An Interpretable Neural Network Approach for Identifying Sources of Predictability on Decadal Timescales in the CESM2-LE

Emily Gordon[1], Elizabeth Barnes[1] and Frances Davenport[2]

[1]Department of Atmospheric Science, Colorado State University
[2]Department of Civil and Environmental Engineering, Colorado State University

CESM Workshop ESPWG Meeting 2023

Can we predict near-term climate trends?

In the future climate, we consider predictability from 2 sources:

- External forcing
- Internal variability

Can we predict near-term climate trends?

In the future climate, we consider predictability from 2 sources:

- **External forcing**
- Internal variability

Predictability from external forcing

Anthropogenic climate change means that global mean temperature will continue to rise

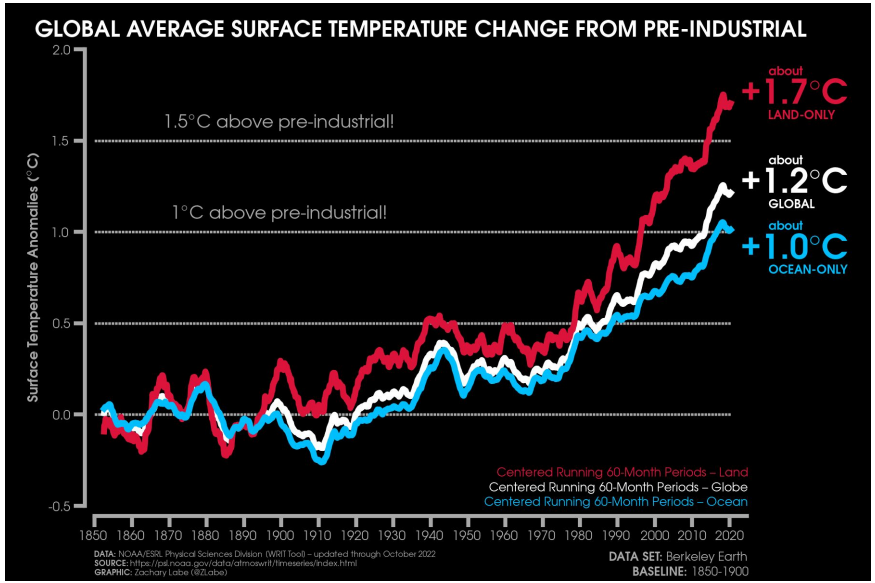


Figure created by Zack Labe, used with permission

Can we predict near-term climate trends?

In the future climate, we consider predictability from 2 sources:

- External forcing
- **Internal variability**

Predictability from internal variability

Large scale ocean patterns like Pacific decadal variability (PDV) and Atlantic multi-decadal variability (AMV) can provide predictability e.g. Meehl et. al 2016, Borchert et. al 2018, Gordon et. al 2021, Gordon and Barnes 2022

Observed PDO/PDV

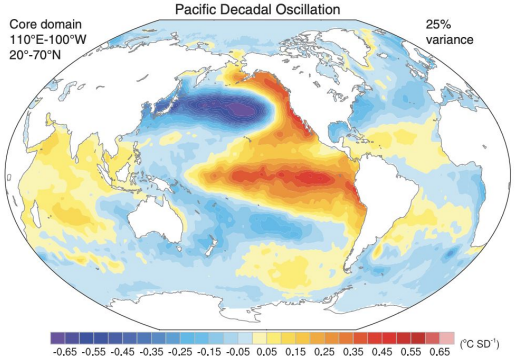


Figure from Trenbert & Fasullo 2013, *Earth's Future*

Observed AMO/AMV

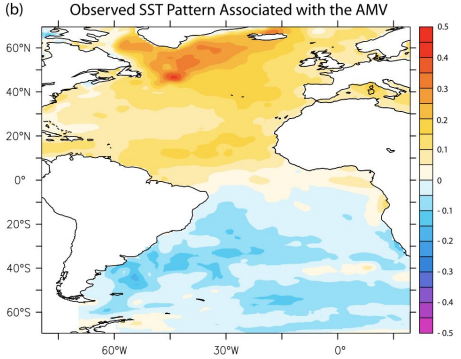
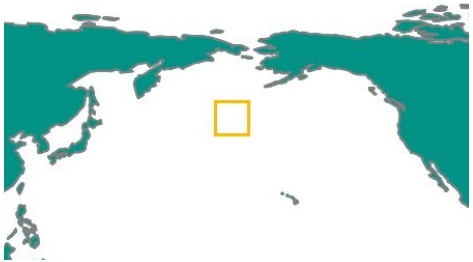
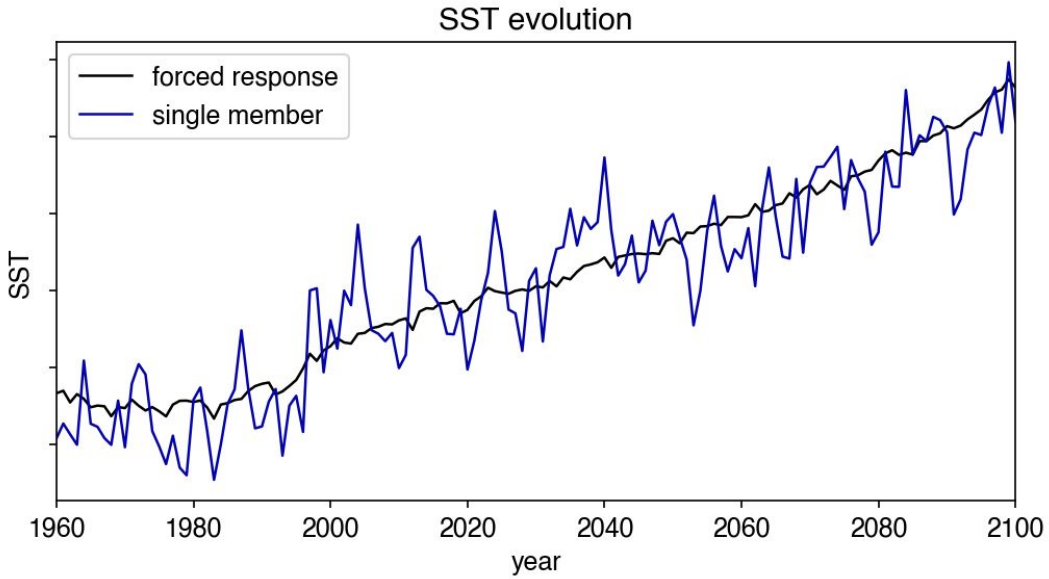


Figure from Zhang et. al 2019, *Reviews of Geophysics*

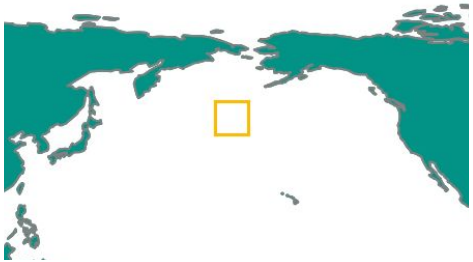
Predicting regional decadal climate trends



- 10 x 10 degree box
- Annual mean SST
- CESM2 LENS from 1960 to 2100

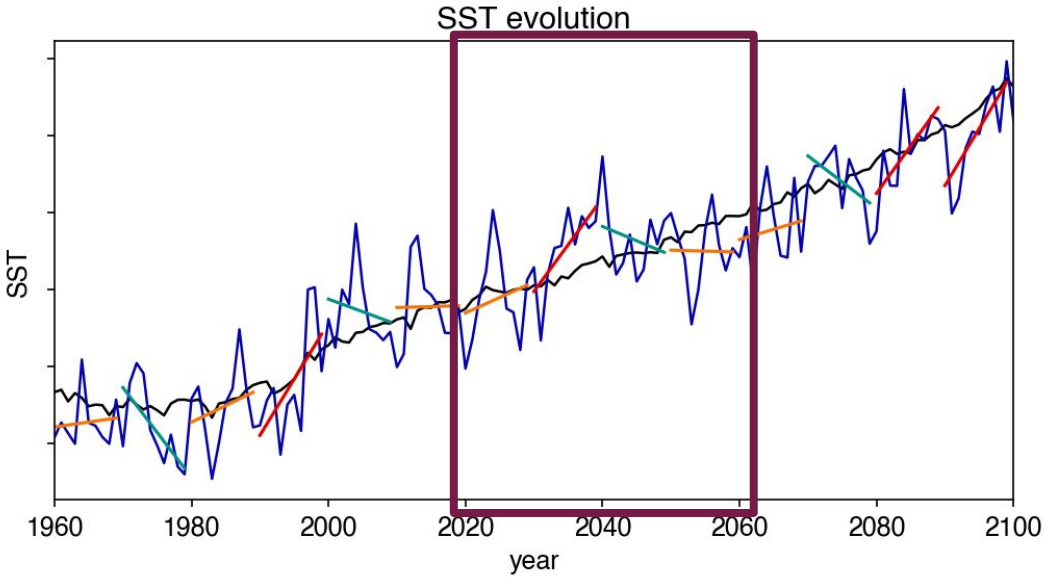


Predicting regional decadal climate trends

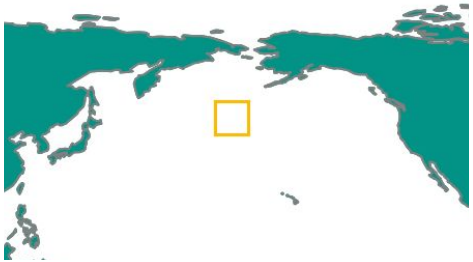


- 10 x 10 degree box
- Annual mean SST
- CESM2 LENS from 1960 to 2100

Regional evolution shows periods of **increased warming** and even **cooling** with high anthropogenic forcing

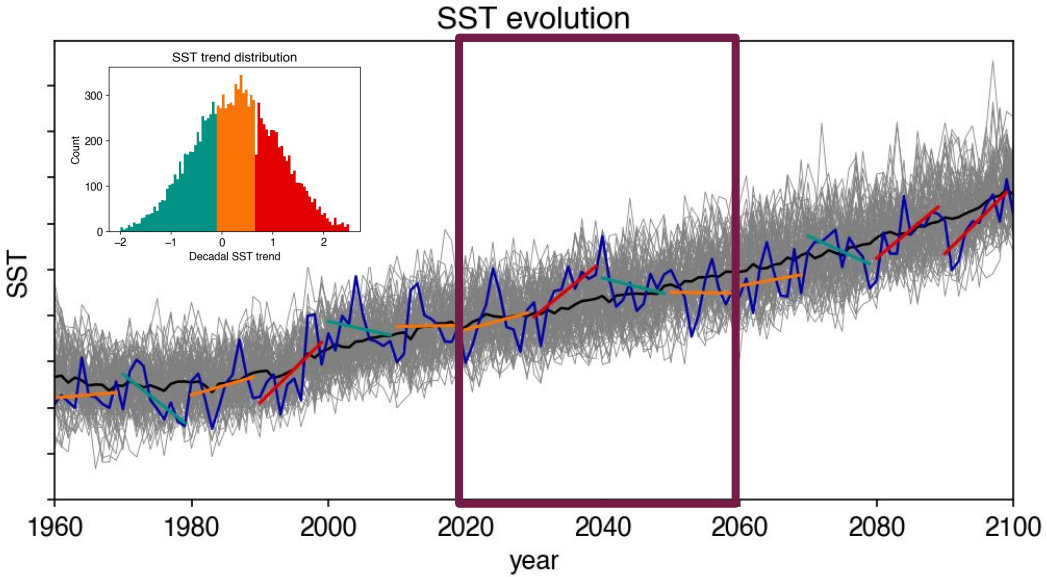


Predicting regional decadal climate trends



- 10 x 10 degree box
- Annual mean SST
- CESM2 LENS from 1960 to 2100

- Calculate decadal (10 year) SST trends
- Build a distribution of these trends from 2020-2050, rank in terciles (thirds)

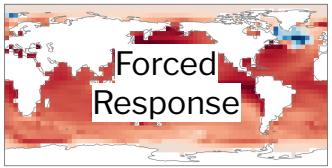


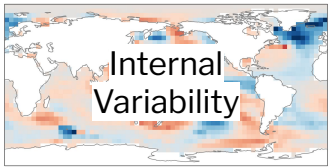
Predicting regional decadal climate trends

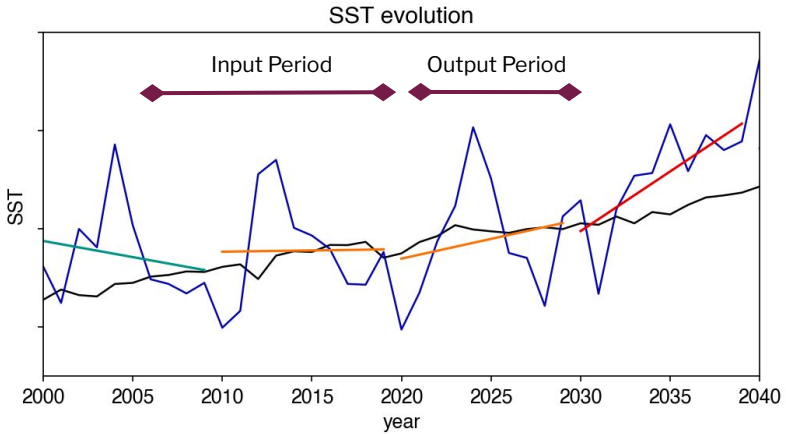
A machine learning approach …

Input full maps of the model **forced response** and **internal variability** averaged over the input period

e.g. in this example, input maps averaged over 2010-2020

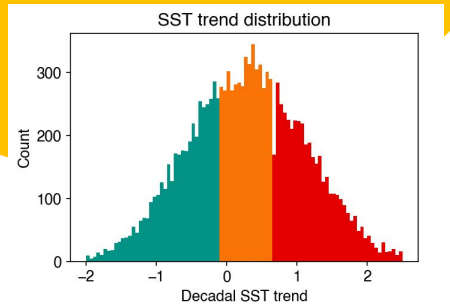






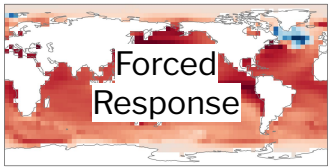
Predicting regional decadal climate trends

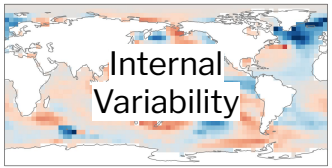
A machine learning approach …

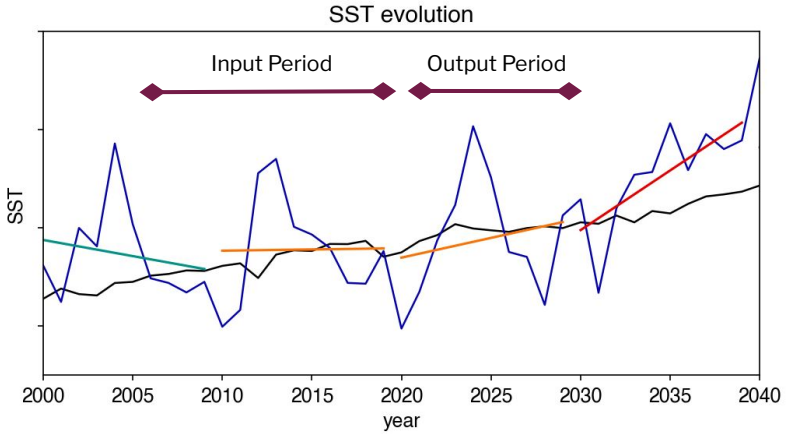


Output a prediction of the future 10 year trend

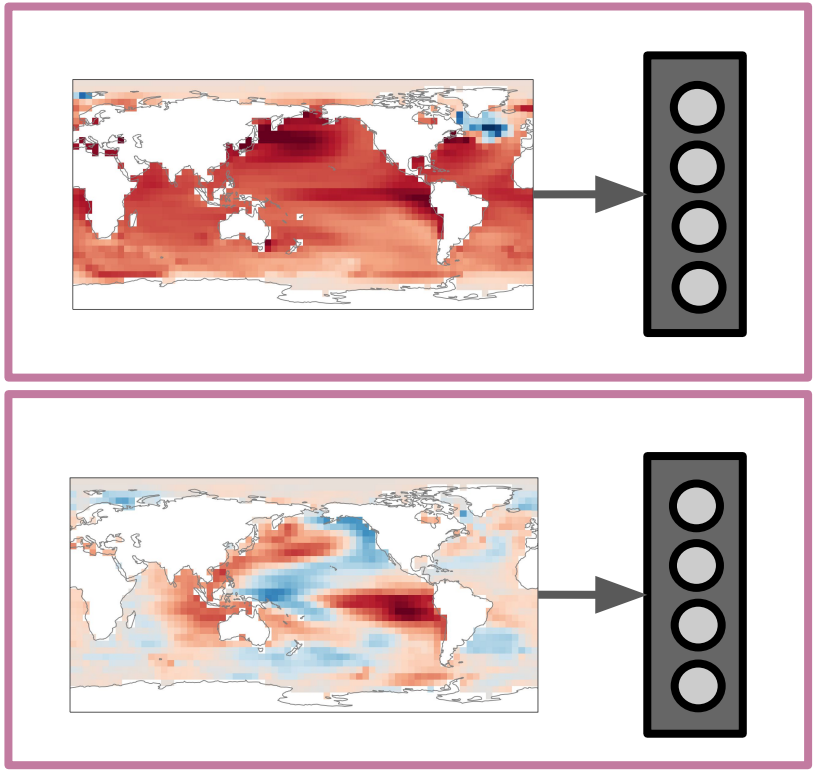
e.g. predict whether the 2020-2030 trend will be in the lower, middle, or upper third





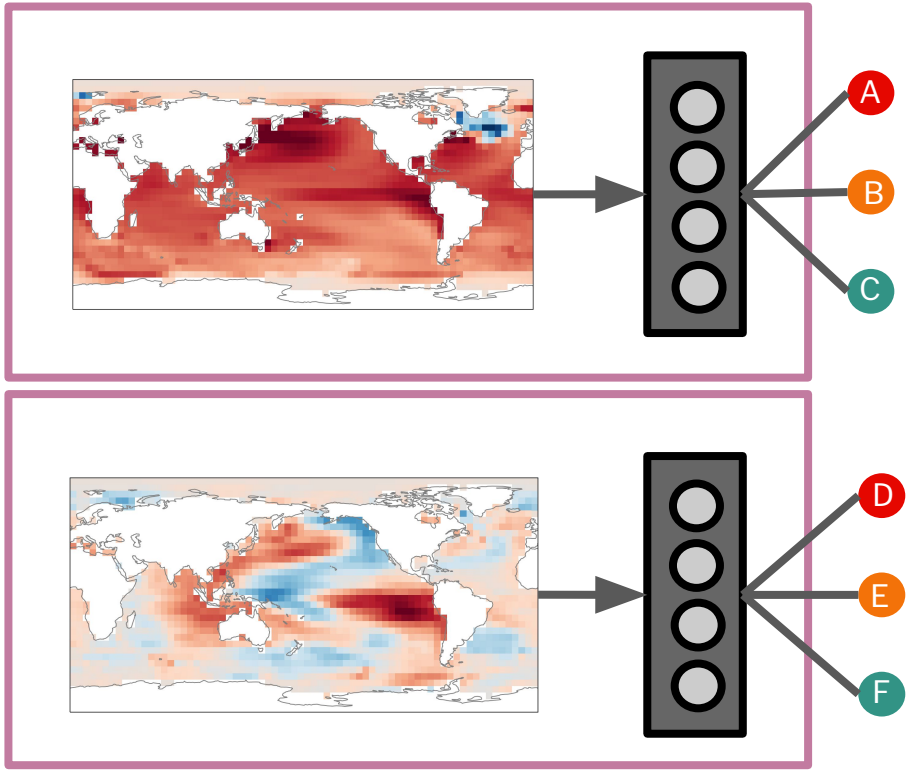


Designing an interpretable neural network



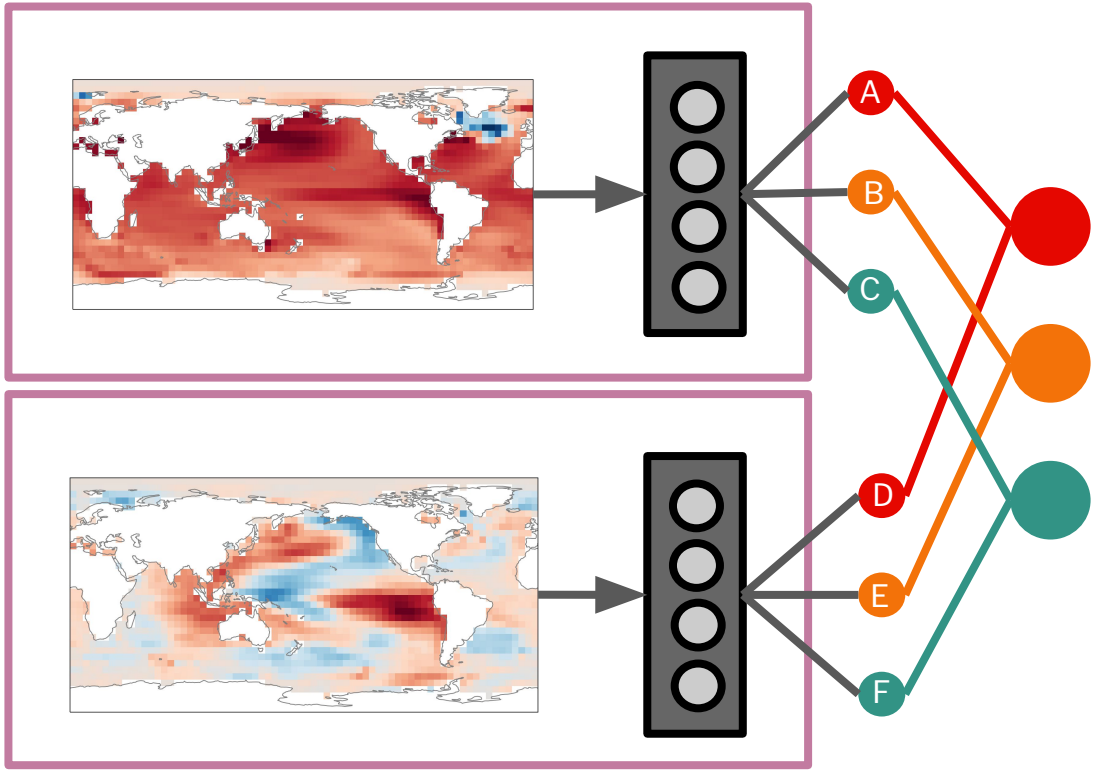
Each map is connected to a separate neural network

Designing an interpretable neural network



Each network outputs its own tercile predictions

Designing an interpretable neural network



Upper Third = A + D

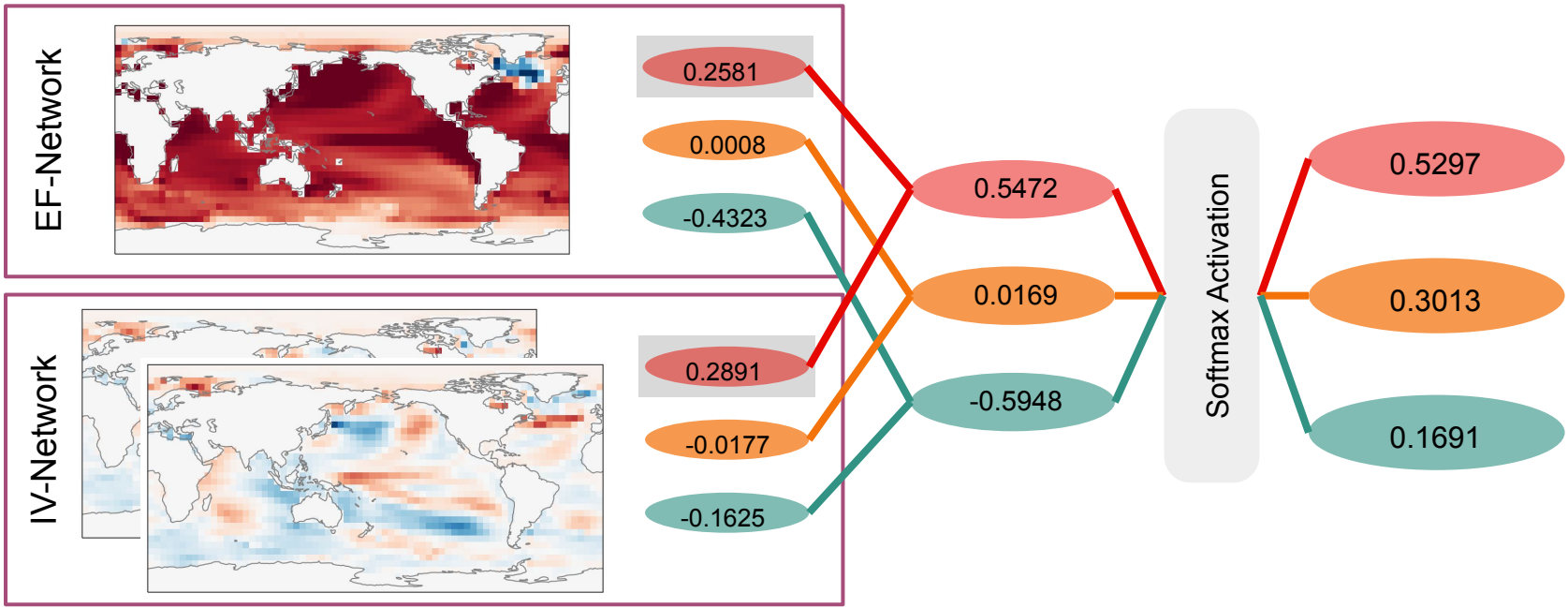
Middle Third = B + E

Lower Third = C + F

Outputs are summed to make the final predictions, highest sum is the ‘winning’ prediction

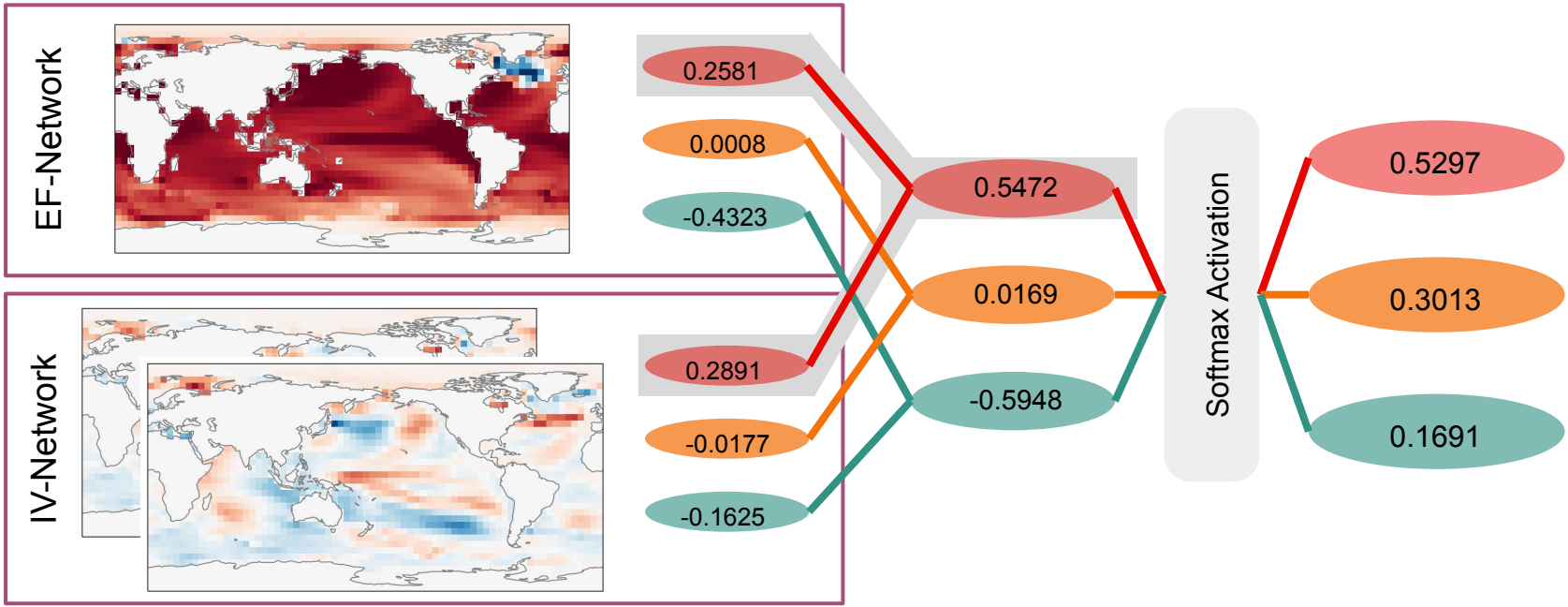
Example Predictions

Predicting 20S-30S, 60W-70W, model year 2045-2054, mem #9



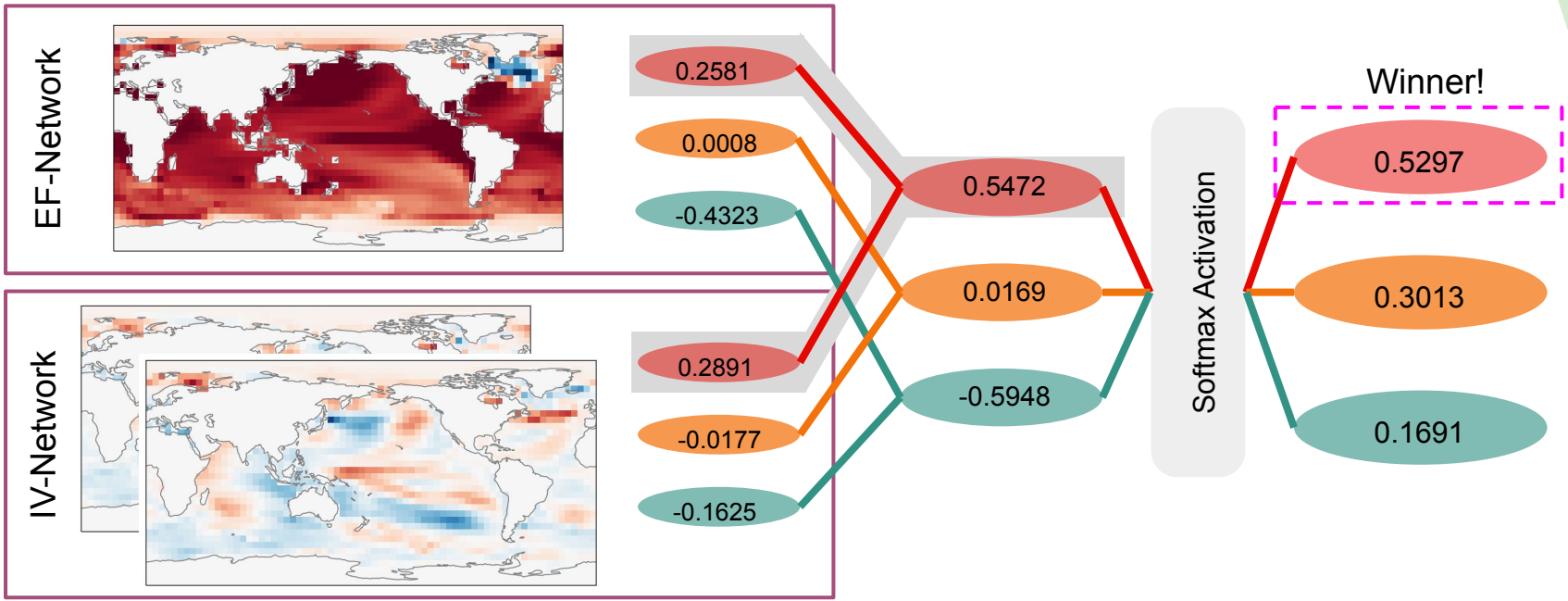
Example Predictions

Predicting 20S-30S, 60W-70W, model year 2045-2054, mem #9



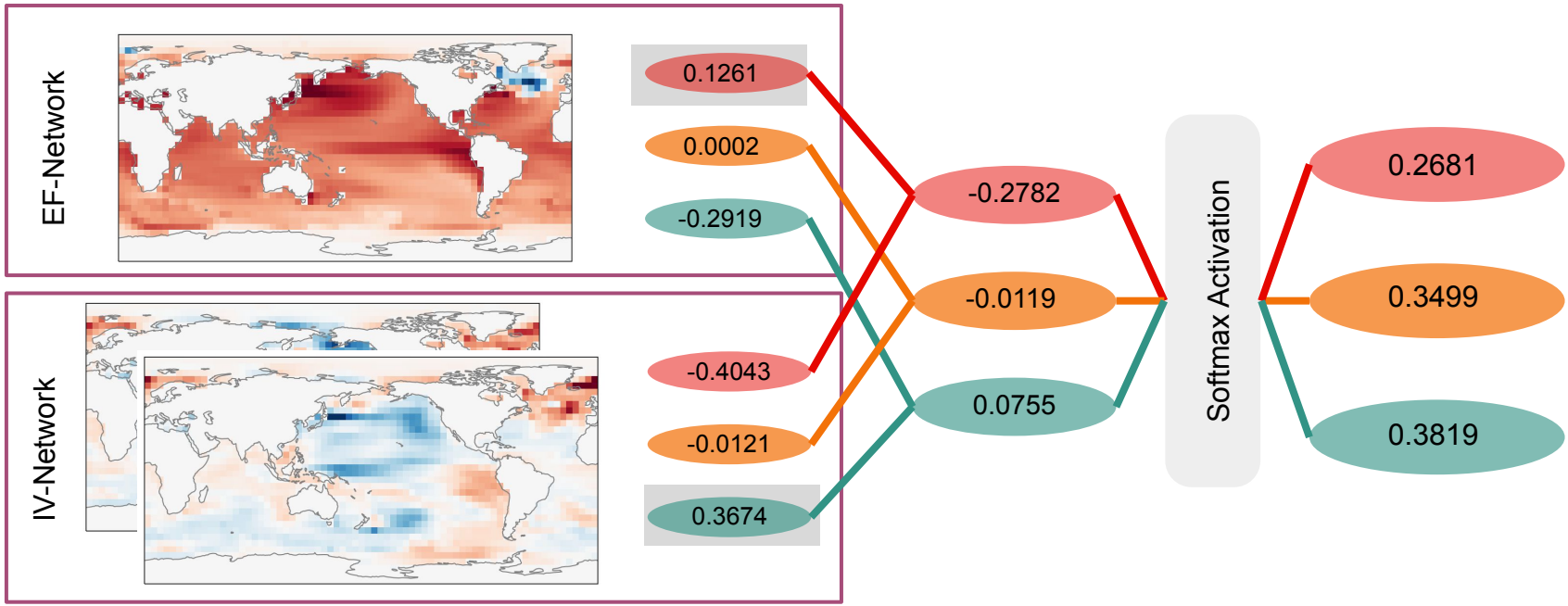
Example Predictions

Predicting 20S-30S, 60W-70W, model year 2045-2054, mem #9



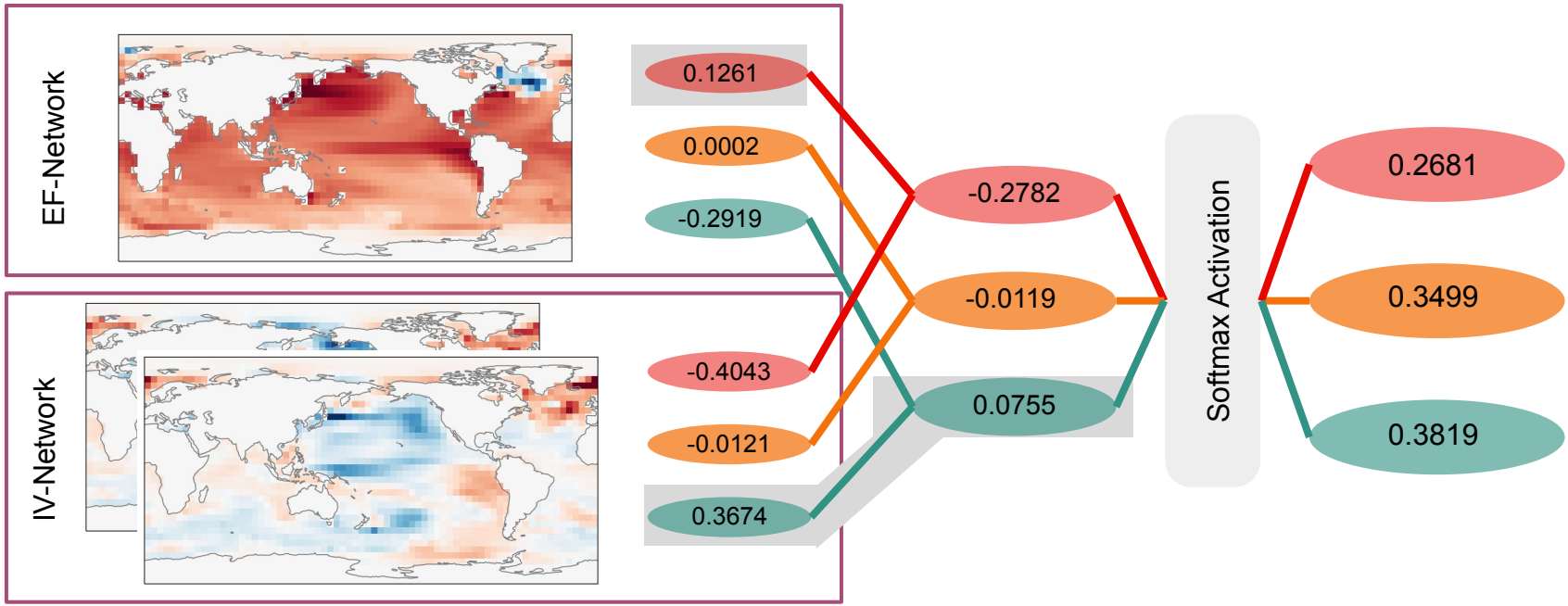
Example Predictions

Predicting 20S-30S, 60W-70W, model year 2033-2042, mem #9



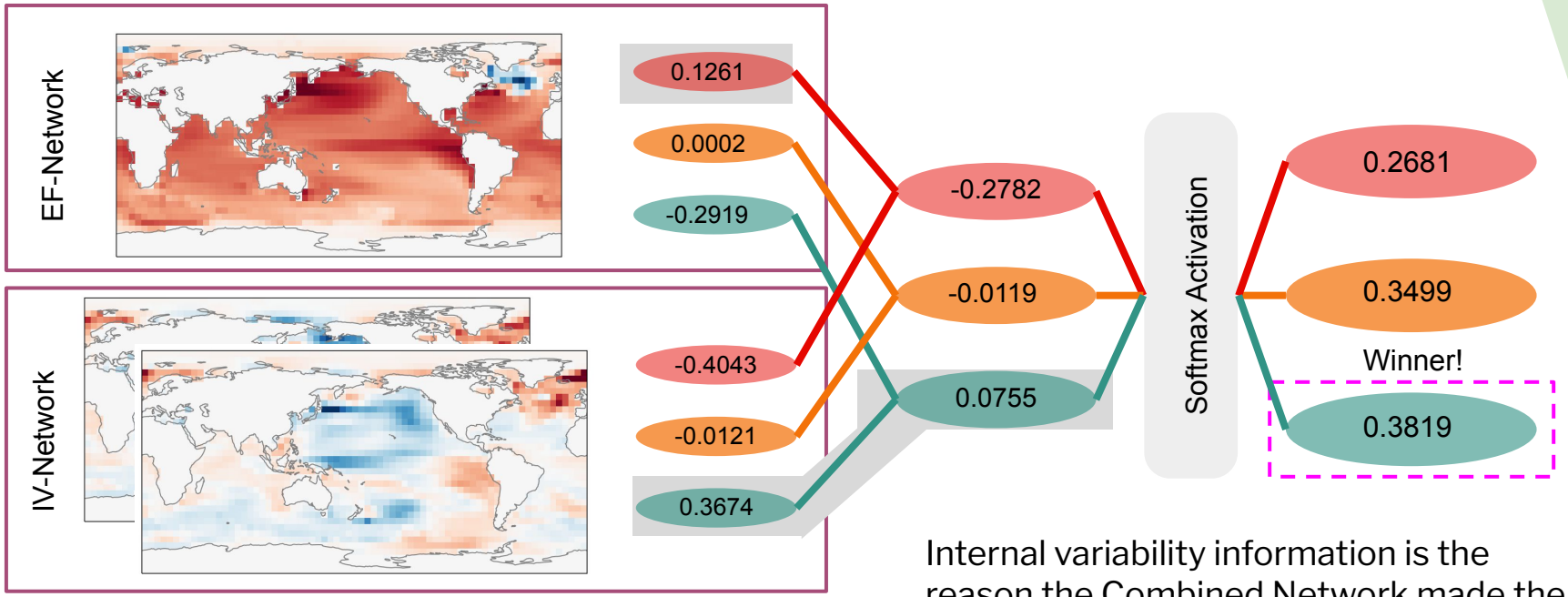
Example Predictions

Predicting 20S-30S, 60W-70W, model year 2033-2042, mem #9



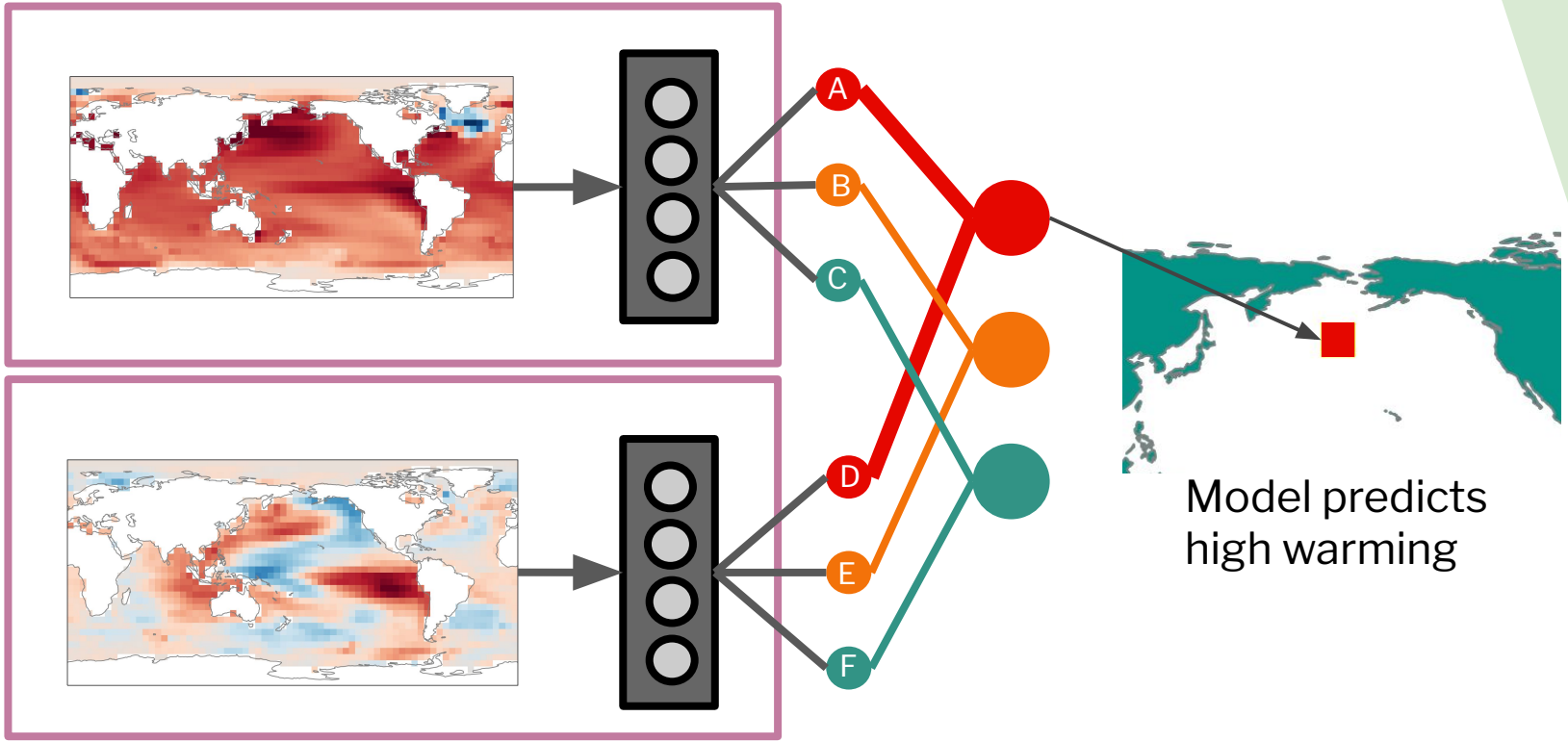
Example Predictions

Predicting 20S-30S, 60W-70W, model year 2033-2042, mem #9

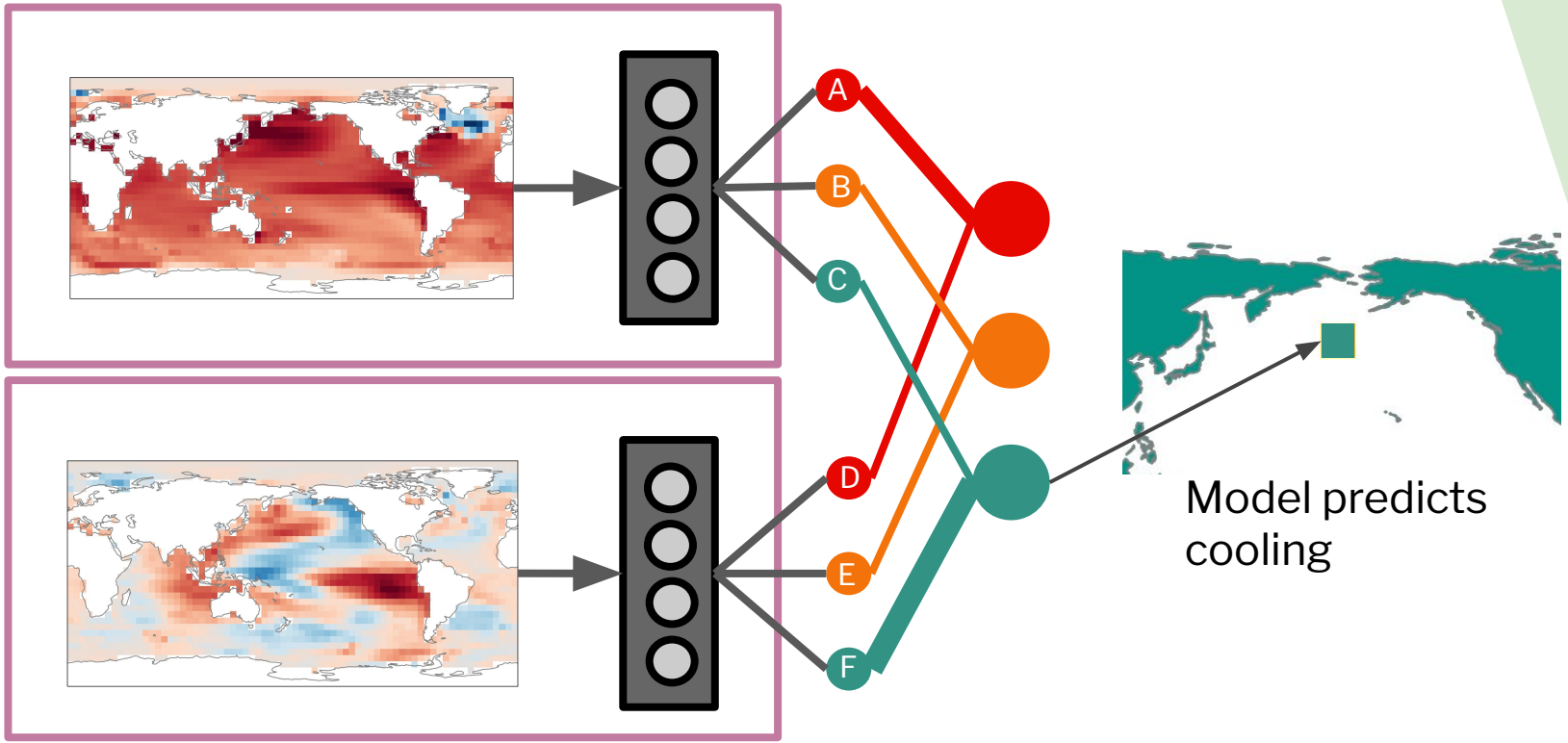


Internal variability information is the reason the Combined Network made the correct prediction!

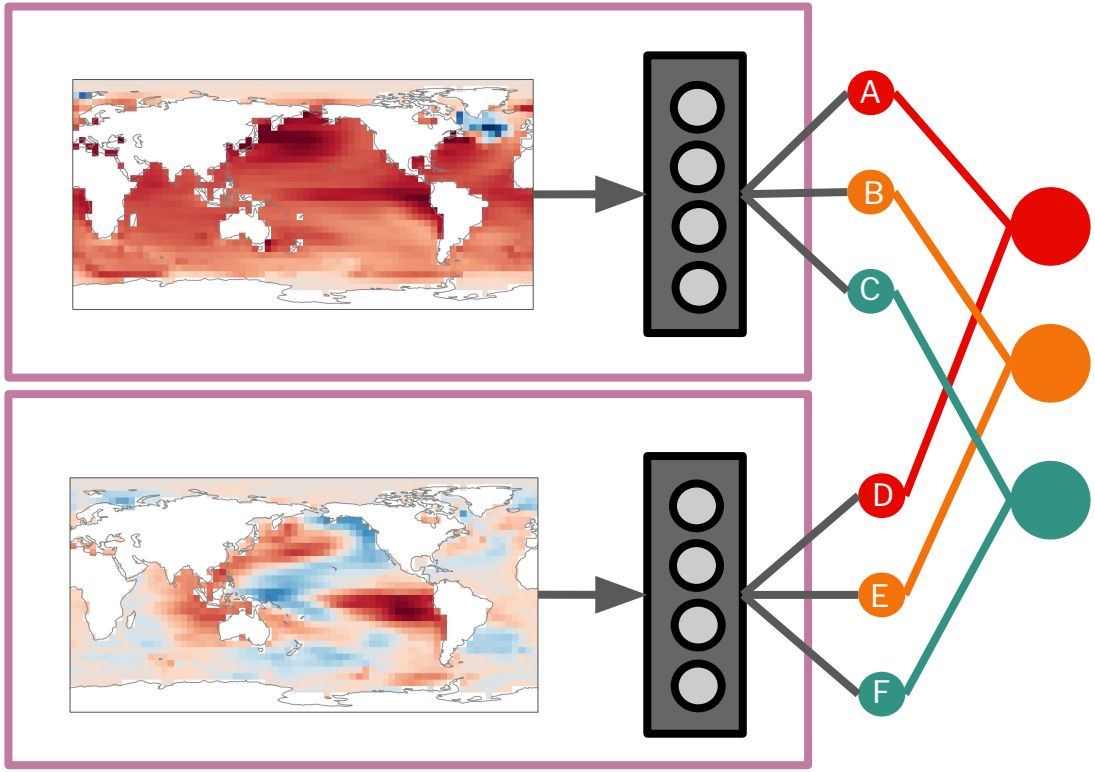
Designing an interpretable neural network



Designing an interpretable neural network

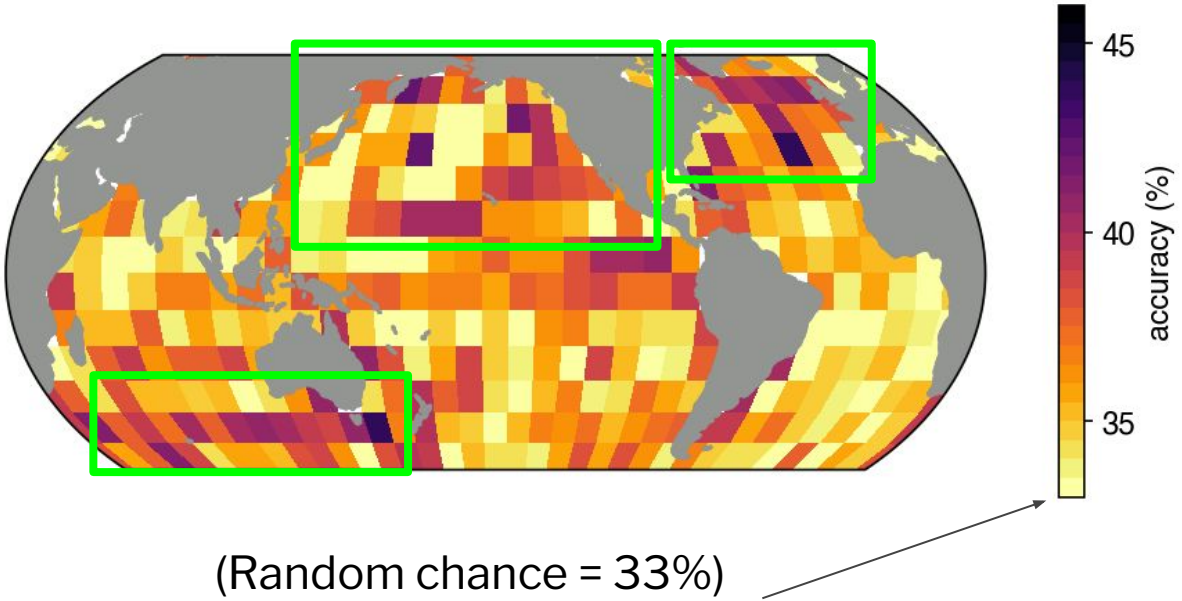


Designing an interpretable neural network



We can calculate the **contribution** of **internal variability** and **external forcing** to a particular prediction

Calculate accuracy over 2020-2050



(Random chance = 33%)

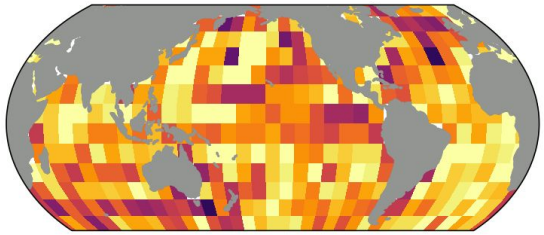
More accurate in North Atlantic, North Pacific and Southern Indian oceans (, Meehl et. al 2017)

✓ Neural network is making sensible predictions

Gordon et. al *in prep.*

Predictability from internal variability vs external forcing over 2020-2050

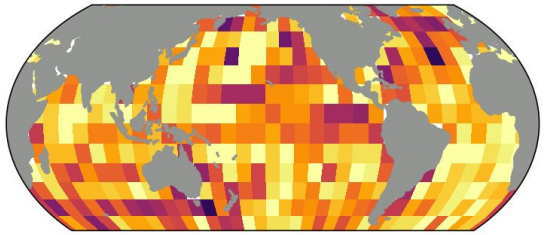
Calculate accuracy at each 10 x 10 location



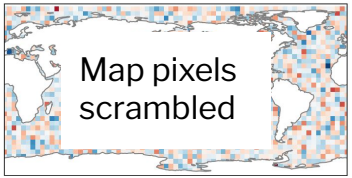
Gordon et. al *in prep.*

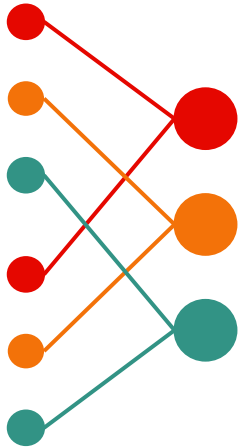
Predictability from internal variability vs external forcing over 2020-2050

Calculate accuracy at each 10 x 10 location



Scramble the pixels of the internal variability (IV) input and input "new" testing into networks

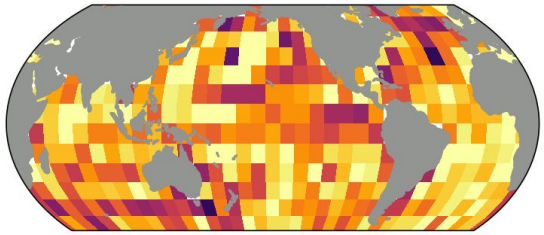




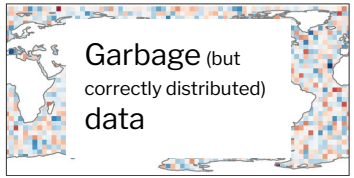
Gordon et. al *in prep.*

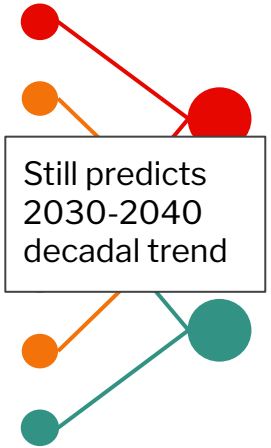
Predictability from internal variability vs external forcing over 2020-2050

Calculate accuracy at each 10 x 10 location



Scramble the pixels of the internal variability (IV) input and input "new" testing into networks

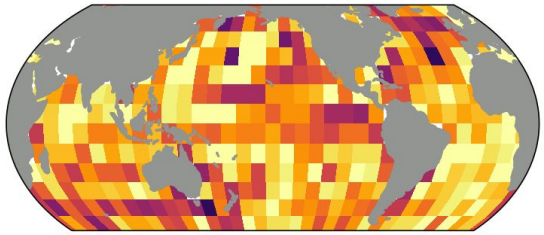




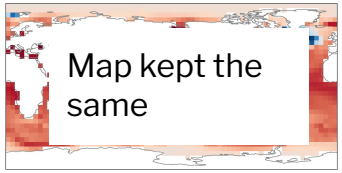
Gordon et. al *in prep.*

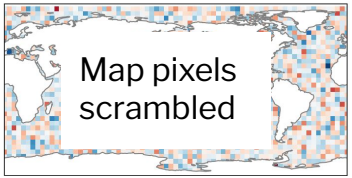
Predictability from internal variability vs external forcing over 2020-2050

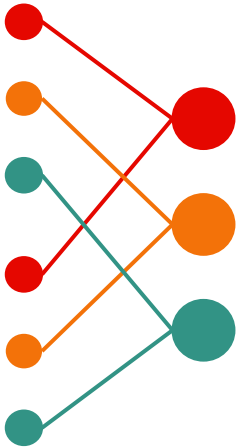
Calculate accuracy at each 10 x 10 location



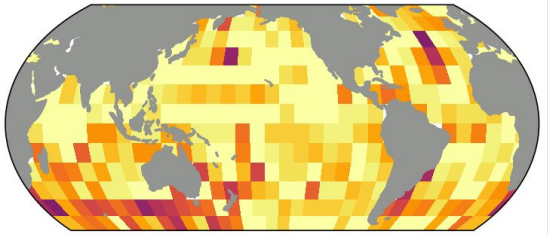
Scramble the pixels of the internal variability (IV) input and input "new" testing into networks





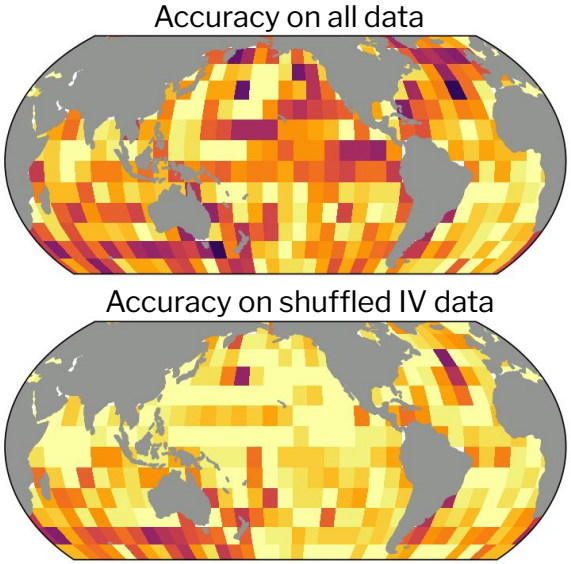


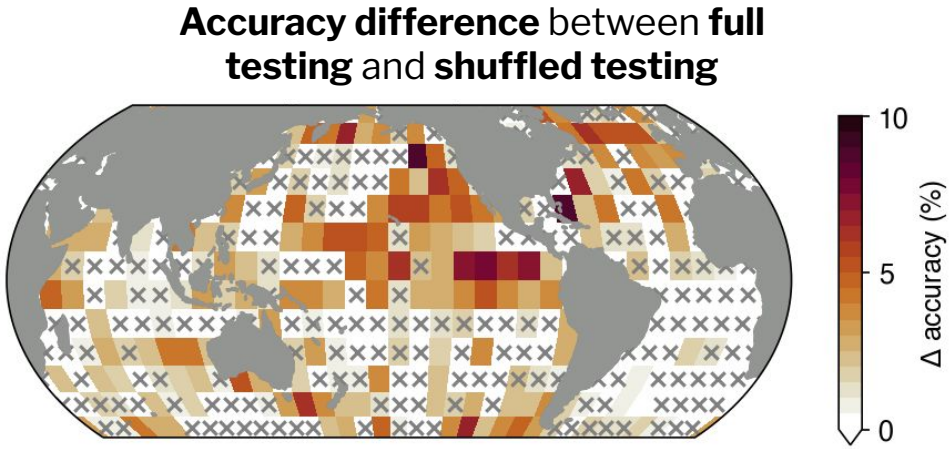
Re-calculate accuracy with scrambled IV data



Gordon et. al *in prep.*

Predictability from internal variability vs external forcing over 2020-2050

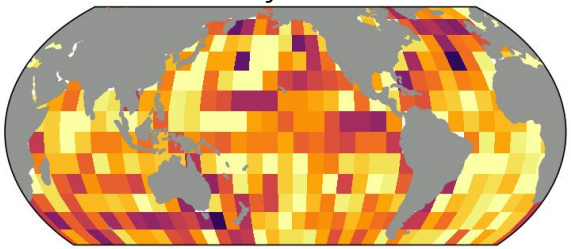




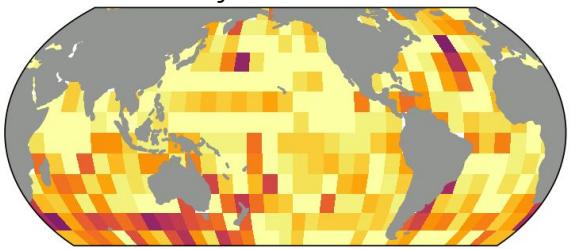
Gordon et. al *in prep.*

Predictability from internal variability vs external forcing over 2020-2050

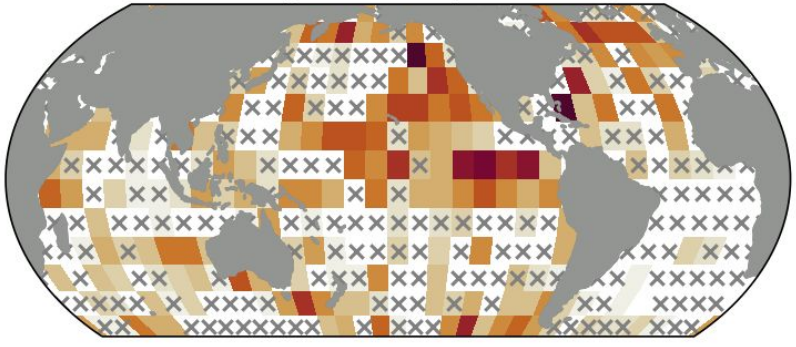
Accuracy on all data



Accuracy on shuffled IV data



Accuracy difference between **full testing** and **shuffled testing**

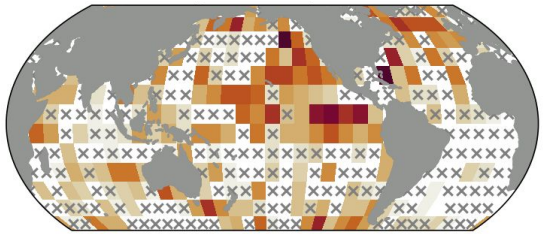


Interpretation: If networks only used the forced response for its predictions, this figure would be zeros everywhere

Key Takeaway: In many regions, **internal variability significantly increase network's accuracy** above external forcing

Gordon et. al *in prep.*

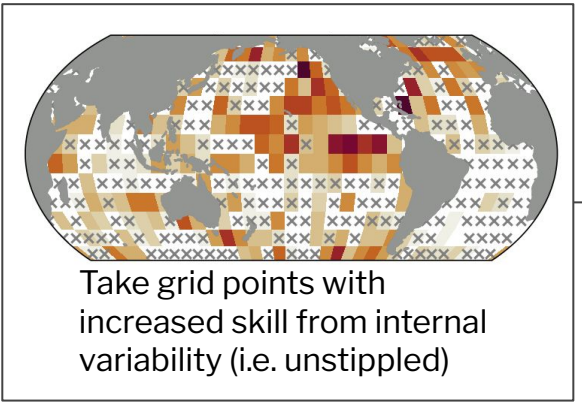
Are some grid points predictable for the same initial state?



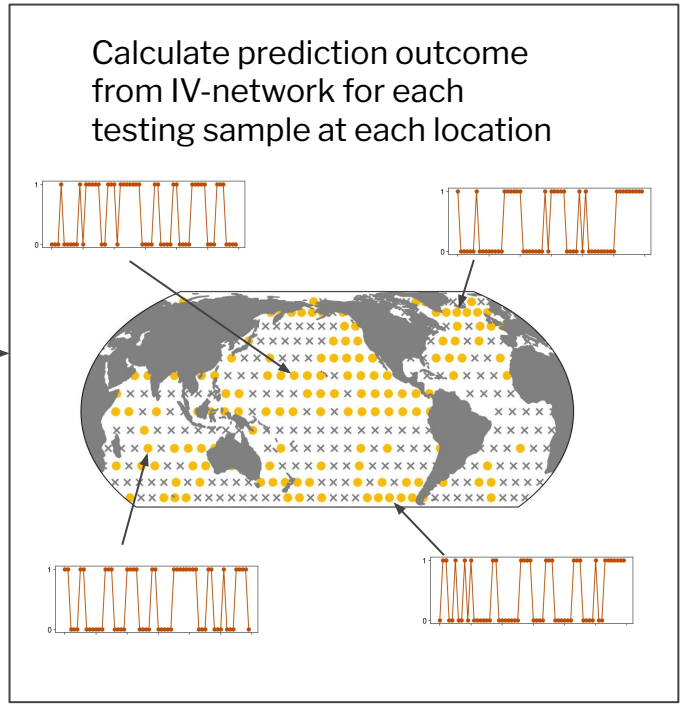
Take grid points with increased skill from internal variability (i.e. unstippled)

Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?

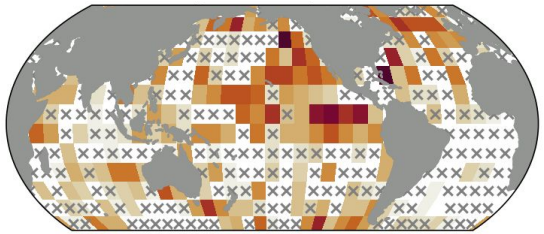


Take grid points with increased skill from internal variability (i.e. unstippled)



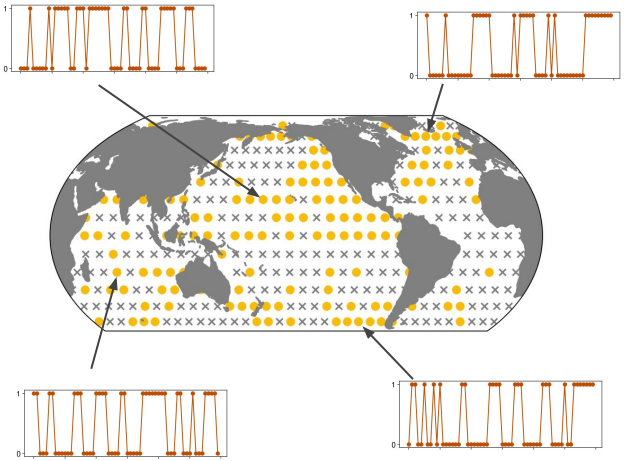
Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?



Take grid points with increased skill from internal variability (i.e. unstippled)

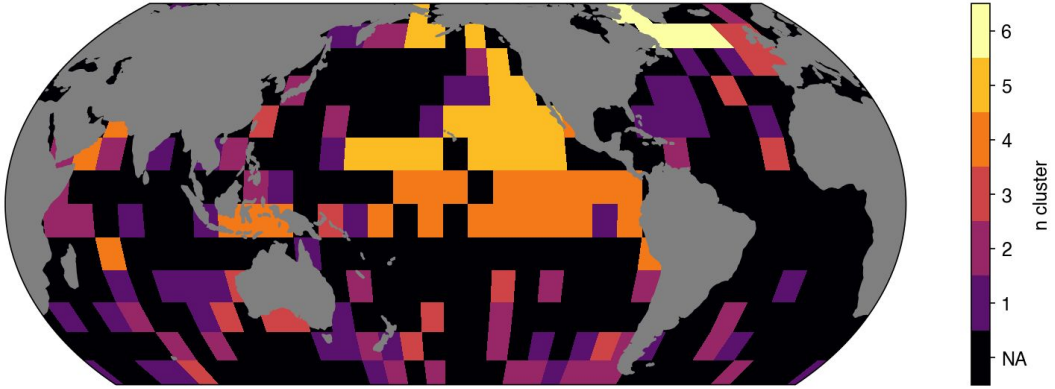
Calculate prediction outcome from IV-network for each testing sample at each location



Kmeans cluster prediction outcome to group grid points where prediction skill covaries with input state

Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?



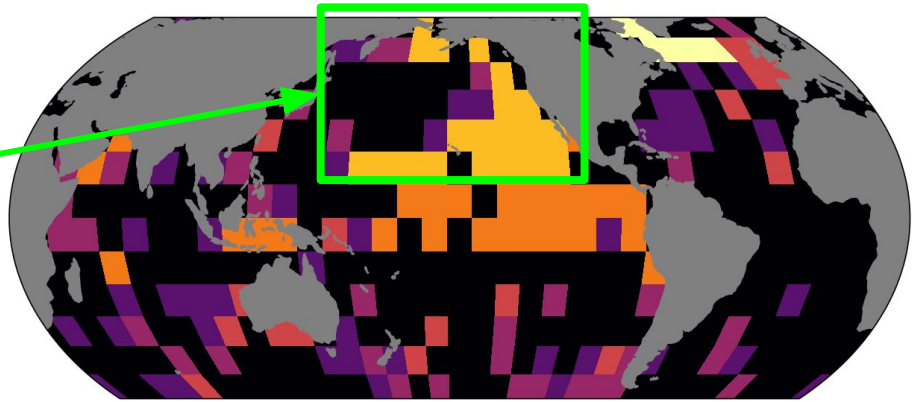
3 obvious spatial regions, 3 maybe not so obvious

Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?

Grid cells resemble PDO horseshoe

Predictability is derived timescale of the PDO

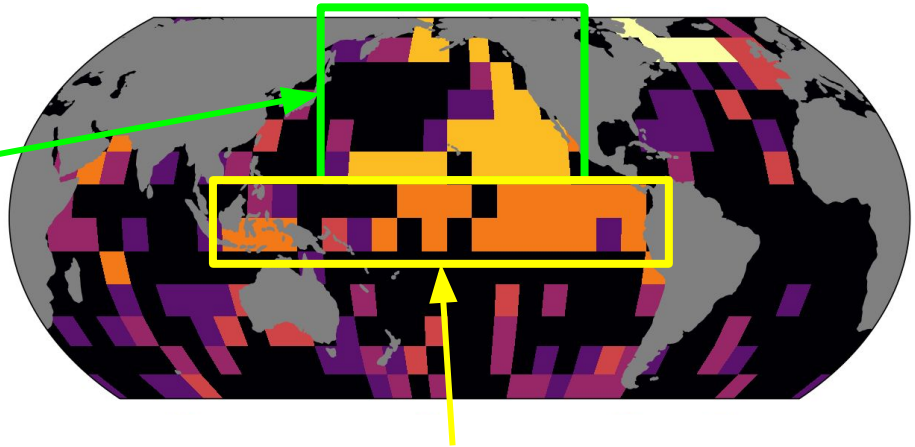


Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?

Grid cells resemble PDO horseshoe

Predictability is derived timescale of the PDO



Tropical Pacific/ENSO region

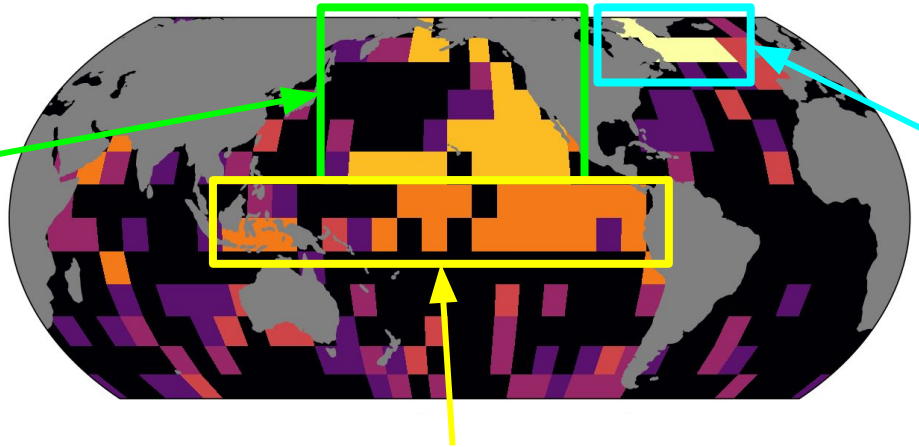
Predictability on decadal timescale likely aliasing of interannual ENSO

Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?

Grid cells resemble PDO horseshoe

Predictability is derived timescale of the PDO



North Atlantic Subpolar Gyre

Predictability is associate with timescale of heat transport into gyre

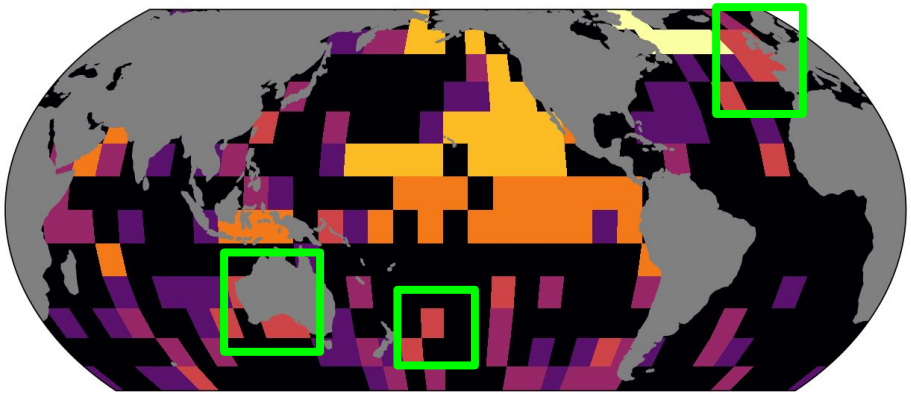
Tropical Pacific/ENSO region

Predictability on decadal timescale likely aliasing of interannual ENSO

Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?

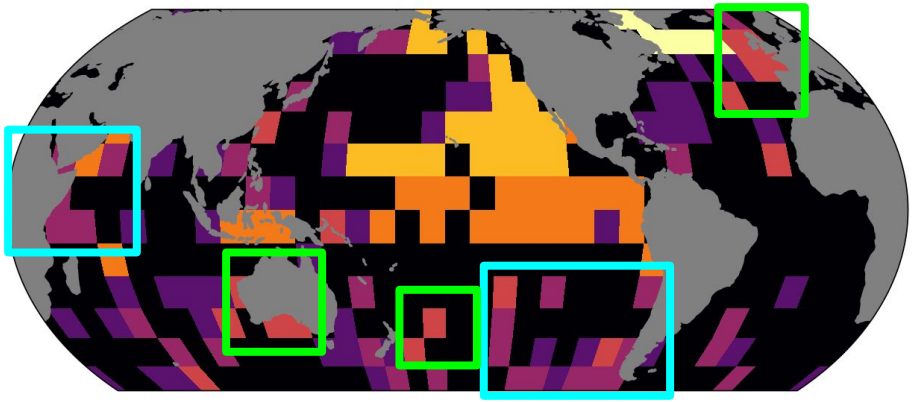
Northeast Atlantic + Australia/New Zealand?



Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?

Northeast Atlantic + Australia/New Zealand?

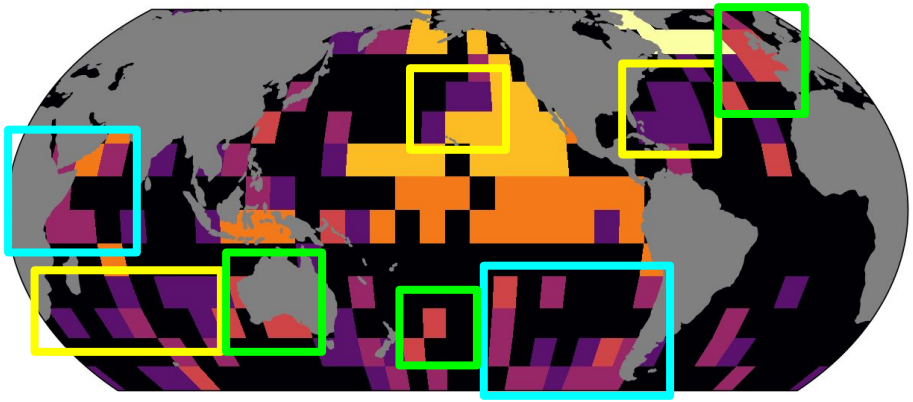


South Pacific + Coastal Indian Ocean?

Gordon et. al *in prep.*

Are some grid points predictable for the same initial state?

Northeast Atlantic + Australia/New Zealand?



Subtropical Atlantic + Subtropical North Pacific + Subtropical Southern Indian?

South Pacific + Coastal Indian Ocean?

Gordon et. al *in prep.*

Takeaways

- Decadal SST trends can be predictable in the CESM2-LE in the near term (2020-2050)
- Internal variability contributes to predictability of SST trends
- Predictability is associated with decadal scale processes in the North Atlantic and North Pacific Oceans

Contact me!

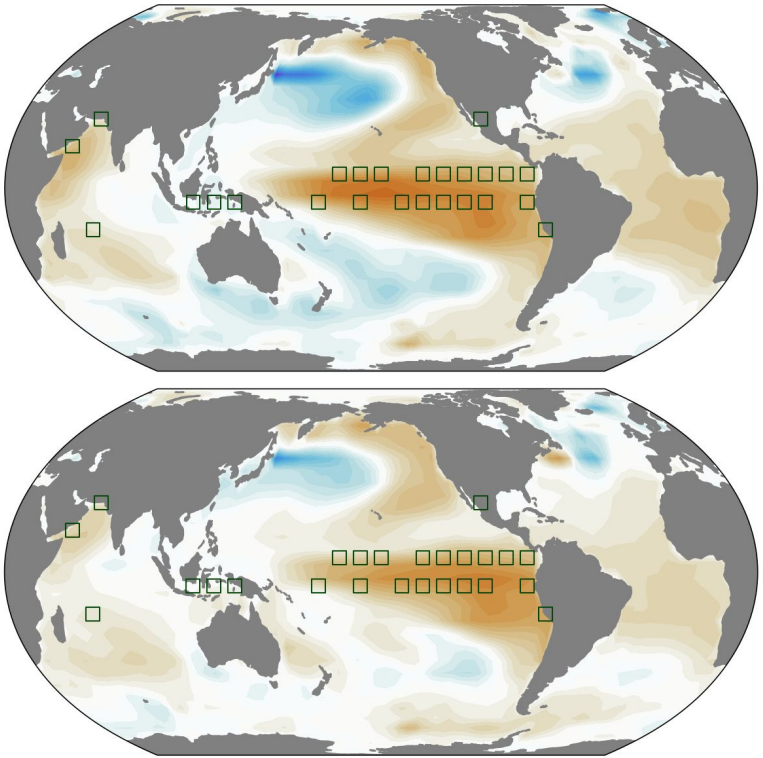
emily.gordon@colostate.edu

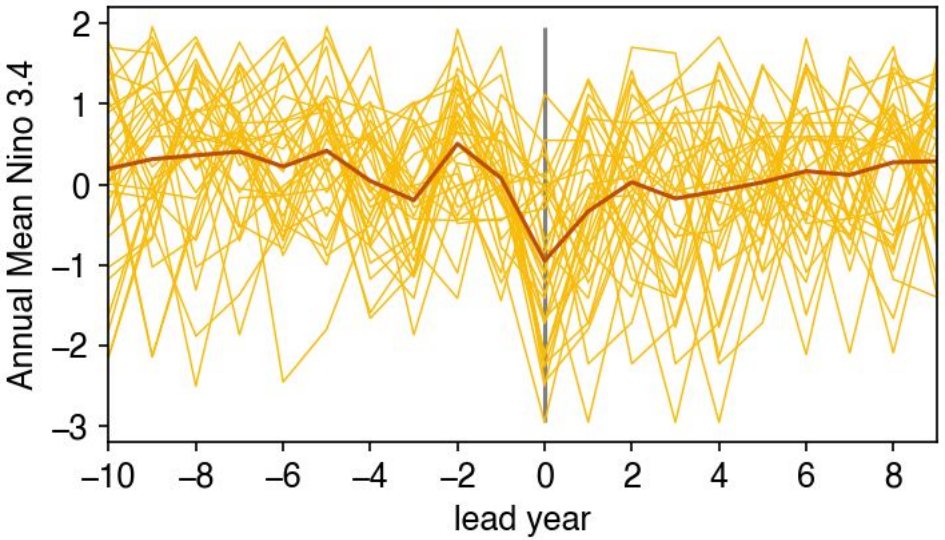
emily-gordy

FULBRIGHT New Zealand
TE TŪĀPAPA MĀTAURANGA O AOTEAROA ME AMERIKA

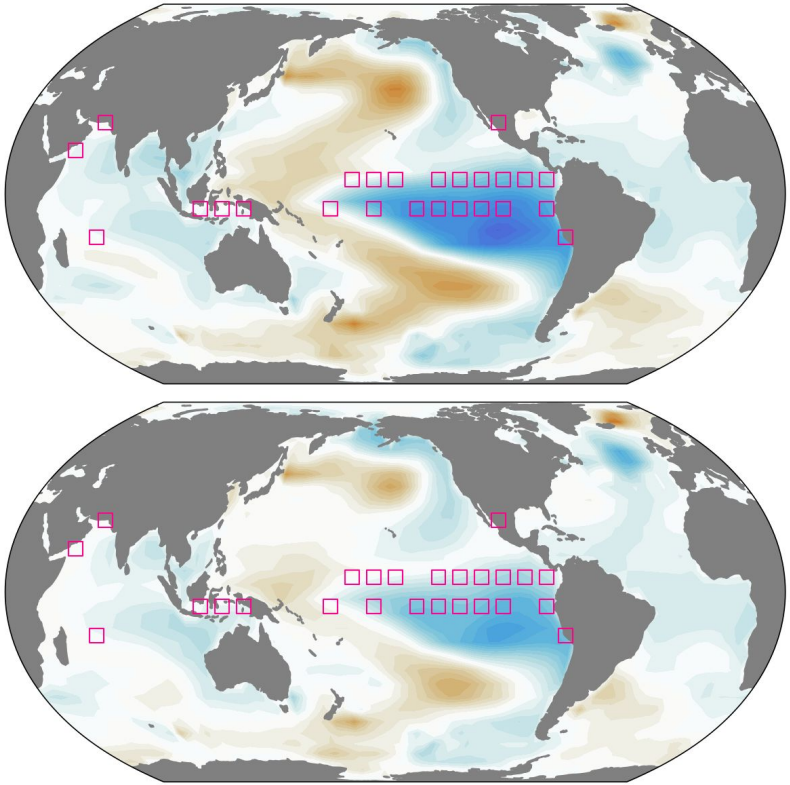
Fin

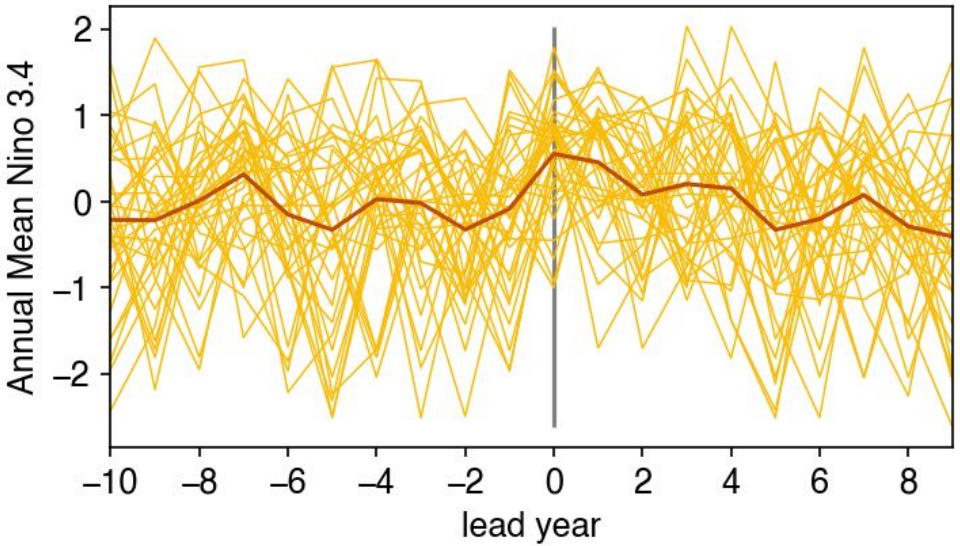
Tropical Pacific, predict positive SST trend



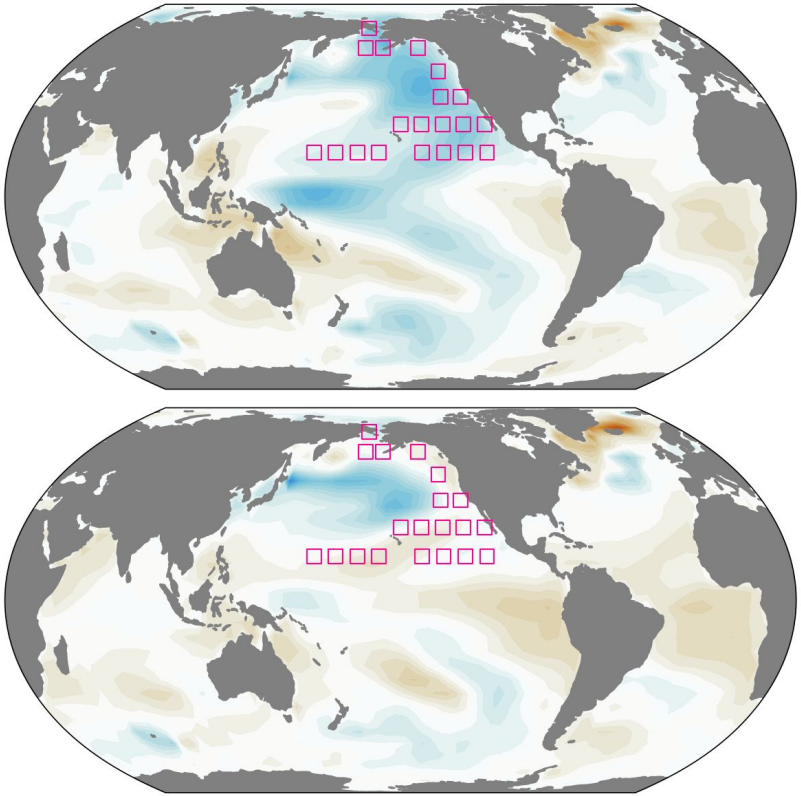


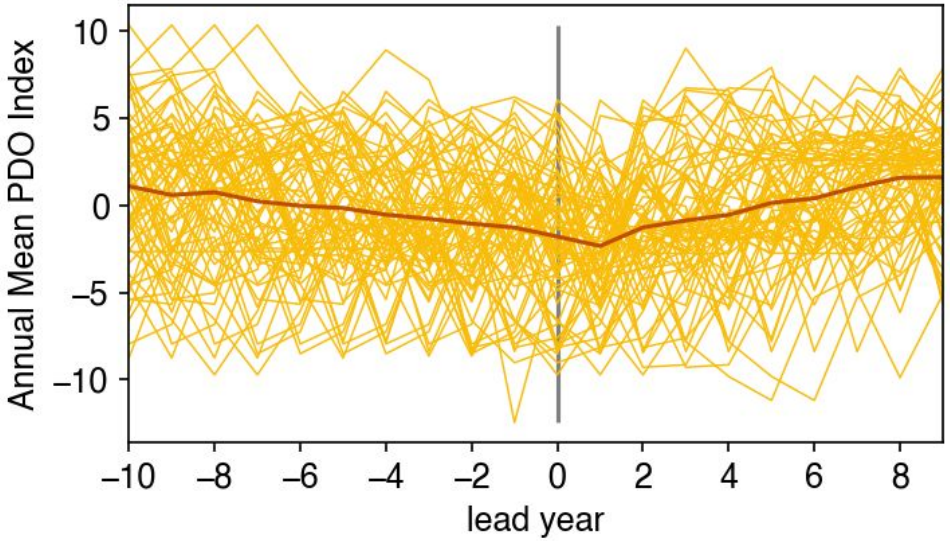
Tropical Pacific, predict negative SST trend



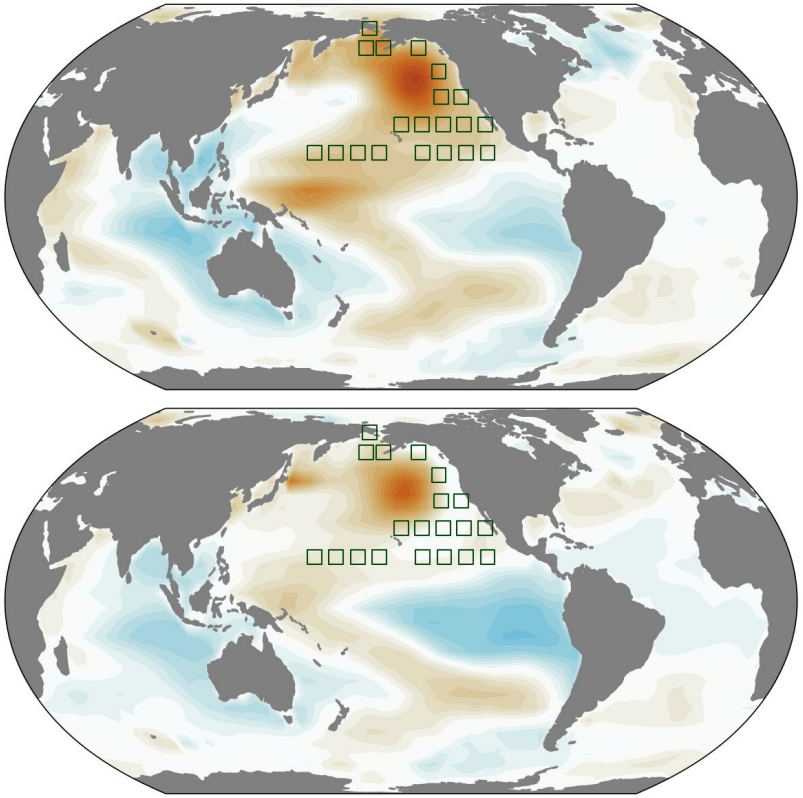


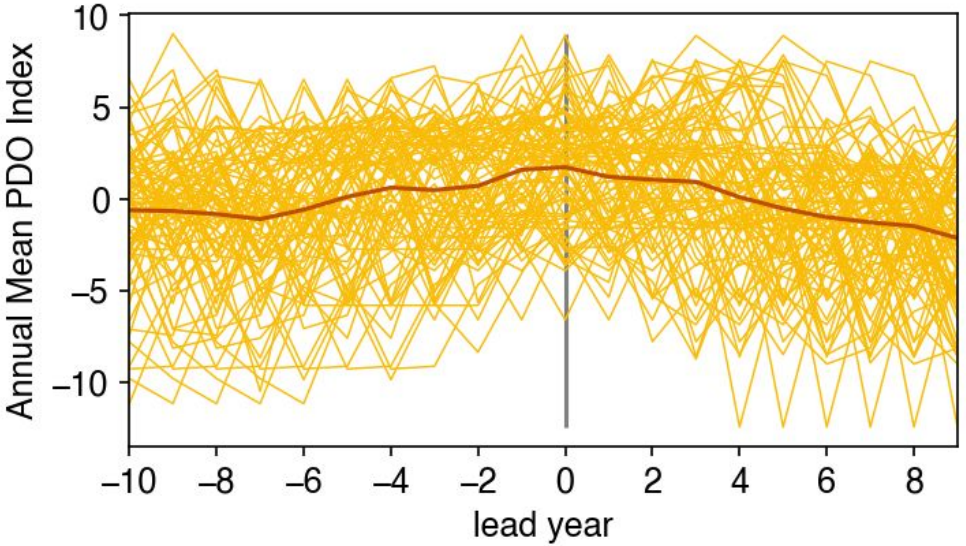
North Pacific, predict negative SST trend



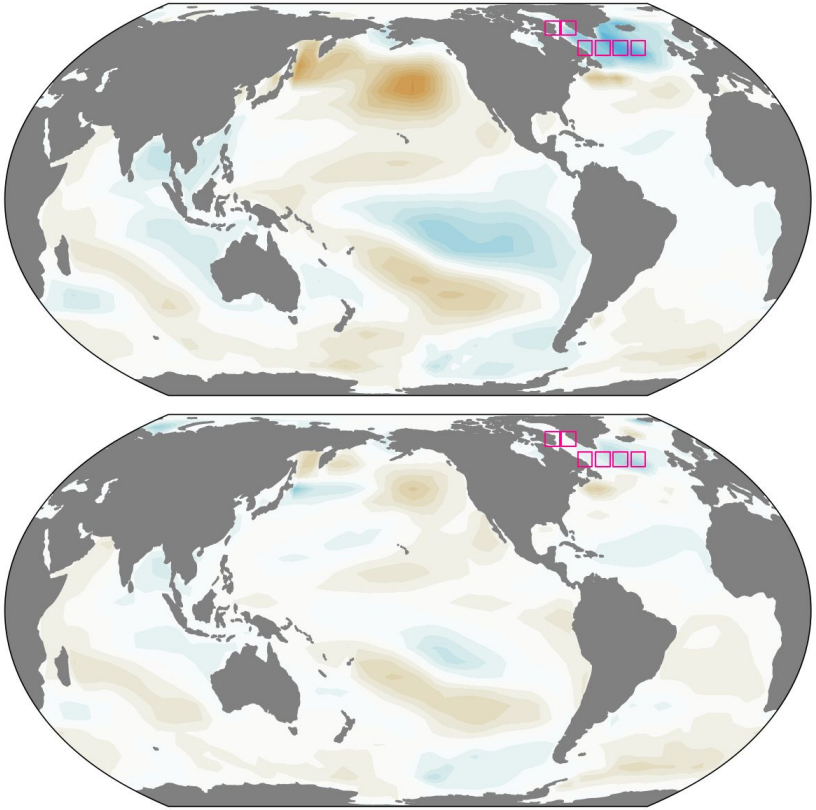


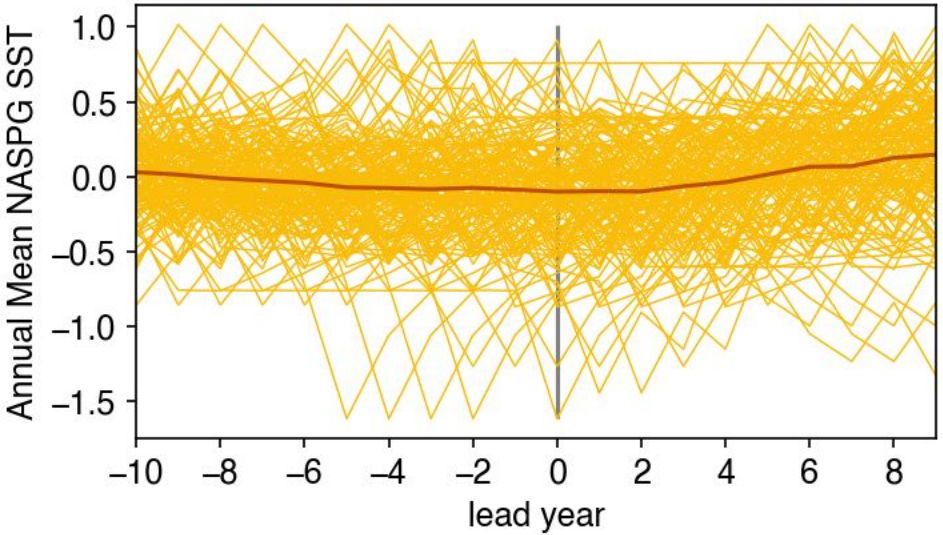
North Pacific, predict positive SST trend





North Atlantic, predict negative SST trend





North Atlantic, predict positive SST trend