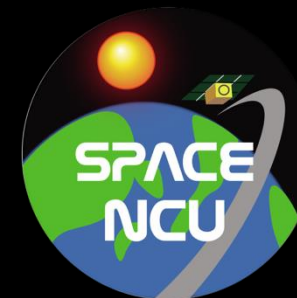


Space Weather Prediction with Graph based Neural Network

2 Sep. 2026, 40th Conference on Weather Analysis and Forecasting

Chung-Yu Shih 施重宇

Prof. Cissi Ying-tsen Lin 林映岑



Communications and Intelligence for the Atmosphere Lab

Department of Space Science and Engineering

National Central University



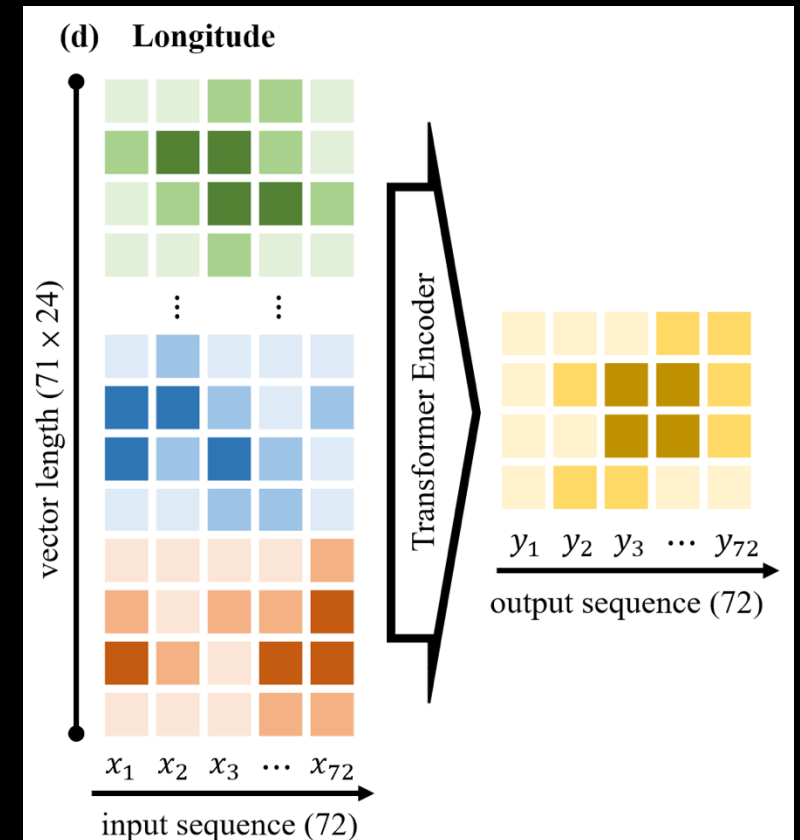
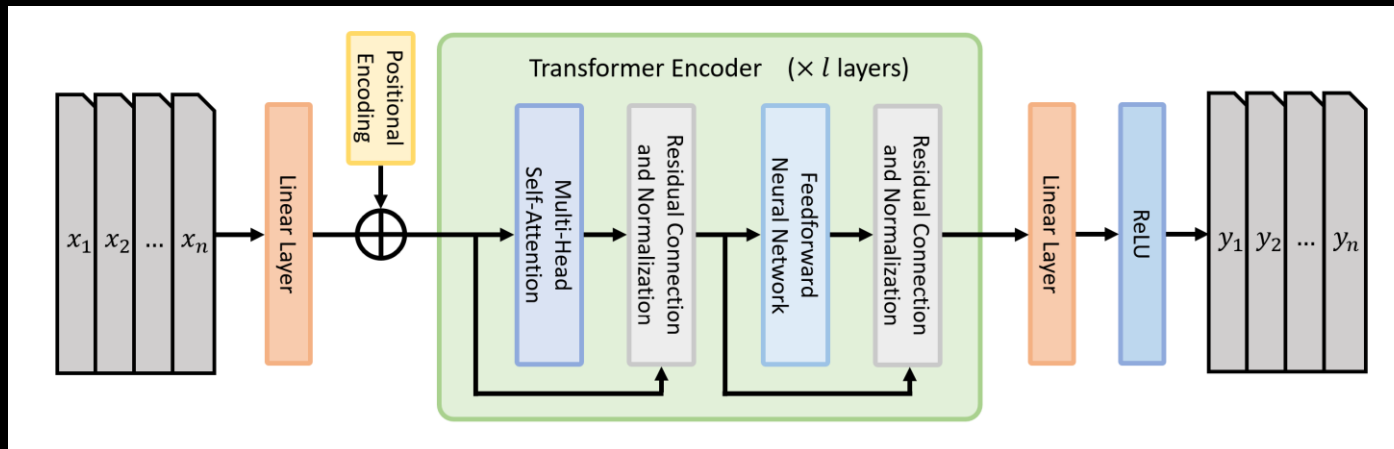
Simulation of Space Weather

- Physical-driven model:
 - Multi-scale interactions (for time and space).
 - Need lots of models coupling (fluid dynamics, rarefied gas dynamics, magnetohydrodynamics).
 - Bad boundary condition and numerical limitation.
 - Time consuming (solving too many model).
- Data-driven (Empirical) model:
 - Over smoothing, no detail and bad prediction at special event.

Model between Physical and Data might be a solution!

About GIM Prediction

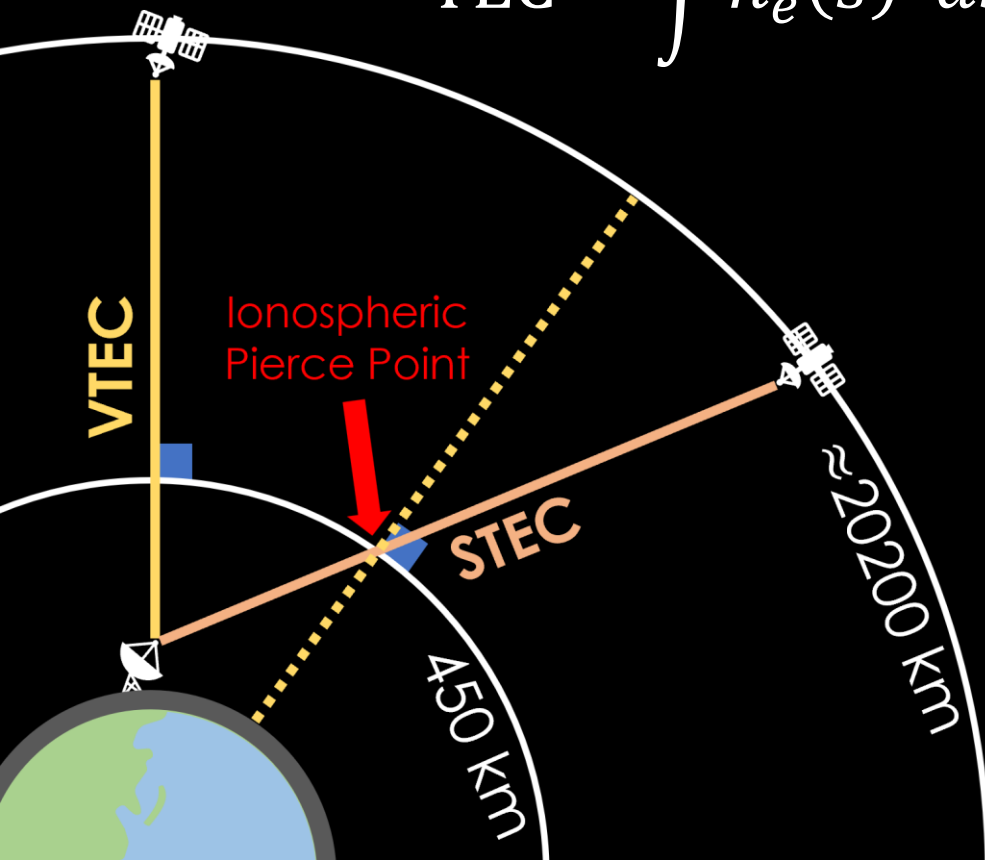
- 2024, our team publish the article on AGU Space Weather Forecasting of global ionosphere maps with multi-day lead time using transformer-based neural networks.
- Based on Transformer Encoder.
(Input: 24 hours history data.
Output: predict future 24 hours.)



TEC and GIM

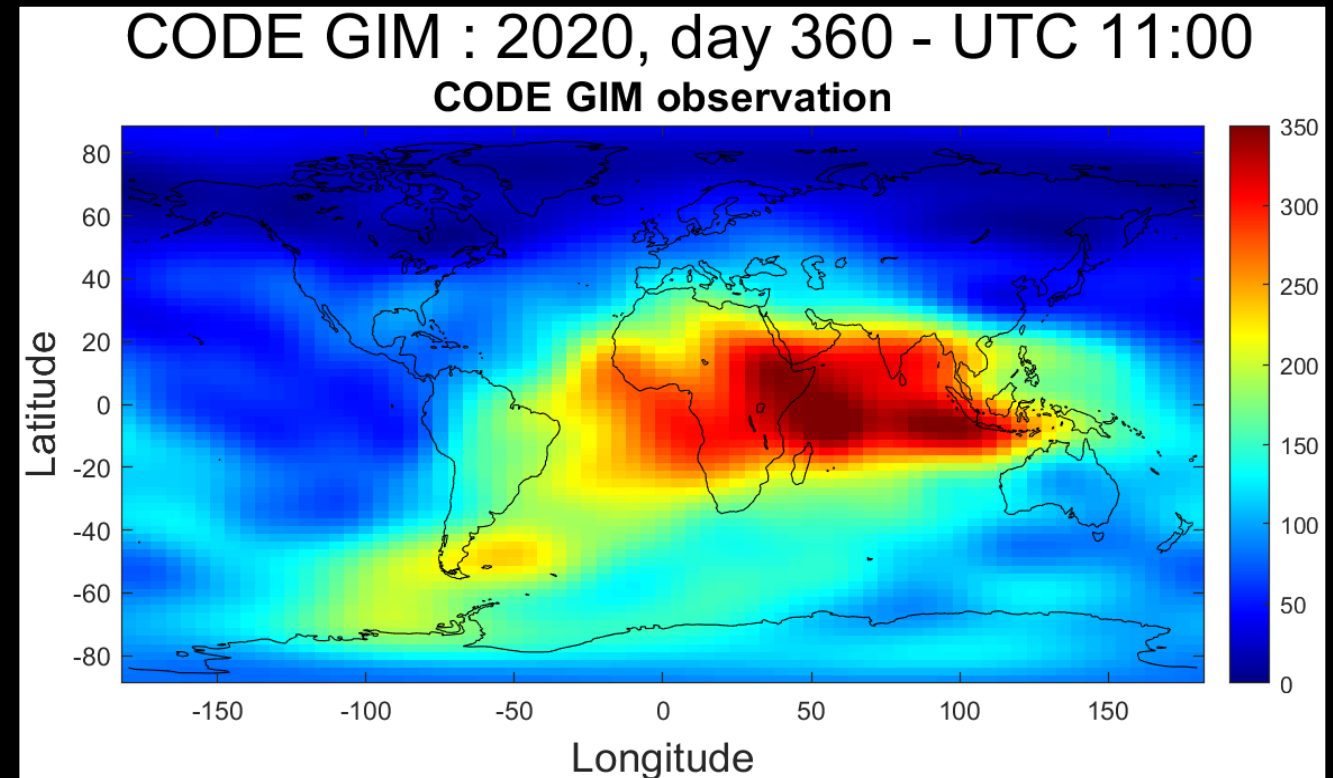
Total Electron Content (TEC)

$$\text{TEC} = \int n_e(s) ds$$



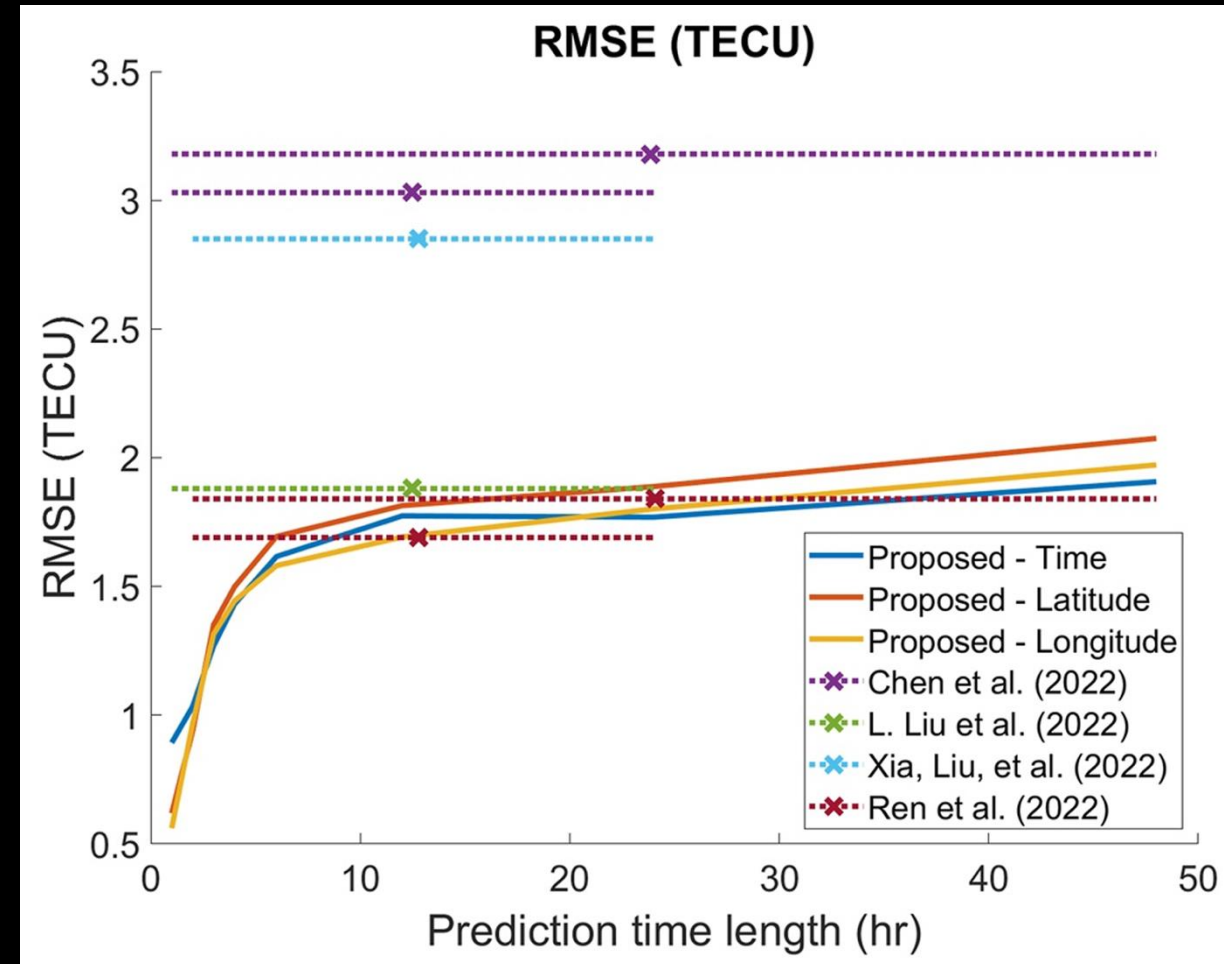
Center for Orbit Determination in Europe (CODE), Global Ionosphere Map (GIM) data.

- Latitude: $-87.5^\circ\text{N} \sim 87.5^\circ\text{N}$ (2.5° , 71 grids).
- Longitude: $-180^\circ\text{E} \sim 180^\circ\text{E}$ (5° , 73 grids).
- Train: 2018, 2019, Test: 2020, 2021.



Previous Works

- Training: 2018 ~ 2019.
Testing: 2020 ~ 2021.
- Special input orientation make the prediction accurate.
- Experiment prove that solar indexes and time information are useless.
- At that time, we want the model **“small and high accuracy”**.



Problem 1: Baseline Method

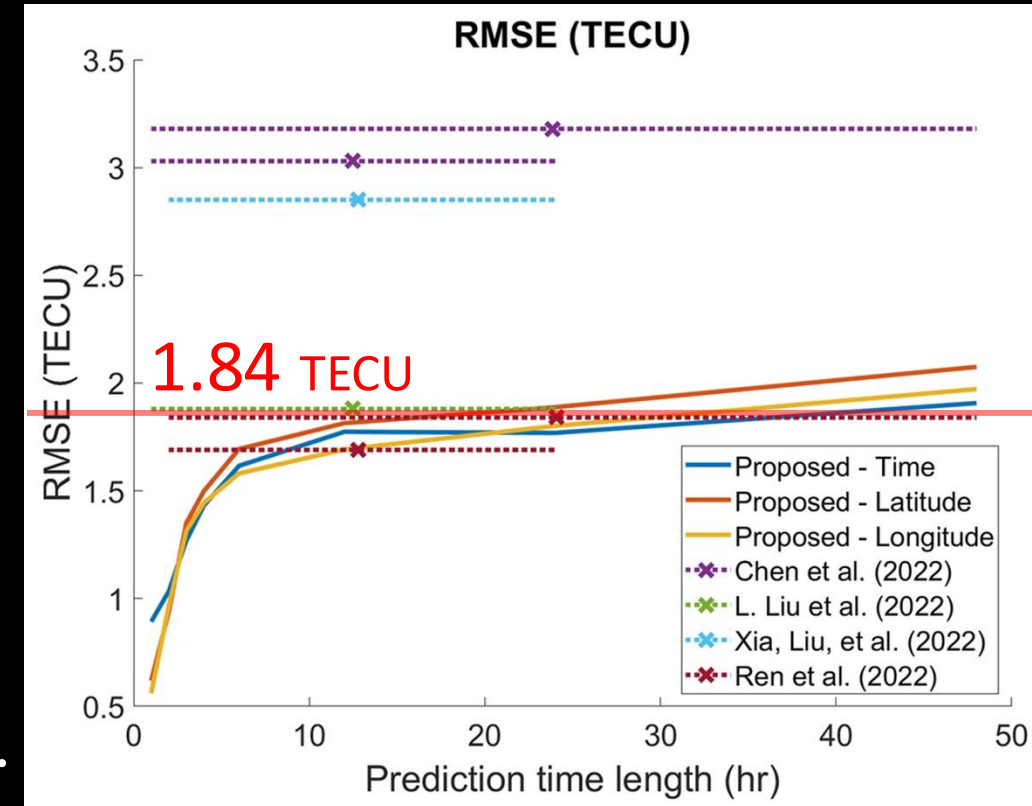
If we want to find the GIM at

2026/9/3 UT12:00

Persistence Method:

(2020 RMSE: 1.84 TECU, 2025 RMSE: 6.81 TECU)

- Use the data 24 hours ago as the prediction. (Take the data at 2026/9/2 UT12:00 GIM as the prediction)
- Base on daily period, simple but effective.



Problem 2: Type of Model

One-shot Forecasting Model (direct multi-step)

$$f_0(\dots, x_{n-1}, x_n) = [x_{n+1}, x_{n+2}, x_{n+3}, \dots]$$

- **High accuracy**, since different forecast horizon have independent relation.
- No error propagation through time.

Iterative Forecasting Model (autoregressive, recursive)

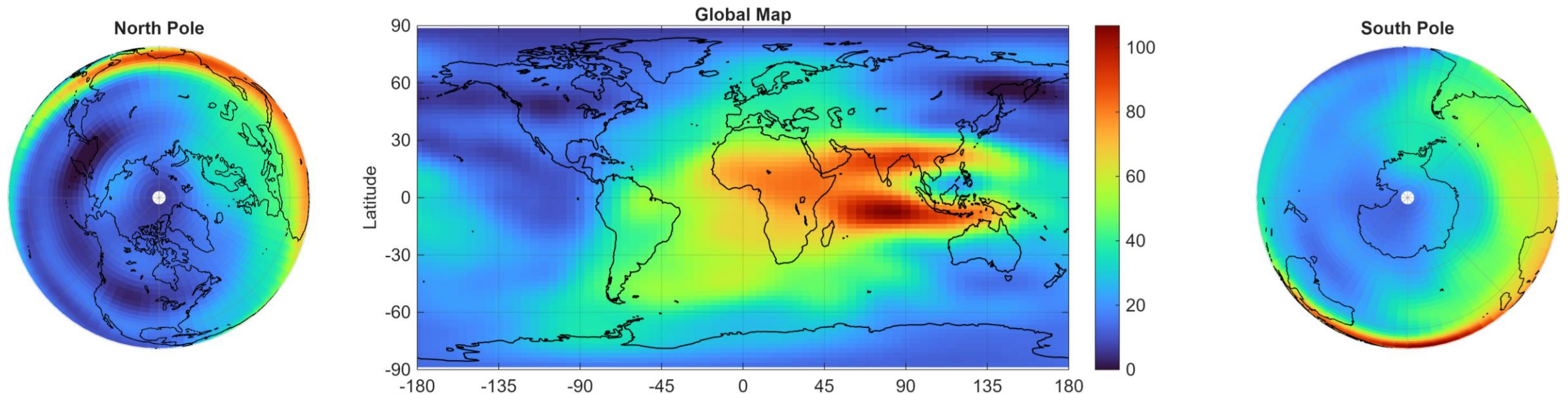
$$f_I(x_n) = x_{n+1}, \quad f_I(f_I(f_I(x_n))) = x_{n+3}$$

- Ensure the model remains **consistent over time**, scalability and robustness.
- **Flexibility** of forecast horizon and model don't need warm up.
- **Markov property**, good for data assimilation.

Problem 3: Mesh System

- CODE GIM use Latitude Longitude grid system ($2.5^\circ \times 5^\circ$).
- **Spatial nonuniform**, polar region has many grid. (71x73 grids).

CODE Global Ionosphere Map at 2025/2/9 UT 12:00



Ultimate Goal!

- Prediction of **Global** total electron content.
- Build an **iterative forecasting** model.
 - Forecast can **start from any time**.
 - Reduce **error propagation** through time.
- The model can work at any parts of **solar cycle**.

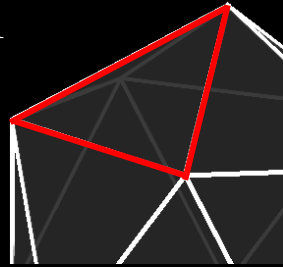
On going

- Include other related information, such as solar indexes, time, etc.
- Working with ensemble data assimilation system.
- Longterm (>24 hours) forecast **without over-smoothing**.

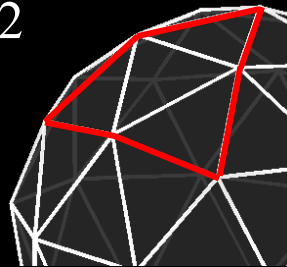
Triangular Mesh on Sphere

Subdivide **Icosahedral**
Spherical Mesh

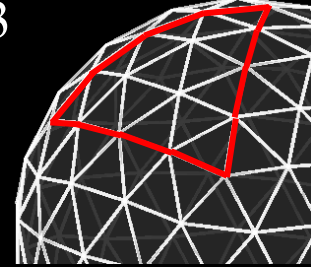
L1



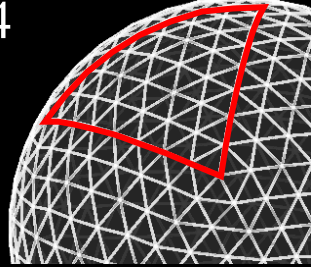
L2



L3



L4

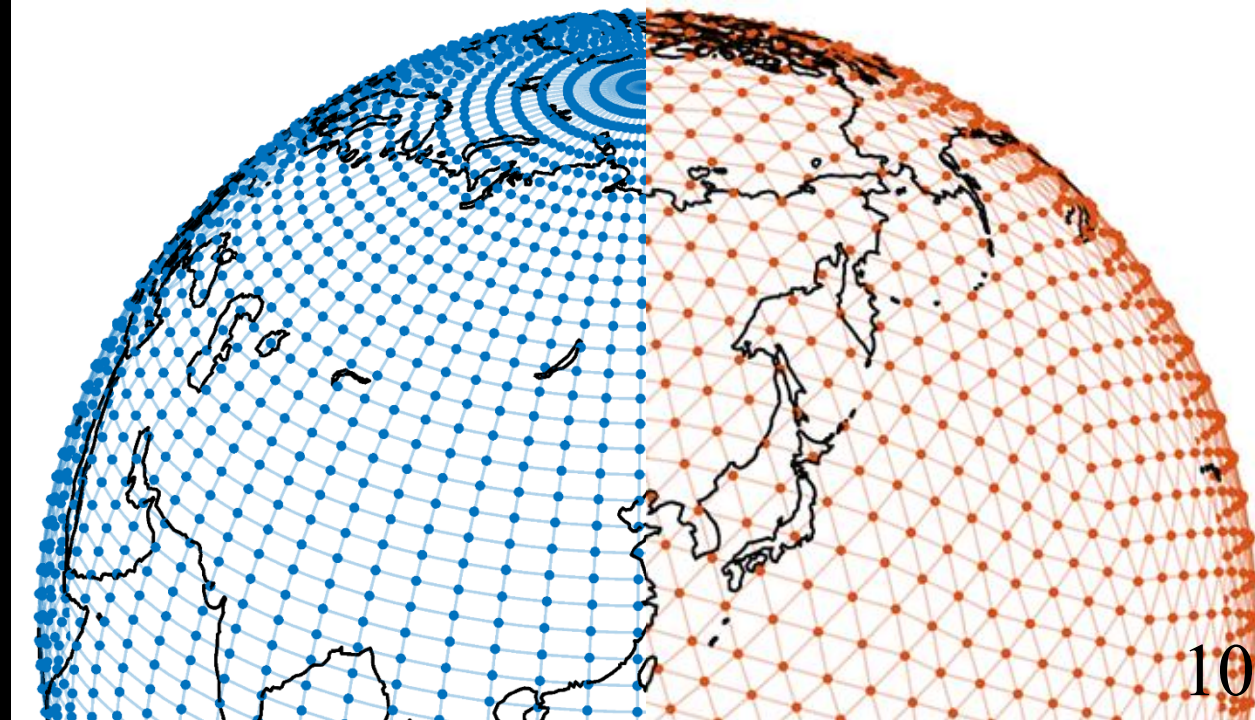


Level	Grid #	Min. Arc	Max. Arc	Avg. Arc*
1	12	63.435°	63.435°	7061.5 km
2	42	31.718°	36.000°	3769.1 km
3	162	15.859°	18.700°	1916.5 km
4	642	7.930°	9.444°	962.3 km
5	2562	3.965°	4.734°	481.7 km
6	10242	1.982°	2.369°	240.9 km
7	40962	0.991°	1.185°	120.5 km

*The average arc length take Earth radius as 6378.1 km.

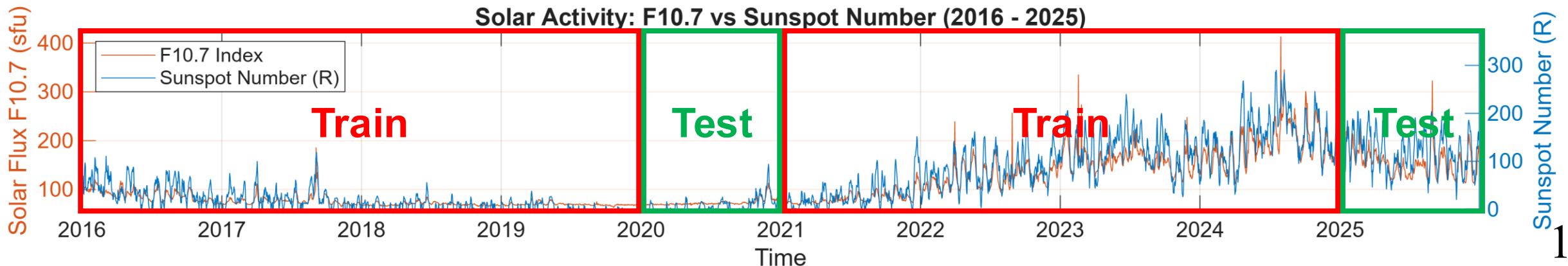
2.5°x5° Lati-Longi Mesh

Level 5 Triangular Mesh



Data & Data Preprocessing

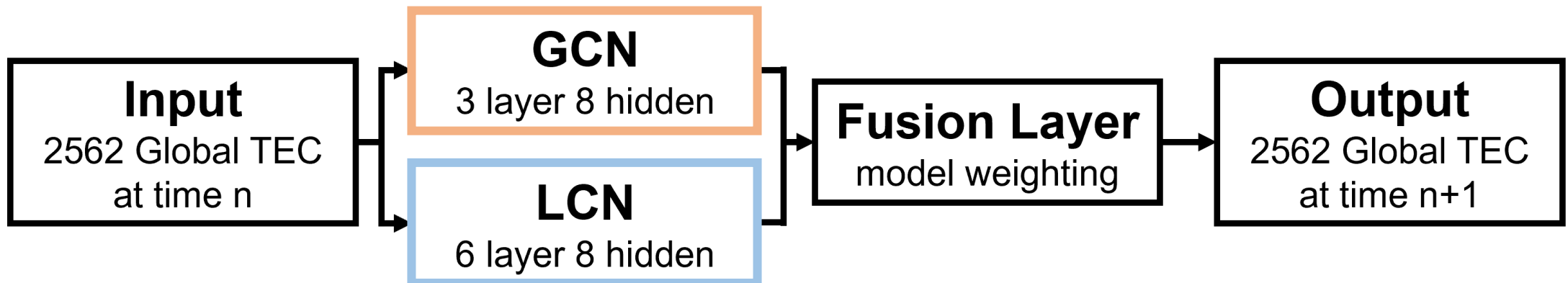
- Center for Orbit Determination in Europe (CODE) GIM data
 - Spatial resolution: Latitude 2.5° x Longitude 5° .
 - Temporal resolution: 1 hour.



Model Architecture

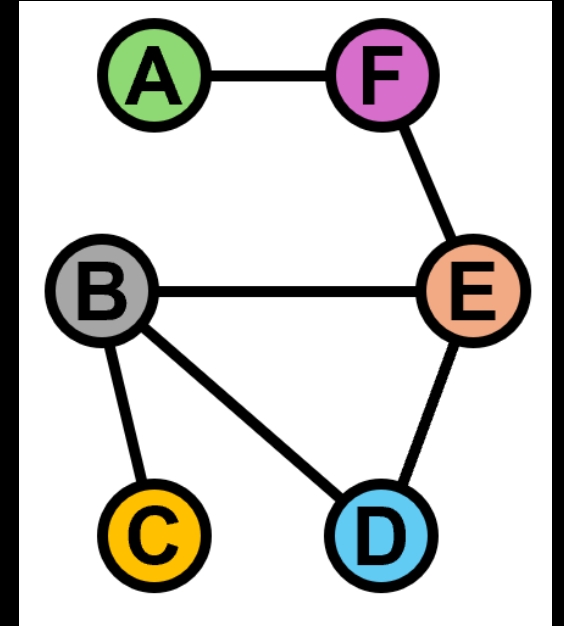
Combine two Graph-based Neural Network to simulate the traditional **diffusion and convection model** in computational fluid dynamics.

- **GCN**: Graph Convolutional Network.
- **LCN**: Locally Connected Network.
- Fusion Layer to connect two different type results.



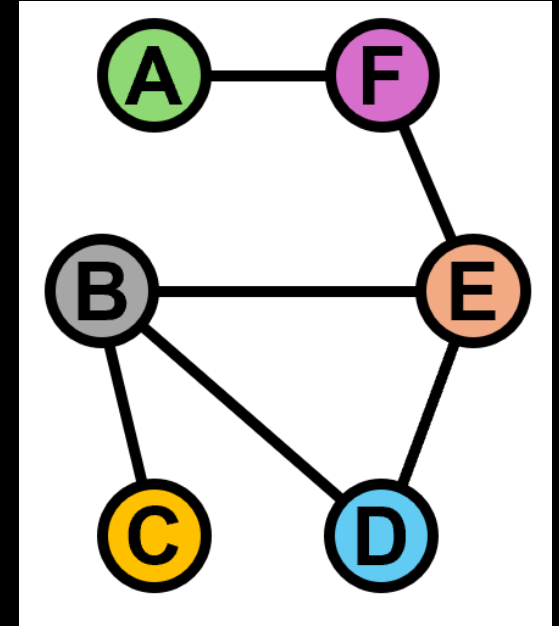
GCN for Diffusion

- GCN: Graph Convolutional Network
- CNN (Convolutional Neural Network) for graph.
- Update each nodes from the neighbors.
- Example: Update nodes B
 - Aggregate features from neighbors C, D, E. (take average etc.)
 - Combine with B's features. (summation)
 - Then use the trained matrix to update.
- Each nodes using same weight (matrix) to update, which means the relation of each nodes are equivalent.
- GCN is small and equivalent for each nodes.



LCN for Convection

- LCN: Locally Connected Network
- Degenerate of FNN (Feedforward Neural Network)
- Using the given graph to build adjacency matrix A .
- Take $A + I$ (self-loop) as the mask of matrix.
- Only train the black element!
(white element is always zero)
- With leaky ReLU as σ .
- **Why symmetric mask?**
Training efficiency!
(two element for one edge?)

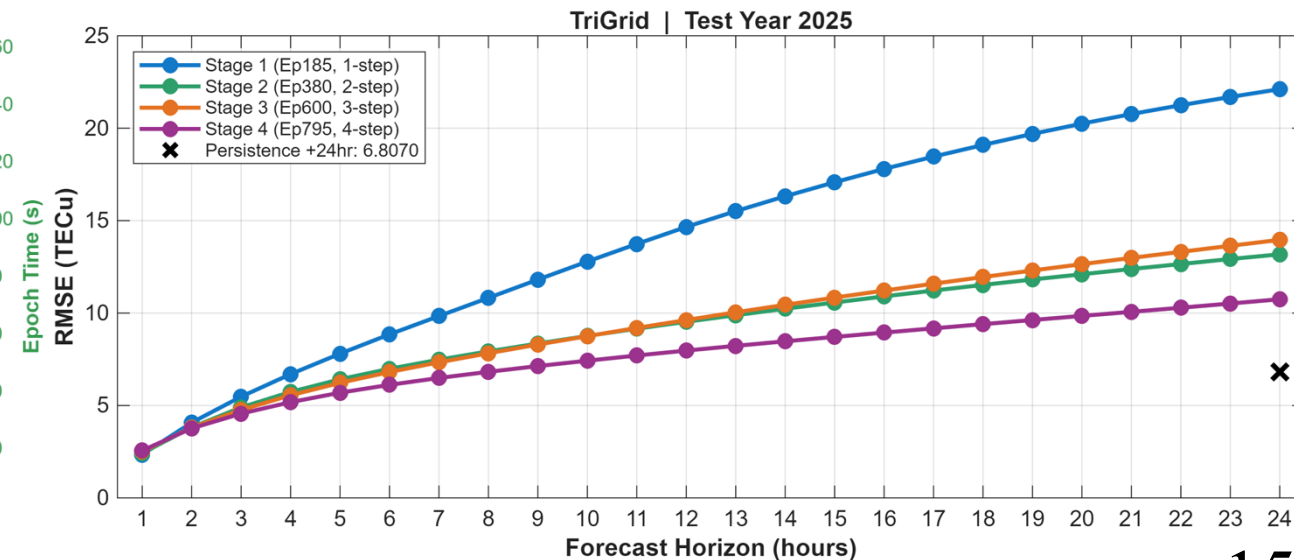
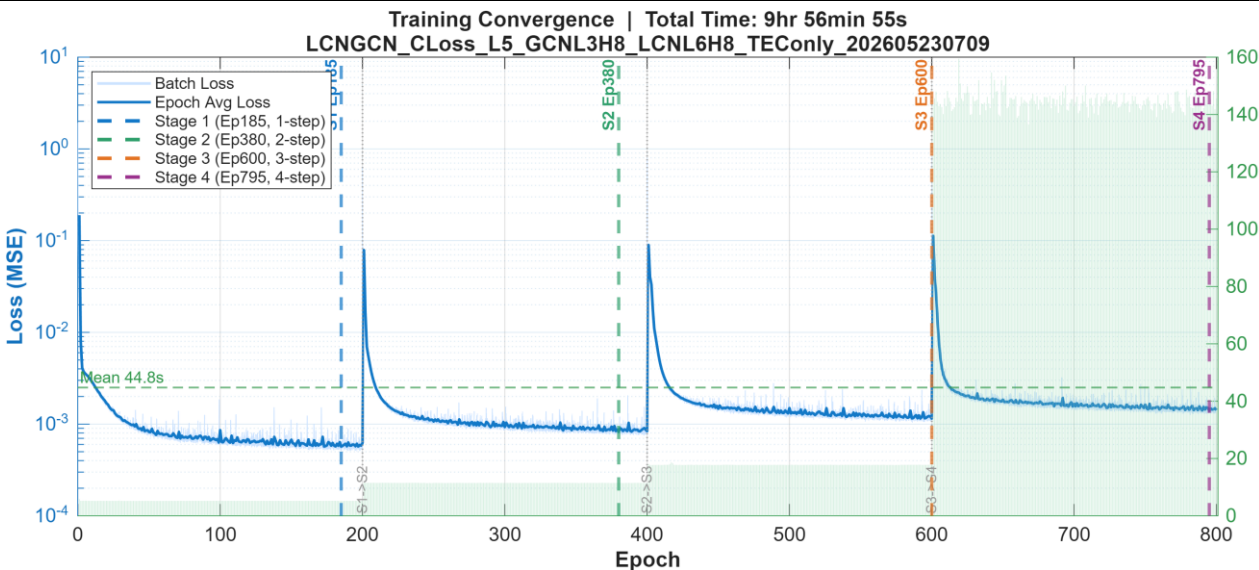


$$\begin{bmatrix} A_o \\ B_o \\ C_o \\ D_o \\ E_o \\ F_o \end{bmatrix} = \sigma \left(\begin{bmatrix} \text{mask} \end{bmatrix} \times \begin{bmatrix} A_i \\ B_i \\ C_i \\ D_i \\ E_i \\ F_i \end{bmatrix} + \begin{bmatrix} \text{self-loop} \end{bmatrix} \right)$$

The diagram illustrates the LCN operation. On the left, a column vector of output nodes $A_o, B_o, C_o, D_o, E_o, F_o$ is shown. This is equal to the sigmoid function σ applied to the product of a 6x6 mask matrix and a column vector of input nodes $A_i, B_i, C_i, D_i, E_i, F_i$. The mask matrix is symmetric and contains black squares at positions (i,j) where $i \neq j$ and (i,j) is an edge in the graph, and black squares on the diagonal representing self-loops. The input vector is also shown with colored circles corresponding to the nodes in the graph. The result of the multiplication and addition is a column vector of 6 black squares, representing the unactivated outputs before the sigmoid function.

Curriculum Learning

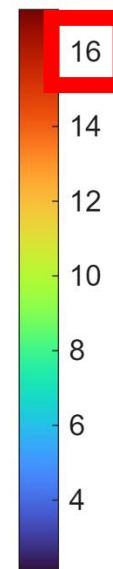
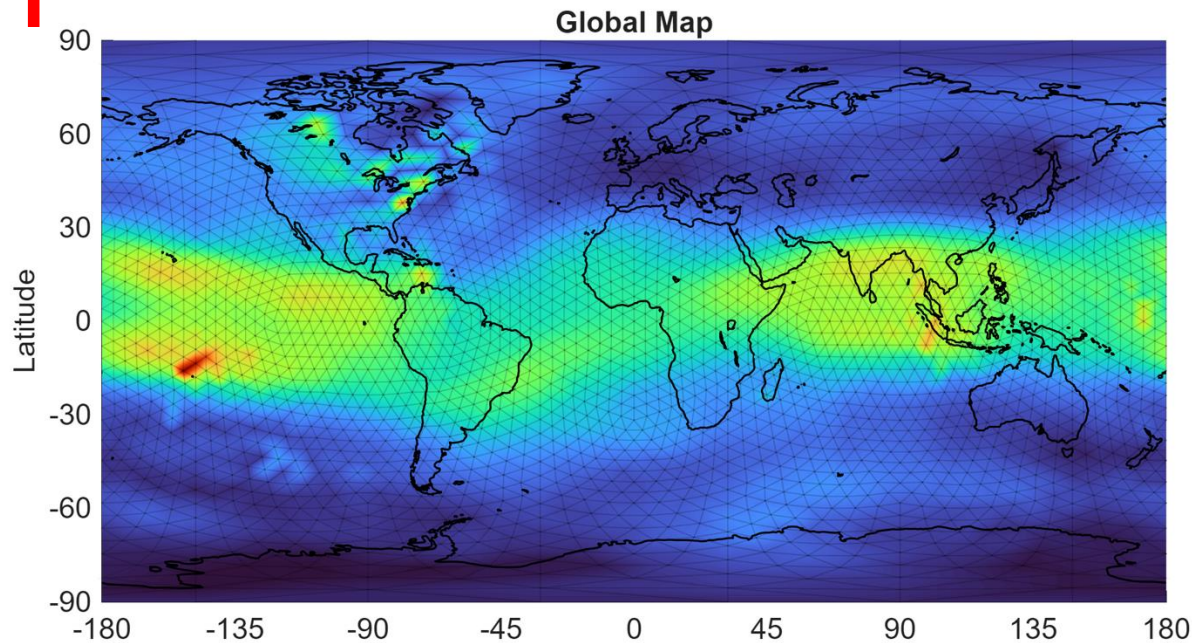
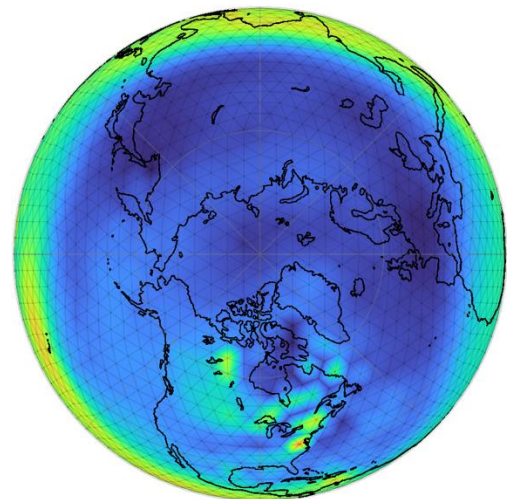
- To **reduce the error propagation**, we apply curriculum learning.
- New training stage we include **longer forecast horizon**.
- Speed up the training progress, backward propagation through time (BPTT) is expansive! (we only apply two stage training)



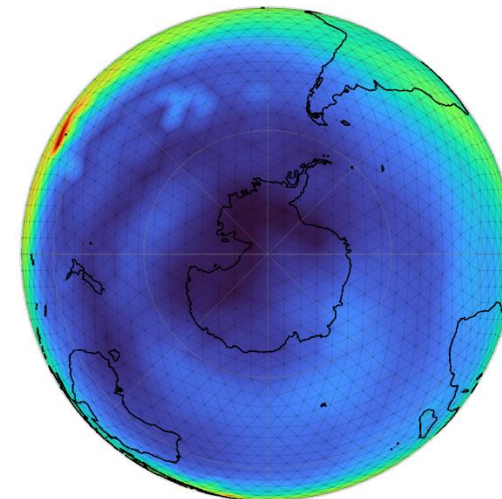
Stage 1

Stage 1 (Ep192) | +24hr RMSE | TriGrid L5 | 2020 [TECu]

North Pole



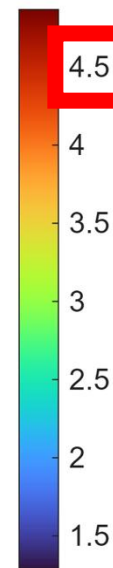
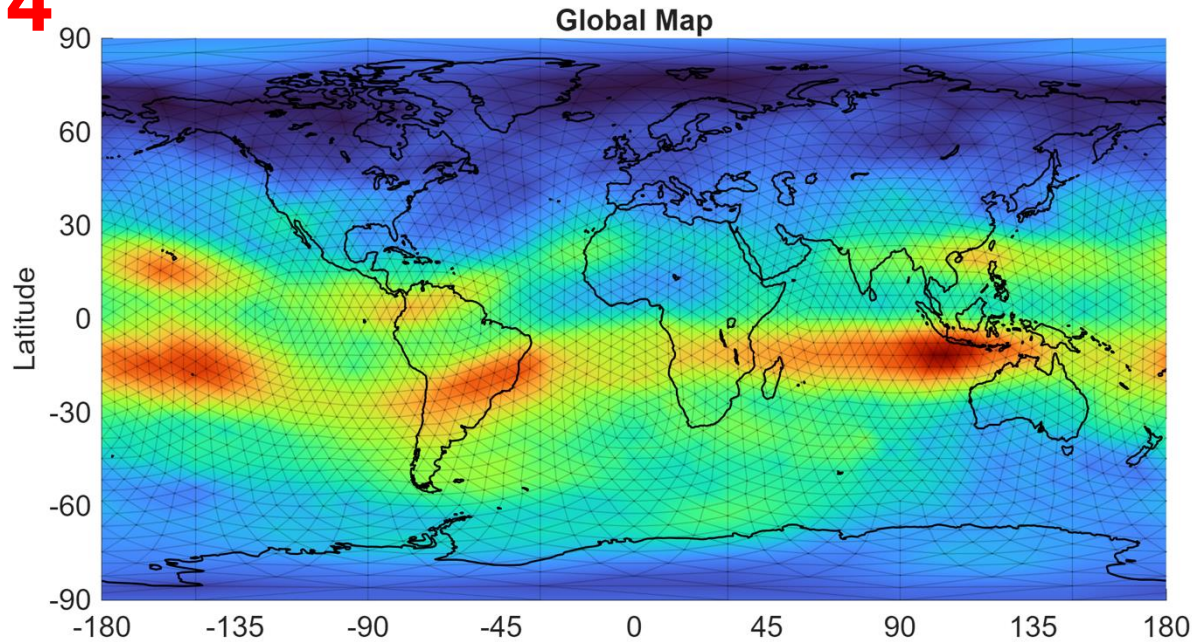
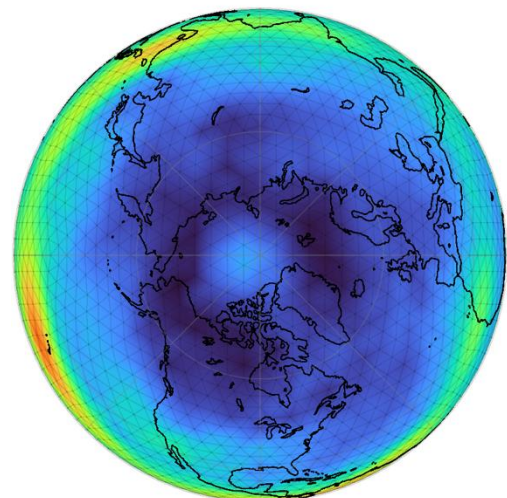
South Pole



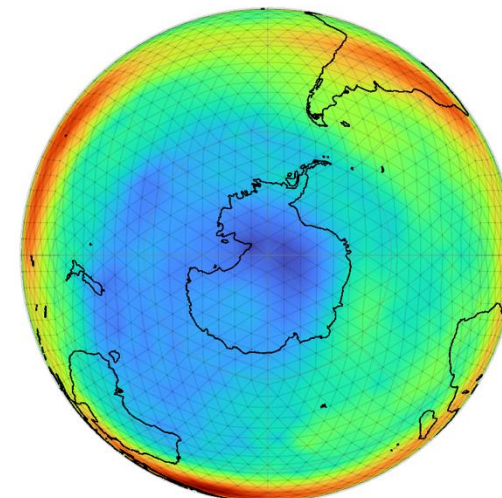
Stage 4

Stage 4 (Ep797) | +24hr RMSE | TriGrid L5 | 2020 [TECu]

North Pole



South Pole



Custom Loss Function

We customize the loss function to solve following problem:

- **Stability of long-term forecast**, prevent error propagation through time.
- **Temporal non-stationary**, can working in different solar activity.

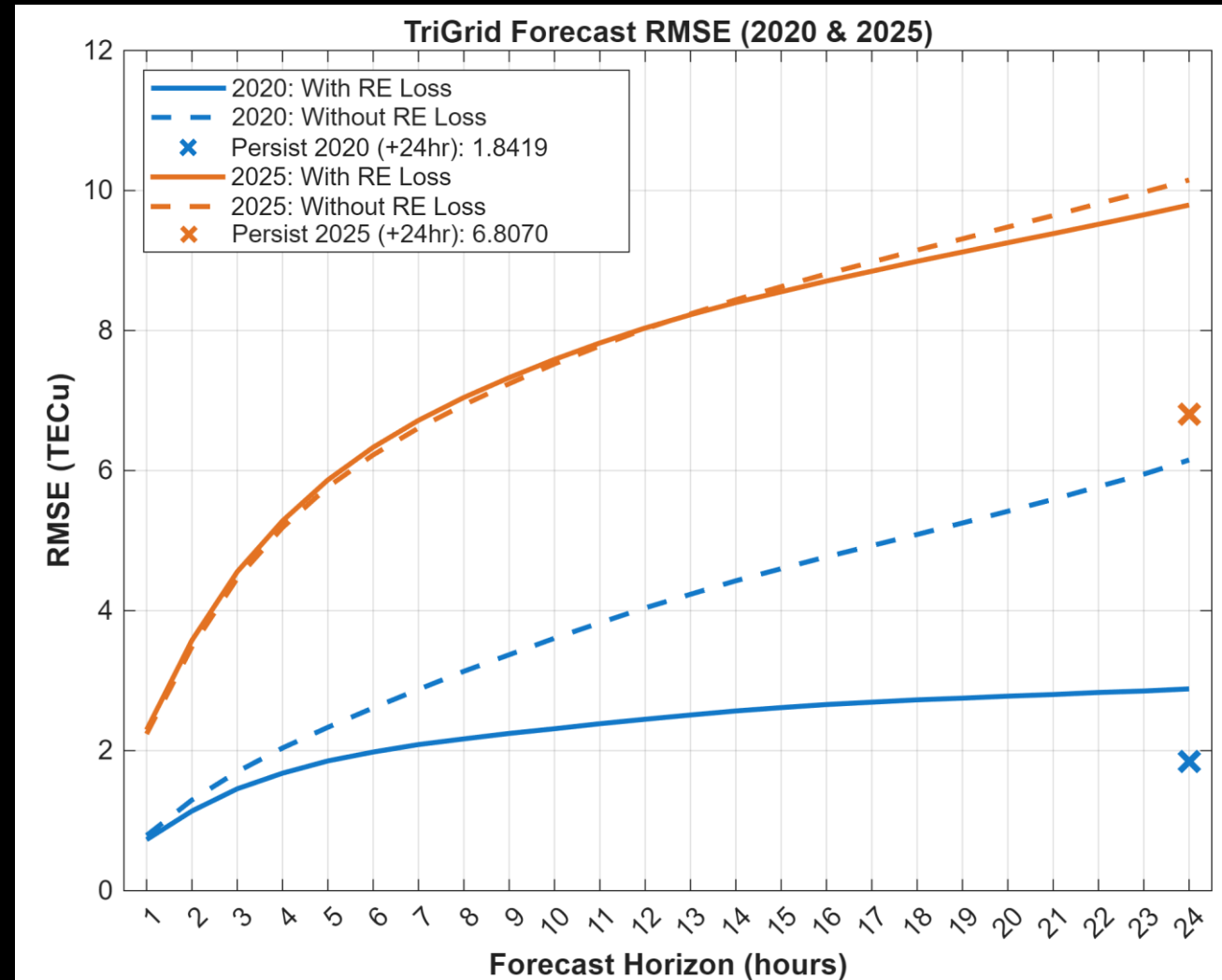
$$\text{Loss} = \sum_{t=1}^T \left(1 - \frac{t-1}{2(T-1)} \right) \left[\underbrace{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}_{\text{MSE (Mean Square Error)}} + w \cdot \underbrace{\frac{1}{n} \sum_{i=1}^n \left(\frac{\hat{y}_i - y_i}{y_i + \varepsilon} \right)^2}_{\text{MSPE with weight } w \text{ (Mean Square Percentage Error)}} \right]$$

Weighting for forecast horizon

Where T is the time steps of forecast, n is the number of spatial grid.
Also we take the $\varepsilon = 5$ TECU to avoid MSPE singularities.

With or Without MSPE

- Using 8 years of data for training cover one solar cycle.
- Without MSPE loss (relative error), the testing set with smaller TEC number have poor performance.
- With MSPE loss can improve the long-term forecast in 2020.



Conclusions and Problems

- Building an iterative forecast model is very hard!
- Triangular mesh can work.
- MSPE loss is useful, one model can work for solar minimum and maximum.
- Curriculum learning makes long term forecast better with only two steps forecast horizon.

Problems

- Need precursor or leading variable. (solar index?)
- CODE GIM is not ground truth, need data assimilation.

Thank you for your attention!

Chung-Yu Shih davidavidshih11@gmail.com



Communications and Intelligence for the Atmosphere Lab
Department of Space Science and Engineering
National Central University