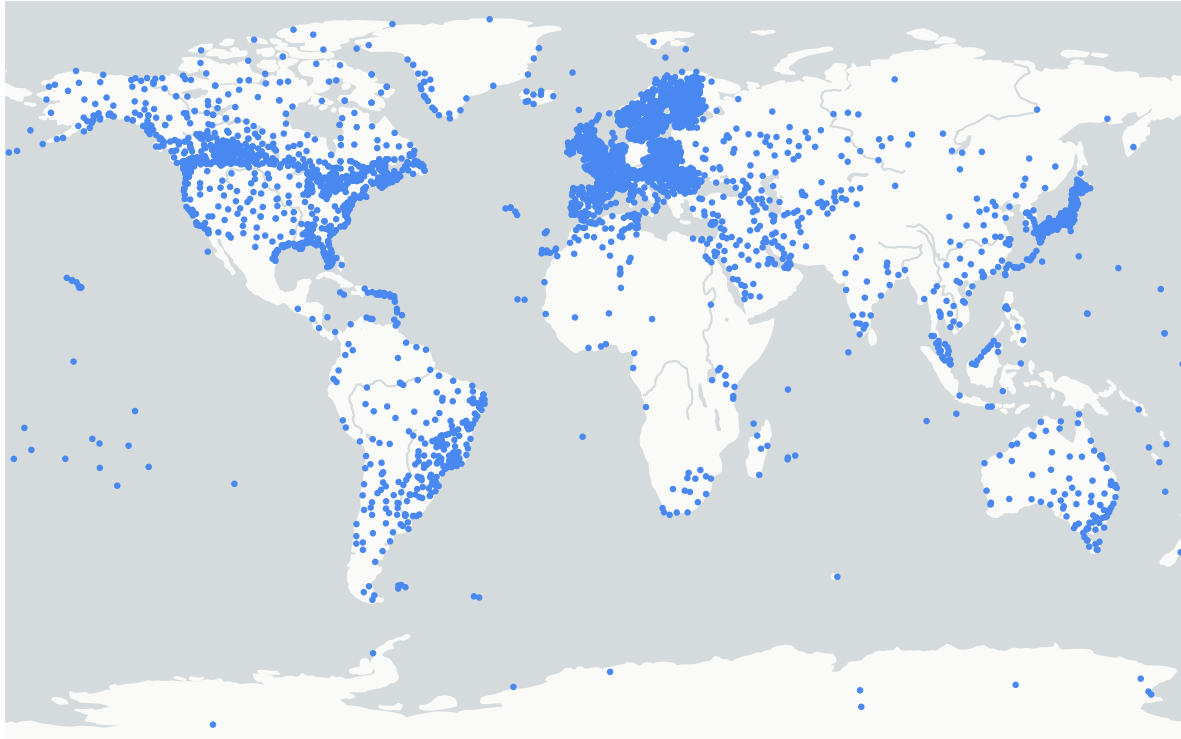
Objective

Real-time collaborative forecasts of worldwide tens of thousands automatic weather stations (prediction the future **0-24 hours** near-surface weather)

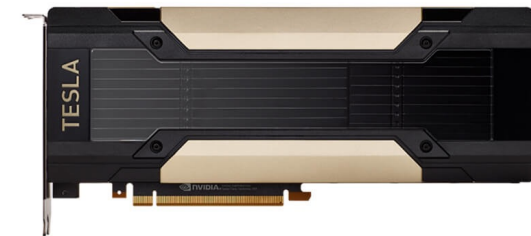


Worldwide Open Problem

Challenge 1: Huge Computation Cost



10,000x computation overhead in both GPU memory and running time



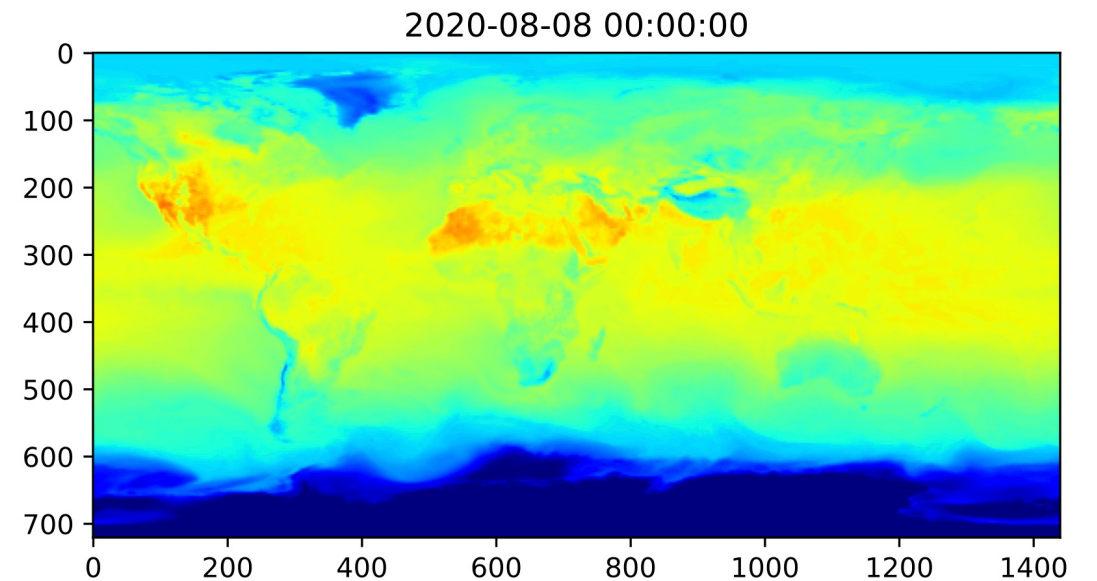
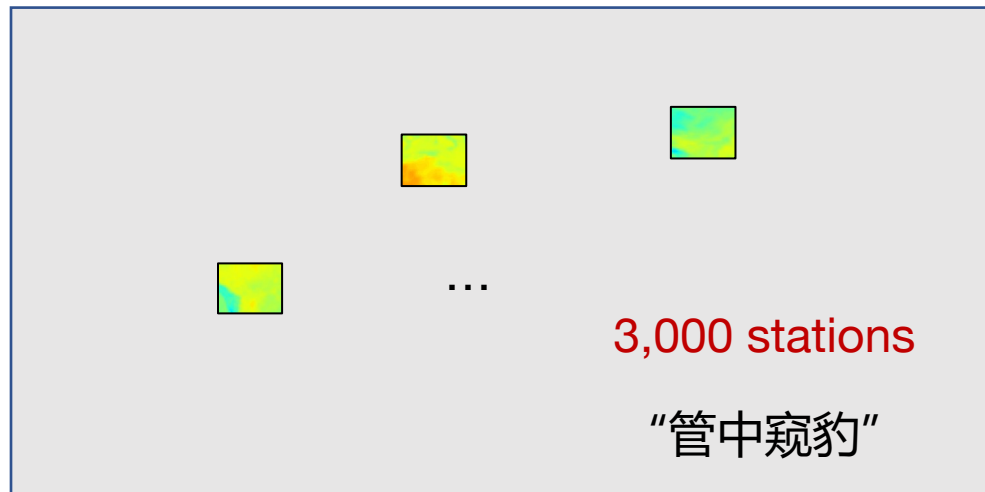
1000 GPU x 400 days

Challenge 2: Partially Observable System

Uncovering the complex system from partial observations

Observations from
scattered weather stations

Global weather system

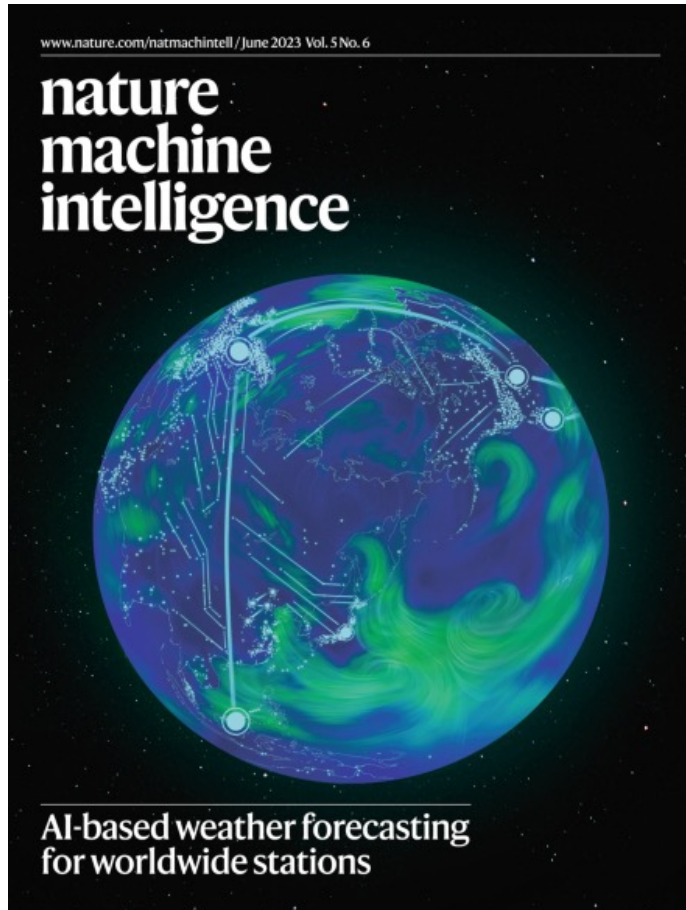


Challenge 3: Near-surface Forecasting



- Extremely complex topography
- Potential chaos effects

Corrformer



Large Meteorology Model for Near-surface Weather

Global collaborative forecasts for the first time

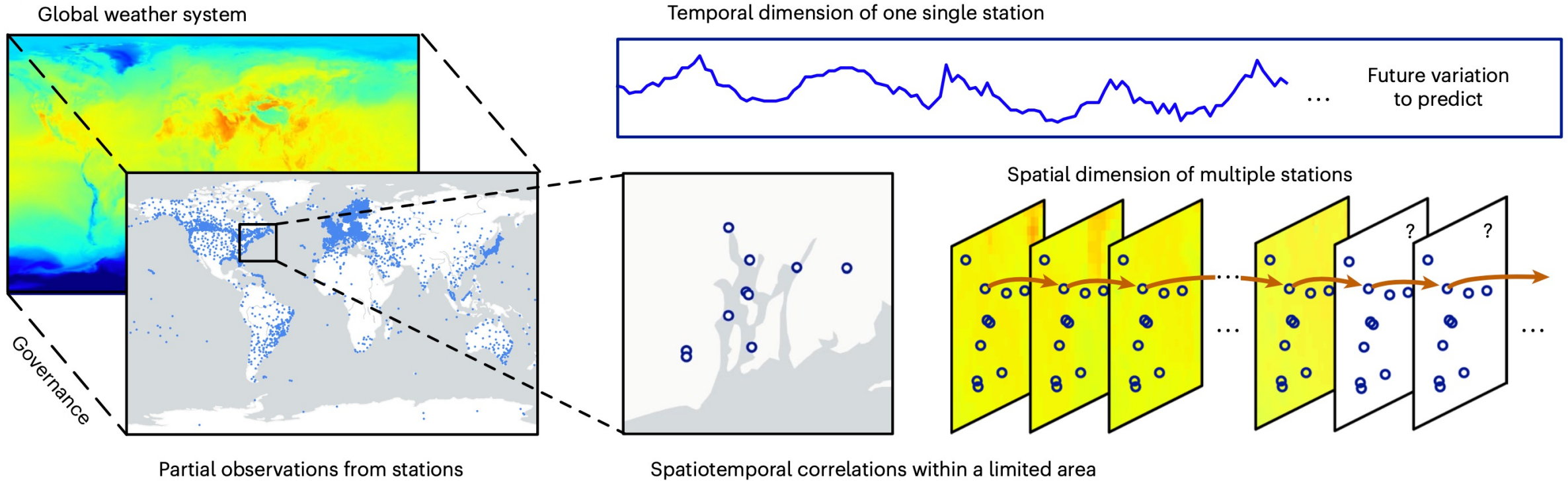
- **Accurate:** Beat EC and GFS in near-surface forecast
- **Efficient:** 1 day for training, 1 second for inference
- **Interpretable:** AI insights for meteorological science

Wu, H. et al. Interpretable weather forecasting for worldwide stations with a unified deep model,

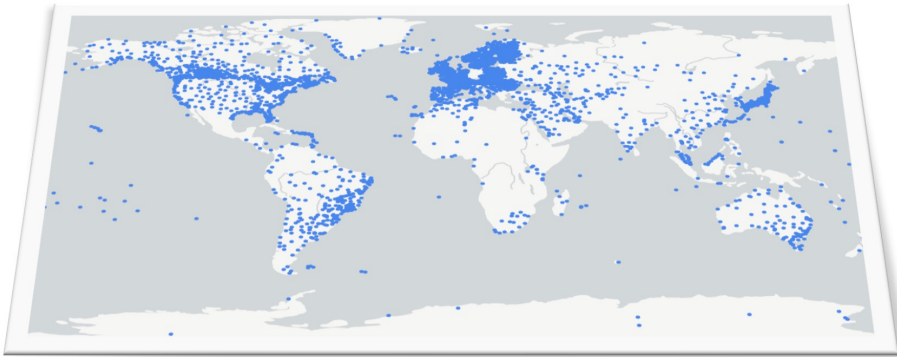
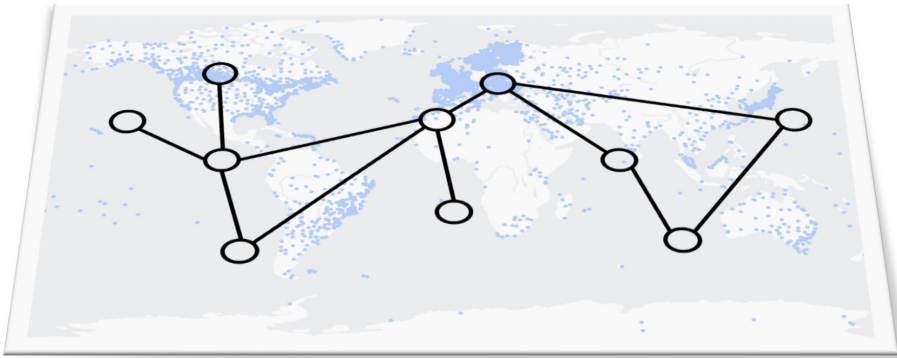
Nature Machine Intelligence 5, 602–611 (2023)

Problem Definition

Capture **complex spatiotemporal correlations** for worldwide weather stations.



Canonical Spatial Modeling



1. Convolution Neural Networks

- *Inapplicable to scattered stations*

2. Graph Neural Networks

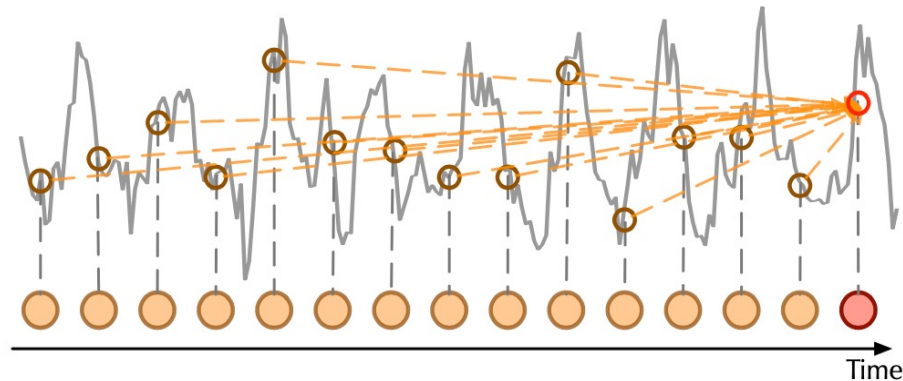
- Massive prior knowledge to construct graph
- Fixed graph vs. ever-changing weather

3. Attention-based models

- Quadratic complexity

How to organize these stations effectively?

Canonical Temporal Modeling



1. Recurrent Neural Networks

- Markov Assumption

2. Attention-based models

- Quadratic complexity
- Scattered steps modeling vs. Continuous weather processes

How to capture dependencies for continuous processes?

Auto-Correlation in Stochastic Process

- Auto-Correlation for wide-sense stationary stochastic process

$$R_{XX}(\tau) = \mathbf{E} \left[\underline{X_{t+\tau}} \underline{\bar{X}_t} \right]$$

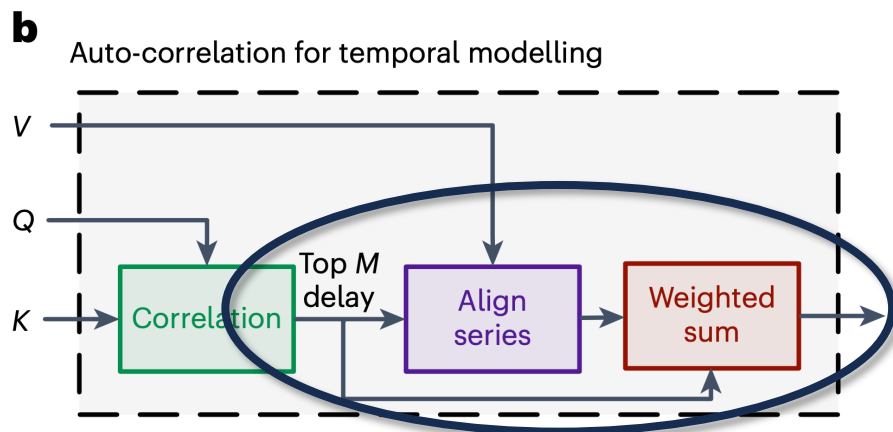
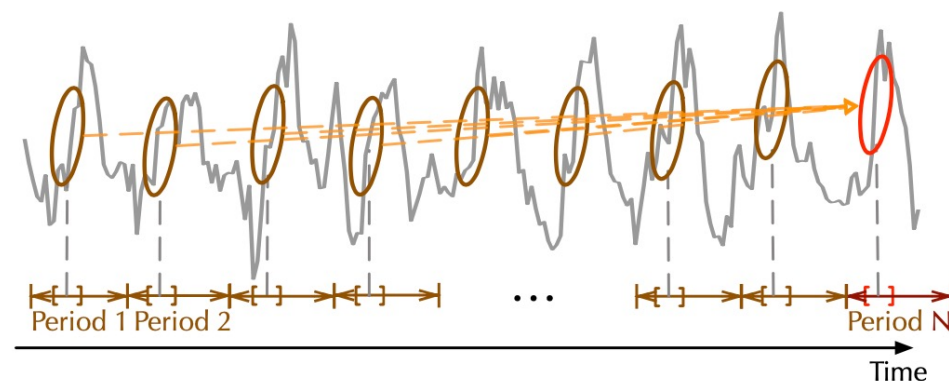
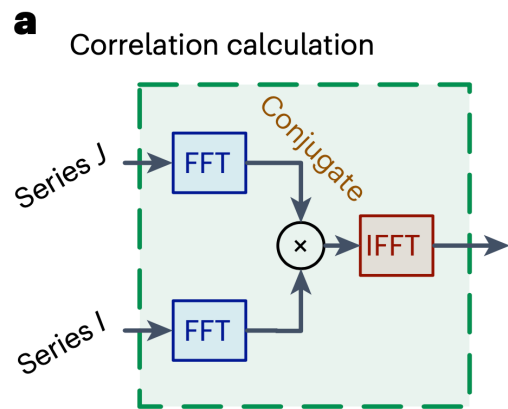
Time Delay Similarity -> Find Period

- Efficient computation with Fast Fourier Transform (FFT)

$$\begin{aligned} F_R(f) &= \text{FFT}[X(t)] \\ S(f) &= F_R(f) F_R^*(f) \\ R(\tau) &= \text{IFFT}[S(f)] \end{aligned}$$

Wiener-Khinchin Theorem

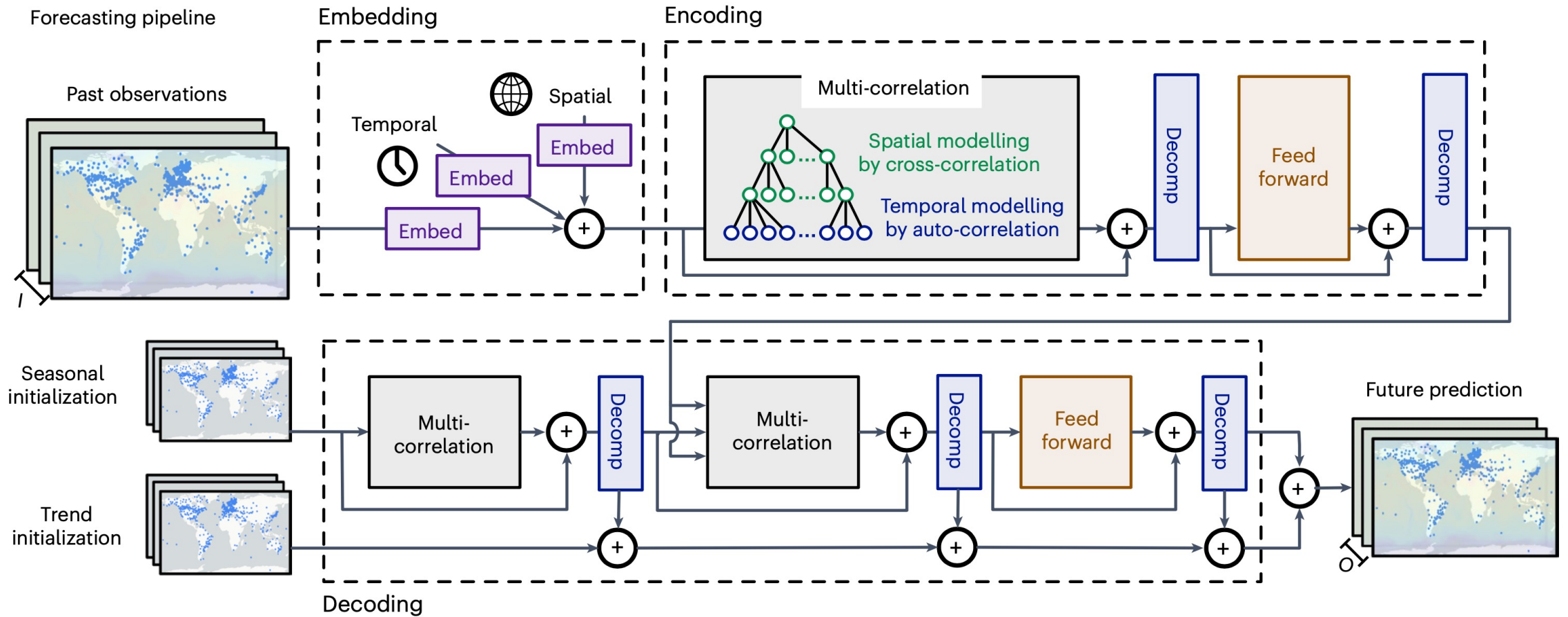
Auto-Correlation for Series-wise Dependencies Modeling



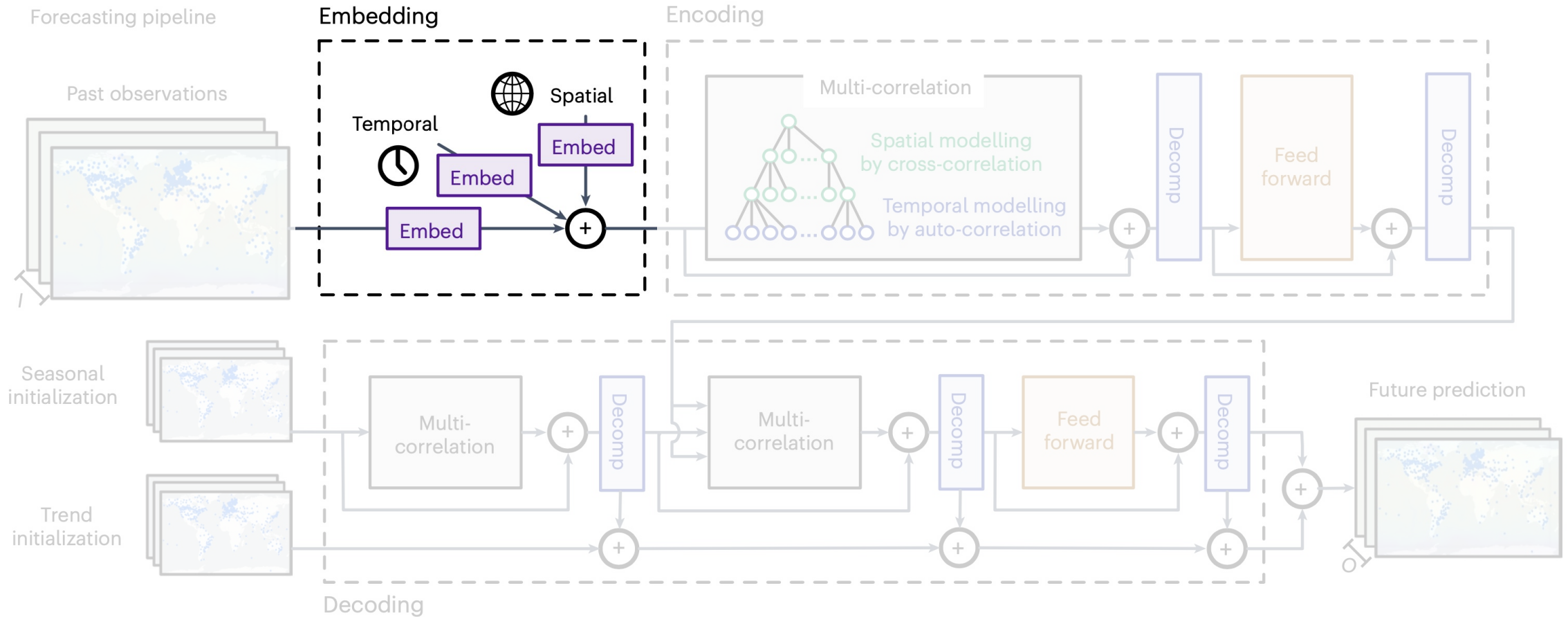
Auto-Correlation Mechanism

- Series-wise dependencies modeling
- Efficiency computation with complexity $O(L \log L)$

Corrformer

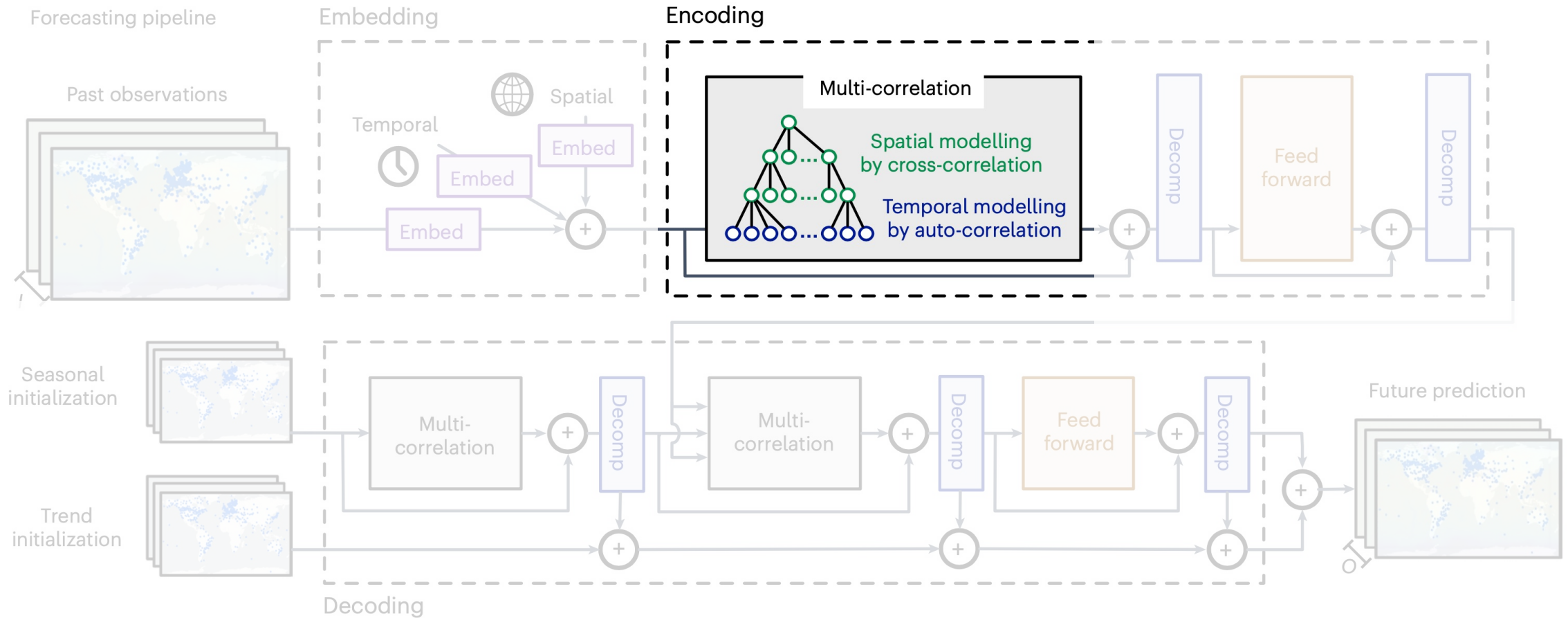


Corrformer



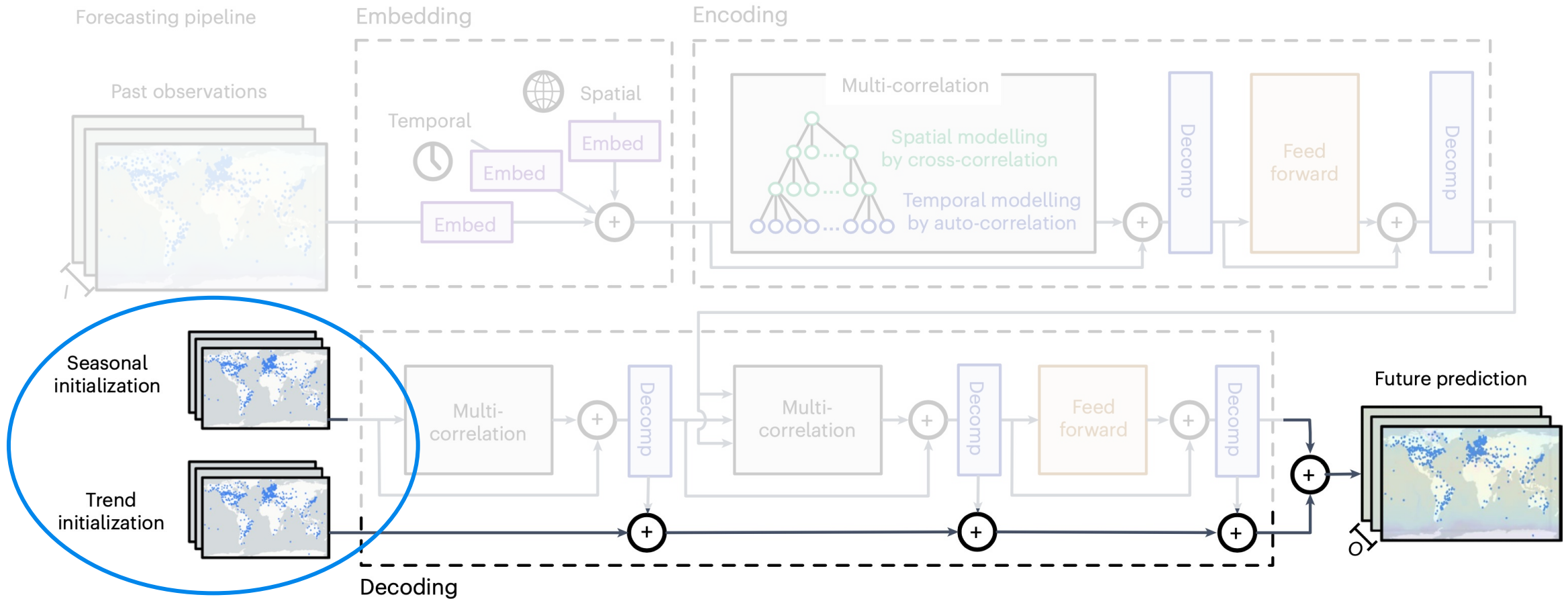
1. Embedding to incorporate **temporal, geographic and topography** information

Corrformer



2. **Multi-correlation** to capture complex spatiotemporal correlations

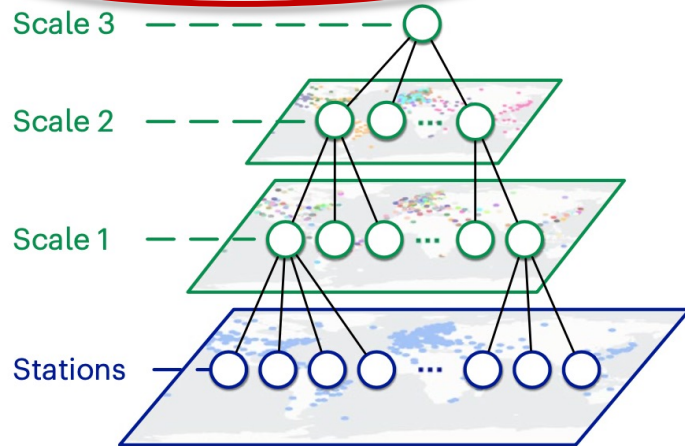
Corrformer



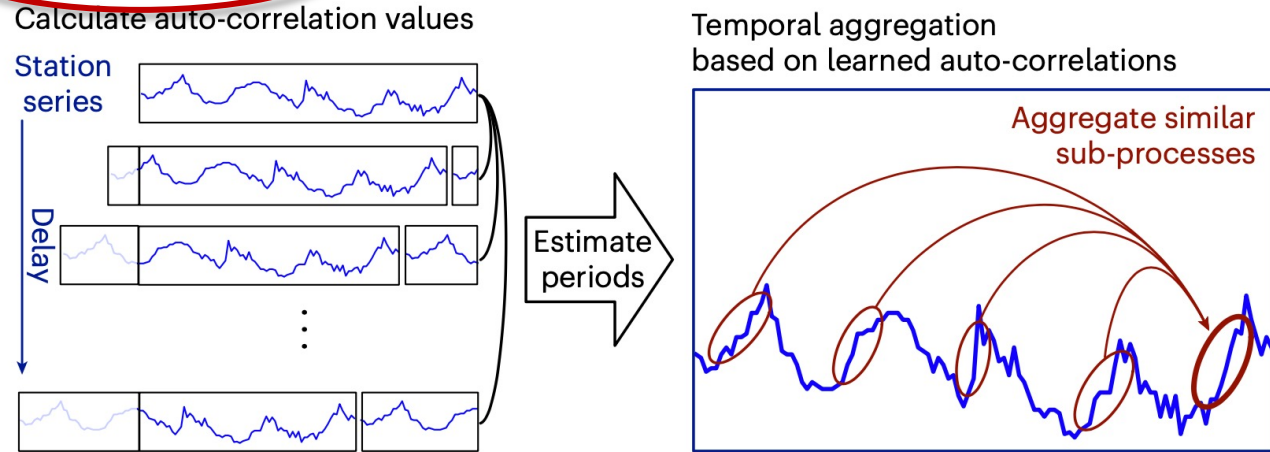
3. Deep decomposition architecture from Autoformer to **enhance stationarity**

Multi-correlation

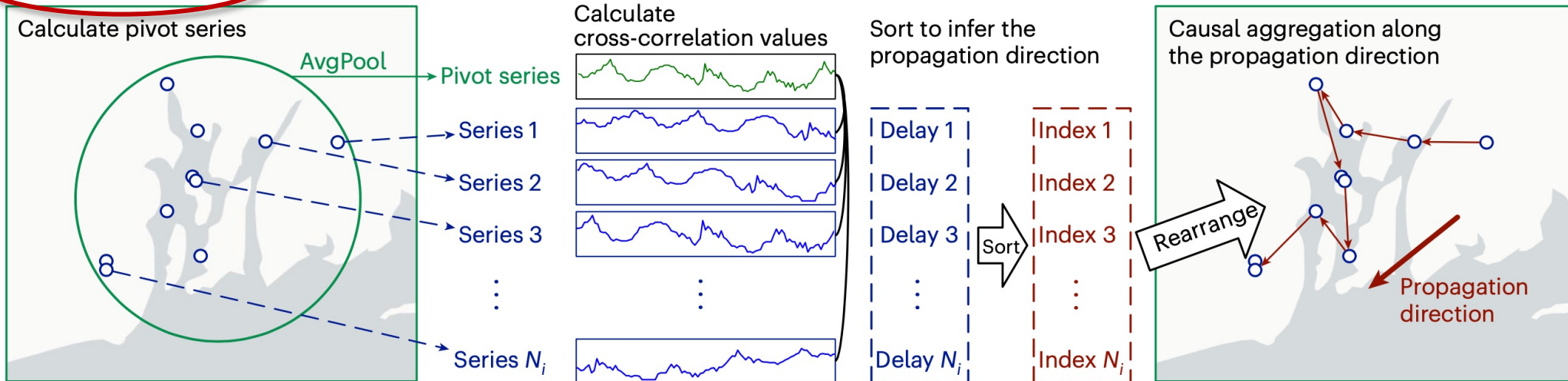
a Tree-based multiscale structure



b Auto-correlation on each leaf node for temporal modelling



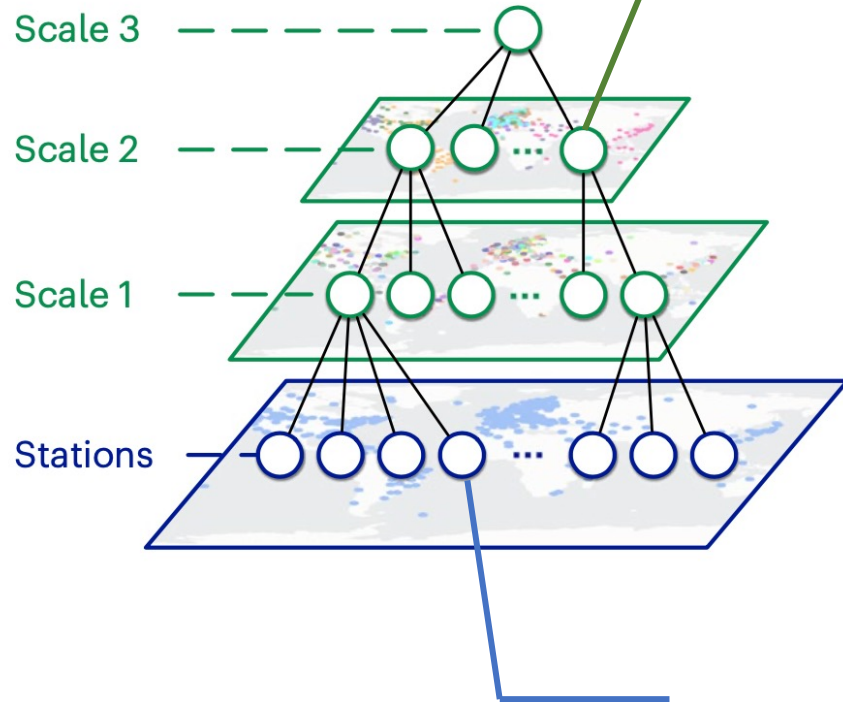
c Cross-correlation on each intermediate node for spatial modelling



Multi-correlation

Intermediate node: summed from child nodes for comprehensive meteorological characteristics of a region

a Tree-based multiscale structure

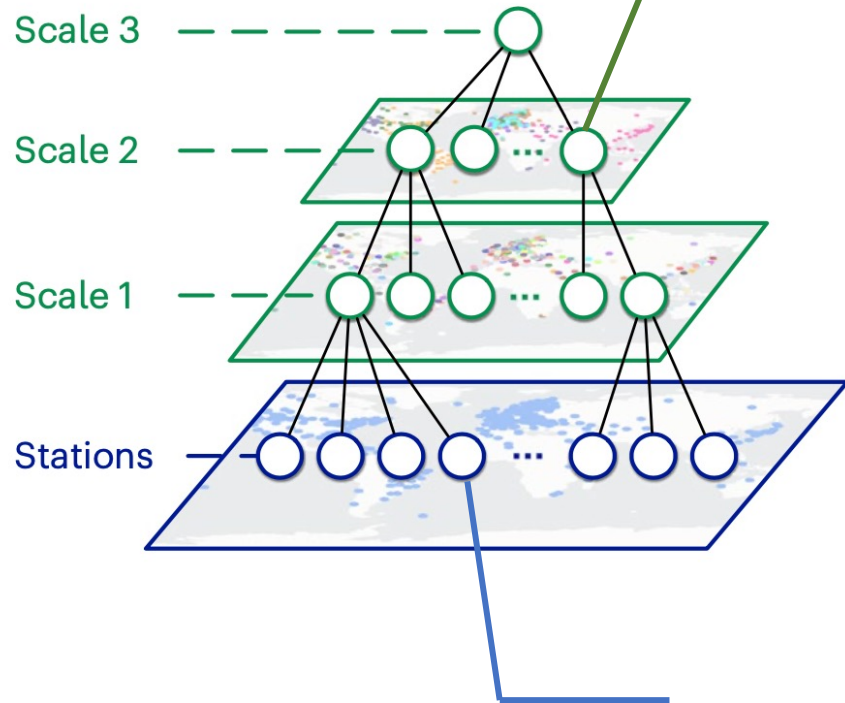


Leaf node: record the real observations

Multi-correlation

Intermediate node: summed from child nodes for comprehensive meteorological characteristics of a region

a Tree-based multiscale structure



Benefits of tree structure

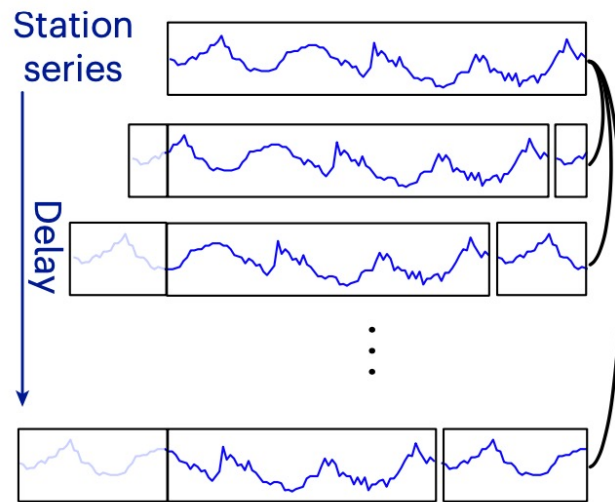
- More flexible for scattered stations
- Inherently model the multi-scale property of weather
- **Provide an elegant framework for spatiotemporal correlation modeling**
- **Significantly reduce computing overhead**

Leaf node: record the real observations

Multi-correlation

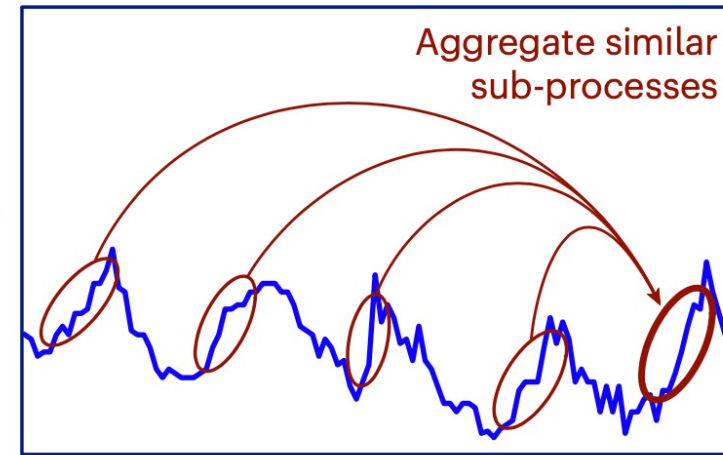
b Auto-correlation on each leaf node for temporal modelling

Calculate auto-correlation values



Estimate periods

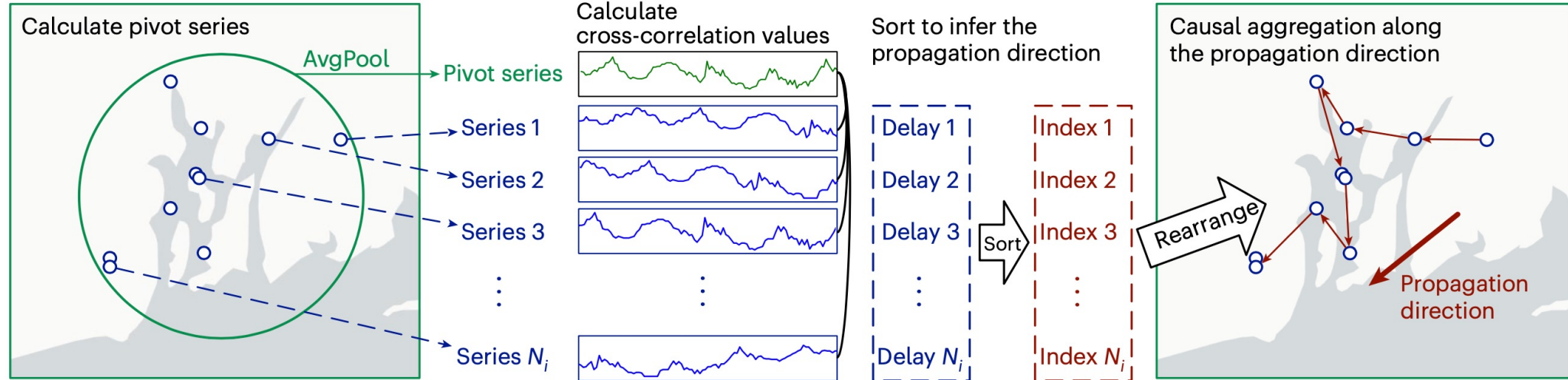
Temporal aggregation based on learned auto-correlations



For the leaf node, calculate Auto-correlation

Multi-correlation

C Cross-correlation on each intermediate node for spatial modelling



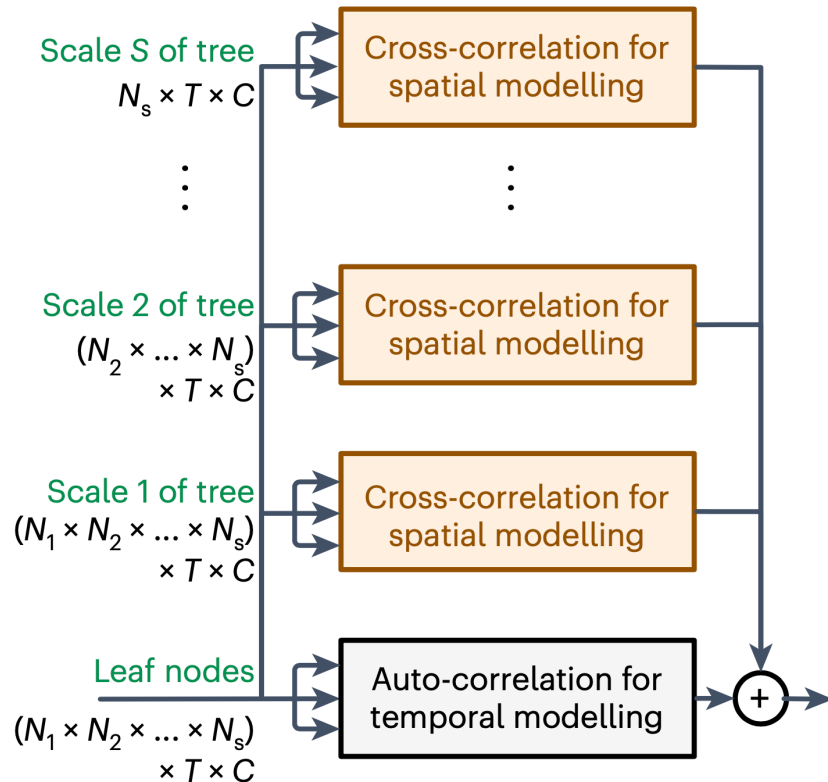
For each intermediate node, calculate Cross-correlation

1. Calculate the temporal delay **between intermediate node and its child nodes**
2. **Sort relative delay values** to infer the propagation direction.
3. **Causal aggregation** along the propagation direction.

Interpretable Evidence

Multi-correlation

Multi-correlation mechanism



Benefits of Multi-correlation

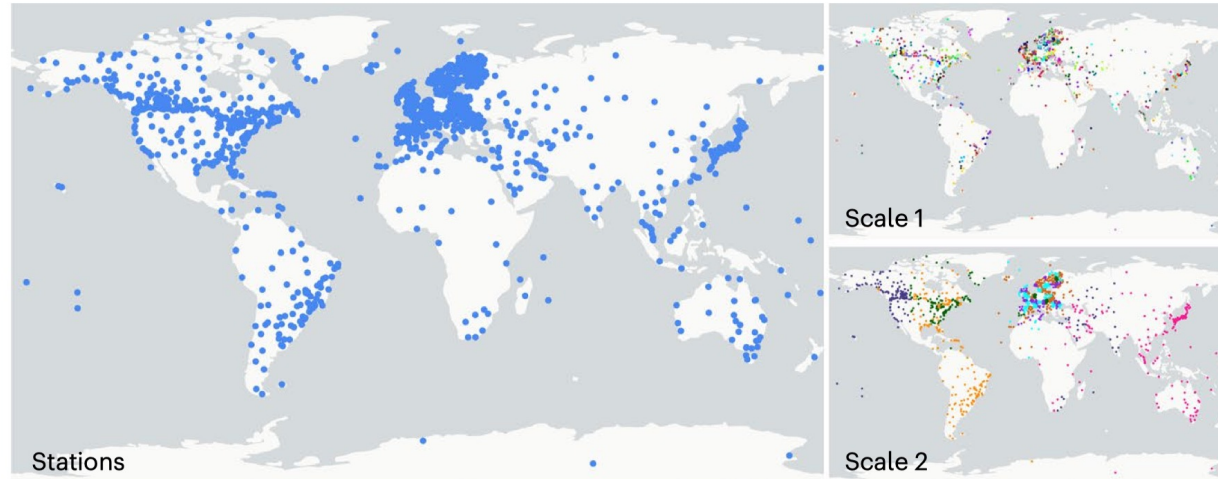
- **Multiscale modeling** for scattered stations
- **Interpretable evidence** for forecasting
- **Efficient computation** (*almost optimal*)

$$O(N^2 L^2) \rightarrow O(NL \log L)$$

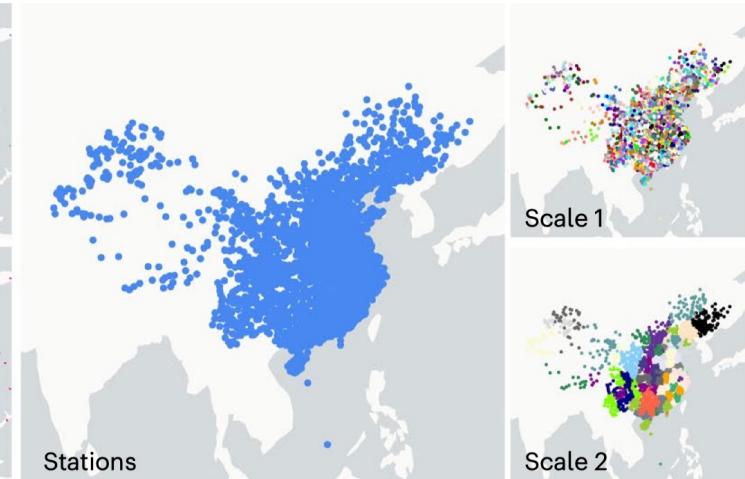
$$O \left(\sum_{i=1}^s \left(\underbrace{N_{i+1} \times \dots \times N_s}_{\text{intermediate nodes at scale } i} \right) \times \left(\underbrace{N_i L \log L}_{\text{cross-correlation}} + \underbrace{N_i \log N_i}_{\text{sort}} + \underbrace{N_i L}_{\text{aggregation}} \right) \right) = O(NL \log L),$$

Experiments

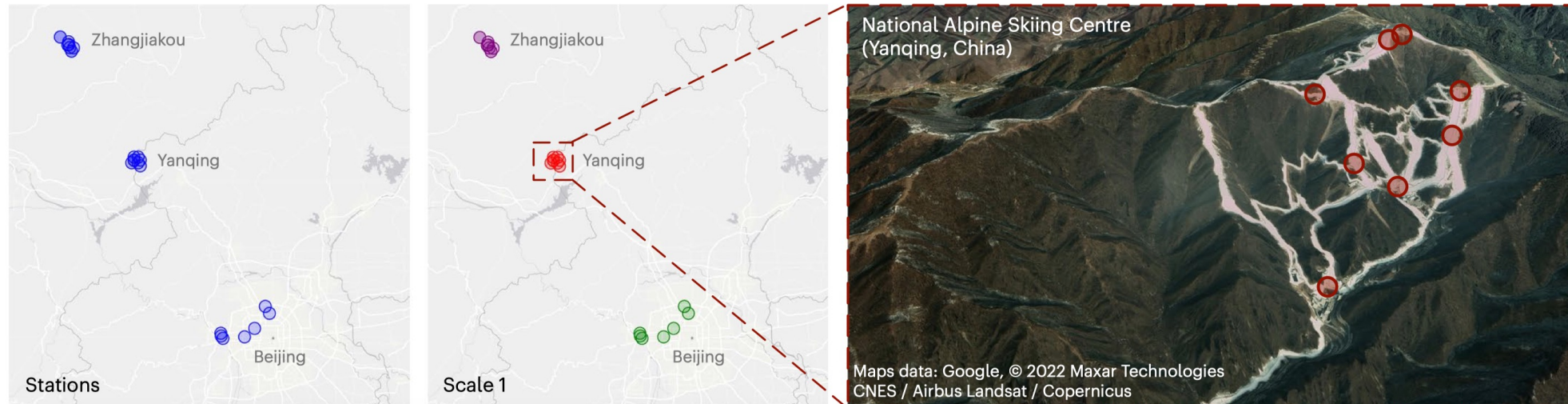
a Global dataset



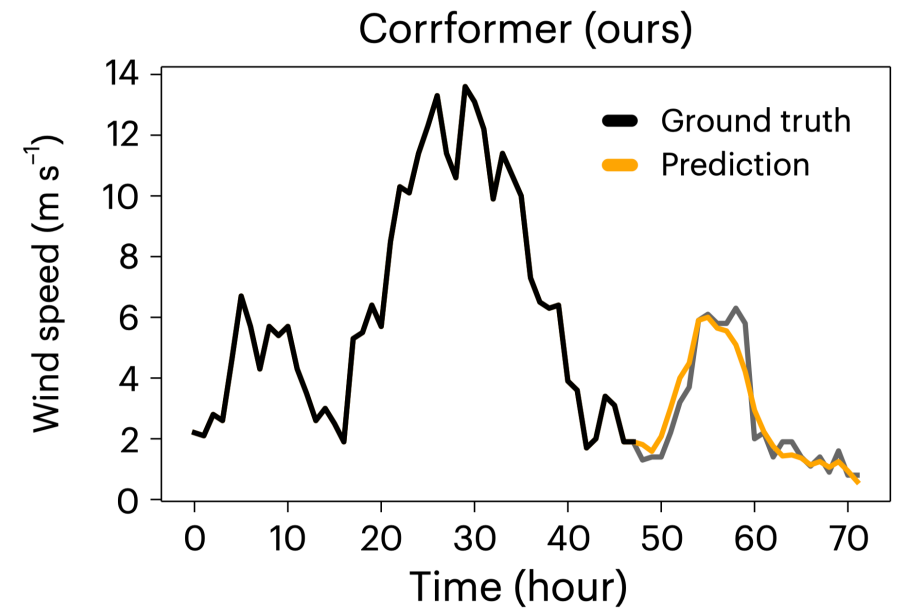
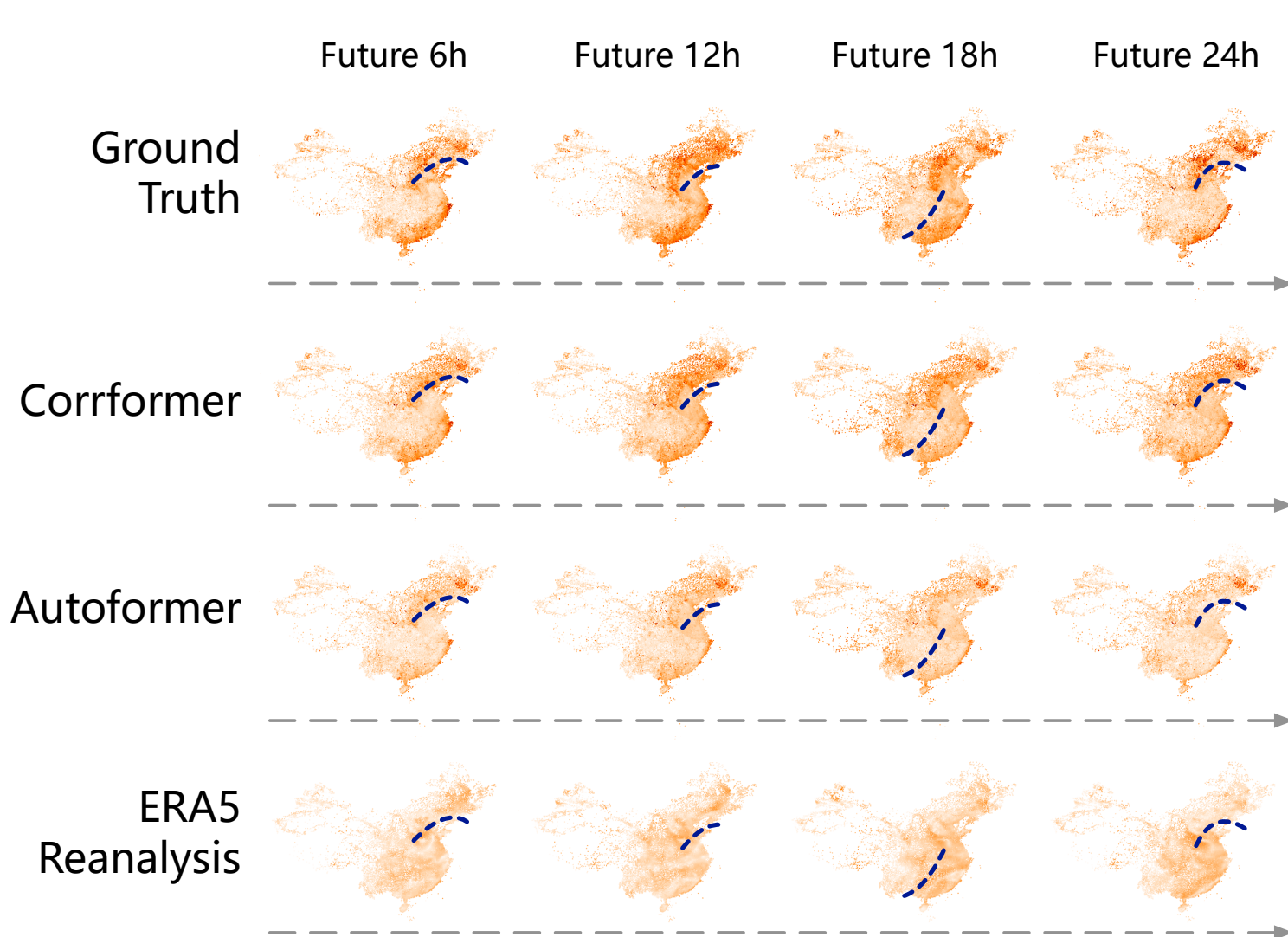
b Regional dataset



c Olympics dataset



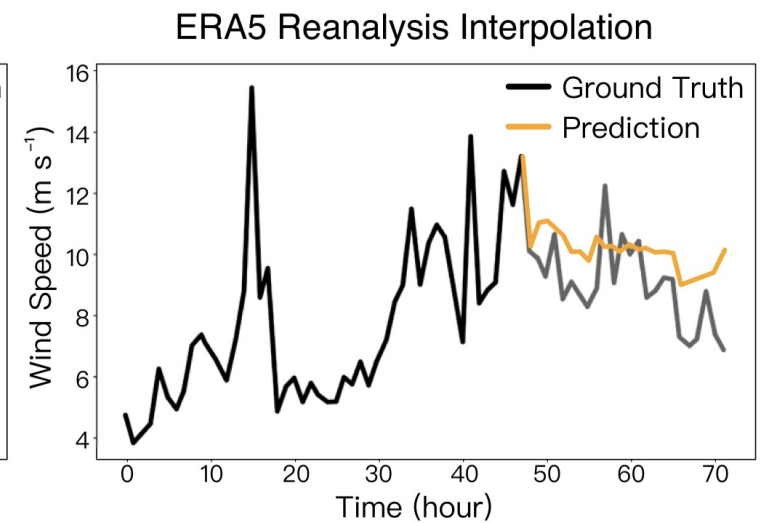
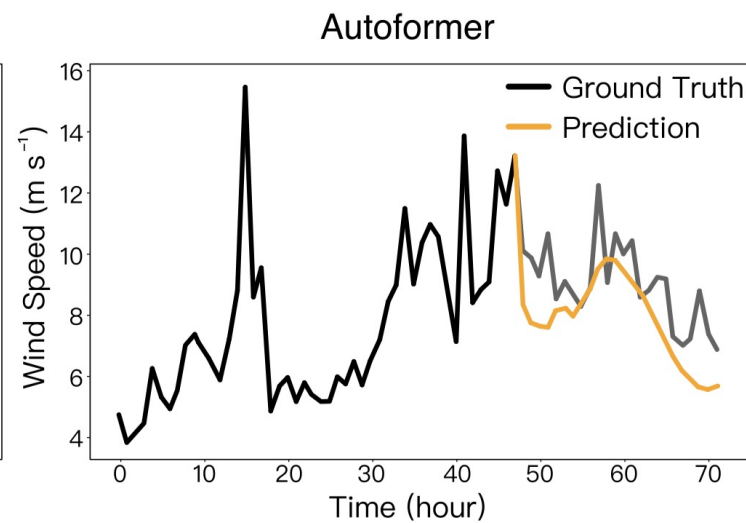
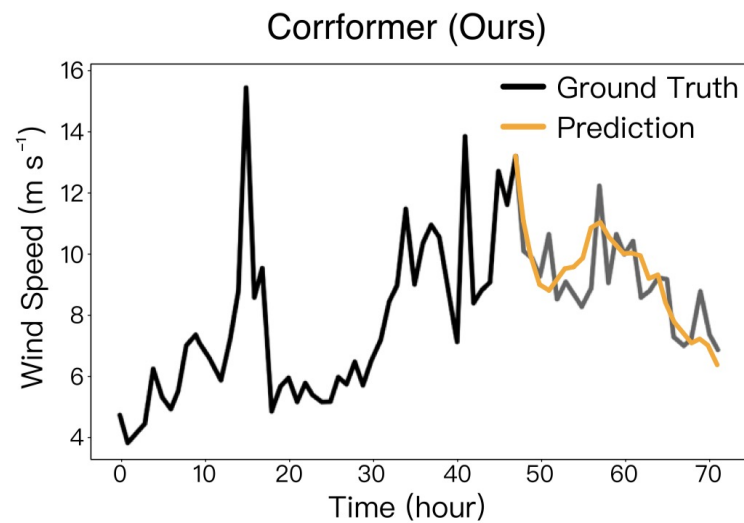
Regional Wind Forecasting



- ✓ Overall **34,040** Stations
- ✓ Training **one day** on **single GPU**, real-time forecast in **1 second**
- ✓ Achieve the **state-of-the-art**

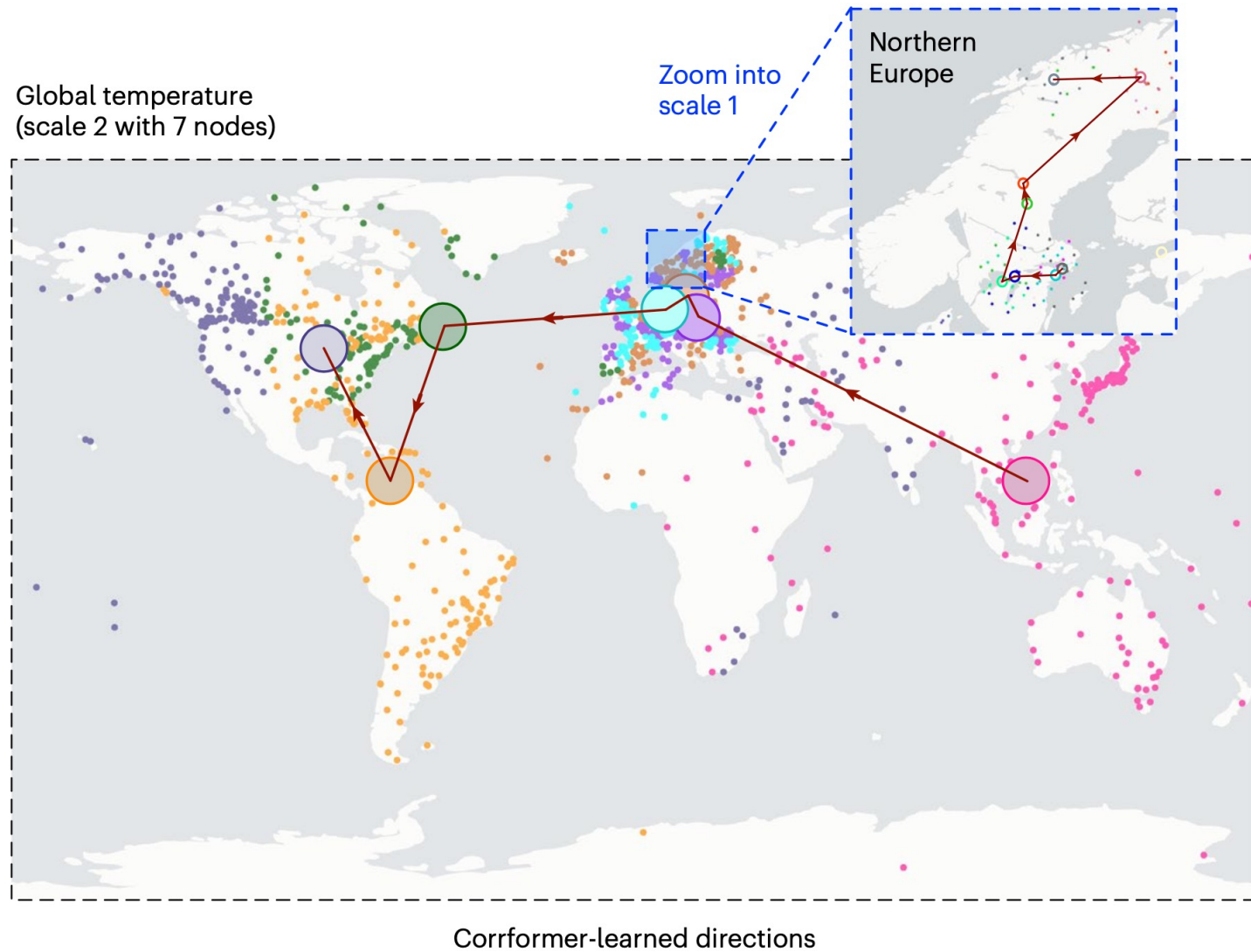
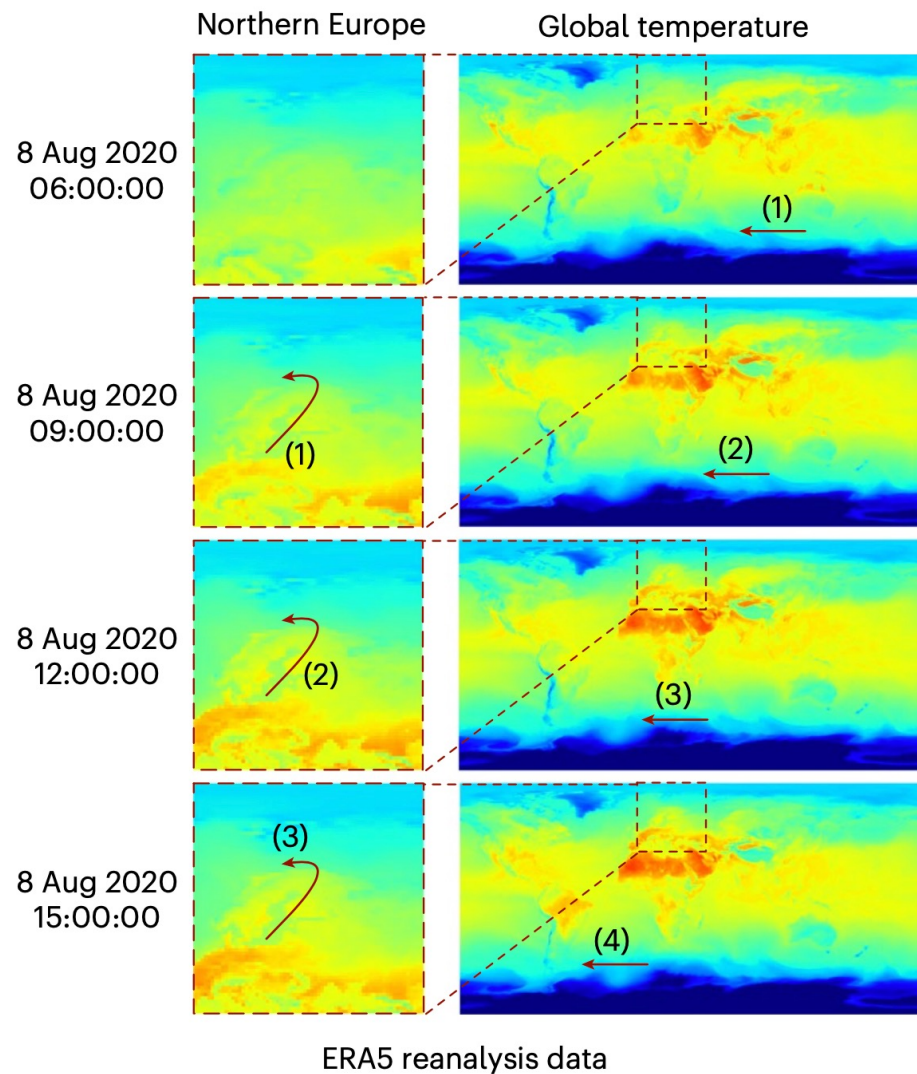
Olympics Wind Forecasting

- Alpine weather is seriously affected by **chaos effects**
- Provide skillful weather forecasting for **extremely complex topography**



Interpretable Forecasting for Global Weather

Propagation direction analysis



Open Source

The screenshot shows the GitHub repository page for 'Corrformer'. At the top, it indicates the repository is public and has 8 watchers, 18 forks, and 104 stars. The main branch is 'main'. A recent commit by 'wuhaixu2016' is shown, updating the README.md file. The file list includes folders like 'data_provider', 'exp', 'layers', 'models', 'scripts', and 'utils', along with files like '.gitignore', 'LICENSE', 'README.md', 'requirements.txt', and 'run.py'. The 'About' section provides a link to a Nature Machine Intelligence article. The 'Releases' section shows the latest version of 'Corrformer' published on April 26. The 'Packages' section indicates no packages are published. The 'Languages' section shows Python at 96.7% and Shell at 3.3%. The 'Suggested Workflows' section includes a 'Pylint' workflow.

| File/Folder | Commit Message | Time Ago |
|------------------|--------------------|--------------|
| data_provider | init | 7 months ago |
| exp | Update exp_main.py | 5 months ago |
| layers | init | 7 months ago |
| models | init | 7 months ago |
| scripts | init | 7 months ago |
| utils | init | 7 months ago |
| .gitignore | Initial commit | 8 months ago |
| LICENSE | Initial commit | 8 months ago |
| README.md | Update README.md | 2 weeks ago |
| requirements.txt | init | 7 months ago |
| run.py | init | 7 months ago |

Corrformer (Nature Machine Intelligence)

In this [paper](#), we present Corrformer with the Multi-Correlation mechanism, which can unify the temporal auto-correlation and spatial correlation in a learned multiscale tree structure.

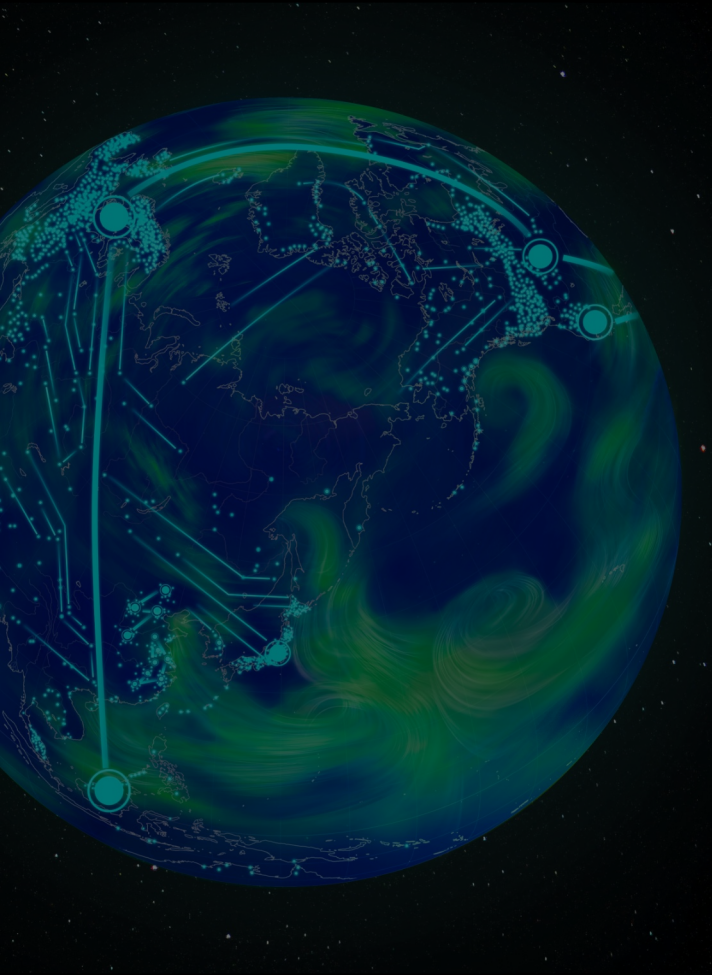
- Corrformer reduces the canonical double quadratic complexity of spatiotemporal modeling to linear in spatial modeling and log-linear in temporal modeling, **firstly achieving collaborative forecasts for tens of thousands of stations within a unified deep model.**
- Corrformer can generate interpretable predictions based on inferred propagation directions of weather processes, **facilitating a fully data-driven AI paradigm for discovering insights for meteorological science.**
- Corrformer yields **state-of-the-art forecasts on global, regional and citywide datasets** with high confidence, beating classical statistical methods, latest deep models, and comparing favorably to numerical methods in near-surface forecasting.



CODE OCEAN

<https://github.com/thuml/Corrformer>

<https://codeocean.com/capsule/0341365/tree/v1>



Thank You

wuhx23@mails.tsinghua.edu.cn