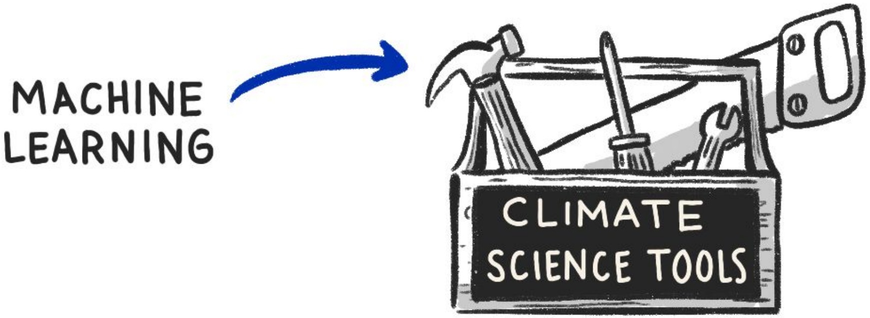
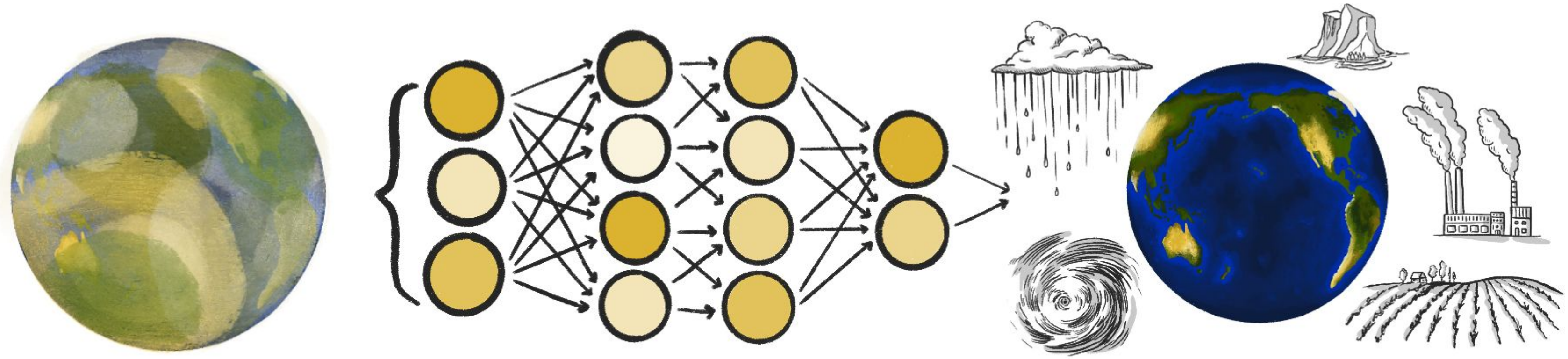
Explainable AI for Climate Projection, Multi-Year Prediction & Discovery

Dr. Elizabeth A. Barnes
Professor, Dept. of Atmospheric Science
Colorado State University

ATMOSPHERIC SCIENCE
COLORADO STATE UNIVERSITY

Ocean Model Development, Data-driven Parameterizations, and Machine Learning in Ocean Models of the Earth System Workshop
Sep 09 2024


MACHINE
LEARNING
CLIMATE
SCIENCE TOOLS



Do it faster/cheaper

Do it better

Learn something new

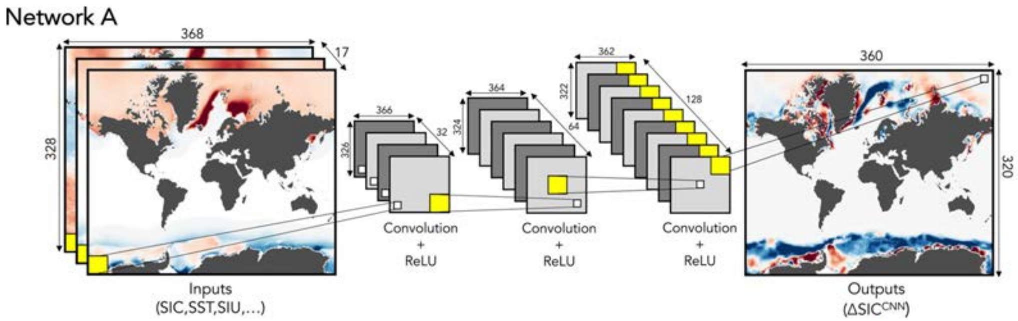
Machine Learning: a tool with many uses.

1

ML for post-processing data [e.g. data compression, data analysis]

Predicting the Errors of Forecast Systems

e.g. Chapman et al. (2019), Cahill et al. (in review), Pan et al. (2021), *Gregory et al. (2023)*



Machine Learning: a tool with many uses.

1

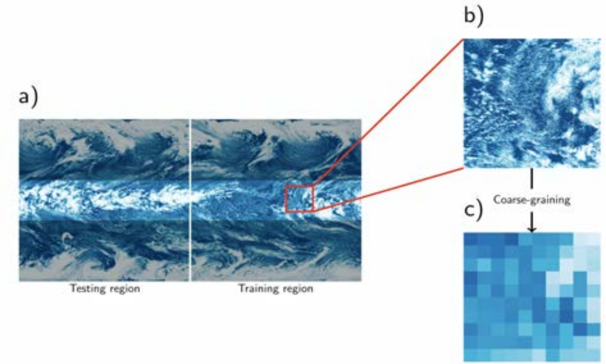
ML for post-processing data
[e.g. data compression, data analysis]

2

ML to improve climate models
[e.g. parameterizations]

Improved Model Parameterizations

e.g. Rasp et al. (2018; PNAS); Schneider et al. (2017; GRL); O'Gorman and Dwyer (2018); Beucler et al. (2020; PRL); Dagon et al. (2020); *Brenowitz and Bretherton (2018)*



Machine Learning: a tool with many uses.

1

ML for post-processing data
[e.g. data compression, data analysis]

2

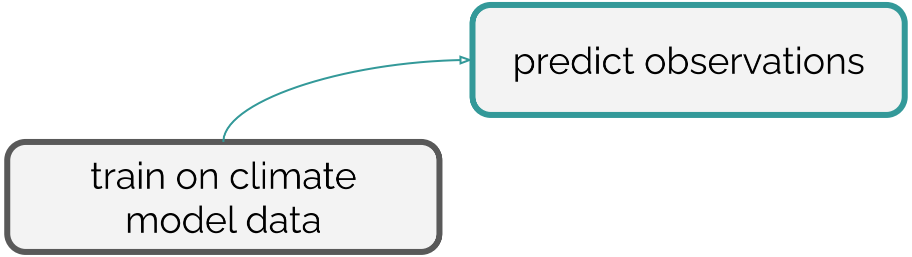
ML to improve climate models
[e.g. parameterizations]

3

Combine disparate datastreams to
explore complex systems

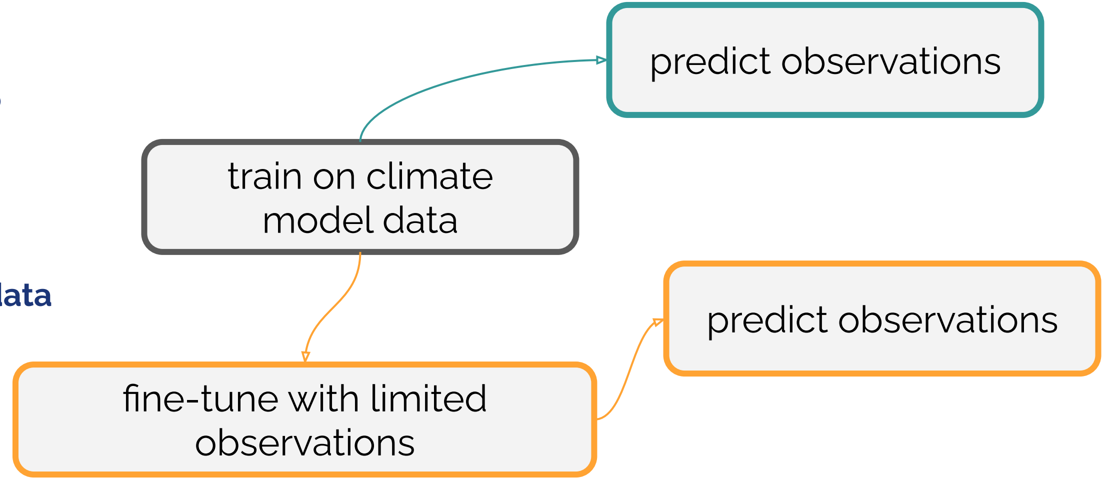
Machine Learning: a tool with many uses.

1. ML for post-processing data
[e.g. data compression, data analysis]
2. ML to improve climate models
[e.g. parameterizations]
3. Combine disparate datastreams to explore complex systems
4. Merging observations and model data



Machine Learning: a tool with many uses.

1. **ML for post-processing data [e.g. data compression, data analysis]**
2. **ML to improve climate models [e.g. parameterizations]**
3. **Combine disparate datastreams to explore complex systems**
4. **Merging observations and model data**



Machine Learning: a tool with many uses.

1

ML for post-processing data
[e.g. data compression, data analysis]

2

ML to improve climate models
[e.g. parameterizations]

3

Combine disparate datastreams to
explore complex systems

4

Merging observations and model data

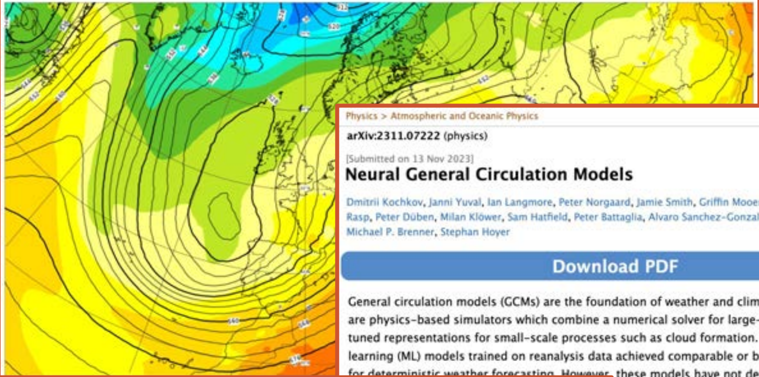
5

Deep-learning weather + climate
emulators

Machine Learning: a tool with many uses.

Experimental: AIFS (ECMWF) ML model: 500 hPa
geopotential height and 850 hPa temperature

Base time: Tue 26 Mar 2024 06 UTC Valid time: Sun 31 Mar 2024 00 UTC (+114h) Area : Europe



Physics > Atmospheric and Oceanic Physics

arXiv:2311.07222 (physics)

[Submitted on 13 Nov 2023]

Neural General Circulation Models

Dmitrii Kochkov, Janni Yuval, Ian Langmore, Peter Norgaard, Jamie Smith, Griffin Mooers, James Lottes, Stephan Rasp, Peter Düben, Milan Klöwer, Sam Hatfield, Peter Battaglia, Alvaro Sanchez-Gonzalez, Matthew Willson, Michael P. Brenner, Stephan Hoyer

Download PDF

General circulation models (GCMs) are the foundation of weather and climate prediction. GCMs are physics-based simulators which combine a numerical solver for large-scale dynamics with tuned representations for small-scale processes such as cloud formation. Recently, machine learning (ML) models trained on reanalysis data achieved comparable or better skill than GCMs for deterministic weather forecasting. However, these models have not demonstrated improved ... for long-term weather and climate ... combines a differentiable solver for ... d show that it can generate forecasts of ... imate on par with the best ML and physics- ... ML models for 1-10 day forecasts, and with the ... recasts ensemble prediction for 1-15 day ... ure, NeuralGCM can accurately track climate ... ultiple decades, and climate forecasts with 140

Physics > Atmospheric and Oceanic Physics

arXiv:2310.02074 (physics)

[Submitted on 3 Oct 2023]

ACE: A fast, skillful learned global atmospheric model for climate prediction

Oliver Watt-Meyer, Gideon Dresdner, Jeremy McGibbon, Spencer K. Clark, Brian Henn, James Duncan, Noah D. Brenowitz, Karthik Kashinath, Michael S. Pritchard, Boris Bonev, Matthew E. Peters, Chris...

Download PDF

Existing ML-based atmospheric models are not suitable for climate predict... long-term stability and physical consistency. We present ACE (AI2 Climate ... parameter, autoregressive machine learning emulator of an existing compr... resolution global atmospheric model. The formulation of ACE allows evalua... such as the conservation of mass and moisture. The emulator is stable for ... conserves column moisture without explicit constraints and faithfully repro... model's climate, outperforming a challenging baseline on over 80% of track... requires nearly 100x less wall clock time and is 100x more energy efficient... model using typically available resources.

Physics > Atmospheric and Oceanic Physics

[Submitted on 22 Mar 2024]

An ensemble of data-driven weather prediction models for operational sub-seasonal forecasting

Jonathan A. Weyn, Divya Kumar, Jeremy Berman, Najeeb Kazmi, Sylwester Klocek, Pete Luferenko, Kit Thambiratnam

We present an operations-ready multi-model ensemble weather forecasting system which uses hybrid data-driven weather prediction models coupled with the European Centre for Medium-range Weather Forecasts (ECMWF) ocean model to predict global weather at 1-degree resolution for 4 weeks of lead time. For predictions of 2-meter temperature, our ensemble on average outperforms the raw ECMWF extended-range ensemble by 4-17%, depending on the lead time. However, after applying statistical bias corrections, the ECMWF ensemble is about 3% better at 4 weeks. For other surface parameters, our ensemble is also within a few percentage points of ECMWF's ensemble. We demonstrate that it is possible to achieve near-state-of-the-art subseasonal-to-seasonal forecasts using a multi-model ensembling approach with data-driven weather prediction models.

Subjects: Atmospheric and Oceanic Physics (physics.ao-ph); Machine Learning (cs.LG)
Cite as: arXiv:2403.15598 [physics.ao-ph]
(or arXiv:2403.15598v1 [physics.ao-ph] for this version)

1 ML for post-processing data
[e.g. data compression, data analysis]

2 ML to improve climate models
[e.g. parameterizations]

3 Combine disparate datastreams to explore complex systems

4 Merging observations and model data

5 Deep-learning weather + climate emulators

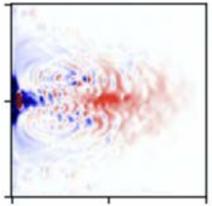
6 Climate change communication

Mila

Machine Learning: a tool with many uses.

ILLUSTRATED BY HEARTWOOD VISUALS

1. ML for post-processing data [e.g. data compression, data analysis]
2. ML to improve climate models [e.g. parameterizations]
3. Combine disparate datastreams to explore complex systems
4. Merging observations and model data
5. Deep-learning weather + climate emulators
6. Climate change communication
7. Learn new things!



$$\hat{\mathbf{S}}_{\mathbf{u}}^{BT} \approx \kappa_{BT}\overline{\nabla}\cdot\begin{pmatrix}\zeta^2-\zeta D & \zeta\tilde{D}\\ \zeta\tilde{D} & \zeta^2+\zeta D\end{pmatrix}$$

Equation Discovery

e.g. *Zanna & Bolton (2020)*

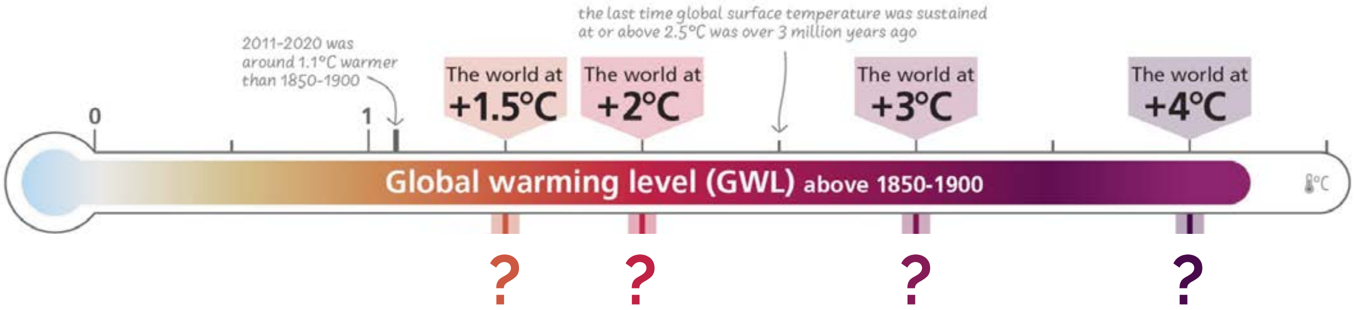
Machine Learning: a tool with many uses.

OUR GOAL:
To develop and implement AI tools to leverage imperfect climate models in support of earth system prediction across time and space.

Climate models provide inaccurate, but invaluable "parallel universes" to mine for information

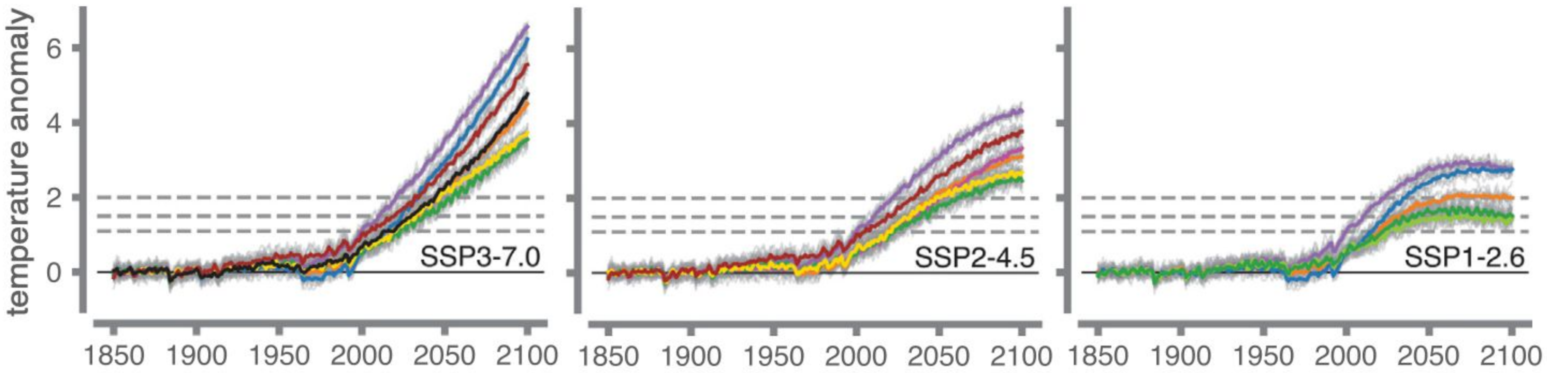
https://futurism.com/parallel-universes-many-worlds-theory

Time Remaining Until Critical Warming Thresholds are Reached



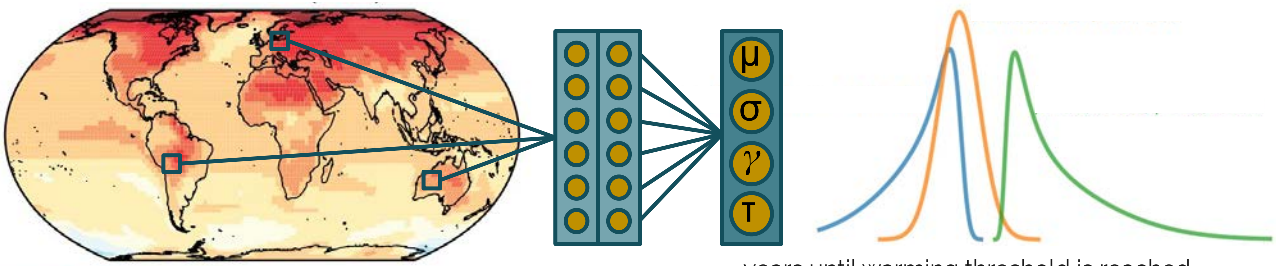
adapted from IPCC AR6 (2023); Summary for Policymakers

Time Remaining Until Critical Warming Thresholds are Reached



Diffenbaugh & Barnes (2023)

Trained on annual maps from 10 realizations from across multiple climate models

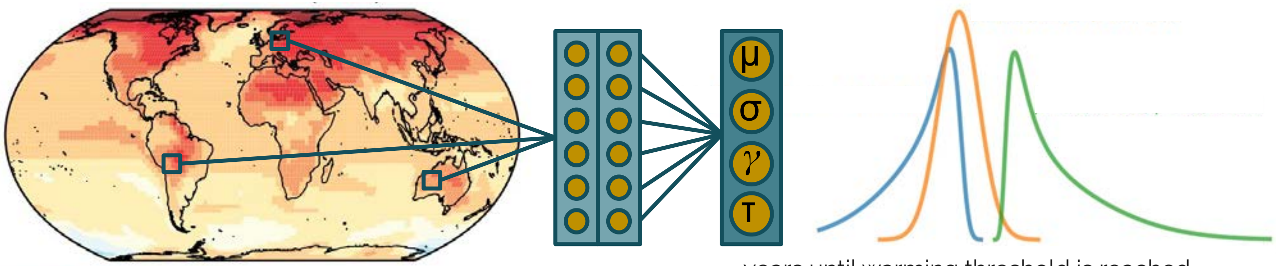


years until warming threshold is reached
+ uncertainties

Train neural network to ingest a single annual temperature map and predict the number of years until a warming threshold is reached

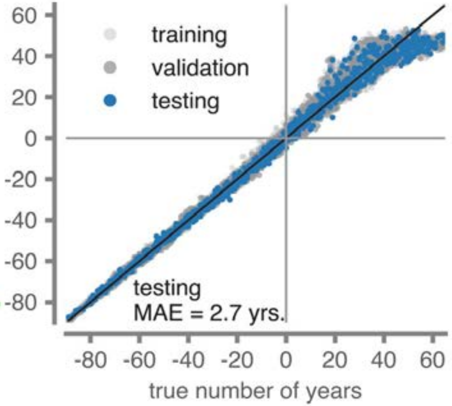
Diffenbaugh & Barnes (2023)

Trained on annual maps from 10 realizations from across multiple climate models



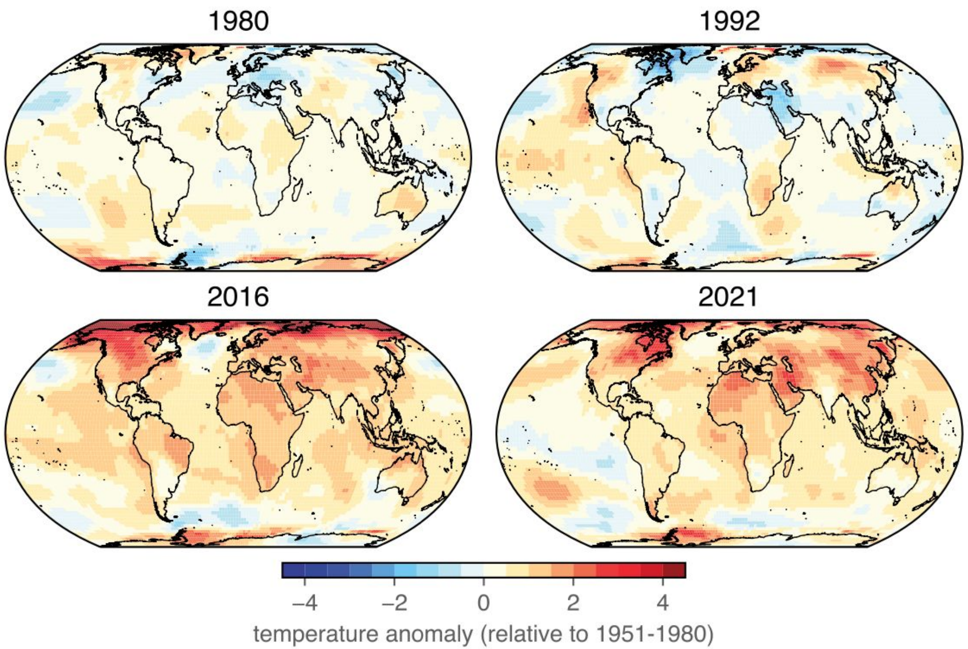
years until warming threshold is reached
+ uncertainties

Climate Model Results



Train neural network to ingest a single annual temperature map and predict the number of years until a warming threshold is reached

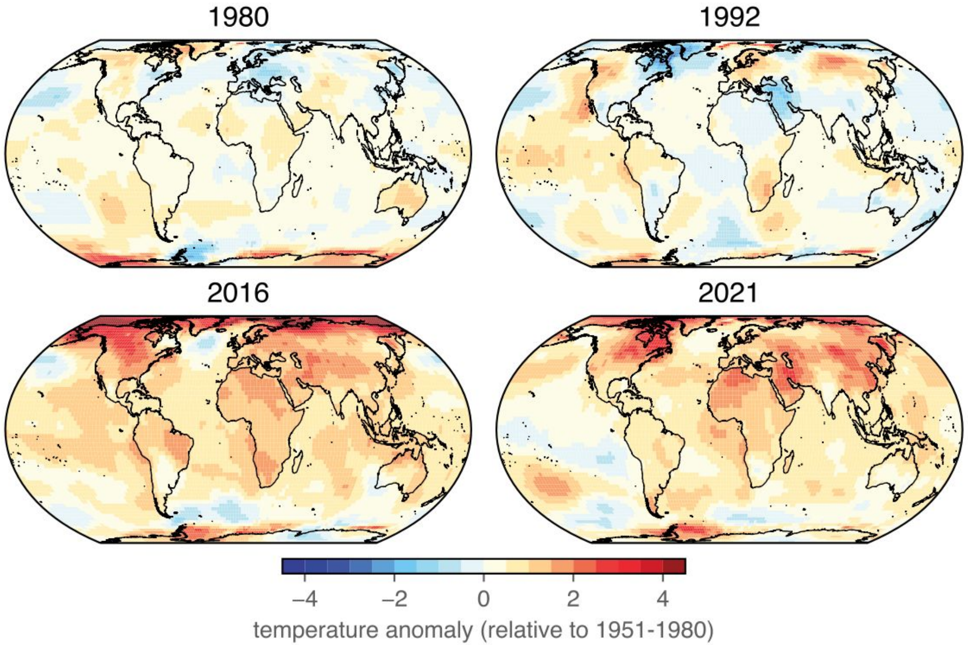
Diffenbaugh & Barnes (2023)

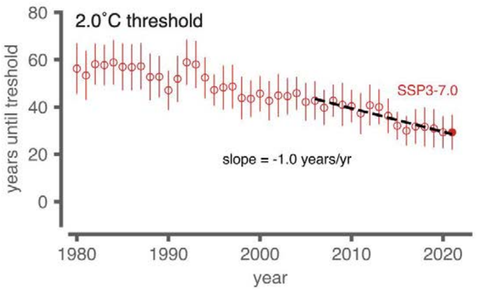


Observations
Berkeley Earth Surface Temperature

Use the trained AI model to predict thresholds based on maps of the observed climate

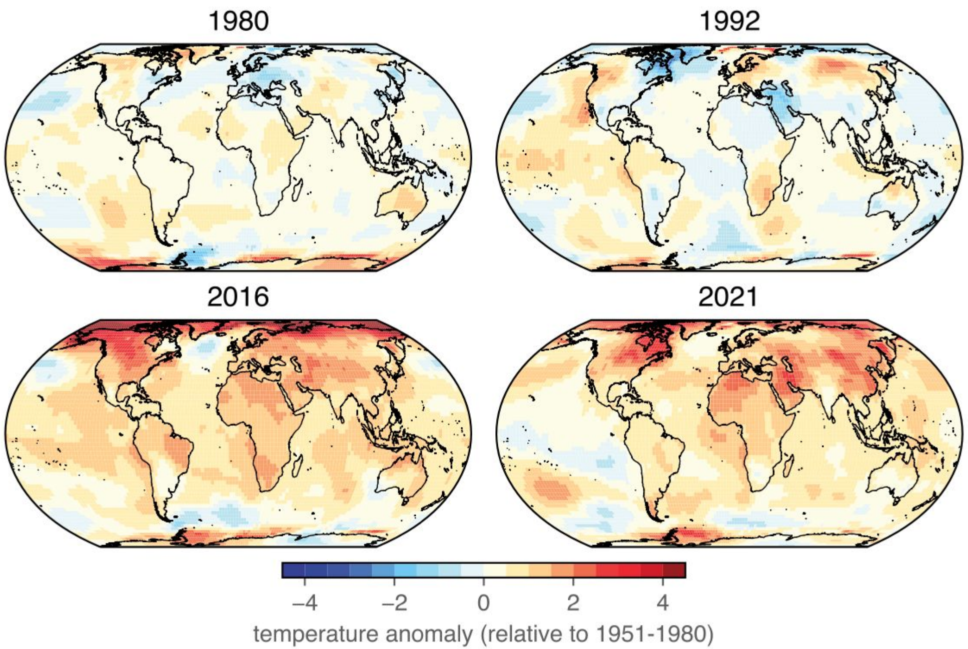
Diffenbaugh & Barnes (2023)

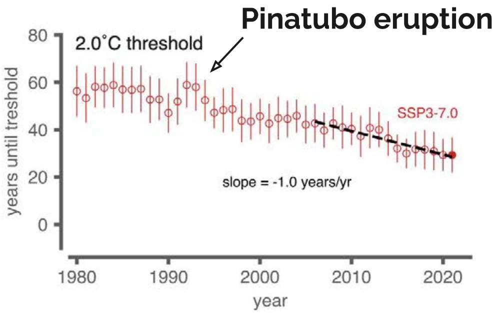




Use the trained AI model to predict thresholds based on maps of the observed climate

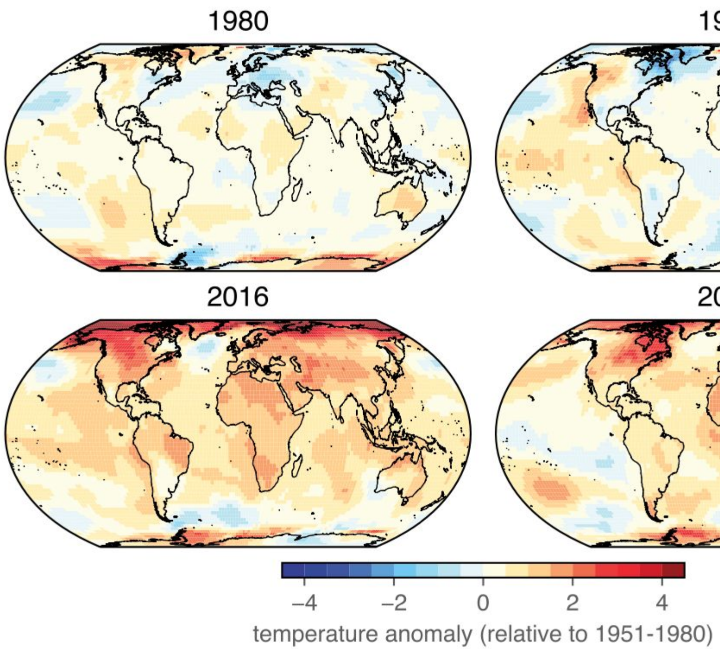
Diffenbaugh & Barnes (2023)

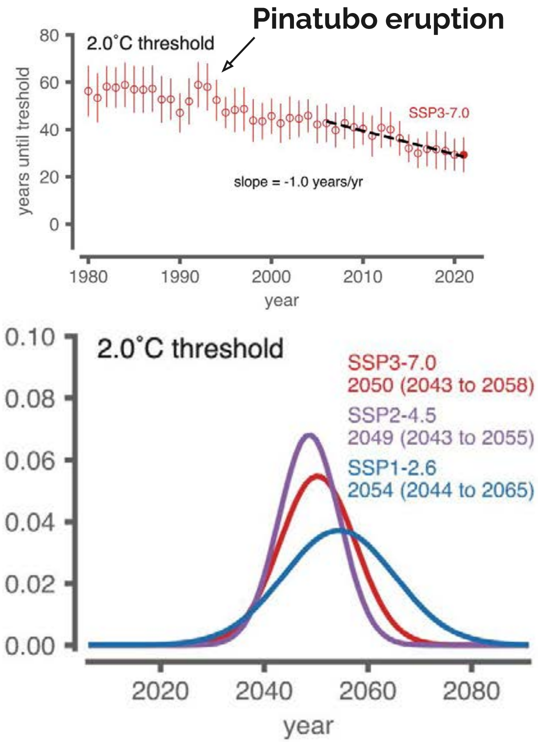




Use the trained AI model to predict thresholds based on maps of the observed climate

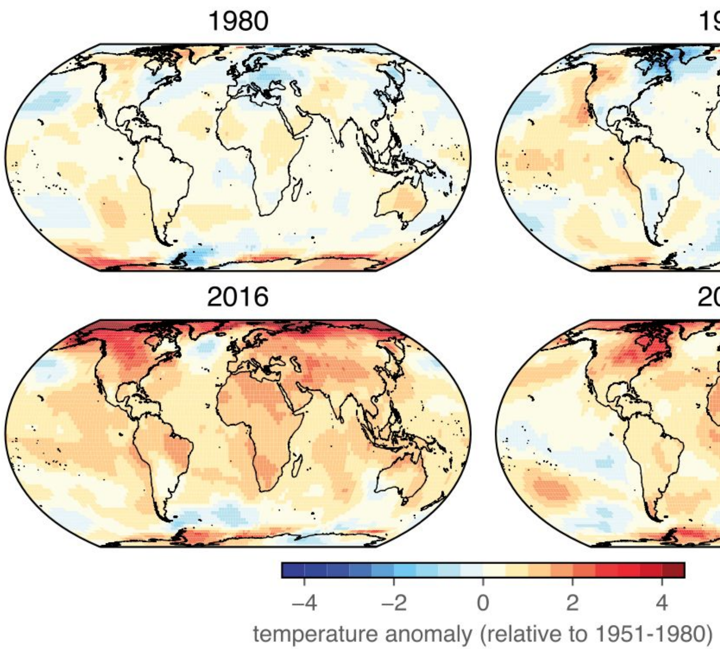
Diffenbaugh & Barnes (2023)

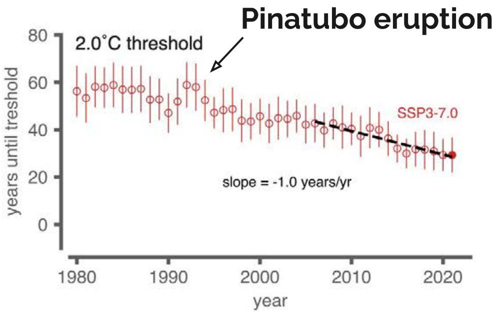


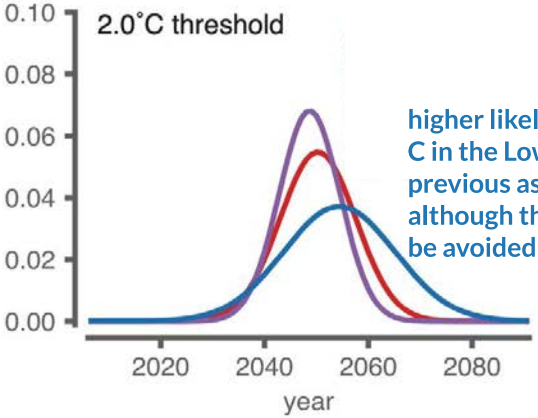


Use the trained AI model to predict thresholds based on maps of the observed climate

Diffenbaugh & Barnes (2023)



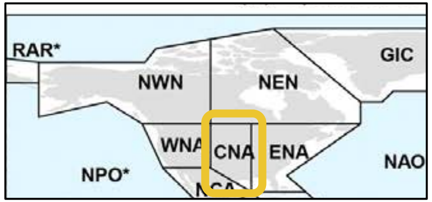




higher likelihood of reaching 2° C in the Low scenario than some previous assessments — although the possibility it could be avoided is not ruled out.

Use the trained AI model to predict thresholds based on maps of the observed climate

Diffenbaugh & Barnes (2023)



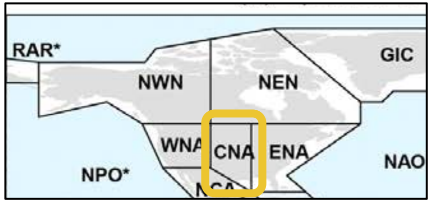
Transfer Learning

predict observations

train on climate
model data

Regional transfer learning provides new insights

Barnes, Diffenbaugh & Seneviratne (2024)



Transfer Learning

predict observations

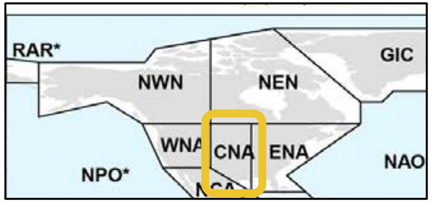
train on climate
model data

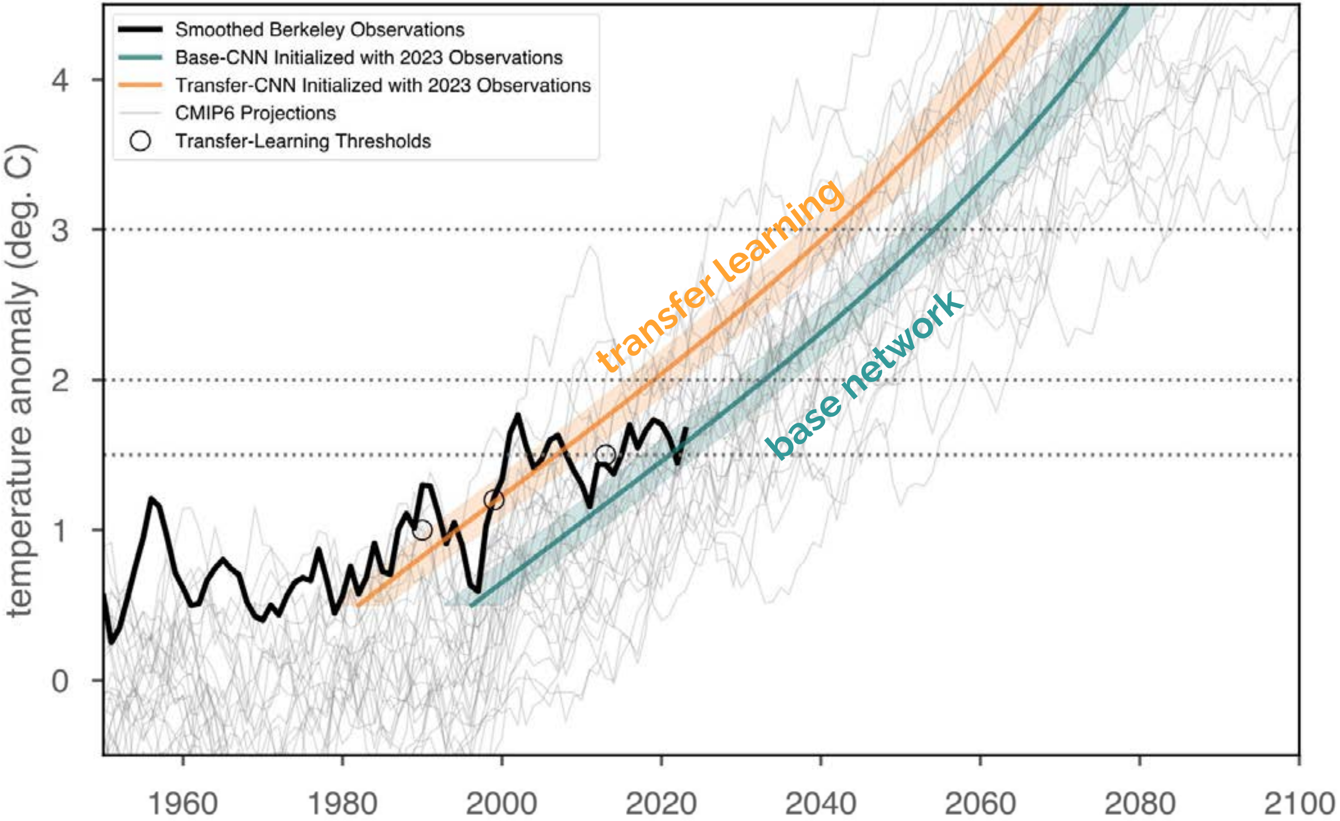
fine-tune with limited
observations (0.8C, 1.0C, 1.2C)

predict observations

Regional transfer learning provides new insights

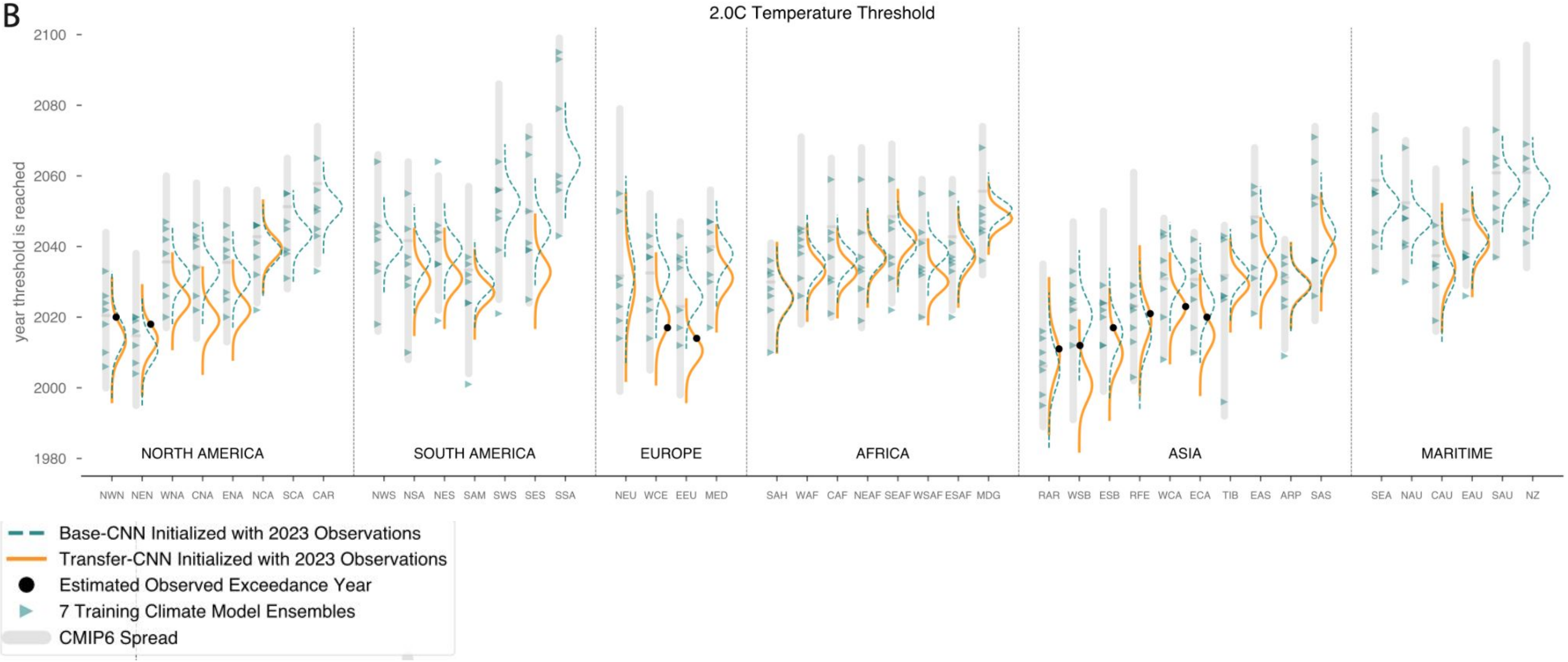

Barnes, Diffenbaugh & Seneviratne (2024)





Regional transfer learning provides new insights


Barnes, Diffenbaugh & Seneviratne (2024)



Regional transfer learning provides new insights


Barnes, Diffenbaugh & Seneviratne (2024)

But how did the network update its prediction?

…what if we could learn which regions of the globe were most relevant to improving the prediction with observations?

Opening the Black Box with XAI

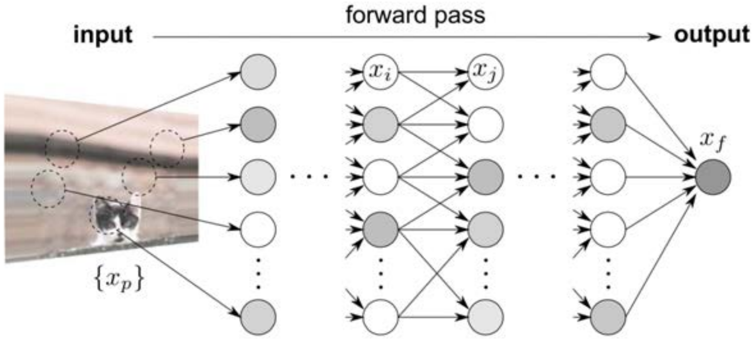
In the past few years multiple papers have come out demonstrating the use of AI explainability methods for earth science

XAI Attribution Methods

Attribution heatmaps are largely consistent with how many climate scientists pose questions

Prediction

of 1 sample



Pr(cat)=.8

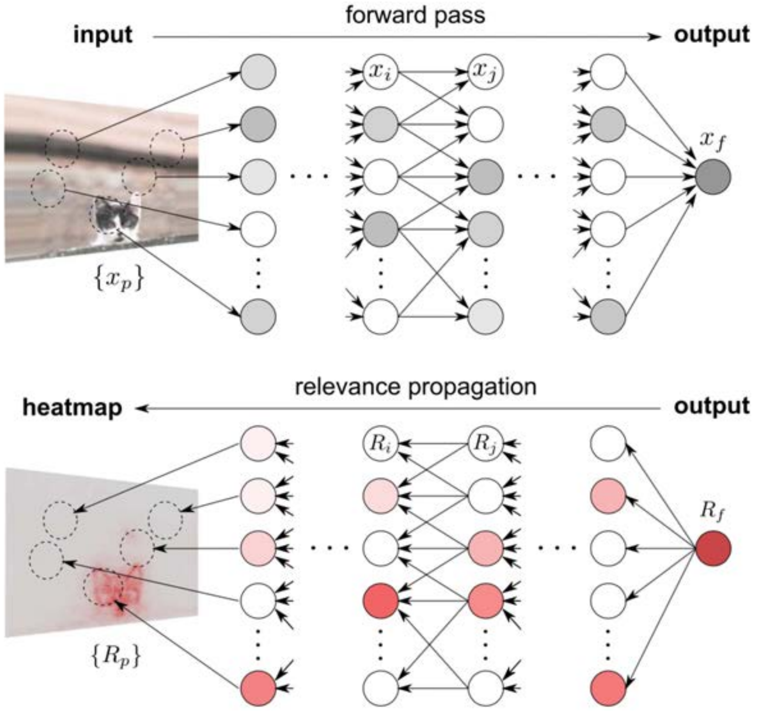
e.g. Montavon et al. (2017), Pattern Recognition; Montavon et al. (2018), Digital Signal Processing

XAI Attribution Methods

Attribution heatmaps are largely consistent with how many climate scientists pose questions

Prediction
of 1 sample

Attribution
of 1 sample



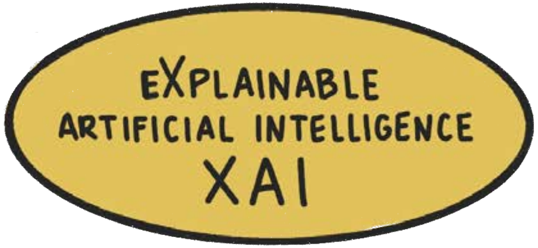
Pr(cat)=.8

Pr(cat)=.8

e.g. Montavon et al. (2017), Pattern Recognition; Montavon et al. (2018), Digital Signal Processing

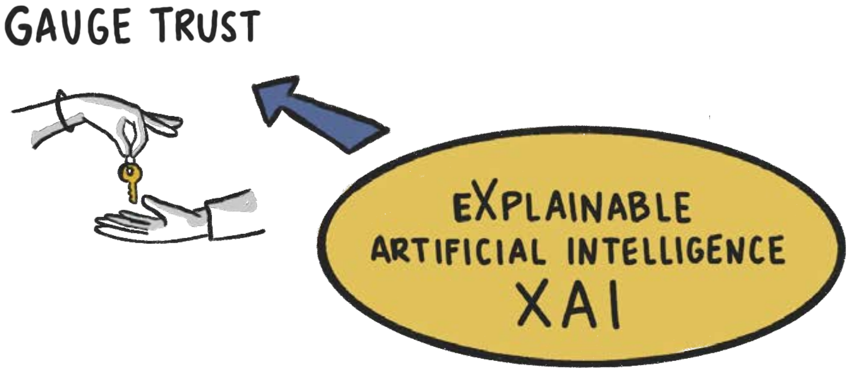
Reasons to care about XAI

A scientist's ultimate goal is typically to understand "why?", but even if you don't care "why?" you should still care about XAI.



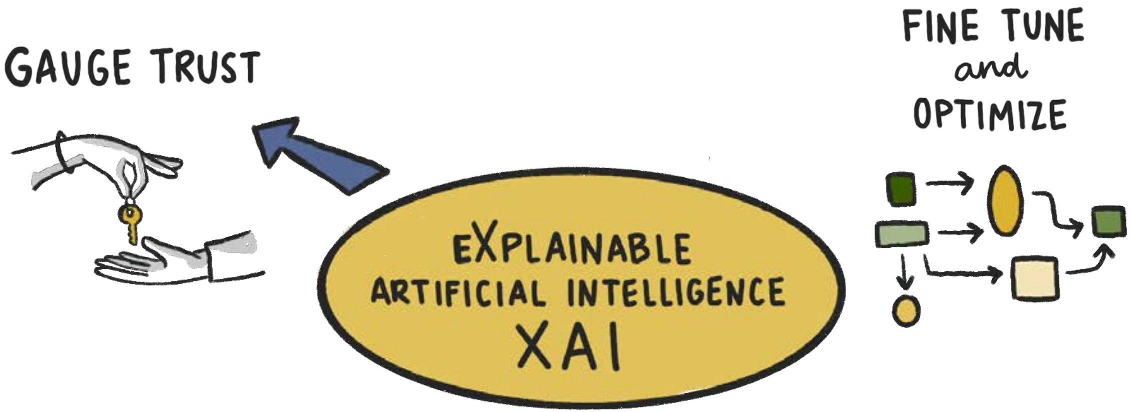
Reasons to care about XAI

A scientist's ultimate goal is typically to understand "why?", but even if you don't care "why?" you should still care about XAI.



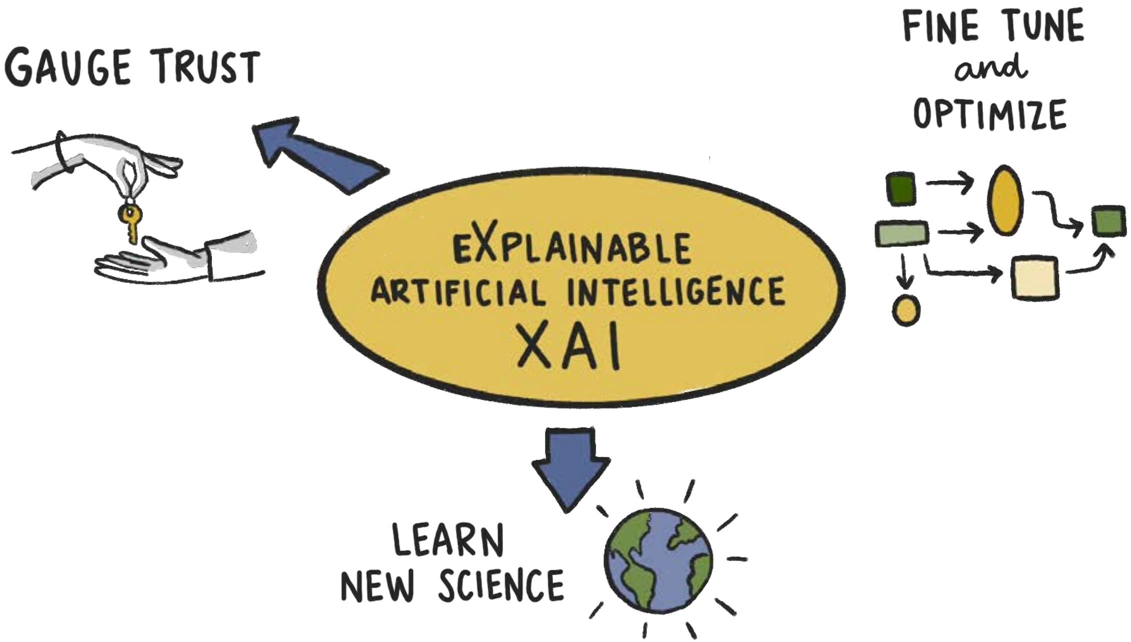
Reasons to care about XAI

A scientist's ultimate goal is typically to understand "why?", but even if you don't care "why?" you should still care about XAI.



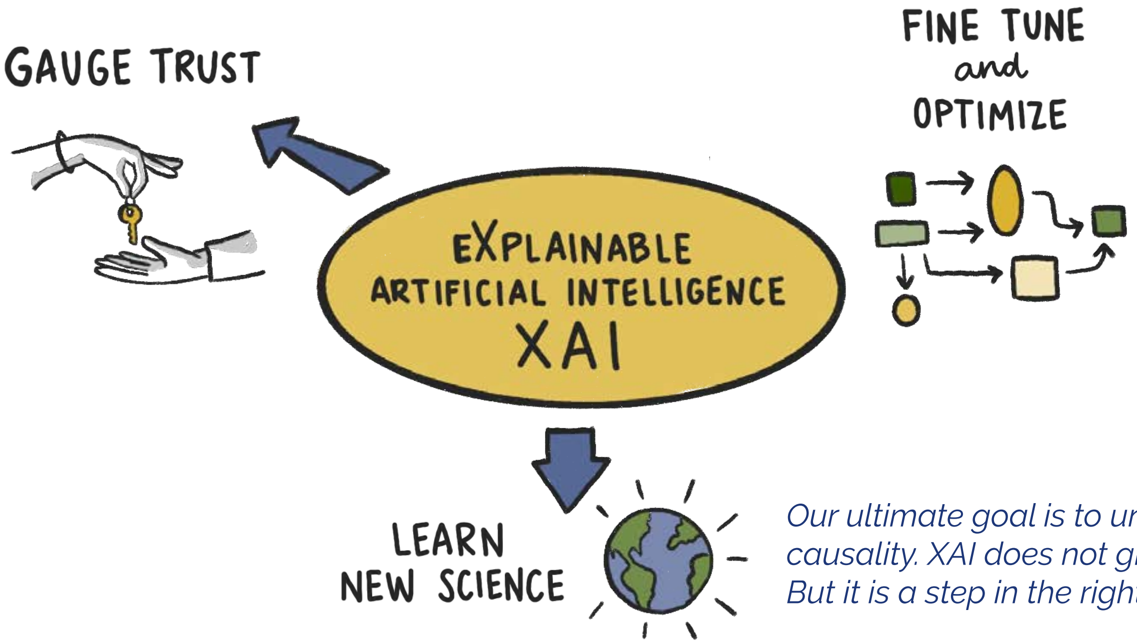
Reasons to care about XAI

A scientist's ultimate goal is typically to understand "why?", but even if you don't care "why?" you should still care about XAI.



Reasons to care about XAI

A scientist's ultimate goal is typically to understand "why?", but even if you don't care "why?" you should still care about XAI.



Our ultimate goal is to understand causality. XAI does not give us this. But it is a step in the right direction.

TIME SUBSCRIBE

TECH • ARTIFICIAL INTELLIGENCE

Chuck Schumer Wants AI to Be Explainable. It's Harder Than It Sounds

9 MINUTE READ

QuantumBlack
AI by McKinsey

Sign in |

OCTOBER 30, 2023

Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

Sec. 8. Protecting Consumers, Patients, Passengers, and Students. (a) Independent regulatory agencies are encouraged, as they deem appropriate, to consider using their full range of authorities to protect American consumers from fraud, discrimination, and threats to privacy and to address other risks that may arise from the use of AI, including risks to financial stability, and to consider rulemaking, as well as emphasizing or clarifying where existing regulations and guidance apply to AI, including clarifying the responsibility of regulated entities to conduct due diligence on and monitor any third-party AI services they use, and emphasizing or clarifying requirements and expectations related to the transparency of AI models and regulated entities' ability to explain their use of AI models.

H.R.6093 - Weather Act Reauthorization Act of 2023

118th Congress (2023-2024) | Get alerts

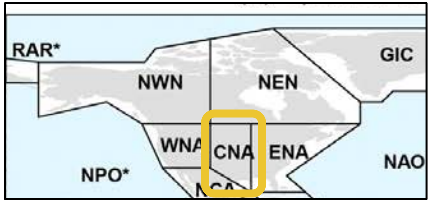
"(c) ARTIFICIAL INTELLIGENCE INVESTMENTS.—*The Under Secretary shall leverage artificial intelligence and machine learning technologies to facilitate, optimize, and further leverage advanced computing to accomplish critical missions of the National Oceanic and Atmospheric Administration by enhancing existing and forthcoming high-performance and cloud computing infrastructure or systems.*

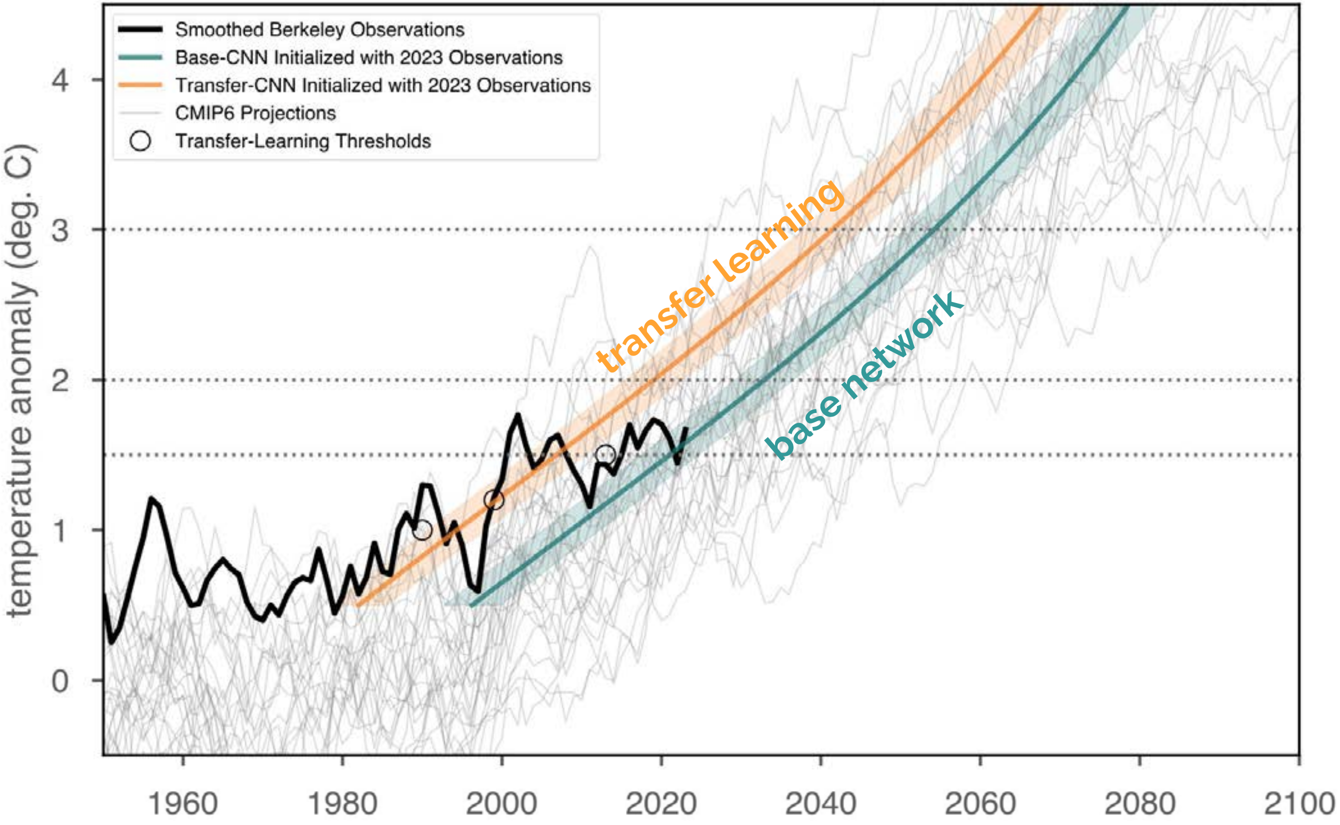
"(d) CENTERS OF EXCELLENCE.—*The Under Secretary may expand, and where applicable establish, centers of excellence to aid the adoption of next-generation artificial intelligence and machine learning enabled advanced computing capabilities. Each such center may carry out activities that include the following:*

The EU AIA mandates that high-risk AI systems must provide clear and comprehensible information about their capabilities and limitations, and that their decision-making process should be transparent and traceable.

Jun 7, 2023

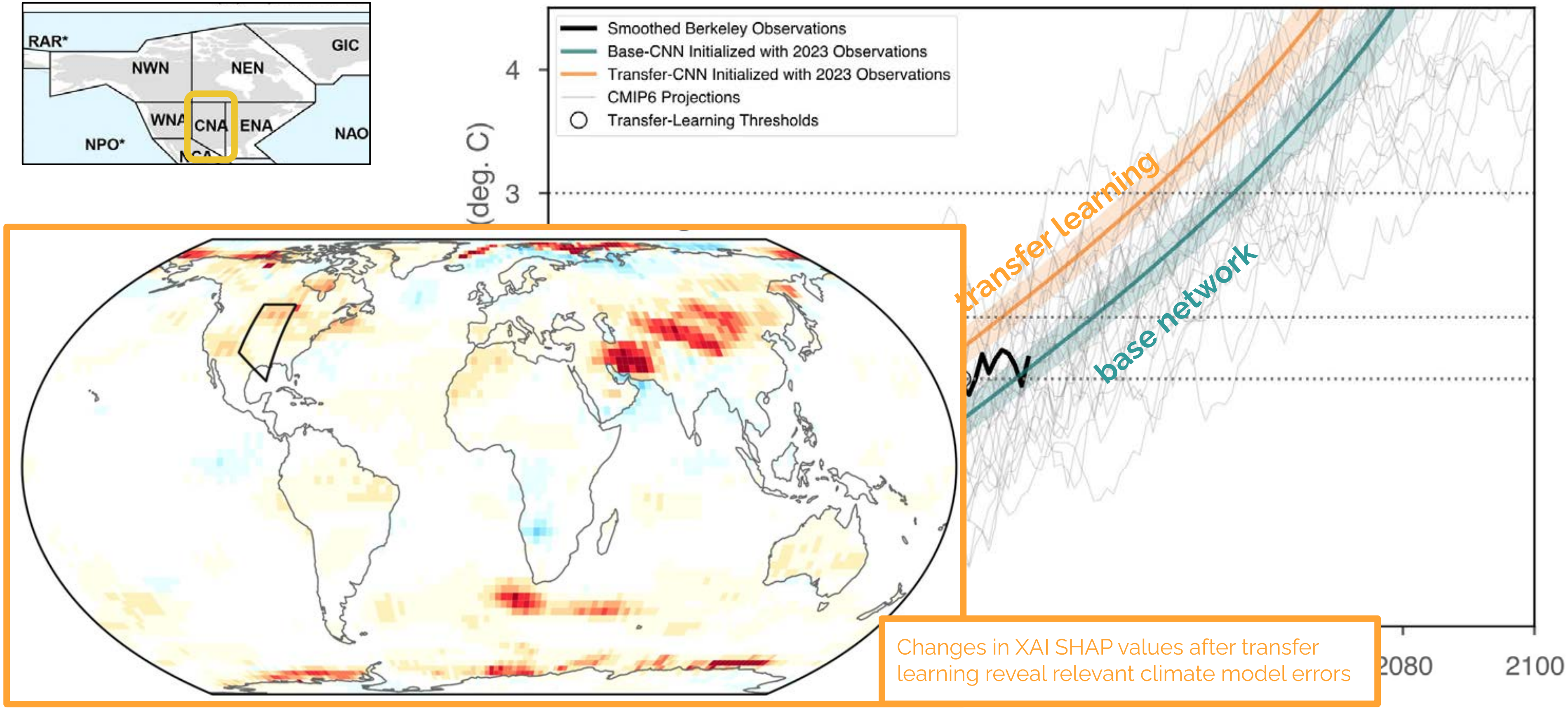
Explainable AI (XAI) will be essential.





Regional transfer learning provides new insights

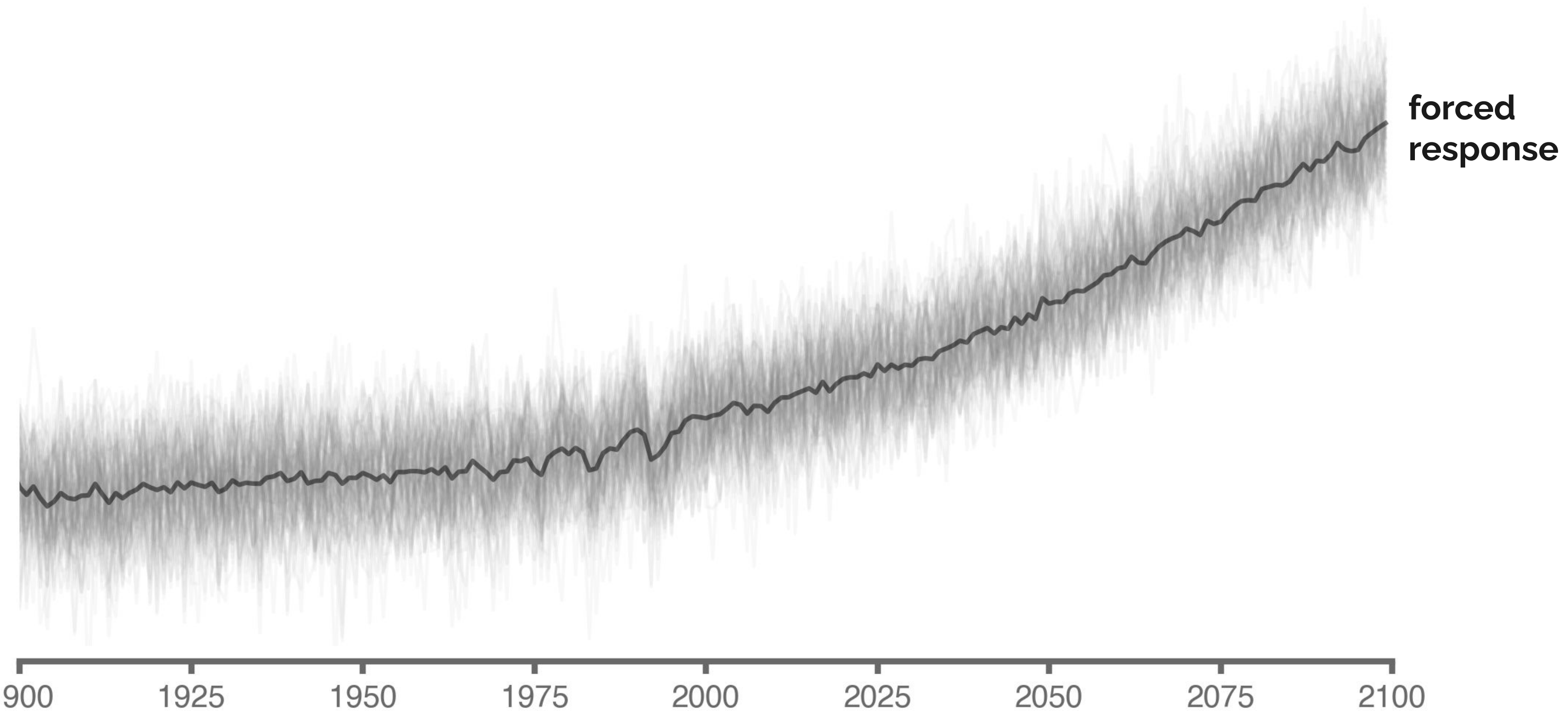
Barnes, Diffenbaugh & Seneviratne (2024)



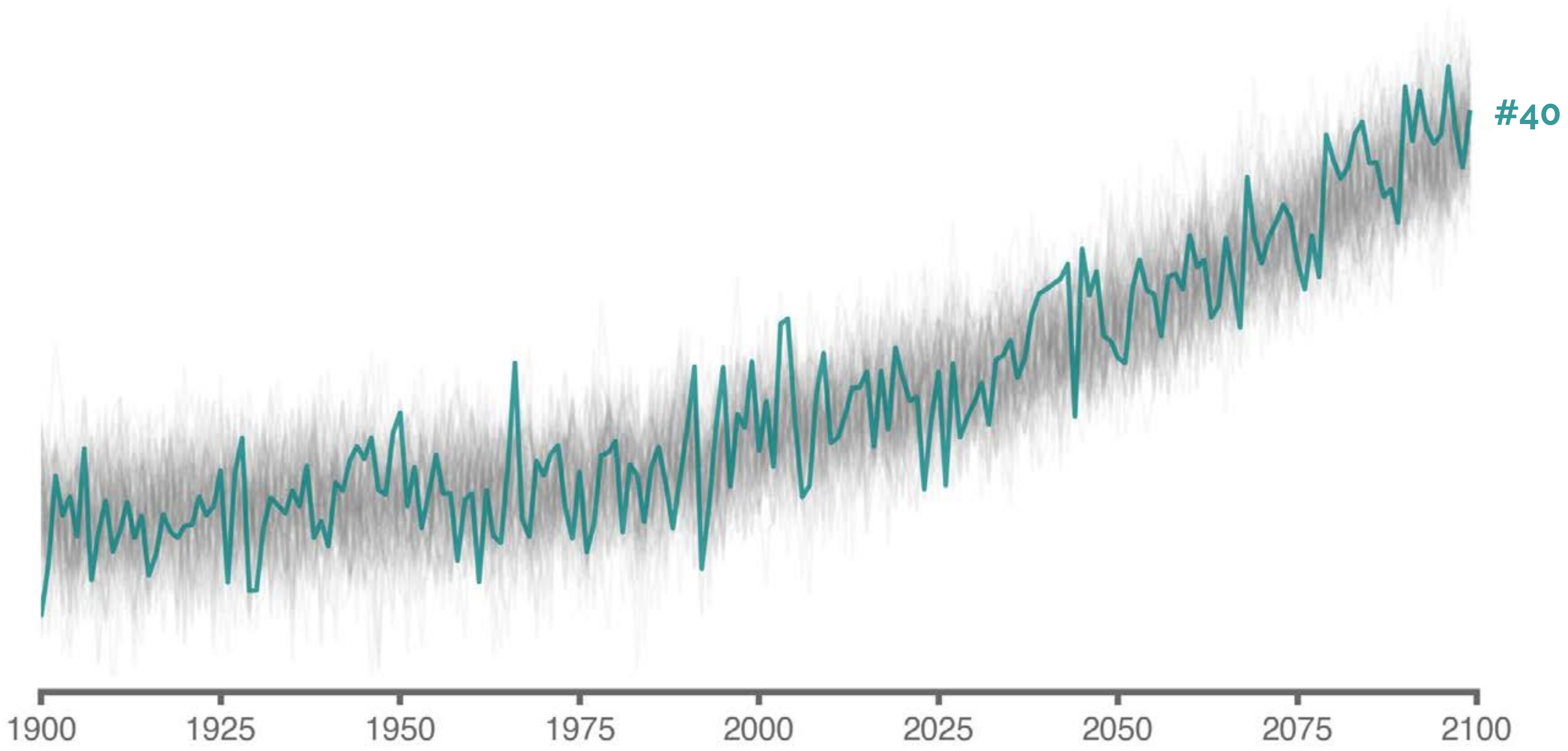
Changes in XAI SHAP values after transfer learning reveal relevant climate model errors

Regional transfer learning provides new insights

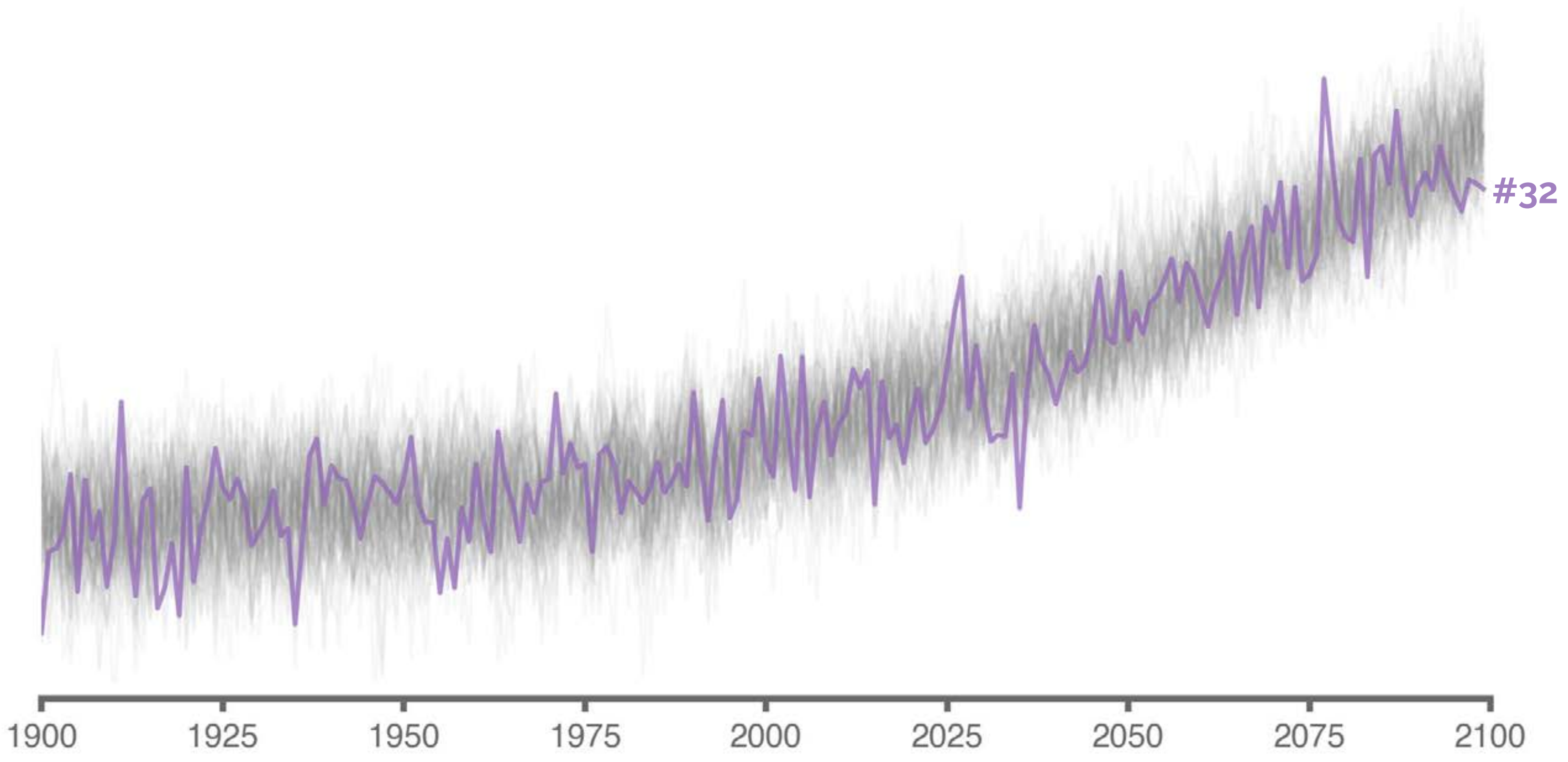
Barnes, Diffenbaugh & Seneviratne (2024)



Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5

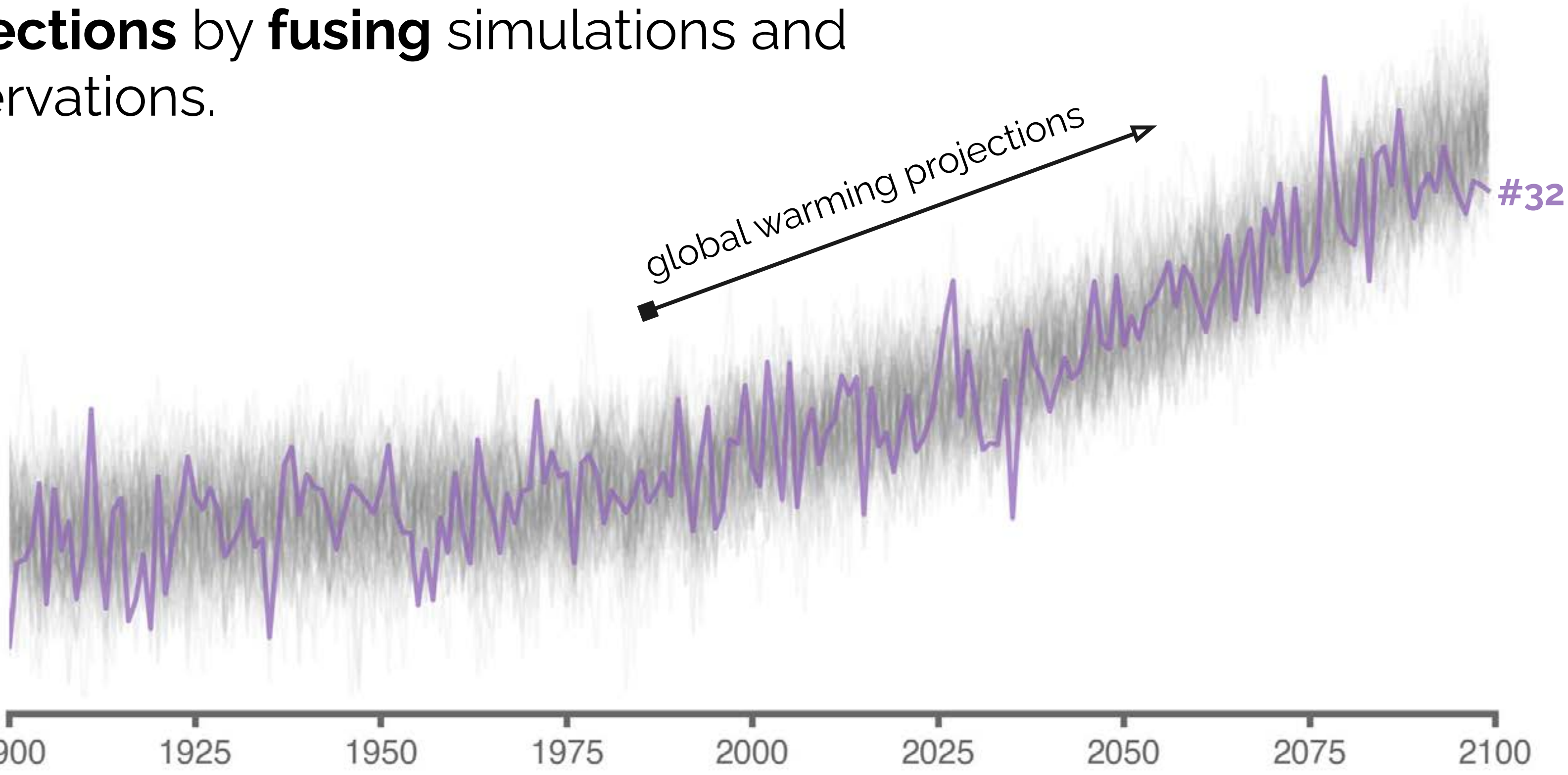


Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5



Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5

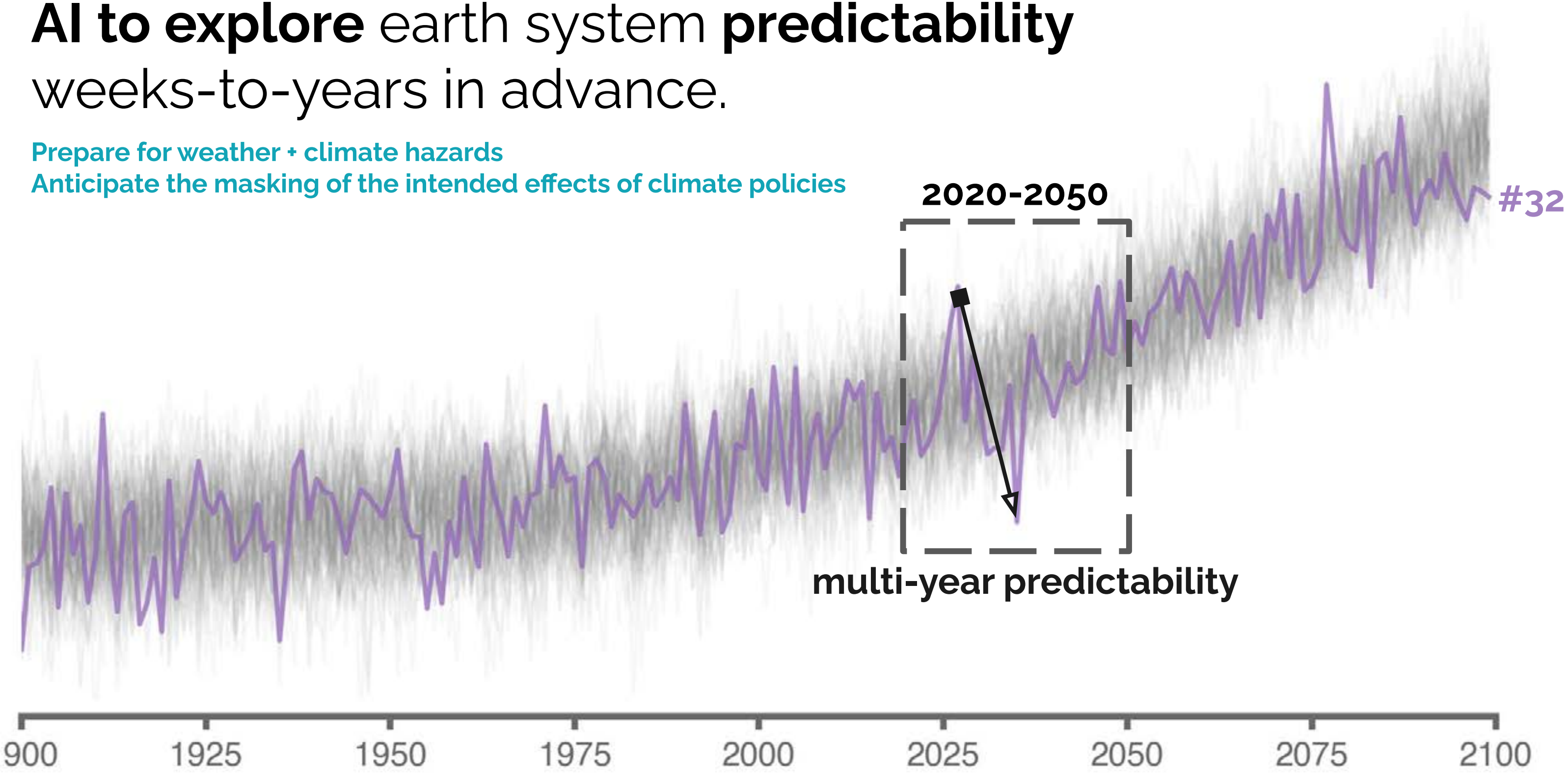
AI to leverage imperfect climate models to better **constrain future projections** by **fusing** simulations and observations.



Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5

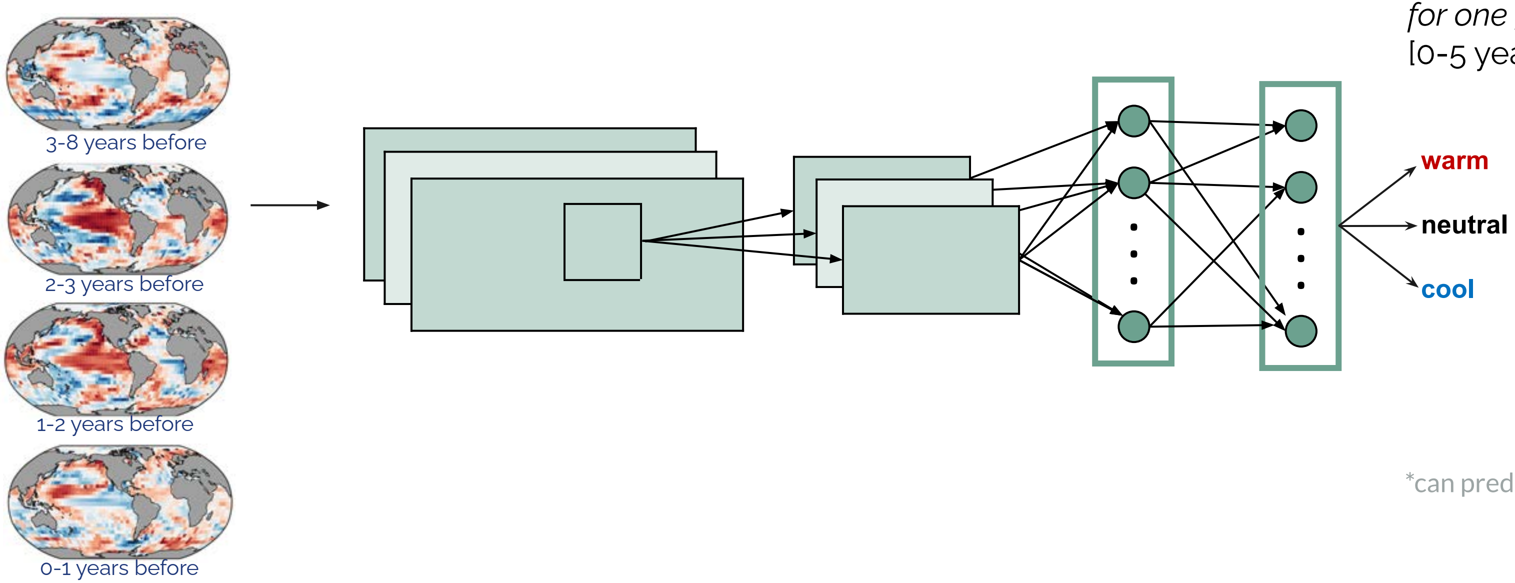
AI to explore earth system **predictability** weeks-to-years in advance.

Prepare for weather + climate hazards
Anticipate the masking of the intended effects of climate policies



Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5

past sea-surface temperatures



future sea surface temperatures*

for one grid point

[0-5 years]

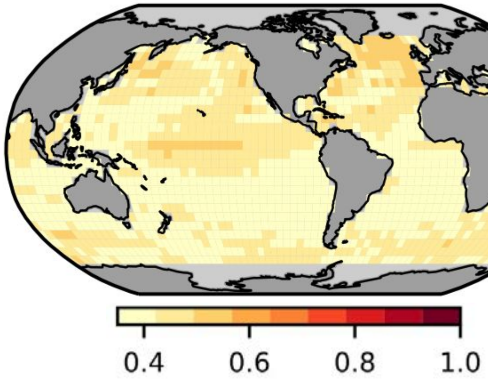
*can predict a range of variables

Predict ocean temperatures 5 years later

Davenport, Barnes & Gordon (2024)

CLIMATE MODEL DATA

Overall Accuracy



Trained on climate model **MPI-ESM-1-2-LR [3,630 years of data]**
Evaluated on climate model **MPI-ESM-1-2-LR**

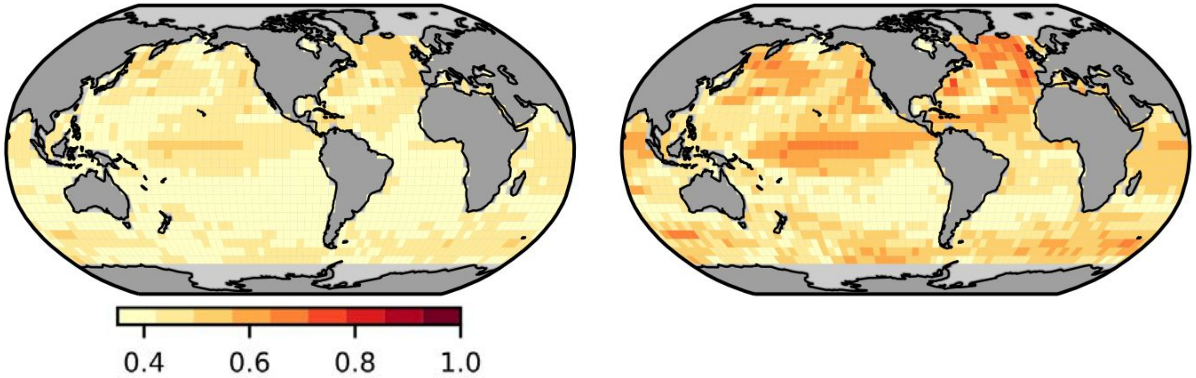
Focusing on when the AI is most confident leads to skillful predictions

Davenport, Barnes & Gordon (2024)

CLIMATE MODEL DATA

Overall Accuracy

Accuracy for 40% most confident predictions



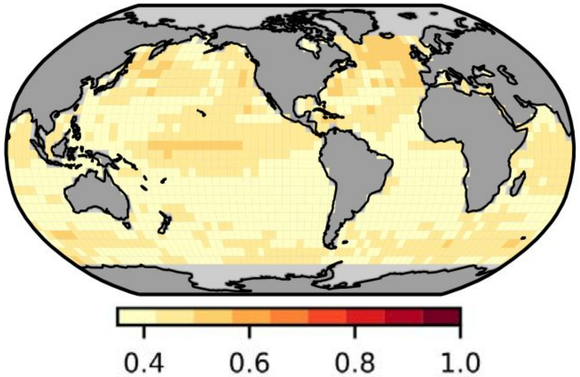
Trained on climate model **MPI-ESM-1-2-LR [3,630 years of data]**
Evaluated on climate model **MPI-ESM-1-2-LR**

Focusing on when the AI is most confident leads to skillful predictions

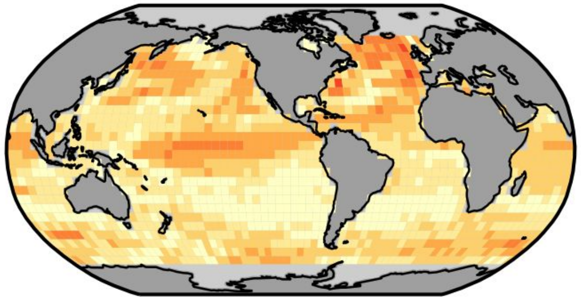

Davenport, Barnes & Gordon (2024)

CLIMATE MODEL DATA

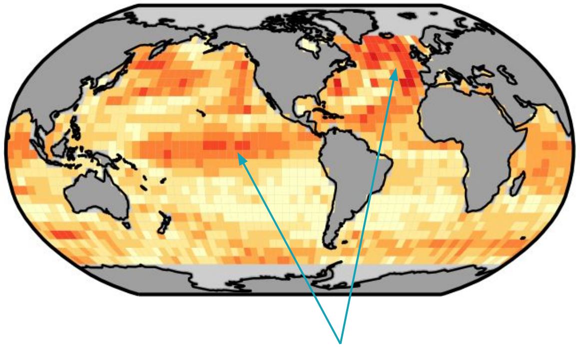
Overall Accuracy



Accuracy for 40% most confident predictions



Accuracy for 20% most confident predictions



>60-80% accuracy!

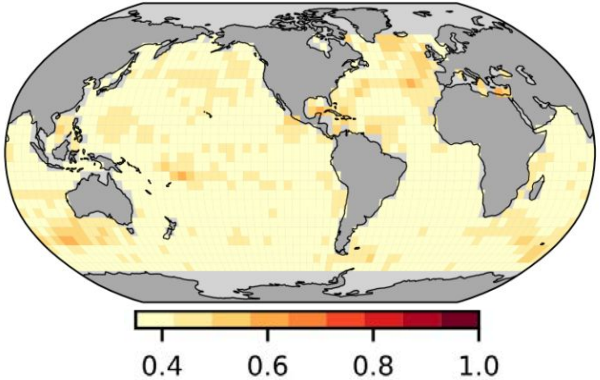
Trained on climate model **MPI-ESM-1-2-LR [3,630 years of data]**
Evaluated on climate model **MPI-ESM-1-2-LR**

Focusing on when the AI is most confident leads to skillful predictions

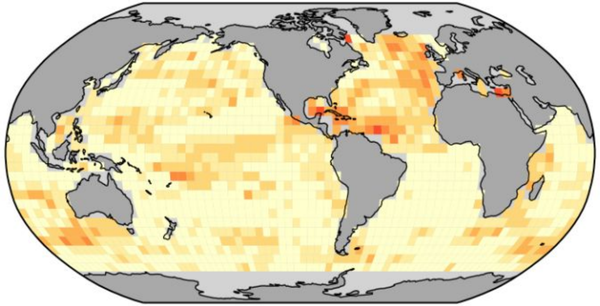

Davenport, Barnes & Gordon (2024)

OBSERVATIONS

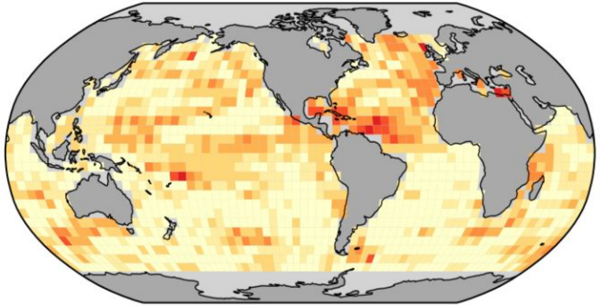
Overall Accuracy



Accuracy for 40% most confident predictions



Accuracy for 20% most confident predictions

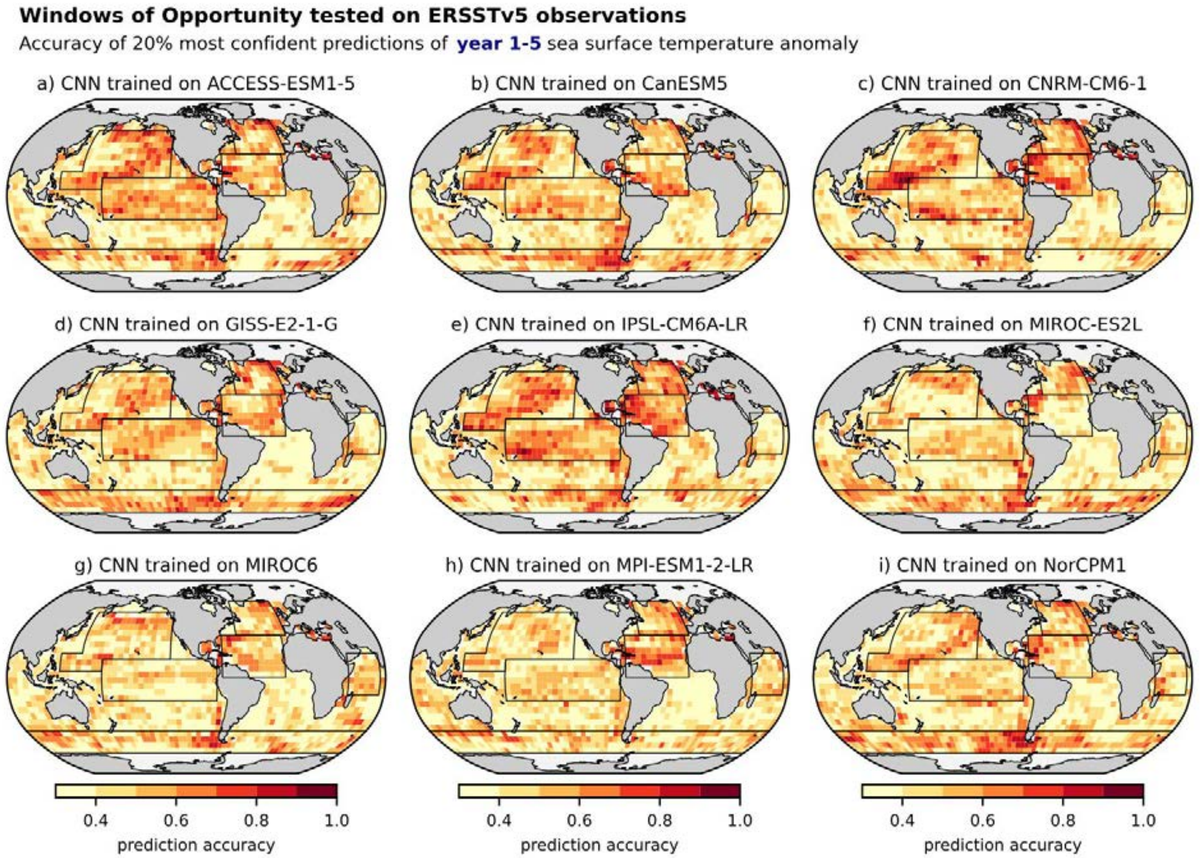


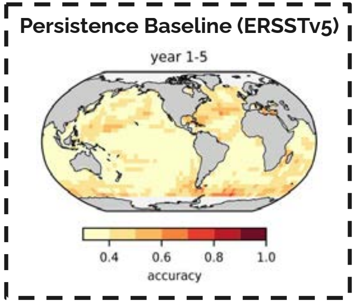
Trained on climate model **MPI-ESM-1-2-LR [3,630 years of data]**
Evaluated on **observations [ERSSTv5; 169 years of data]**

Leveraging climate model data provides skillful predictions of the real world



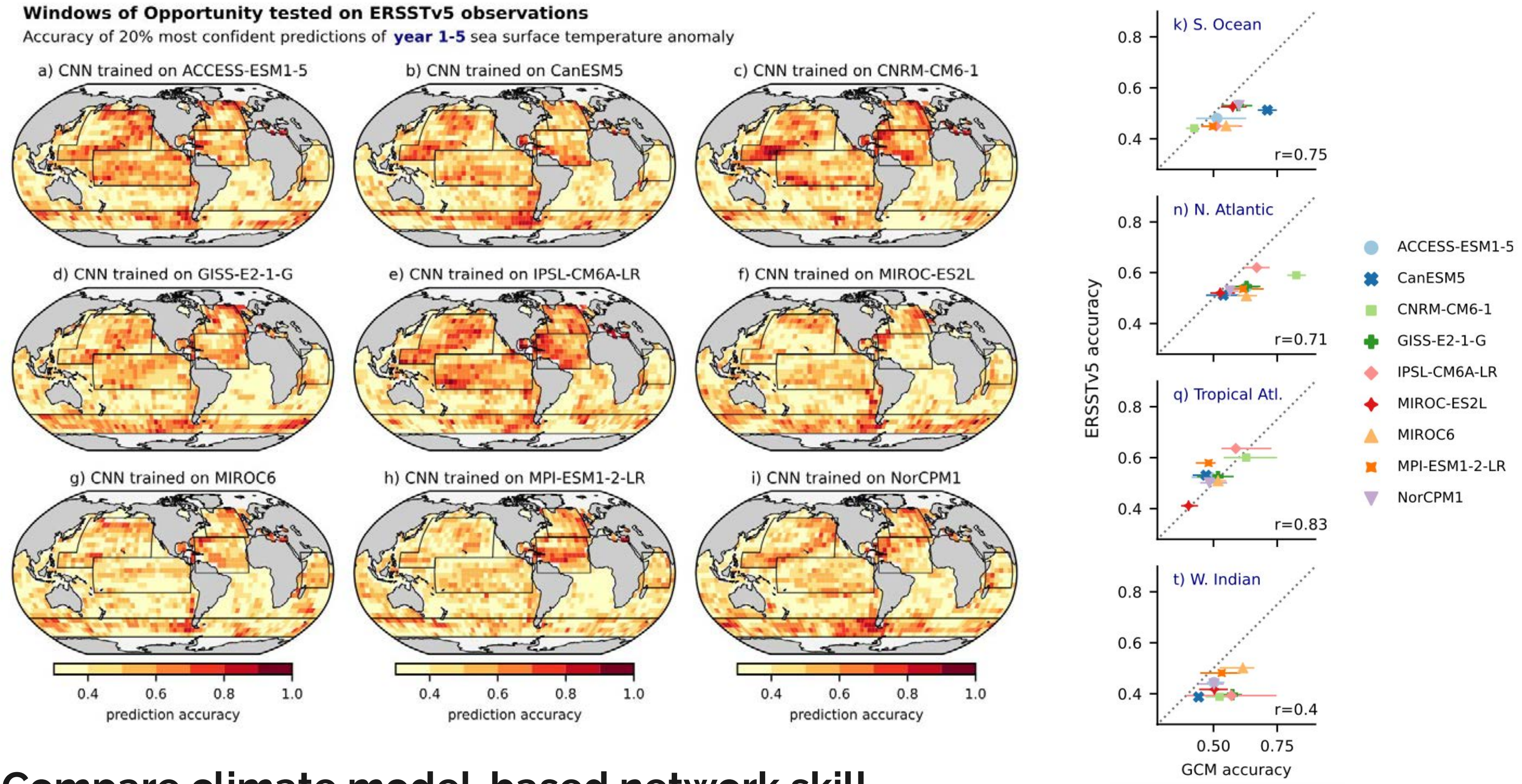
Davenport, Barnes & Gordon (2024)





Compare climate model-based network skill

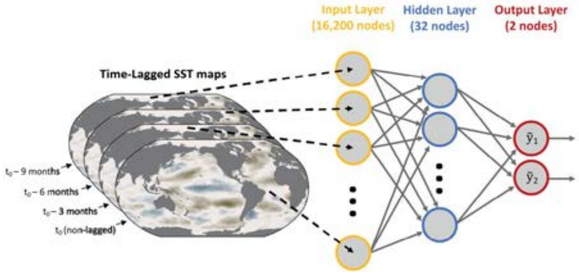

Davenport, Barnes & Gordon (2024)



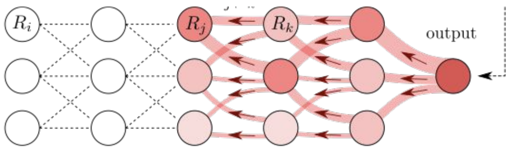
Compare climate model-based network skill

Davenport, Barnes & Gordon (2024)

Predicting 5-year average surface temperature at each grid point
Applied to 1200 years of CESM2 control simulation



Prediction of average land surface temperature anomaly over the next 5 years

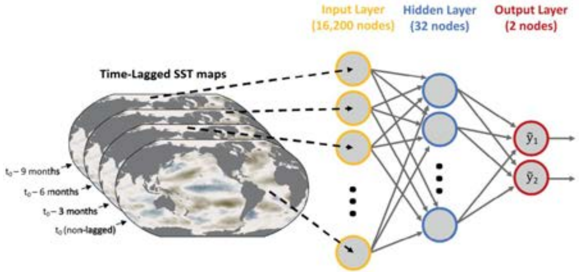


XAI

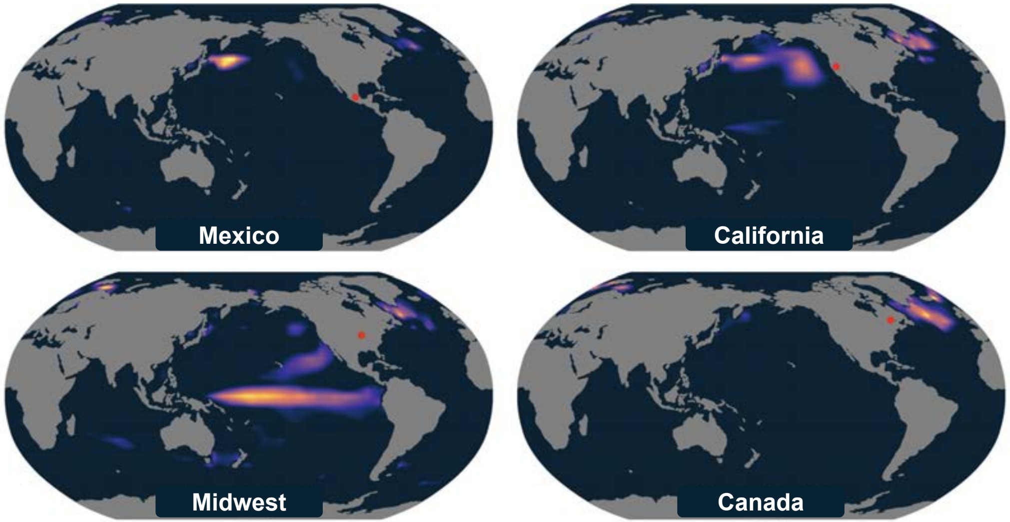
XAI reveals sources of predictability that vary in time and space

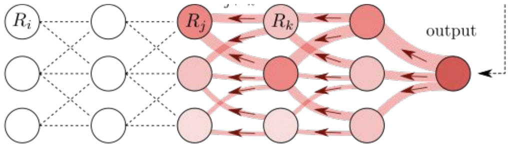
Toms, Barnes & Hurrell (2021)

Predicting 5-year average surface temperature at each grid point
Applied to 1200 years of CESM2 control simulation



Prediction of average land surface temperature anomaly over the next 5 years



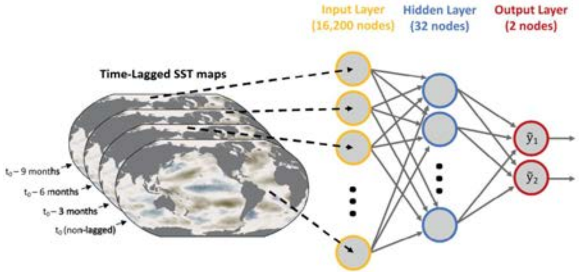


XAI

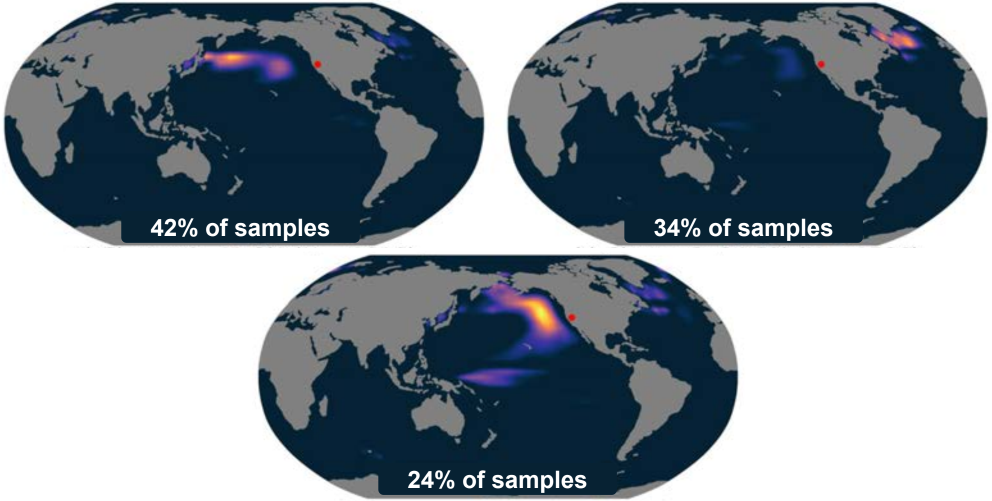
XAI reveals sources of predictability that vary in time and space

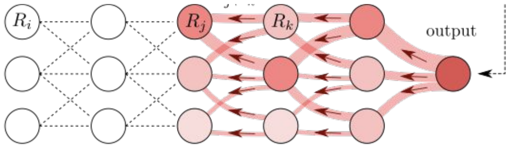
Toms, Barnes & Hurrell (2021)

Predicting 5-year average surface temperature at each grid point
Applied to 1200 years of CESM2 control simulation



Prediction of average land surface temperature anomaly over the next 5 years





XAI

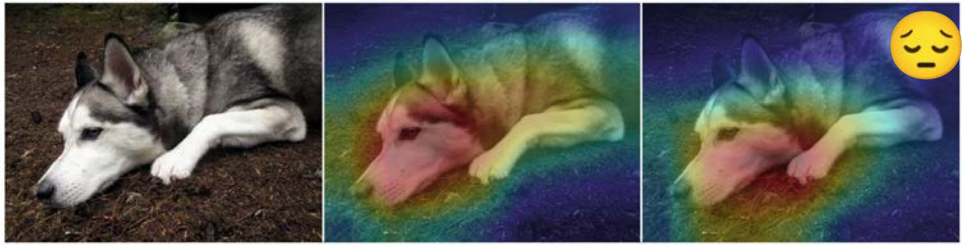
XAI reveals sources of predictability that vary in time and space


Toms, Barnes & Hurrell (2021)

PERSPECTIVE
https://doi.org/10.1038/s42256-019-0048-x
nature machine intelligence

Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead

Cynthia Rudin



Rudin (2019)

Interpretable AI

Explainable AI tells us ***where***, but not ***how***.
Interpretable AI **explicitly incorporates the decision-making process** into its structure.
Models are **interpretable by design.**

Our Current Goal:
work toward building AI models that mimic scientific human reasoning to improve intrinsic interpretability

Interpretable AI

Explainable AI tells us ***where***, but not ***how***.
Interpretable AI **explicitly incorporates the decision-making process** into its structure.
Models are **interpretable by design.**

Climate Model

Atmospheric Predictability as Revealed by ~~Naturally Occurring~~ Analogues

EDWARD N. LORENZ

Dept. of Meteorology, Massachusetts Institute of Technology, Cambridge, Mass.[1]

(Manuscript received 2 April 1969)

ABSTRACT

Two states of the atmosphere which are observed to resemble one another are termed *analogues*. Either state of a pair of analogues may be regarded as equal to the other state plus a small superposed "error." From the behavior of the atmosphere following each state, the growth rate of the error may be determined.

Five years of twice-daily height values of the 200-, 500-, and 850-mb surfaces at a grid of 1003 points over the Northern Hemisphere are procured. A weighted root-mean-square height difference is used as a measure of the difference between two states, or the error. For each pair of states occurring within one month of the same time of year, but in different years, the error is computed.

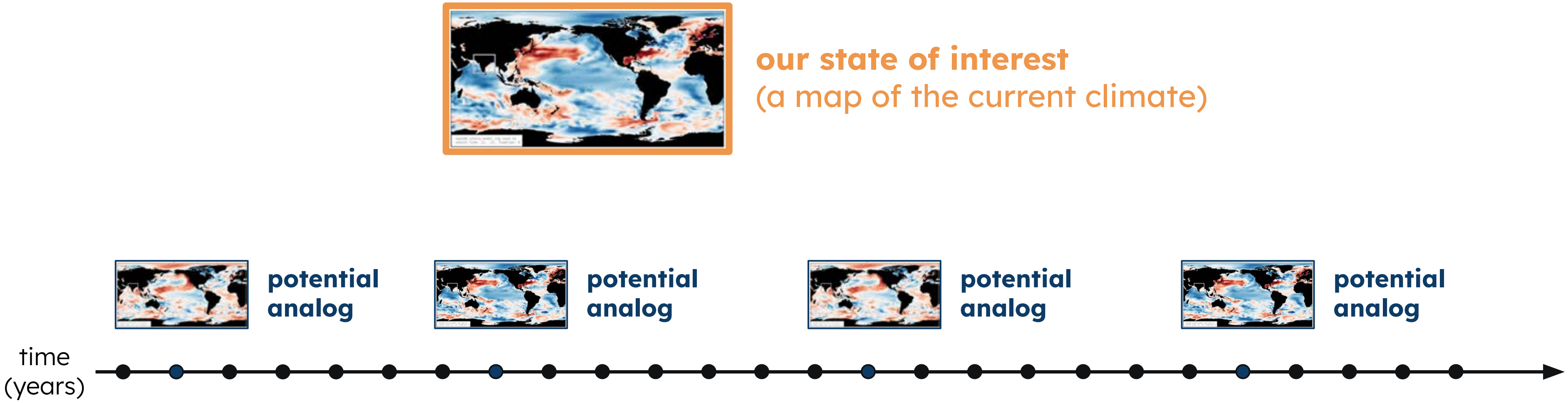
There are numerous mediocre analogues but no truly good ones. The smallest errors have an average doubling time of about 8 days. Larger errors grow less rapidly. Extrapolation with the aid of a quadratic hypothesis indicates that truly small errors would double in about 2.5 days. These rates may be compared with a 5-day doubling time previously deduced from dynamical considerations.

The possibility that the computed growth rate is spurious, and results only from having superposed the smaller errors on those particular states where errors grow most rapidly, is considered and rejected. The likelihood of encountering any truly good analogues by processing all existing upper-level data appears to be small.

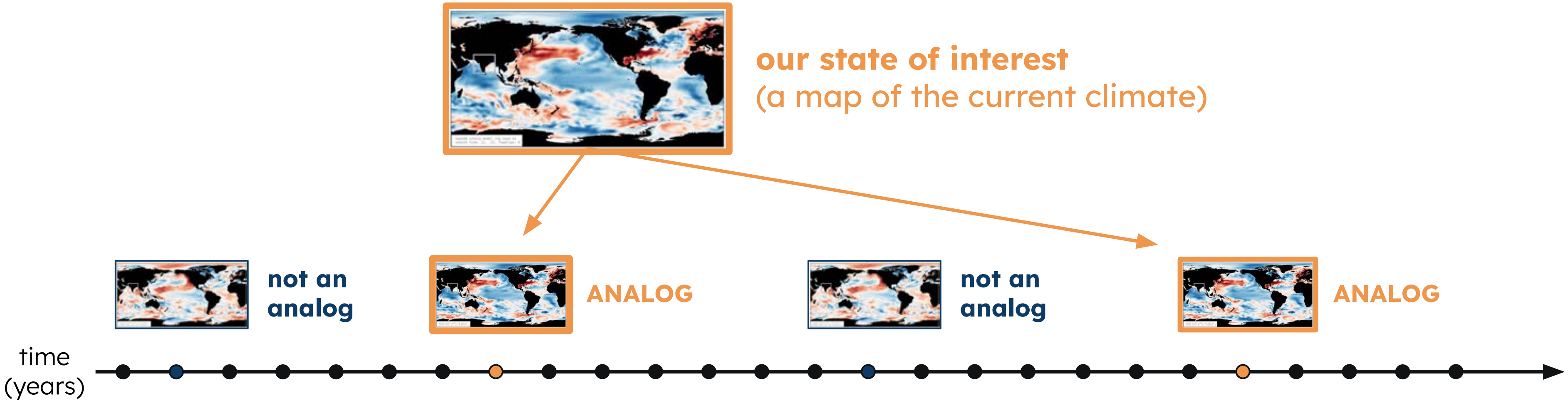
Method of Analogs



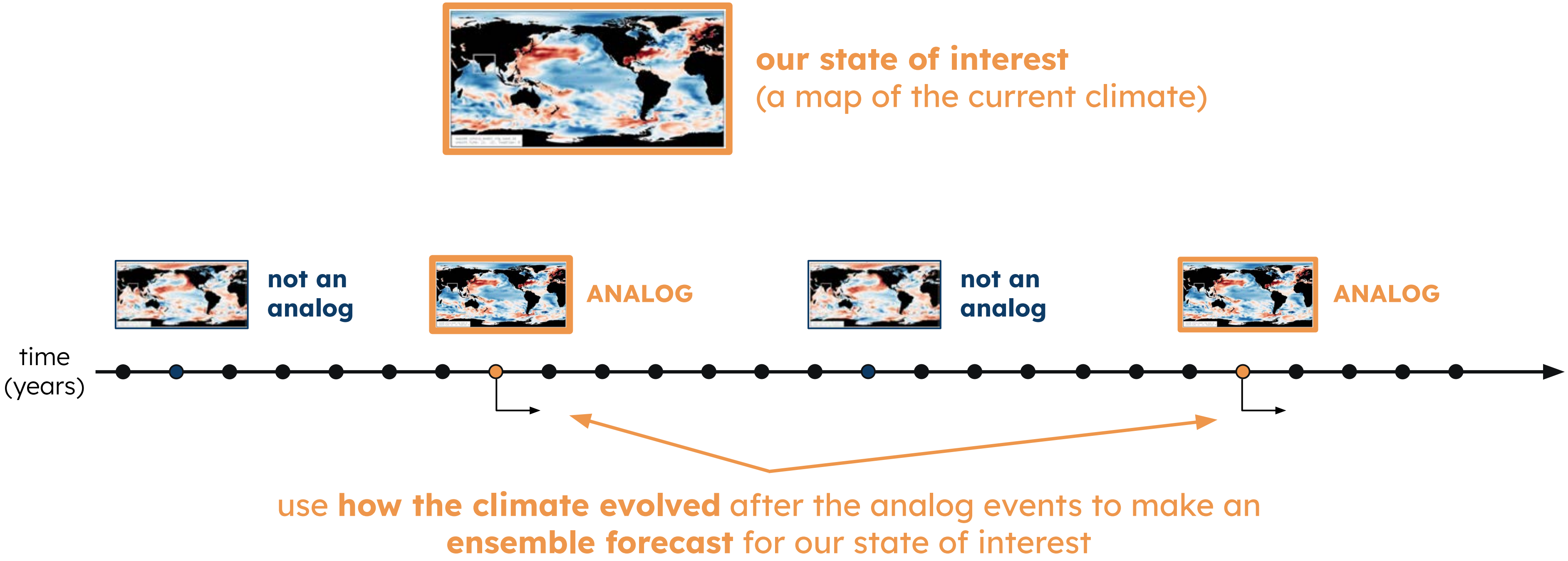
Barnes et al. (2022)
Rader and Barnes (2024)
Gordillo and Barnes (under review)
Fernandez and Barnes (in prep)



Analogs for seasonal-to-decadal prediction

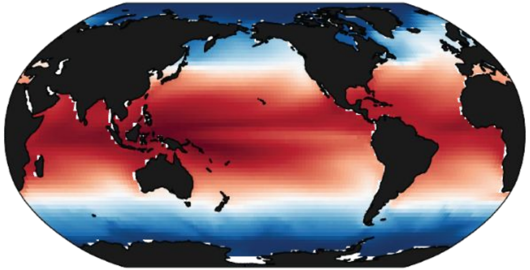


Analogs for seasonal-to-decadal prediction

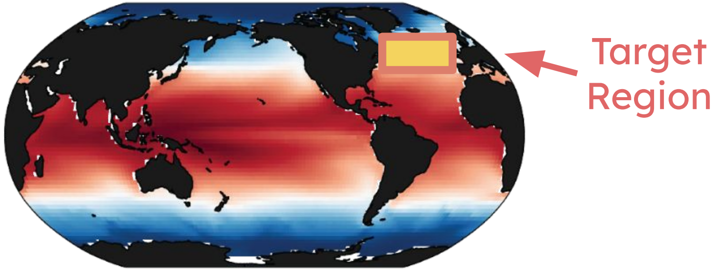


Analogs for seasonal-to-decadal prediction

Use:
SST in Year 0-4



To Predict:
North Atlantic SST in Year 5-9

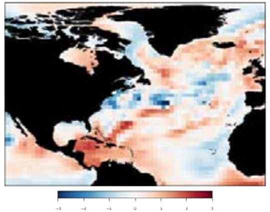


** a perfect model application using the MPI Grand Ensemble*

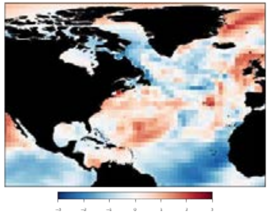
Analog Forecasts for North Atlantic Sea Surface Temperature

Traditionally, if two states **“look similar”** it means they have the smallest difference over the entire globe, or a predefined region.

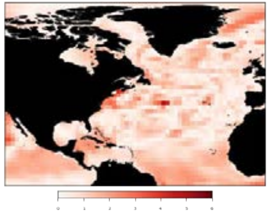
abs(



−



) =

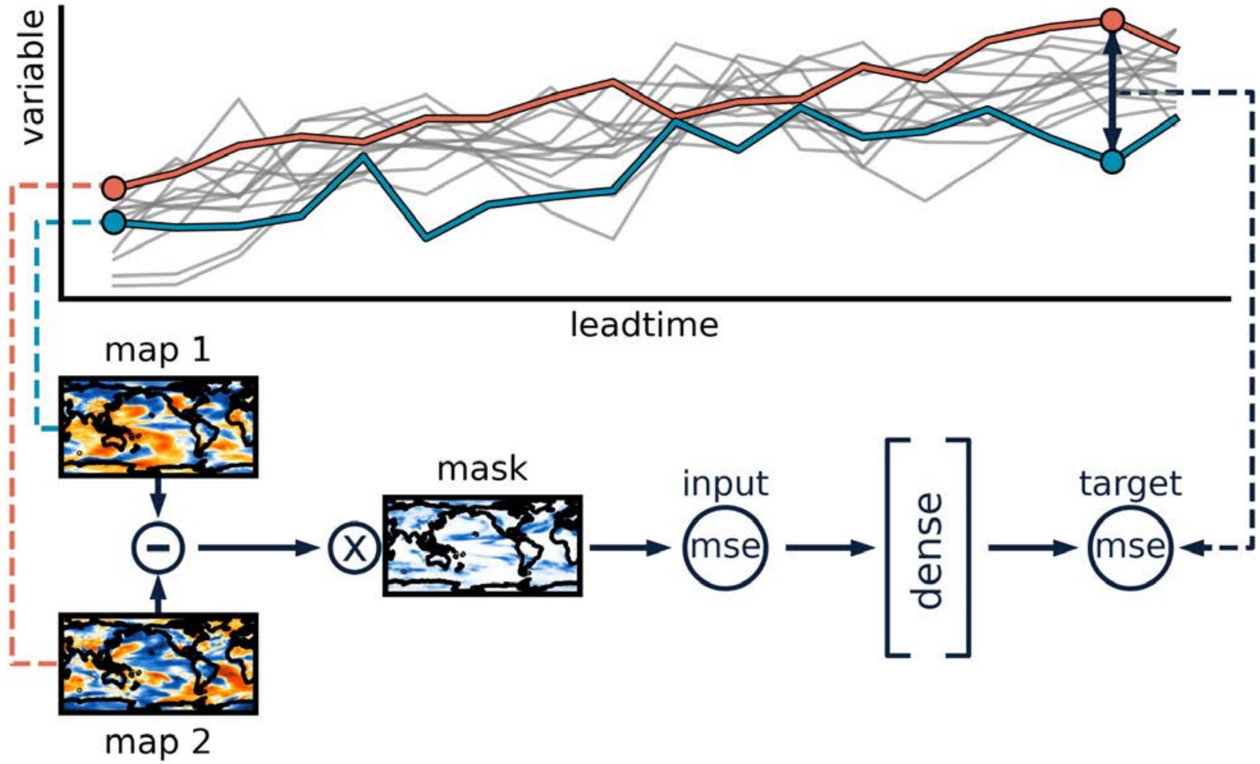


state of interest

potential analog

difference

This assumes that every region is equally important for determining how the climate system will evolve. **We can do better.**



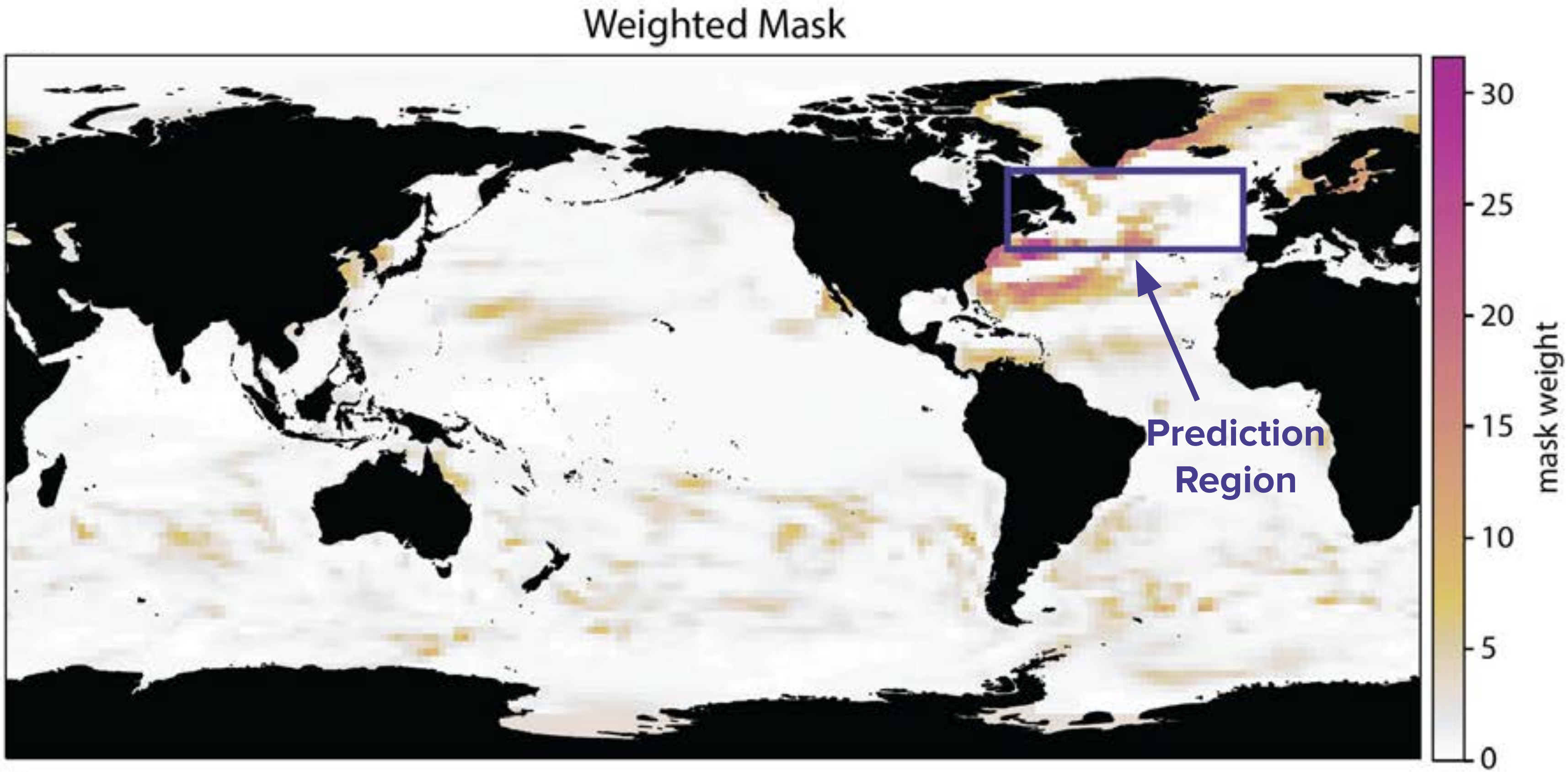
Use AI to learn regions most relevant for a "good analog"

Fernandez and Barnes (in prep)

Take this mask and use it to determine the “best analogs” in the standard, non-AI, way.

Use AI to learn regions most relevant for a “good analog”

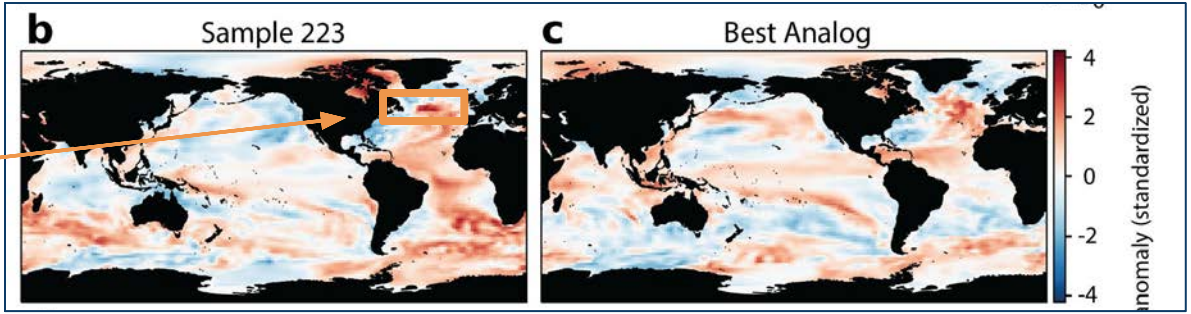
Fernandez and Barnes (in prep)



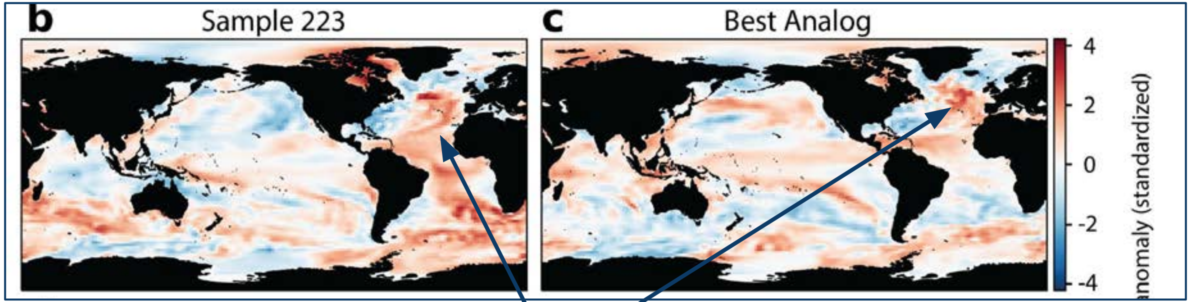
Use AI to learn regions most relevant for a “good analog”

Rader and Barnes (2024)

Predicting 5-yr SST in the North Atlantic



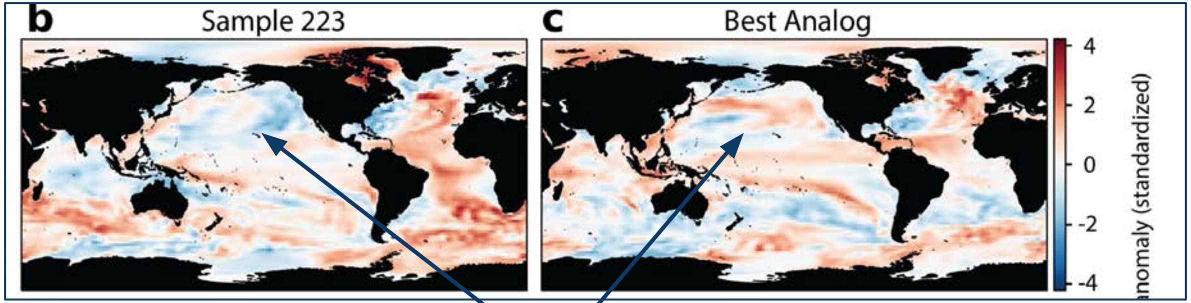
Identify the analogs with the weighted mask



Some regions look similar

Identify the analogs with the weighted mask

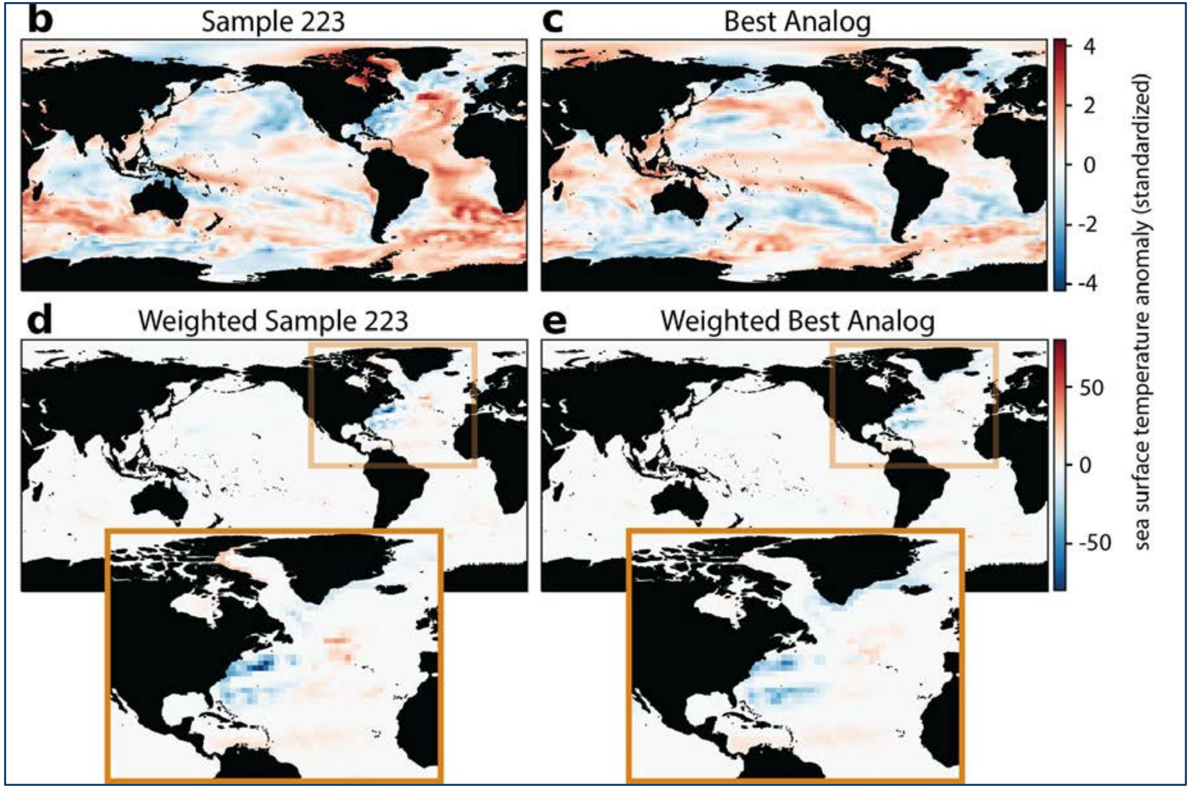
Rader and Barnes (2024)



Some regions look different

Identify the analogs with the weighted mask

Rader and Barnes (2024)

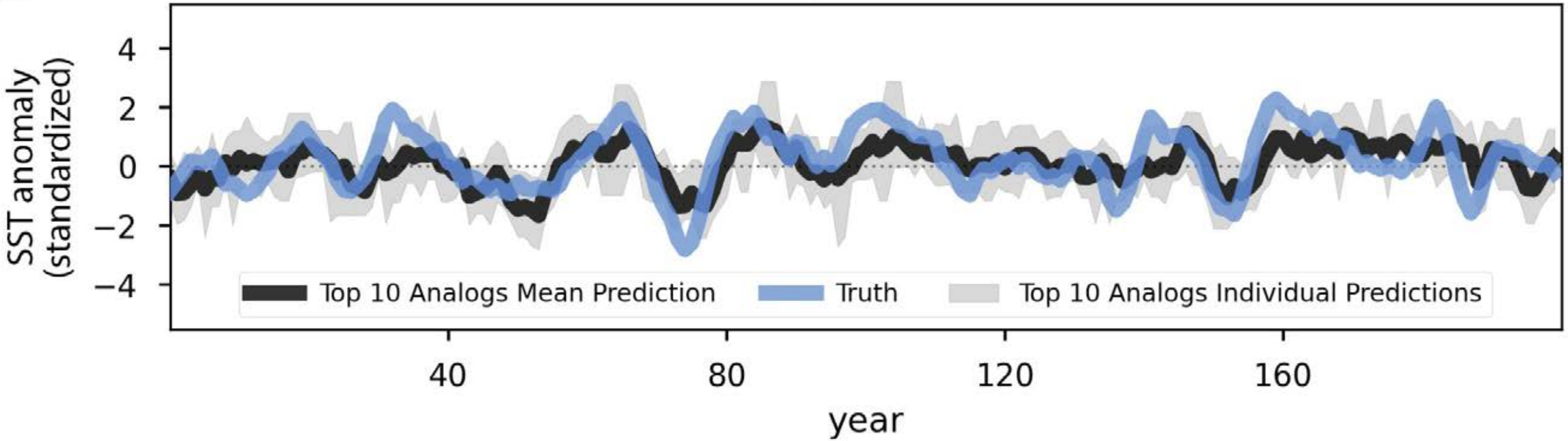


Maps look similar in the precursor regions

Identify the analogs

Rader and Barnes (2024)

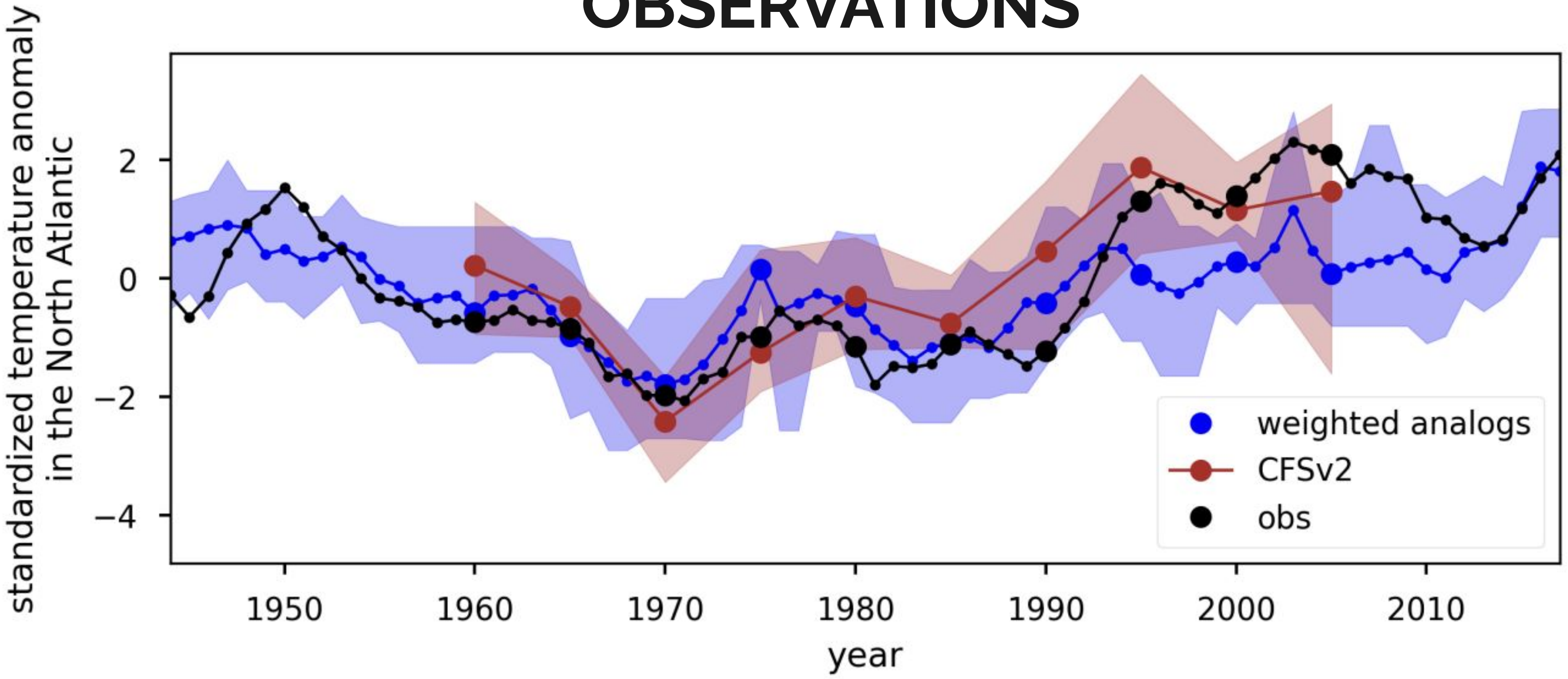
CLIMATE MODEL RESULTS



Forecasts are skillful

Rader and Barnes (2024)

OBSERVATIONS



These optimized analogs rival the skill of dynamical models

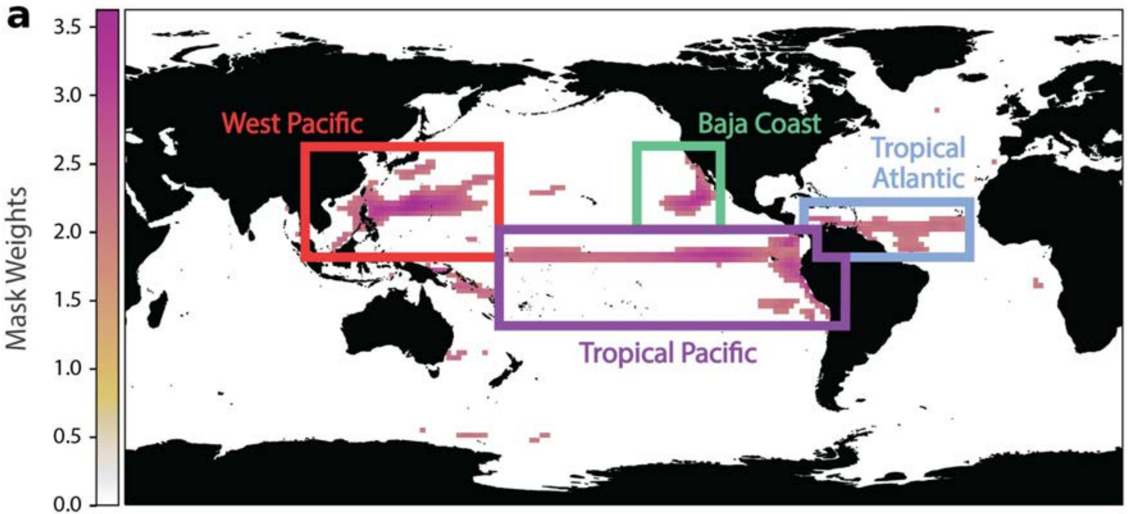
Rader and Barnes (2024)

This approach is interpretable AI!

While one could train a massive black-box AI model to make these predictions, the benefit of analogs is that they are easily understood.

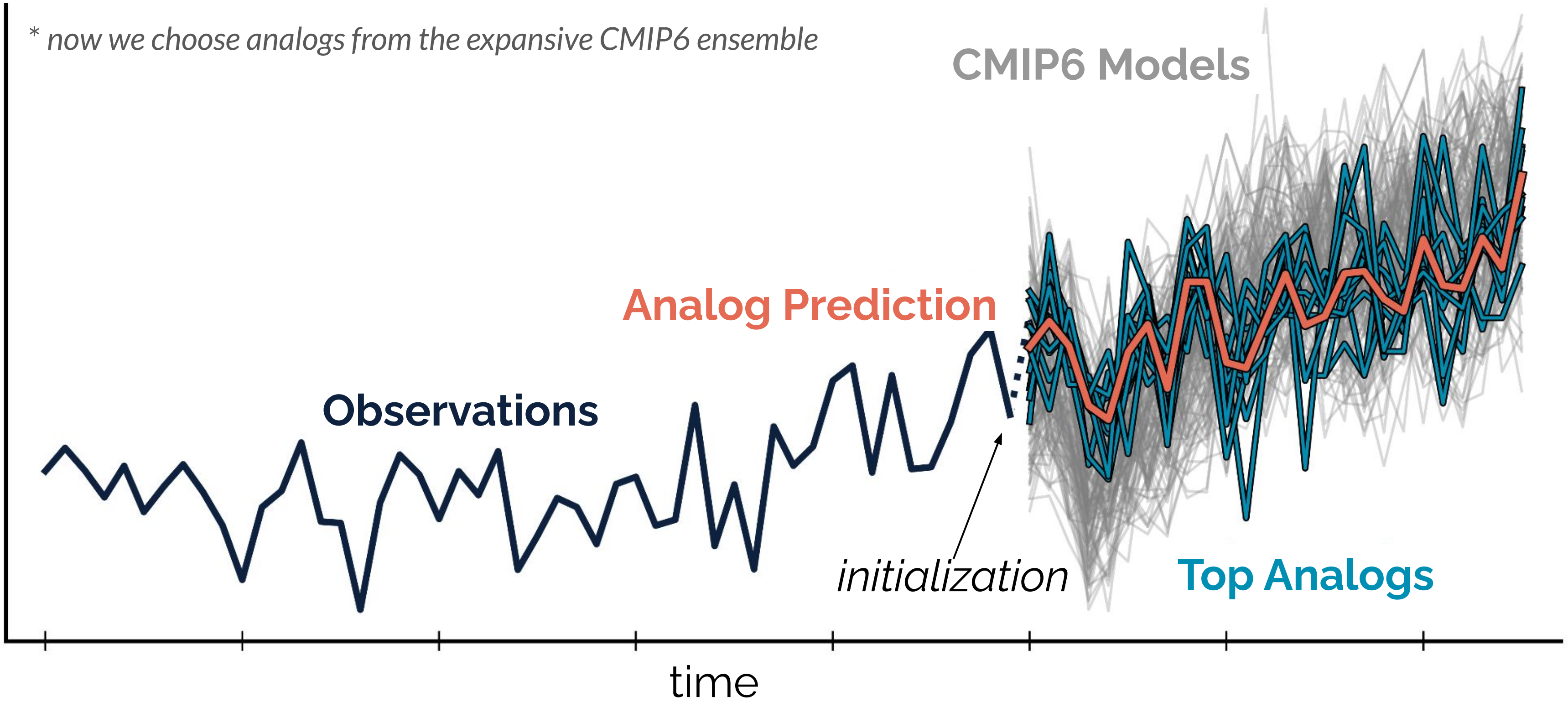
AI is merely used to learn the optimal mask for choosing the analogs. The masks are the “XAI”.

Learned ENSO Prediction Mask



Not a new Idea, but renewed interest!
e.g. Lorenz (1969),
Menary et al. (2021),
Lou et al. (2023),
Mahmood et al. (2021),
Befort et al. (2020),
Ding et al. (2018, 2019),
Toride et al. (2024)

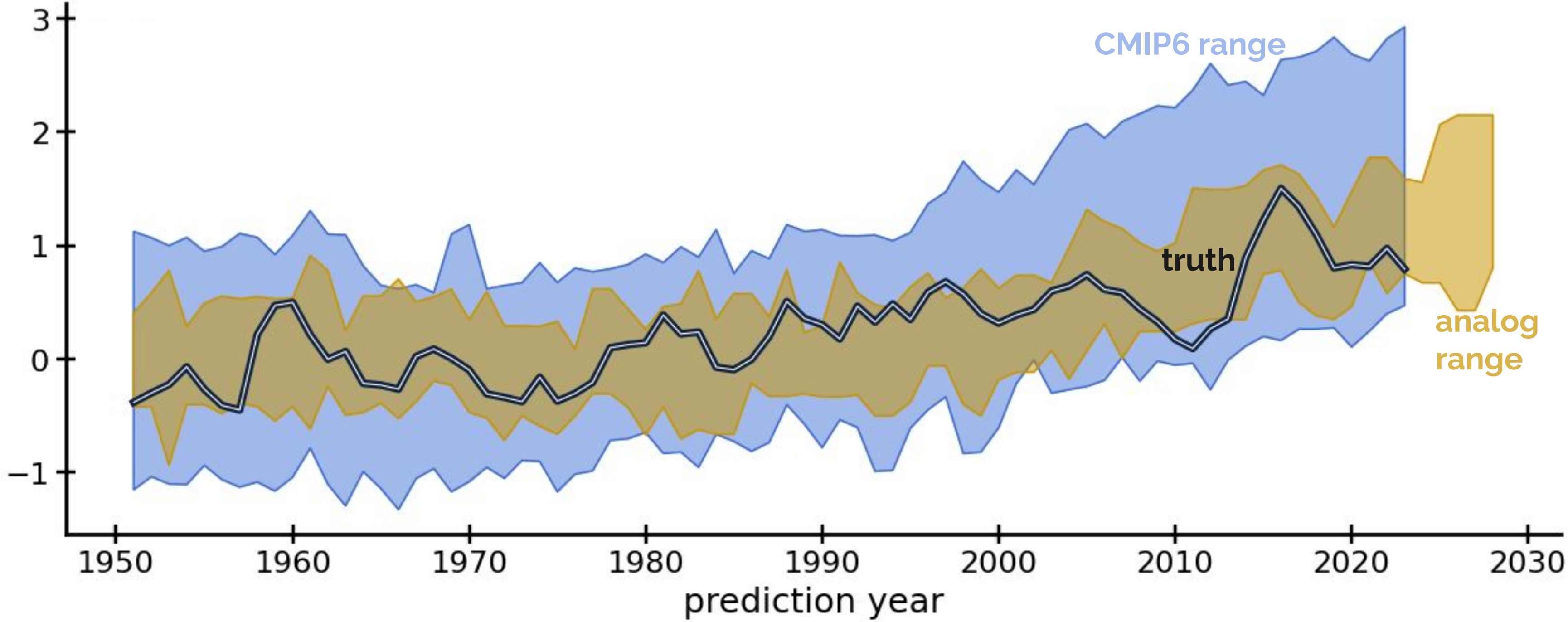
Rader and Barnes (2024)



Analogs can help constrain projections → initialized predictions

Fernandez and Barnes (in prep)

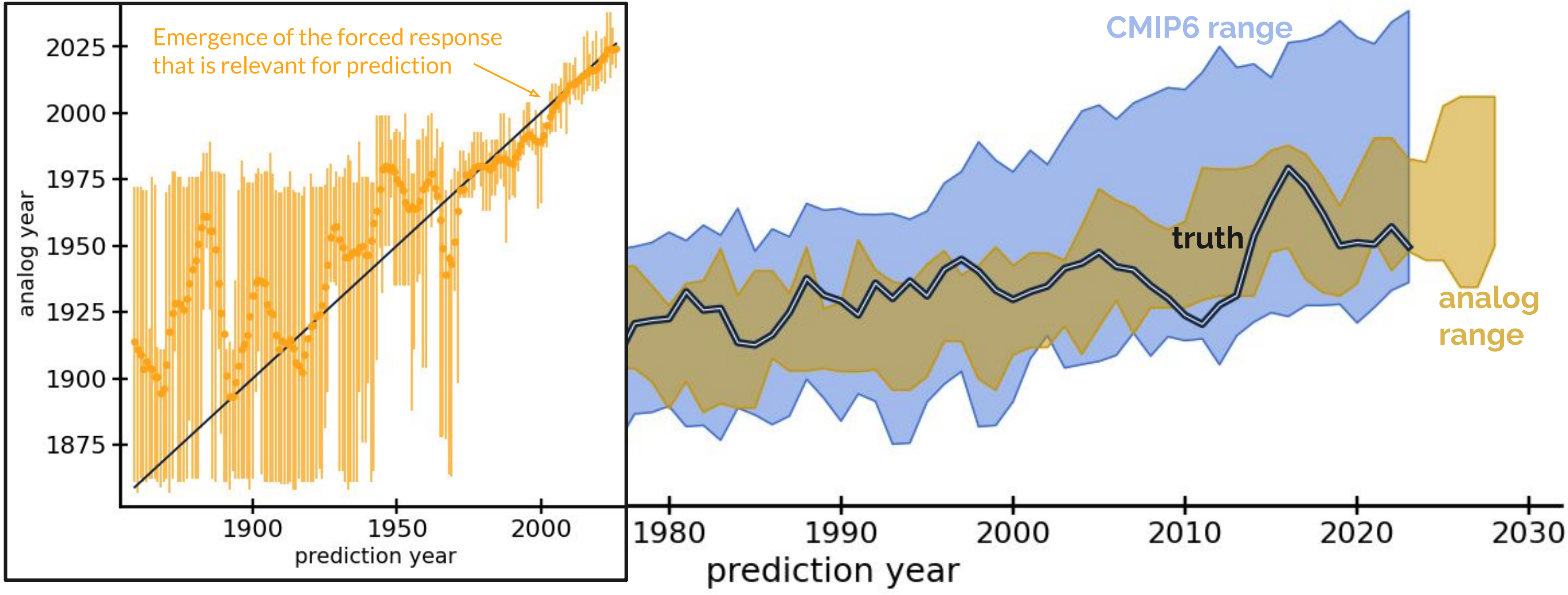
Yrs 5-7 Surface Temperature Anomalies Western U.S.



Analogs can help constrain projections → initialized predictions

Fernandez and Barnes (in prep)

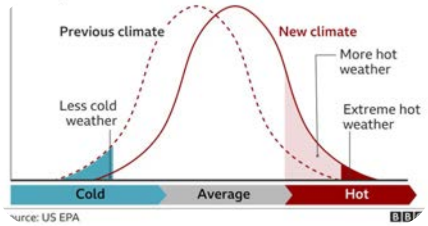
Yrs 5-7 Surface Temperature Anomalies
Western U.S.



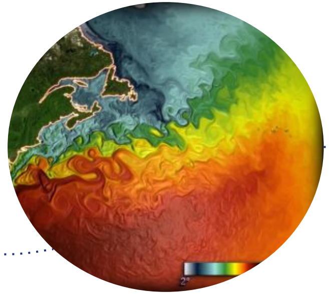
Analogs can help constrain projections → initialized predictions

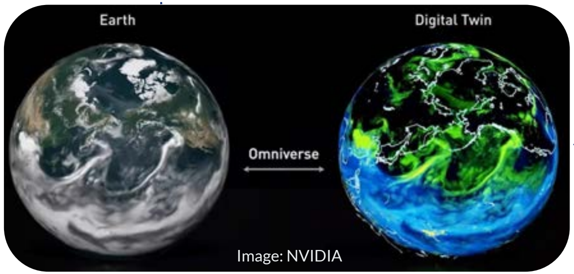
Fernandez and Barnes (in prep)

Climate change influence the performance of AI models



Generative AI for downscaling, ensembling, data assimilation...





AI for weather & climate model replacement

These are exciting times! XAI will be key to advancing science.

Thank you.

eabarnes@colostate.edu
https://barnes.atmos.colostate.edu
github: eabarnes1010