Tropical and Midlatitude S2S Prediction using UFS and Machine Learning

Eric D. Maloney, Elizabeth Barnes, Jack Cahill, Zaibeth Carlo Frontera, Yu-Cian Tsai

Colorado State University





Objectives

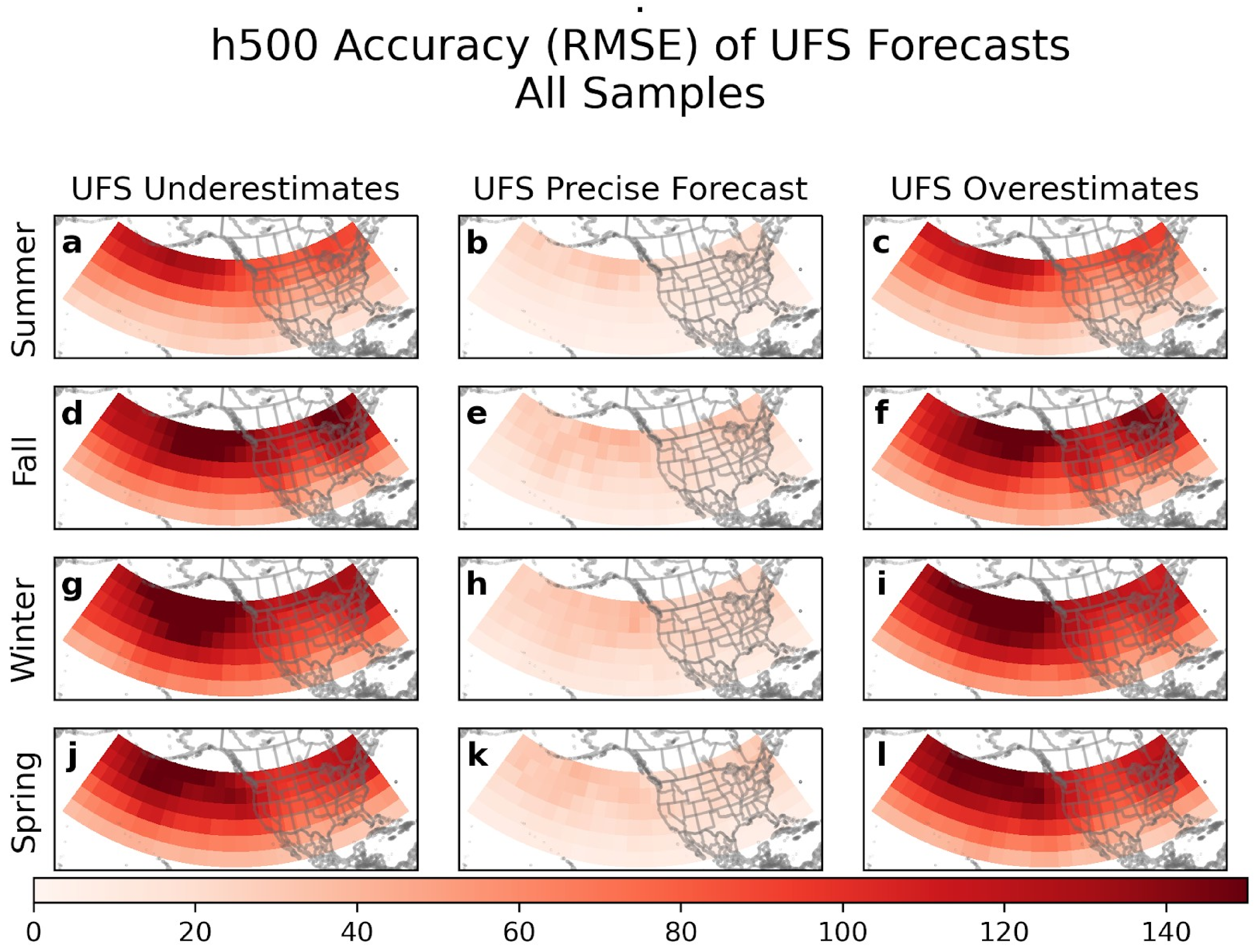
Part 1. Use ML to predict errors in subseasonal North American geopotential height forecasts in GEFS hindcasts

Part 2. Demonstrate that several prototypes of the UFS produce common subseasonal prediction errors over the tropical east Pacific and Atlantic, affecting the conditions that modulate tropical cyclones in these basins

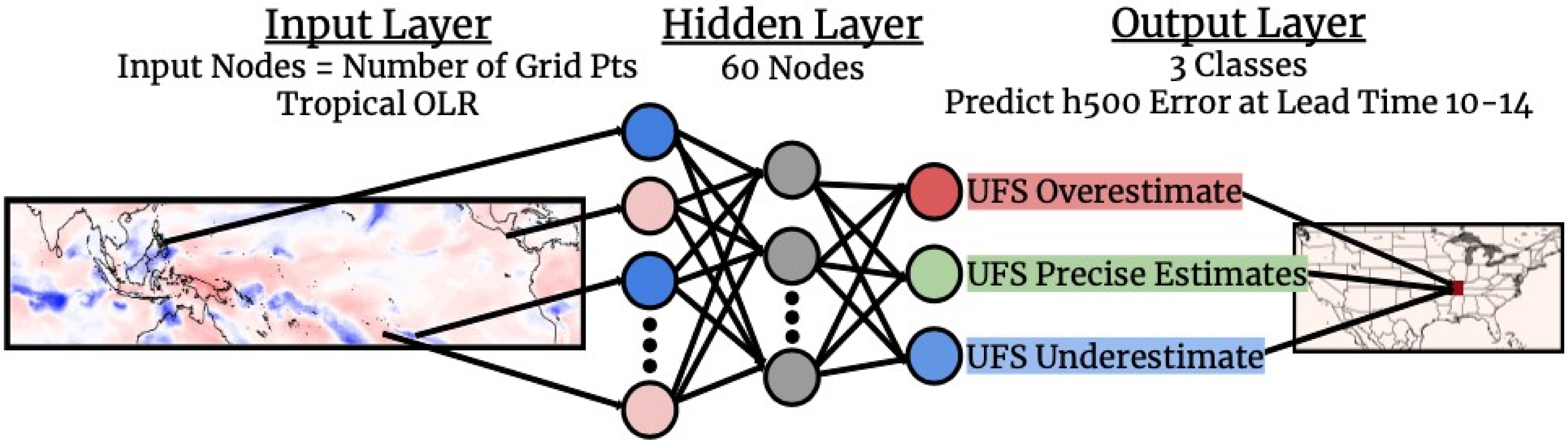
Part 3. Neural network (NN) model utilizing ENSO and MJO indices and other local environmental information used to predict east Pacific and Atlantic cyclogenesis

Part 1: Predicting Forecasts Errors in GEFS ("Errors of Opportunity")

- Global Ensemble Forecast System (GEFSv12)
- The reforecasts, comprising a total of 1042 samples, were initiated every 7 days from 01/05/2000 to 12/18/2019
- Integrated out to 35 days. Our study focuses on **<u>lead times of 10-14 days</u>**

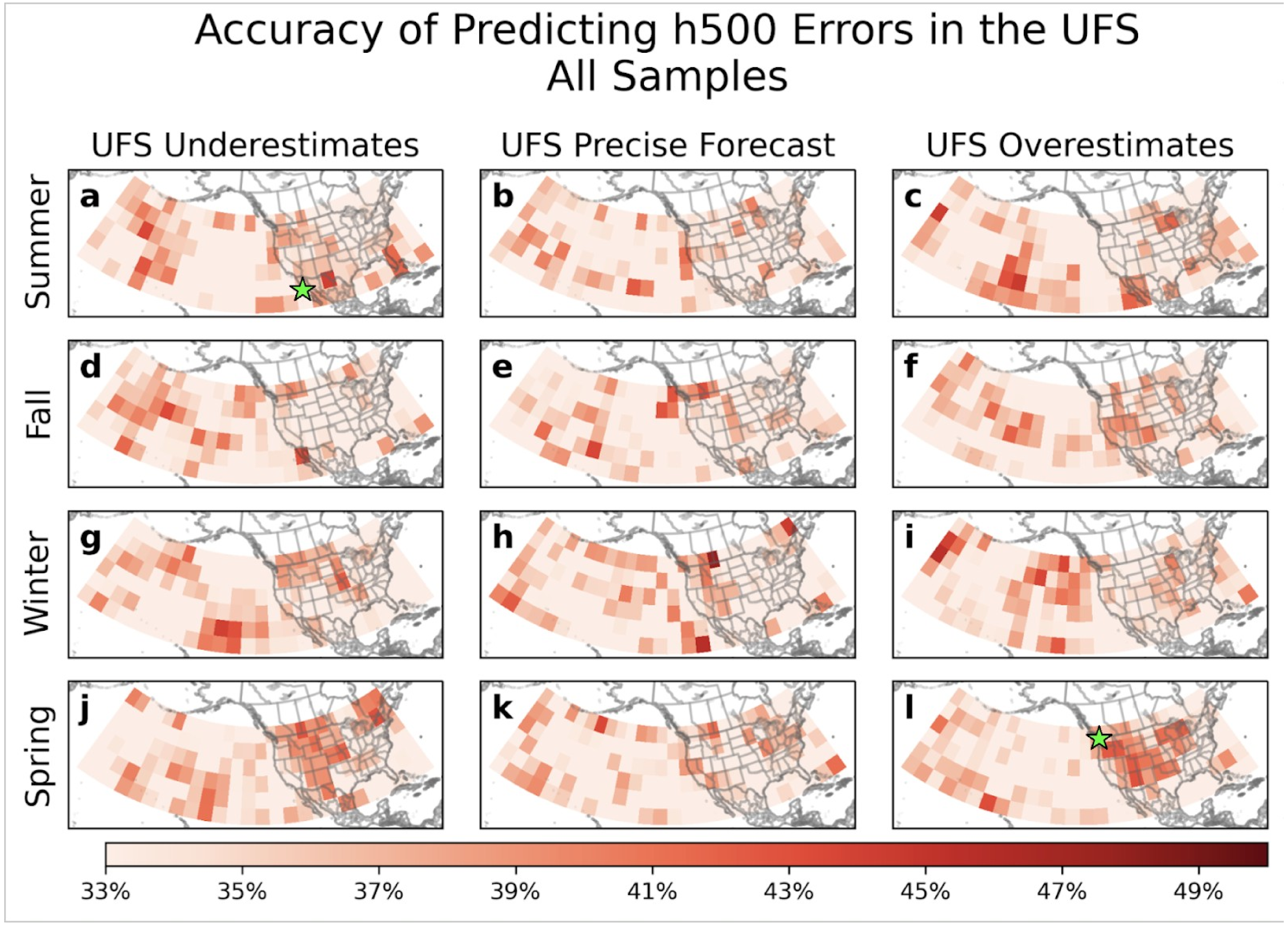


Using Neural Networks to Predict Errors in the Unified Forecast System (UFS) on S2S Timescales



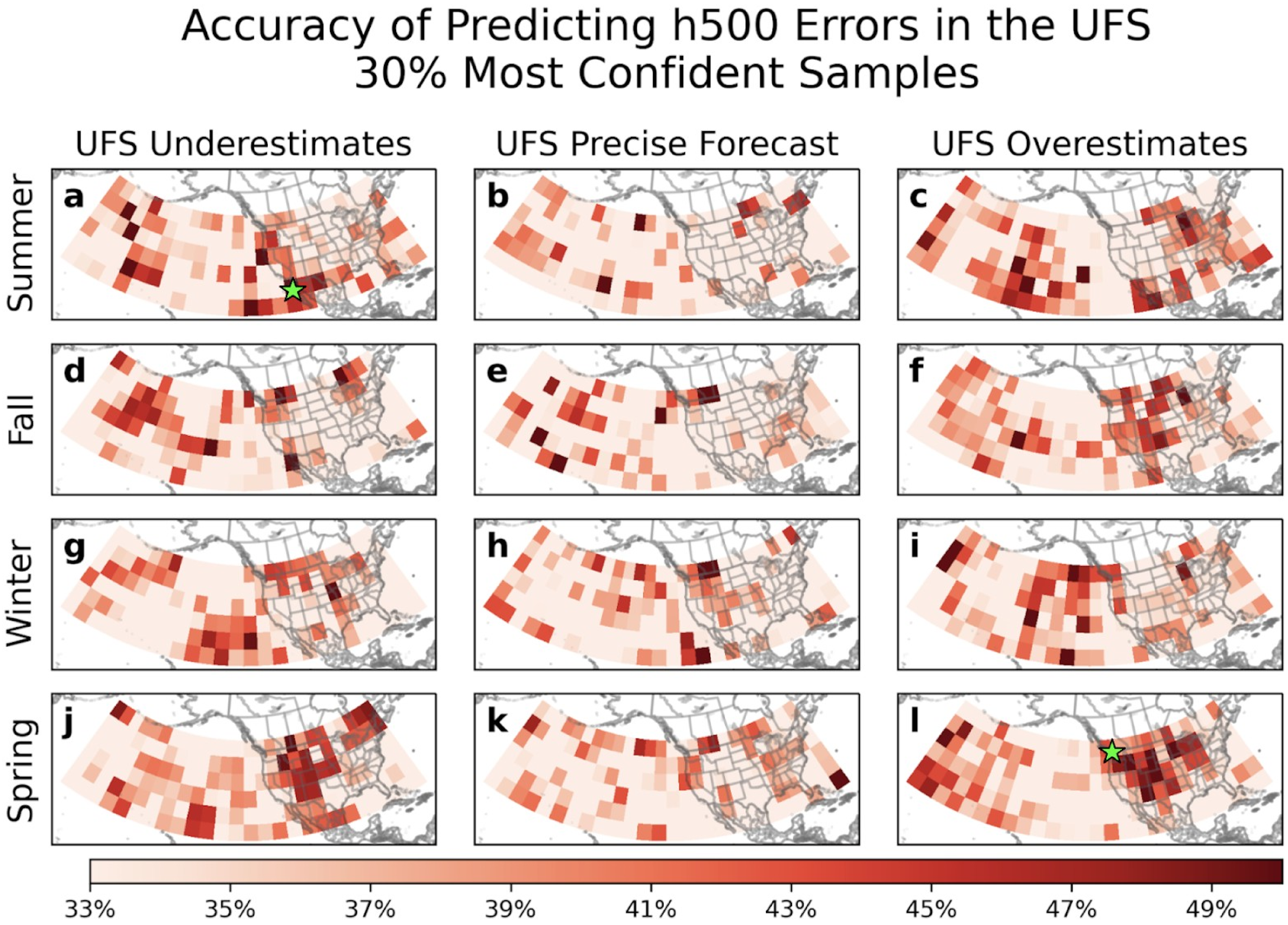
Cahill, Barnes, Maloney, Harr, Madaus, Sain (2024), *WAF*

Using Neural Networks to Predict Errors in the Unified Forecast System (UFS) on S2S Timescales



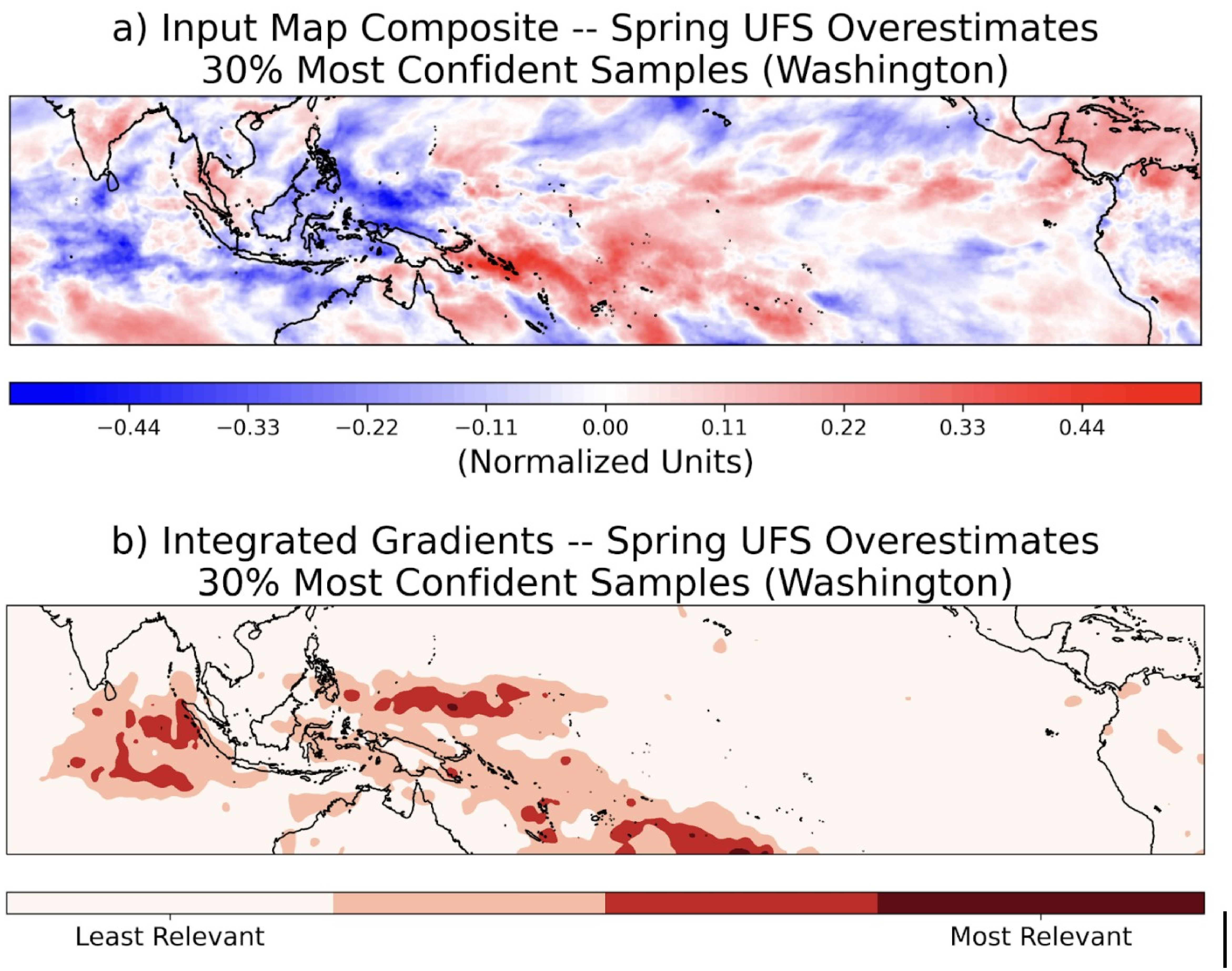
Cahill, Barnes, Maloney, Harr, Madaus, Sain (2024), *WAF*

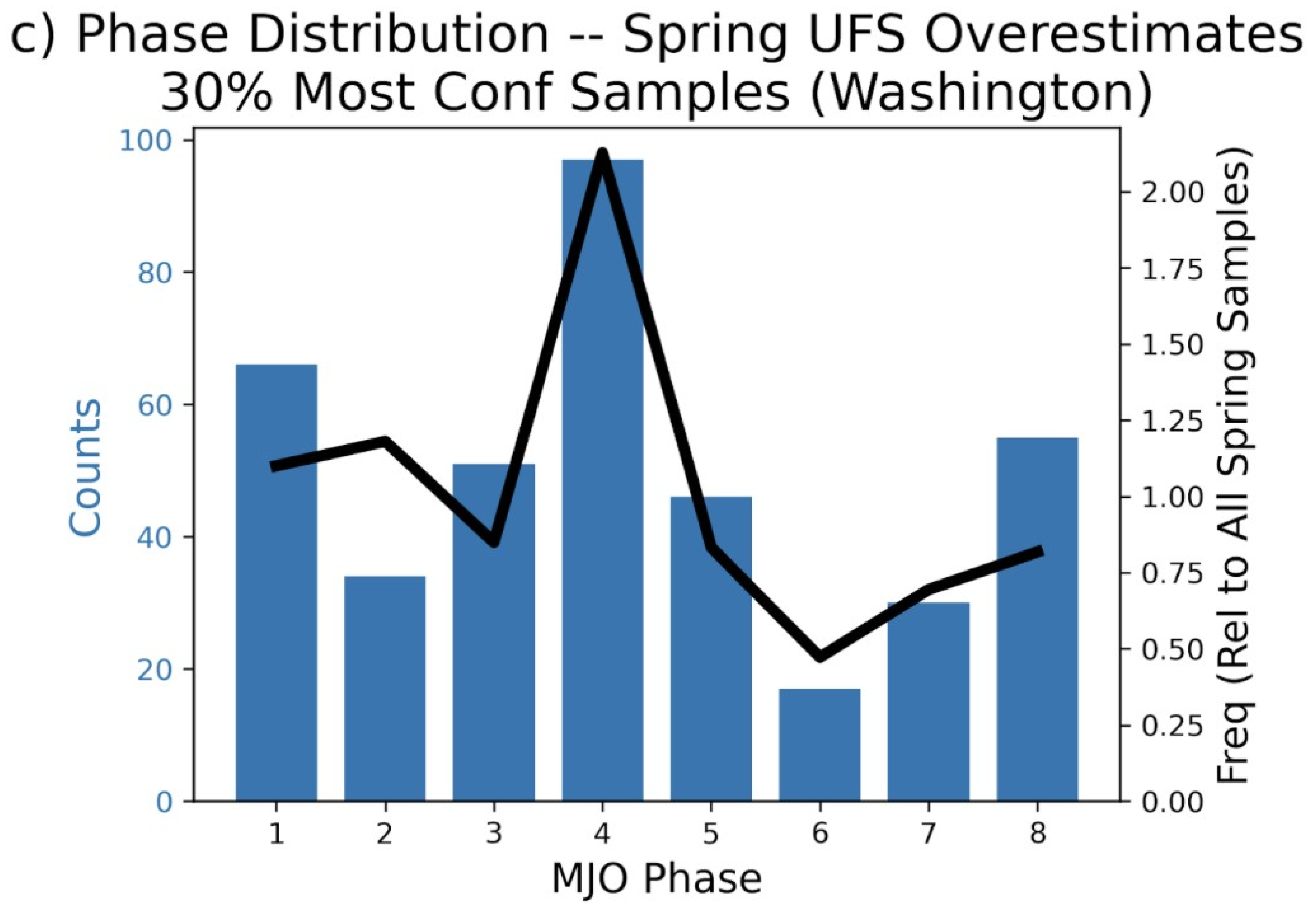
Using Neural Networks to Predict Errors in the Unified Forecast System (UFS) on S2S Timescales



Cahill, Barnes, Maloney, Harr, Madaus, Sain (2024), *WAF*

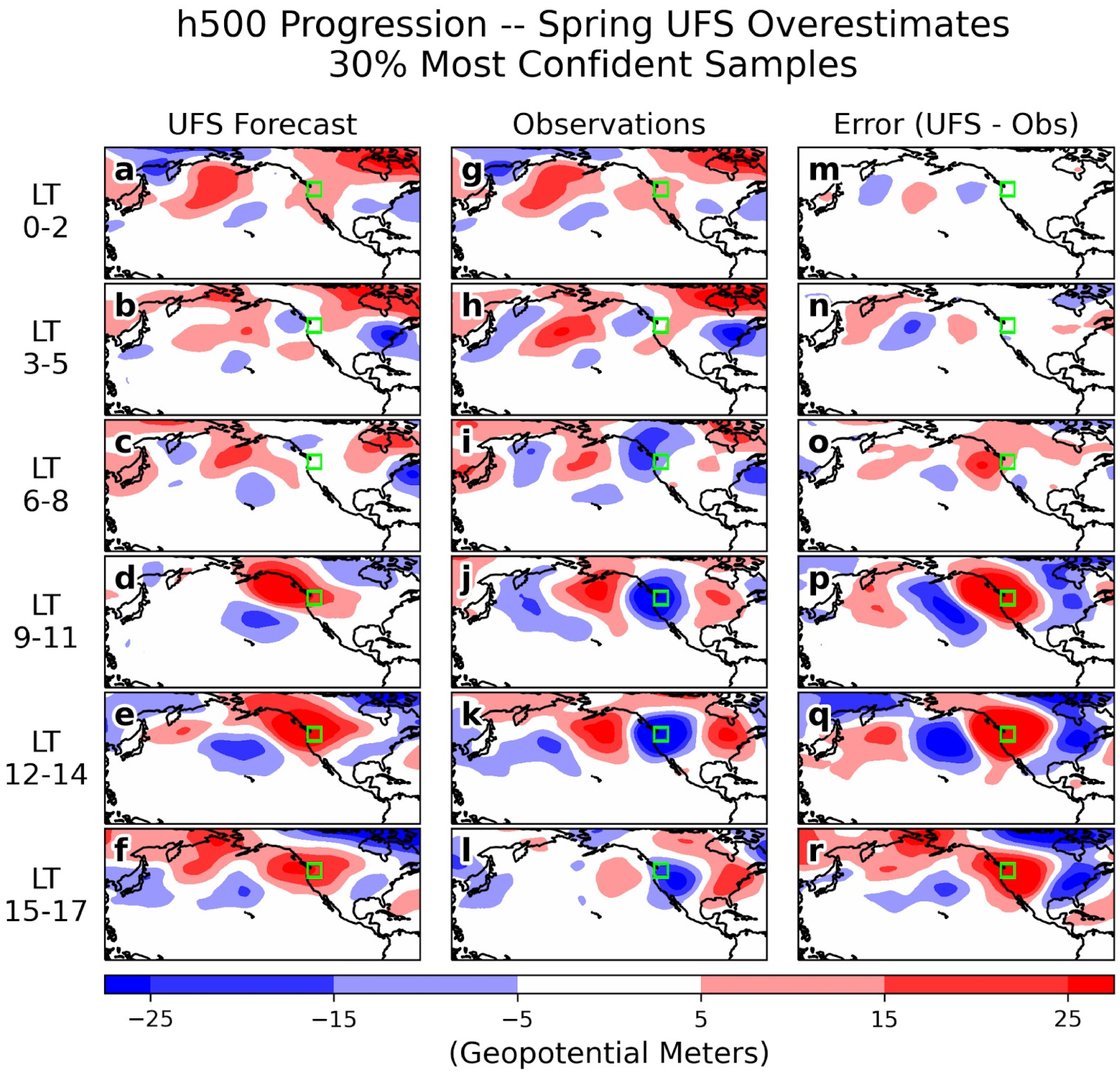
Using Neural Networks to Predict Errors in the Unified Forecast System (UFS) on S2S Timescales



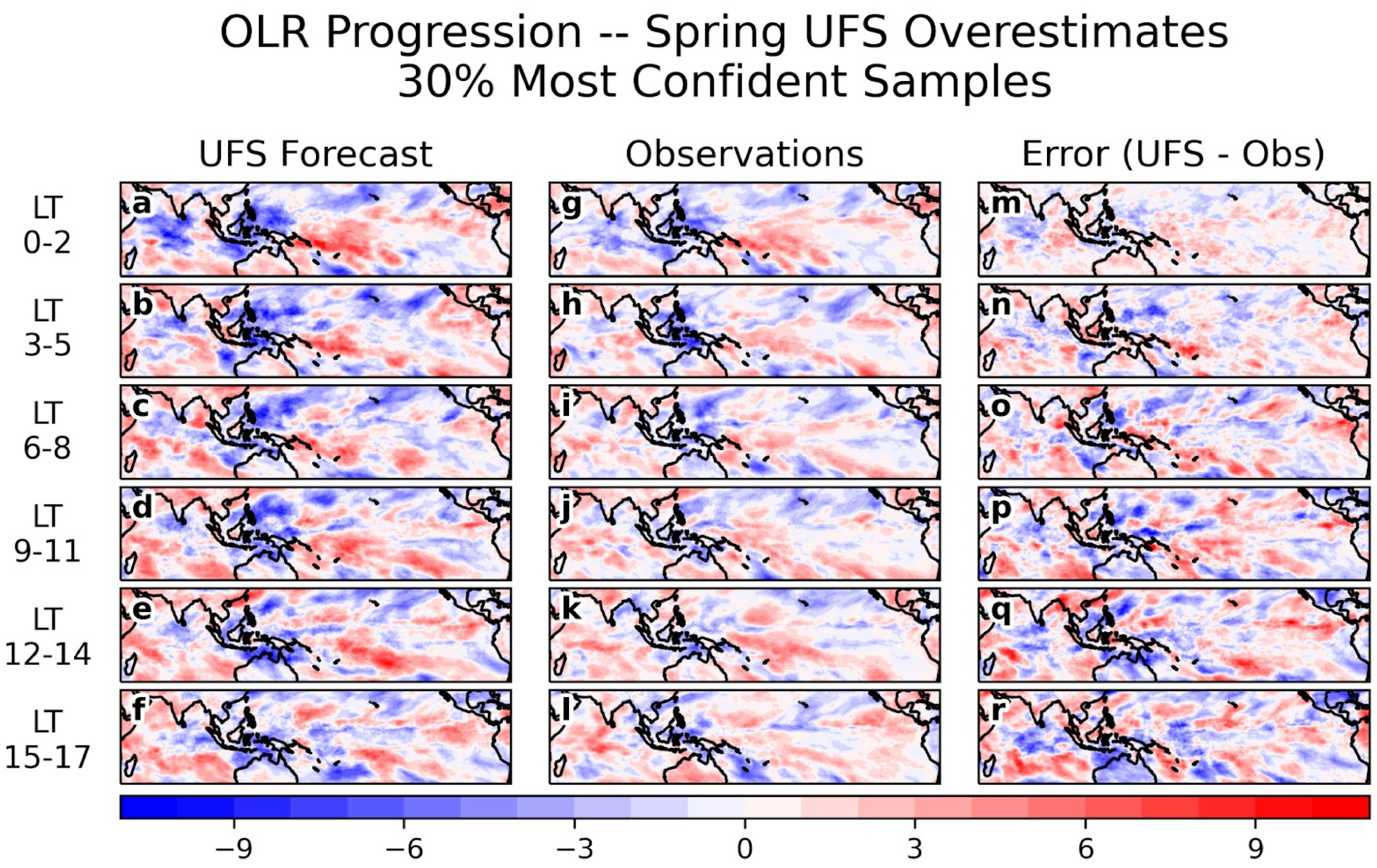


Cahill, Barnes, Maloney, Harr,
Madaus, Sain (2024), *WAF*

Using Neural Networks to Predict Errors in the Unified Forecast System (UFS) on S2S Timescales



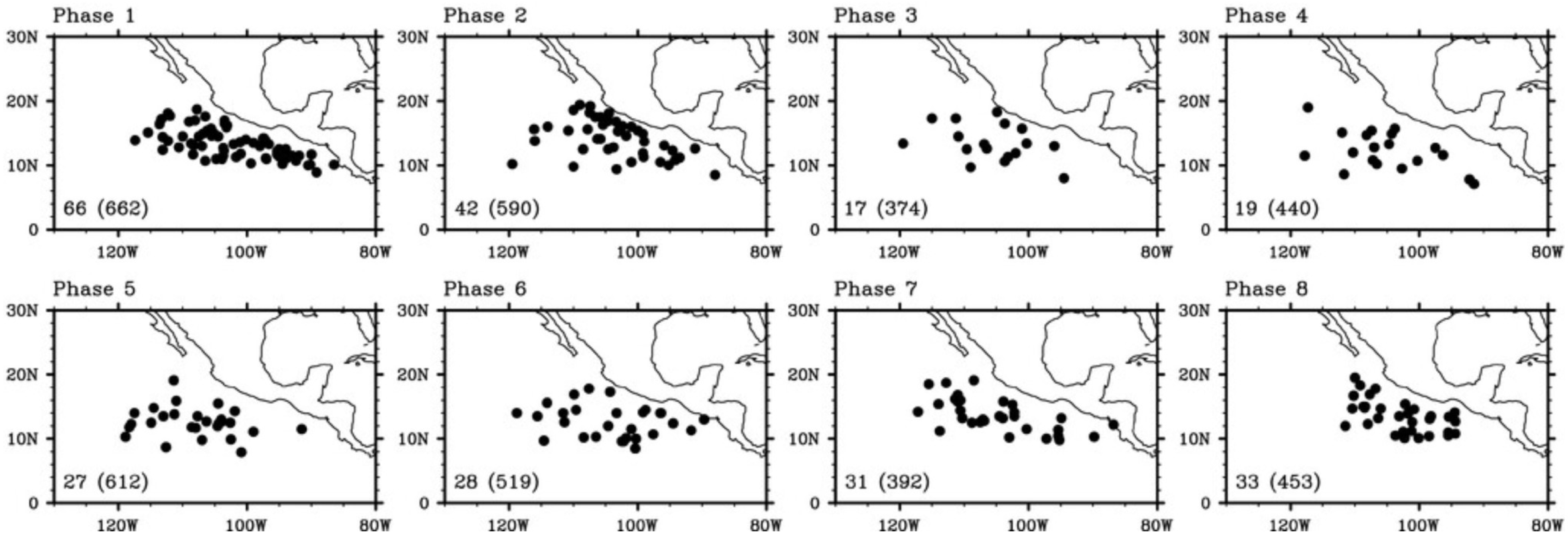
Associated with MJO Phase 4



Cahill, Barnes, Maloney, Harr, Madaus, Sain (2024), *WAF*

Part 2: Subseasonal TC Genesis Prediction

TC Genesis vs. MJO Phase



Henderson and Maloney (2013)

UNIFIED FORECAST SYSTEM RESEARCH TO OPERATIONS PROJECT

(UFS S2S PROTOTYPES 5-8)

Data Source: https://vlab.noaa.gov/web/ufs-r2o/dataproducts

Re-forecast period	Forecast lead	Frequency of initialization
April 2011 to March 2018 (7 years)	6 hourly output, out to 35 days	1st and 15th of each month

- The re-analysis dataset, ERA5, is utilized to compare with the UFS model.

- The All-season Real-time Multivariate (RMM) MJO Index (Wheeler-Hendon) is obtained from the The Centre for Australian Weather and Climate Research.

  Data Source: http://www.bom.gov.au/climate/mjo/

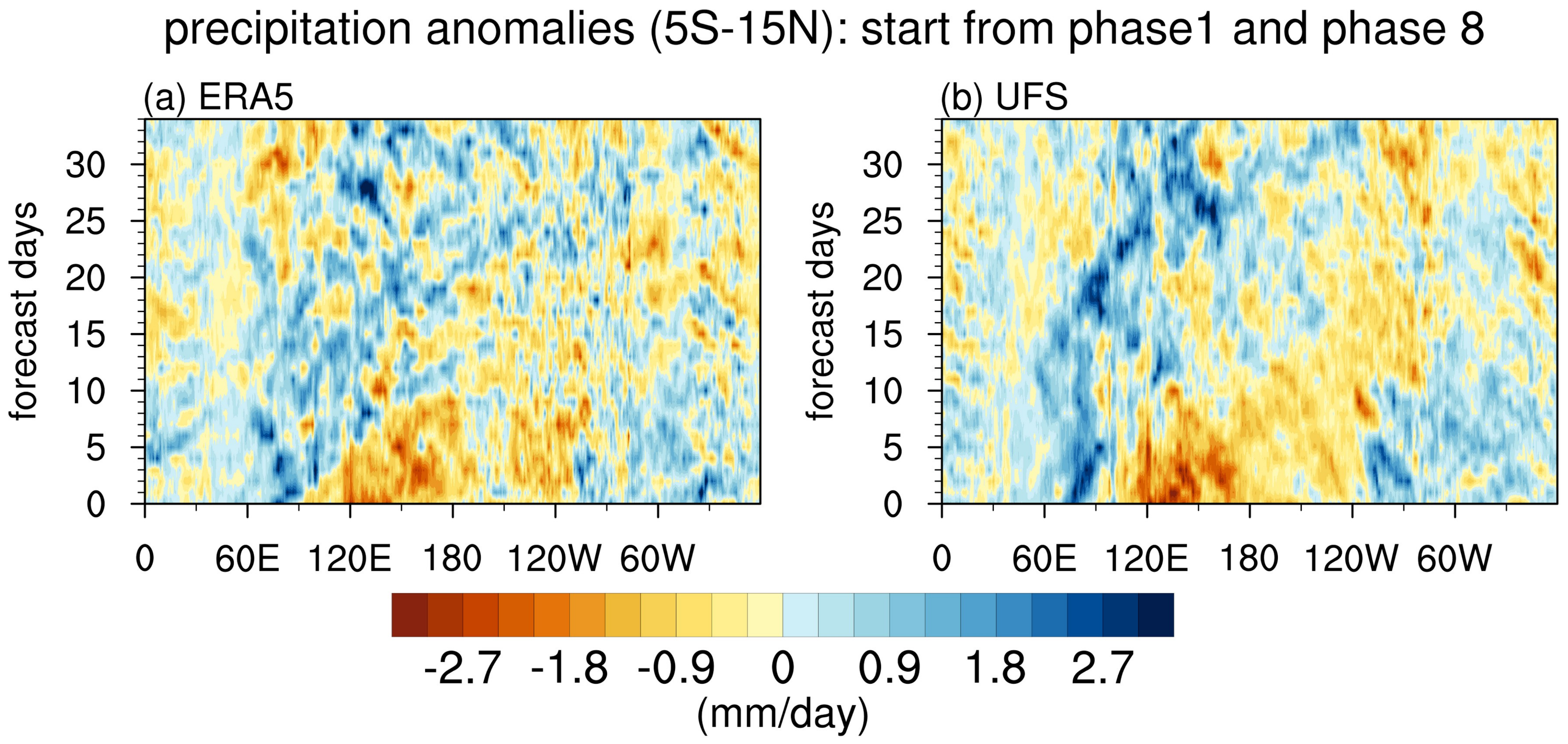
- The TC information is obtained from the International Best Track Archive for Climate Stewardship (IBTrACS) dataset.

  Data Source: https://www.ncei.noaa.gov/products/international-best-track-archive

Yu-Cian Tsai

Phase 1, 8 (avg): Composite

22 cases



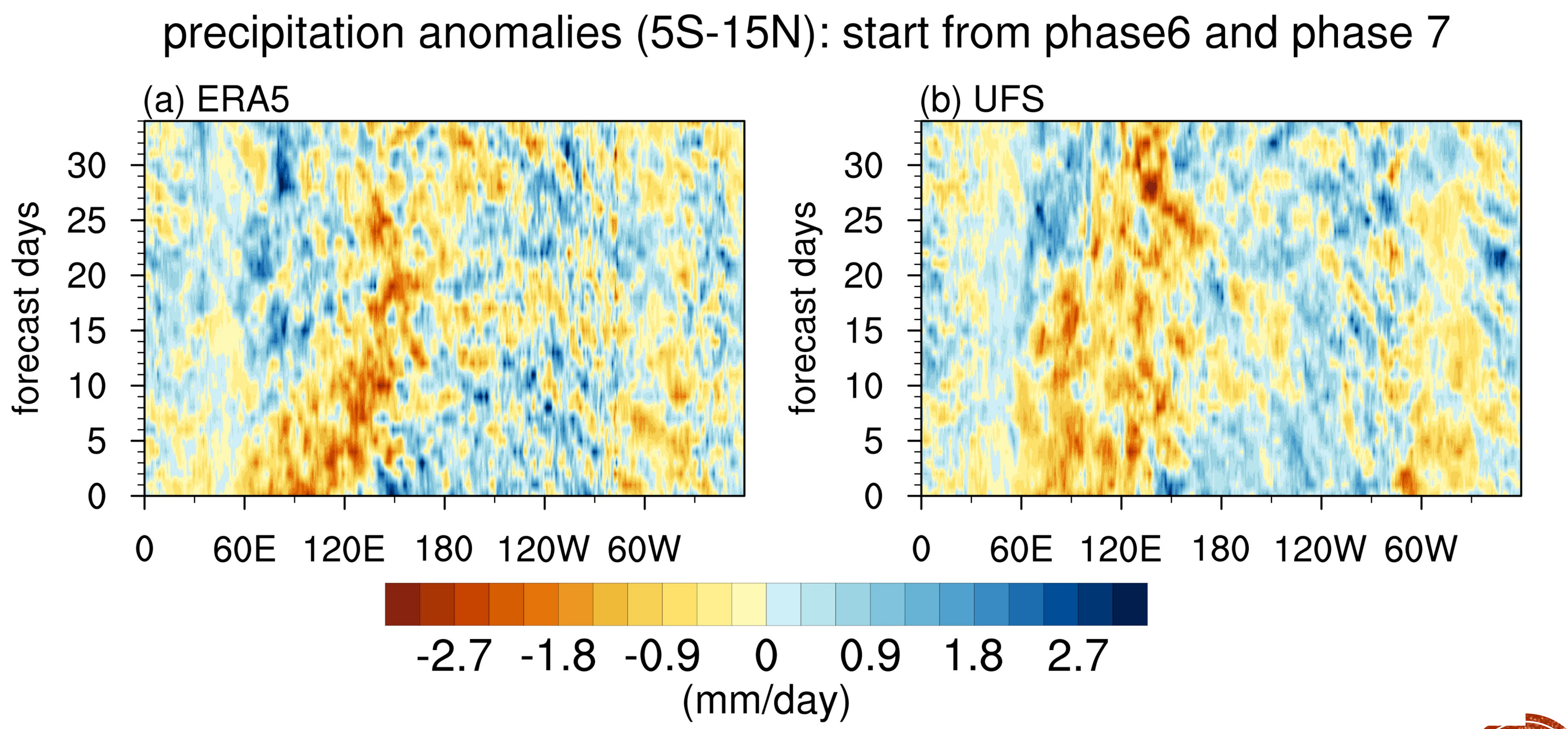
Boreal Summer Only

Model MJO Propagates Too Slowly and Is Too Strong?

Produces timing and strength errors of MJO precipitation in east Pacific

Phase 6, 7 (avg): Composite

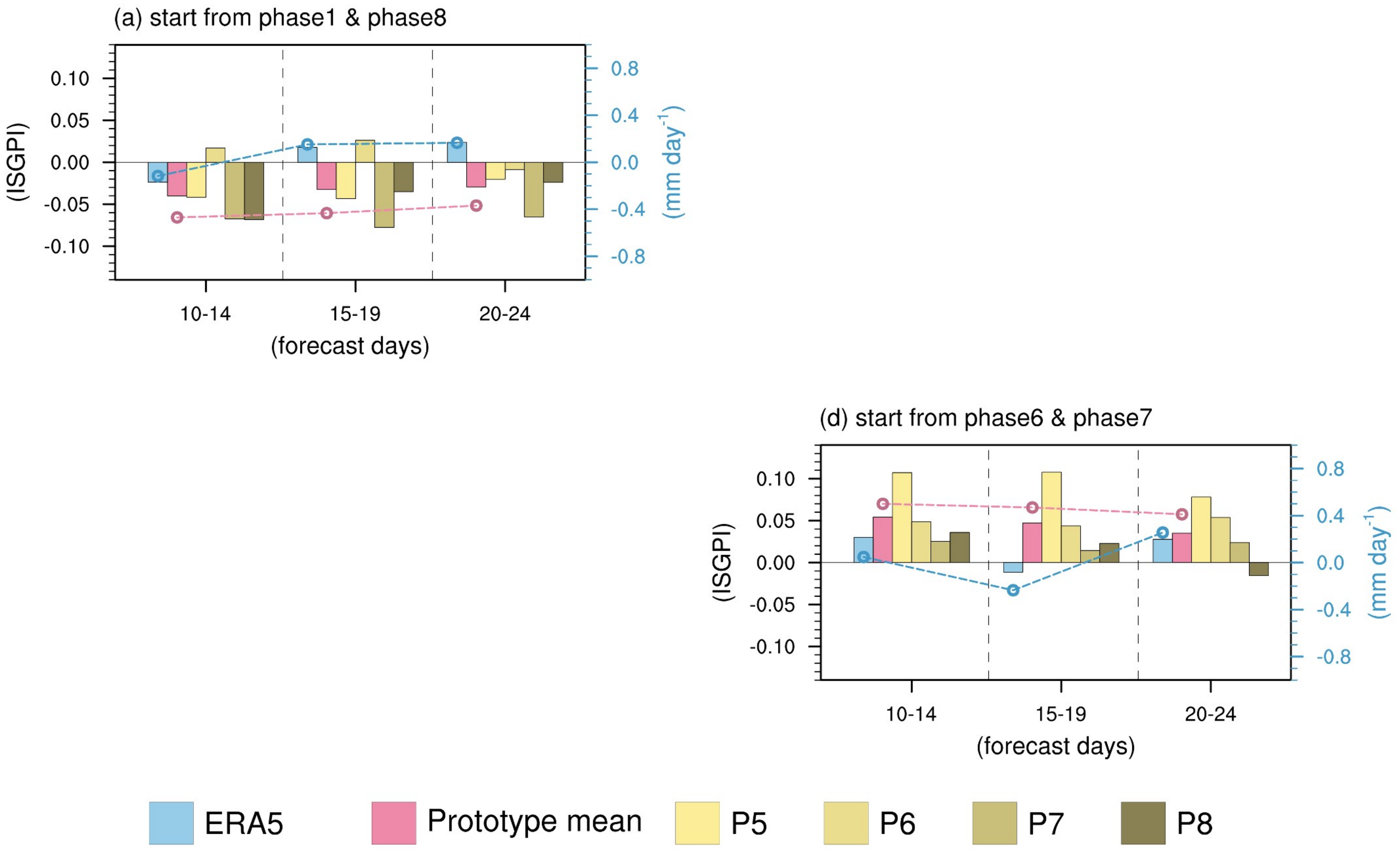
20 cases



Genesis Potential Forecasts

$$(-0.67) \times \omega_{500} + (0.24) \times f\zeta_{r850} + (-0.02) \times V_{zs}$$

wavenumber 1-10 ISGPI and Precipitation anomalies (10-20N, 90-120W)



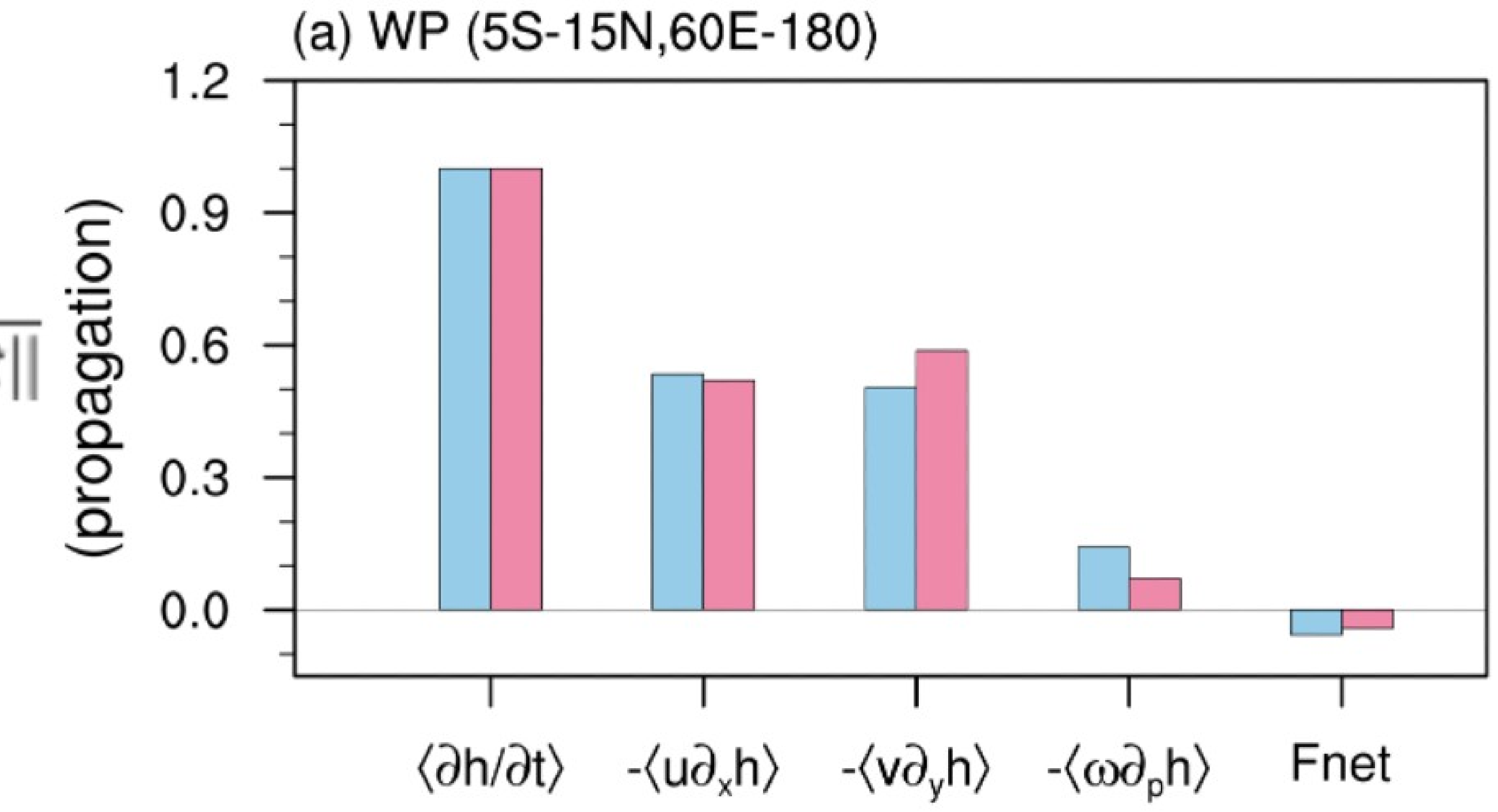
Units: TC genesis per day per 10°x10° grid

Using intraseasonal genesis potential index: Moon et al. (2018)

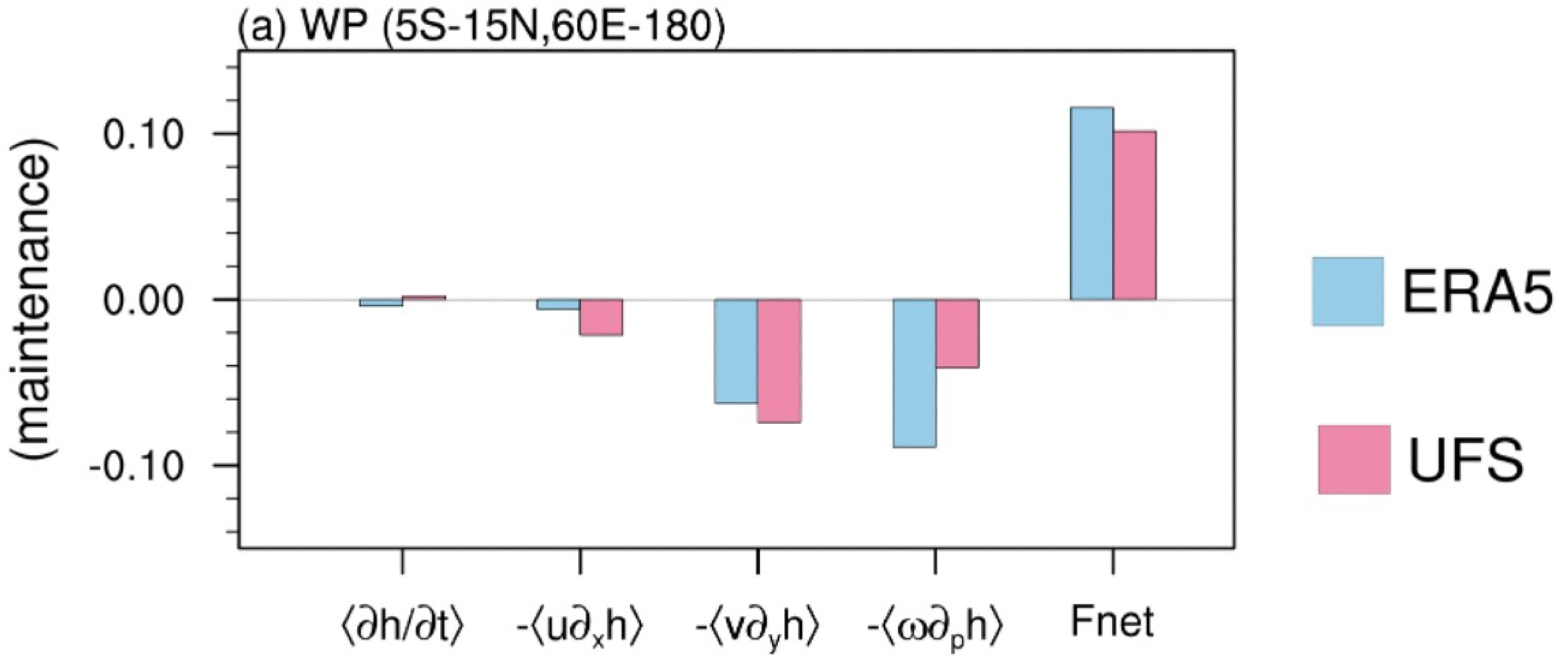
Moist Static Energy Variance Budget (West Pacific)

Subseasonal to seasonal (day10~day24)

$$S_p(F)=\frac{||F\cdot \partial P'/\partial t||}{L_v||\partial P'/\partial t\cdot \partial P'/\partial t||}$$

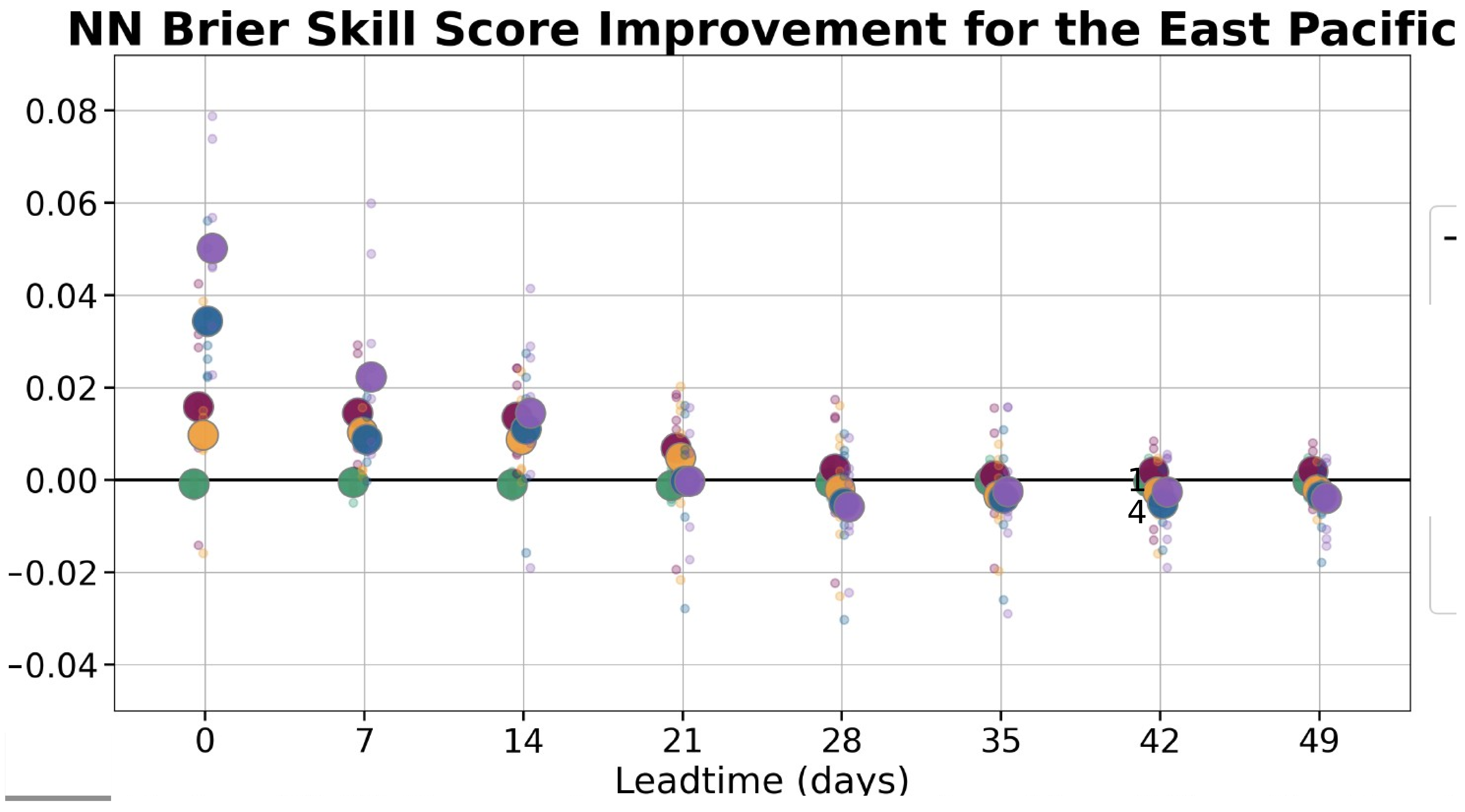


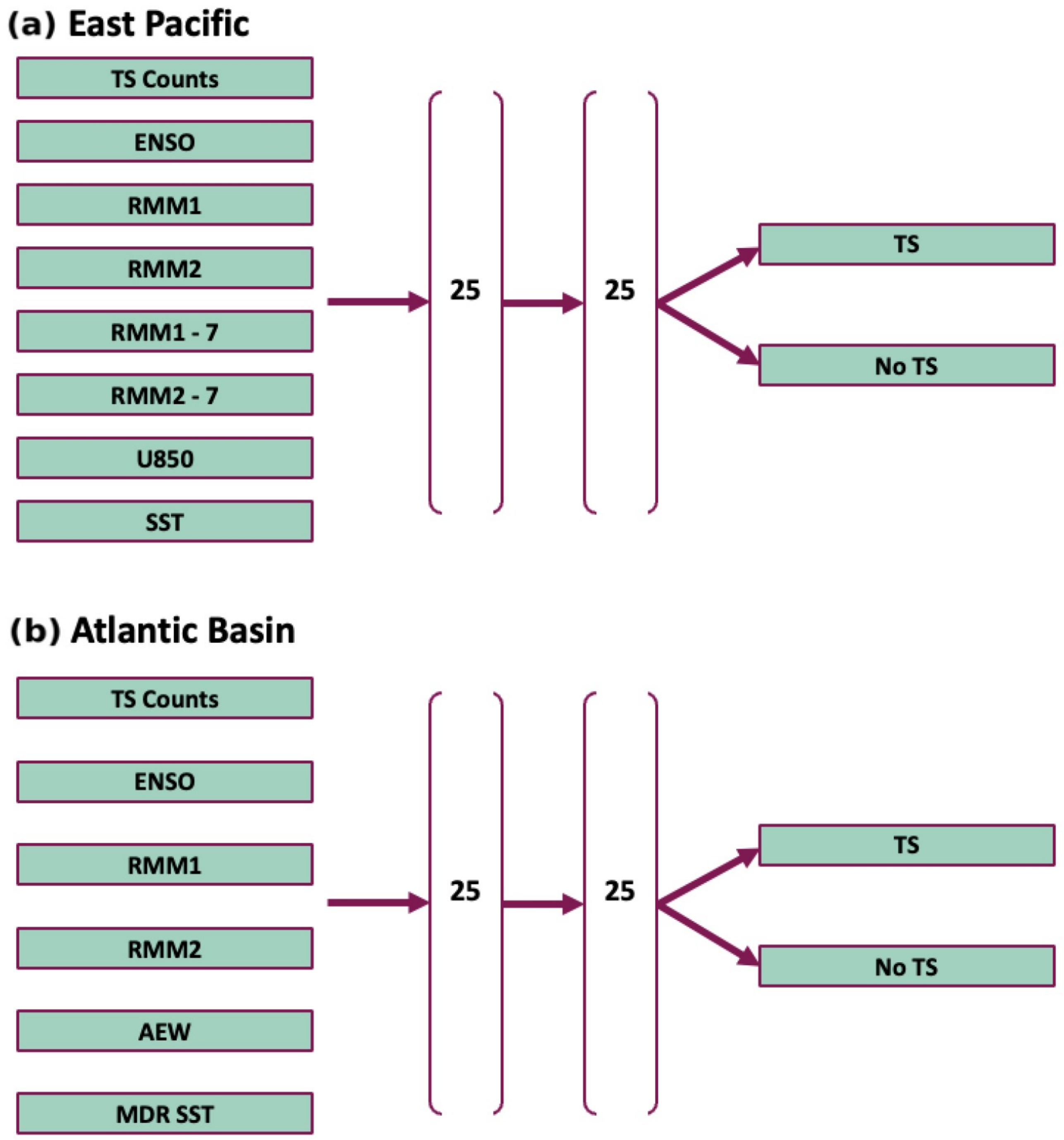
$$S_m(F)=\frac{||F\cdot P'||}{L_v||P'\cdot P'||}$$



e.g. Wing and Emanuel (2014)

Part 3: ENSO and MJO as Predictors Lead to east Pacific TC Cyclogenesis Prediction Skill





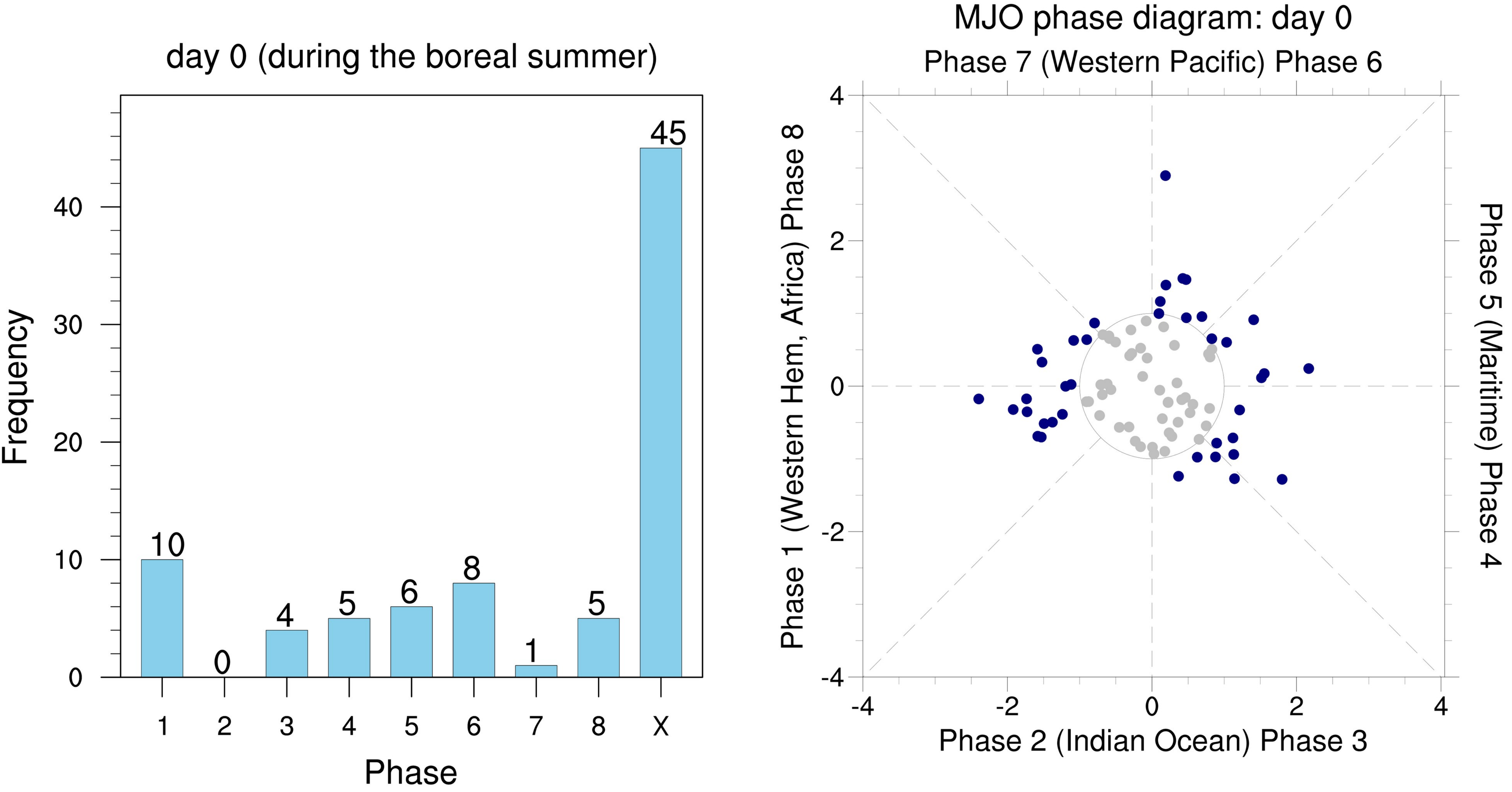
Carlo Frontera, Barnes, Maloney (2024), *WAF*

Conclusions

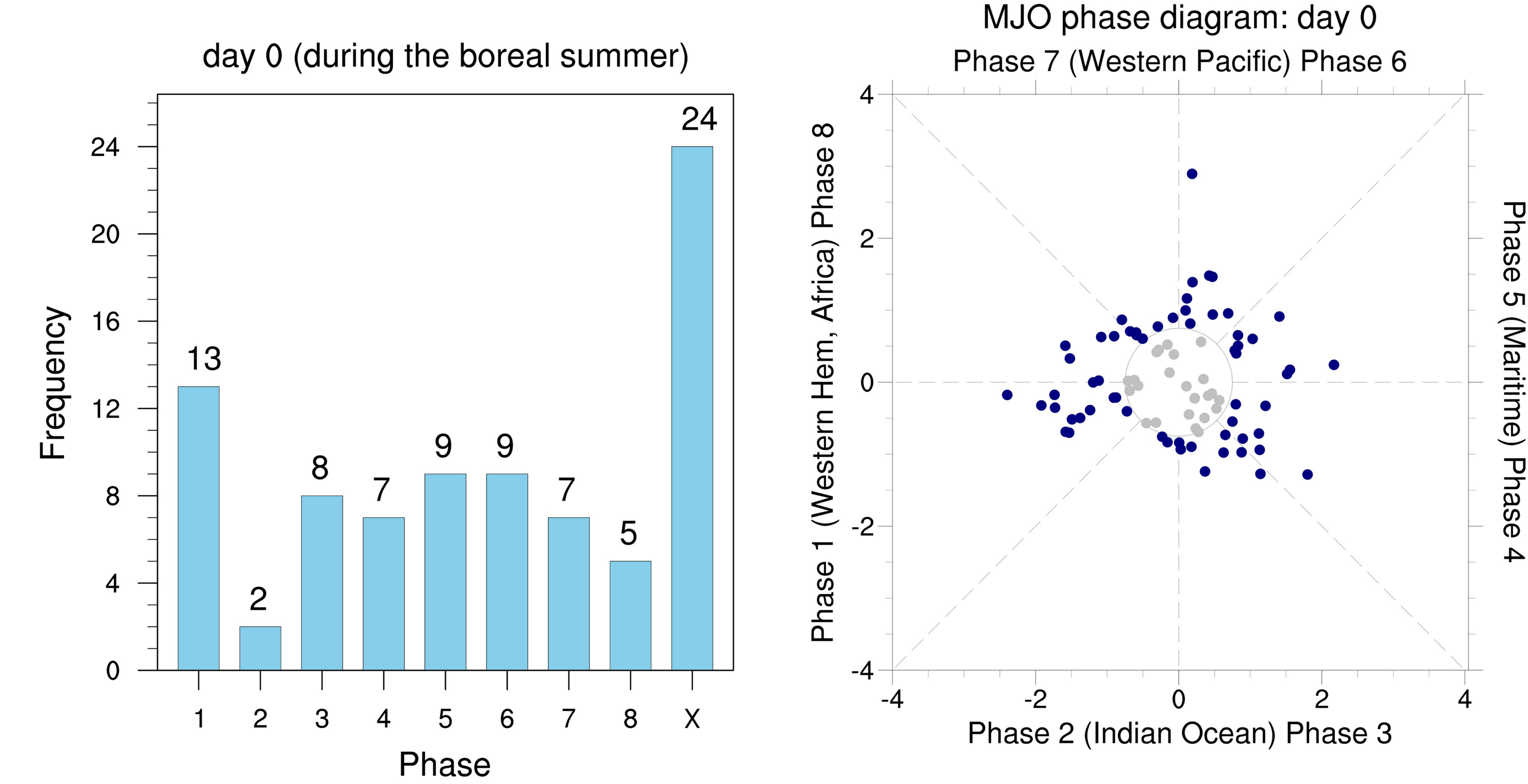
- A NN model accurately identifies underestimates of Spring 2-week forecast geopotential heights in the Pacific Northwest in MJO phase 4 stemming from the UFS's failure to correctly forecast teleconnection patterns.
- Phases 1 and 2 high accuracy during summer
- When the UFS is initiated in MJO phases with a strong dipole of convection across the Maritime Continent, prominent subseasonal UFS forecast errors result in the Western Hemisphere that affect cyclogenesis predictions
- NN model utilizing ENSO and MJO indices and other local environmental information are also used to predict east Pacific and Atlantic

Thanks

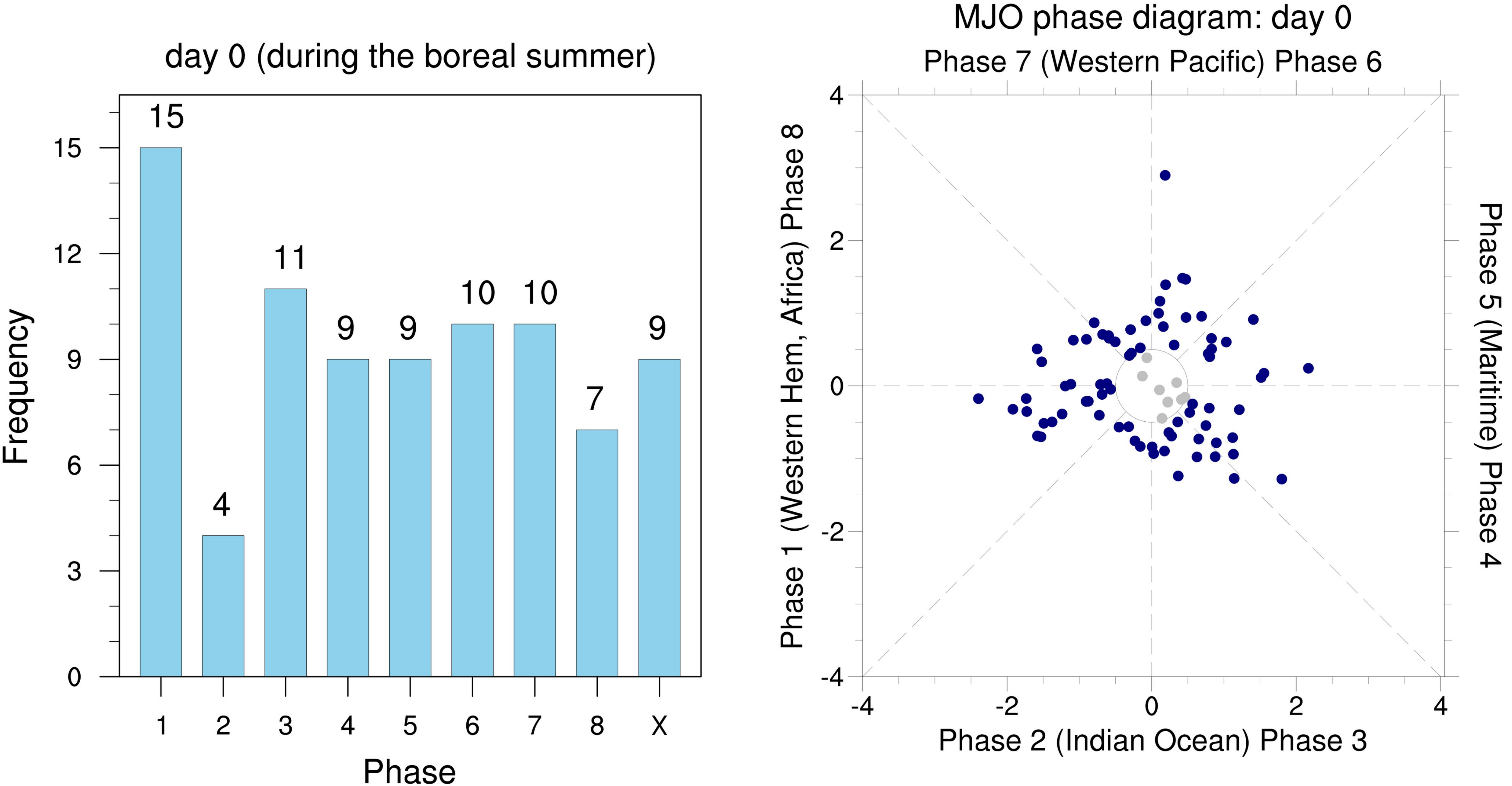
(a) RMM Index >= 1



(b) RMM Index >= 0.75



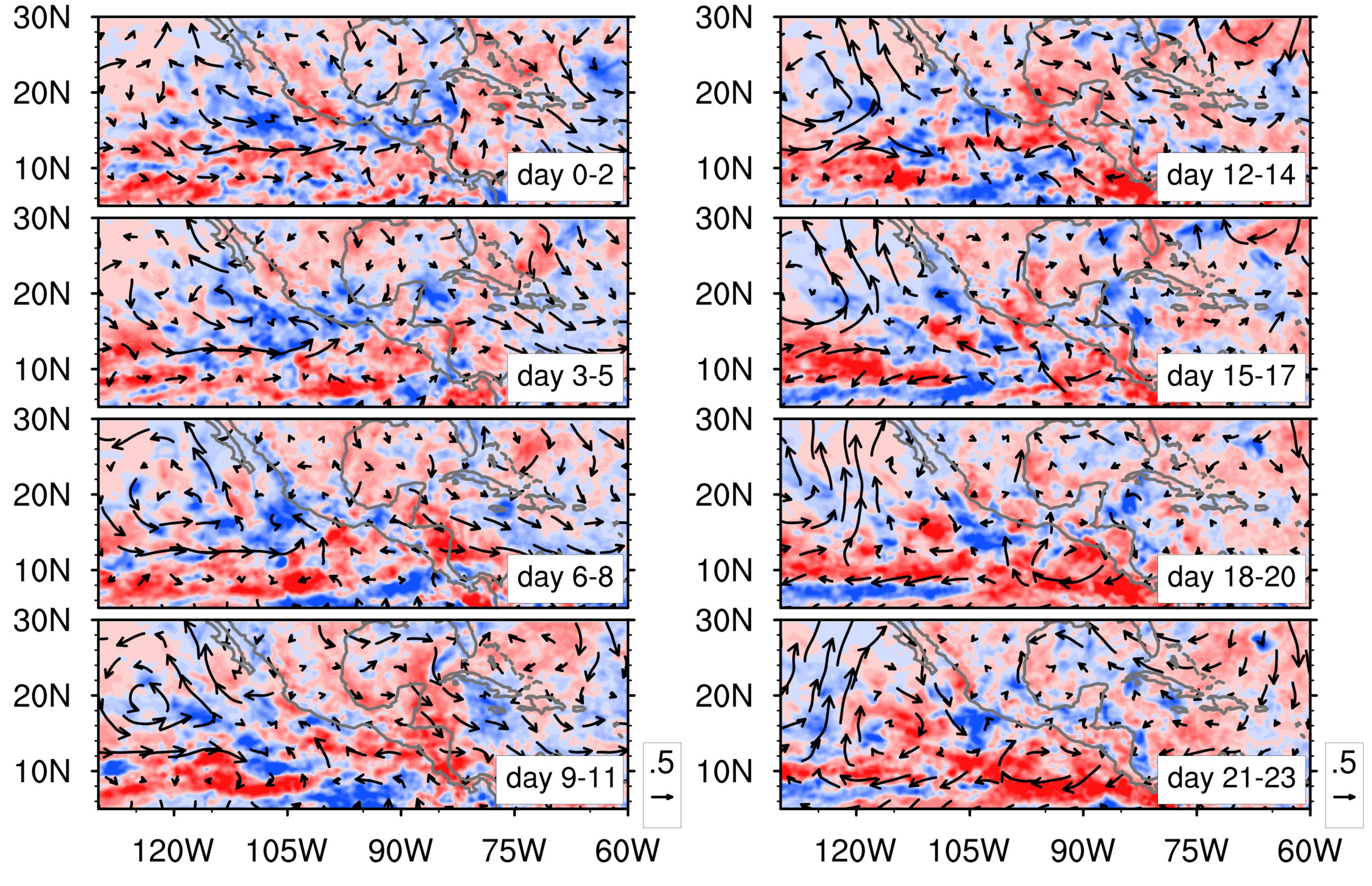
(c) RMM Index >= 0.50



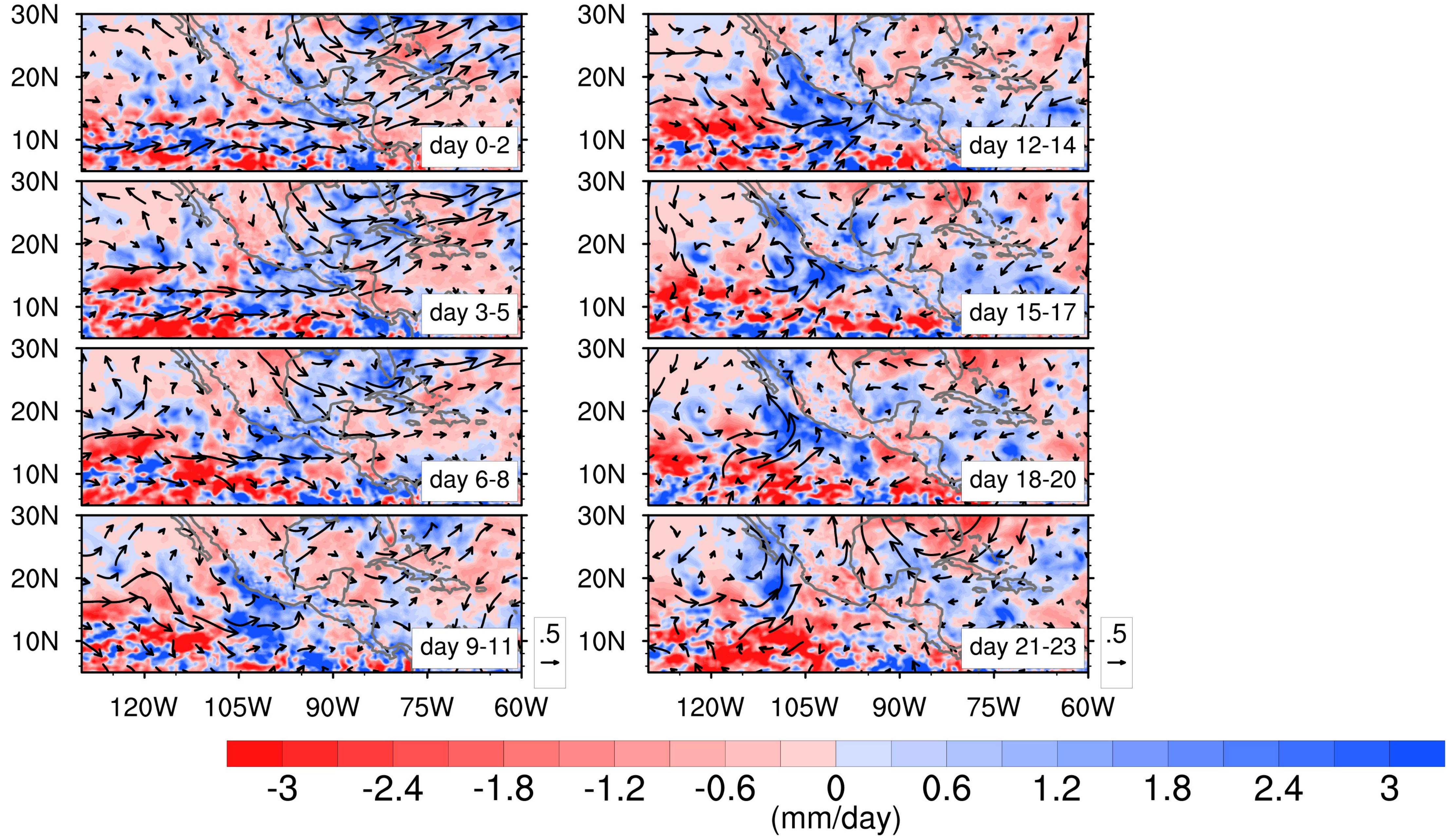
- First, we generate statistics of the MJO phases at model initialization days (day 0) for different RMM index thresholds.
- “X” means that MJO is inactive (RMM index < threshold) at those initialization days.
- To obtain a large enough sample size of MJO cases starting at different phases, we decided to set the threshold of RMM index to 0.5.

Boreal Summer Only

UFS: MJO start from phase 1 and phase 8 (precipitation & UV850)



ERA5: MJO start from phase 1 and phase 8 (precipitation & UV850)



e.g. After Phases 1 and 8, UFS wants to too strongly suppress east Pacific precipitation and produces easterly low-level anomalies

Would unrealistically suppress east Pacific cyclogenesis.

AI Architecture

- input layer consists of OLR across the tropics resulting in a total of 48,581 vectorized and normalized grid points/input nodes. To account for the large number of input nodes we apply a “dropout layer” following the input layer

- one hidden layer of 60 nodes

- rectified linear unit is used as the activation function. Subsequently, the output layer is composed of three nodes, representing the three classes of error: UFS underestimates, UFS precise estimates, and UFS overestimates.

- a softmax activation function is applied to the output layer which remaps the values of the three-node output such that they sum to one.

- The largest value of the three nodes is then defined as the network’s predicted class.

- Additionally, we associate the value of the winning class with its predicted value, which we call “model confidence”. This “model confidence” quantifies the neural network's certainty in its classification decision for each sample. A higher value, closer to 1, indicates a stronger confidence in the network's prediction, suggesting that the model perceives clear, definitive features in the data that align with the predicted class.

- Conversely, a lower confidence value, like 0.6, implies uncertainty and less distinct features in the data for that winning class. The batch size is set to 32 and the neural network is run for 100 epochs with a learning rate of 0.0001.

AI Architecture

- The network is trained on 16 years of data (834 or 833 samples depending on the presence of leap years) and the remaining 4 years of data (208 or 209 samples) are used as testing data.

- In an attempt to ameliorate potential issues of a small testing set, a five-fold cross-validation technique is employed. This approach incorporates each consecutive testing fold, resulting in 1042 testing samples for analysis.

- To ensure robustness, the neural network is run for six different random initialization seeds of starting weights for each training-testing fold. Unless otherwise stated, our analysis shown here is performed solely on the testing data averaged across all cross-validation folds and seeds. This average is computed after all folds and seeds are run.

- This setup is applied to all 156 grid points across the North Pacific and continental United States such that each location is trained using 30 different networks.

AI Performance

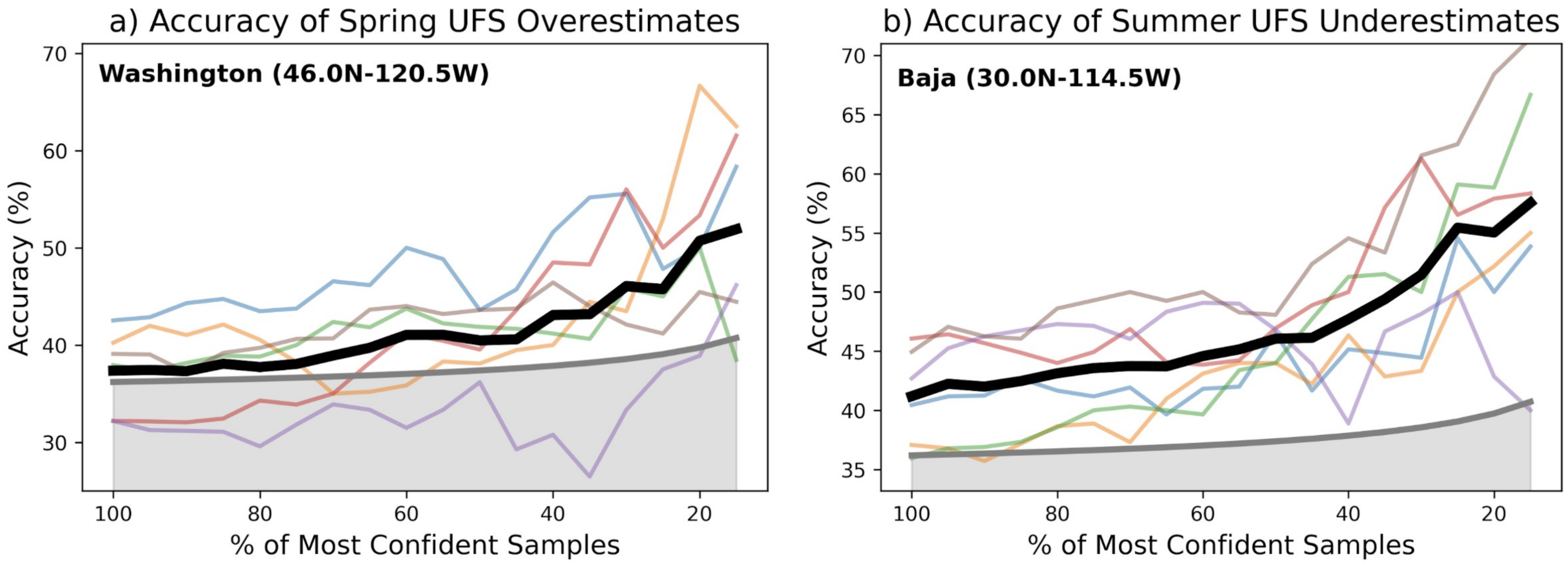


Figure 3: Neural network accuracy as a function of confidence for a singular location in (a) Washington State (46°N - 120.5°W) and (b) Baja California (30°N - 114.5°W). Light colored lines represent the six neural networks with cross-validation included, the thick black line represents the neural network average across the six networks, and the gray region represents the 95% confidence interval based on random chance using the number of samples at each ‘percent most confident’.