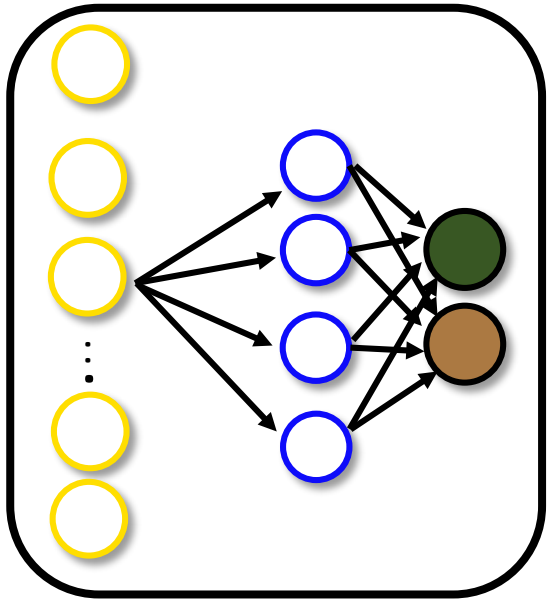
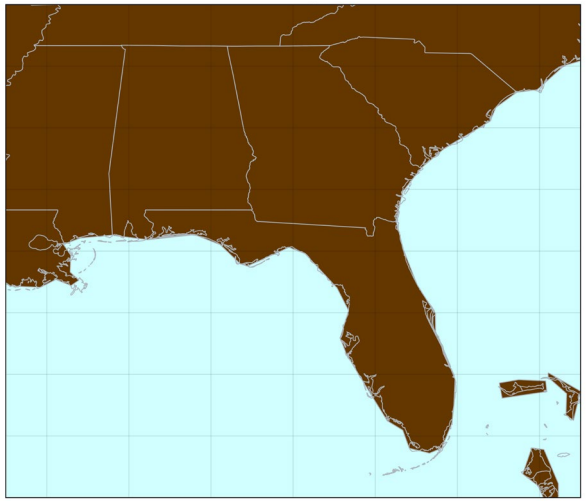
Understanding Predictability of Daily Southeast US Precipitation using Explainable Machine Learning





Kathy Pegion

George Mason University, AOES

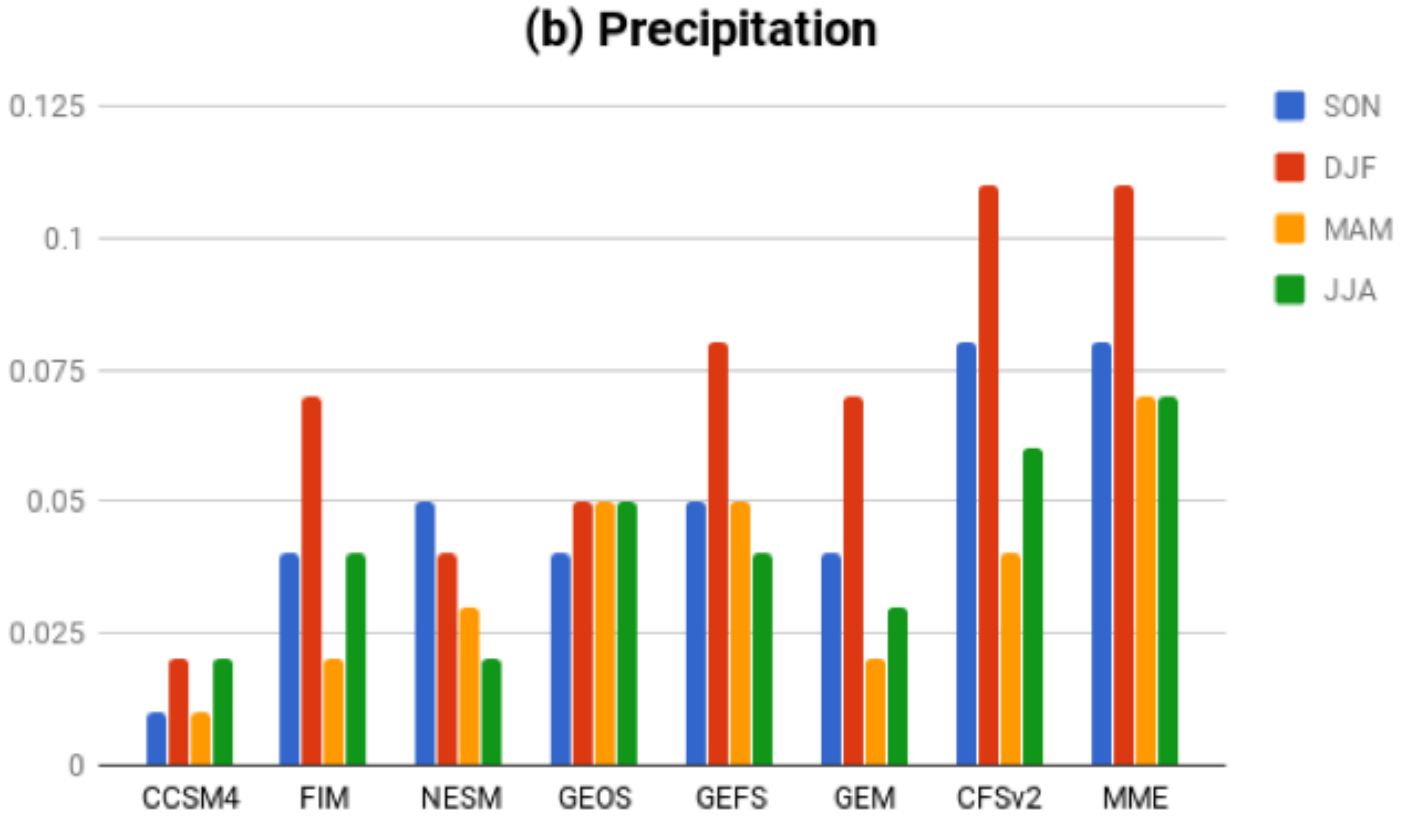
Atmospheric, Oceanic, and Earth Sciences

Pegion, K., B. Kirtman, and E. Becker,
submitted to Artificial Intelligence for the Earth System

COLA

SubX Average Anomaly Correlation

North America Week 3-4

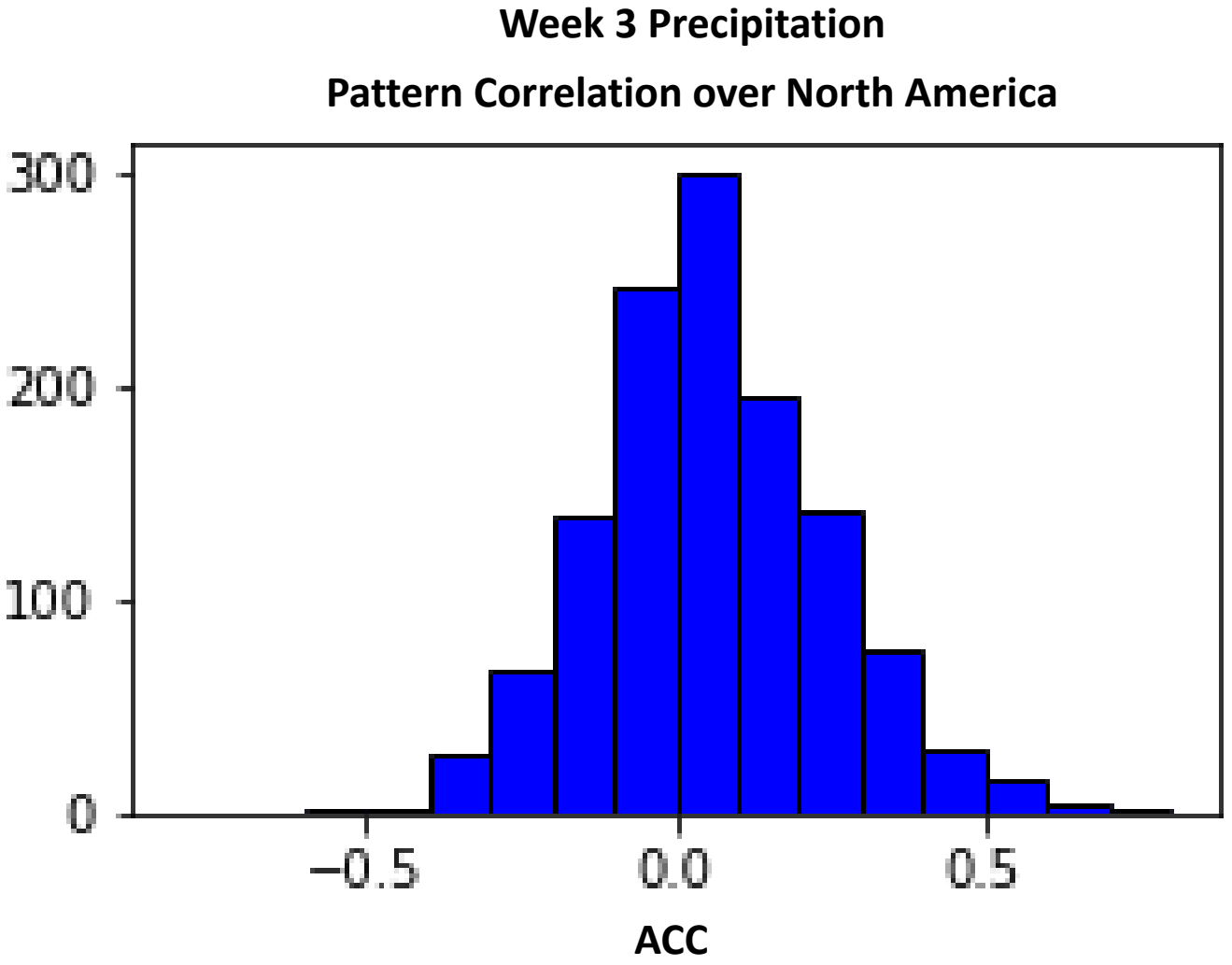


Current prediction skill of subseasonal precipitation is poor

Are these low skill forecasts useful?

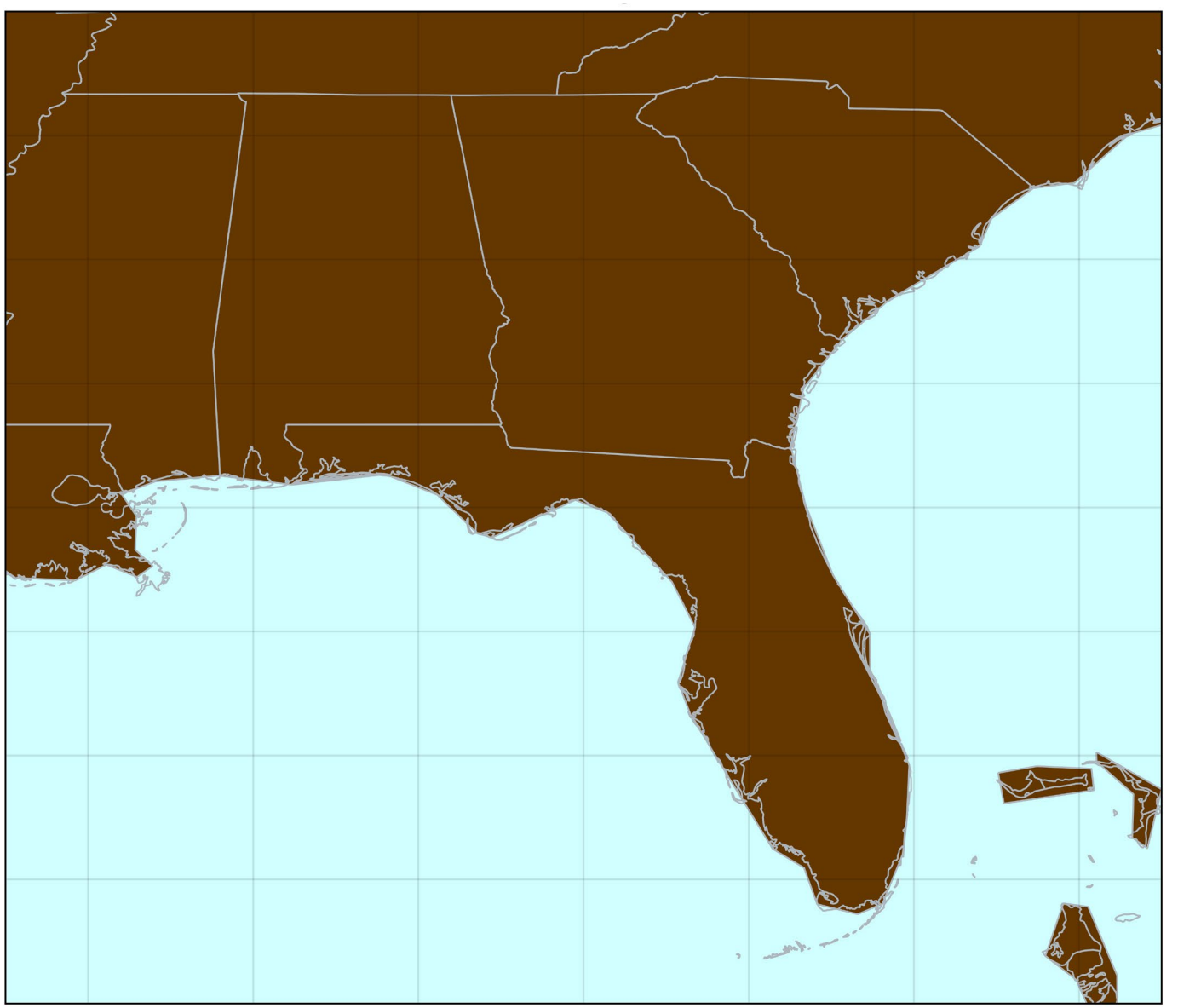
- Depends on the user & application
- Some regions have more skill
- Some times are more skillful

Forecasts of Opportunity



How can we make better and more useful S2S precipitation forecasts?

- Identify forecasts of opportunity
- Predictability problem: need to better understand sources of predictability



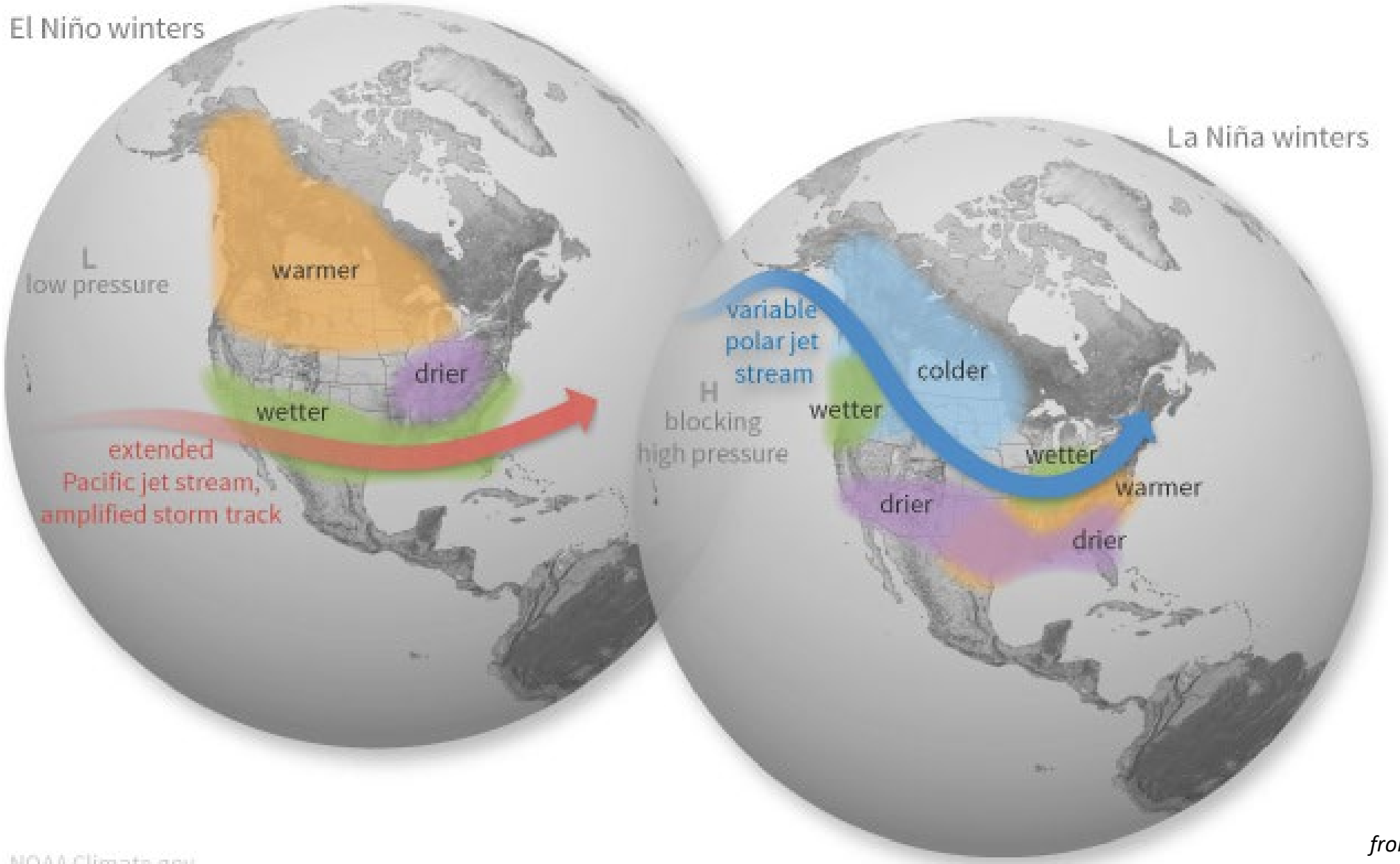
Predictor
AMO
PDO
Niño3.4
NAO
MJO Phase
MJO Amplitude
NASH Phase
NASH Amplitude
PNA Weather Regimes
MLSO

Sources of Predictability

- Lots of potential sources of predictability
- How do we disentangle them?

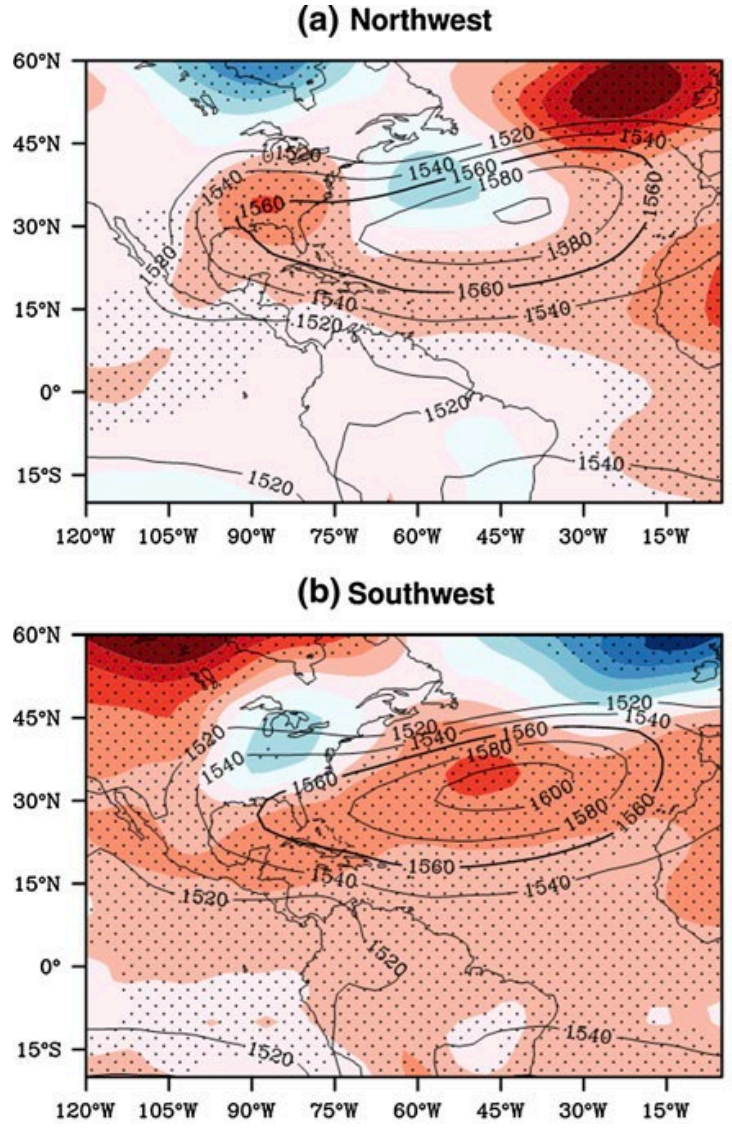
Explainable Machine Learning

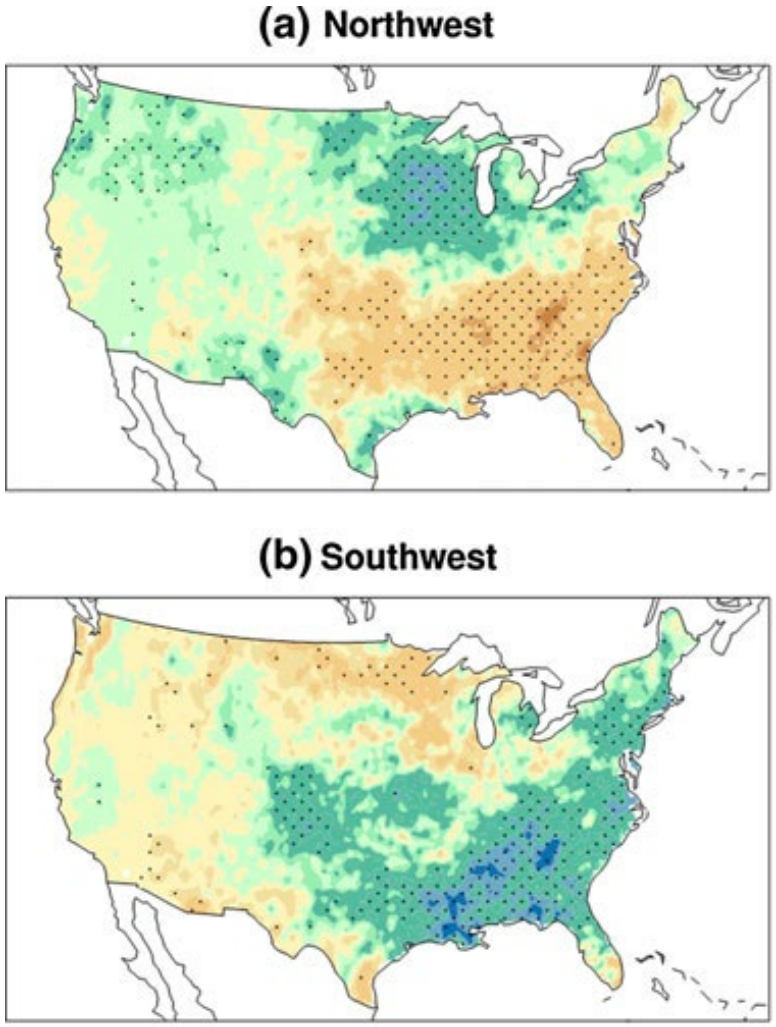
How does ENSO Impact SEUS Precipitation?

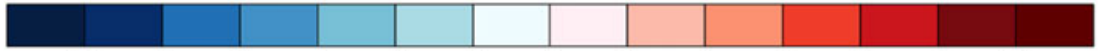


from NOAA Climate.gov

How does the NASH Impact SEUS Precipitation?





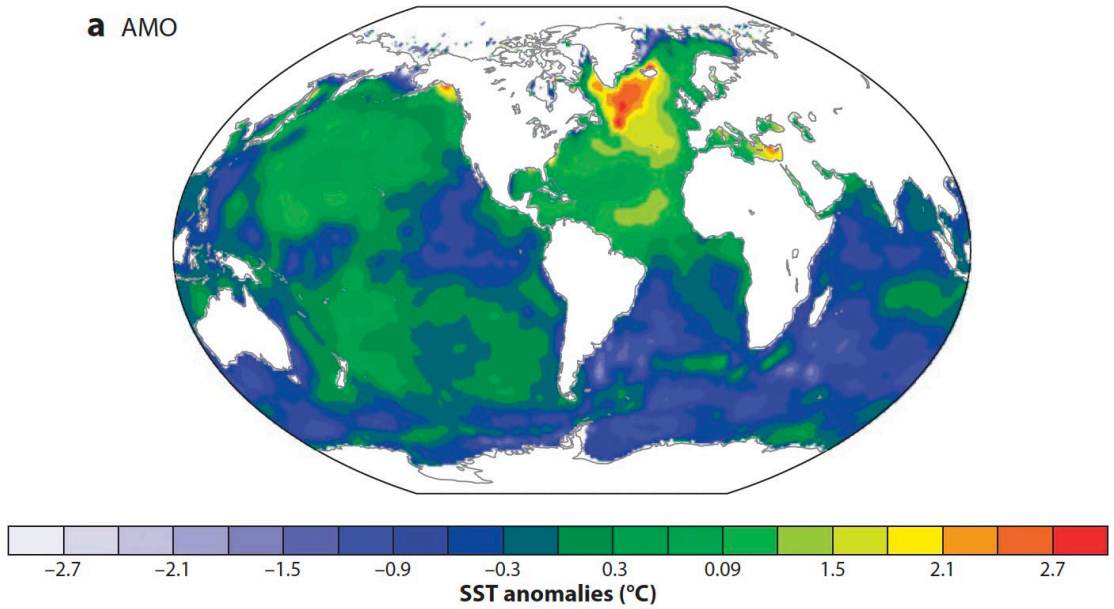


gpm

mm day^{-1}

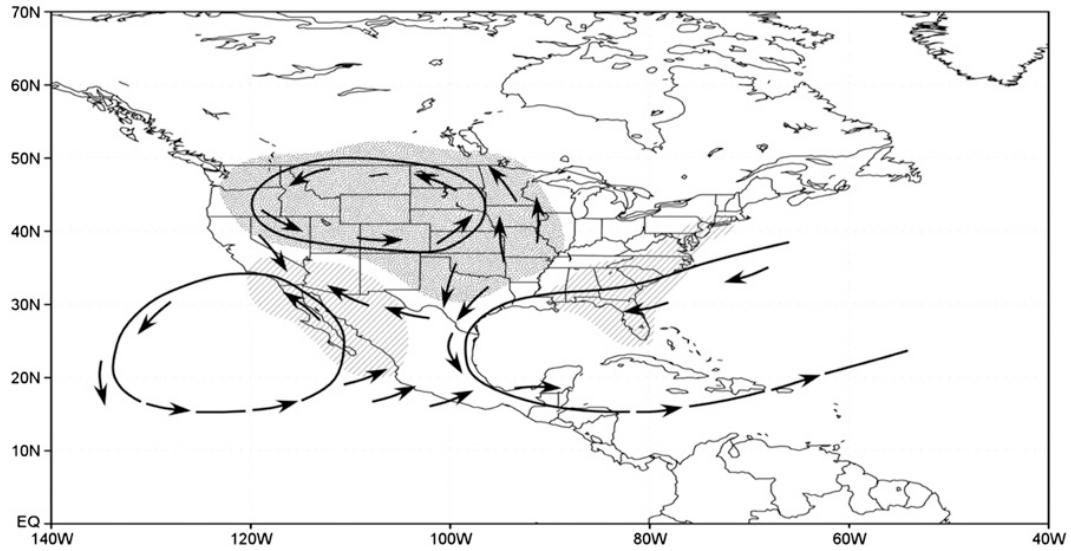
from Li et al. 2012, Clim. Dyn.

How does the AMO Impact SEUS Precipitation?

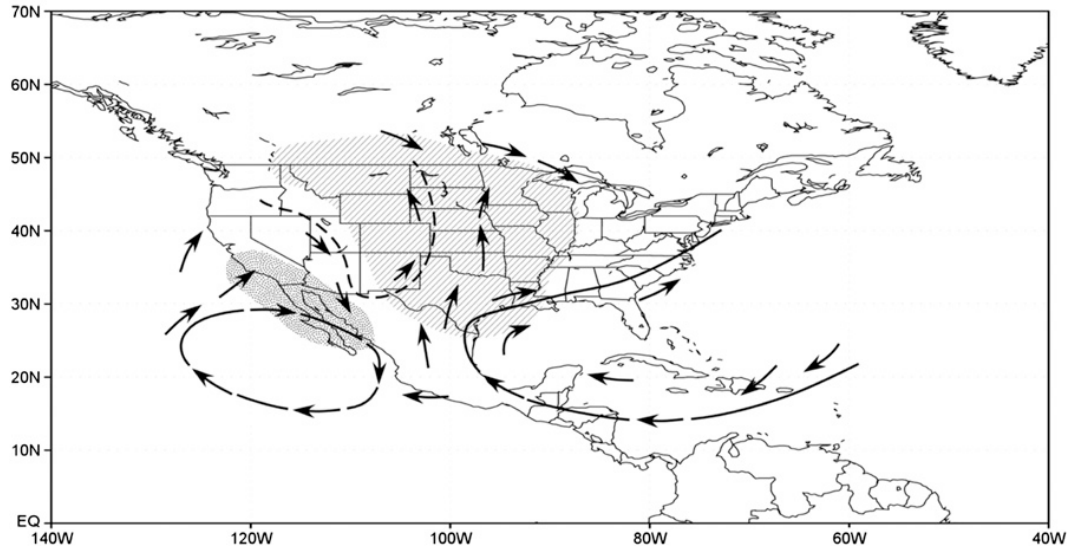


from Deser at al. 2010, Ann Rev Marine Sci.

a) Three-cell anomalous circulation in lower troposphere during warm phase



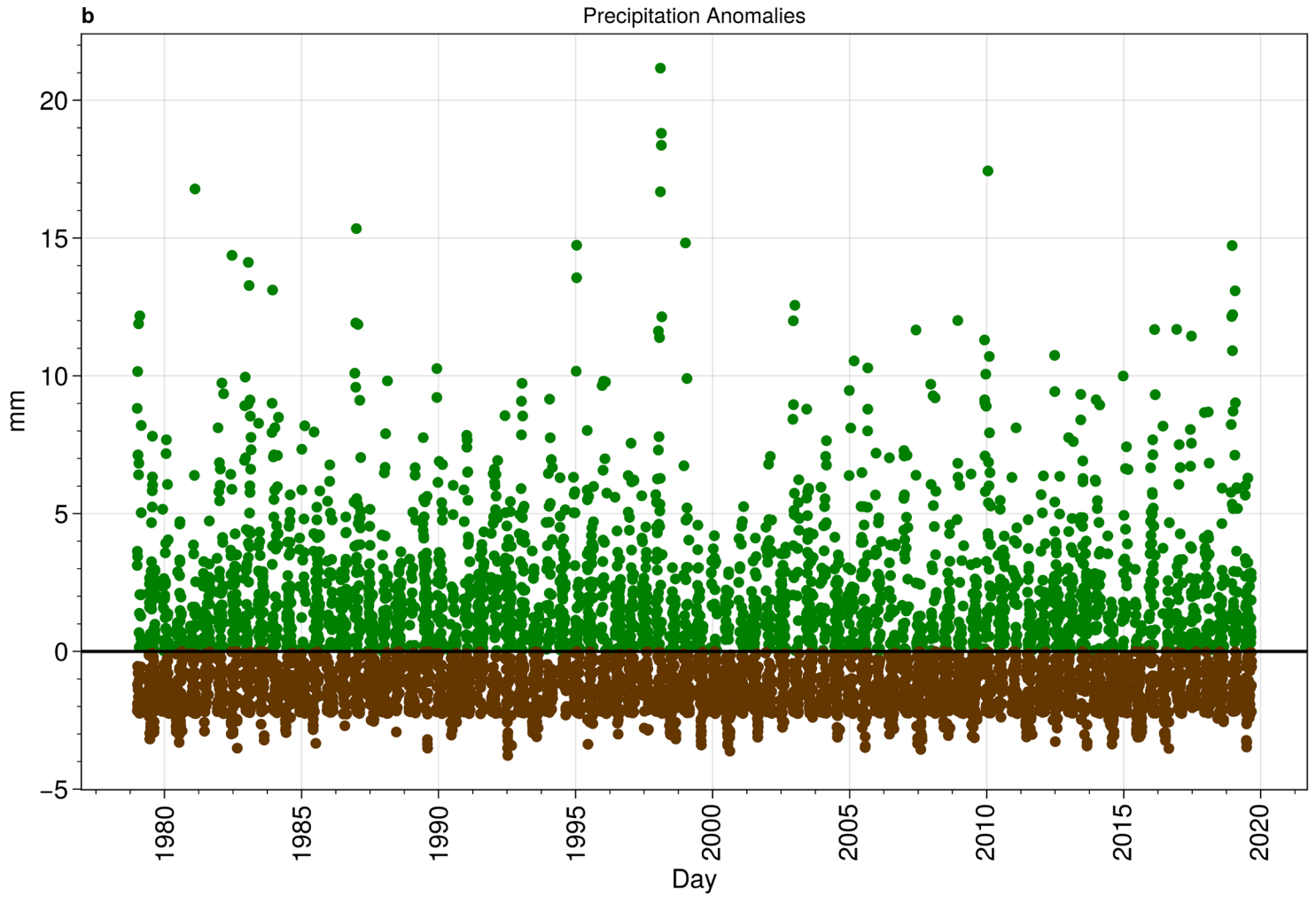
b) Three-cell anomalous circulation in lower troposphere during cold phase

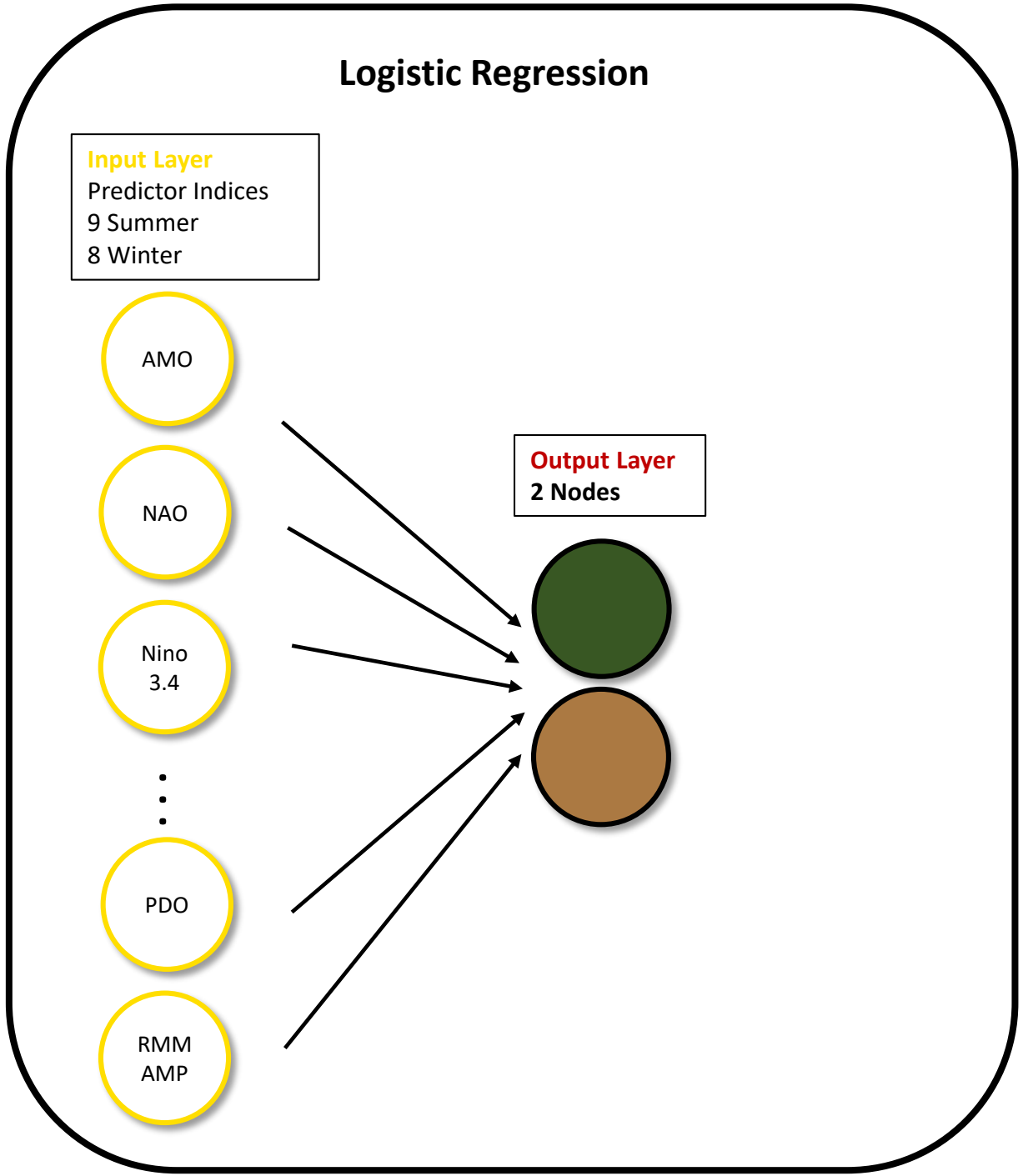


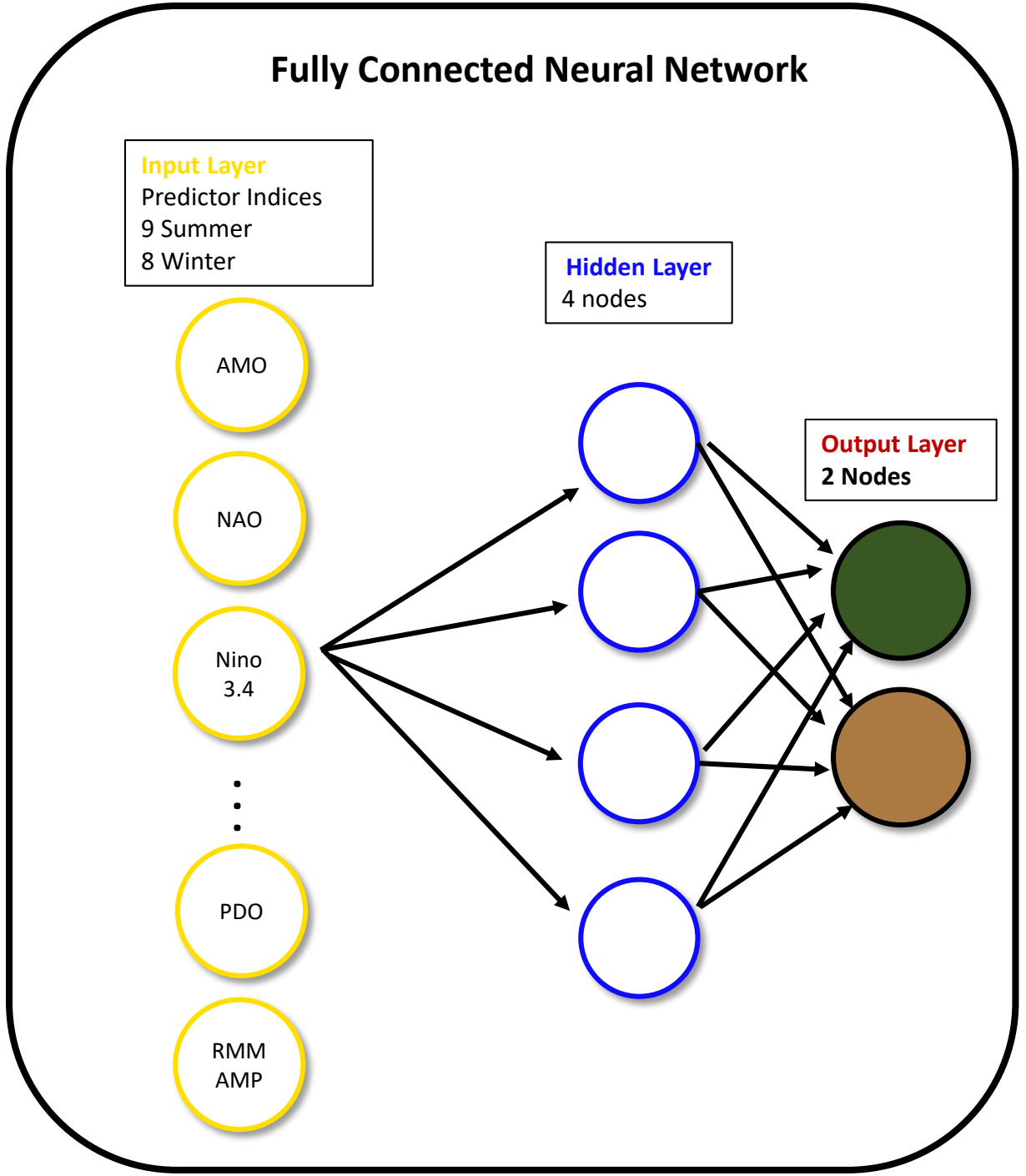
from Hu et al., J.Clim 2011

How well can we predict the sign of SEUS Precipitation anomalies if we know the predictors perfectly?

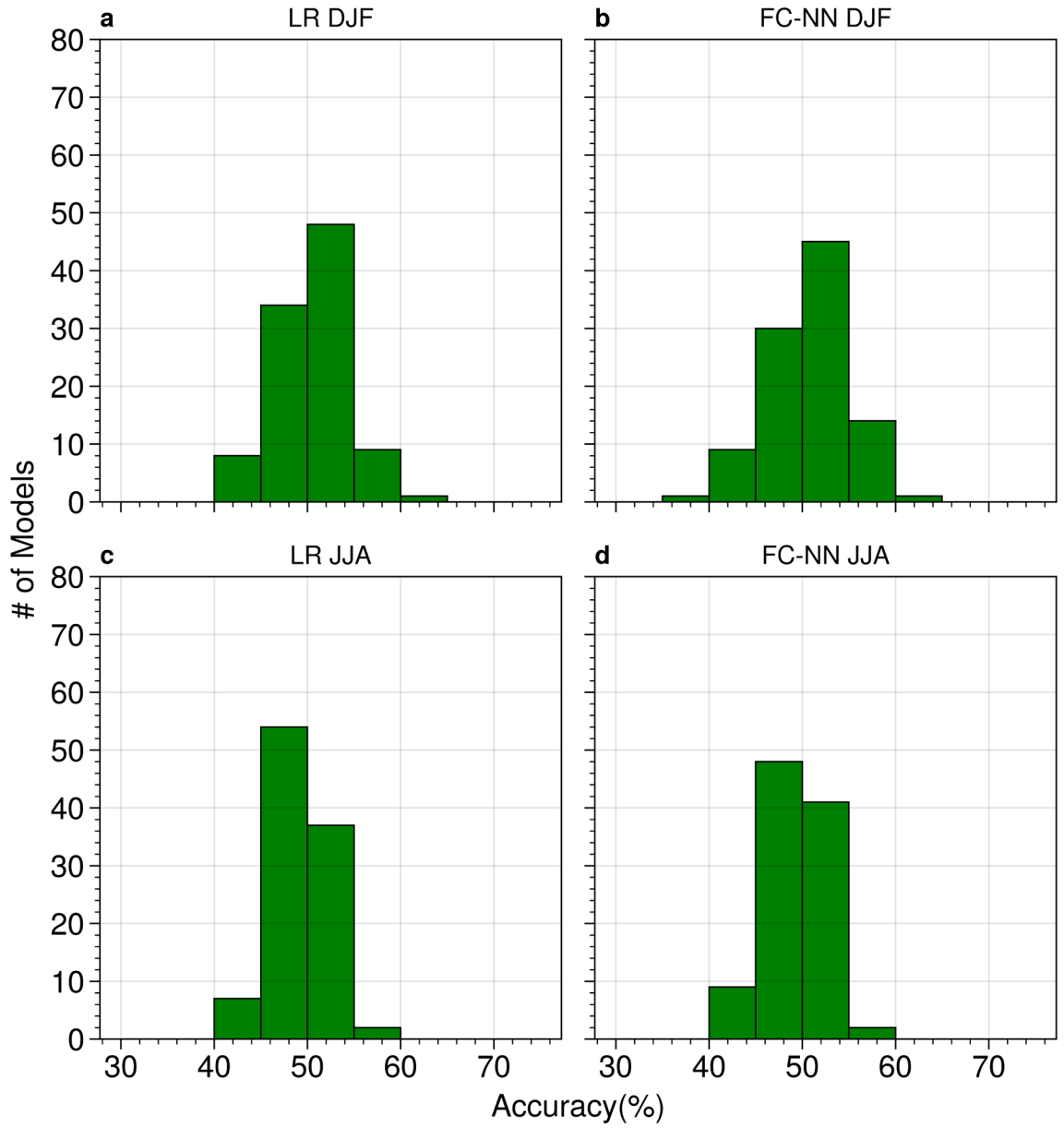
- Large-scale climate indices as predictors
- Gridded fields as predictors



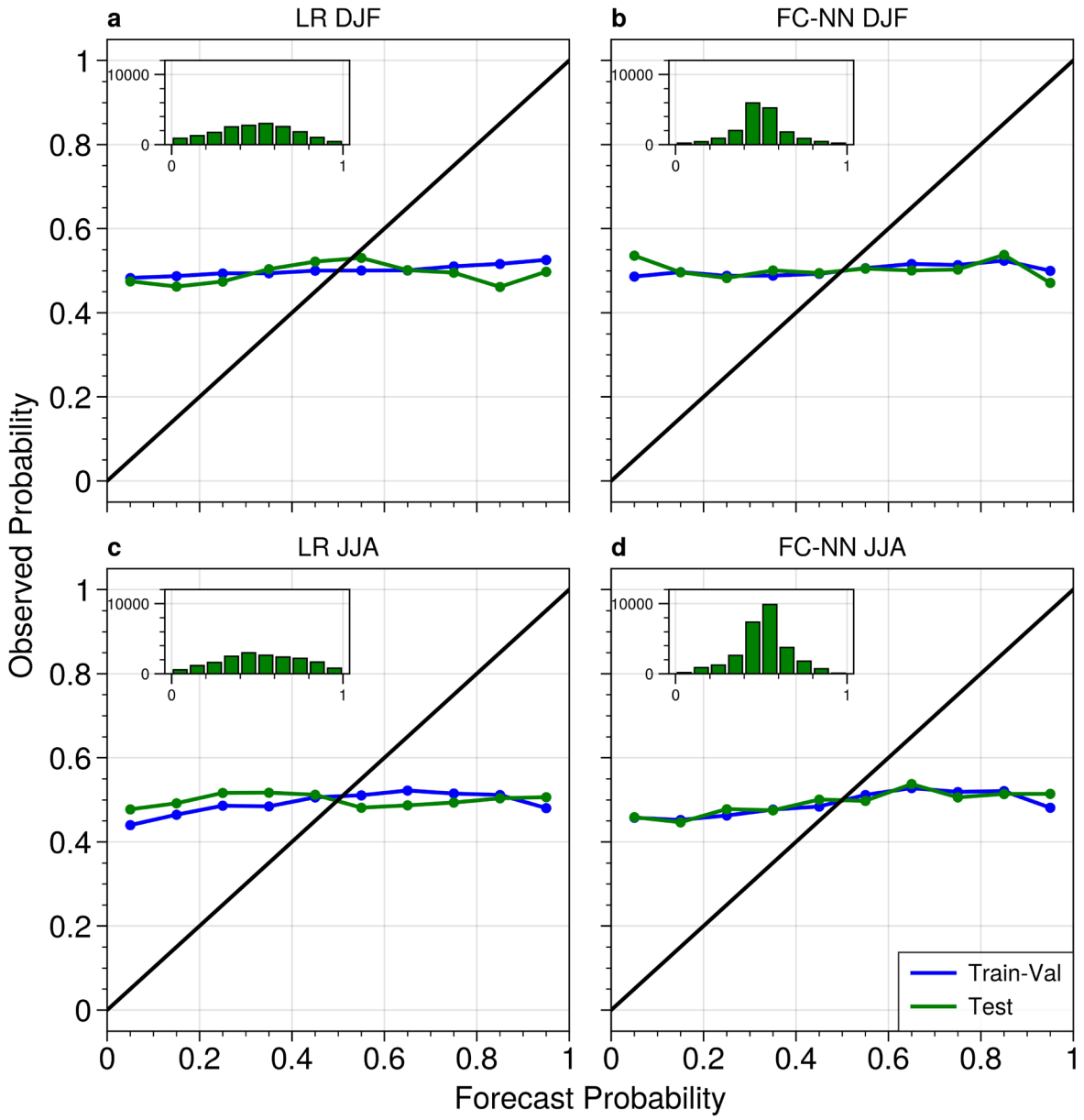

Logistic Regression
Input Layer
Predictor Indices
9 Summer
8 Winter
AMO
NAO
Nino
3.4
PDO
RMM
AMP
Output Layer
2 Nodes


Fully Connected Neural Network
Input Layer
Predictor Indices
9 Summer
8 Winter
Hidden Layer
4 nodes
AMO
NAO
Nino
3.4
PDO
RMM
AMP
Output Layer
2 Nodes

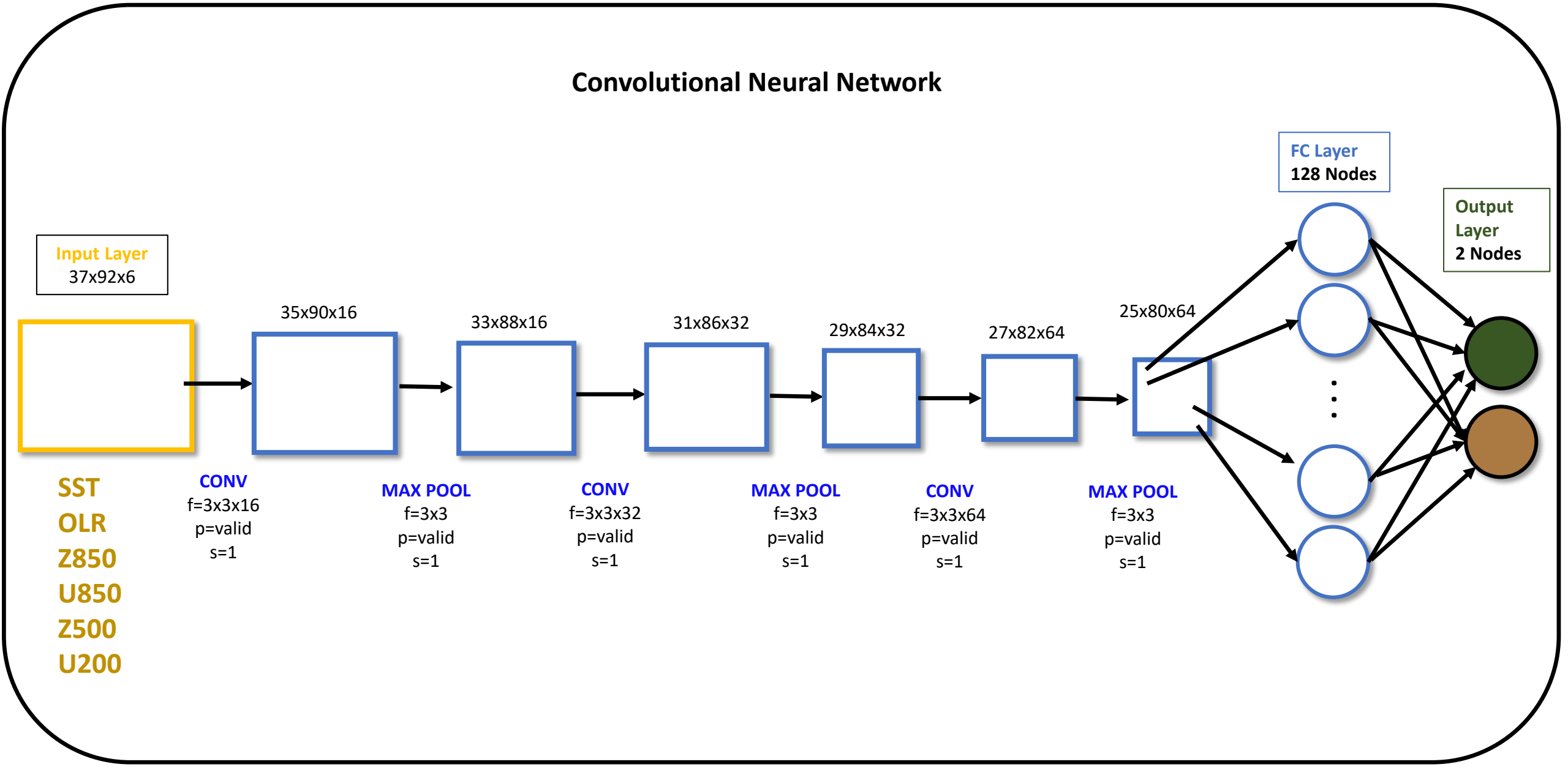
Models are no more accurate than flipping a coin



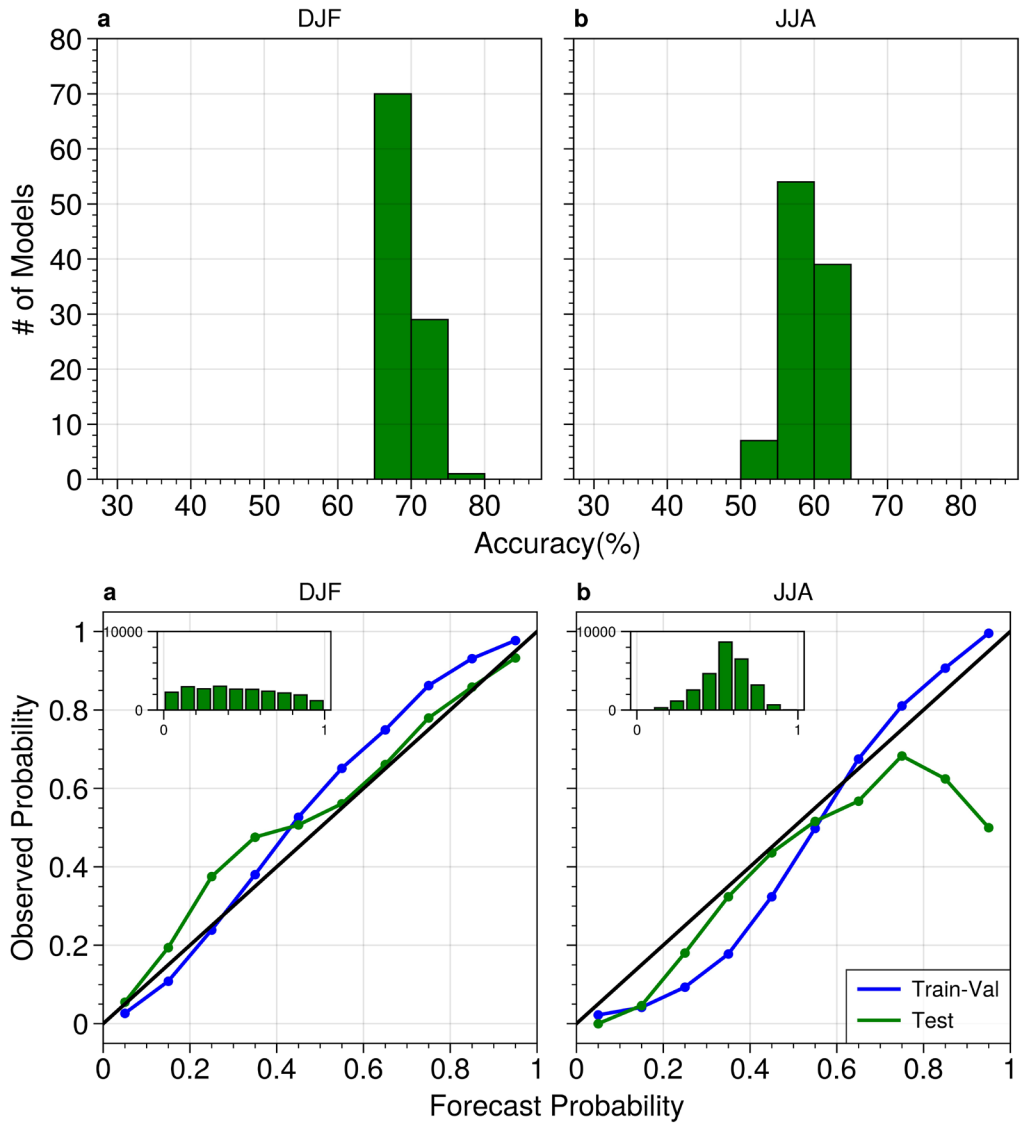
Model probabilities are not reliable



from Pegion et al., submitted to AIES

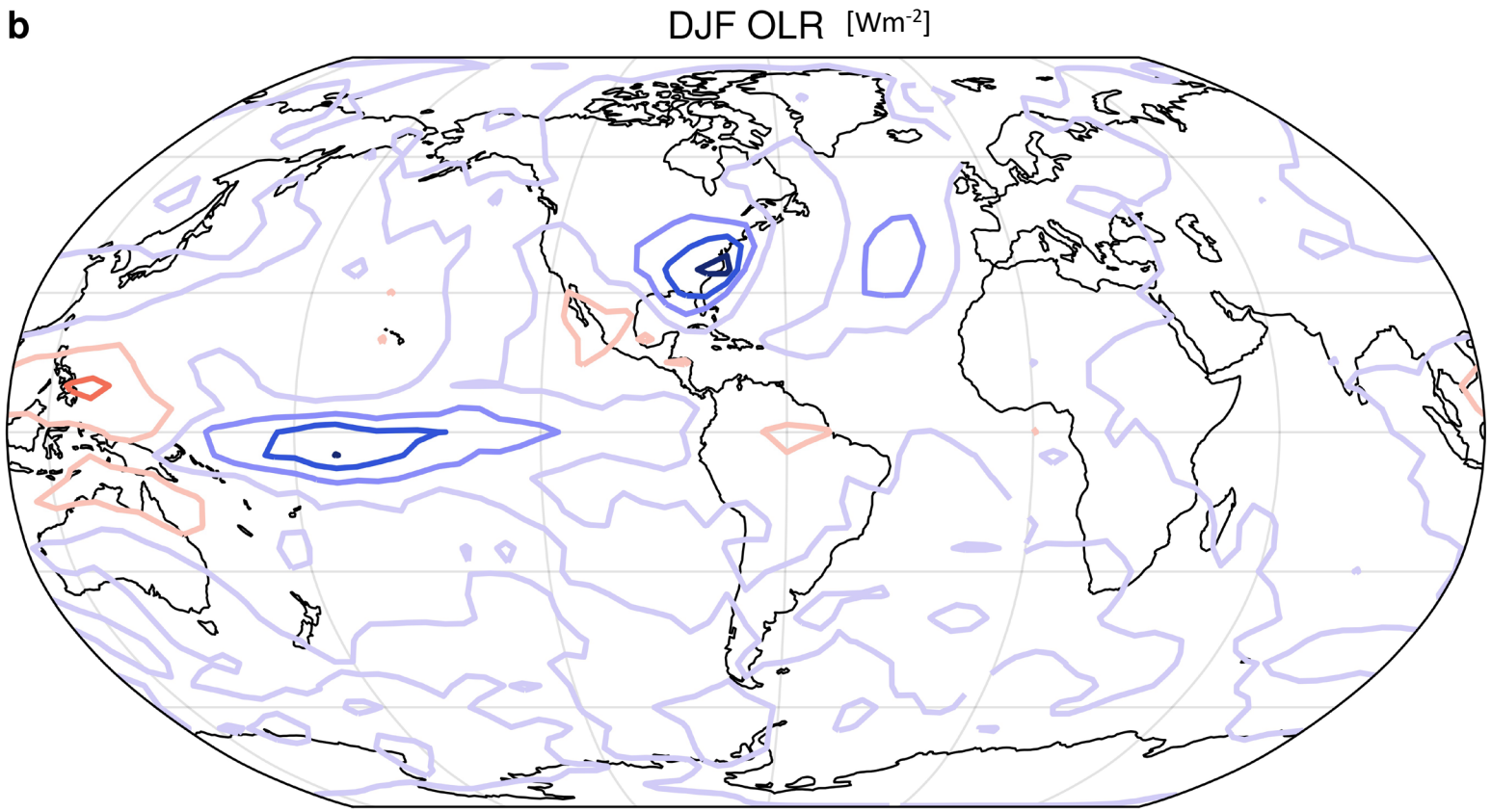

Convolutional Neural Network
Input Layer
37x92x6
SST
OLR
Z850
U850
Z500
U200
35x90x16
CONV
f=3x3x16
p=valid
s=1
33x88x16
MAX POOL
f=3x3
p=valid
s=1
31x86x32
CONV
f=3x3x32
p=valid
s=1
29x84x32
MAX POOL
f=3x3
p=valid
s=1
27x82x64
CONV
f=3x3x64
p=valid
s=1
25x80x64
MAX POOL
f=3x3
p=valid
s=1
FC Layer
128 Nodes
Output
Layer
2 Nodes

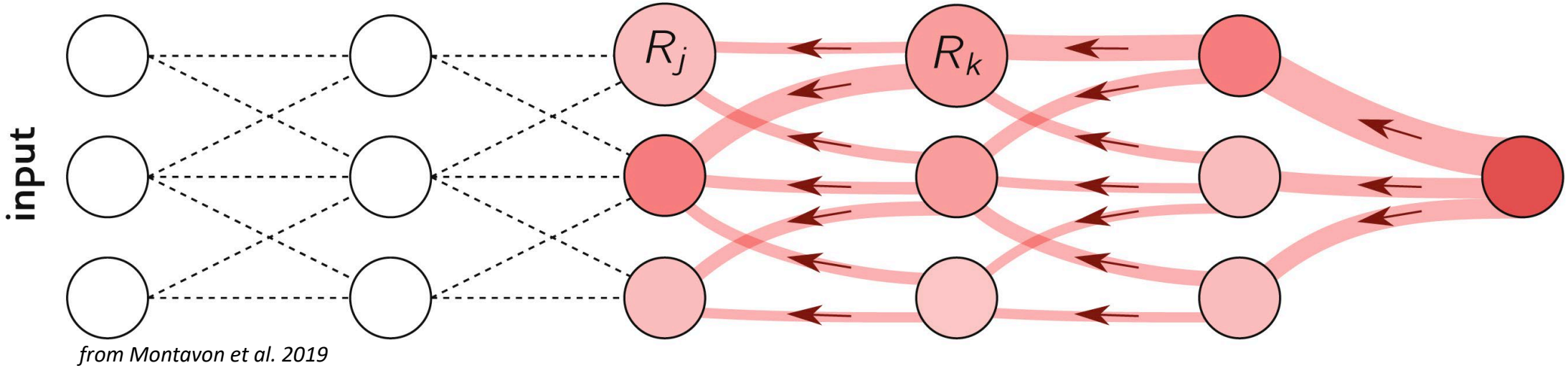
- Forecasts are more accurate
- Forecasts are reliable
- Probabilities identify forecasts of opportunity



Positive SEUS Precipitation Anomalies

Confident (>=80%) & Correct Forecasts



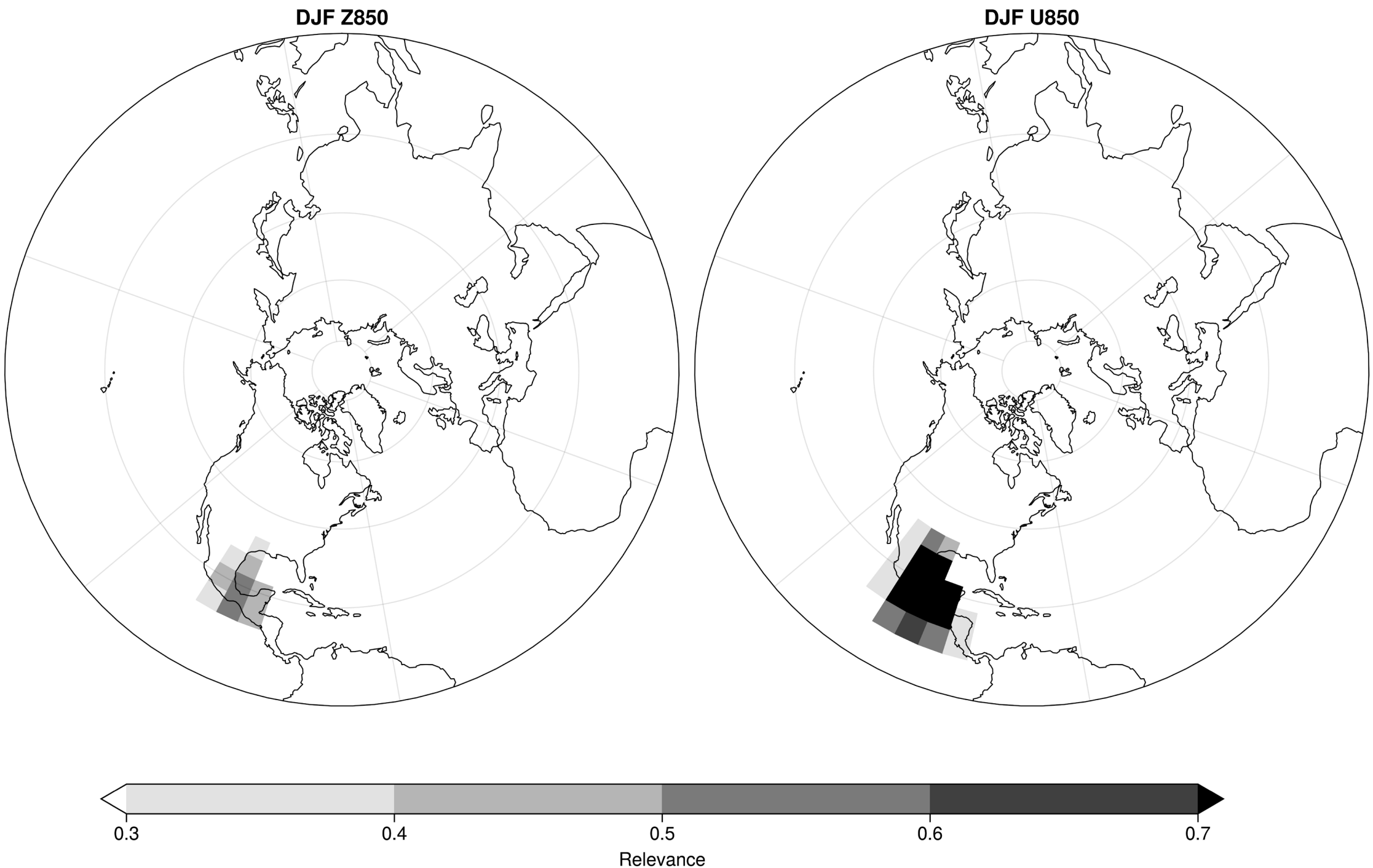


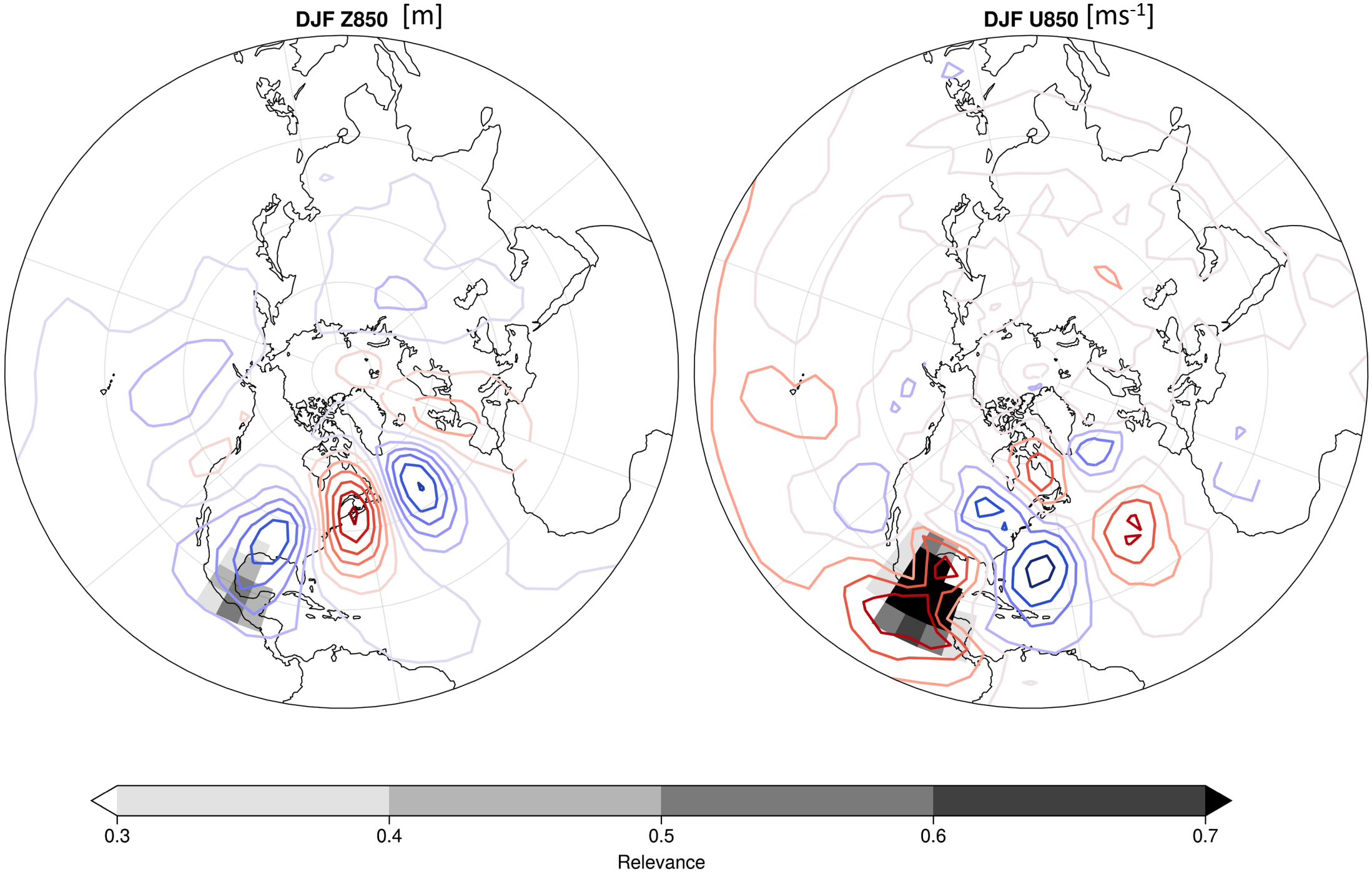
from Montavon et al. 2019

Explainable ML Methodology Layerwise Relevance Propagation (LRP)

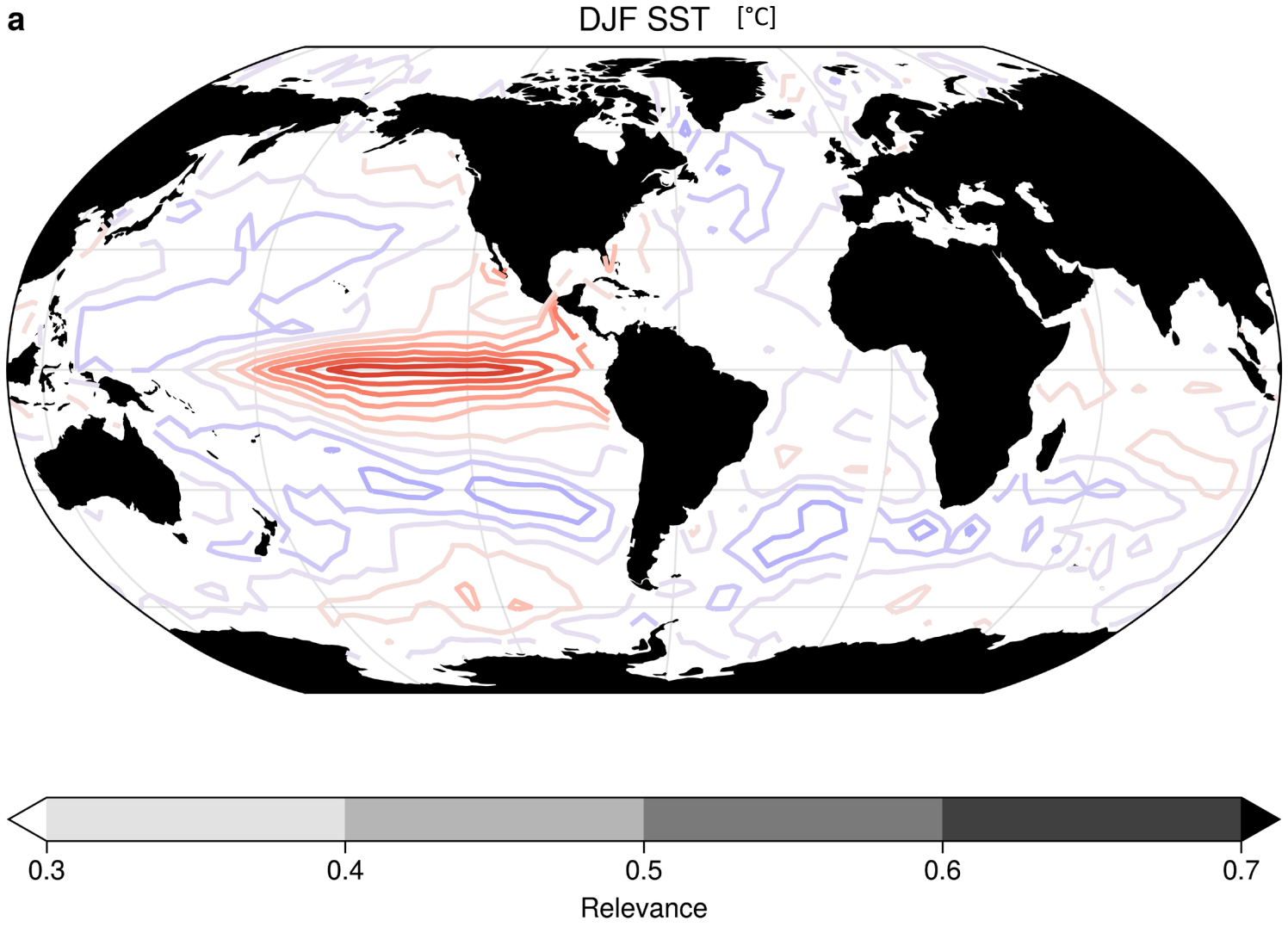
- Get a heat map that shows us the relevant variables and gridpoints
- Applied to each forecast individually
- Consistently relevant predictors across many inputs are our "sources of predictability"

Bach et al. 2015, Montavon et al. 2019, Mayer and Barnes 2021, Toms et al. 2020, 2021


DJF Z850
DJF U850
0.3
0.4
0.5
0.6
0.7
Relevance

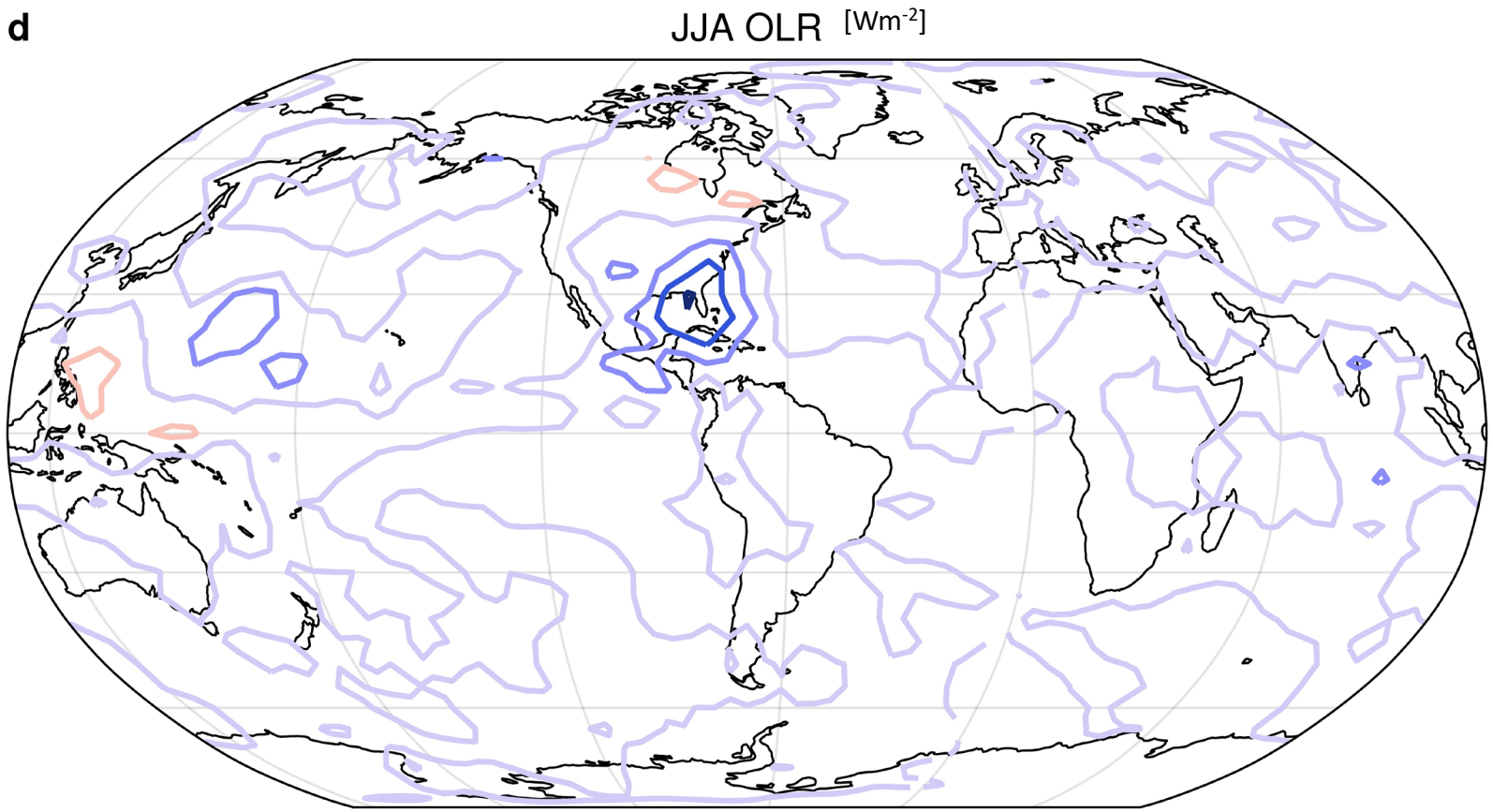

DJF Z850 [m]
DJF U850 [ms-1]
0.3
0.4
0.5
0.6
0.7
Relevance

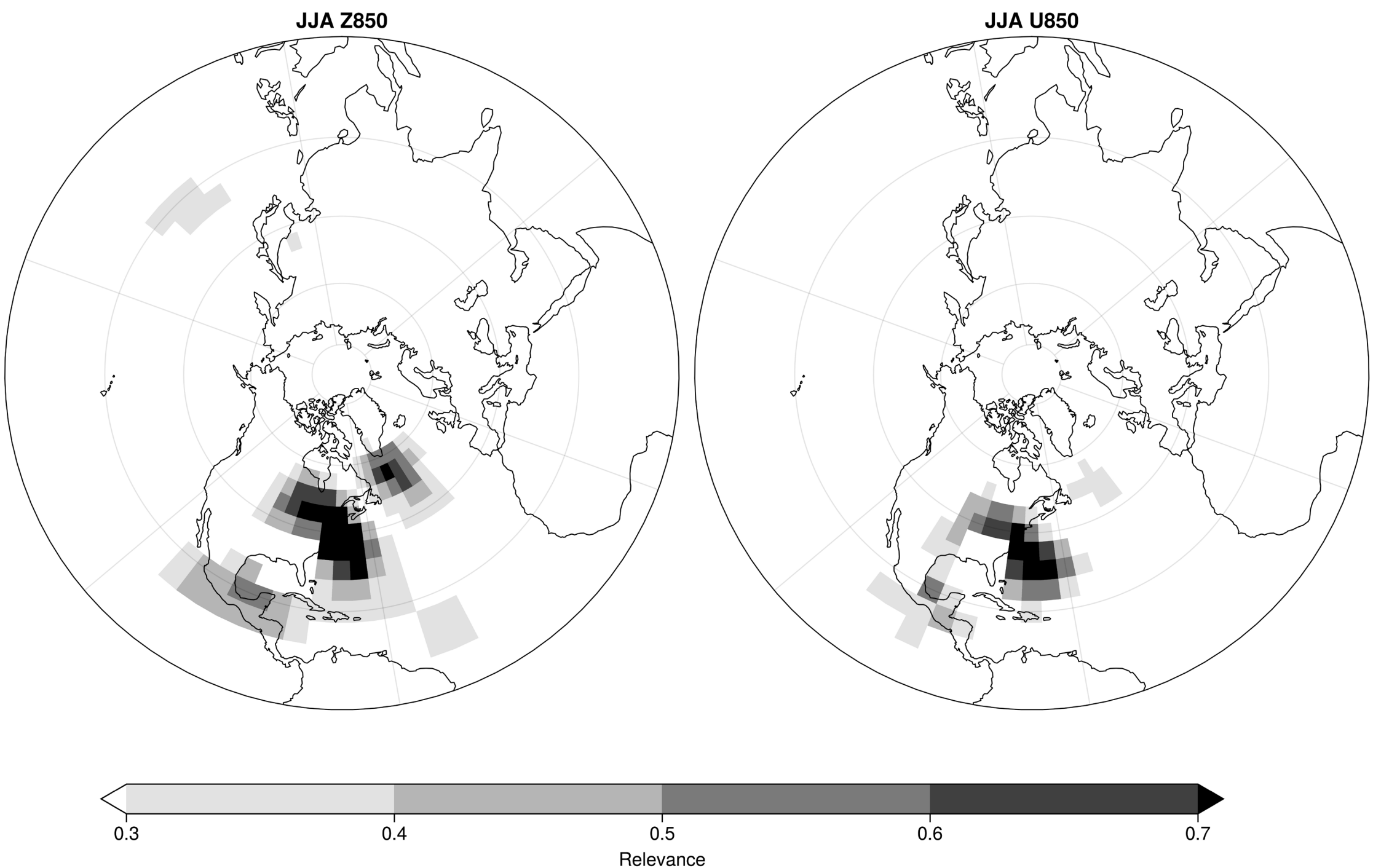
Winter is Related to El Niño

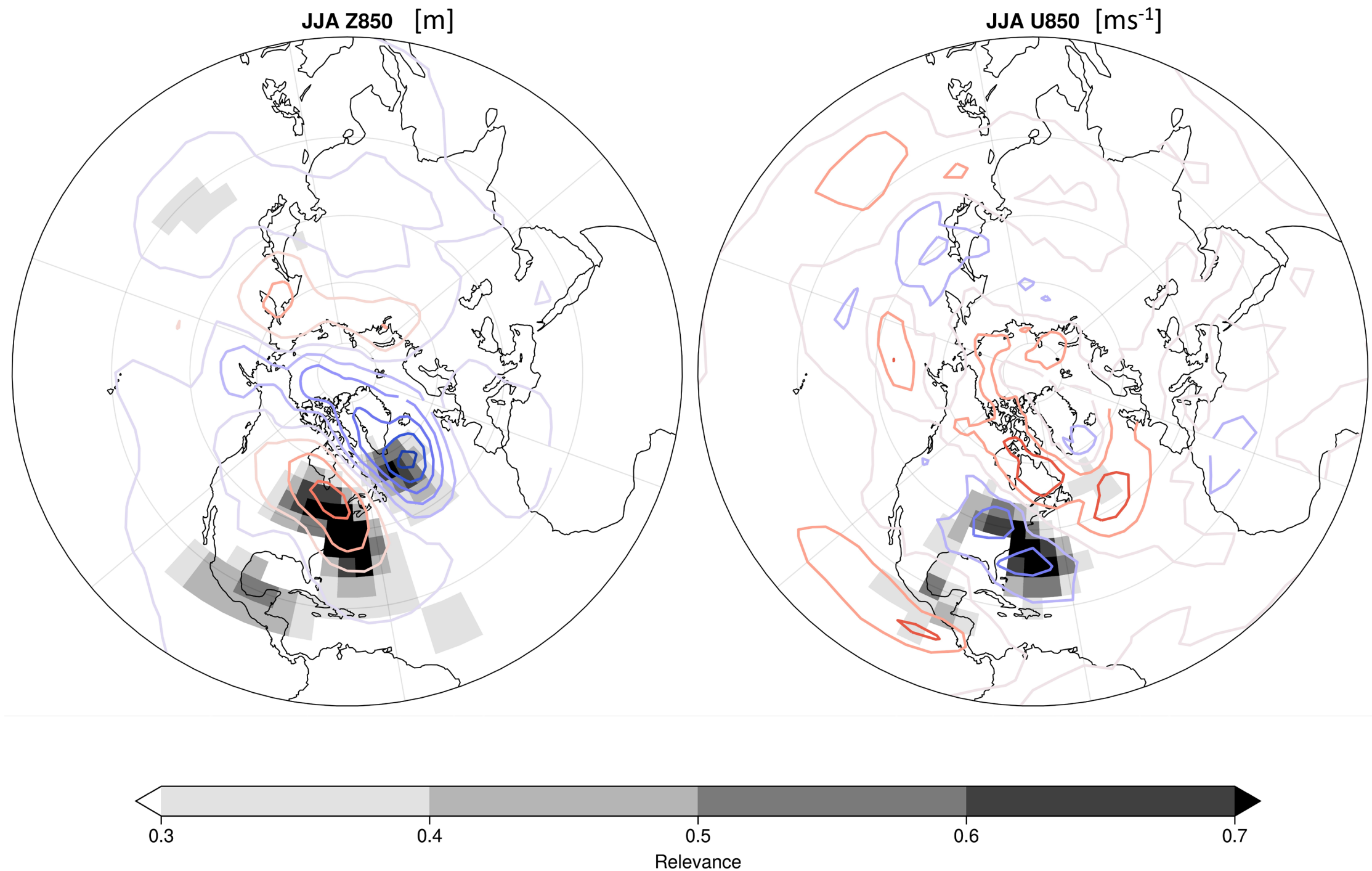


Positive SEUS Precipitation Anomalies

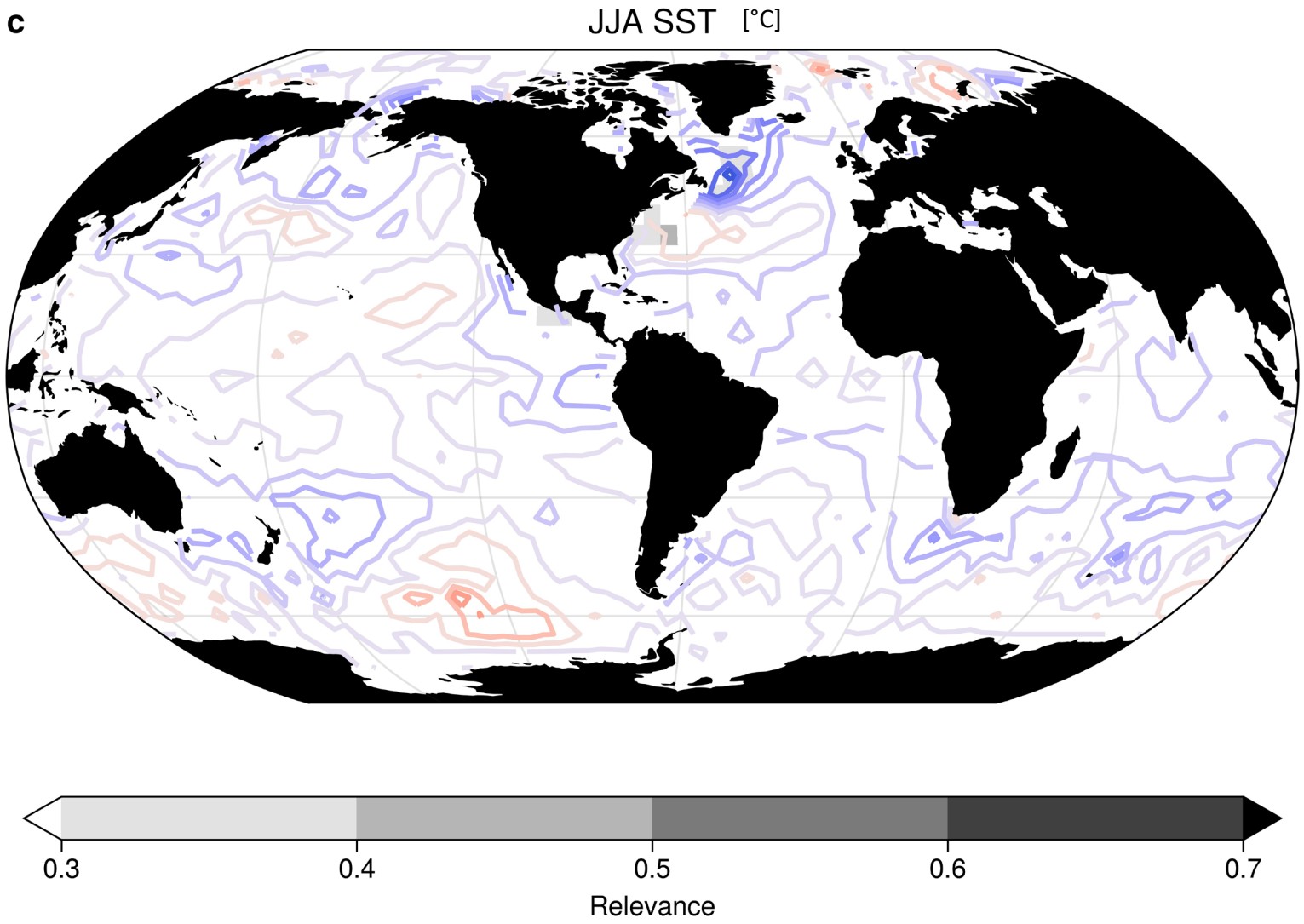
Confident (>=80%) & Correct Forecasts



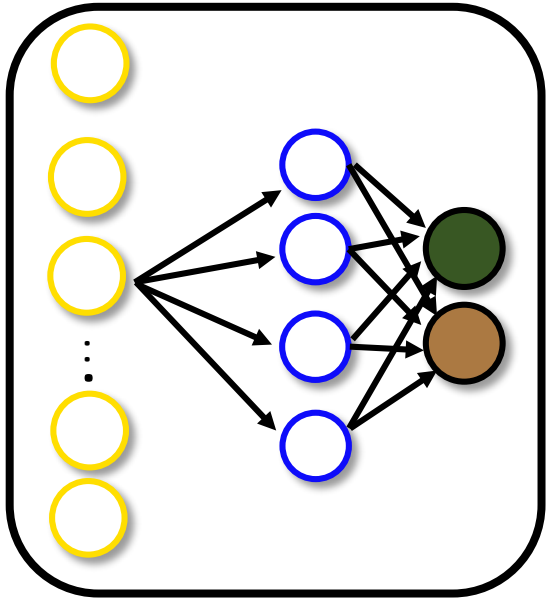

JJA Z850
JJA U850
0.3
0.4
0.5
0.6
0.7
Relevance

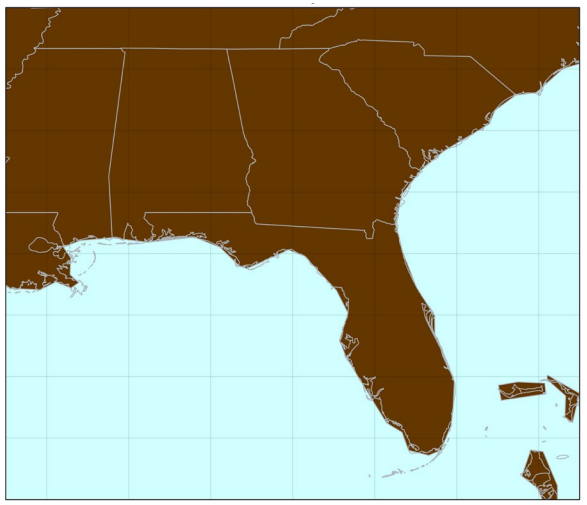

JJA Z850 [m]
JJA U850 [ms-1]
0.3
0.4
0.5
0.6
0.7
Relevance

Summer is Related to North Atlantic SST



Summary & Conclusions





- Using indices of large-scale climate phenomena is insufficient to skillfully or reliably predict the sign of daily precipitation anomalies in the SEUS.
- Global gridded fields as predictors are accurate and reliable allowing us to identify forecasts of opportunity
- The CNN can identify physically relevant relationships between large scale climate and daily SEUS precipitation
- Local circulation is most important
- Winter: related to ENSO
- Summer: related to NASH-N Atl SST

Next Steps

Expand to other regions
Use what we have learned about the SST as predictors for summer to perform sensitivity re-forecast experiments with the NCAR model.
Real prediction with a lead-time to understand how the most relevant predictors change over a subseasonal to seasonal forecast.
Use predicted values of large-scale climate fields from the SubX models to understand how errors in these fields lead to errors in precipitation forecasts.
Predict extreme precipitation events during a week or season.