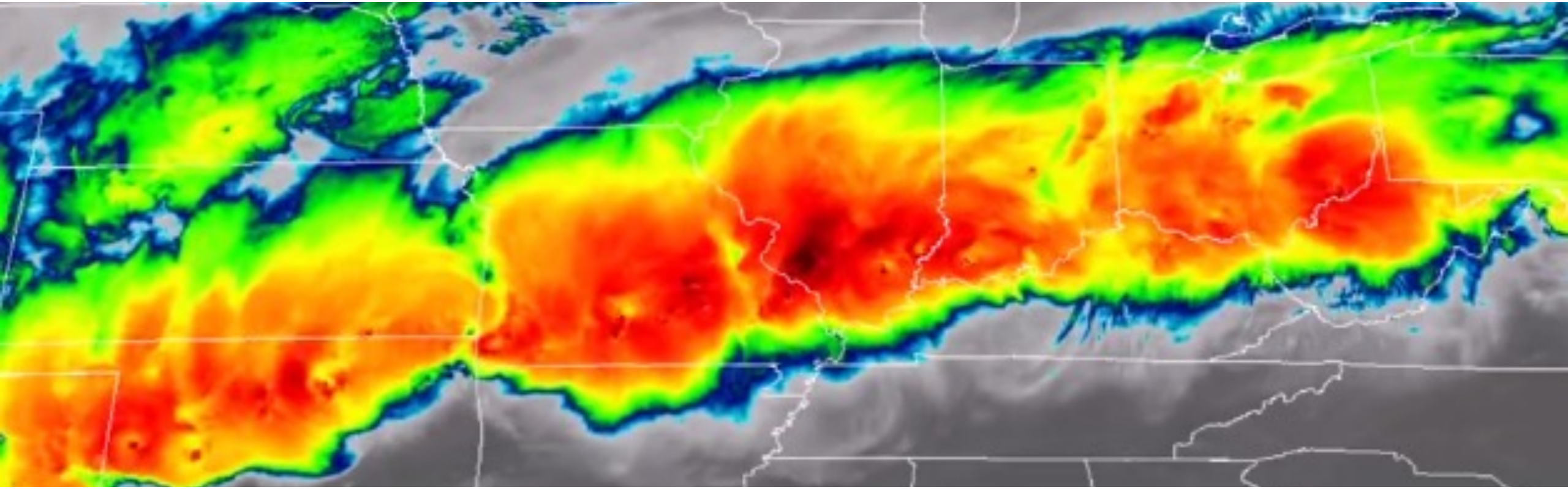
Predicting Subseasonal Extreme Precipitation Events Across the United States



Jason C. Furtado, Ty Dickinson, and Michael Richman
University of Oklahoma School of Meteorology

CIWRO Workshop on S2S Prediction for High-Impact Weather Events

ICER #1663840

7 October 2022

PRES²iP

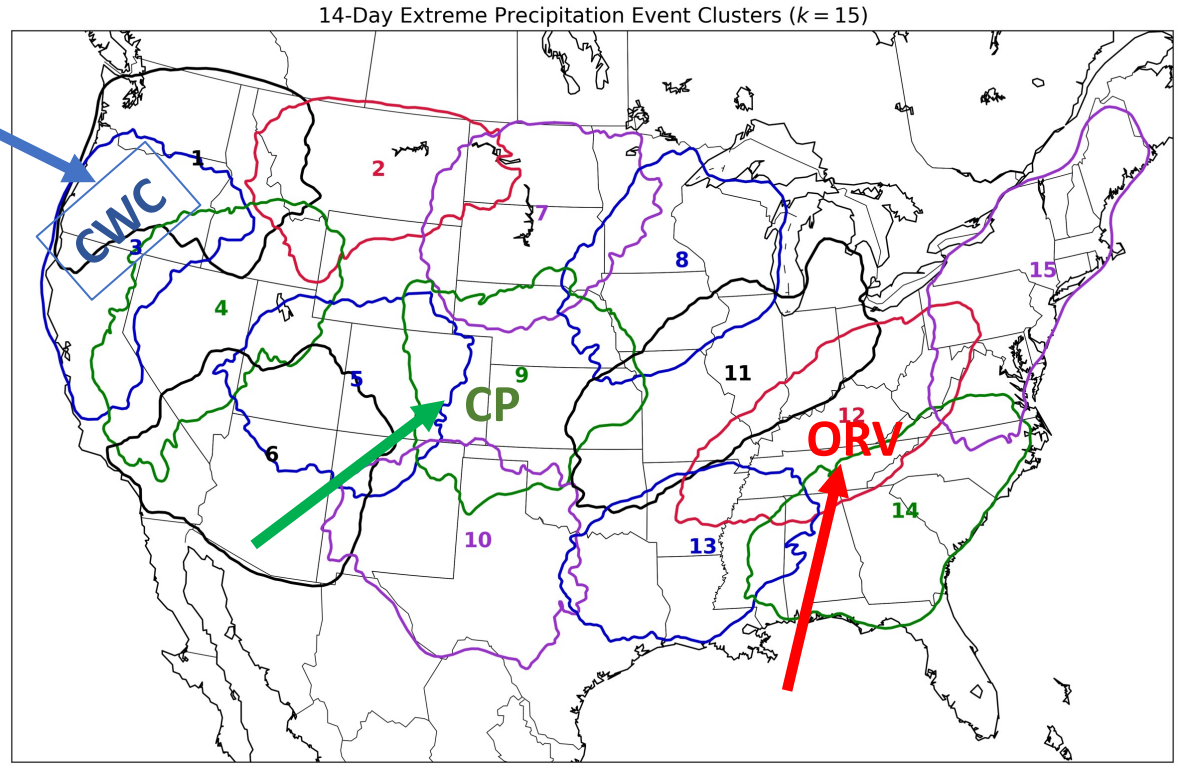
Prediction of Rainfall Extremes at Sub-Seasonal to Seasonal Periods

Goals of the Project:

1. Define databases of S2S extreme events.
2. Quantify statistical and dynamical links between S2S extreme events and synoptic-scale and global scale precursors.
3. Improve capability to predict S2S extreme events.
4. Increase communication between research scientists and stakeholder communities.

S2S Extreme Precipitation Event Database

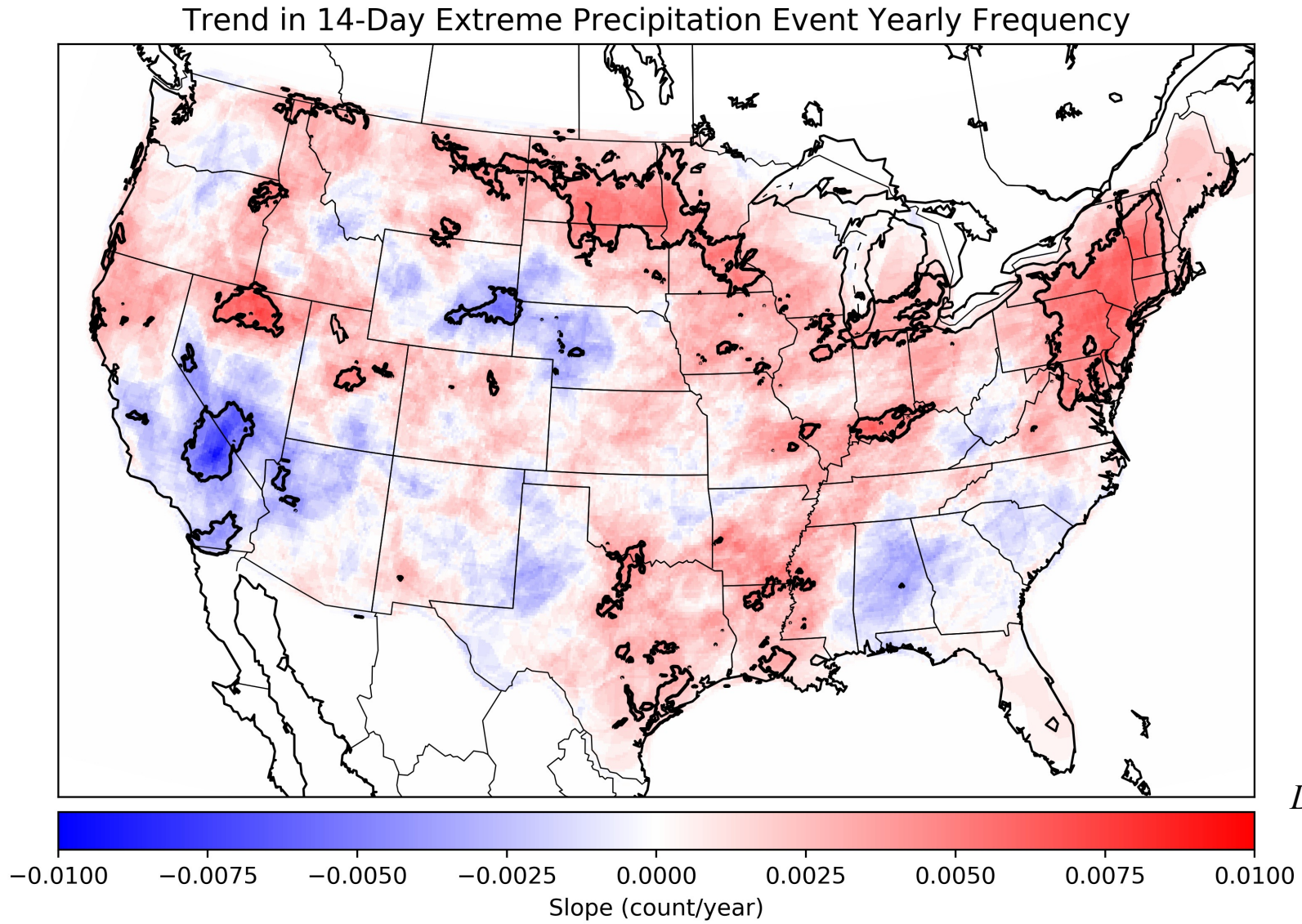
- Created in *Dickinson et al.* [2021] for 14-day events (also have 30-day events available).
  - Totals exceed 99th percentile, more than 7 days of above normal daily precipitation, have areal extent ≥ 200,000 km^2.
- Events clustered by region using *k*-means clustering.
- 1950 – near present



Dickinson et al. [2021]

Downloadable database:
https://github.com/tydickinson29/PRES2iPpy/tree/master/pres2ippy/databases

Trends in 14-Day Extreme Precipitation Events



Dickinson et al. [2021]

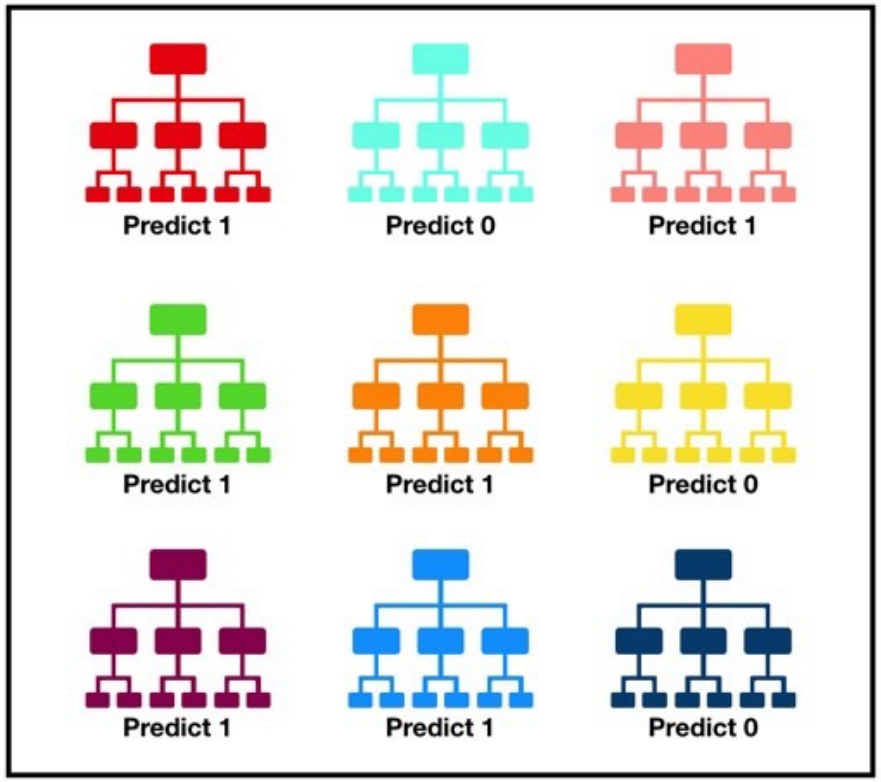
PRES²iP

Prediction of Rainfall Extremes at Sub-Seasonal to Seasonal Periods

Goals of the Project:

1. Define databases of S2S extreme events. ✅
2. **Quantify statistical and dynamical links between S2S extreme events and synoptic-scale and global scale precursors.**
3. **Improve capability to predict S2S extreme events.**
4. Increase communication between research scientists and stakeholder communities.

Toward An ML Approach for Forecasting 14-Day Extreme Precipitation Events

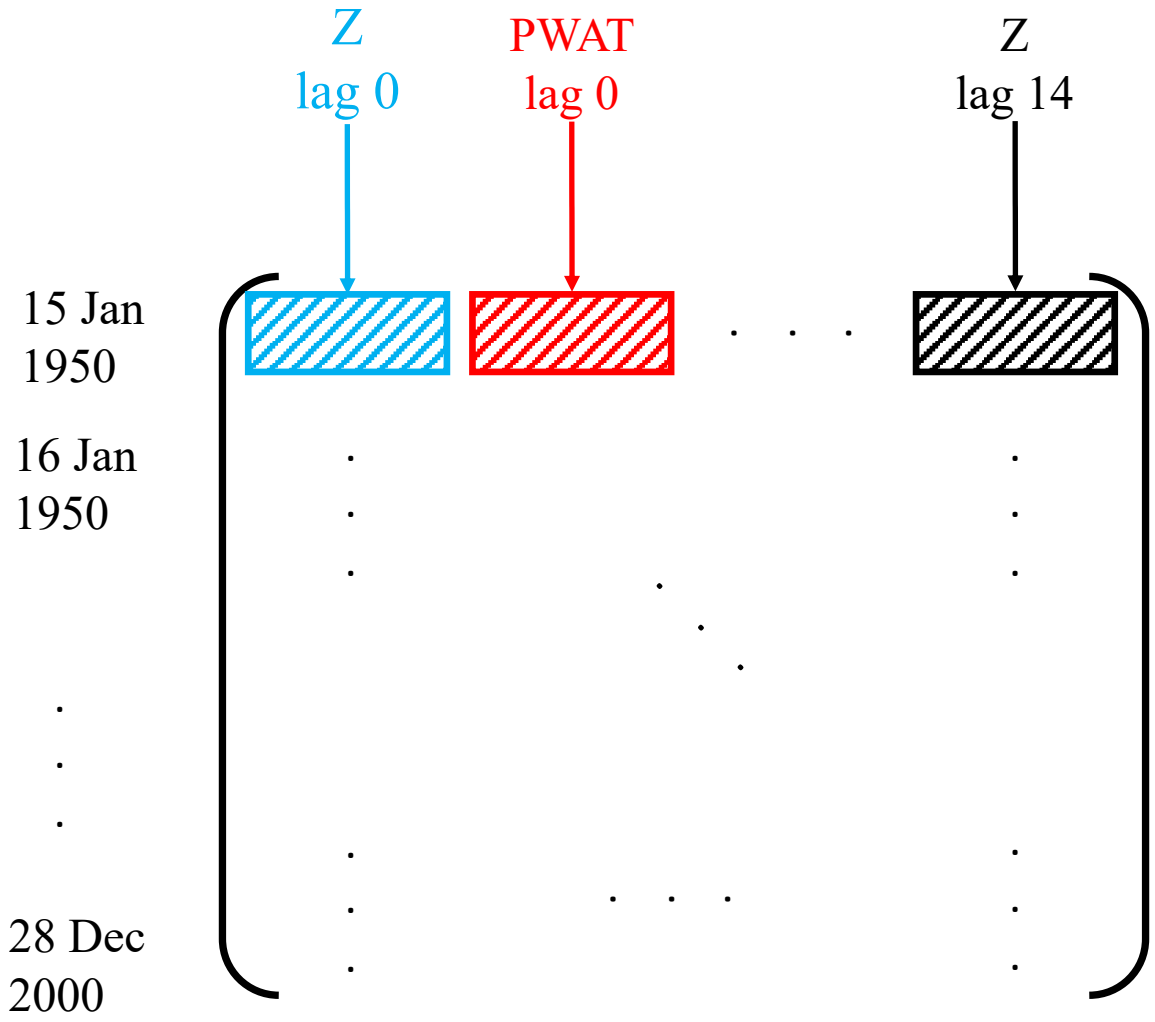


Research Questions:

1. How well can a random forest approach classify days as being extreme or not extreme?
2. Which atmospheric variables (and where) are most important in classifying extreme versus non-extreme days?

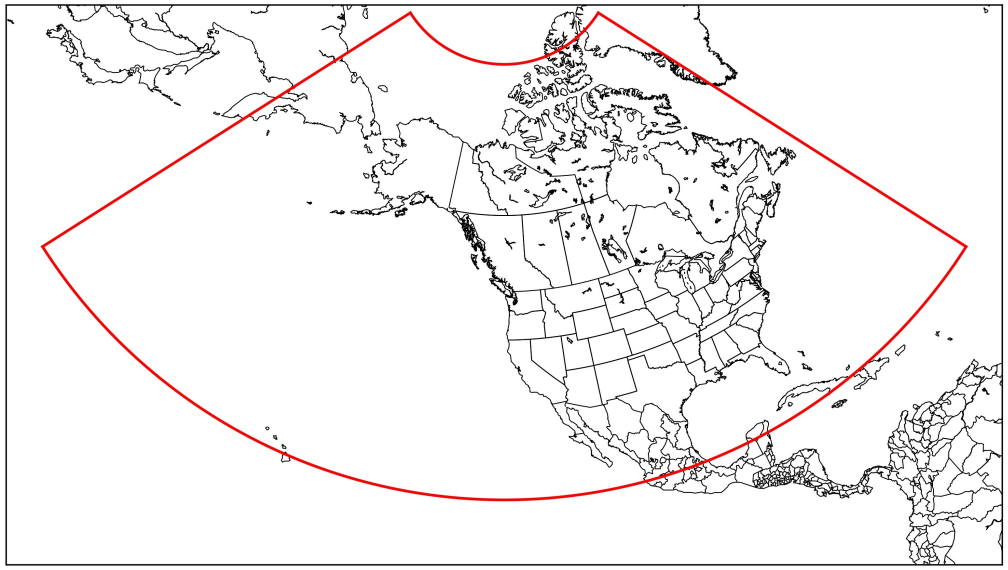
Data

- ERA-5 reanalysis – daily- 1.5° lat/lon grid. Linearly detrended.
- **Predictors** (7-day centered running mean):
  - GPH - Averaged [850, 300] hPa (*Lags 0, 7, 14*)
  - Precipitable water (PWAT) (*Lag 0*)
  - *u* & *v* winds - Averaged [850, 300] hPa (*Lag 0*)
  - SLP (*Lag 0*)
- **Predictand**: 1 (day is in an extreme event), 0 (it's not).
- **Training:** 1950-2000; **Testing:** 2001-2018



Workflow

- Domain: [20°, 80°] N; [160°, 310°] E
- 100 trees in forest
- Gini impurity criterion
- 10-fold cross-validation used for optimizing hyperparameters.



Critical Success Index (CSI)

$$CSI = \frac{A}{A + B + C}$$

		Observed	
		Yes	No
Forecast	Yes	A	B
	No	C	D

Optimized Model Accuracies

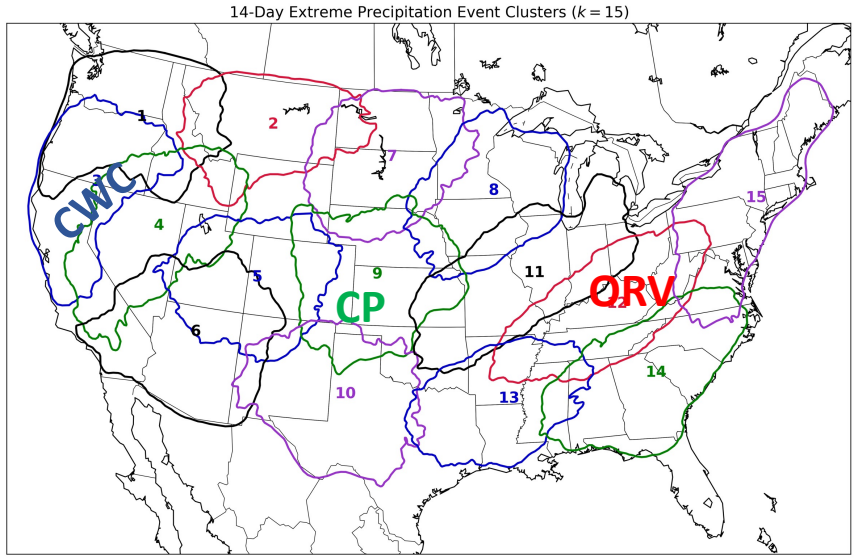
Training

Region	Hits	False Alarms	Misses	CSI
CWC	357	735	6	0.325
CP	268	201	3	0.568
ORV	320	281	3	0.530

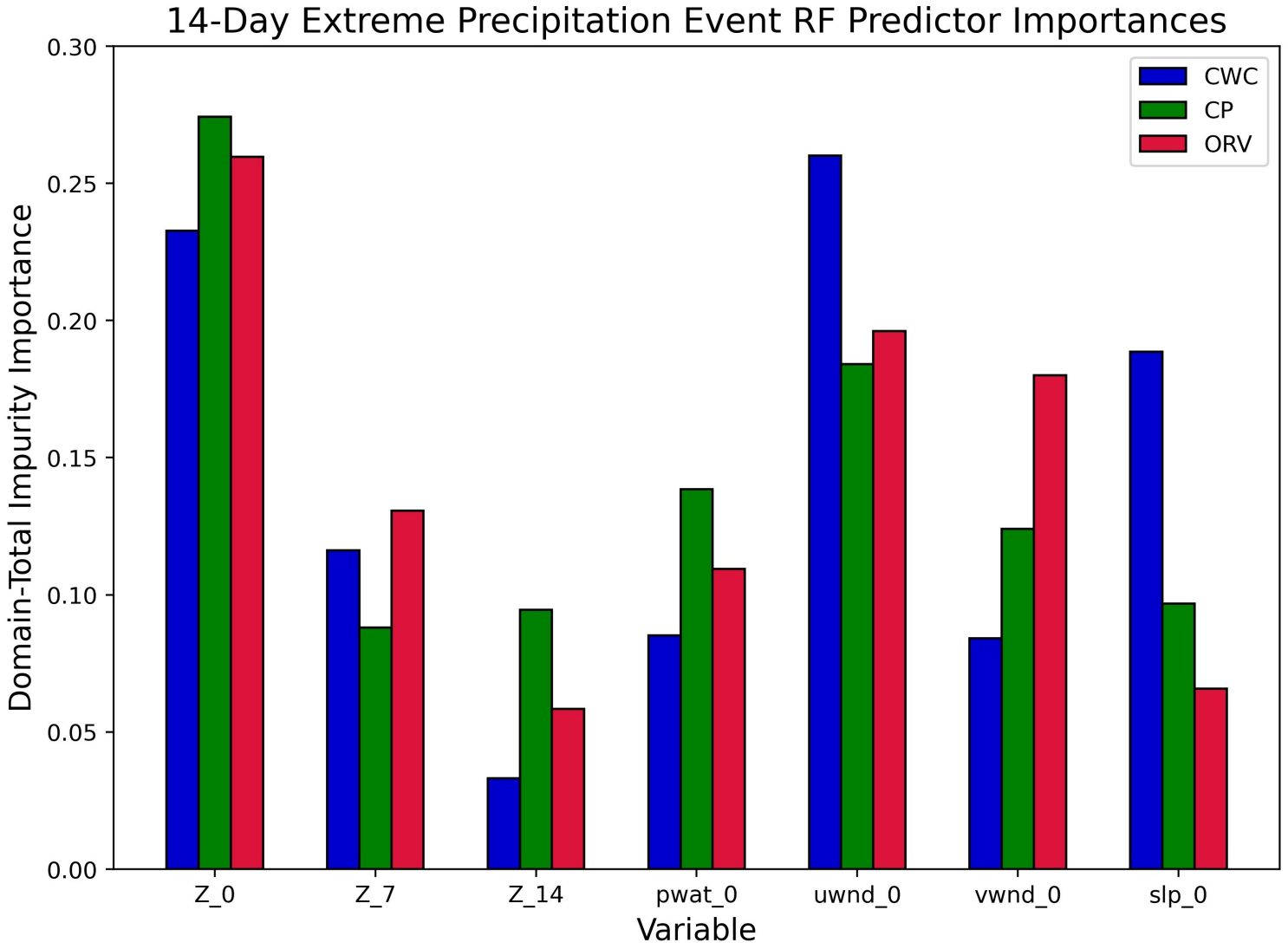
Testing

Region	Hits	False Alarms	Misses	CSI
CWC	34	311	22	0.093
CP	7	67	51	0.056
ORV	21	106	154	0.075

- **CWC**: Lots of hits but a lot of false alarms
- **CP**: Opposite of CWC
- **ORV**: Similar to CP



Which Predictors are the Most Important?



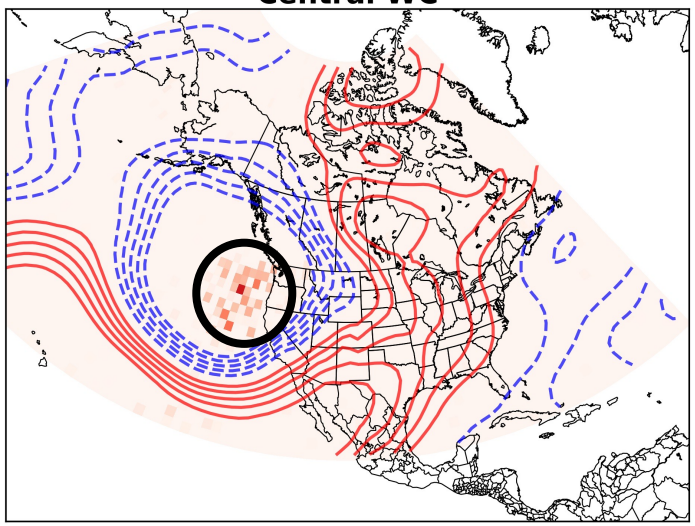
Lag 0 Geopotential Height Importance

More Important

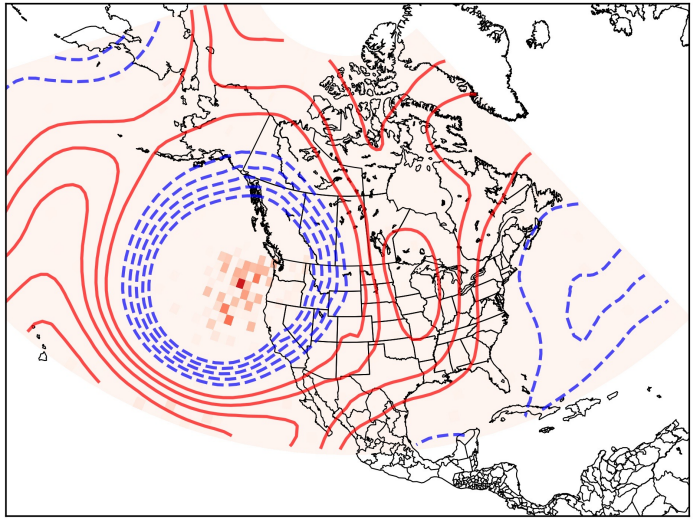
Less Important

Central WC

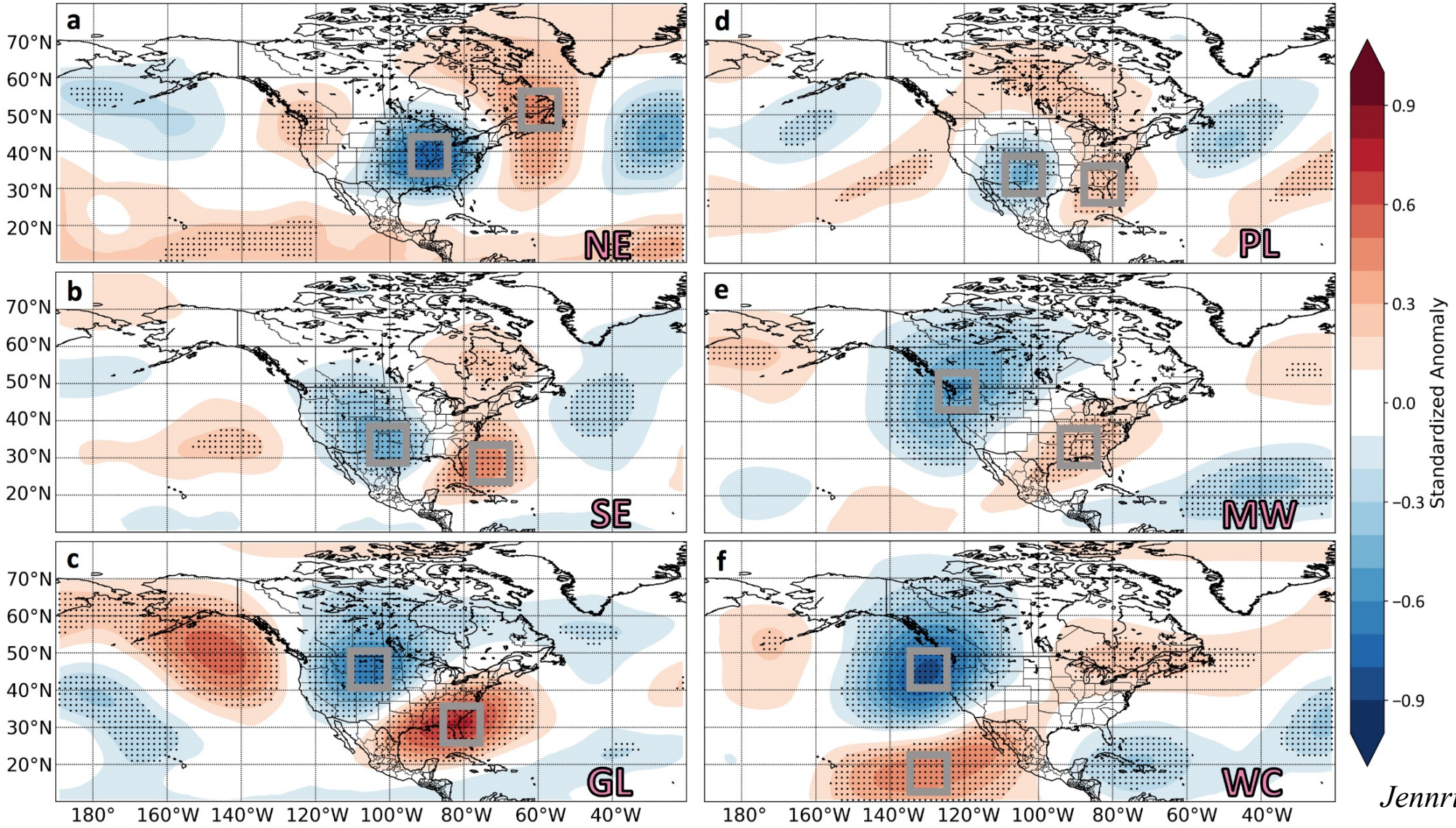
Hits



FAs



14-Day Extreme Event Total Composites



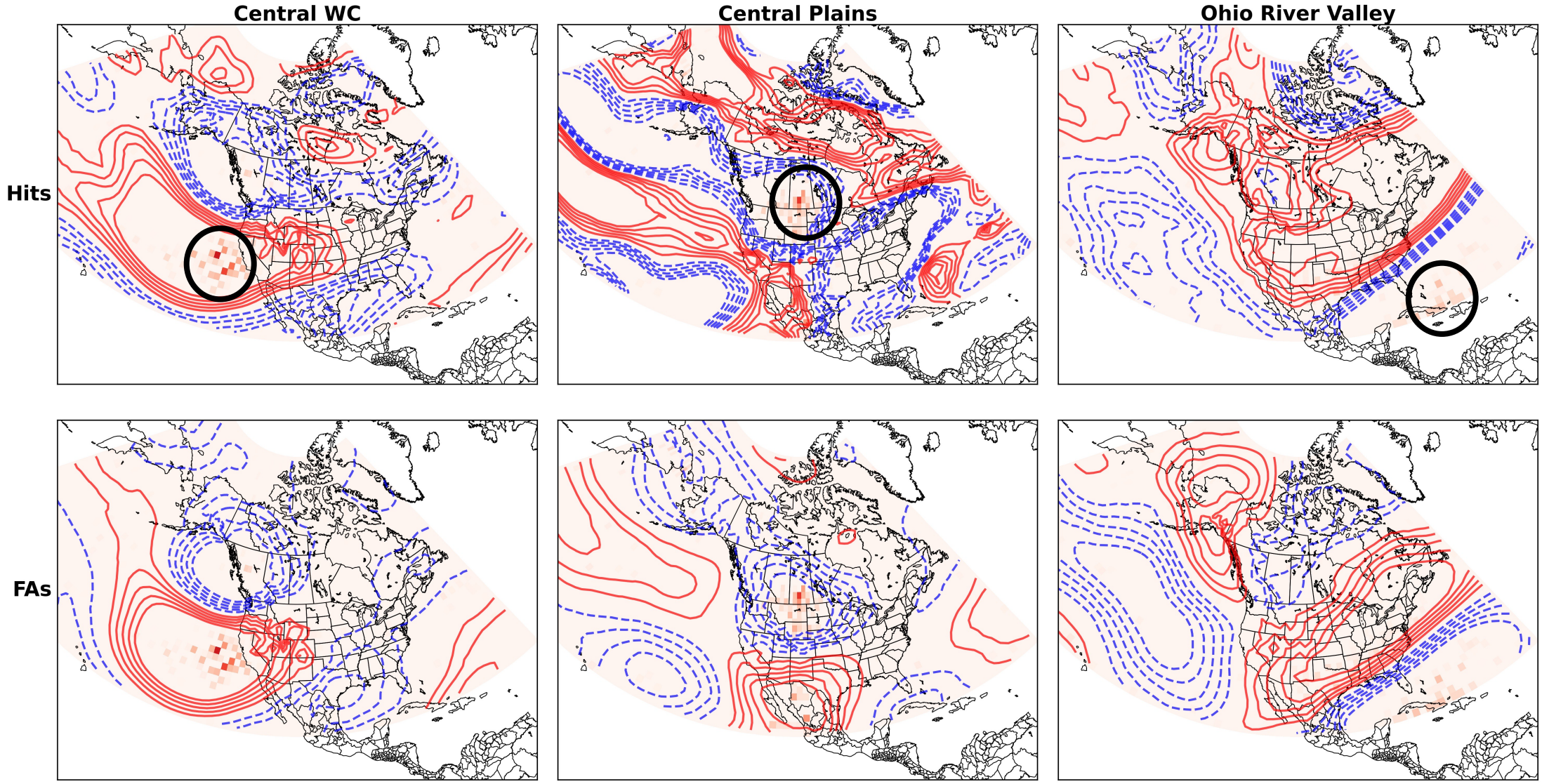
500 hPa Z500 Standardized Anomalies

Jennrich et al. [2020]

Lag 0 Zonal Wind Importance

More Important

Less Important



Take-Home Messages

- Our random forest model currently has some skill in differentiating 14-day extreme precipitation event and non-event days. Much more to interpret!

- **Most important features:** (1) Upstream trough (preceding and up to start date) / downstream ridge (develops during event); (2) Orientation (and strength) of the polar and subtropical jet streams.

- Currently testing a deep neural network to produce probabilistic forecasts for 14-day extreme precipitation events.

Research Group Website: https://ifurtado.org

THANK YOU!

Email: jfurtado@ou.edu