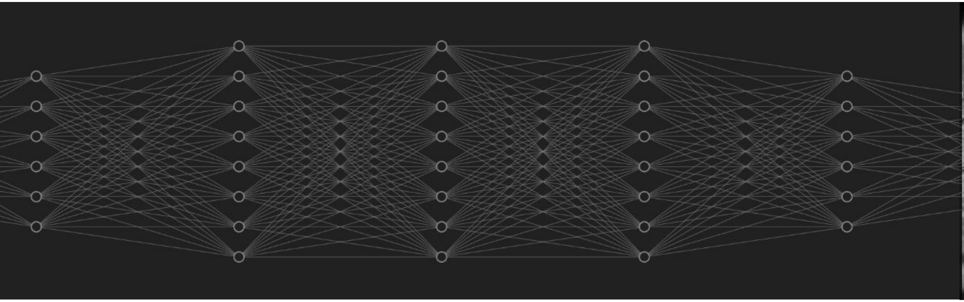
An Earth-system-oriented view of the S2S predictability of North American weather regimes using ML



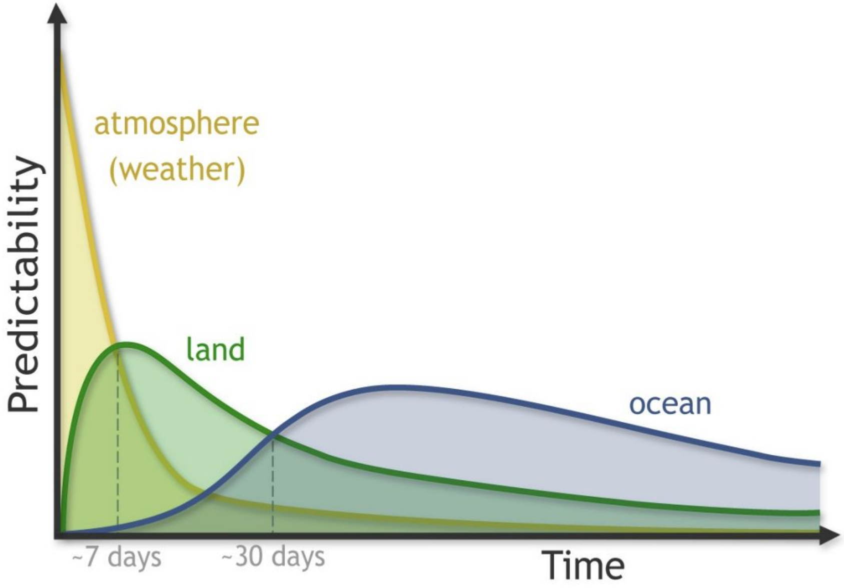

Jhayron Pérez-Carrasquilla and Dr. Maria Molina
Department of Atmospheric and Oceanic Science

Research question

Where does atmospheric predictability come from?

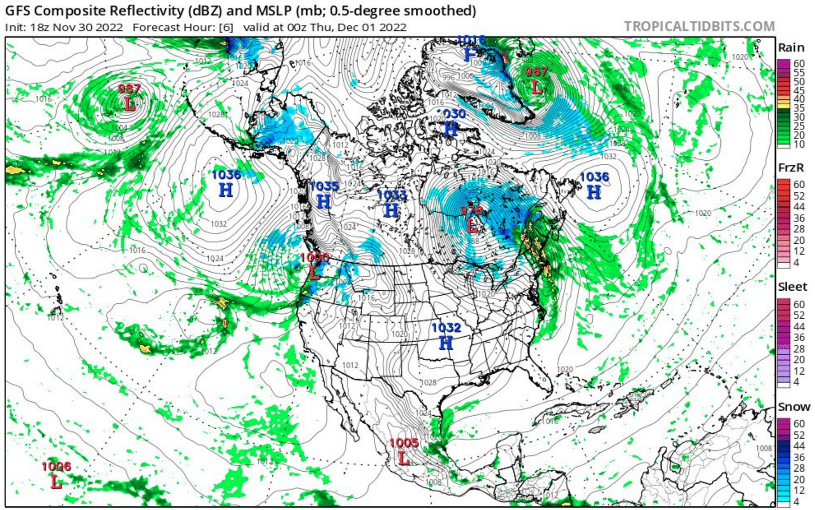
What variables or Earth system components?

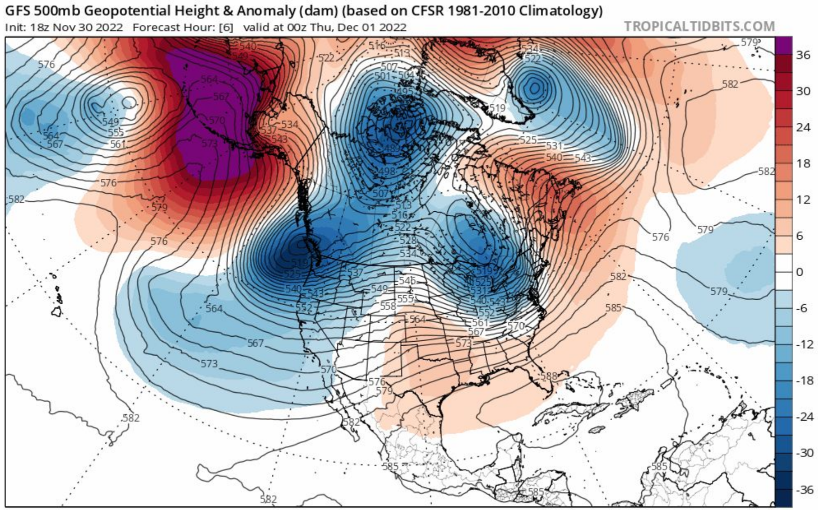
What processes?



Lang, A. L., Pegion, K., & Barnes, E. A. (2020). Introduction to special collection:"Bridging weather and climate: subseasonal-to-seasonal (S2S) prediction". Journal of Geophysical Research: Atmospheres.

What to predict? (How to "simplify" the atmosphere?)





Small-scale features

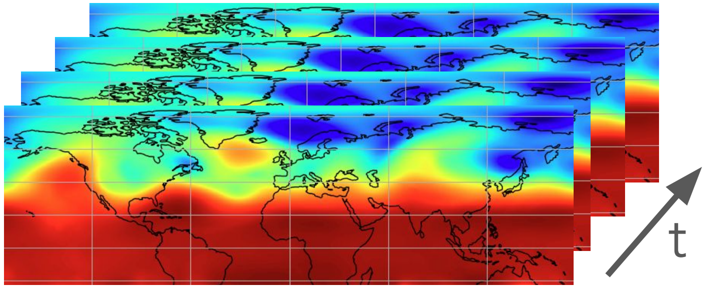
Large-scale features

Weather Regimes are easier to follow and understand and also affect the surface.

Images from: https://www.tropicaltidbits.com/

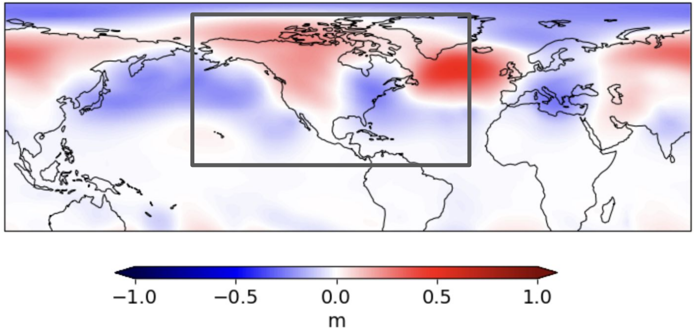
Computing year-round Weather Regimes

1) Bi-weekly 500hPa Geopotential height (Z500) 1981-2020



2) Extracted region of interest, remove annual cycle and linear trends

Z500 Anomalies

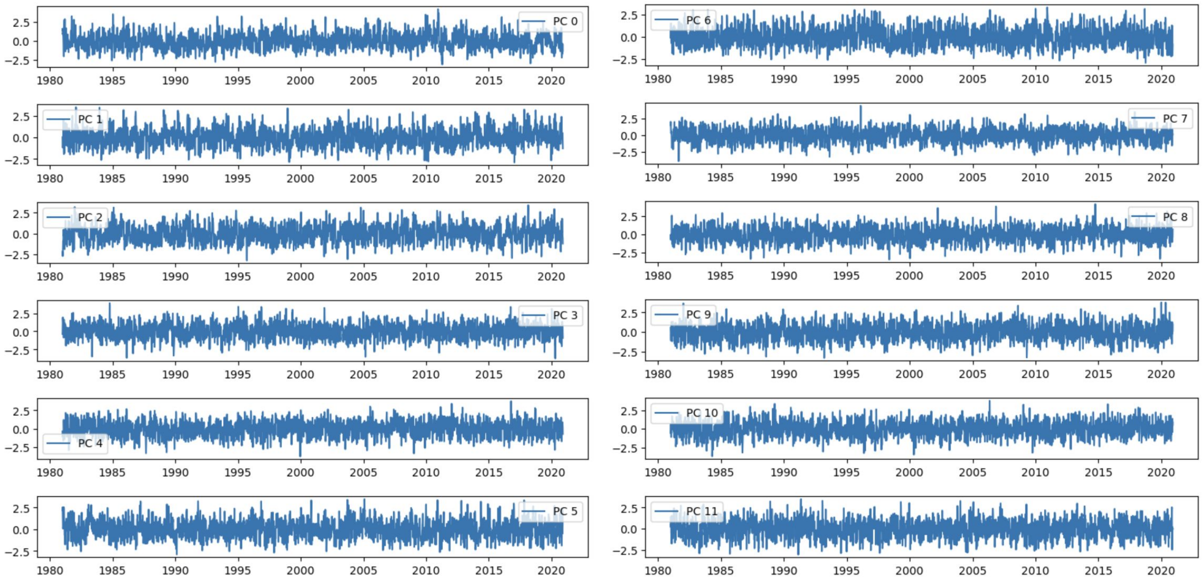


Molina, M. J., J. H. Richter, A. A. Glanville, K. Dagon, J. Berner, A. Hu, and G. A. Meehl, 2023. Subseasonal representation and predictability of North American weather regimes using cluster analysis. Artificial Intelligence for the Earth Systems, 1–54.

Lee, Simon H., Michael K. Tippett, and Lorenzo M. Polvani. "A new year-round weather regime classification for North America." Journal of Climate 36.20 (2023): 7091-7108.

Computing year-round Weather Regimes

3) Dimensionality reduction: 12 first PCs (85% of variance)



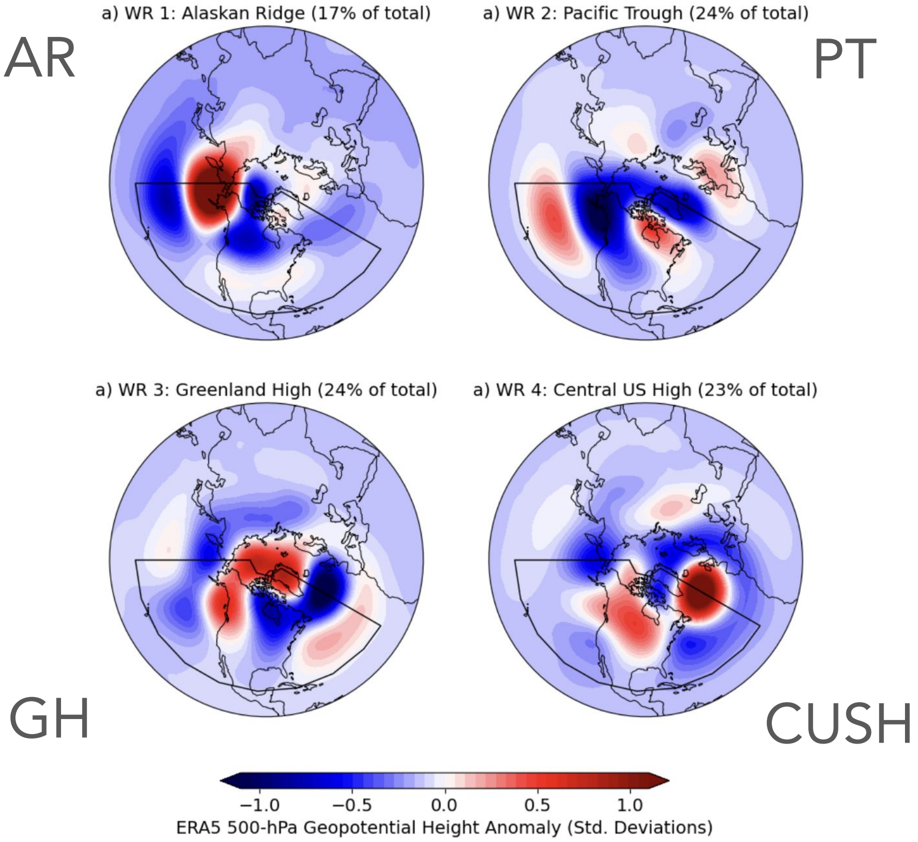
k-means clustering: extract 4 categories

Molina, M. J., J. H. Richter, A. A. Glanville, K. Dagon, J. Berner, A. Hu, and G. A. Meehl, 2023. Subseasonal representation and predictability of north american weather regimes using cluster analysis. Artificial Intelligence for the Earth Systems, 1–54.

Lee, Simon H., Michael K. Tippett, and Lorenzo M. Polvani. "A new year-round weather regime classification for North America." Journal of Climate 36.20 (2023): 7091-7108.

Computing year-round Weather Regimes

4) 5 classes: 4 Weather regimes + 1 No WR class (near climatology)



Molina, M. J., J. H. Richter, A. A. Glanville, K. Dagon, J. Berner, A. Hu, and G. A. Meehl, 2023. Subseasonal representation and predictability of north american weather regimes using cluster analysis. Artificial Intelligence for the Earth Systems, 1–54.

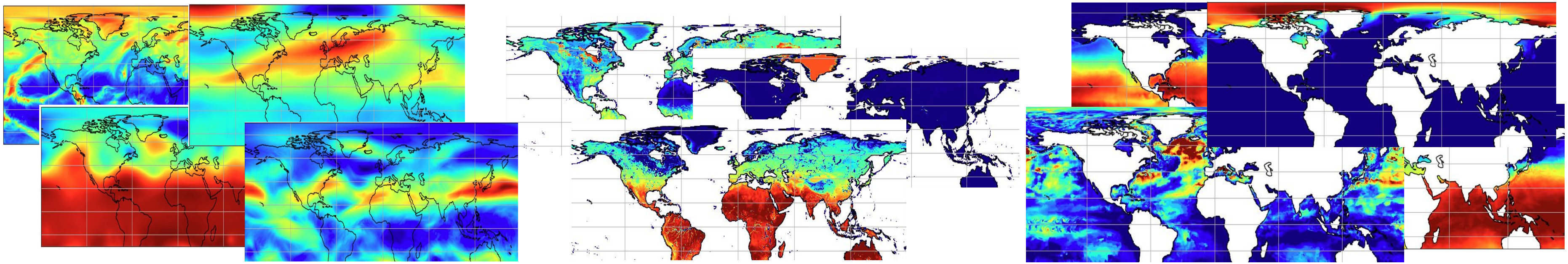
Lee, Simon H., Michael K. Tippett, and Lorenzo M. Polvani. "A new year-round weather regime classification for North America." Journal of Climate 36.20 (2023): 7091-7108.

Predictors representing the initial state of the Earth system

Atmosphere (ERA5) - 4 variables: Z500hPa, U10hPa and U200hPa, and OLR.

Land (ERA5) - 9 variables: Soil integrated moisture and heat for different depths, and snow depth.

Ocean (SODA) - 10 variables: OHC for different depths, SSH, SST, MLD, ice properties.



Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

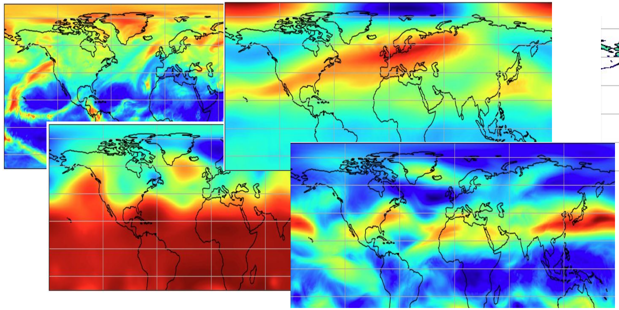
Predictors representing the initial state of the Earth system

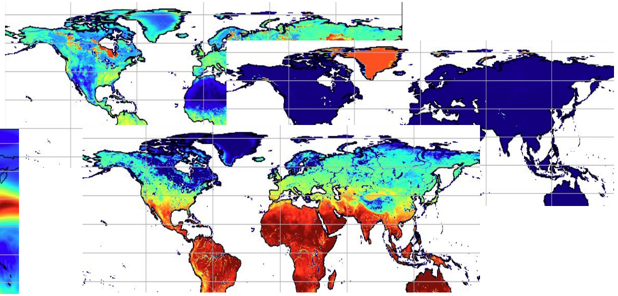
Removed climatology

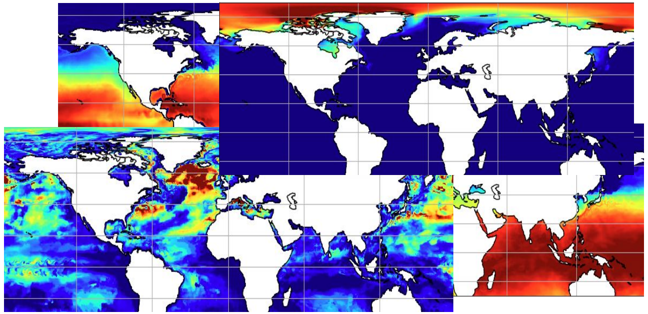
Removed trends

Dimensionality reduction with PCA: For each variable, we extracted the first 20 PCs

Trained different models with each variable individually



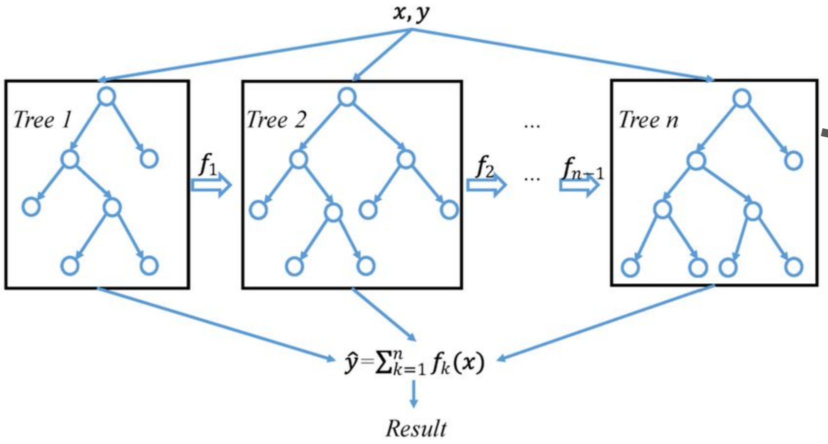




Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

ML framework for each variable

XGBoost

20 PCs from one variable →  → 1 of 5 classes: weather regimes

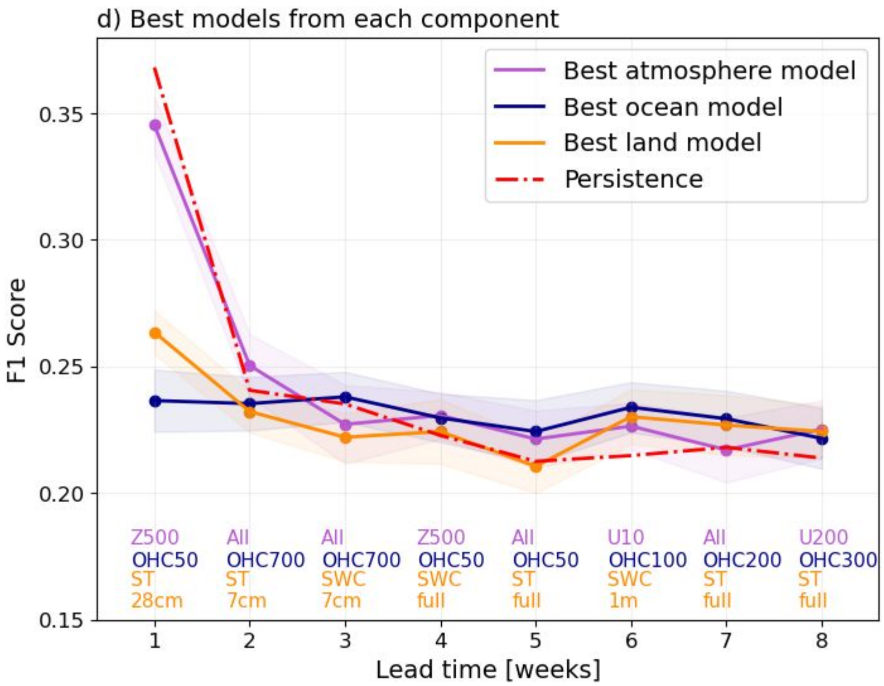
Not looking for the best performance but for a fair comparison among models.

Friedman, J. H., 2001: Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189–1232.

Molina, M. J., and Coauthors, 2023b: A review of recent and emerging machine learning applications for climate variability and weather phenomena. Artificial Intelligence for the Earth Systems, 1–46.609

Fatima, S., A. Hussain, S. B. Amir, S. H. Ahmed, and S. M. H. Aslam, 2023: Xgboost and random forest algorithms: An in depth analysis. Pakistan Journal of Scientific Research, 3 (1), 26–31.

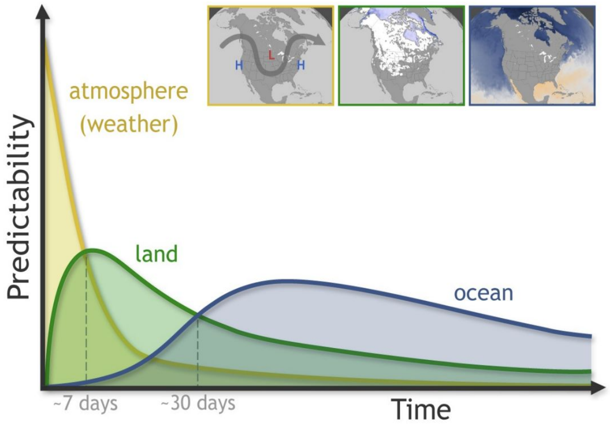
Models' comparison: Which component is more important?



Ocean OHC

Land Integrated soil moisture and soil temperature

Atmosphere Combined model

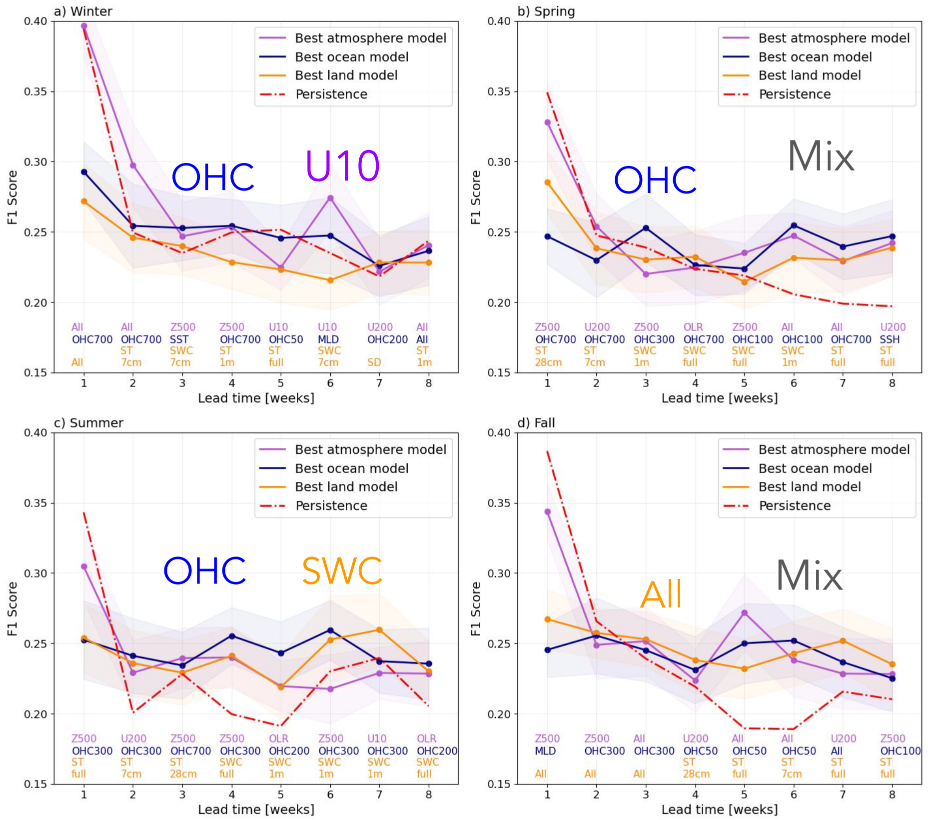


Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Seasonal differences

Winter

Spring



Summer

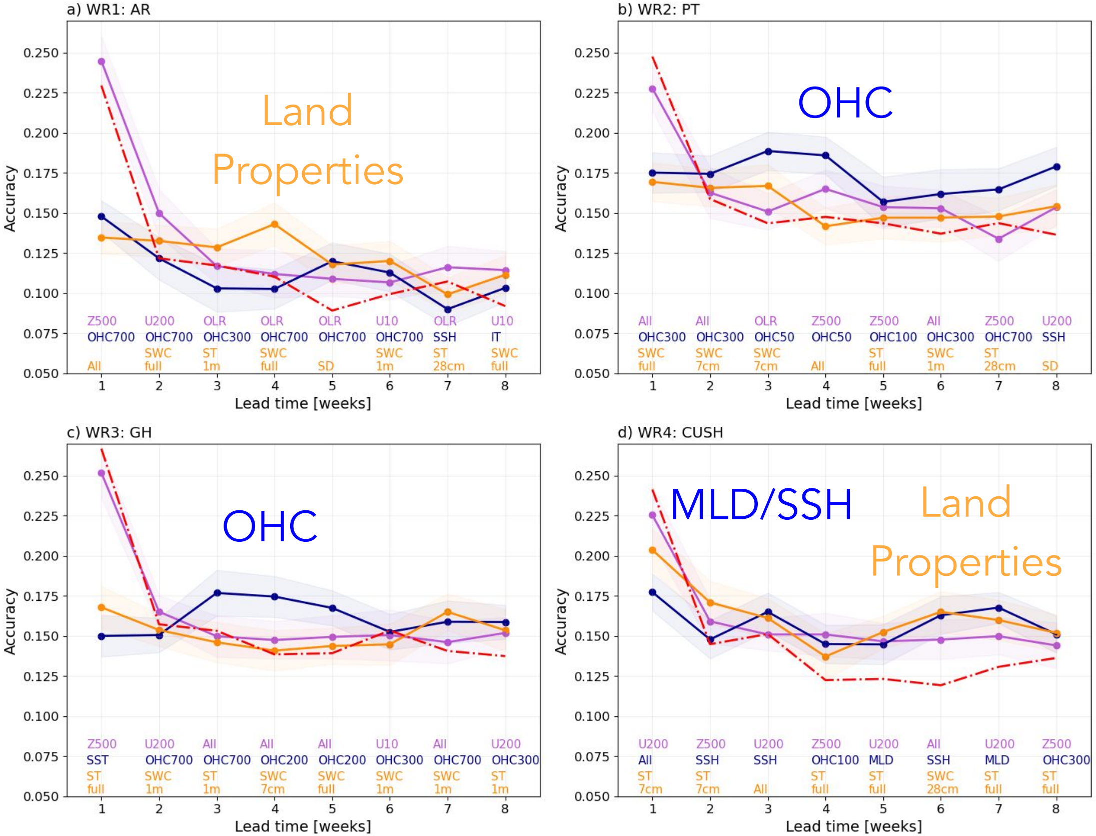
Fall

Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Differences depending on the WR of interest

Alaskan Ridge

Pacific Trough

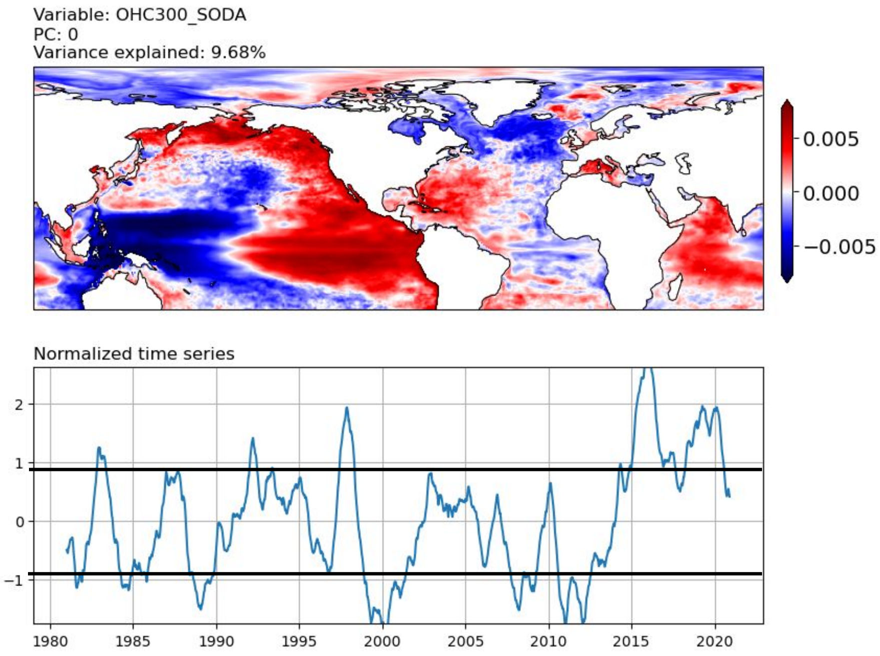


Greenland High

Central-US High

Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

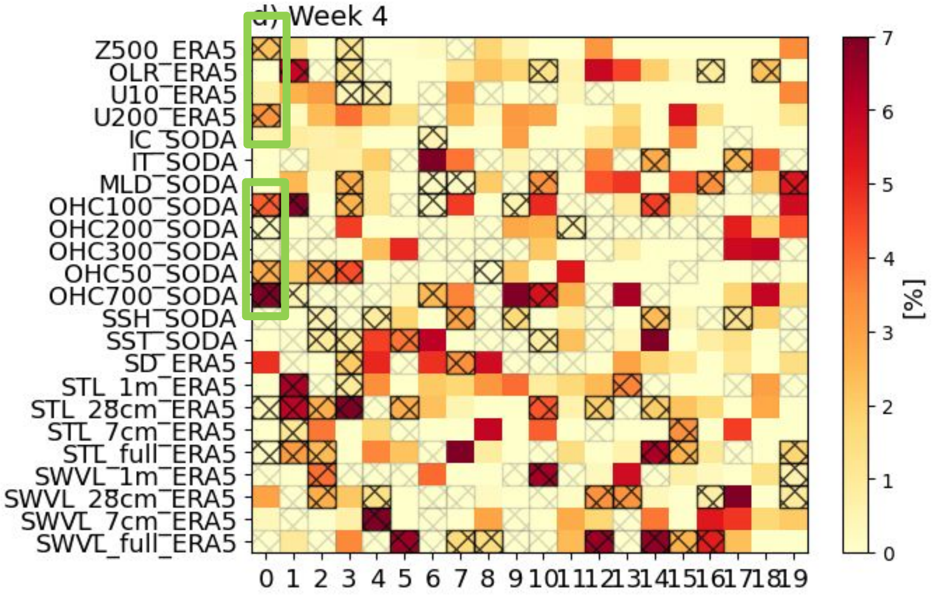
Analysis of the Principal Components



What spatial patterns are relevant from each variable?

Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Which Principal Components are important?



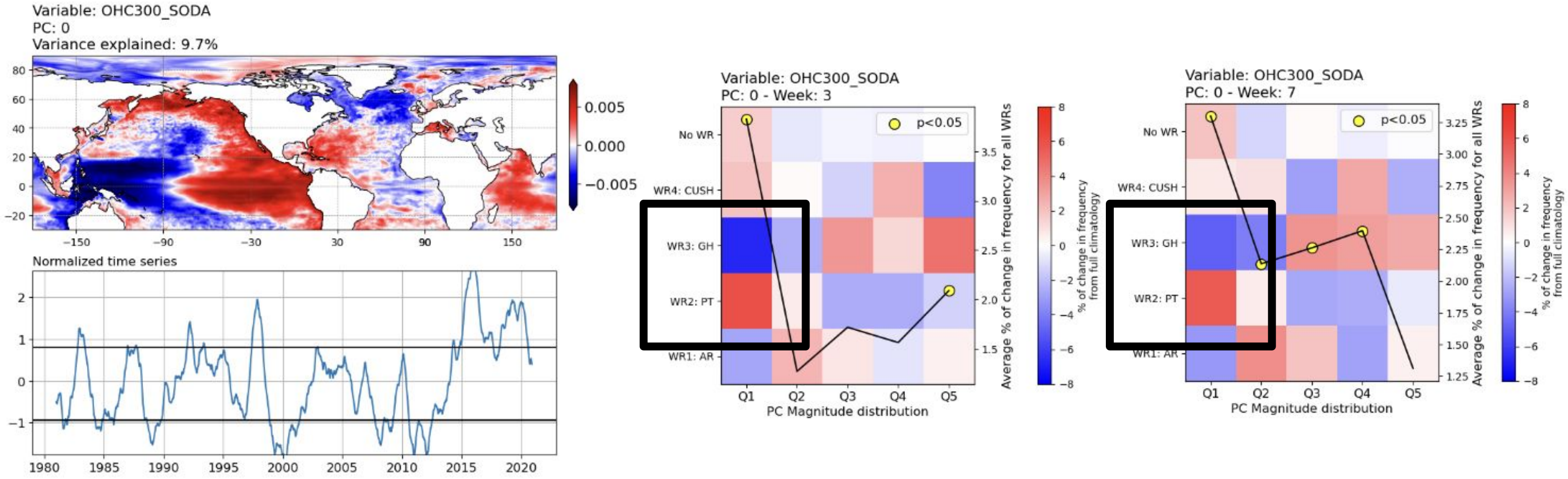
2 Conditions (black hatching):

- The PC time series being anomalously high or low produced a statistically significant change in the distribution of future WRs.

- The magnitude of the PC time series is higher than average when correct predictions are made.

Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

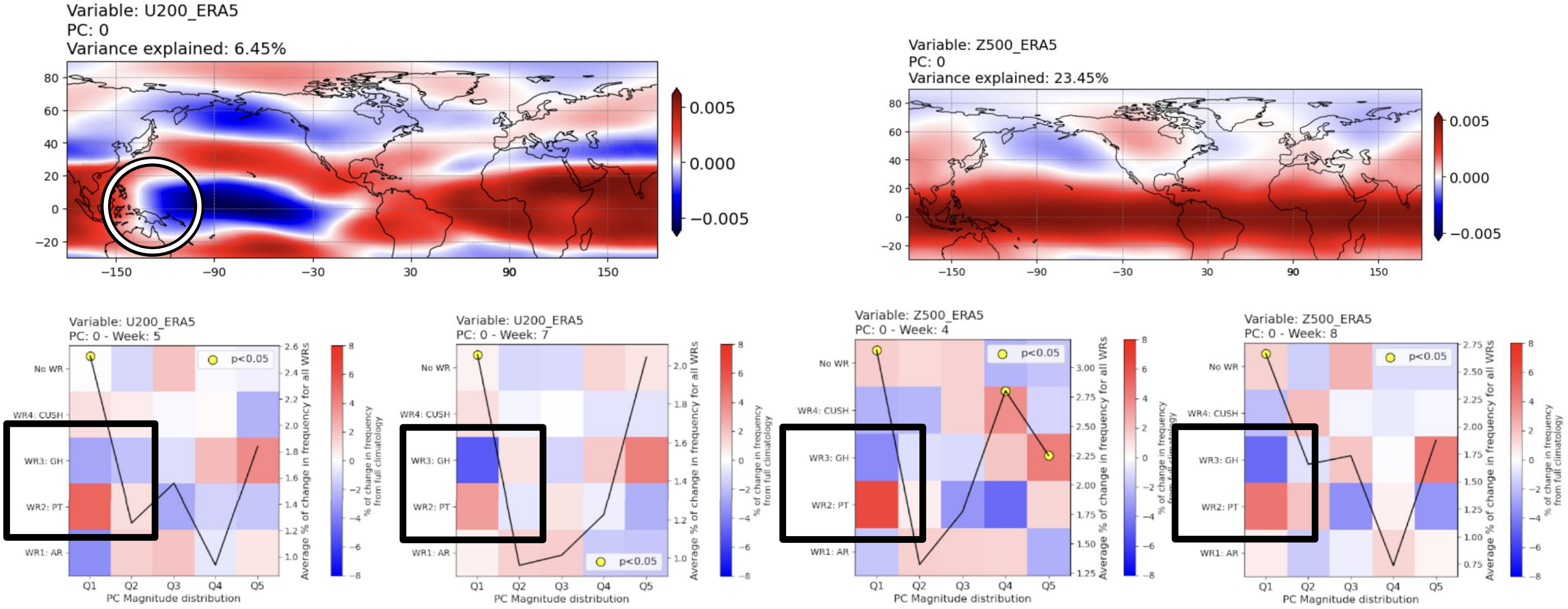
ENSO pattern



Same signal seen in OHC at different depths, sea-surface height, sea surface temperature.

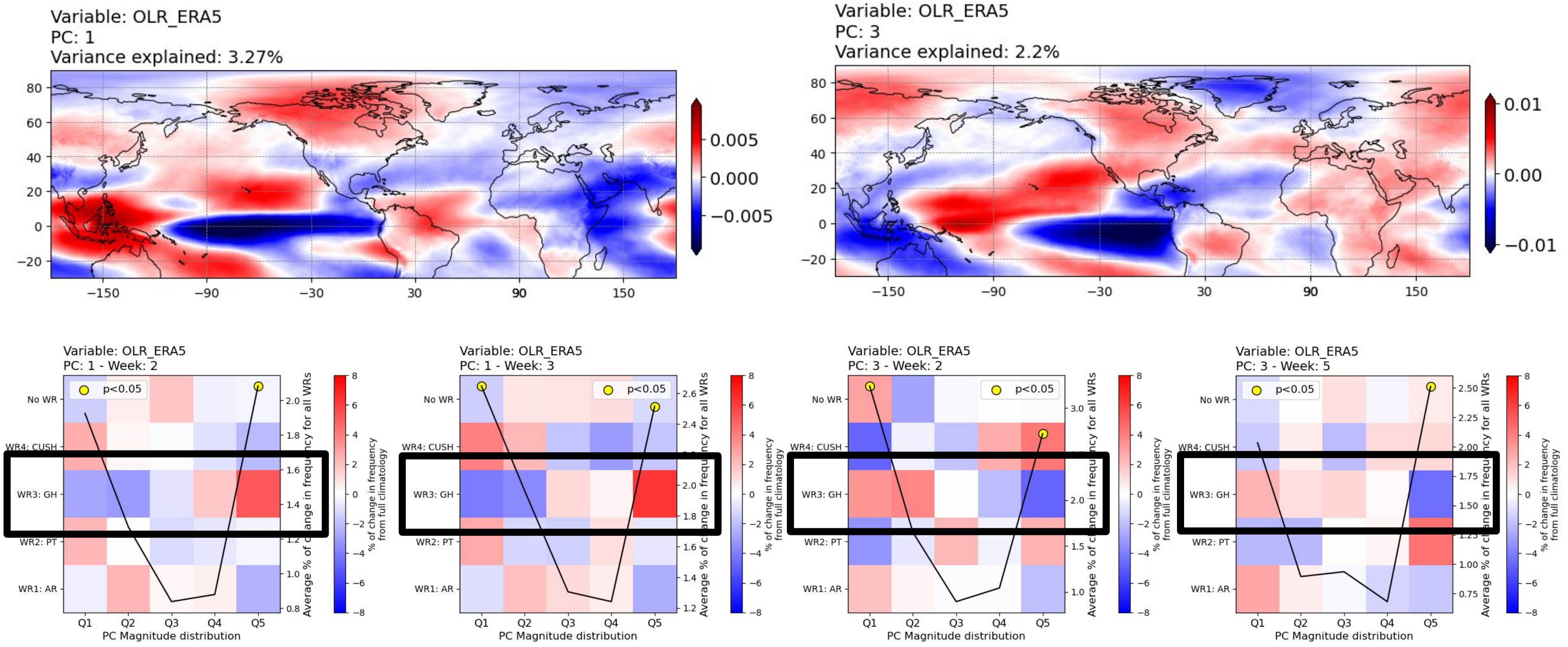
Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Atmospheric Associated Pattern



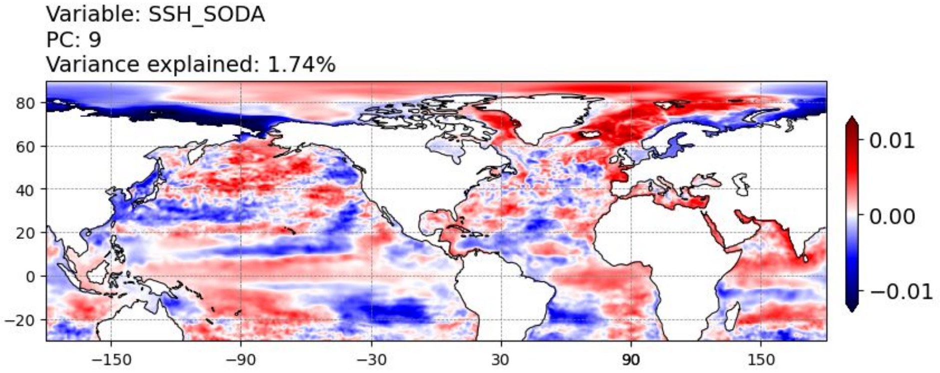
Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

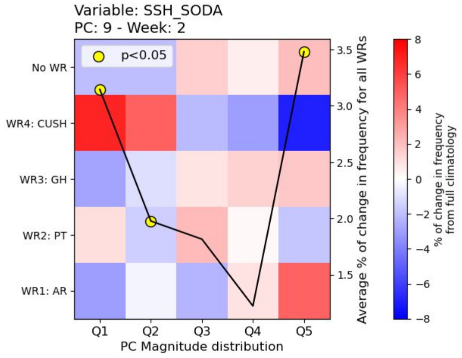
The effect changes depending on where convection is located

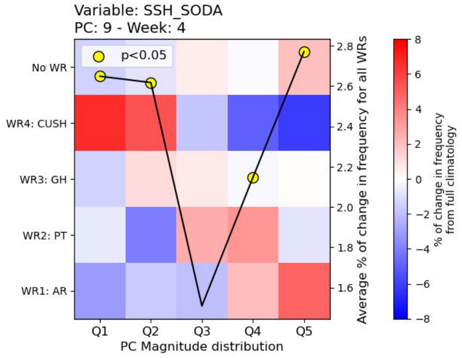


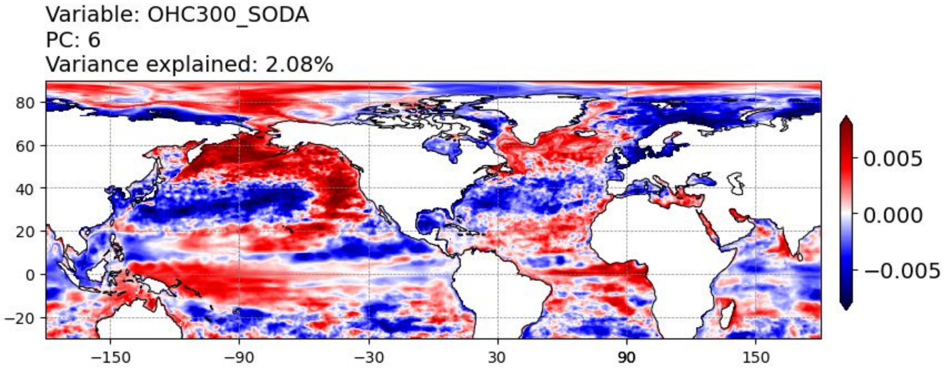
Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

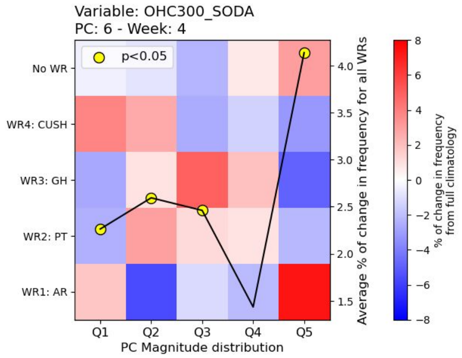
More from the ocean

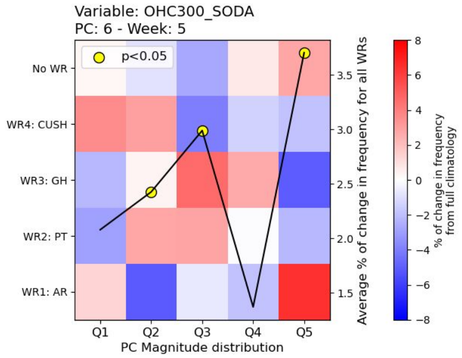






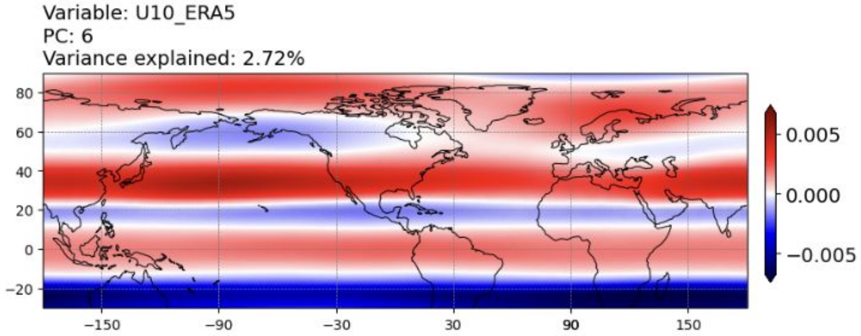


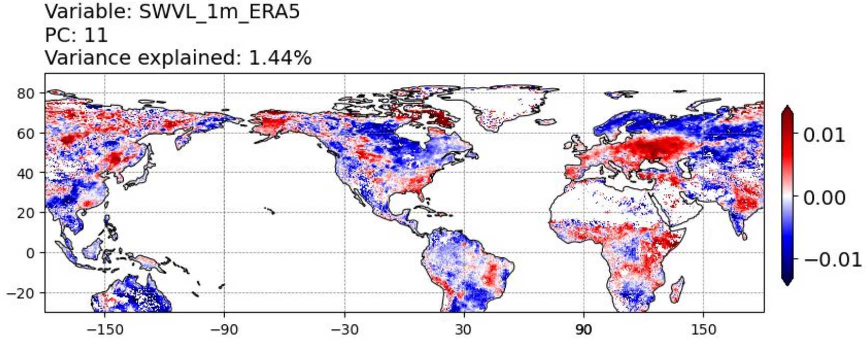


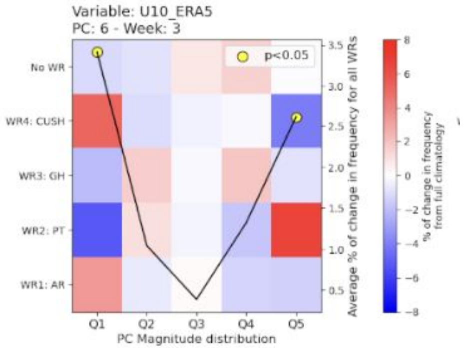


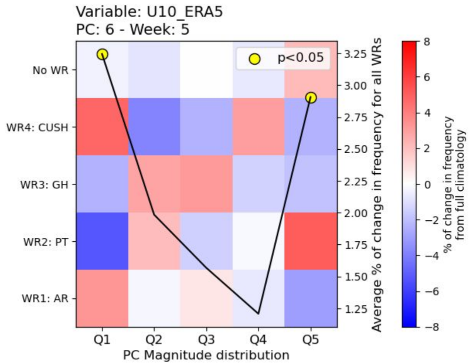
Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

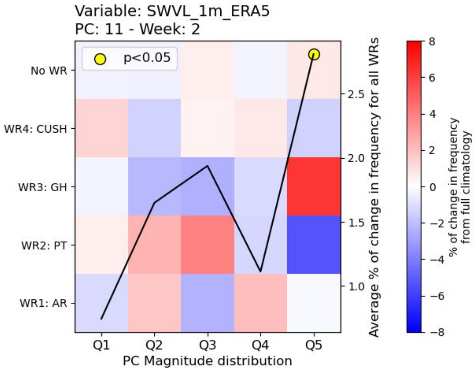
More from the atmosphere and land

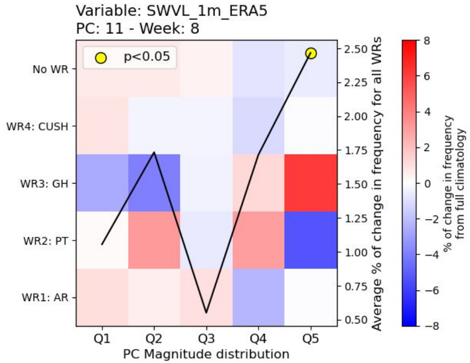












Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost

Jhayron S. Pérez-Carrasquilla[a] and Maria J. Molina[a,b]

[a] *Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA*

[b] *National Center for Atmospheric Research, Boulder, Colorado, USA*

UMD | climate & extremes data science

NCAR CISL

Advanced Research Computing

Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Summary

jhayron@umd.edu

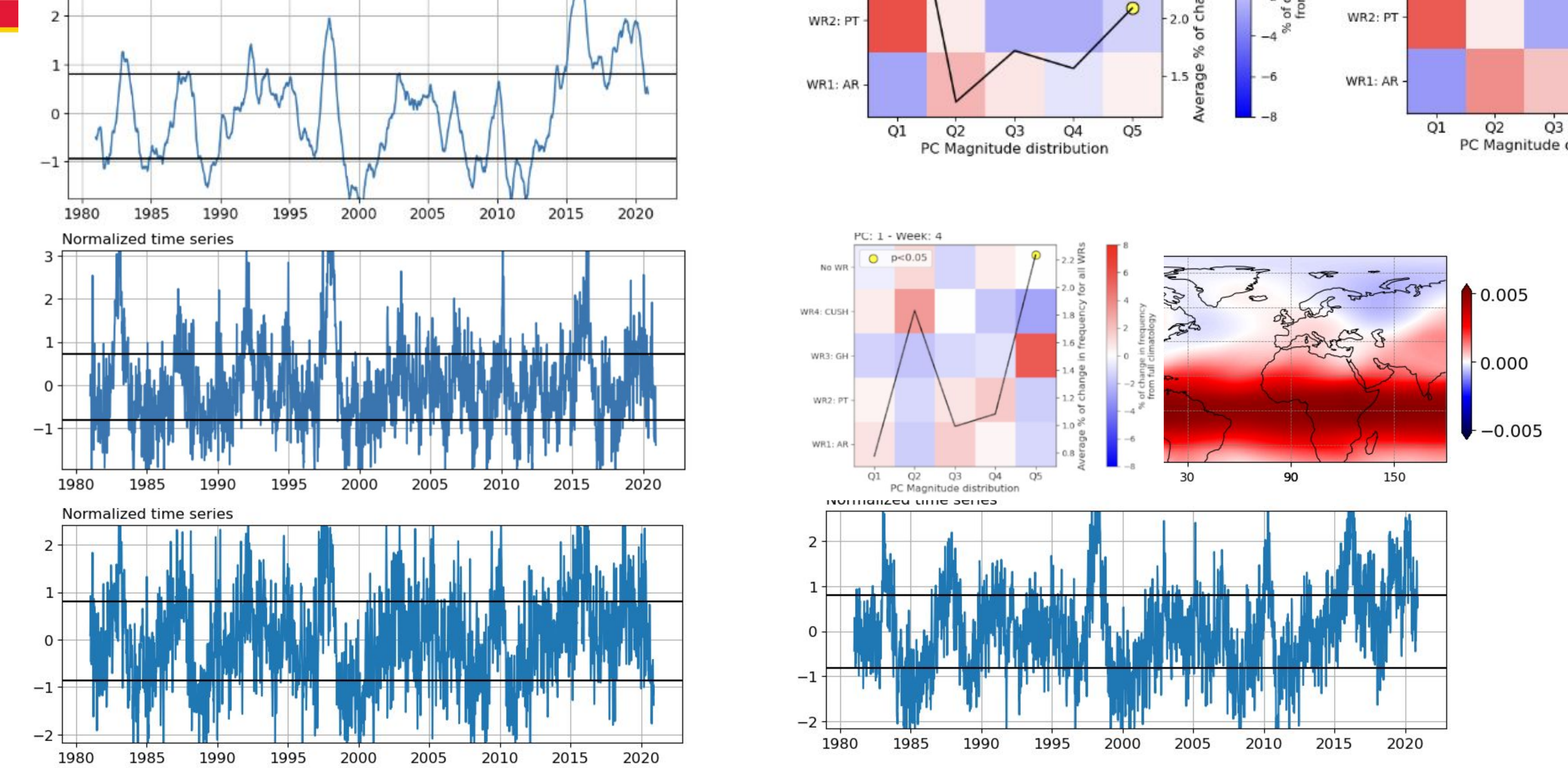
ML useful not only for prediction but for producing complementary scientific knowledge.

The ocean's role (OHC) appears to be the predominant source of information even from early stages.

ENSO signal, specifically La Niña seems to be the greatest forecast of opportunity for Greenland High and Pacific Trough.

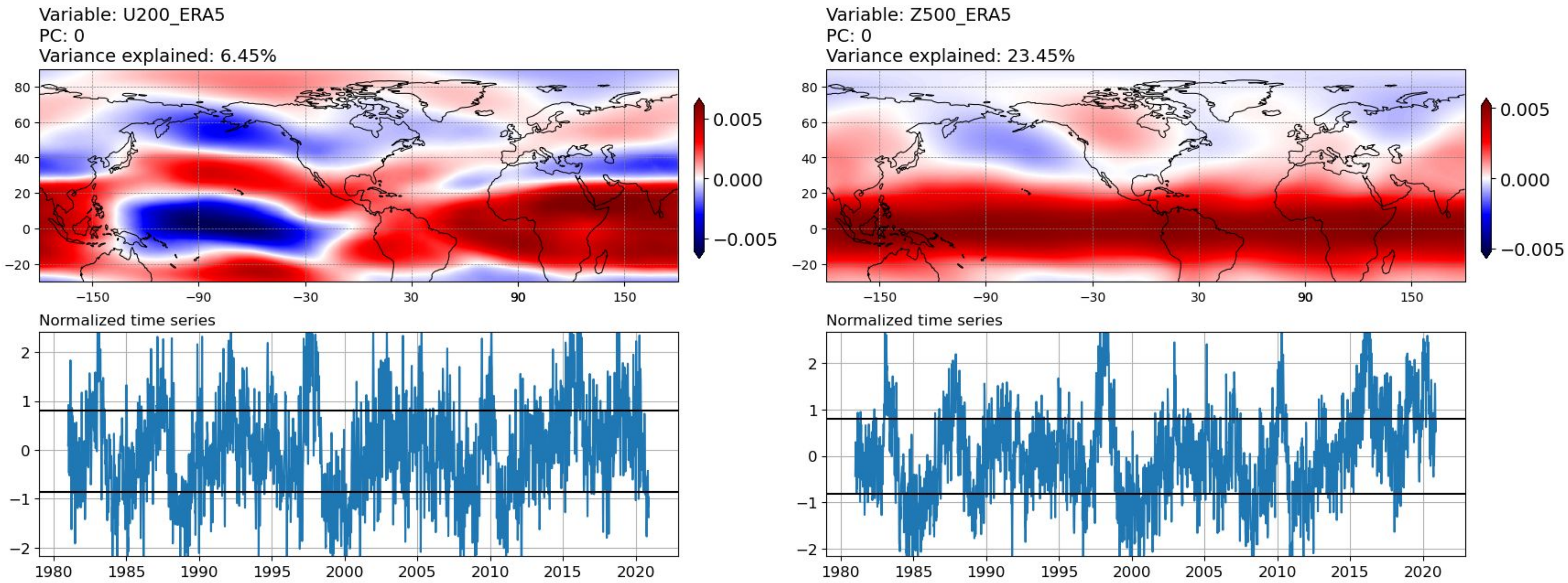
Specific changes in stratospheric winds and soil moisture are associated with changes in the likelihood of occurrence of WRs.

Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).



Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Atmospheric Associated Pattern

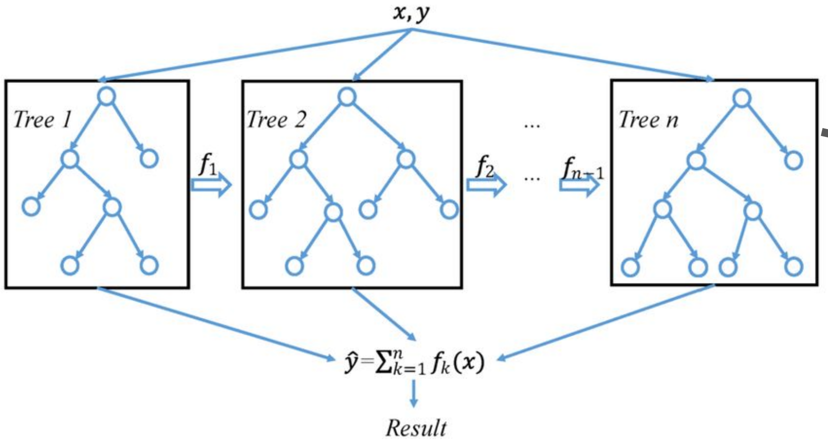


Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

ML framework for each variable

XGBoost

20 PCs from one variable



1 of 5 classes: weather regimes

Advantages: Widely used, similar performance to DL, low computational cost, fewer hyperparameters and lower sensitivity to them

Friedman, J. H., 2001: Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189–1232.

Molina, M. J., and Coauthors, 2023b: A review of recent and emerging machine learning applications for climate variability and weather phenomena. Artificial Intelligence for the Earth Systems, 1–46.609

Fatima, S., A. Hussain, S. B. Amir, S. H. Ahmed, and S. M. H. Aslam, 2023: Xgboost and random forest algorithms: An in depth analysis. Pakistan Journal of Scientific Research, 3 (1), 26–31.

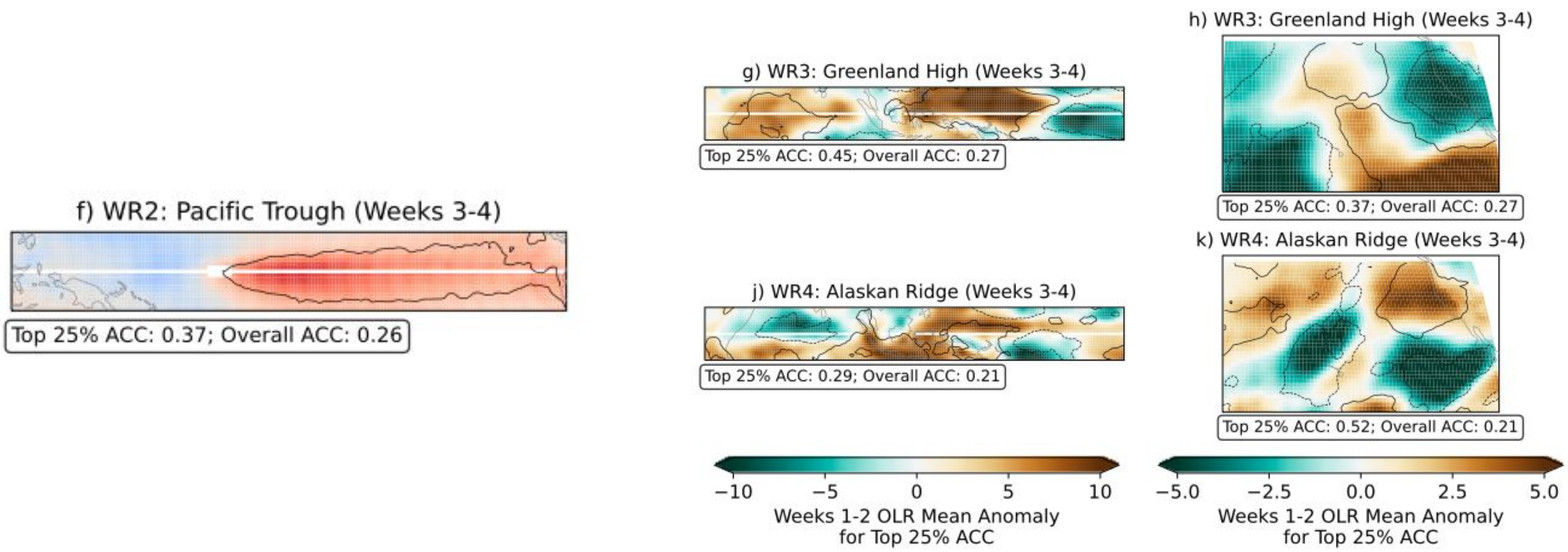
Previous work on WRs predictability

"Large-scale OLR and SST patterns upstream of North America across the subtropics and tropics can shed light on precursor mechanisms associated with weather regimes that can result in anomalous temperature, precipitation, and extremes"

"Skillful representation of upstream precursor patterns during weeks 1–2 within CESM2 can contribute to skillful prediction of weeks 3–4 weather regimes over North America "

Molina, M. J., J. H. Richter, A. A. Glanville, K. Dagon, J. Berner, A. Hu, and G. A. Meehl, 2023. Subseasonal representation and predictability of north american weather regimes using cluster analysis. Artificial Intelligence for the Earth Systems, 1−54.

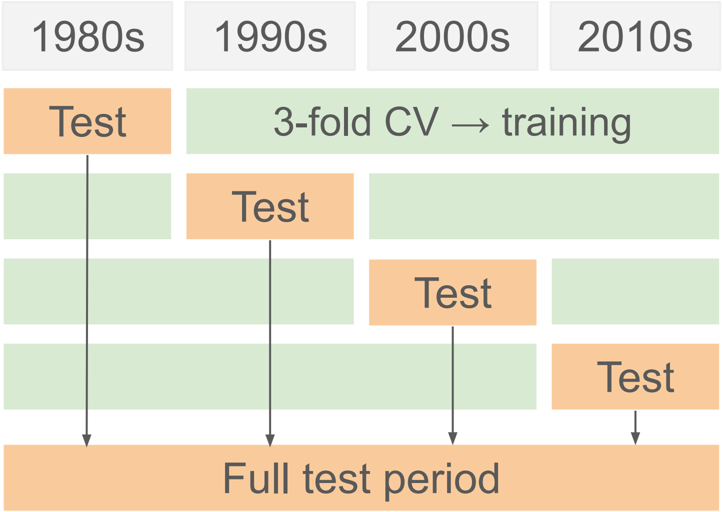
Previous work on WRs predictability



Molina, M. J., J. H. Richter, A. A. Glanville, K. Dagon, J. Berner, A. Hu, and G. A. Meehl, 2023. Subseasonal representation and predictability of north american weather regimes using cluster analysis. Artificial Intelligence for the Earth Systems, 1–54.

Bayesian hyperparameter optimization for xgboost and cross-testing/cross-validation

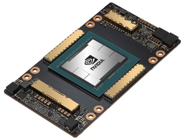
Aiming to find **robust** and **fair** models.




```
#### Tree specific hyperparameters ####
'max_depth': (3,12), #maximum depth of tree
'min_child_weight': (1,50), #minimum sum of instance weight needed in a child,
#prevents the creation of too small leaves
'subsample': (0.1, 1), ## percentage of samples used for each tree construction
'colsample_bytree': (0.1, 1), ## percentage of features used for each tree construction.
'colsample_bylevel': (0.1, 1),## percentage of features used for each split/level.
#### Learning task-specific hyperparameters ####
'learning_rate': (0.01, 0.3), #step size shrinkage usage in updates
'gamma':(0, 3), #minimum loss redution required to make a further partition on a leaf node of tree
'reg_lambda':(0, 10), #L2 regularization term on weights
'reg_alpha':(0, 10),#L1 regularization term on weights
#### General ####
'class_weights':(0, 1), #use class weights True or False
```

176 models x 4 test-folds x 3 cv-folds
x ~100 trials =

211,200 trinings

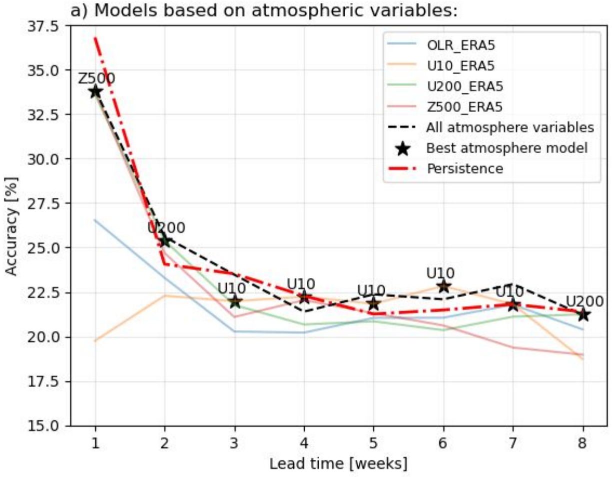


DEPARTMENT OF ATMOSPHERIC & OCEANIC SCIENCE

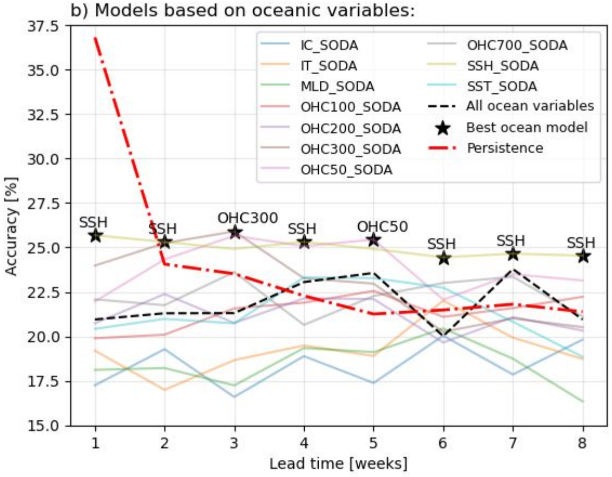
UMD | climate & extremes data science

NCAR CISL

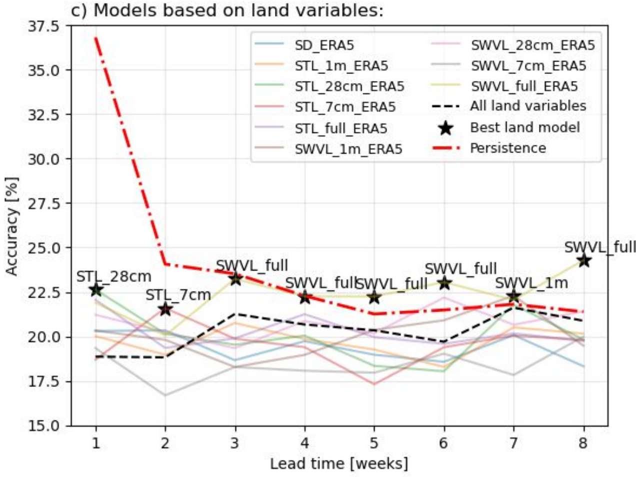
Overall results vs. persistence



Not much improvement over persistence, **stratospheric winds** performing the best.

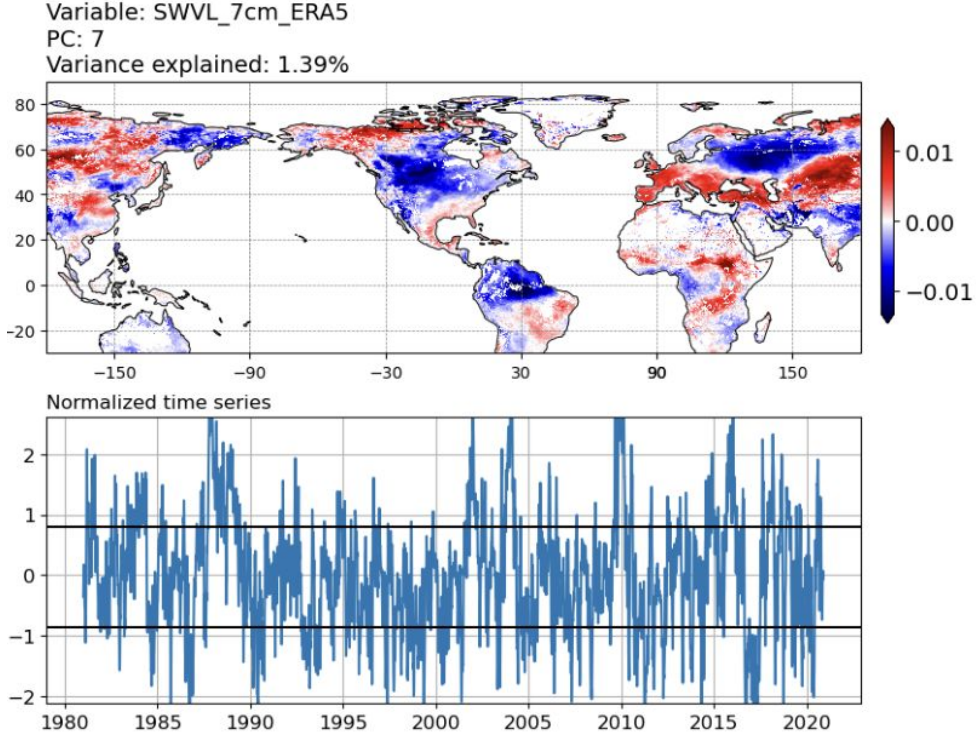


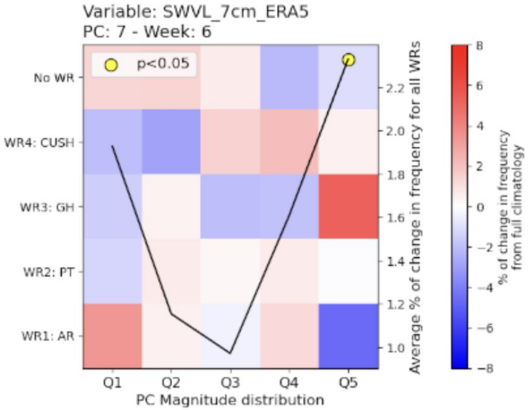
Better than persistence after week 2, **SSH and OHC** performing the best.

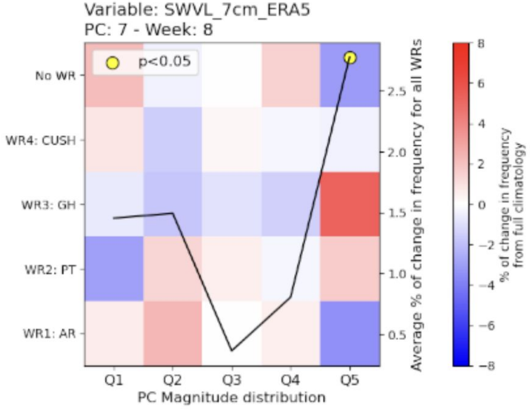


Small but increasing improvement over persistence after week 4, **soil moisture integrated down to 2.89m** performing the best.

Other precursors

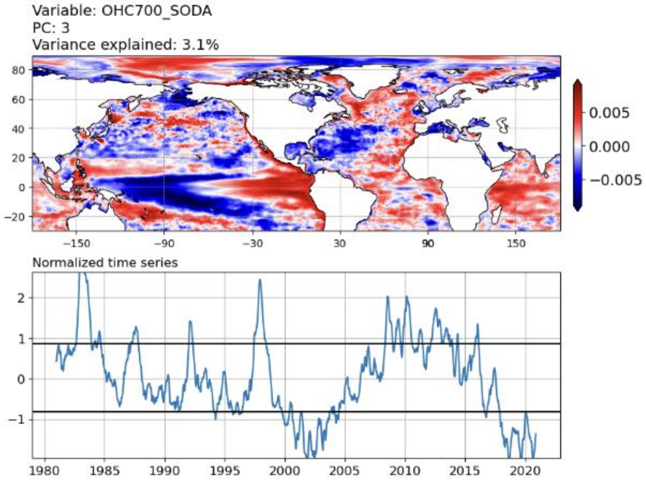


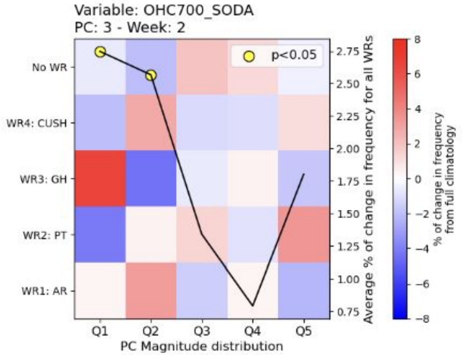


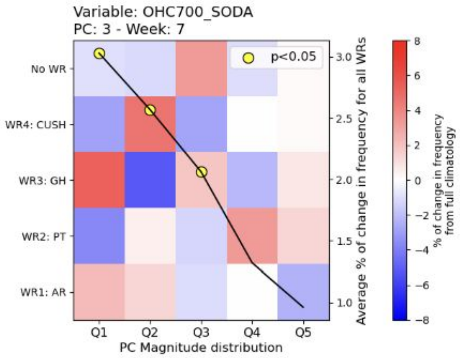


Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Oceanic predictors: Tropics





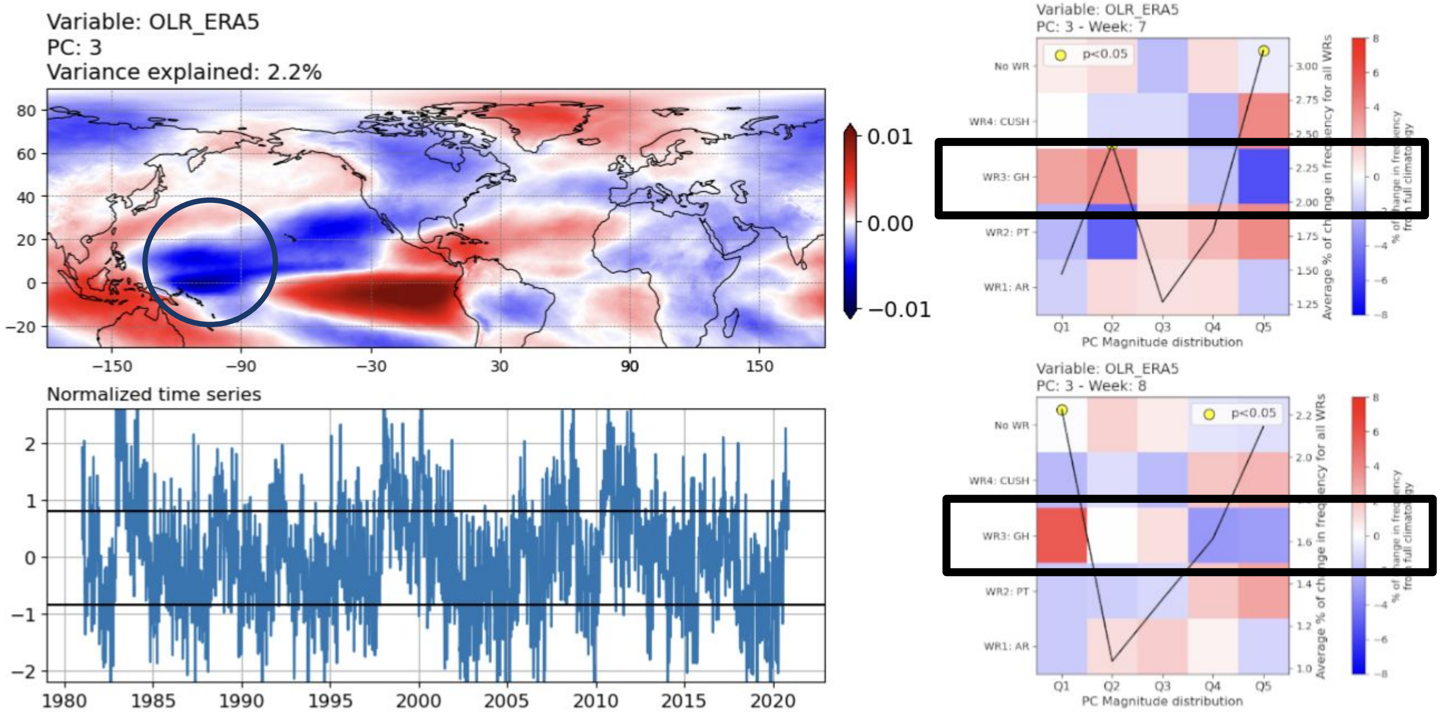


Same signal seen in OHC at different depths and sea-surface height.

Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

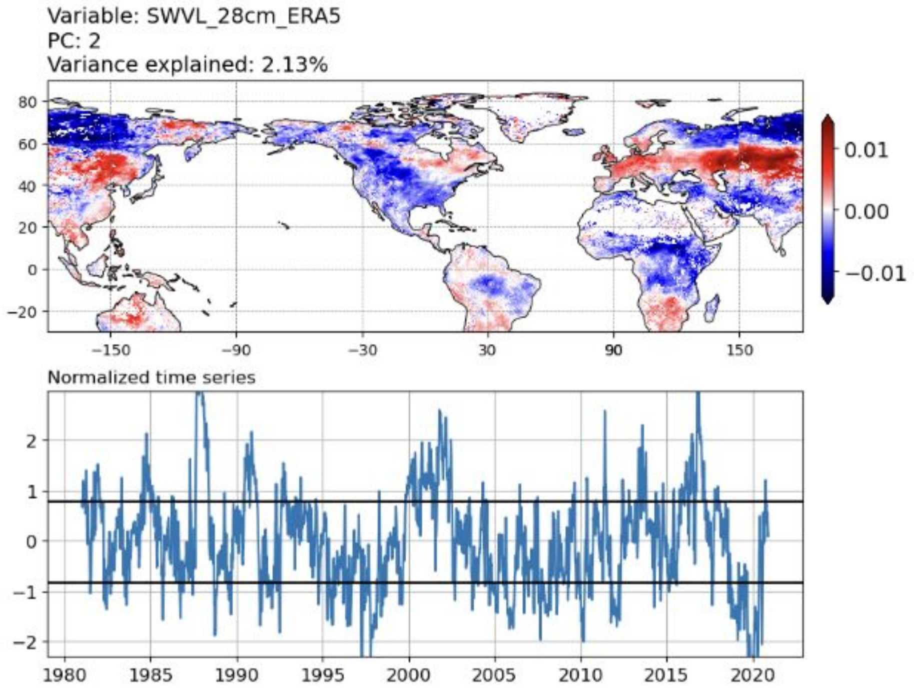
The effect changes depending on where convection is located

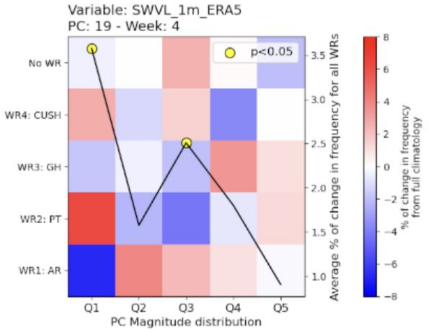
Predictability may "last" more if convection is more spread over the Pacific.

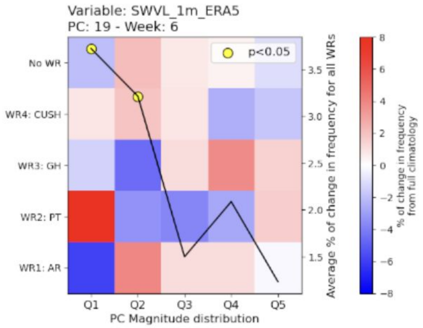


Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Other precursors

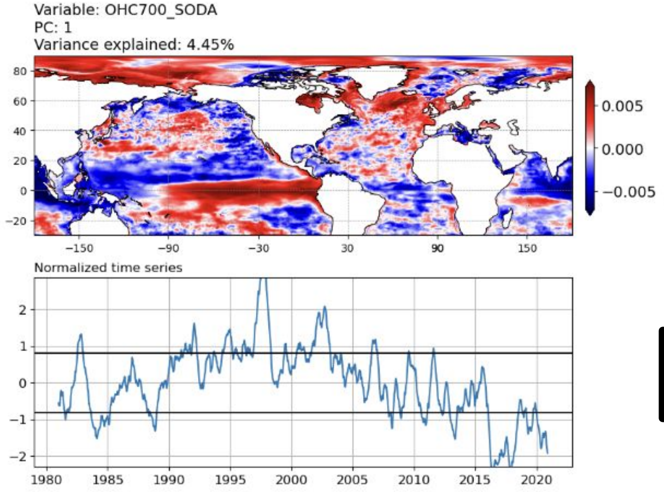


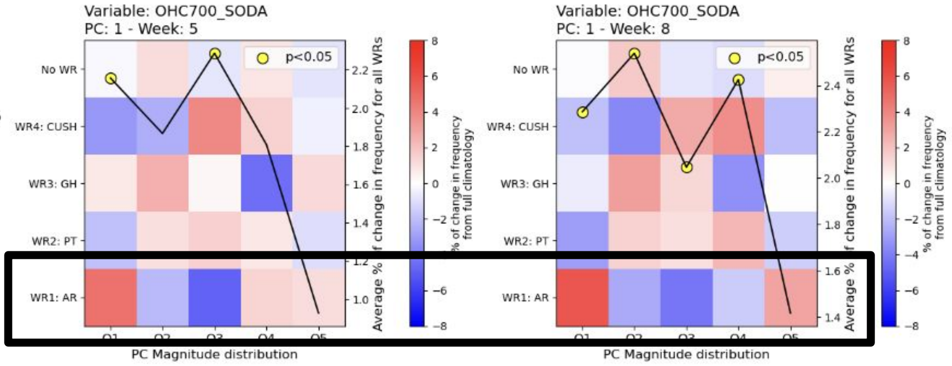




Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

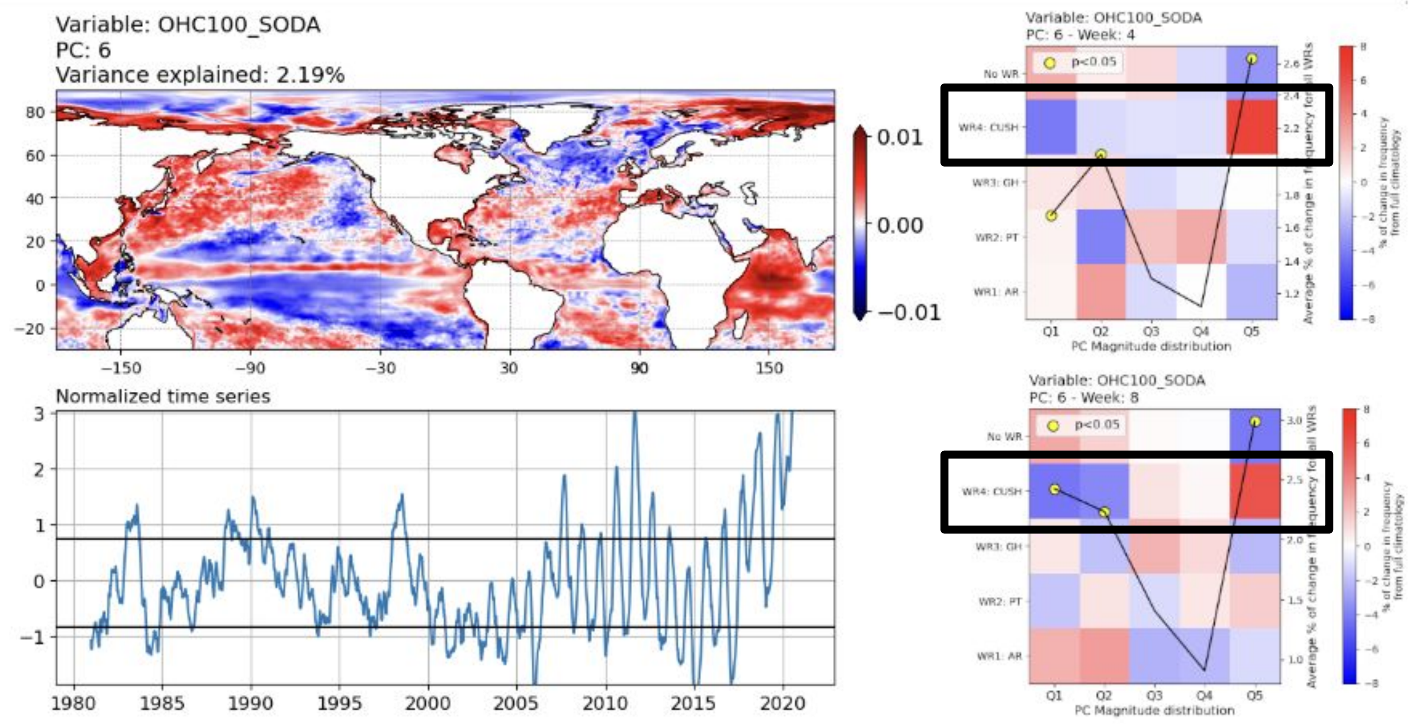
Oceanic predictors: ENSO + Arctic Ocean





Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).

Oceanic predictors: Indian Ocean



Perez-Carrasquilla, J., Molina, M. J. 2023. An Earth-System-Oriented View of the S2S Predictability of Weather Regimes using XGBoost. Artificial Intelligence for the Earth Systems (in preparation).