**4th International Workshop on Waves, Storm Surges, and Coastal Hazards Incorporating the 18th International Waves Workshop**

CLSO-SZU
COLLEGE OF LIFE SCIENCES AND OCEANOGRAPHY
生命与海洋科学学院

September 2025 · Santander

# Physics-guided Deep Learning for Wave Modelling: From Single Points to Global Fields


[1] Haoyu Jiang (Haoyujiang@szu.edu.cn)

[2] Yuhao Song, Xinxin Wang, Tianxiang Gao, Yuan Zhang

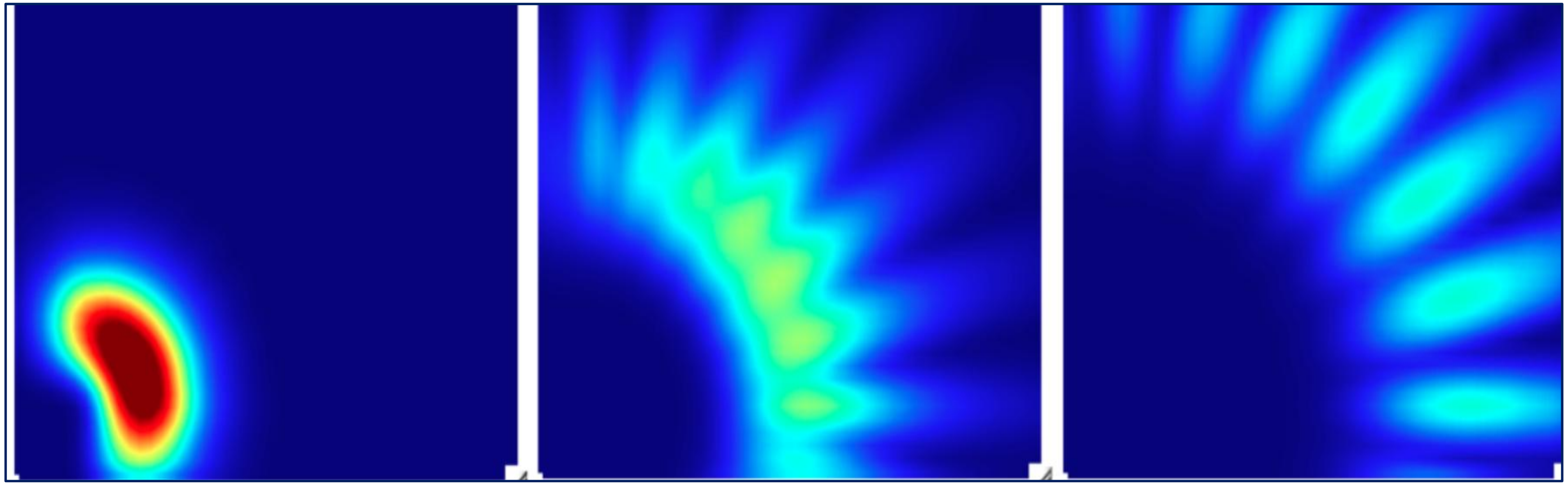
Department of Oceanography, Shenzhen University, Shenzhen, China

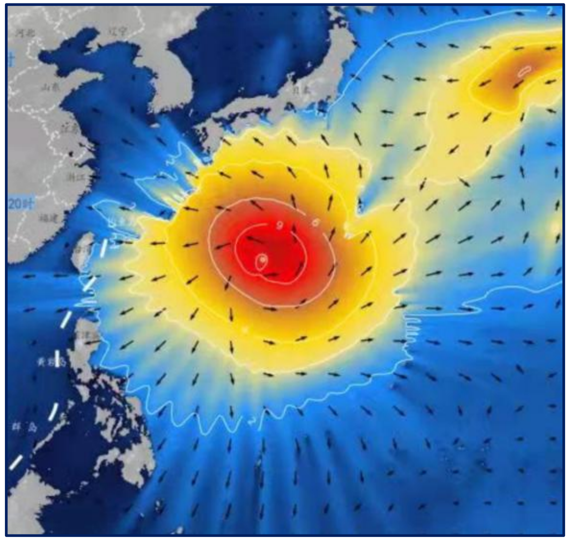
College of Marine Science and Technology, China University of Geosciences, Wuhan, China

# Shortcomings of numerical wave model

- Computationally expensive ➞ challenging for time-sensitive and resource-constrained scenarios.
- Accuracy is somewhat limited by incomplete physical representations and numerical effects, e.g., Garden Sprinkler Effect (GSE).

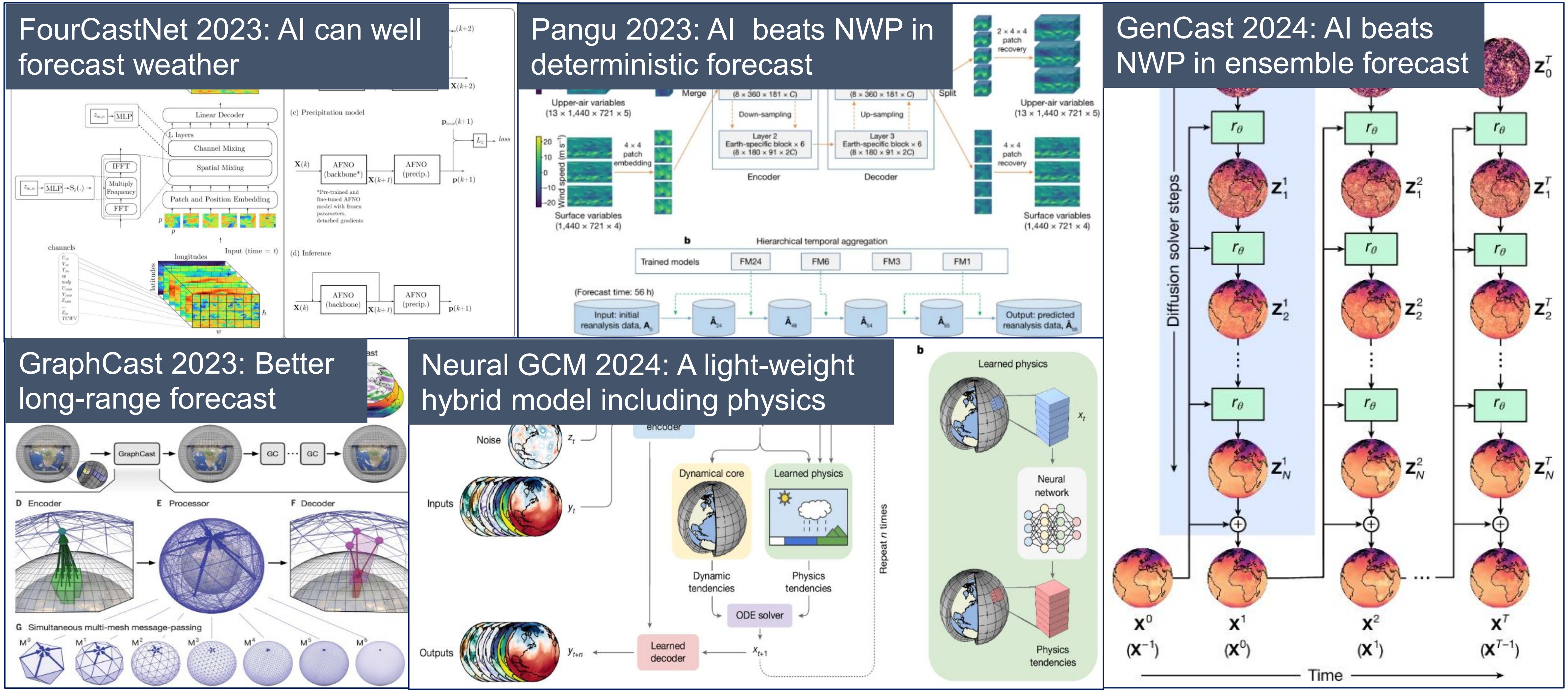
$$\frac{\partial N}{\partial t} + \nabla_x \cdot \dot{\mathbf{x}} N + \frac{\partial}{\partial k}\dot{k}N + \frac{\partial}{\partial \theta}\dot{\theta}N = \frac{S}{\sigma}$$





# A new direction: AI-based model

- AI-based weather forecasting models have achieved success : Similar (or even better) accuracy than numerical models with much lower computational costs



# AI for wave modelling

In recent years, many studies used deep learning for time series prediction of ocean waves:

Fan et al. 2020, Ni & Ma 2020, Huang & Dong 2021, Gao et al. 2021, Zhou et al. 2021, Feng et al. 2022, Song et al. 2023, Minuzzi & Farina 2023 ,Chen et al. 2023,

"DL can capture the nonlinear variability of wave, so they can be used for wave forecasting…" , **but, really ?**

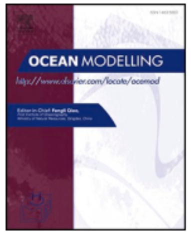
Ocean Modelling 189 (2024) 102364

ELSEVIER

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod




Research article

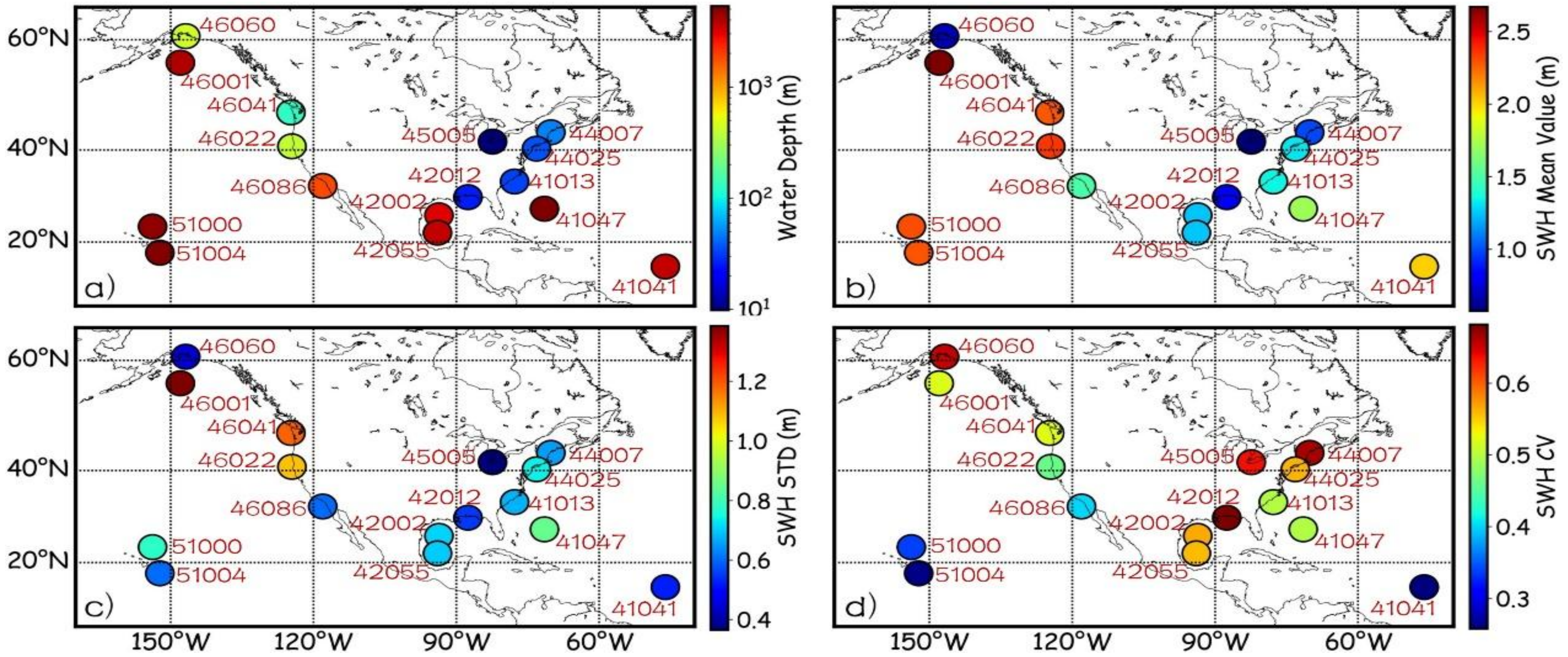
Comment on papers using machine learning for significant wave height time series prediction: Complex models do not outperform auto-regression

Haoyu Jiang [a,b,c,*], Yuan Zhang [a], Chengcheng Qian [a,b,d], Xuan Wang [a,e]

[a] Hubei Key Laboratory of Marine Geological Resources, China University of Geosciences, Wuhan, PR China
[b] Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao, PR China
[c] Shenzhen Research Institute, China University of Geosciences, Shenzhen, PR China
[d] Department of Information and Files, North China Sea Marine Forecasting Center of State Oceanic Administration, Qingdao, PR China
[e] School of Physics and Electronic Information, Weifang University, Weifang, PR China

# NDBC Buoy Data

- 16 NDBC buoys with different water depths, different regions, different wave climate properties.




**2016: Model training | 2017: Model Testing**

# Selected machine learning models

➢ Five models were selected:

| Liner Auto-regression (AR) | XGBoost (XGB) | Fully-connected ANN | LSTM | WaveNet |
|---|---|---|---|---|
| Simplest machine learning model | Widely-used tree model | Simplest deep learning model | Typical recurrent neural network | 1-D convolutional module |

**Input Length**

**6h/24h**

**Forecast Length**

**1/3/6/12/18/24/36/48/72h**

a)

| T-*m* | ··· | T-2 | T-1 | T0 | T+1 | T+2 | ··· | T+*n* |
|---|---|---|---|---|---|---|---|---|
| SWH | SWH | SWH | SWH | SWH | SWH | SWH | SWH | SWH |

**m-h input**
**n-h forecast**

b)

| T-24 | ··· | T-2 | T-1 | T0 | T+1 | T+2 | ··· | T+6 |
|---|---|---|---|---|---|---|---|---|
| SWH | SWH | SWH | SWH | SWH | SWH | SWH | SWH | SWH |

**24-h input**
**6-h forecast**

c)

| T-6 | ··· | T-2 | T-1 | T0 | T+1 | T+2 | ··· | T+24 |
|---|---|---|---|---|---|---|---|---|
| SWH | SWH | SWH | SWH | SWH | SWH | SWH | SWH | SWH |

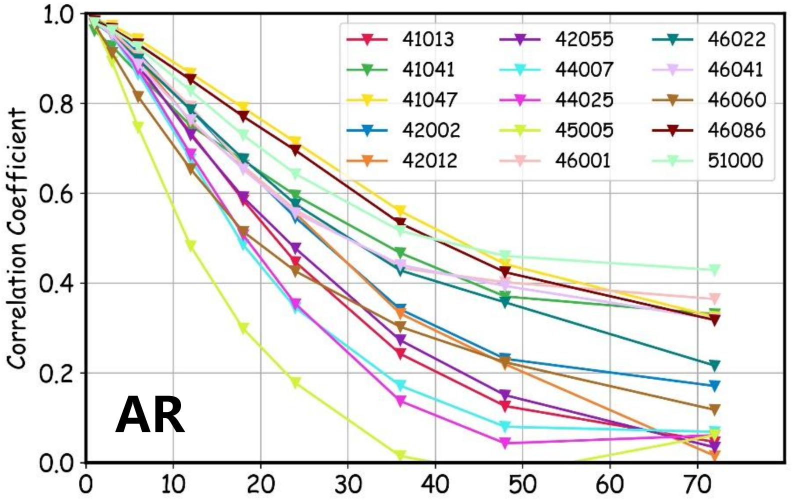
**6-h input**
**24-h forecast**

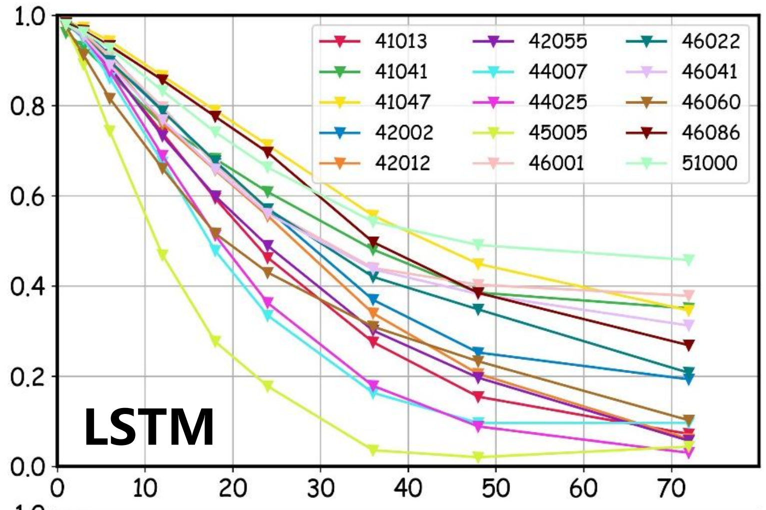
| Model Input | Data Not Used | Model Output |
|---|---|---|

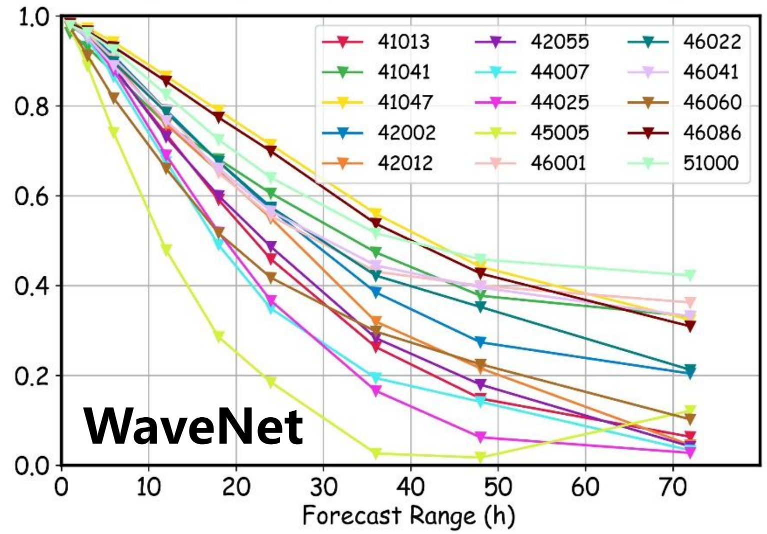
**Schematic of I/O**

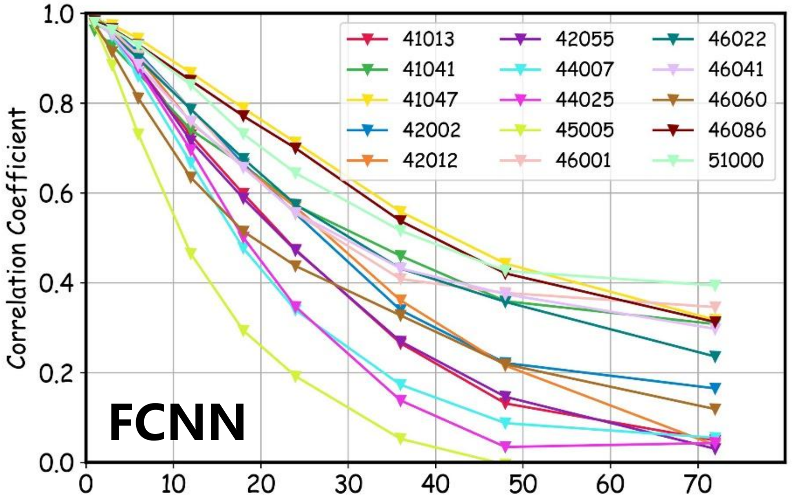
# Time series forecast is probably not a good wave modelling tool

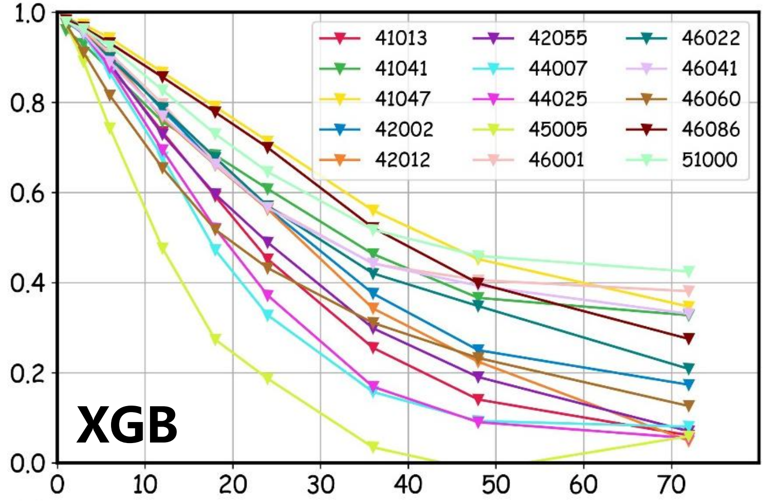
**The variation of correlation coefficents with forecast time for five models (different plots) at 16 buoy locations (different colors).**



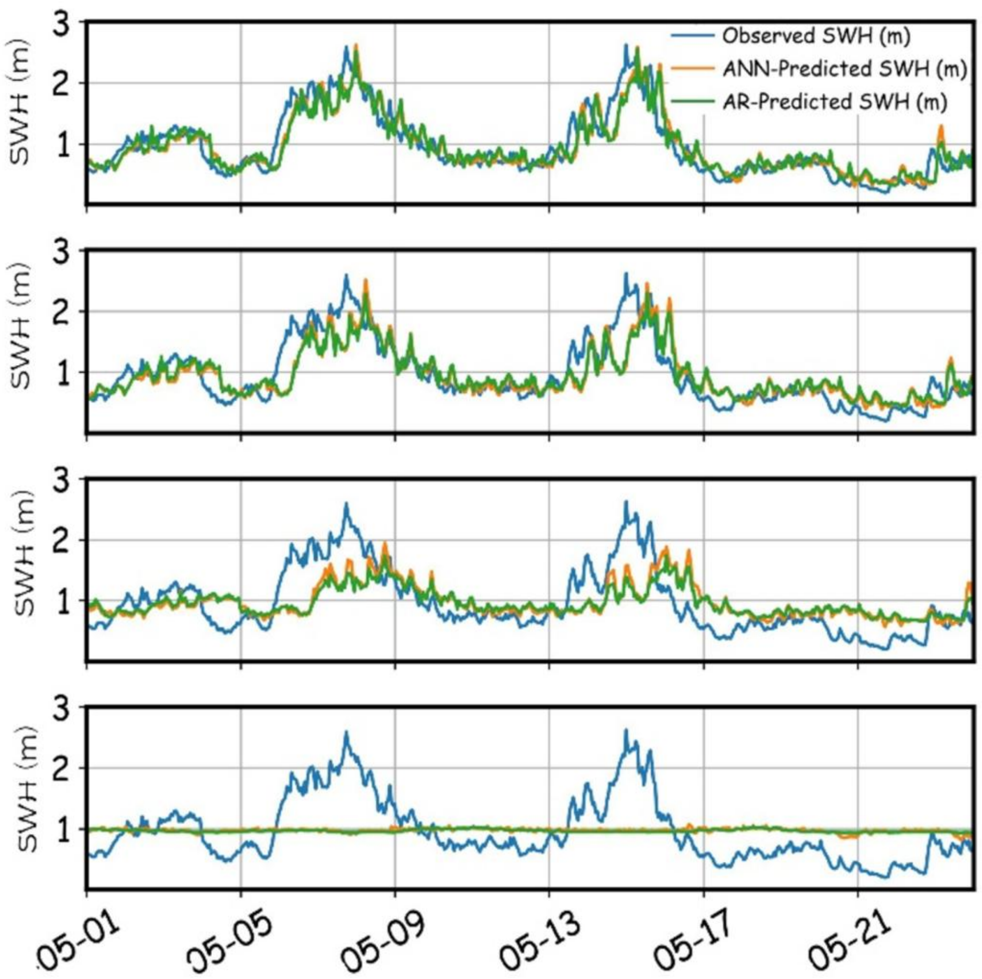




**The performance of different models shows NO DIFFERENCE 😂**

**All show POOR performance…**

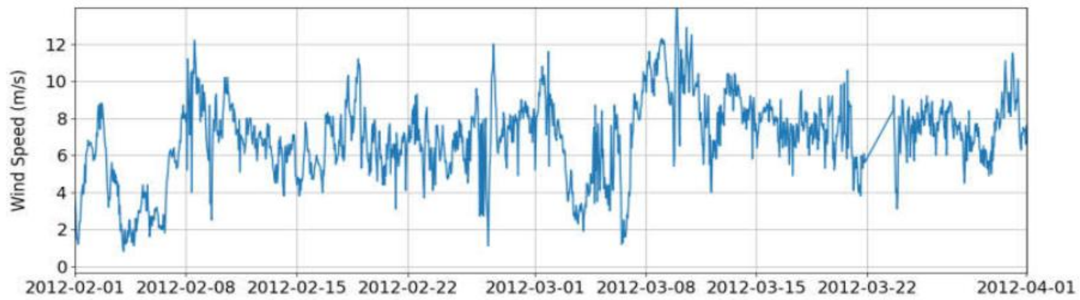
# Time series forecast is probably not a good wave modelling tool

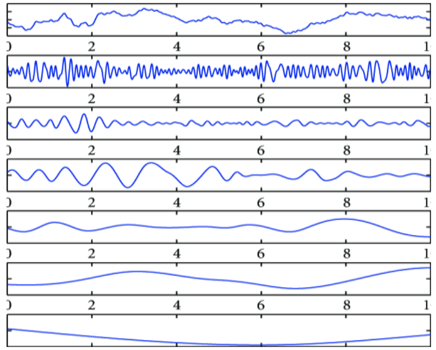



**Inertial Forecast**

**Mean Value**

What the model has learned is simply turning from an inertial forecast to the mean value with the increase of forecast time.



Decompose the signal using EMD/VMD
then
make prediction on components?

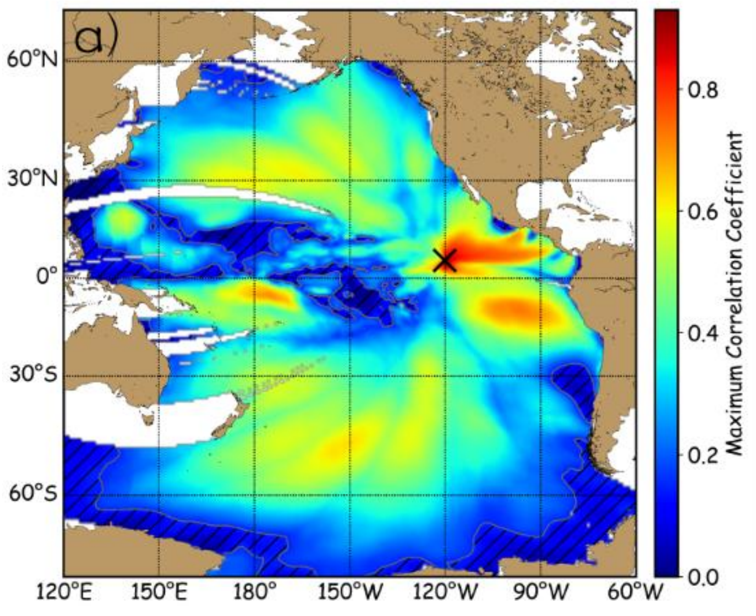
Many papers declare they obtain better results using this method

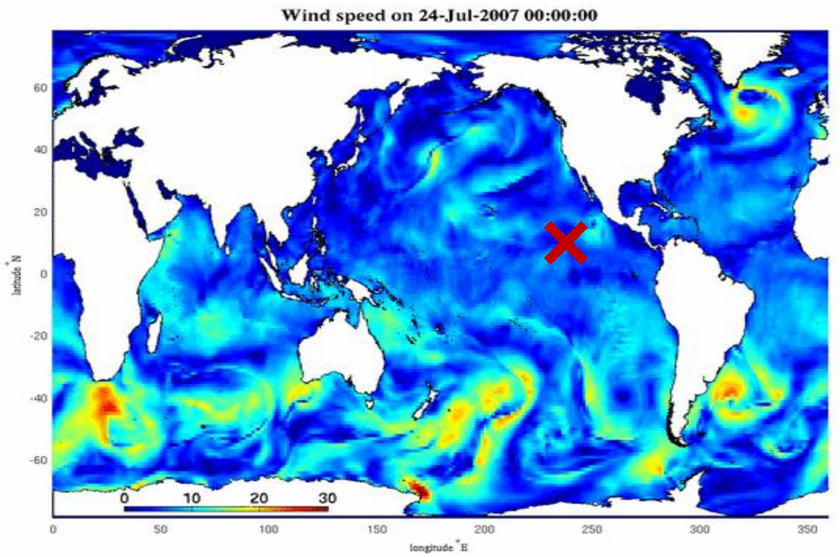
Data leakage!

Test set are actually unknown, and should NOT be decomposed.

# AI wave modelling: A different physics-guided strategy

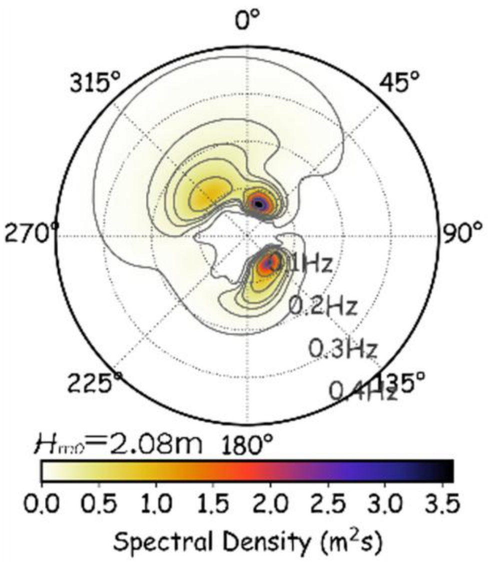
- **Physics-guided:** Using known physics knowledge to help the selection of model I/O
- Wave modelling is more of a forcing problem than an initial value problem, **forcing is necessary.**
- Waves are either generated by local current winds or by remote historical winds.
- Correlation between wind (current/historical, local/remote) and wave is strongly nonlinear because of complicated physical processes, e.g., generation, propagation, interaction
- Using AI to find the nonlinear correlations between wind and waves?




**Wind Input**

**Deep Learning?**




**Spectra/IWP Output**

# Data: Input & Output

- **Data：ERA5**

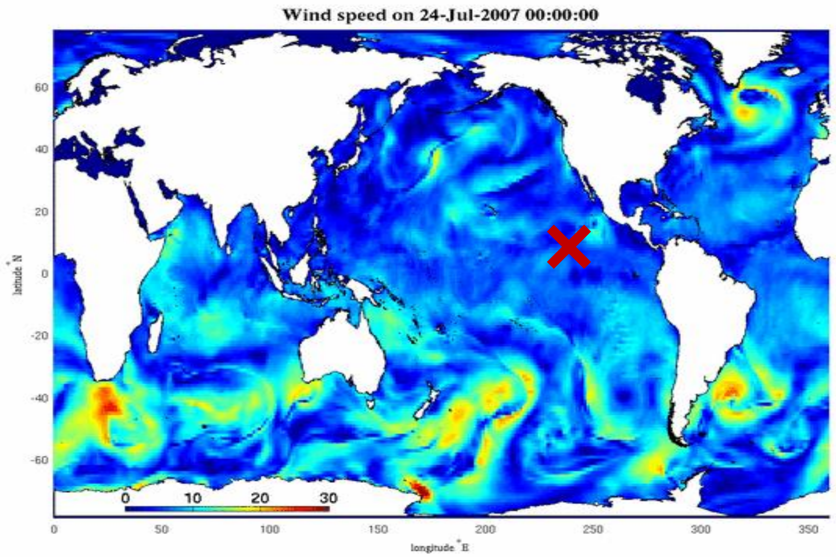
  Input: ERA5 Wind（Global，0.25°×0.25°×1h）

  Output: Directional Wave Spectra（Several Points）

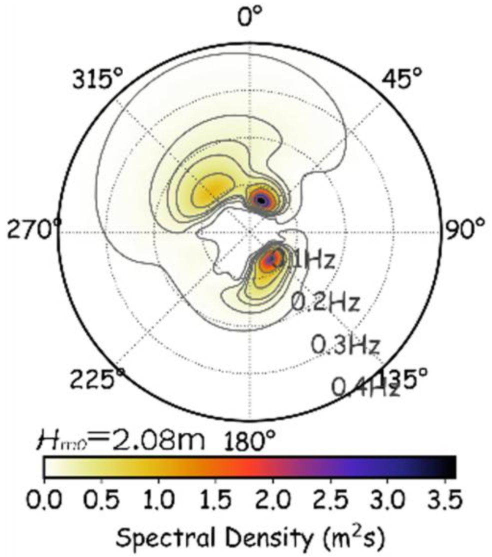
**Why buoy data are not used?**

1. Buoy spectra are not that reliable
2. Model has more data

- **“Brute Force” - One model for one point.**




**AI**




**INPUT：** Global wind for the past 10 days
**720(lon)×320(lat)×80(time)×2（U/V)**

**OUTPUT：** DWSs at a Given Point
**24（Dir）×31（Freq)**

**BIG**
**input data**
**&**
**SMALL**
**sample size**

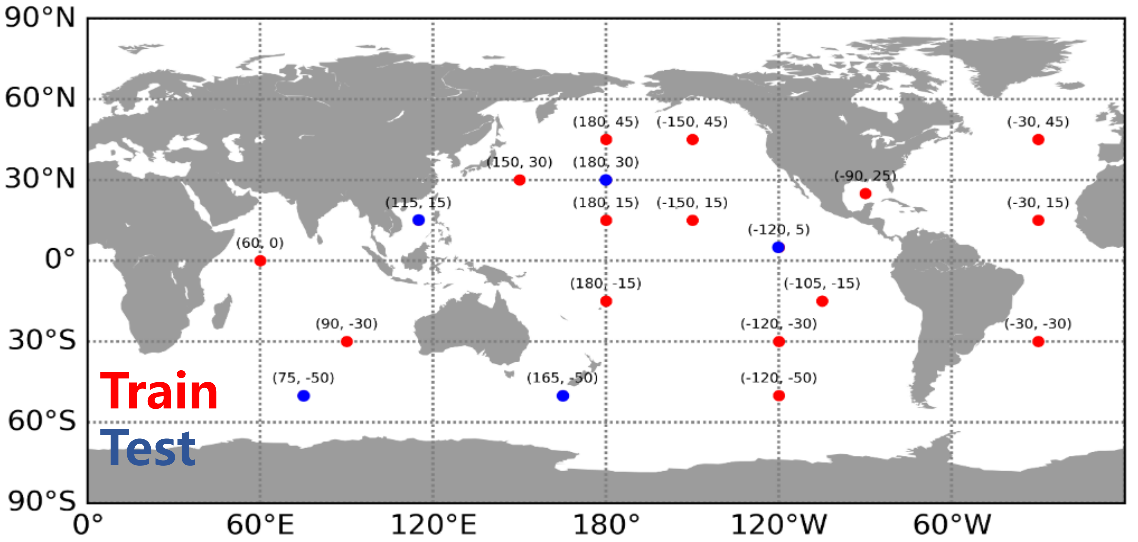
**Difficult for Model Training!**

# Grid re-organization

- **Data: ERA5**

  ERA5 Wind (Global, 0.25°×0.25°×1h)

  Directional Wave Spectra (Several Points)




- **Re-organizing the grid according to the generation/propagation process of wind-seas/swells**

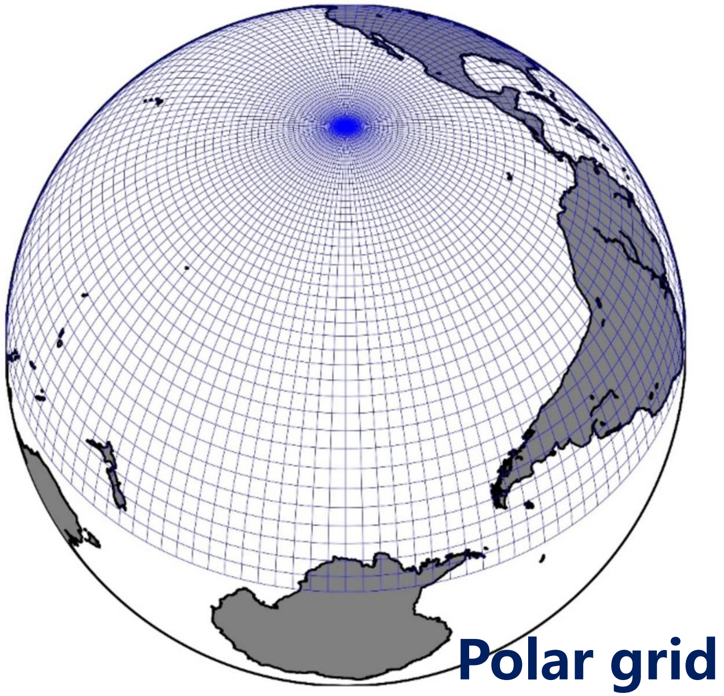
  Reorganize the data to a **POLAR GRID centered at the target point:**

  1. Centro-symmetric: Same input structure for different points
     → **One model for all points**
  2. Different spectral directions share the same physics
     → **One model for all directions**

  **The complexity of the model input is greatly reduced.**

  Directional resolution: 3°

  Distance resolution: $50\text{km}\times1.05^n$ (Near/Far: High/Low Resolution)



# Feature Engineering

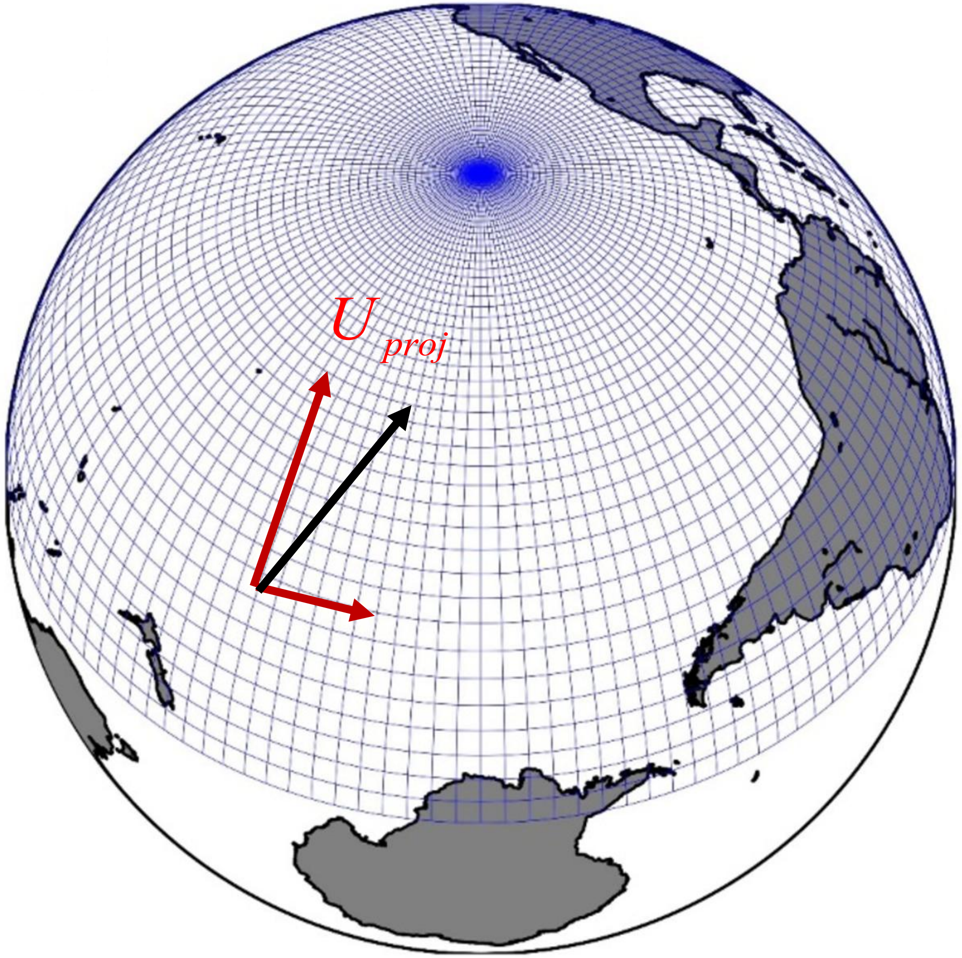
- **Feature Engineering of Input**

  Instead of using U & V, we use the projection of wind speed (& its square) along the great circle to the target point. (The great circle are actually the "meridian" of the polar grid)

  The projection is set to 0 when it is negative or when the great circle passes lands

$$U_{proj} = Max[0, U\cos(\theta_U - \theta_A)\delta]$$

$$U^2_{proj} = Max[0, U^2\cos(\theta_U - \theta_A)\delta]$$



# Model Input & Output

➢ **Input Vector & Model Output**

**OUTPUT**

The spectral densities at each direction as OUTPUT

[ ERA5 directional wave spectra has 24 directions (15° each) ]

**Format:** 31 spectral densities at a given direction.

**INPUT**

For the spectral densities at a direction, the wind features for the past 10 days in the corresponding 15° sector where the waves come from are used as input

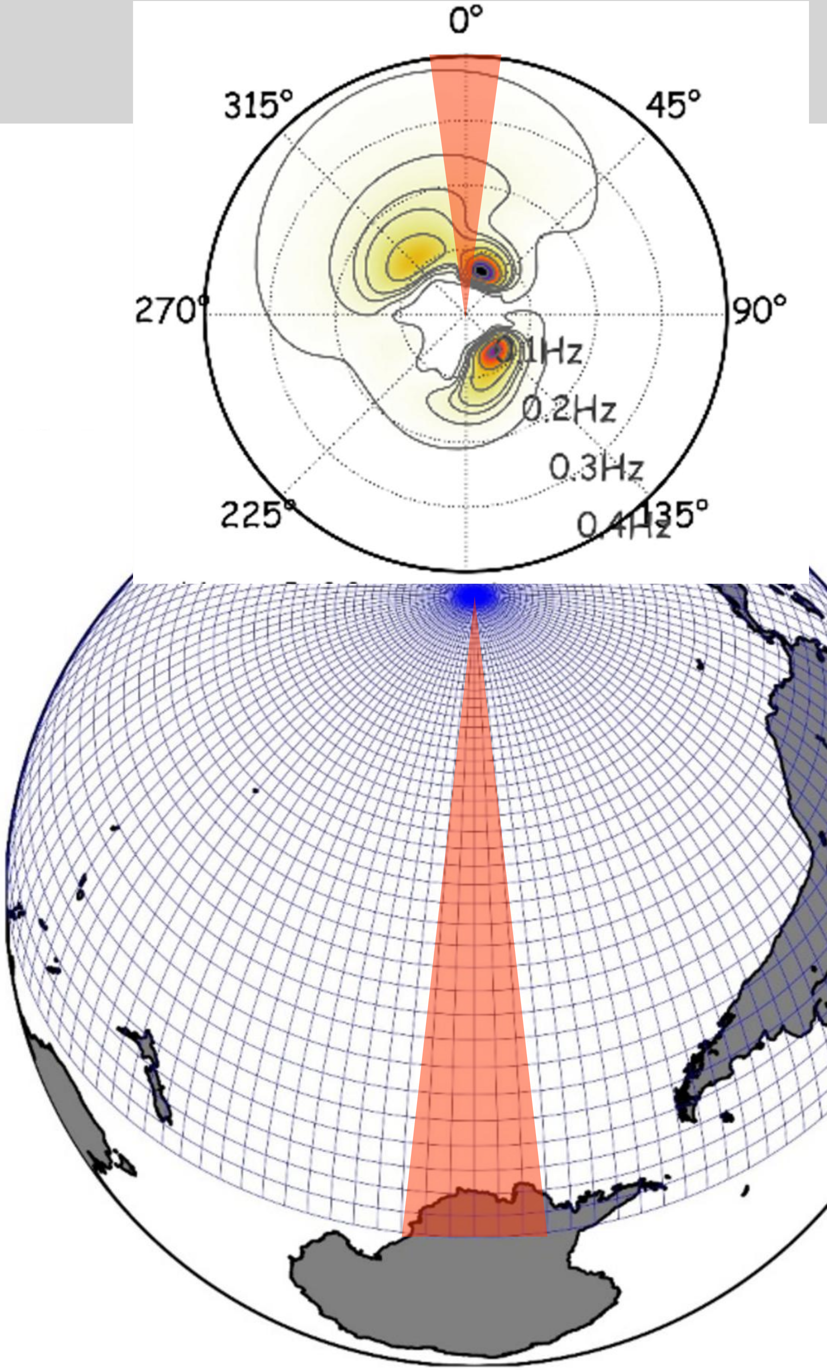
**Format:**

2(Features)×5(3°sectors)×51(distance grid)×80(time)

**BIG input data** → **Input array size was reduce for 24 times**

**&**

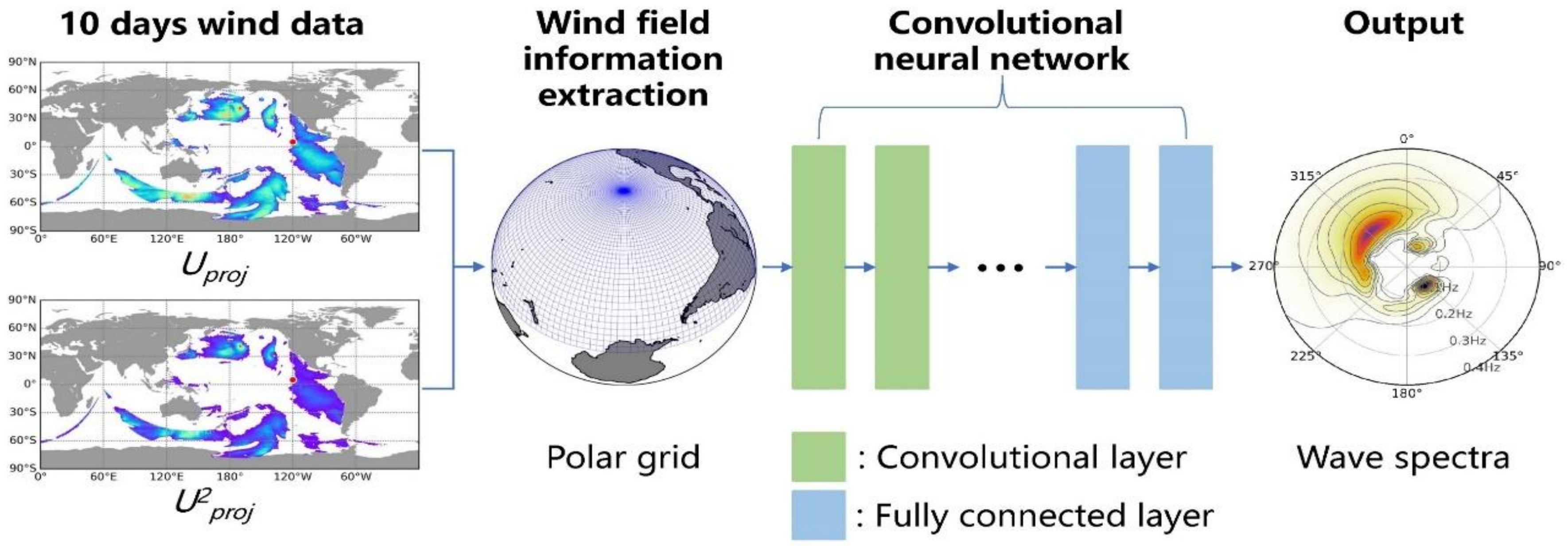
**SMALL sample size** → **Sample size was increased for 24 times**



# Model Architecture

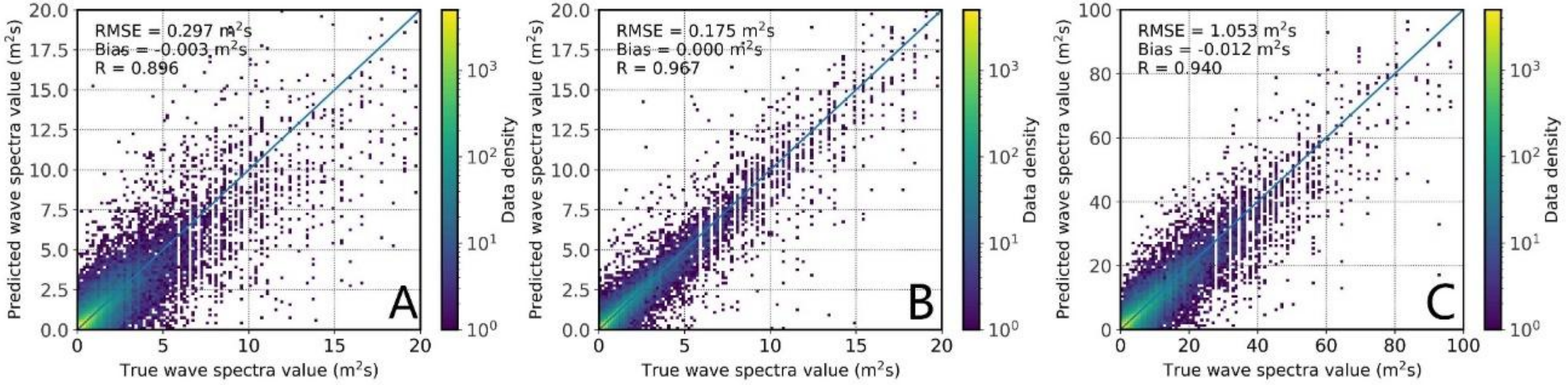
➢ **Model Selection：CNN**

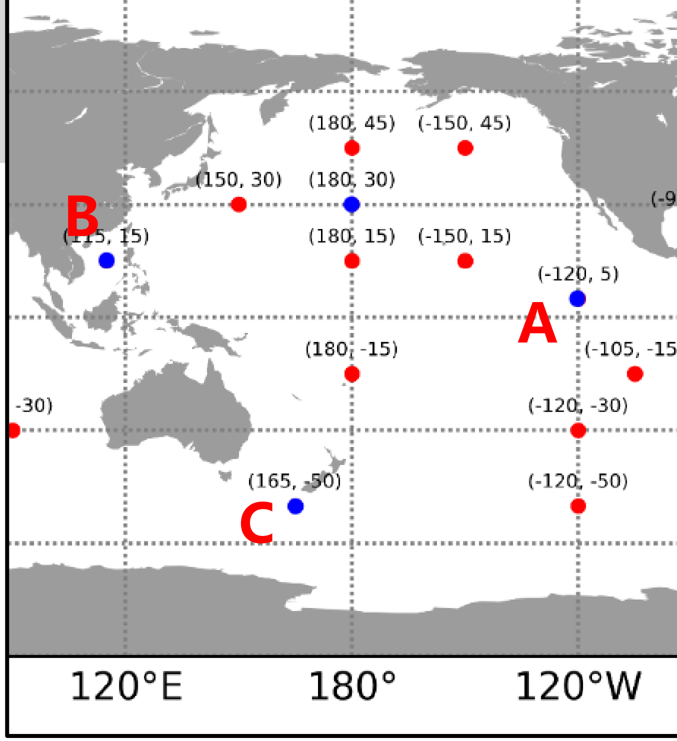
- 8 × 2D convolutional layers+4 × fully-connected layers
- 31 parallel networks, one for each frequency (no need to balance the loss)
- Kernel Size：3×3，No. of kernels：64→128→512→256→128→64→16→8
- Fully-connection sturcture：120→32→8→1
- MSELoss / ReLU / Padding / Maxpooling / Adam / EarlyStopping / Dropout



# Model Performance

## ➢ The agreement in spectral density

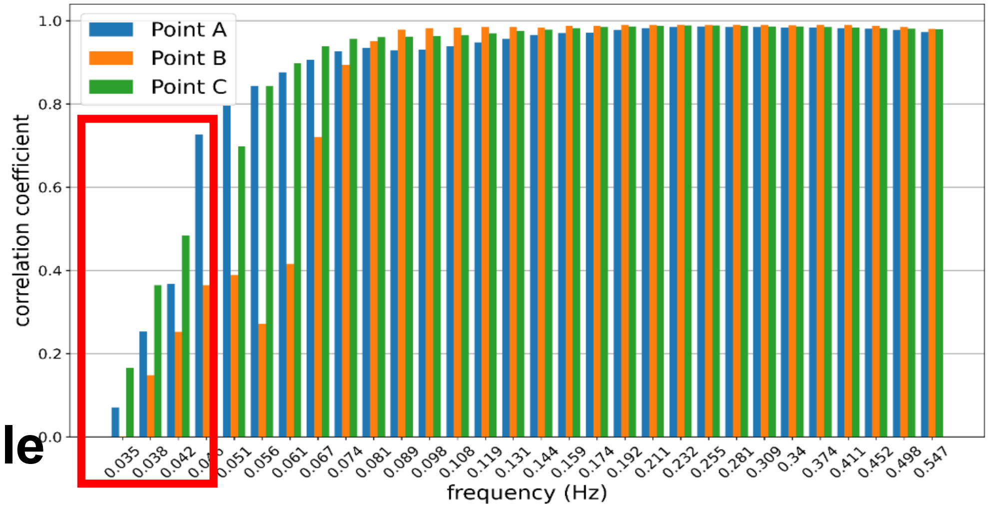


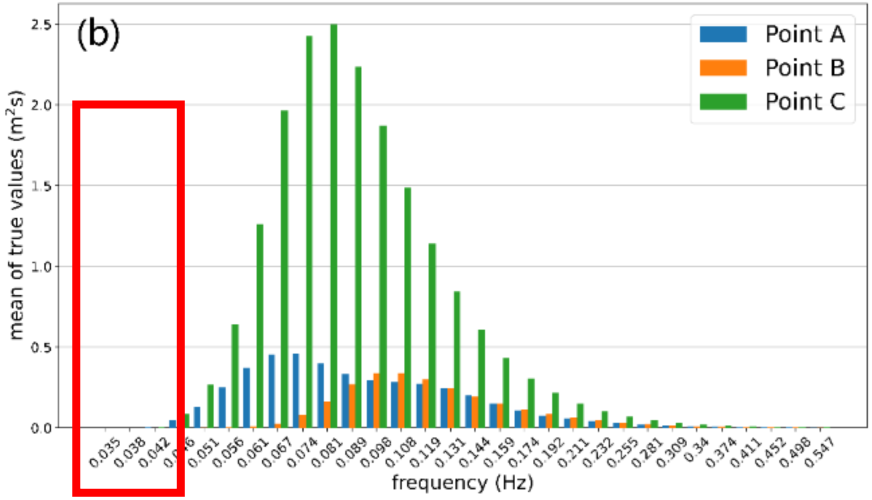

**Good overall agreement**

**Better performance for high-frequencies**

**Low correlations are in frequencies with negligible spectral densities**



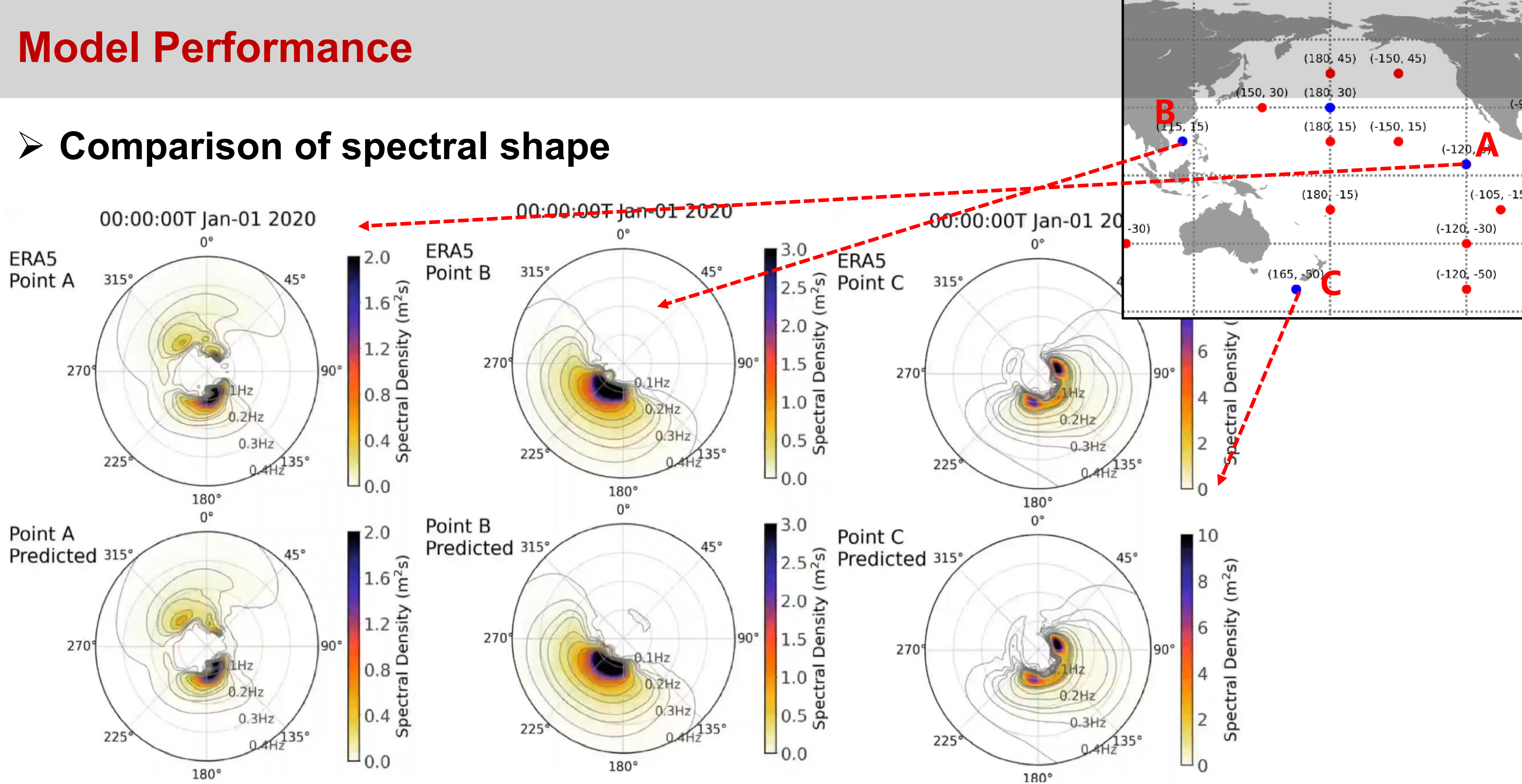

Correlation coefficients between predicted & ERA5 spectral densities




Mean 1-direction wave spectra

# Model Performance

➢ **Comparison of spectral shape**




Both local wind-seas and swells coming from thousands kilometers away are well modelled

# Model Performance

➢ **Comparison of Integral Wave Parameters**

$$SWH = 4.04\sqrt{m_0}$$

$$MWP = m_0 / m_2$$

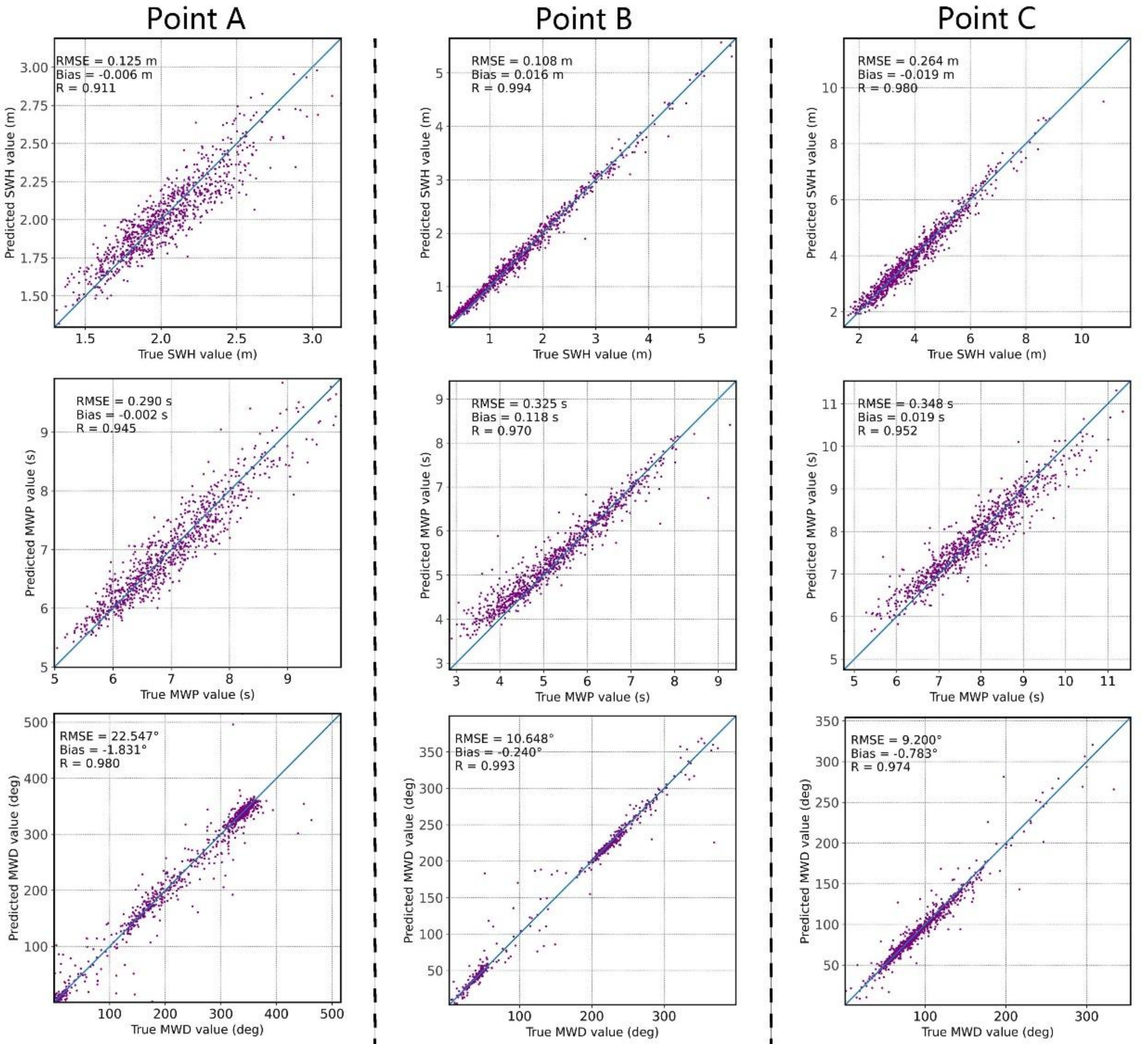
$$MWD = \arctan\left(SF / CF\right)$$

$$m_i = \iint f^i S\left(f, \theta\right) df d\theta$$

$$SF = \iint \sin\theta \mathrm{g} S\left(f, \theta\right) df d\theta$$

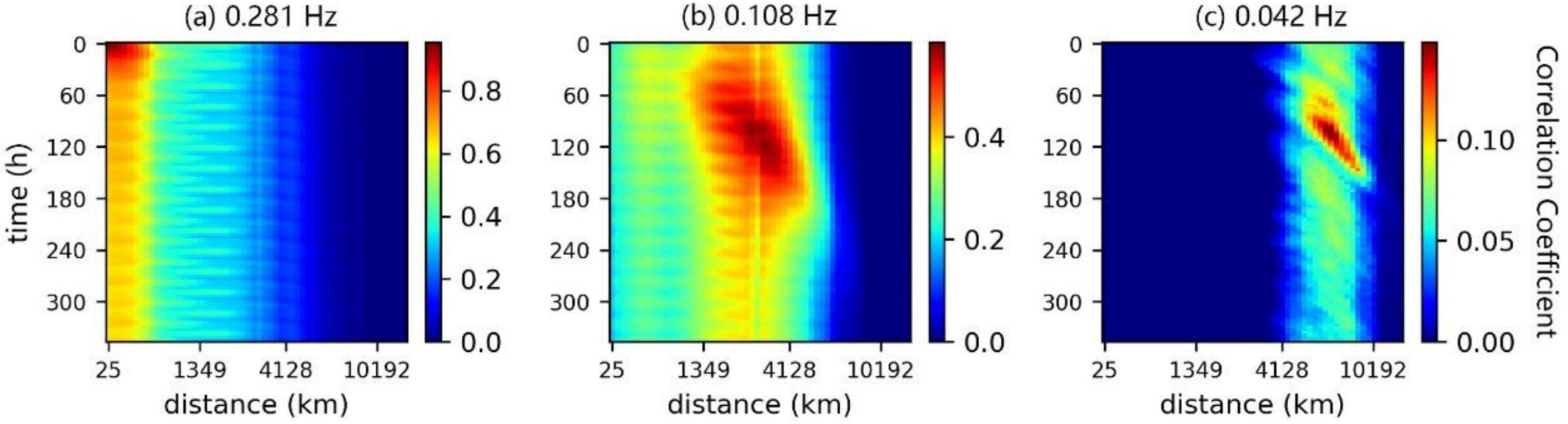
$$CF = \iint \cos\theta \mathrm{g} S\left(f, \theta\right) df d\theta$$

**Also in good agreement**



# Discussion: How the AI model works?

**Correlation coefficients between $U^2_{proj}$ and spectral densities at (a) 0.28 Hz, (b) 0.11 Hz, and (c) 0.04 Hz with different spatial-temporal distances to Point A.**




- **High-freq:** Equilibrium range. Spectral densities are highly correlated to local wind (r>0.9). It is easy to find a good fit.
- **Mid-freq:** Dominated by wind forcing, dissipation, and wave-wave interaction. The spectral densities are the result of integral of wind along wave propagation. The empirical relation between wind speed/fetch/duration and wave spectra is re-learned by deep learning.
- **Low-freq:** More challenging. Wave growth + swell fast propagation (frequency dispersion, angular spreading, dissipation, etc.). The DNN can learn some parts, but the error also becomes larger.

# Downscaling to coastal wave spectra also by AI

- NO nearshore physical processes considered in the above model, so not suitable for nearshore wave spectra’s modelling.
- In coastal areas, wave dynamics are mainly influenced by the bathymetry and coastal morphology.
- Once the Directional Wave Spectra (DWSs) at the open ocean boundary are known, the DWSs at various locations along the coast are almost determined.
- Correlation between open ocean and coastal DWSs is nonlinear and complicated.
- Again, AI is good at digging such correlation from the data !

# Downscaling to coastal wave spectra also by AI

➢ **Study Region： Southern California Bight**

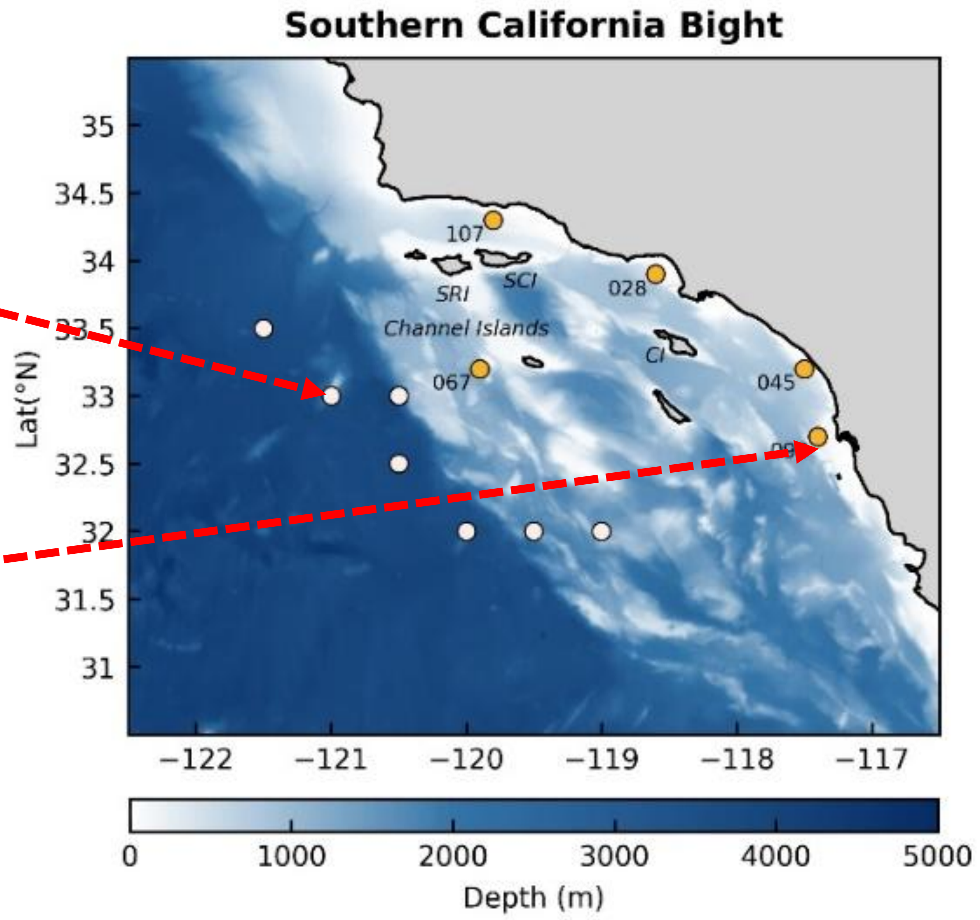
➢ **Data： IOWAGA (ST4)**

**Inputs:**

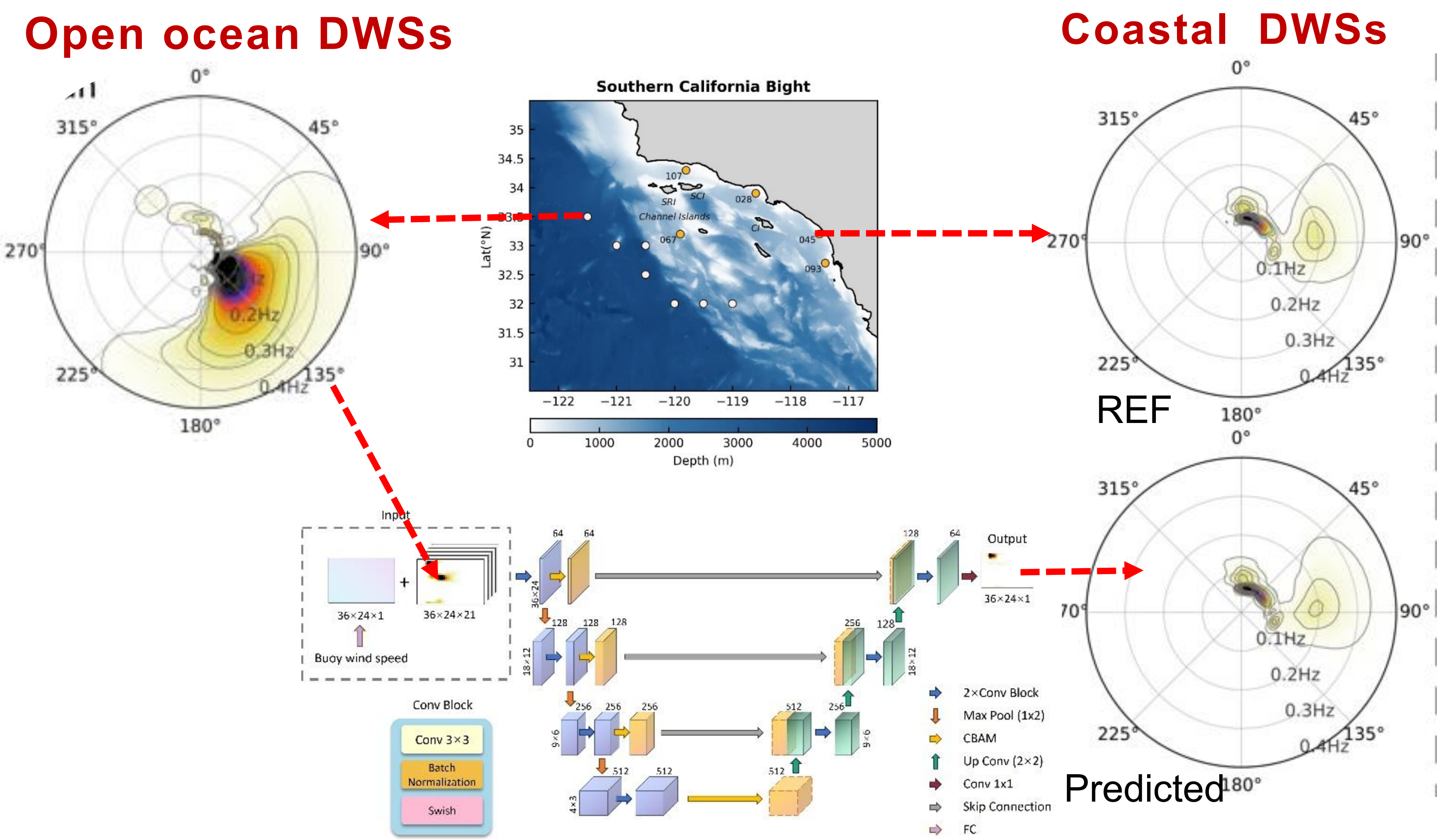
DWSs at **7** points [White Points] in the open boundary
×**3** time steps (T-0/-3/-6)
+ wind vector in the target location
[**(7×3+1)×24×36** arrays]

**Outputs:**

DWSs at the target location(s) [Orange locations]
[**24×36** arrays]



# AI Model Architecture




- Model selection:

  **Unet ++**

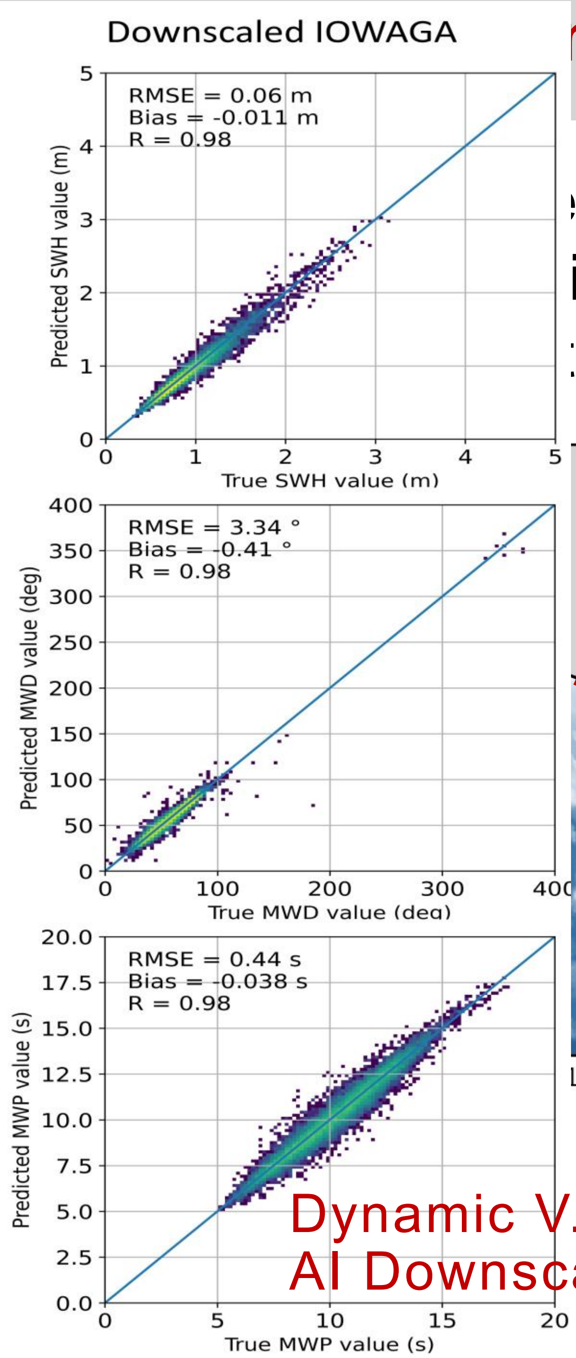
Key: Skip connection for different resolutions.

- Training
- 1996–2015
  ~1h in a 3080Ti

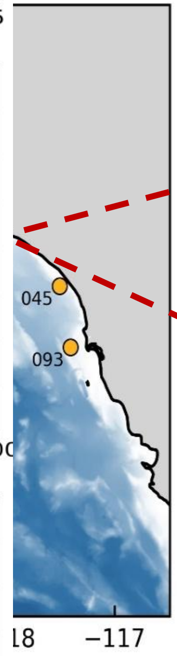
- Testing:
- 2016–2021
  ~10 seconds to finish the computing

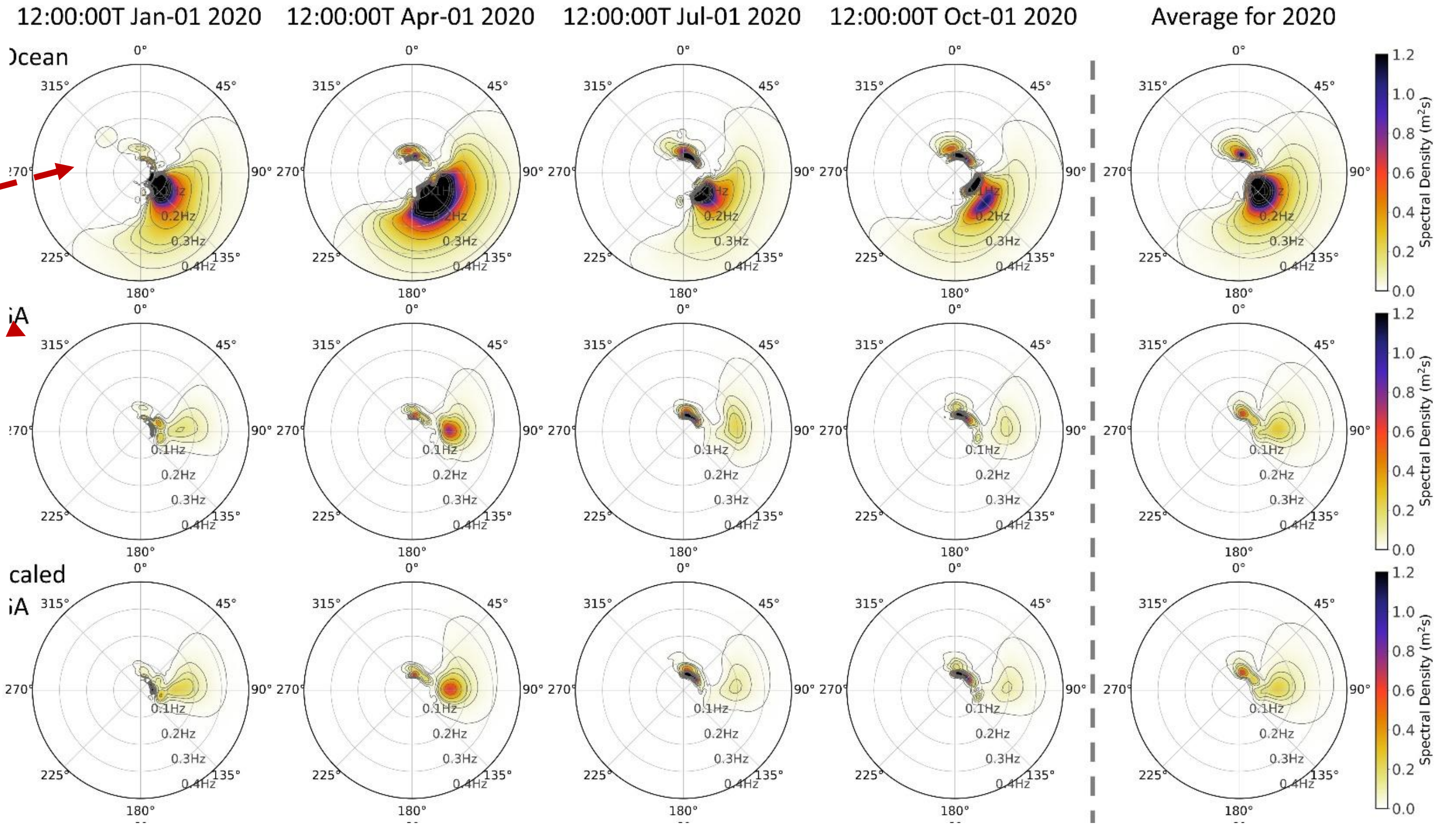
# nce

ence of coastal topography, the shapes of DWSs in nearshore ignificantly from those in open oceans, even when the two tively close (~250 km apart).




Dynamic V.S.
AI Downscaling




**Open boundary DWSs**

**Dynamic Down-scaling**

**AI Down-scaling**

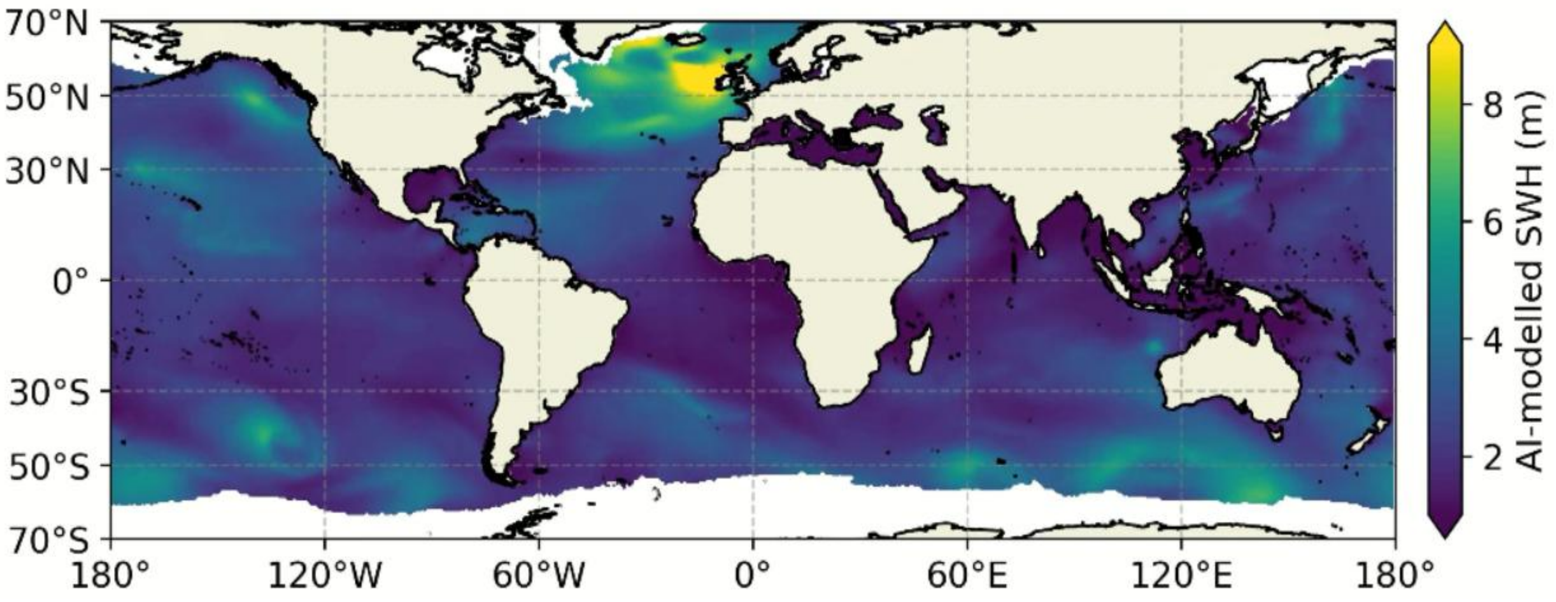
# From Single-point to Global Fields

If global historical (~10 days) & current winds are known, the spatial distributions of any wave parameters are almost determined at the current moment [Omitting ocean currents and sea ices].

**INPUT:** Global wind for the past 10 days ➝ **AI** ➝ **OUTPUT:** DWSs at a point (a 2-D array)

➝ **AI** ➝ **OUTPUT:** Global Field of Wave Parameter (also a 2-D array) ?

- Modelled wave parameter: **SWH**
- (One of) The most important wave parameter
- Global high-quality satellite observations
  ➝ Good training targets

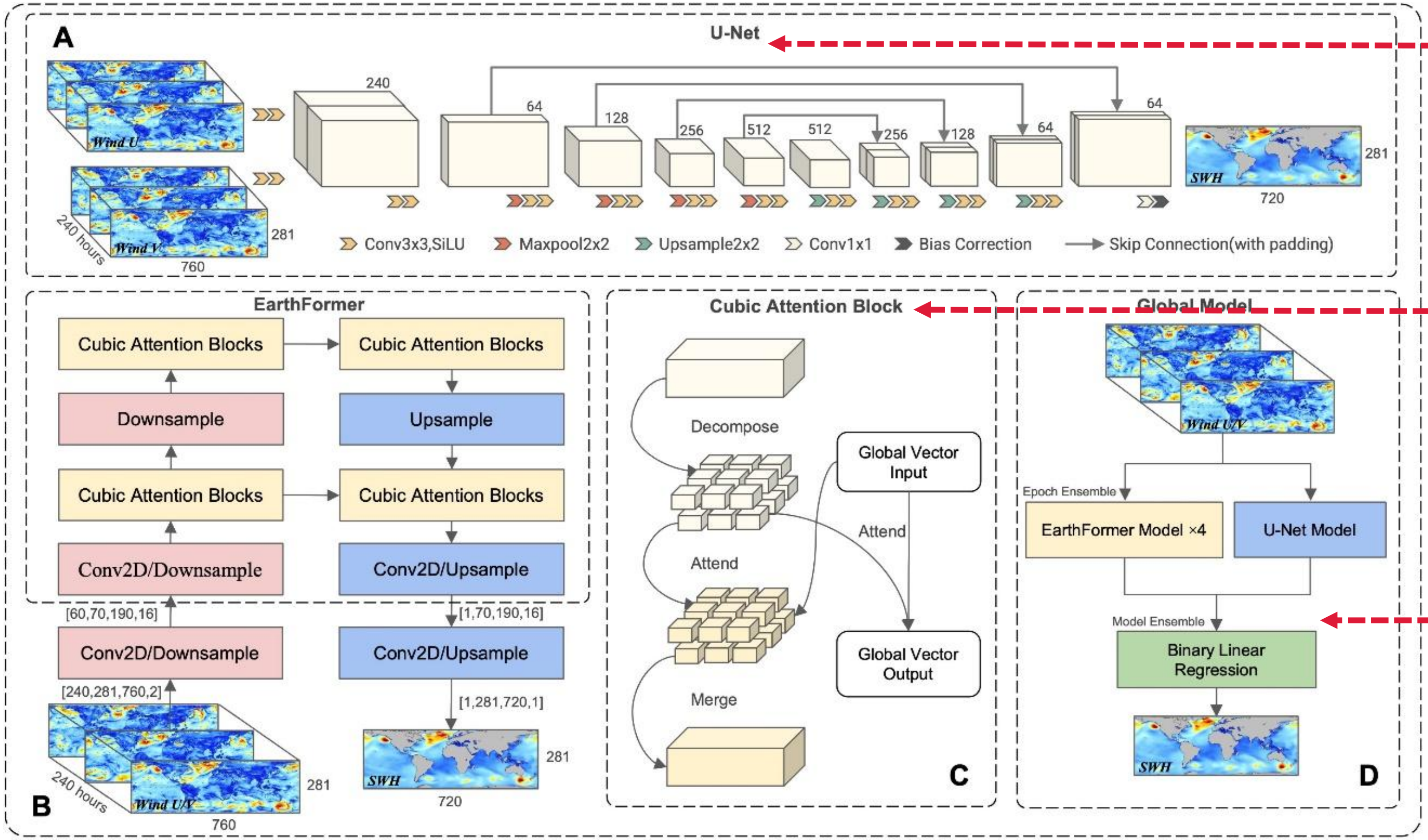


# From Single-point to Global Fields

## Pre-training

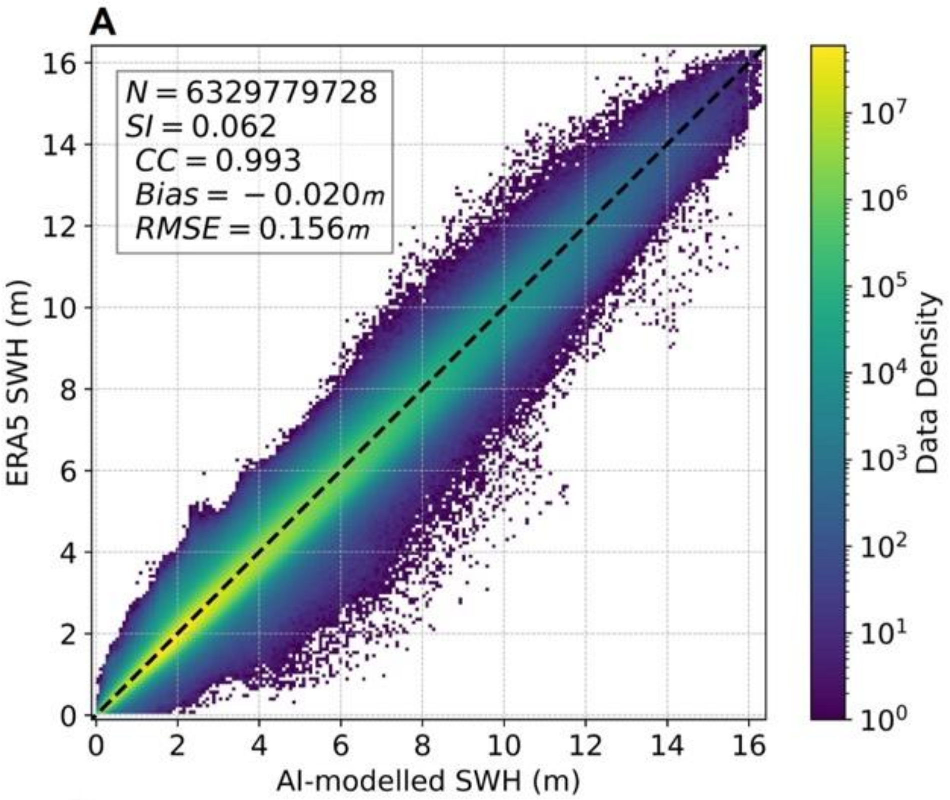
**INPUT:** ERA5 winds (U &V) of 10 days **OUTPUT:** ERA5 SWH

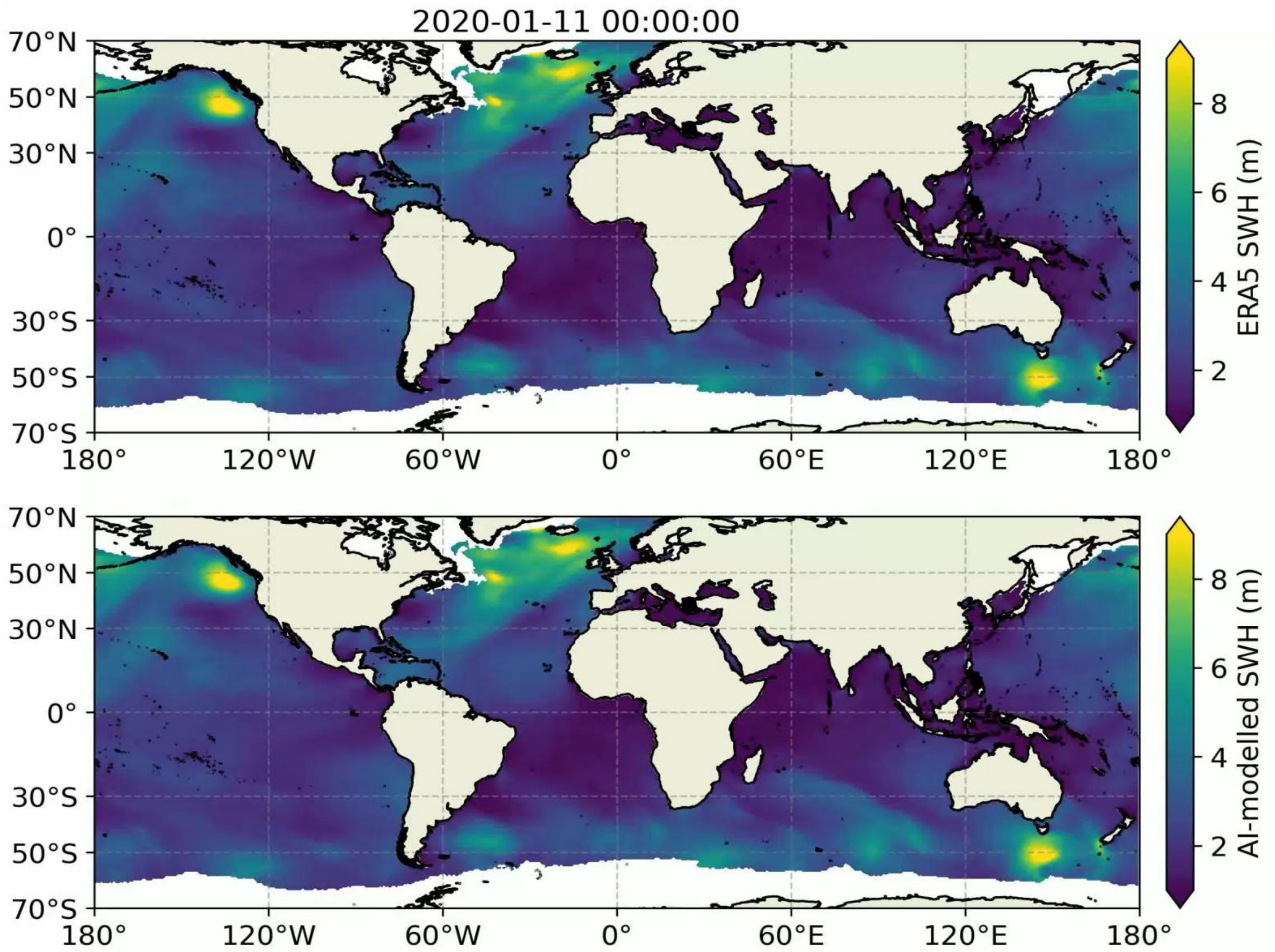
- 2000-2017: Train | 2022: Validation | 1997-1999 & 2018-2020: Testing




- Unet: Local attention by convolutional module to capture local correlation
- Earthformer: Global attention for tele-connection between remote winds and swells
- Model ensemble to obtain better results

# Results for the pre-trained model: V.S. ERA5 SWH test set

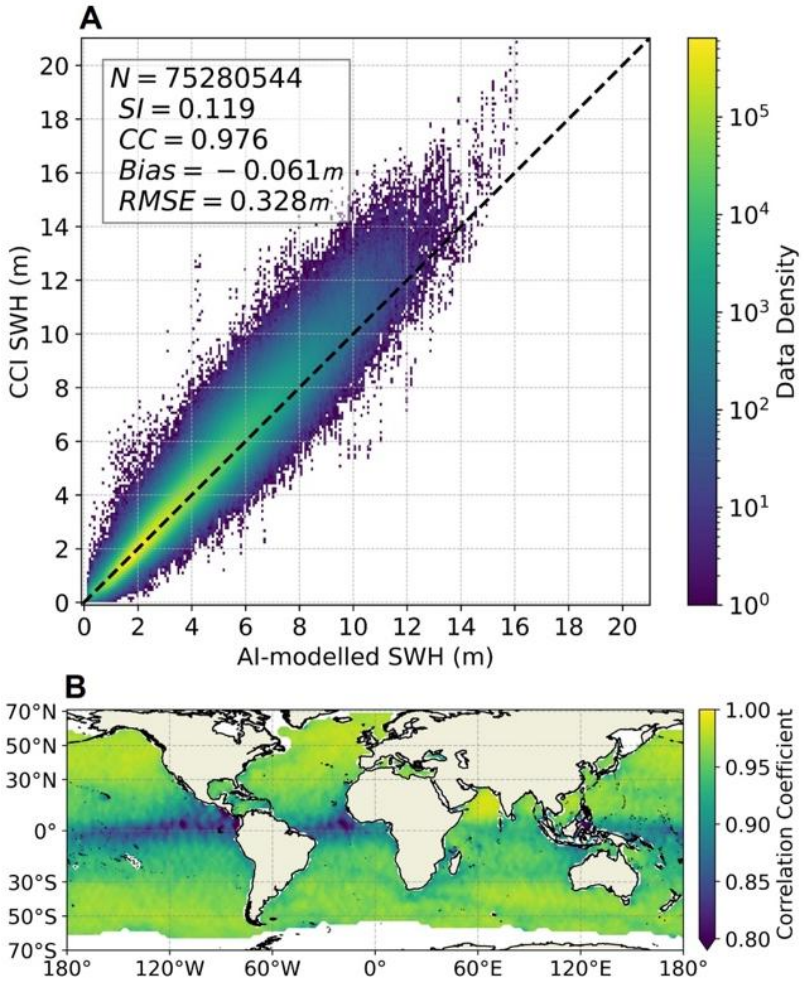


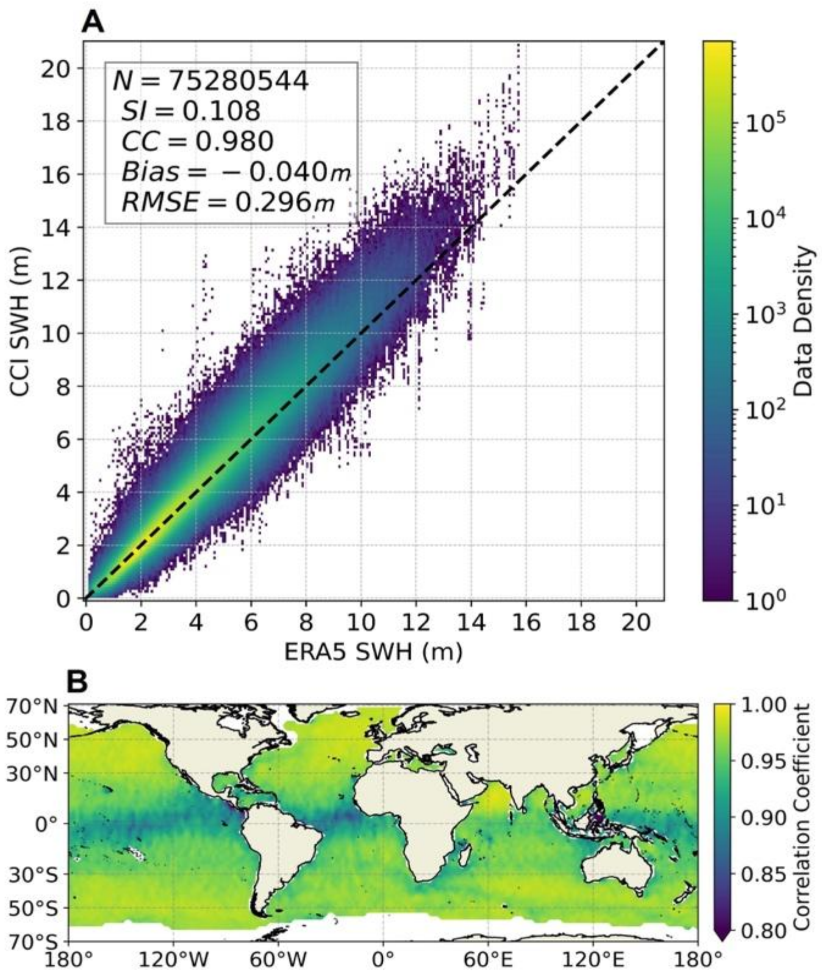

- Good agreement with ERA5 SWH

# Results for the pre-trained model: V.S. Altimeter Data

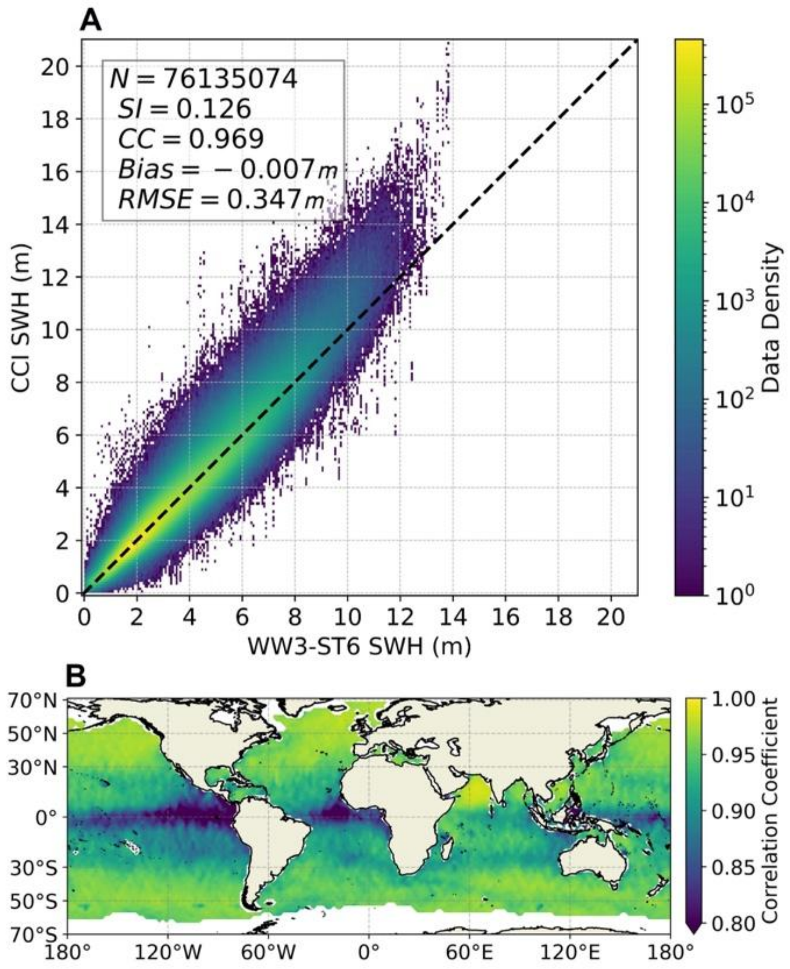
- Only Topex & Sentinel-3A are used for evaluation as they are not assimilated into ERA5




**AI vs Altimeter**




**ERA5 vs Altimeter**




**WW3-ST6 vs Altimeter**

# Transfer Learning using Altimeter Data

- The CCI-sea state dataset was merged with the WW3-ST6 data using an objective analysis method from 2016 to 2020 (At least four satellite in orbit during this period) to generate the target dataset.
  [2017-2019 for training, 2016 for validation, 2020 for testing]

- During the fine-tuning process, the parameters of the encoder layers of the AI model were fixed, and only the decoder layers were updated through back-propagation.

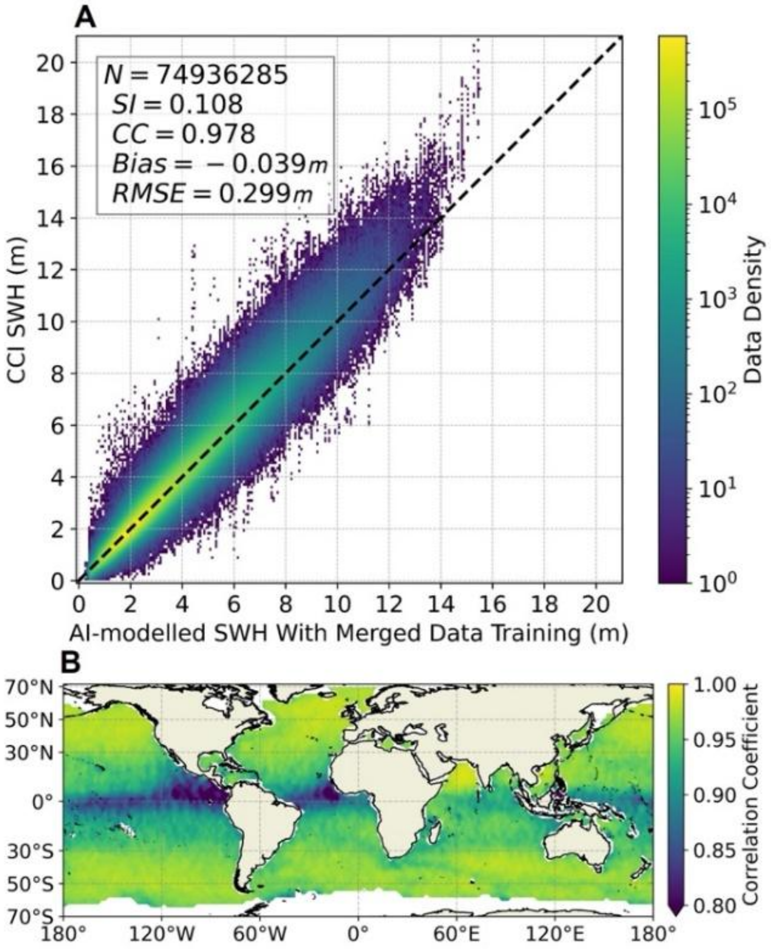
$$Loss = \frac{1}{n}\sum_{t=1}^{Time}\sum_{j=1}^{Lat}\sum_{i=1}^{Lon}\left[W_{i,j,t}\,g\left(y_{i,j,t} - x_{i,j,t}\right)\cos\theta_j\right]^2 \qquad (S5)$$

$$W_{i,j,t} = 1/\left[R\left(i,j,t\right)+c\right] \qquad (S6)$$

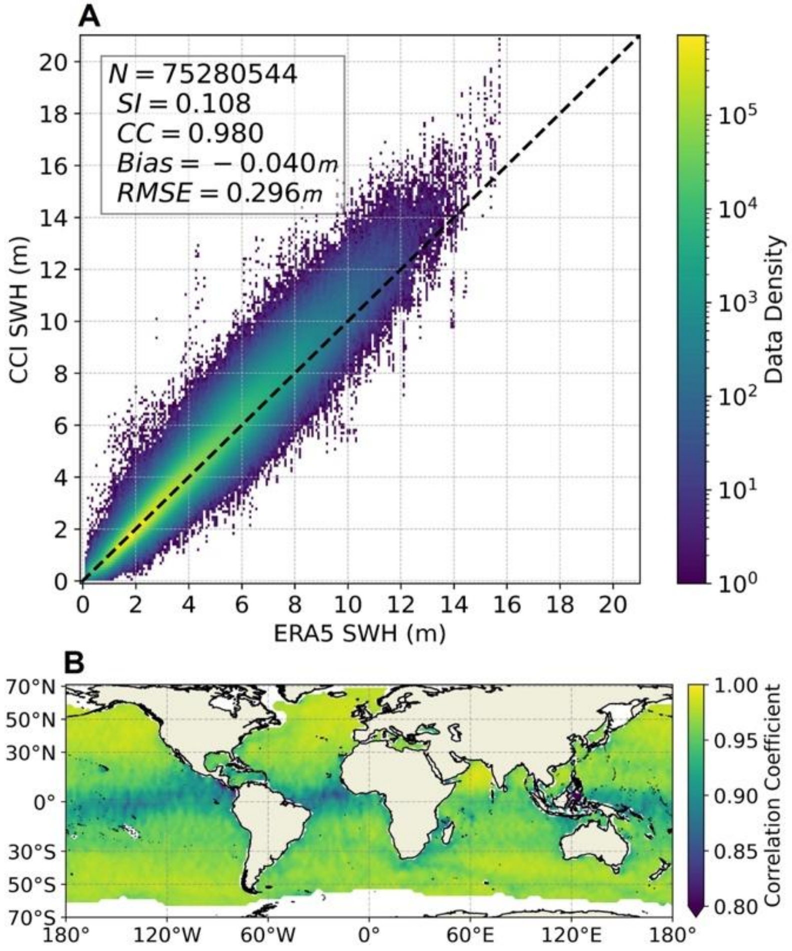
Transfer learning: To addresses the problem of "small" datasets when the target "small" dataset shares similar characteristics with a "large" dataset.

# Results for the fine-tuned model: V.S. Altimeter Data

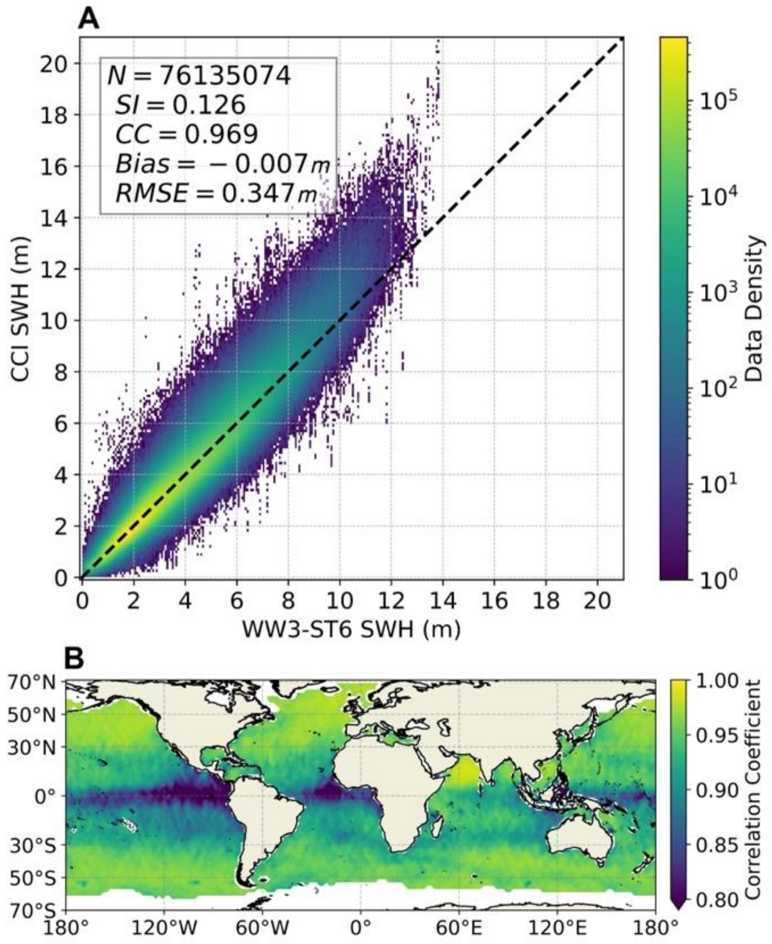
- Only Topex & Sentinel-1A are used for evaluation as they are not assimilated into ERA5




AI vs Altimeter

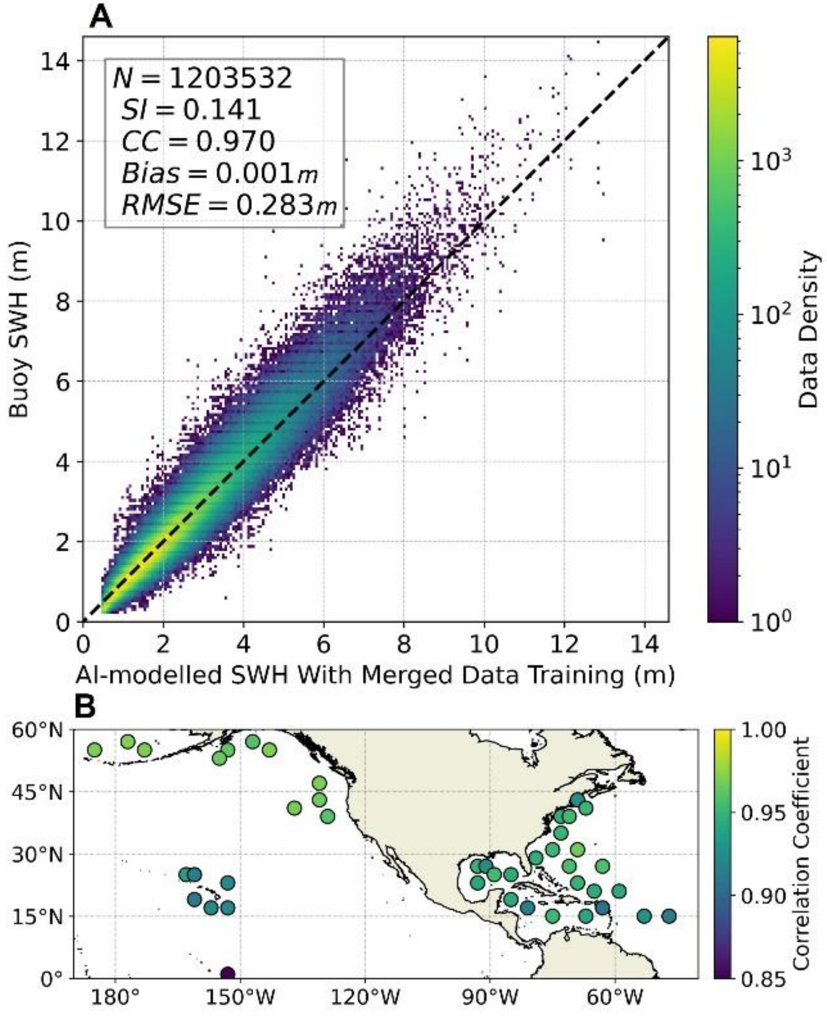



ERA5 vs Altimeter

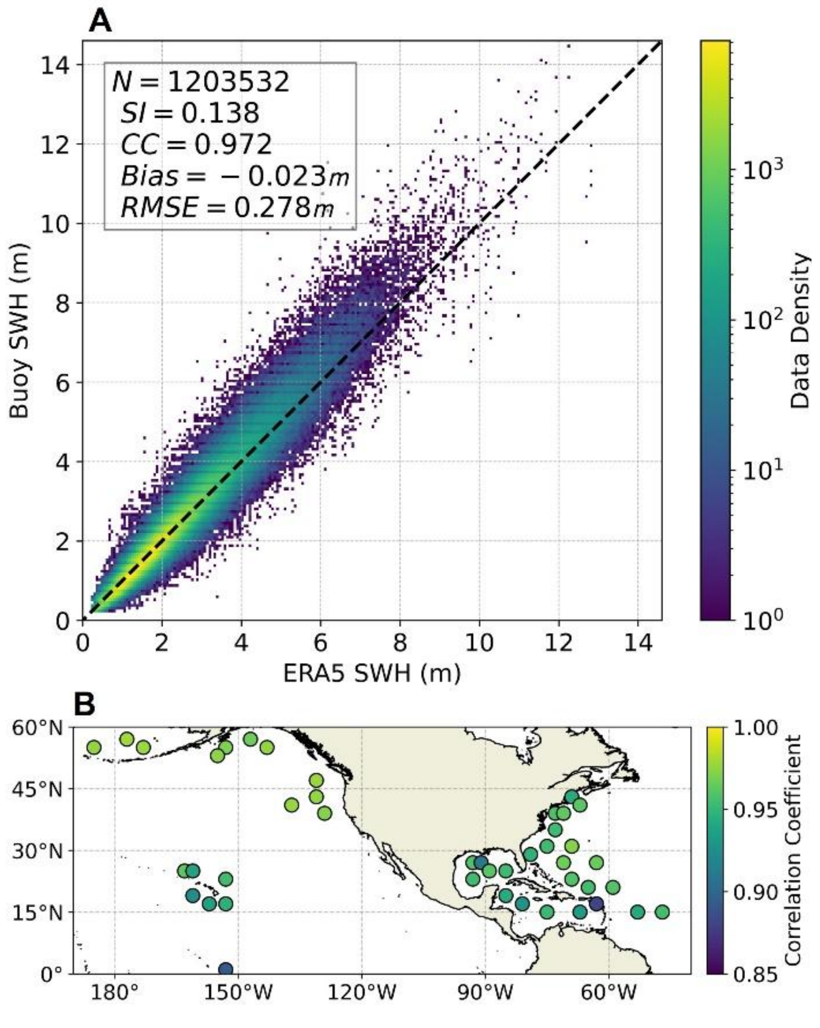



WW3-ST6 vs Altimeter

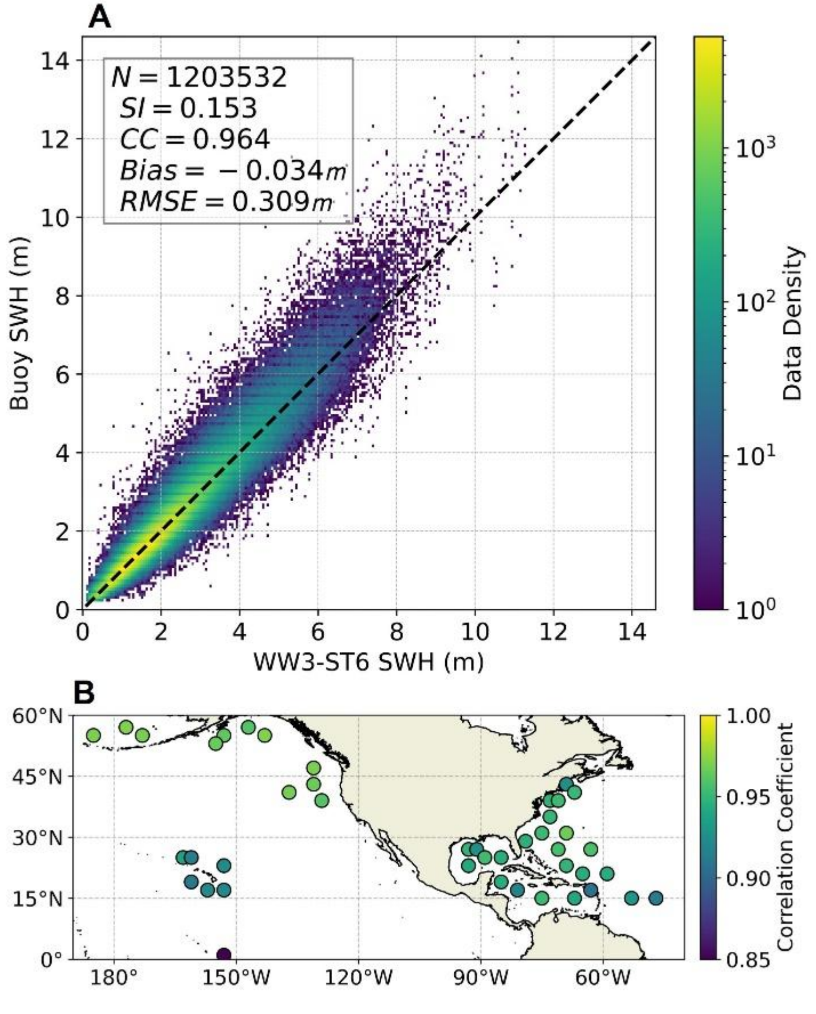
# Results for the fine-tuned model: V.S. NDBC buoy data




**AI vs NDBC**




**ERA5 vs NDBC**




**WW3-ST6 vs NDBC**

# Summary

Haoyujiang@szu.edu.cn

- Machine learning-based time series prediction methods are useless for wave modelling.
- Using known knowledge of wave physics to guide the section of AI model input/output.

E.g., Using winds over the past several days to predict current wave conditions.

- Several AI models are established for wave modelling (with good performance):

**Model A:** For modelling single-point Directional Wave Spectra (DWSs) in open oceans
Physics: Correlation between DWSs and local/remote winds due to wave growth & propagation.

**Model B:** Downscaling open ocean DWSs to single-point coastal DWSs
Physics: Complex but fixed mapping relation due to the stable bathymetry & coastal morphology.

**Model C:** Modelling global field of SWH
Physics: Same as A, waves are either generated by local current winds or remote historical winds

1. Physics-guided Deep Learning for Skillful Wind-wave Modelling, *Science Advances*, 2024.
2. Comment on papers using machine learning for significant wave height time series prediction: Complex models do not outperform auto-regression, *Ocean Modelling*, 2024.
3. Statistical downscaling of coastal directional wave spectra using deep learning, *Coastal Engineering*, 2024.
4. A Deep Learning-Based Approach for Empirical Modelling of Single-Point Wave Spectra in Open Oceans, *Journal of Physical Oceanography*, 2023.

**THANK YOU!**