

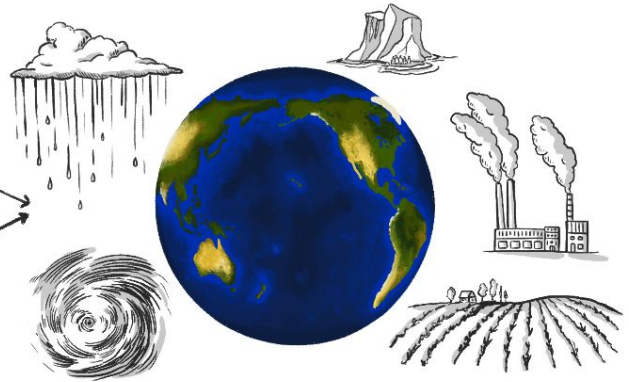
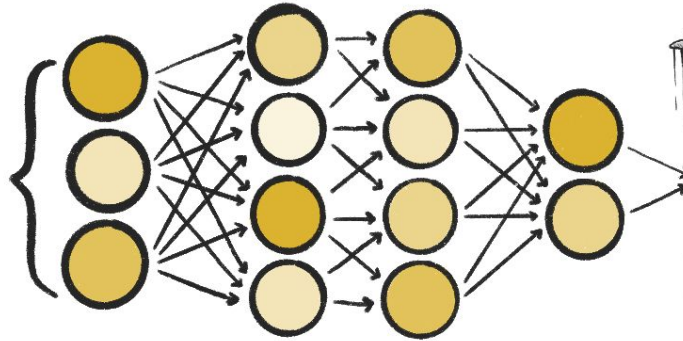
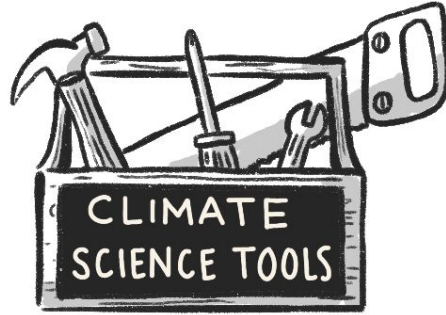
Explainable AI for Climate Projection, Multi-Year Prediction & Discovery

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



ATMOSPHERIC SCIENCE
COLORADO STATE UNIVERSITY

MACHINE
LEARNING



Do it faster/cheaper

Do it better

Learn something new

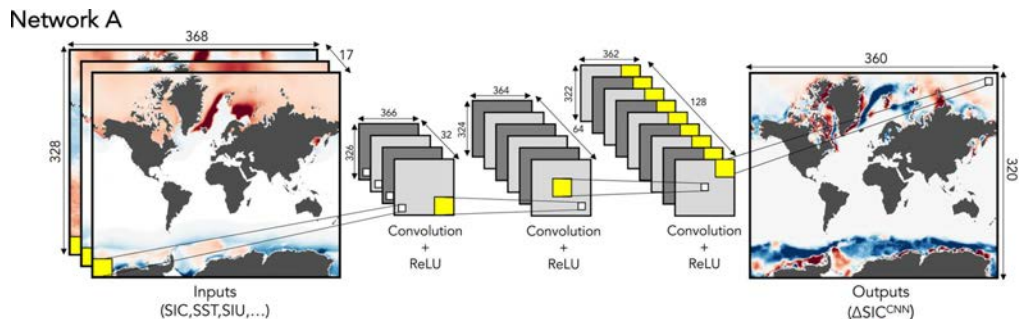
Machine Learning: a tool with many uses.

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

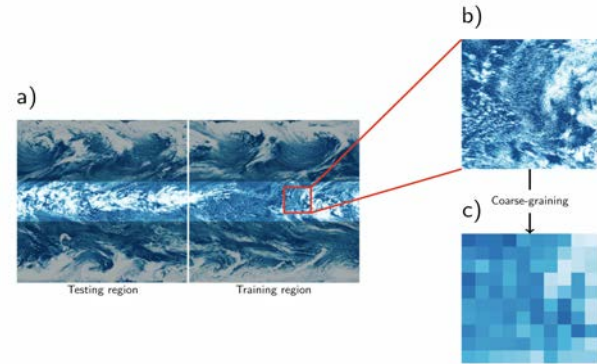
ML for post-processing data
[e.g. data compression, data analysis]

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

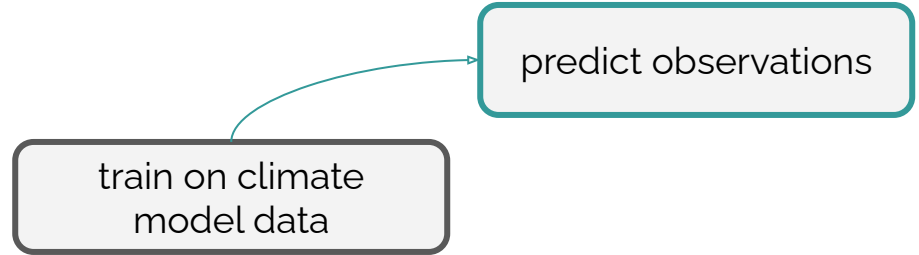
ML to improve climate models
[e.g. parameterizations]

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

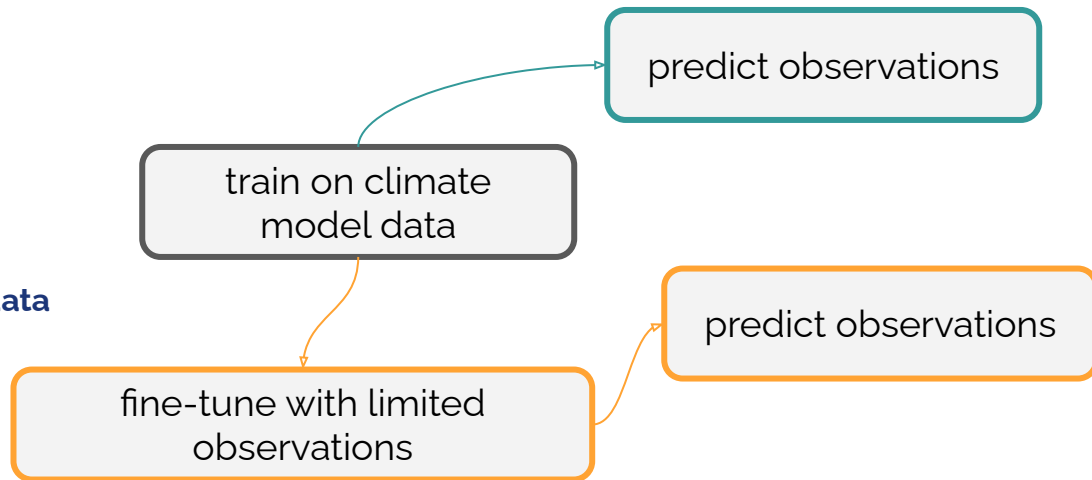
ML to improve climate models
[e.g. parameterizations]

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

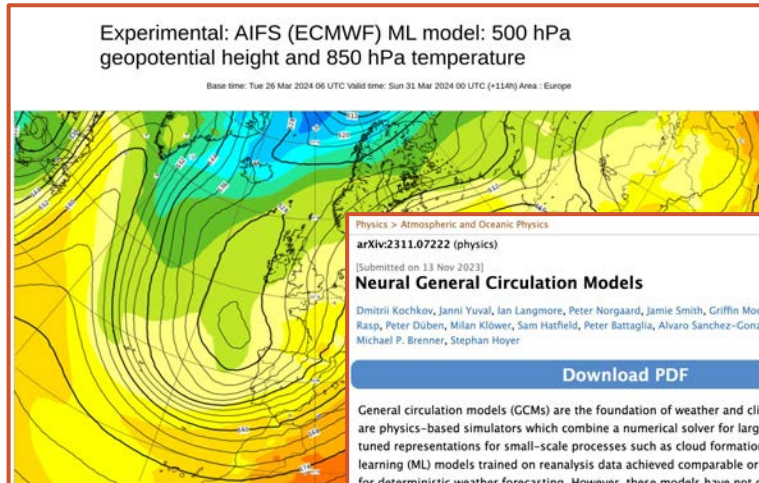
Combine disparate datastreams to
explore complex systems

4

Merging observations and model data

5

Deep-learning weather + climate
emulators



Physics > Atmospheric and Oceanic Physics

arXiv:2311.07222 (physics)

[Submitted on 13 Nov 2023]

Neural General Circulation Models

Dmitri 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

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 prediction long-term stability and physical consistency. We present ACE (AI2 Climate parameter, autoregressive machine learning emulator of an existing comprehensive resolution global atmospheric model. The formulation of ACE allows evaluation such as the conservation of mass and moisture. The emulator is stable for conserves column moisture without explicit constraints and faithfully reproduces model's climate, outperforming a challenging baseline on over 80% of tracks 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 near-seasonal forecasting

Jonathan A. Weyn, Divya Kumar, Jeremy Berman, Najeeb Kazmi, Sylwester Klocek, Pete Lufrenko, 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)

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

6

Climate change communication



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

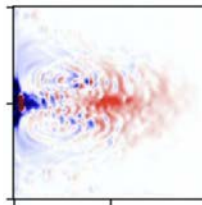
Deep-learning weather + climate
emulators

6

Climate change communication

7

Learn new things!



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

Equation Discovery

e.g. Zanna & Bolton (2020)

Machine Learning: a tool with many uses.

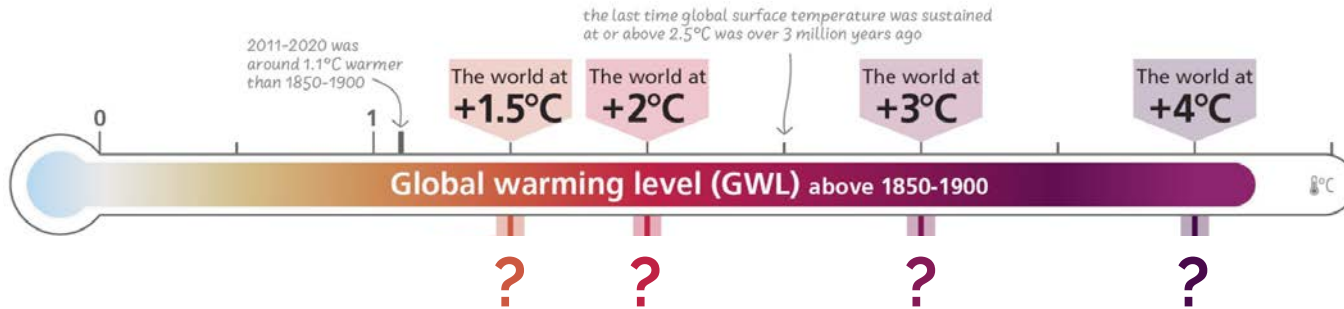
OUR GOAL:

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

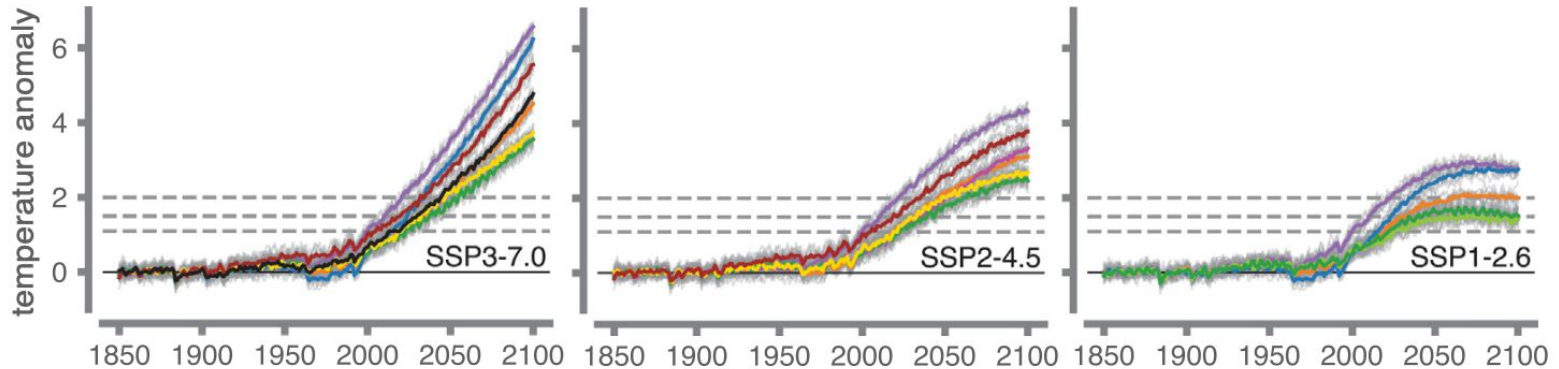
Climate models provide inaccurate, but invaluable “parallel universes” to mine for information



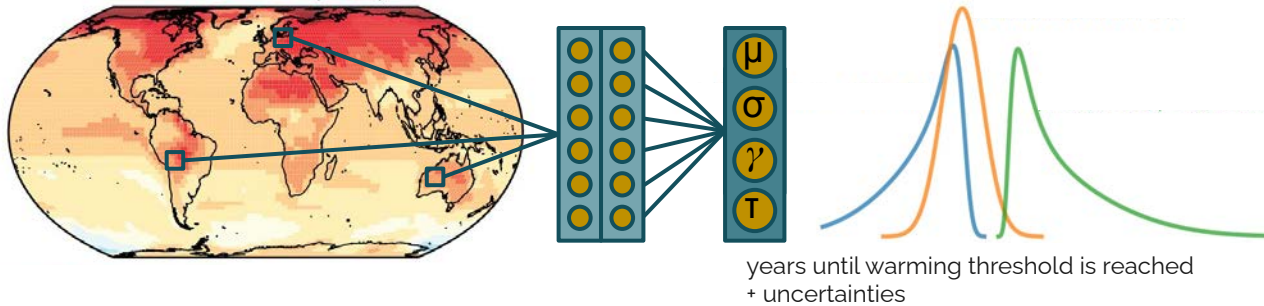
Time Remaining Until Critical Warming Thresholds are Reached



Time Remaining Until Critical Warming Thresholds are Reached



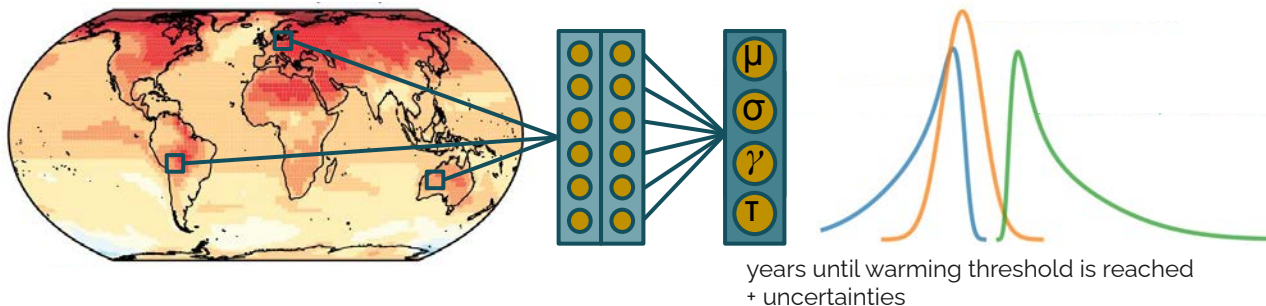
Trained on annual maps from 10 realizations from across multiple climate models



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

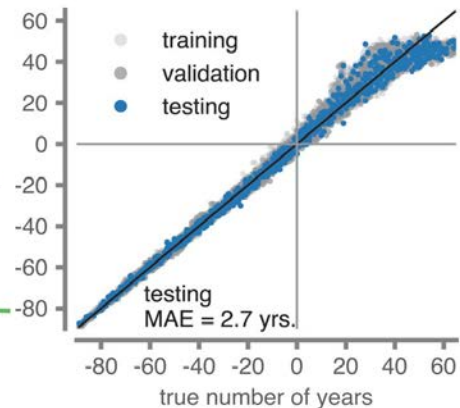


Trained on annual maps from 10 realizations from across multiple climate models



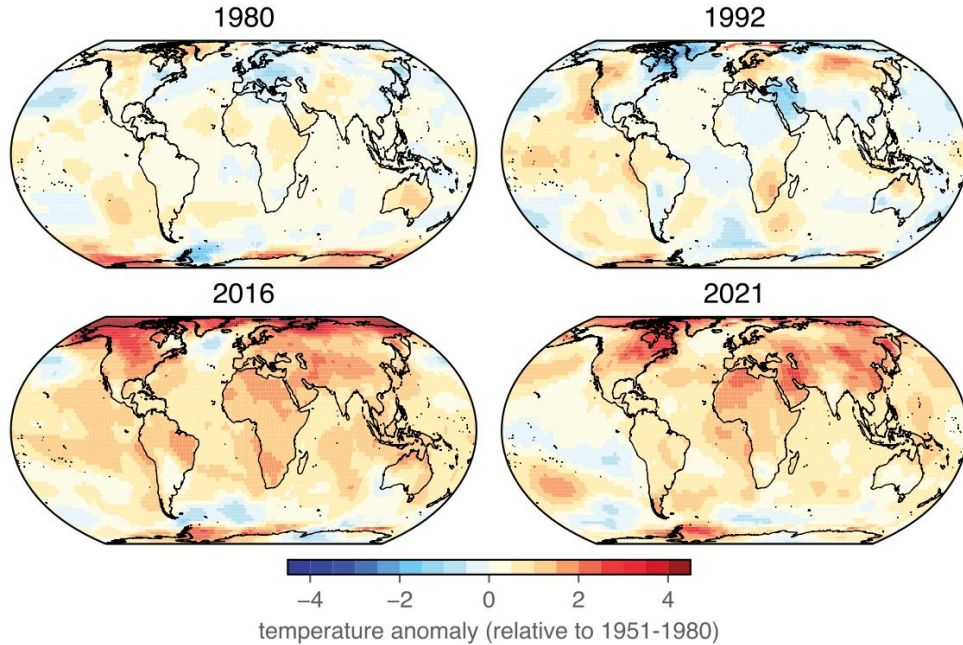
years until warming threshold is reached
+ uncertainties

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

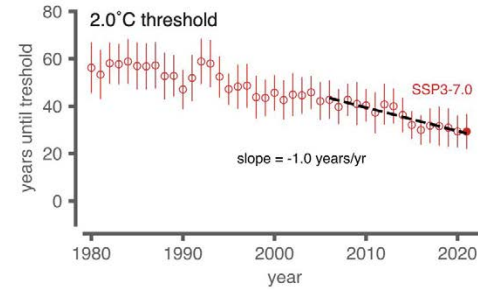
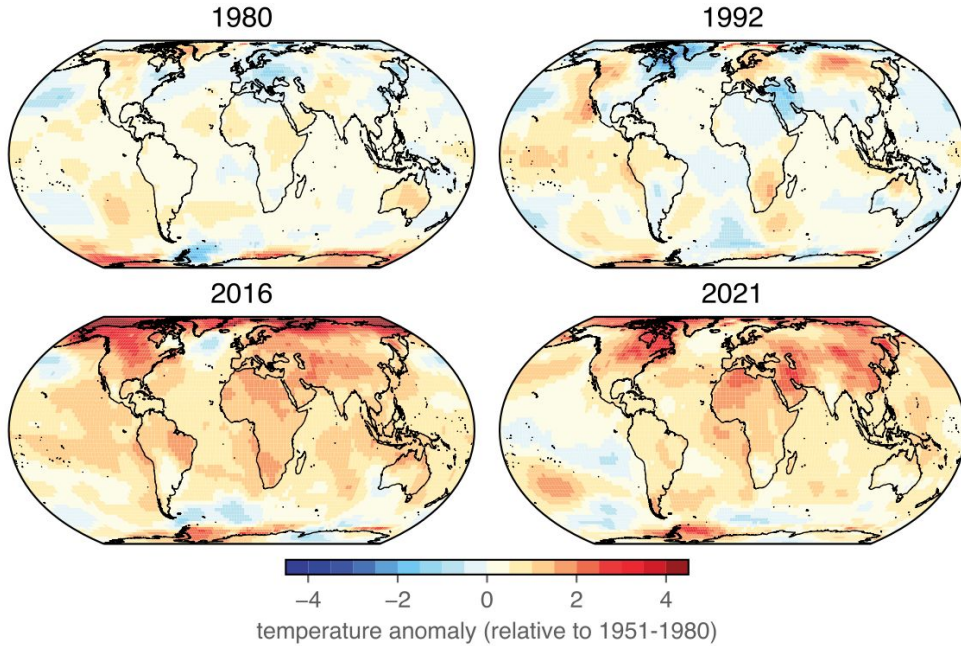




Observations
Berkeley Earth Surface Temperature

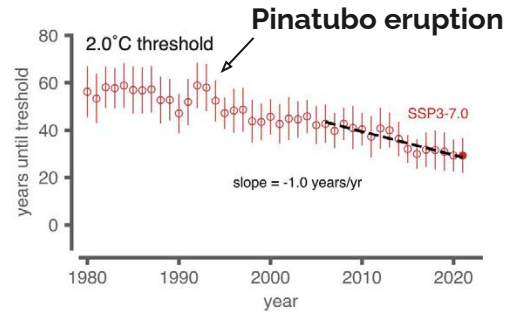
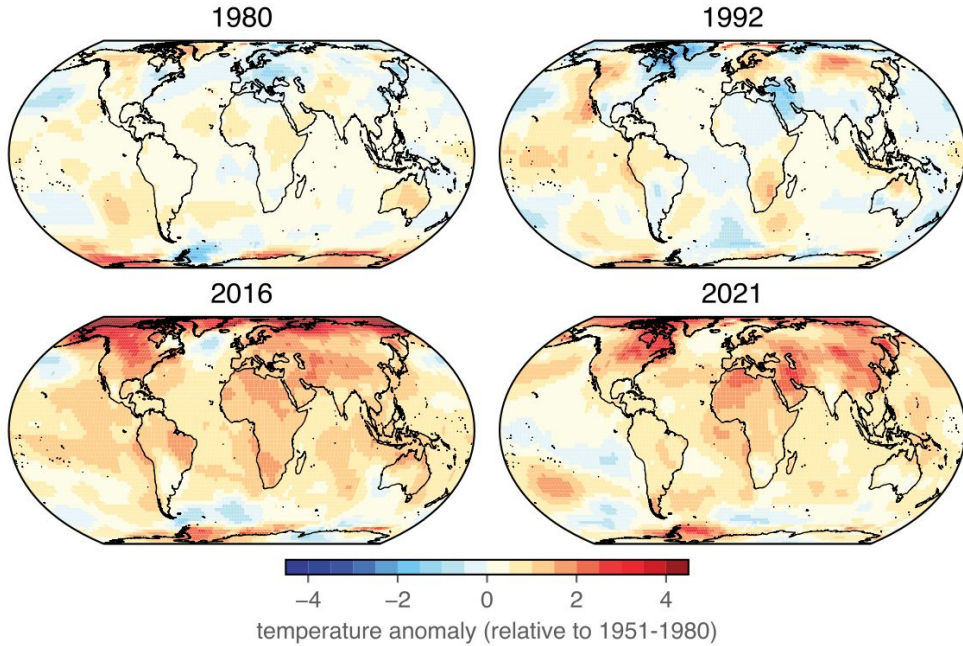
Use the trained AI model to predict thresholds based on maps of the observed climate





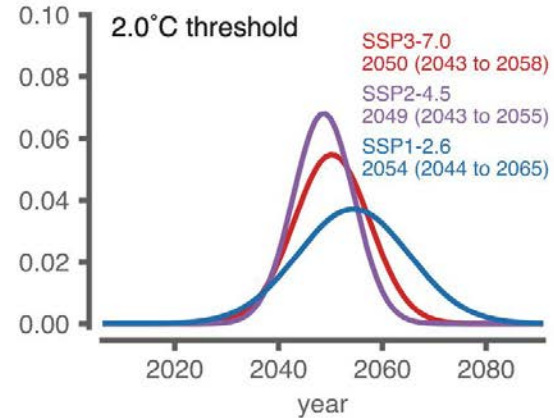
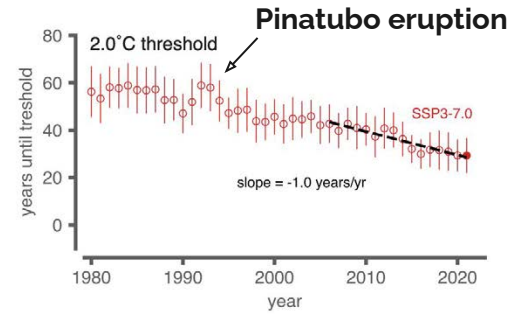
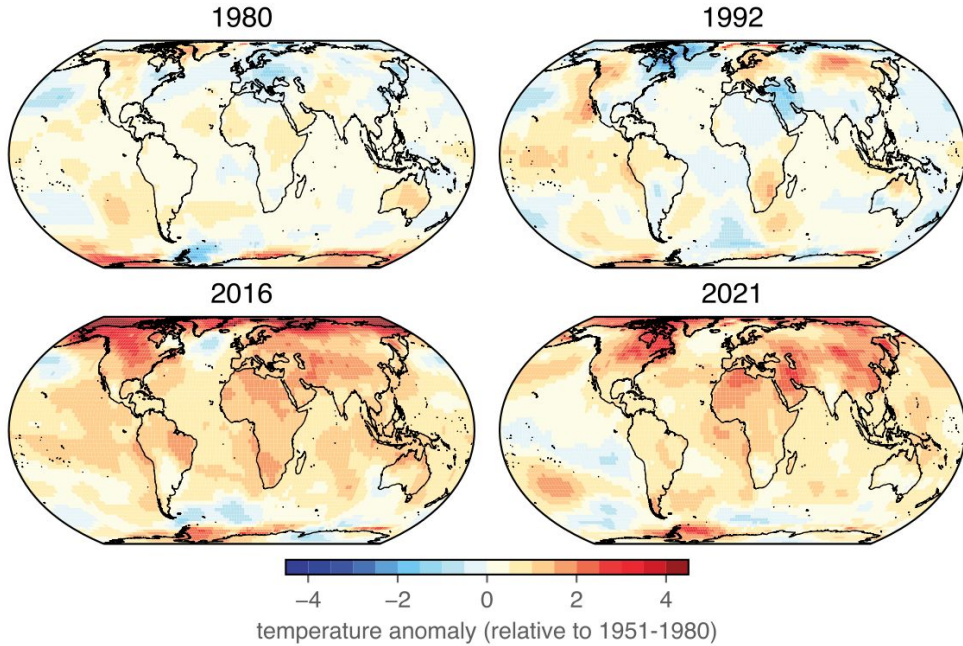
Use the trained AI model to predict thresholds based on maps of the observed climate





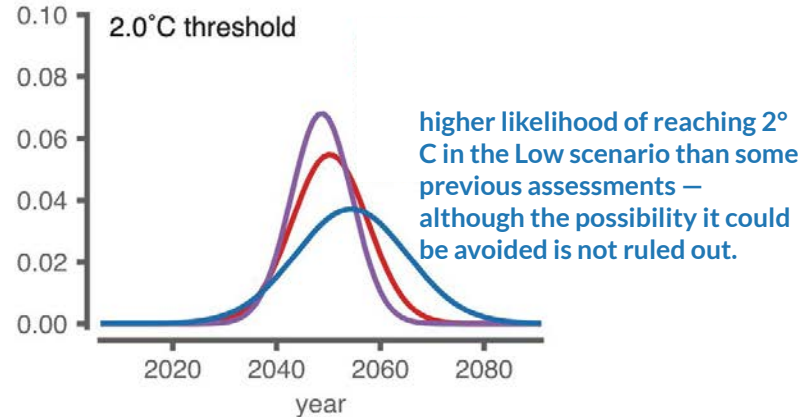
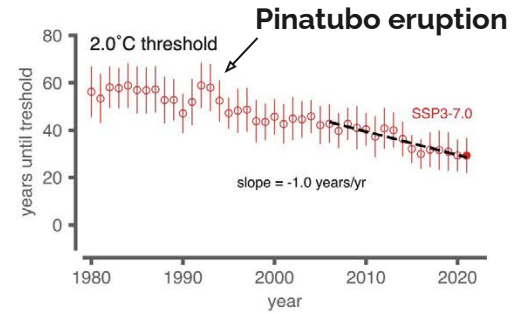
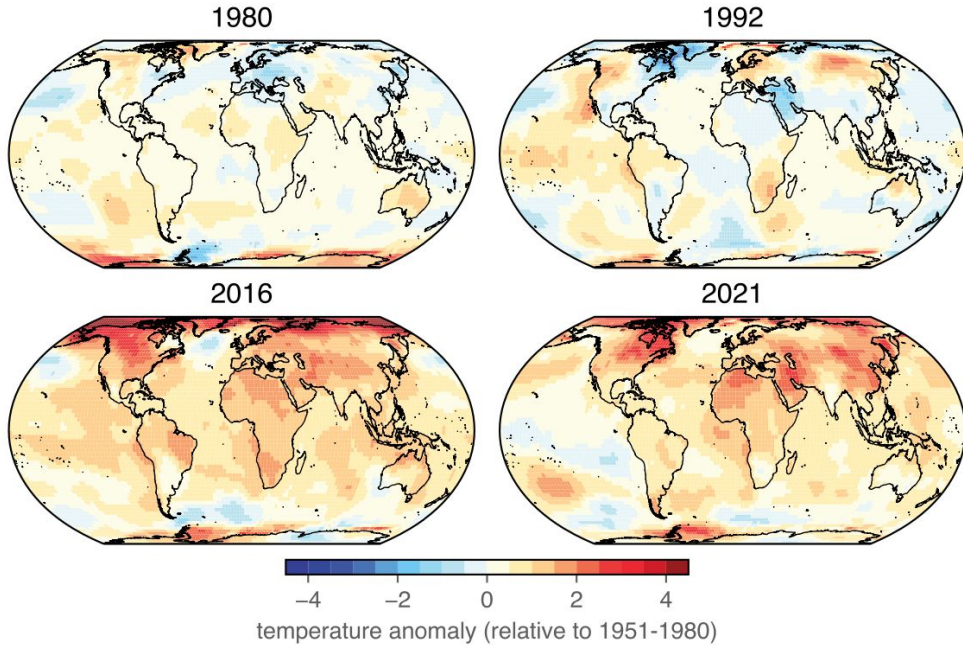
Use the trained AI model to predict thresholds based on maps of the observed climate





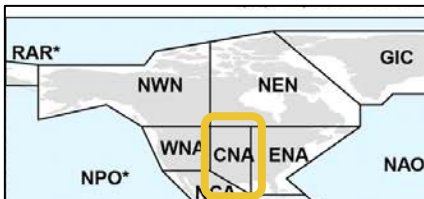
Use the trained AI model to predict thresholds based on maps of the observed climate



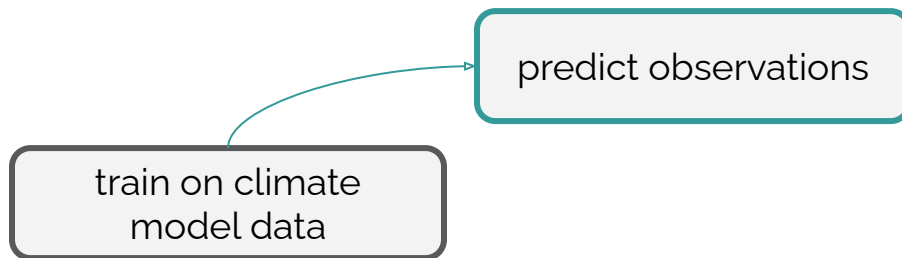


Use the trained AI model to predict thresholds based on maps of the observed climate



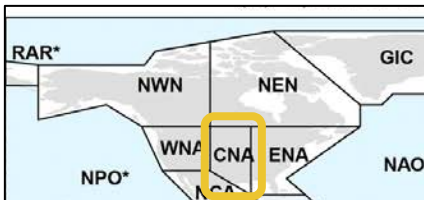


Transfer Learning

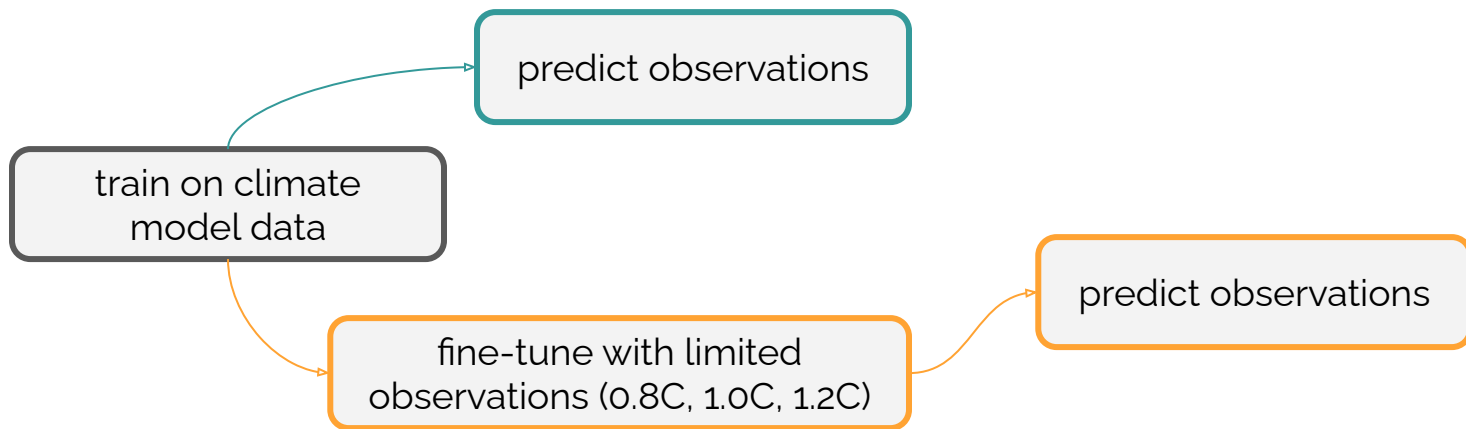


Regional transfer learning provides new insights



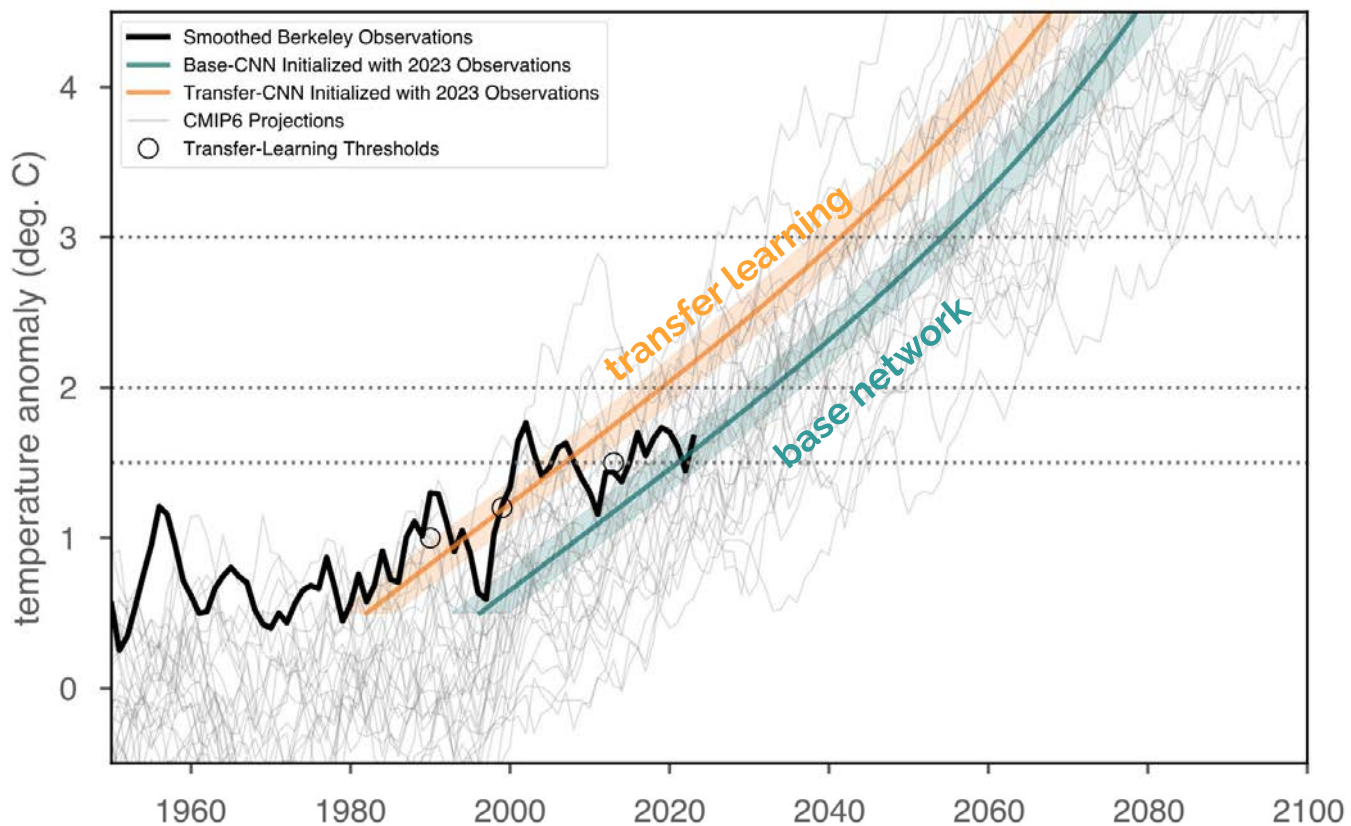
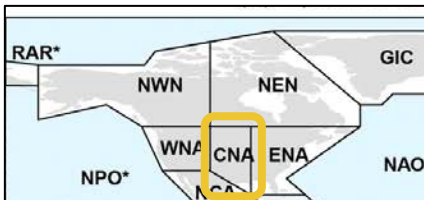


Transfer Learning



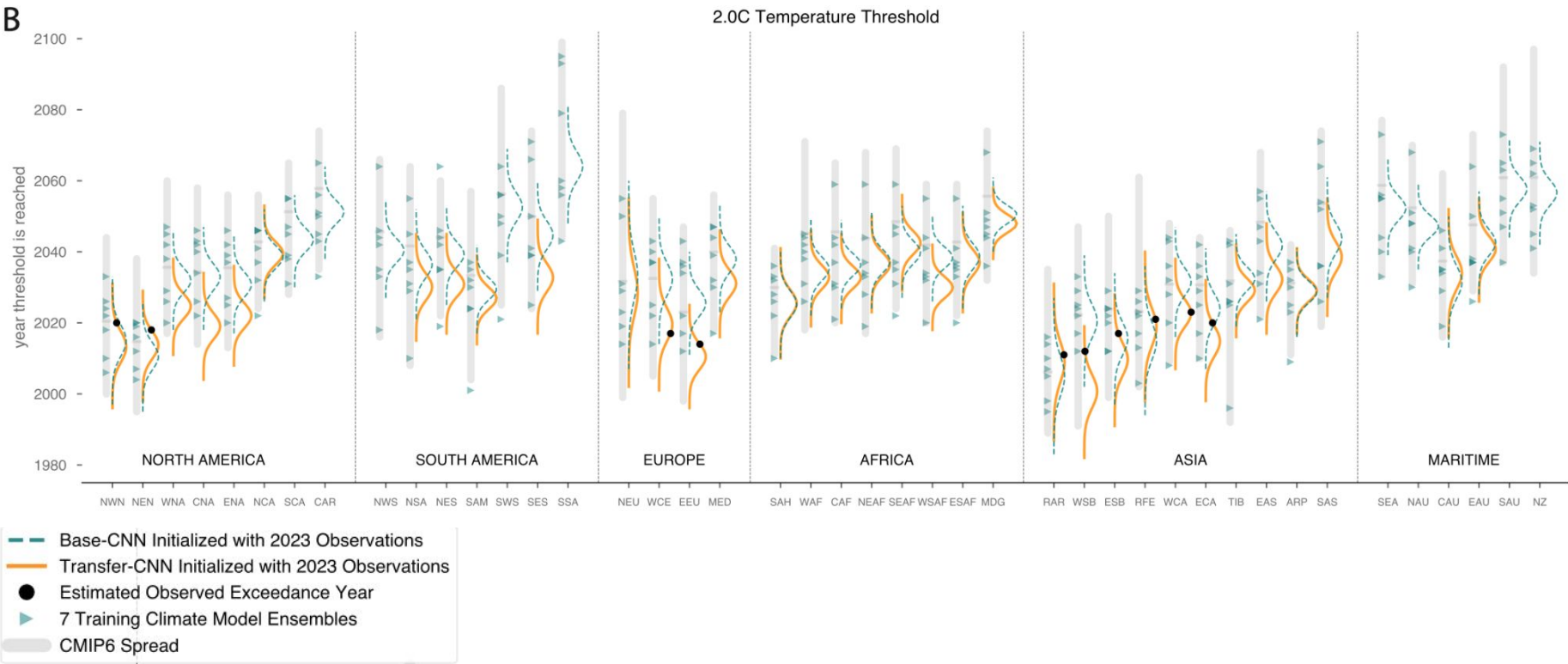
Regional transfer learning provides new insights





Regional transfer learning provides new insights



B

Regional transfer learning provides new insights



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

MAKING THE BLACK BOX MORE TRANSPARENT

Understanding the Physical Implications of Machine Learning

AMY MCGOVERN, RYAN LAGROZOFF, DAVID JOHN GAGNE II, G. EUJENGINEN, KIMBERLY L. ELMKE, CAMERON R. HOEYER, AND TRAVIS SMITH

Machine learning model interpretation and visualization focusing on meteorological domains are introduced and analyzed.

Machine learning (ML) and deep learning (DL; LeCun et al. 2015) have recently achieved breakthroughs across a variety of fields, including the world's best Go player (Silver et al. 2016, 2017), medical diagnosis (Rakhlenko et al. 2018), and galaxy

classification (Doleman et al. 2015). Simple forms of ML (e.g., linear regression) have been used in meteorology since at least the 1950s (Malone 1955), and ML has been used extensively to forecast convective hazards since the mid-1990s. Kitzmiller et al. (1995) use linear regression to forecast the probability of tornadoes, large hail, or damaging wind; Jillett et al. (1997) use linear regression to forecast hail probability and size; Marzban and Stumpf (1996, 1998) use neural networks to forecast the probability of tornadoes and damaging wind, respectively; and Marzban and Witt (2001) use neural networks to forecast hail size. Gagne et al. (2013, 2017a) use random forests to forecast hail probability at 1-day lead time; McGovern et al. (2014) and Williams (2014) use random forests

AFFILIATIONS: McGovern and Jenkinson—University of Oklahoma, Norman, Oklahoma; Lagrozoff—Cooperative Institute for Mesoscale Meteorological Studies, and University of Oklahoma, Norman, Oklahoma; Gagne—National Center for Atmospheric Research, Boulder, Colorado; Elmke and Smith—Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma; Hoyer—School of Meteorology, University of Oklahoma, Norman, Oklahoma
CORRESPONDING AUTHOR: Amy McGovern
amcgovern@ou.edu

Environmental Data Science (2023), 1, e8, 1–17
doi:10.1017/eds.2023.7

METHODS PAPER

Neural network attribution methods for problems in geoscience: A novel synthetic benchmark dataset

Antonis Mamelakis^{1,*}, Imme Ebert-Uphoff^{2,3}, and Elizabeth A. Barnes¹

¹Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA
²Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, Colorado, USA
³Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado, USA
*Corresponding author. E-mail: amamelakis@rams.colostate.edu

Received: 29 November 2021; Revised: 05 April 2022; Accepted: 28 April 2022

Keywords: attribution benchmark; Explainable Artificial Intelligence; geosciences; ground truth; neural networks; regression problems

Abstract

Despite the increasingly successful application of neural networks to many problems in the geosciences, their complex and nonlinear structure makes the interpretation of their predictions difficult, which limits model trust and does not allow scientists to gain physical insights about the problem at hand. Many different methods have been introduced in the emerging field of Explainable Artificial Intelligence (XAI), which aims at attributing the network's prediction to specific features in the input domain. XAI methods are usually assessed by using benchmark datasets (such as MNIST or ImageNet for image classification). However, an objective, theoretically derived ground truth for the attribution is lacking for most of these datasets, making the assessment of XAI in many cases subjective. Also, benchmark datasets specifically designed for problems in geosciences are rare. Here, we provide a framework, based on the use of additively separable functions, to generate attribution benchmark datasets for regression problems for which the ground truth of the attribution is known a priori. We generate a large benchmark dataset and train a fully connected network to learn the underlying function that was used for simulation. We then compare estimated heatmaps from different XAI methods to the ground truth in order to identify examples where specific XAI methods perform well or poorly. We believe that attribution benchmarks as the ones introduced herein are of great importance for further application of neural networks in the geosciences, and for more objective assessment and accurate implementation of XAI methods, which will increase model trust and assist in discovering new science.

JULY 2024
BOMMER ET AL.
Finding the Right XAI Method—A Guide for the Evaluation and Ranking of Explainable AI Methods in Climate Science
PHILINE LOU BOMMER,^{1,2} MARLENE KRETSCHMER,^{1,2} ANNA HEDSTRÖM,^{2,3} DILYARA BAREEVA,^{2,4} MARINA M.C. HOHN^{2,5,6,7}

¹ *Understandable Machine Intelligence Lab, Technical University Berlin, Berlin, Germany*
² *Department of Data Science, ATR, Potsdam, Germany*
³ *Leipzig Institute for Meteorology, University of Leipzig, Leipzig, Germany*
⁴ *Department of Meteorology, University of Reading, Reading, United Kingdom*
⁵ *Institute of Computer Science – University of Potsdam, Potsdam, Germany*
⁶ *Berlin Institute for the Foundations of Learning and Data, Berlin, Germany*
⁷ *Machine Learning Group, UiT The Arctic University of Norway, Tromsø, Norway*

(Manuscript received 25 August 2023, in final form 8 March 2024, accepted 19 March 2024)

ABSTRACT: Explainable artificial intelligence (XAI) methods shed light on the predictions of machine learning algorithms. Several different approaches exist and have already been applied in climate science. However, usually missing ground truth explanations complicate their evaluation and comparison, subsequently impeding the choice of the XAI method. Therefore, in this work, we introduce XAI evaluation in the climate context and discuss different desired explanation properties, namely, robustness, faithfulness, randomization, complexity, and localization. To this end, we chose previous work as a case study where the decade of annual-mean temperature maps is predicted. After training both a multilayer perceptron (MLP) and a convolutional neural network (CNN), multiple XAI methods are applied and their skill scores in reference to a random uniform explanation are calculated for each property. Independent of the network, we find the XAI methods such as Integrated Gradients, layerwise relevance propagation, and input times gradients exhibit considerable robustness, faithfulness, and complexity while sacrificing randomization performance. Sensitivity methods, gradient SmoothGrad, NoiseGrad, and FusionGrad, match the robustness skill but sacrifice faithfulness and complexity for the randomization skill. We find architecture-dependent performance differences regarding robustness, complexity, and localization skills of different XAI methods, highlighting the necessity for research task-specific evaluation. Overall, our work offers an overview of different evaluation properties in the climate science context and shows how to compare and benchmark different explanation methods, assessing their suitability based on strengths and weaknesses, for the specific research problem at hand. By that, we aim to support climate researchers in the selection of a suitable XAI method.

SIGNIFICANCE STATEMENT: Explainable artificial intelligence (XAI) helps to understand the reasoning behind the prediction of a neural network. XAI methods have been applied in climate science to validate networks and provide new insight into physical processes. However, the increasing number of XAI methods can overwhelm practitioners, making it difficult to choose an explanation method. Since XAI methods' results can vary, uniform choices might cause misleading conclusions about the network decisions. In this work, we introduce XAI evaluation to compare and assess the performance of explanation methods based on five desirable properties. We demonstrate that XAI evaluation reveals the strengths and weaknesses of different XAI methods. Thus, our work provides climate researchers with the tools to compare, analyze, and subsequently choose explanation methods.

Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability

Benjamin A. Toner¹, Elizabeth A. Barnes^{2,3}, and Imme Ebert-Uphoff^{2,3}

¹ *Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA*
² *Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO, USA*
³ *Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA*

Abstract: Neural networks have become increasingly prevalent within the geosciences, although a common limitation of their usage has been a lack of methods to interpret what the networks learn and how they make decisions. As such, neural networks have often been used within the geosciences to most accurately identify a desired output given a set of inputs, with the interpretation of what the network learns used as a secondary metric to ensure the network is making the right decision for the right reason. Neural network interpretation techniques have become more advanced in recent years, however, and we therefore propose that the ultimate objective of using a neural network can also be the interpretation of what the network has learned rather than the output itself. We show that the interpretation of neural networks can enable the discovery of scientifically meaningful connections within geoscience problems. In particular, we use two methods for neural network interpretation called backword optimization and layerwise relevance propagation, both of which project the decision pathways of a network back onto the original input dimensions. To the best of our knowledge, LRP has not yet been applied to geoscientific research, and we believe it has great potential in this area. We show how these interpretation techniques can be used to initially identify scientifically meaningful information from neural networks to apply them to common climate patterns. These results suggest that combining interpretable neural networks with novel scientific hypotheses will open the door to many new areas of neural network related geoscience research.

Plain Language Summary: Neural networks, a form of machine learning, have become

BAMS ISSUES EARLY ONLINE RELEASE COLLECTIONS FOR AUTHORS

RESEARCH ARTICLE | 31 AUGUST 2020
Evaluation, Tuning and Interpretation of Neural Networks for Working with Images in Meteorological Applications
Imme Ebert-Uphoff and Elizabeth A. Barnes
Bull. Amer. Meteor. Soc. 1–18
<https://doi.org/10.1175/BAMS-D-20-0097.1>

Split-Screen PDF © Share © Cite © Get Permissions

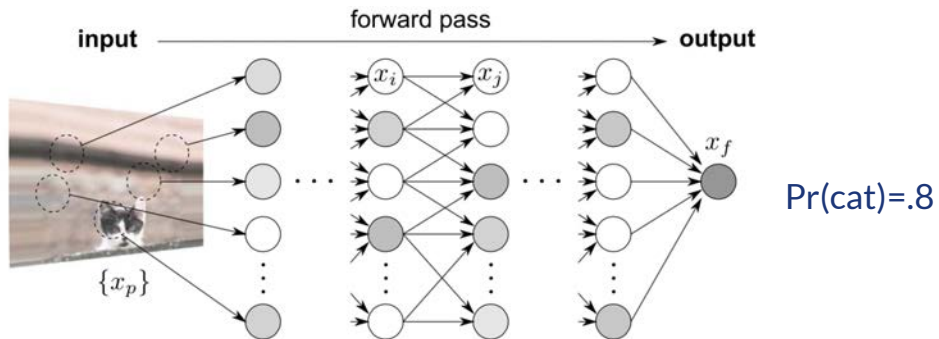
Capsule:
This article discusses strategies for the development of neural networks (aka deep learning) for meteorological applications. Topics include evaluation, tuning and interpretation of neural networks for working with meteorological images.

Abstract:
The method of neural networks (aka deep learning) has opened up many new opportunities to utilize remotely sensed images in meteorology. Common applications include image classification, e.g., to determine whether an image contains a tropical cyclone, and image-to-image translation, e.g., to emulate radar imagery for satellites that only have passive channels. However, there are yet many open questions regarding the use of neural networks for working with meteorological images, such as best practices for evaluation, tuning and interpretation. This article highlights several strategies and practical considerations for neural network development that have not yet received much attention in the meteorological community, such as the concept of receptive fields, underutilized meteorological performance measures, and methods for neural network interpretation, such as synthetic experiments and layer-wise relevance propagation. We also consider the process of neural network interpretation as a whole, recognizing it as an iterative meteorologist-driven discovery process that builds on experimental design and hypothesis generation and testing. Finally, while most work on neural network interpretation in meteorology has so far focused on networks for image classification tasks, we expand the focus to also include networks for image-to-image translation.

XAI Attribution Methods

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

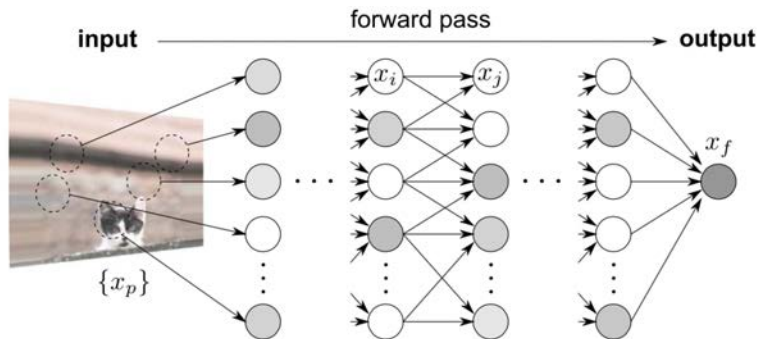
Prediction
of 1 sample



XAI Attribution Methods

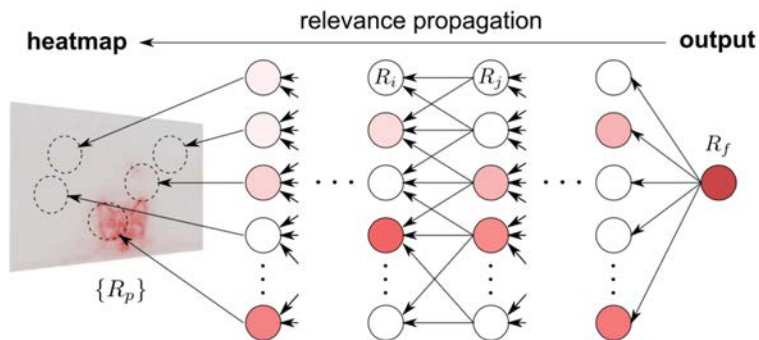
Attribution heatmaps are largely consistent with how many climate scientists pose questions

Prediction
of 1 sample



$\Pr(\text{cat})=.8$

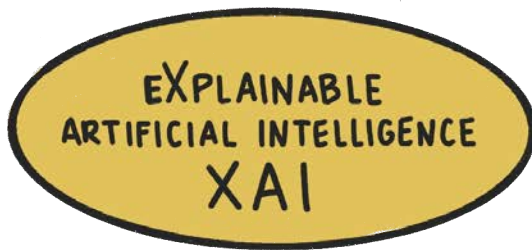
Attribution
of 1 sample



$\Pr(\text{cat})=.8$

Reasons to care about XAI

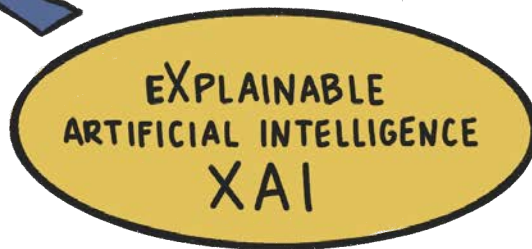
A scientist's ultimate goal is typically to understand "why?", but even if you don't care "why?" you should still care about XAI.



Reasons to care about XAI

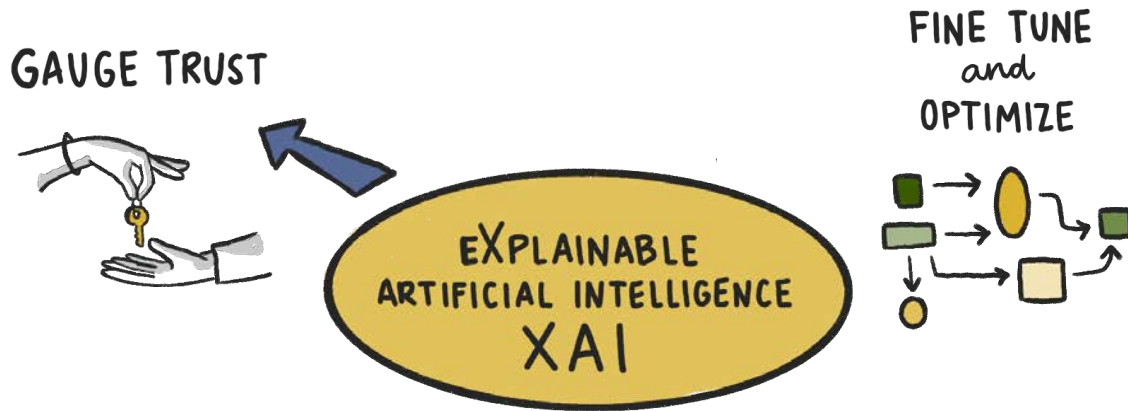
A scientist's ultimate goal is typically to understand "why?", but even if you don't care "why?" you should still care about XAI.

GAUGE TRUST



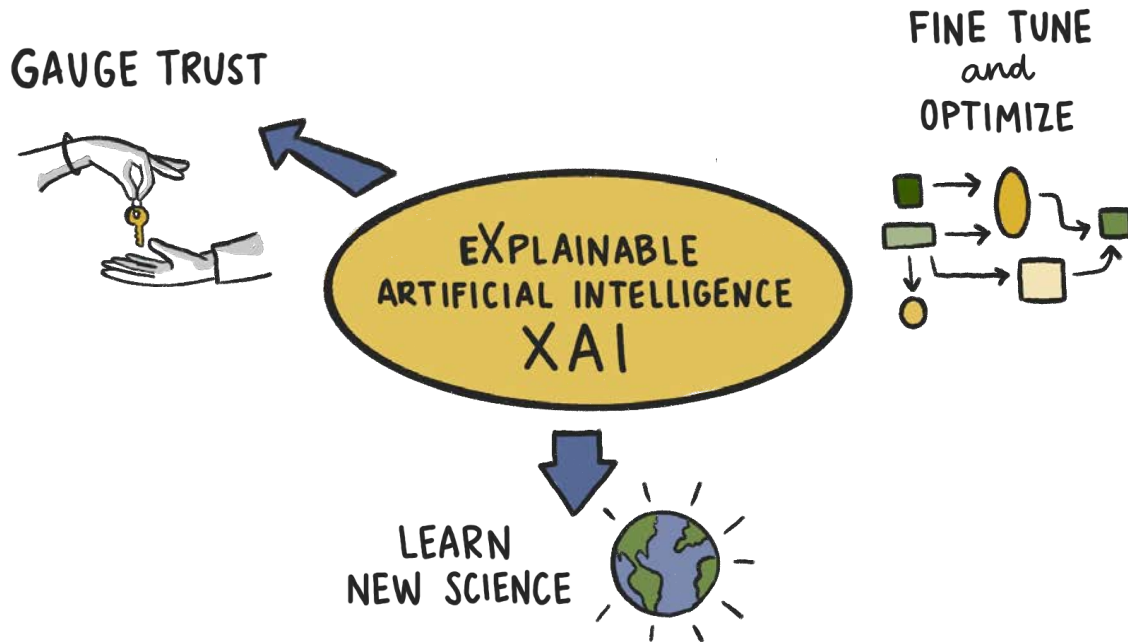
Reasons to care about XAI

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



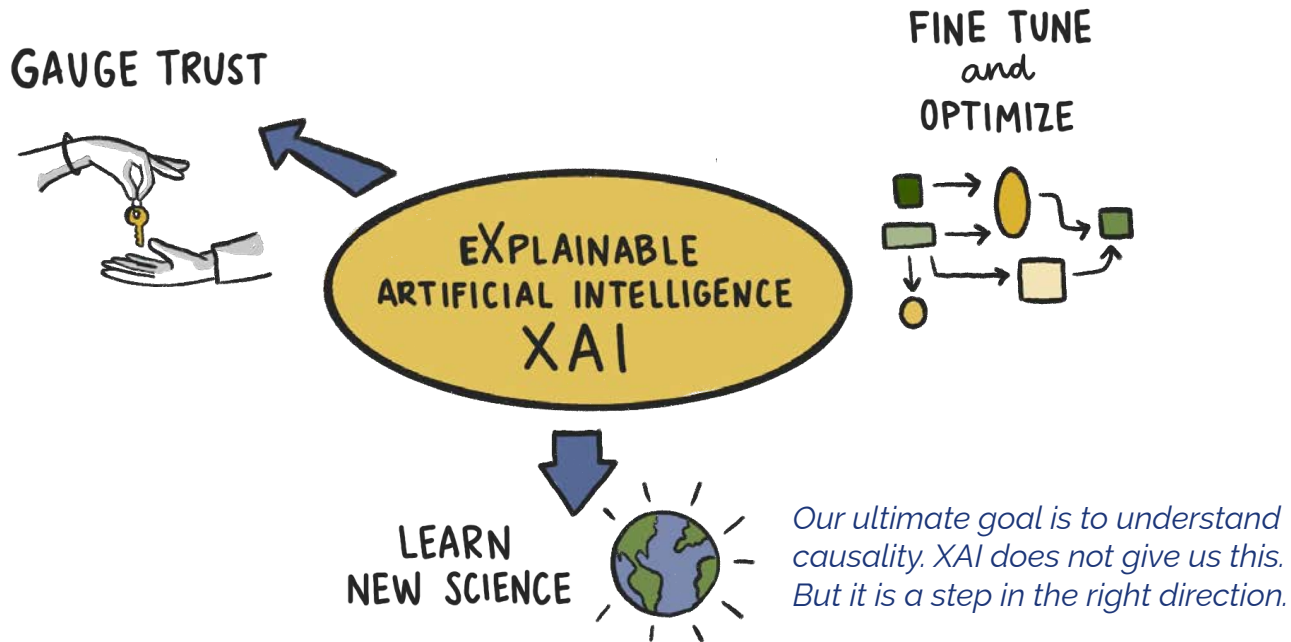
Reasons to care about XAI

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.



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 |



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:

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

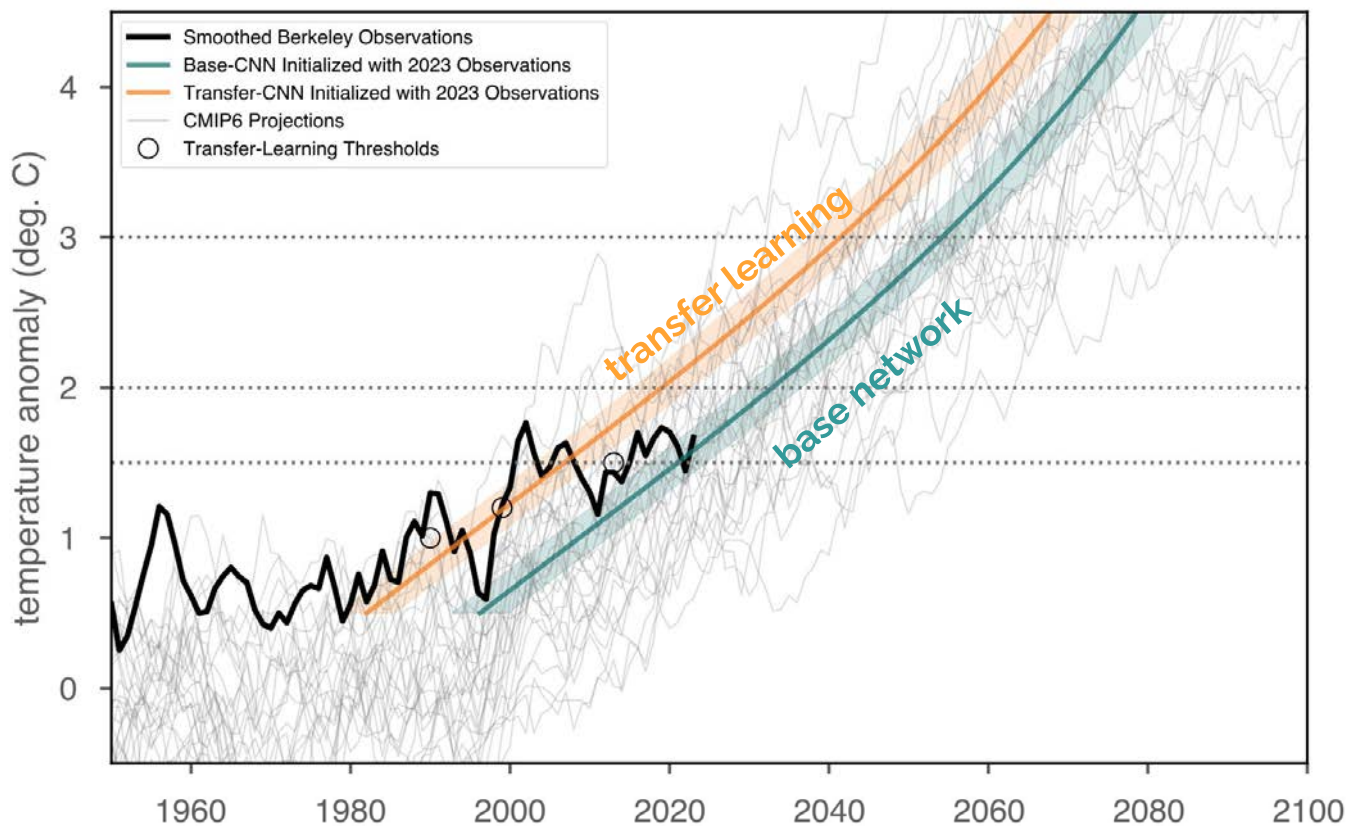
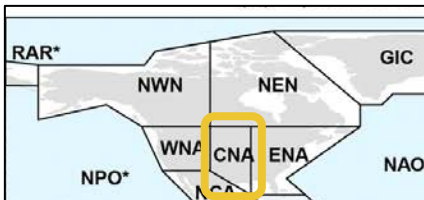
Sec. 8. Protecting Consumers, Patients, Passengers, and Students. (a) Independent regulatory agencies are encouraged, as they deem appropriate, to consider using their full range of authorities to protect American consumers from fraud, discrimination, and threats to privacy and to address other risks 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.**

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

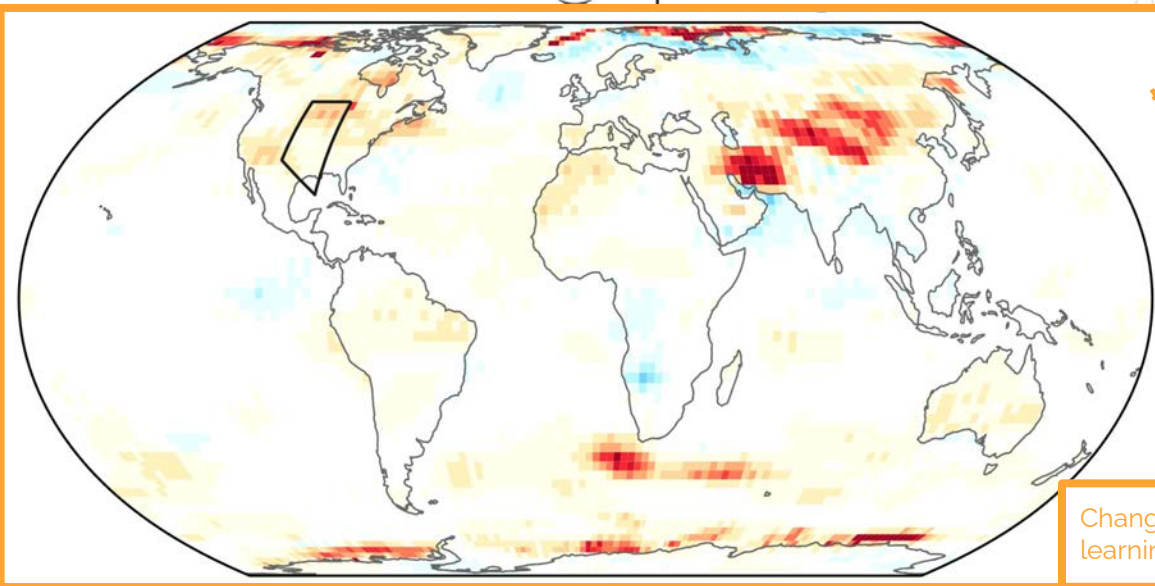
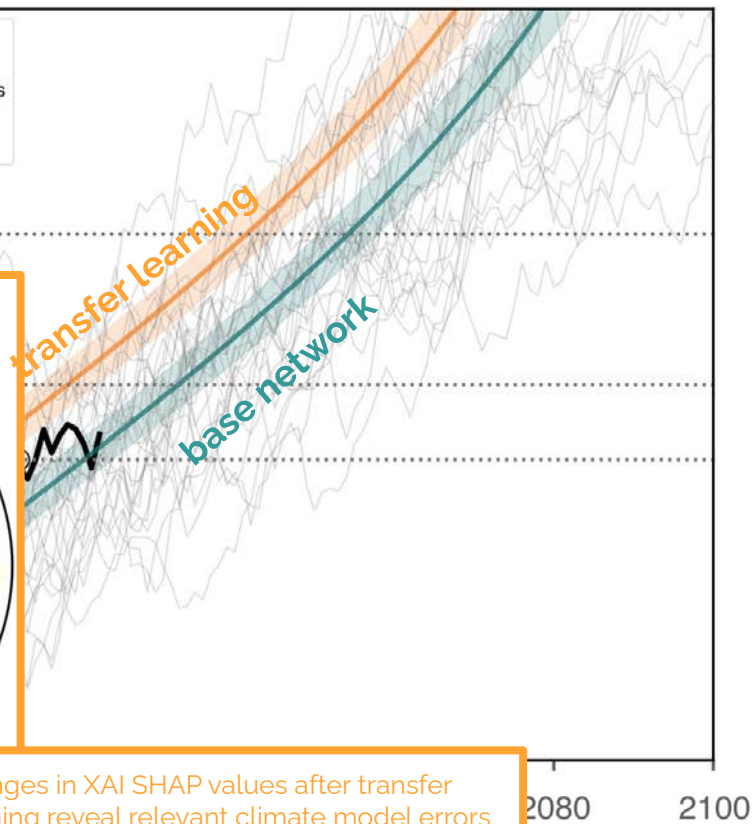
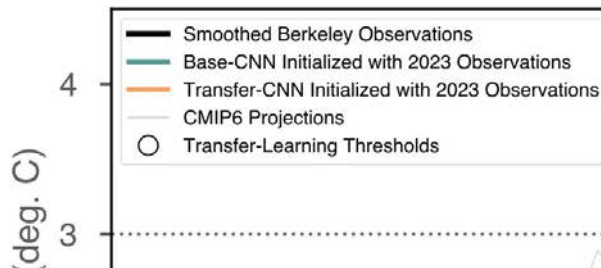
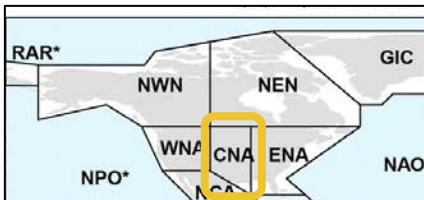


Explainable AI (XAI) will be essential.



Regional transfer learning provides new insights

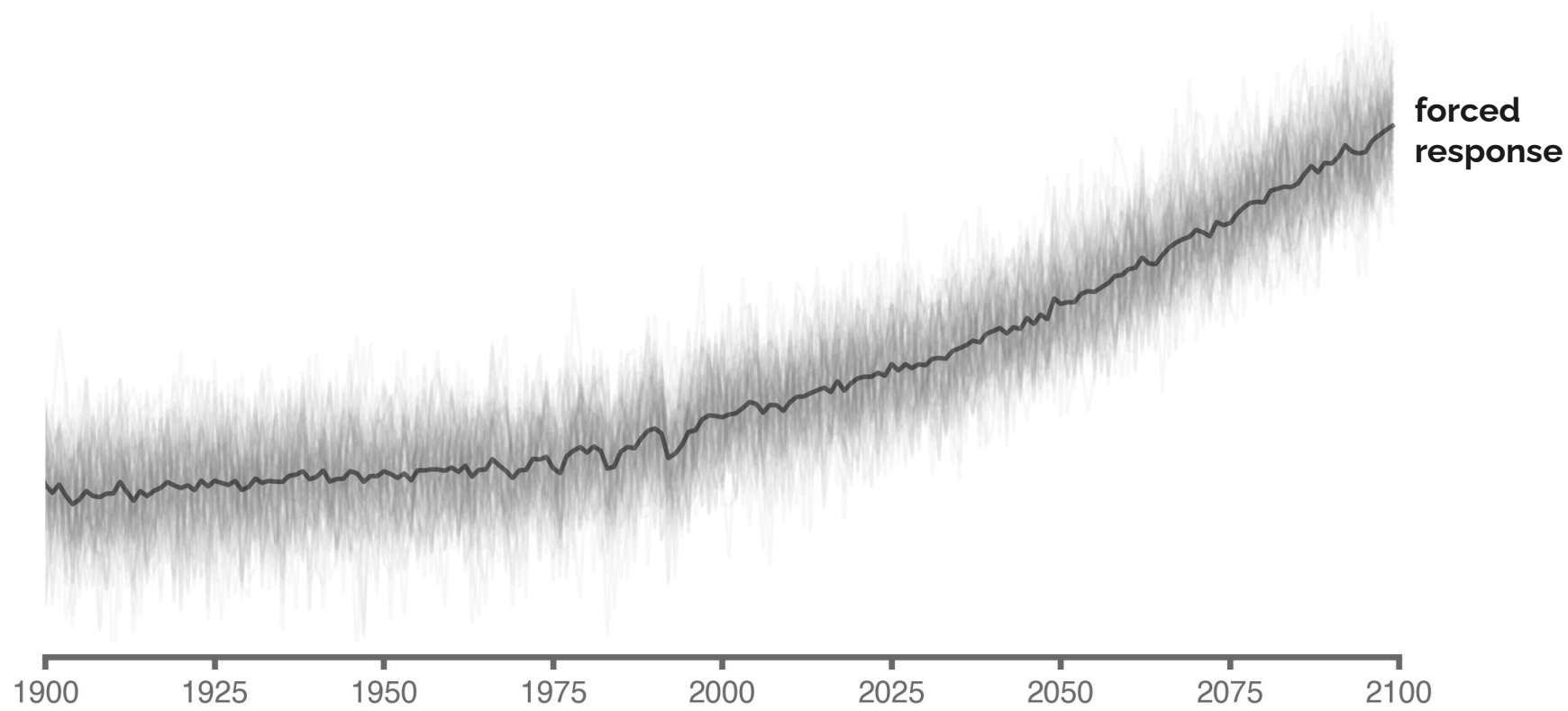




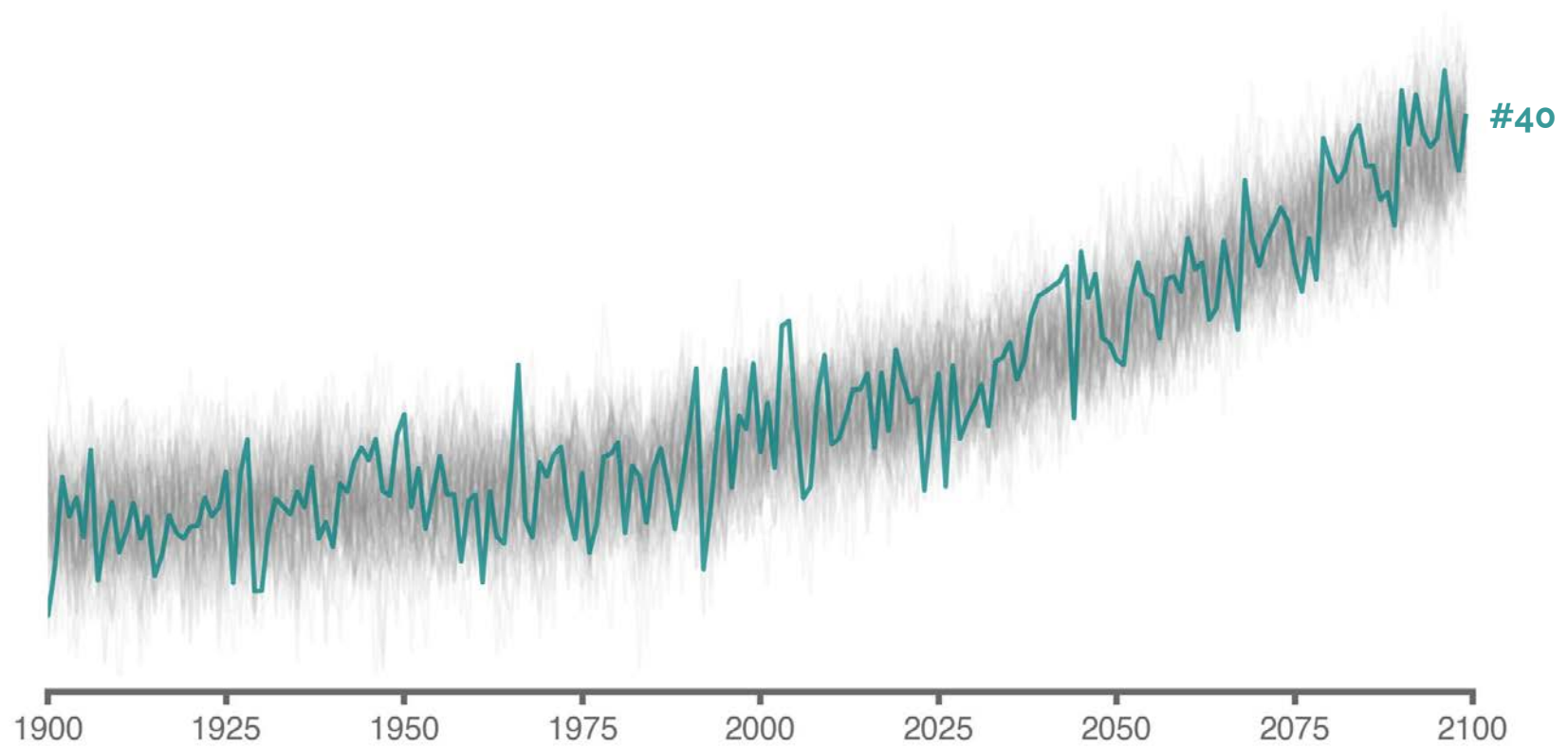
Changes in XAI SHAP values after transfer learning reveal relevant climate model errors

Regional transfer learning provides new insights

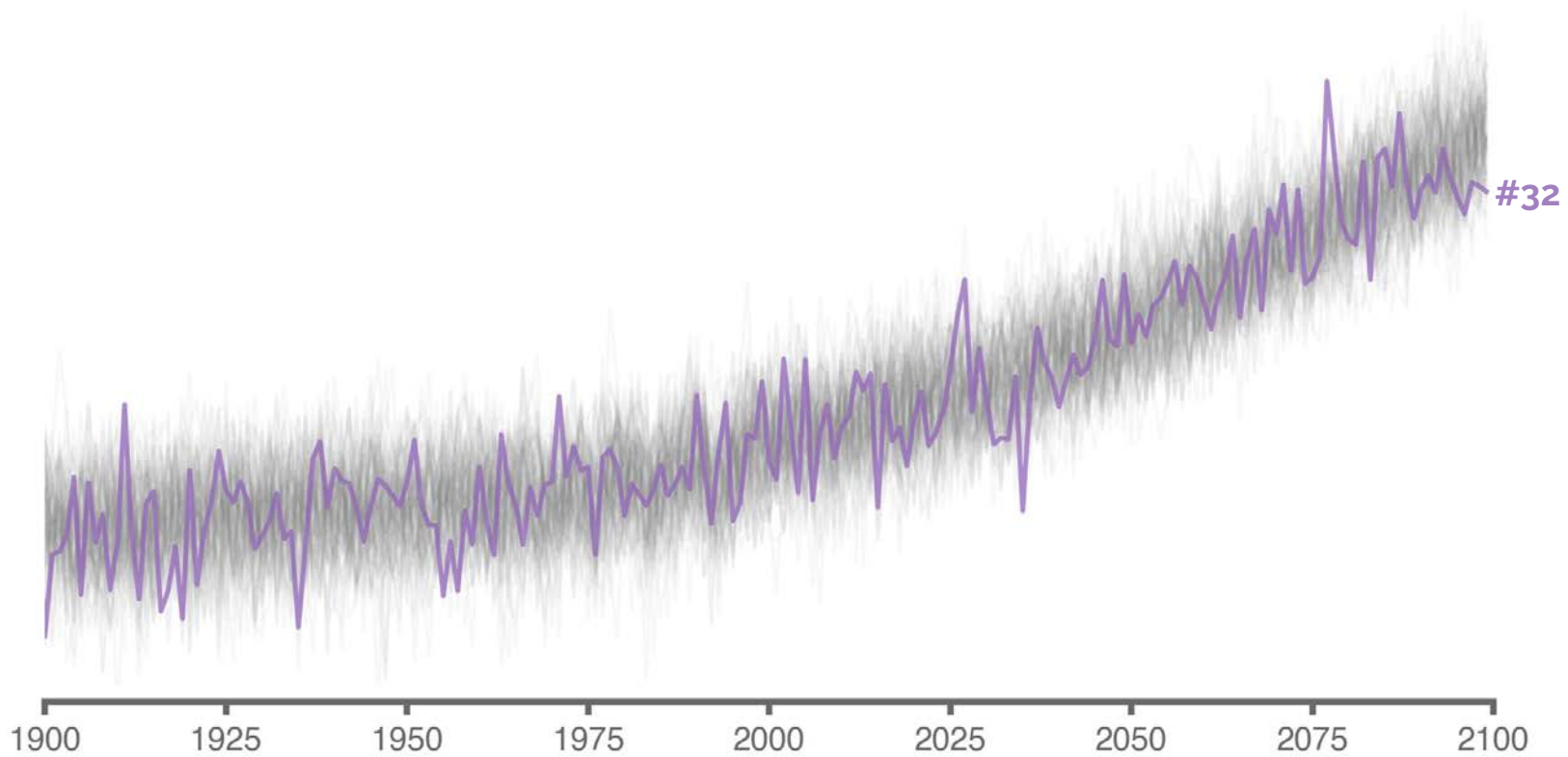




Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5

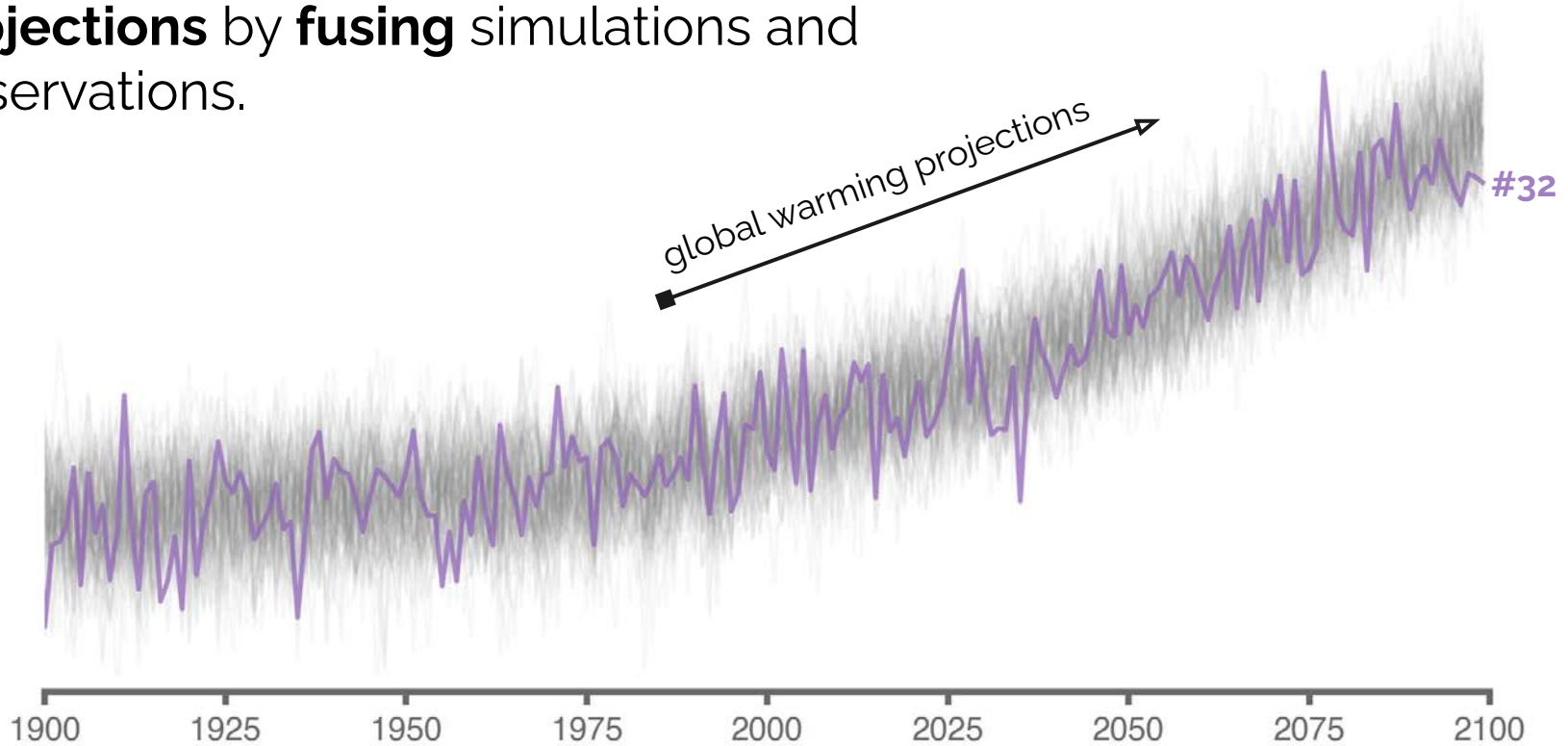


Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5



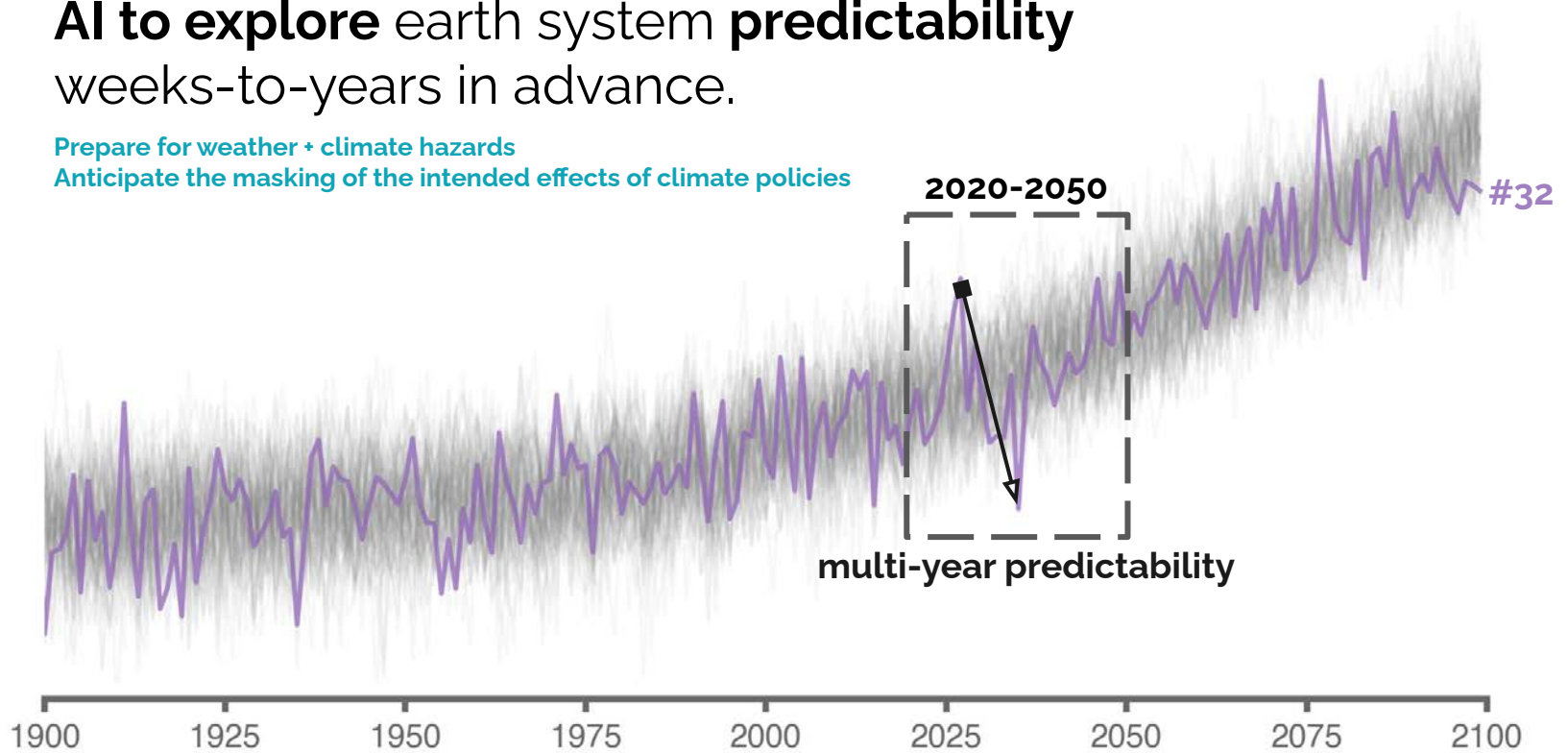
Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5

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

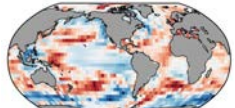


AI to explore earth system **predictability** weeks-to-years in advance.

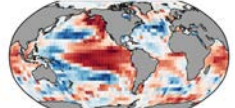
Prepare for weather + climate hazards
Anticipate the masking of the intended effects of climate policies



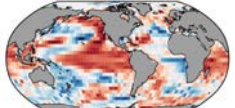
past sea-surface temperatures



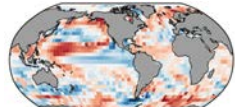
3-8 years before



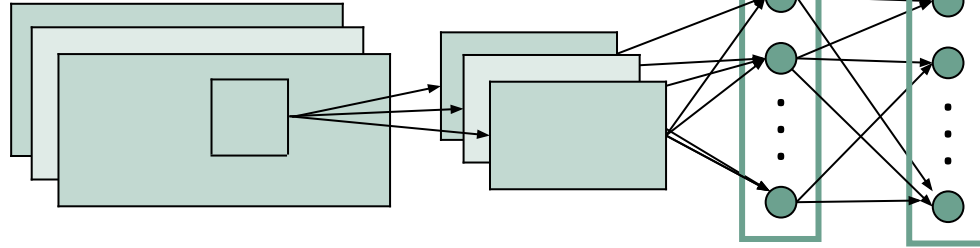
2-3 years before



1-2 years before



0-1 years before



future sea surface temperatures*
for one grid point
[0-5 years]

warm

neutral

cool

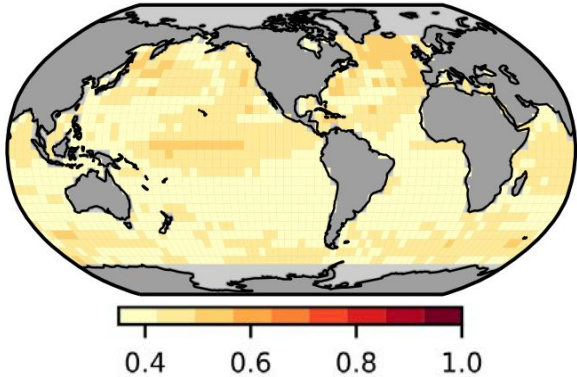
*can predict a range of variables

Predict ocean temperatures 5 years later



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](#)

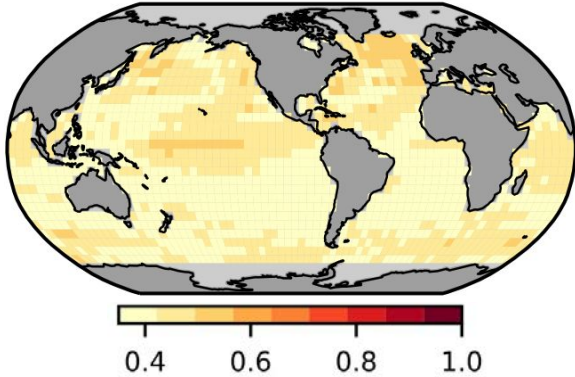
Focusing on when the AI is most confident leads to skillful predictions



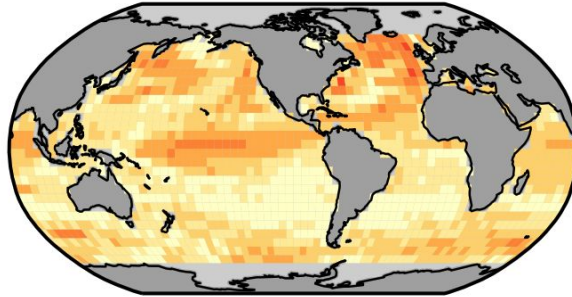
Davenport, Barnes & Gordon (2024)

CLIMATE MODEL DATA

Overall Accuracy



Accuracy for 40% most confident predictions



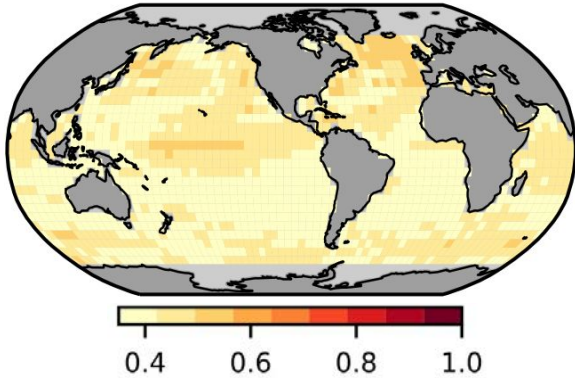
Trained on climate model [MPI-ESM-1-2-LR](#) [3,630 years of data]
Evaluated on climate model [MPI-ESM-1-2-LR](#)

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

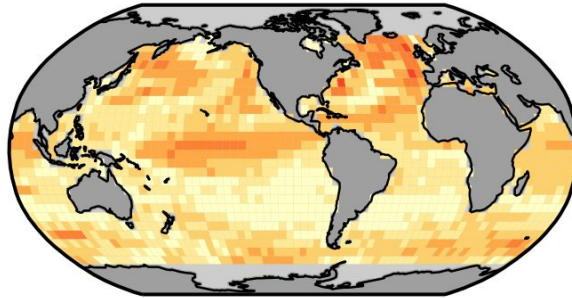


CLIMATE MODEL DATA

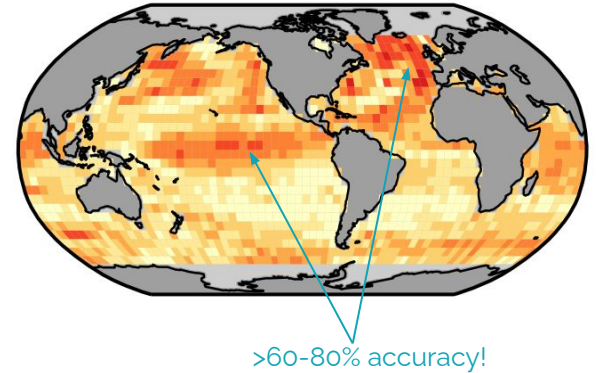
Overall Accuracy



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 climate model [MPI-ESM-1-2-LR](#)

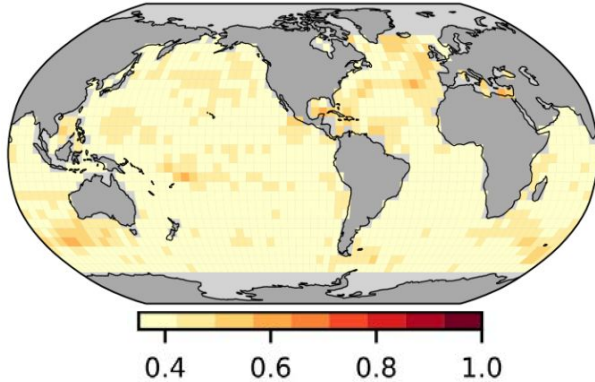
Focusing on when the AI is most confident leads to skillful predictions



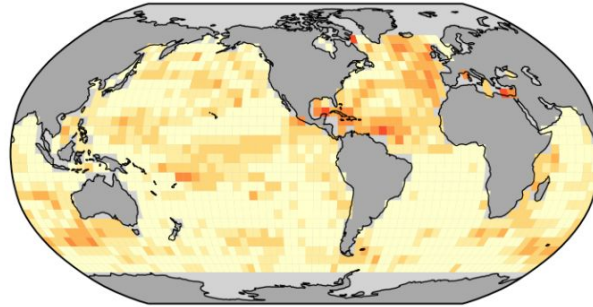
Davenport, Barnes & Gordon (2024)

OBSERVATIONS

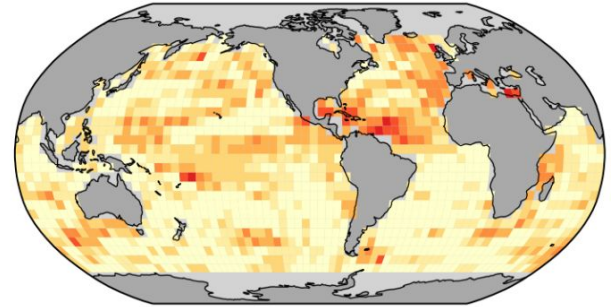
Overall Accuracy



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]

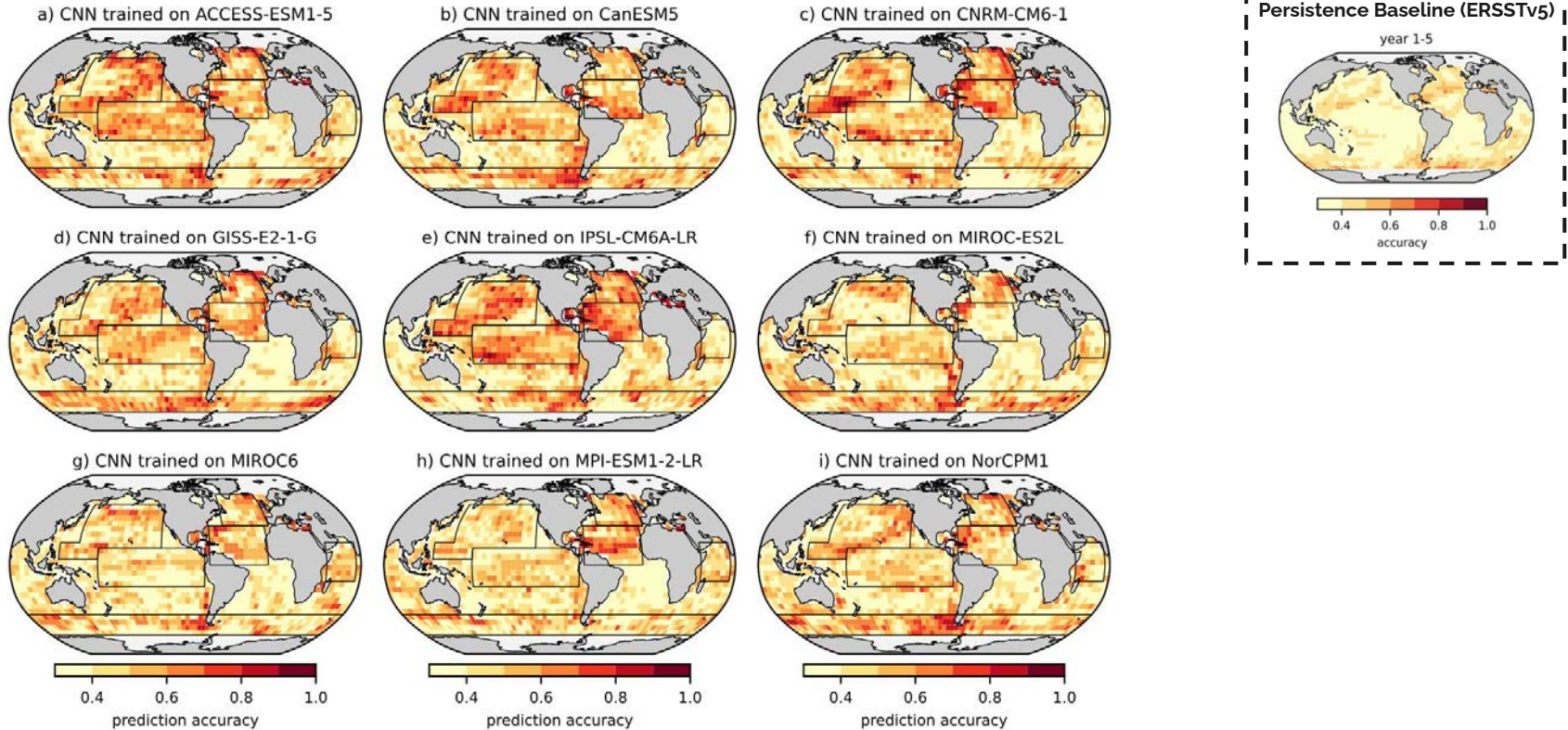
Leveraging climate model data provides skillful predictions of the real world



Davenport, Barnes & Gordon (2024)

Windows of Opportunity tested on ERSSTv5 observations

Accuracy of 20% most confident predictions of **year 1-5** sea surface temperature anomaly

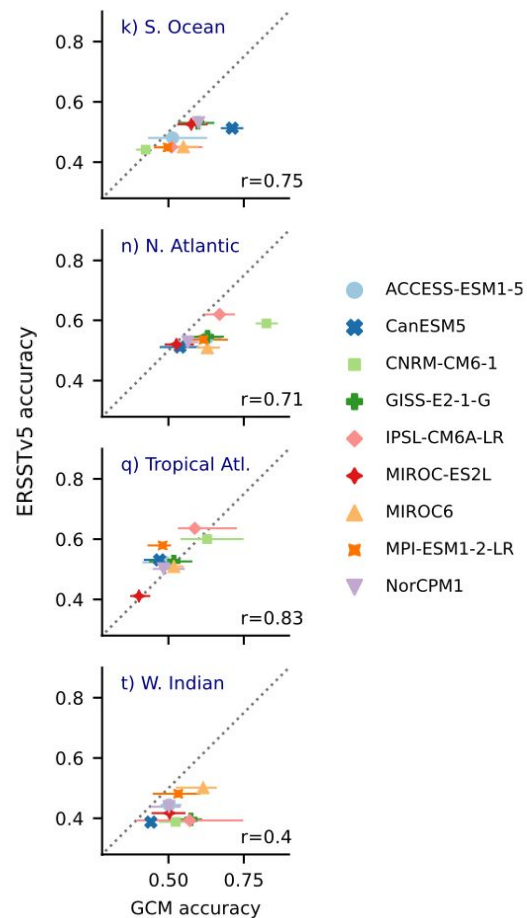
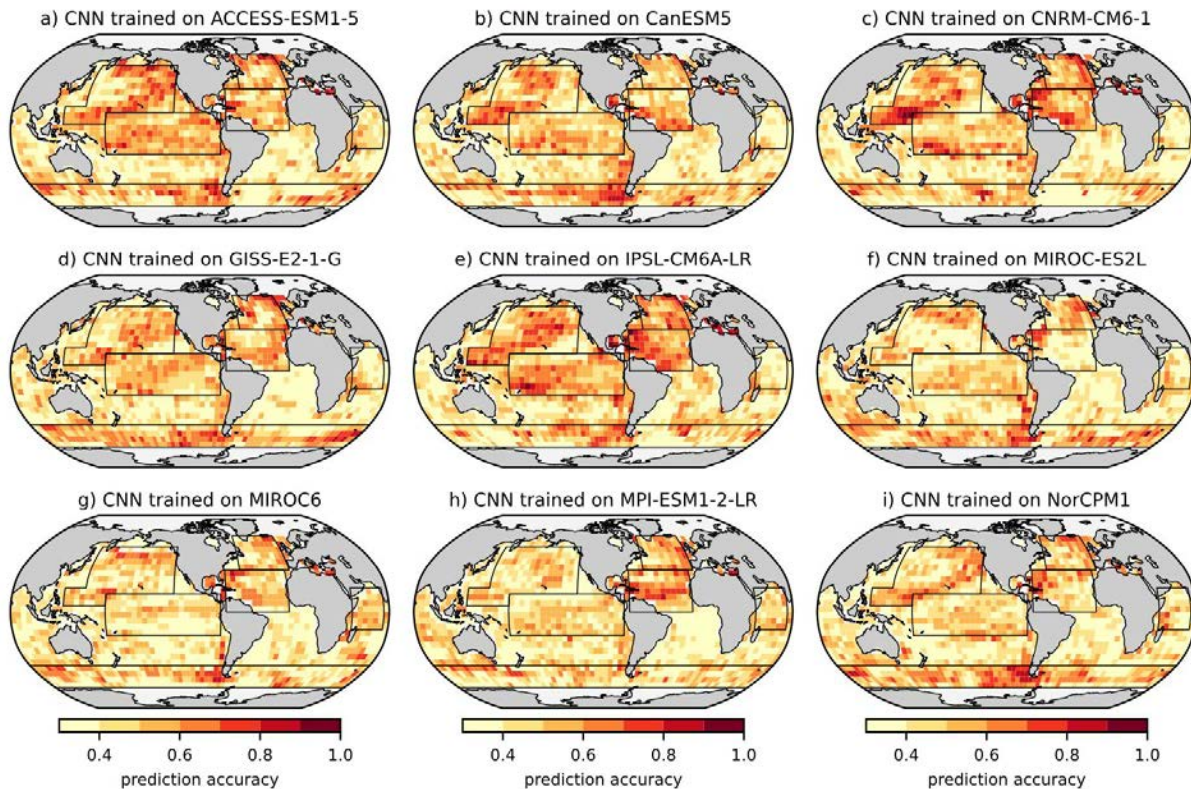


Compare climate model-based network skill



Windows of Opportunity tested on ERSSTv5 observations

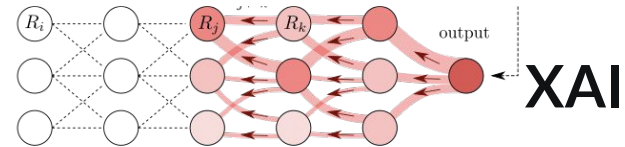
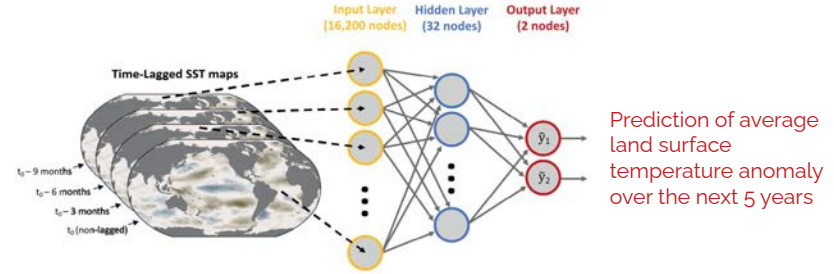
Accuracy of 20% most confident predictions of **year 1-5** sea surface temperature anomaly



Compare climate model-based network skill



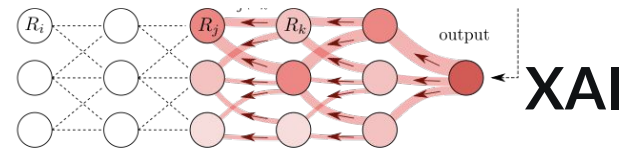
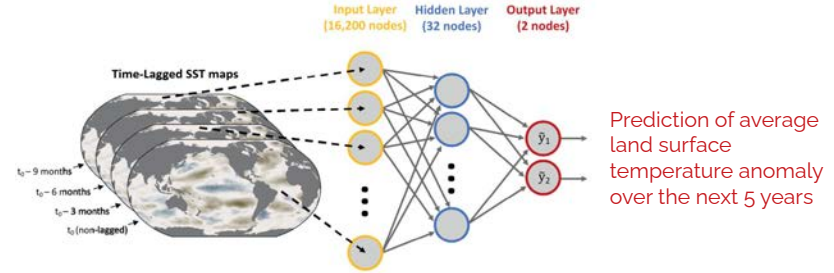
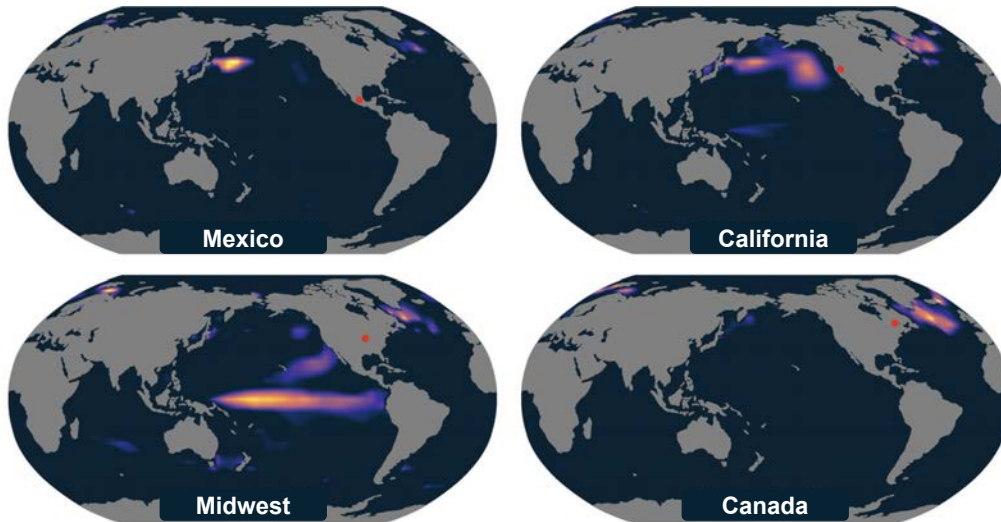
Predicting 5-year average surface temperature at each grid point
Applied to 1200 years of CESM2 control simulation



XAI reveals sources of predictability that vary in time and space



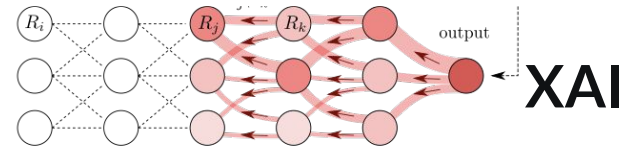
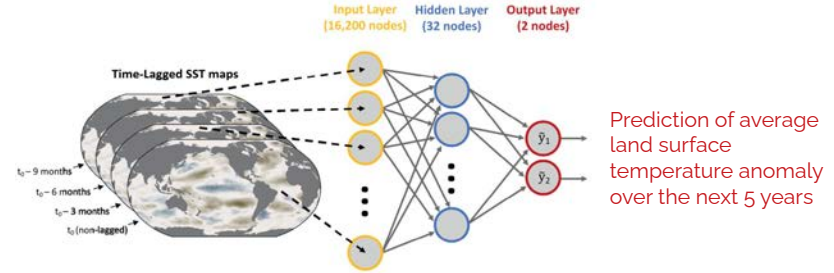
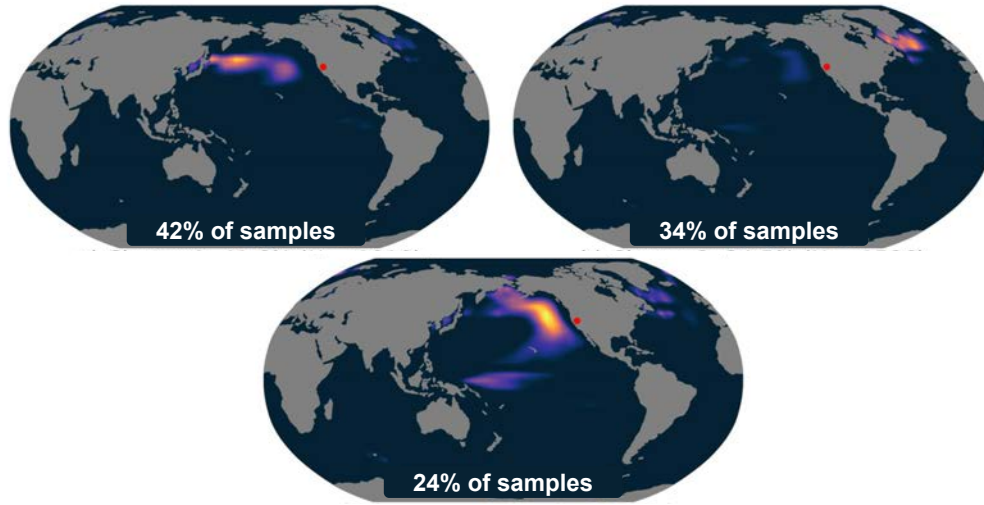
Predicting 5-year average surface temperature at each grid point
Applied to 1200 years of CESM2 control simulation



XAI reveals sources of predictability that vary in time and space



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



XAI reveals sources of predictability that vary in time and space



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

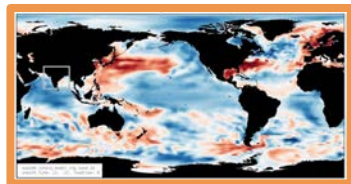


Barnes et al. (2022)

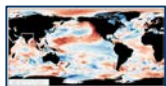
Rader and Barnes (2024)

Gordillo and Barnes (under review)

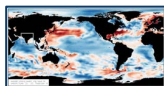
Fernandez and Barnes (in prep)



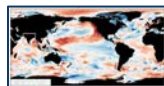
our state of interest
(a map of the current climate)



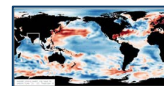
potential analog



potential analog



potential analog

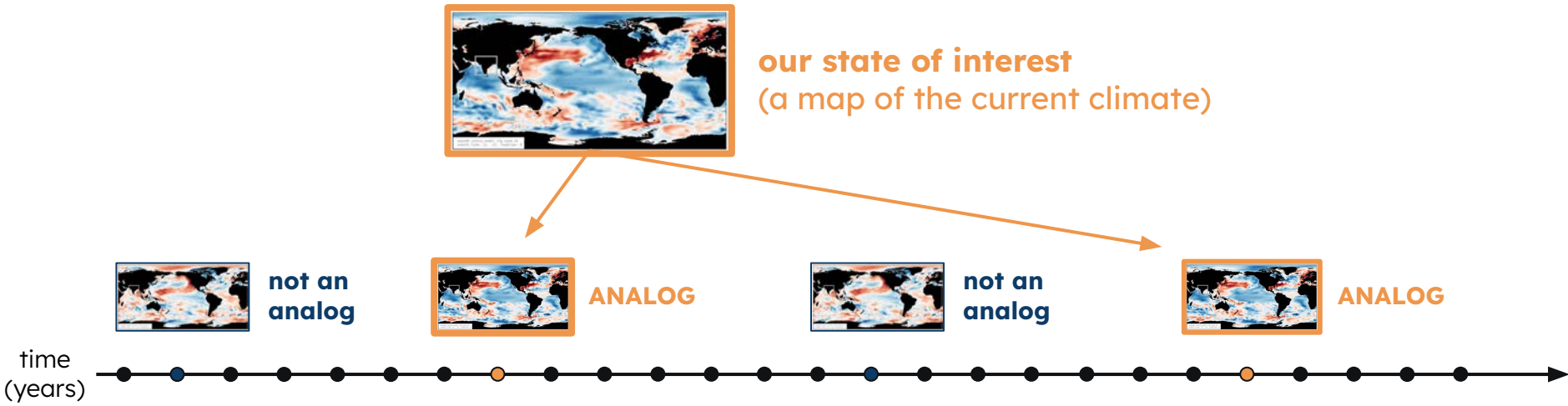


potential analog

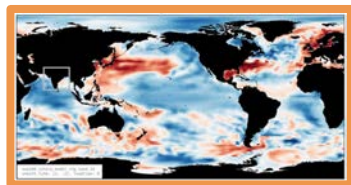
time
(years)



Analog for seasonal-to-decadal prediction



Analogs for seasonal-to-decadal prediction



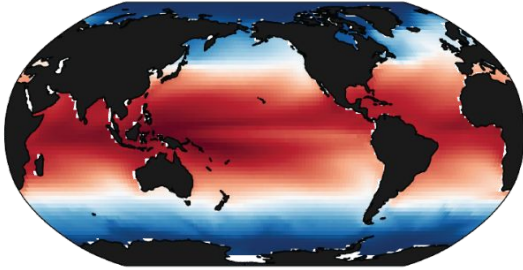
our state of interest
(a map of the current climate)



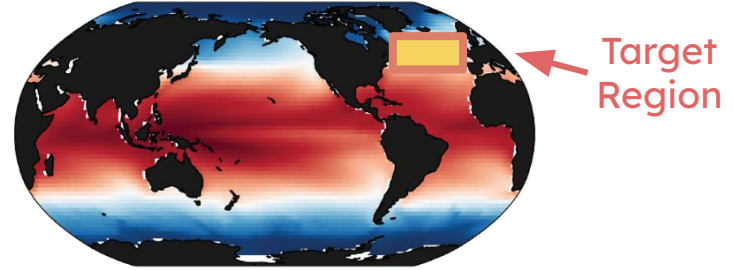
use **how the climate evolved** after the analog events to make an **ensemble forecast** for our state of interest

Analogues for seasonal-to-decadal prediction

Use:
SST in Year 0-4



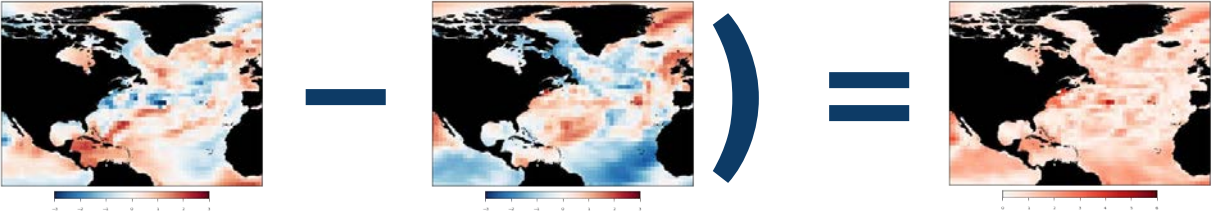
To Predict:
North Atlantic SST in Year 5-9



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

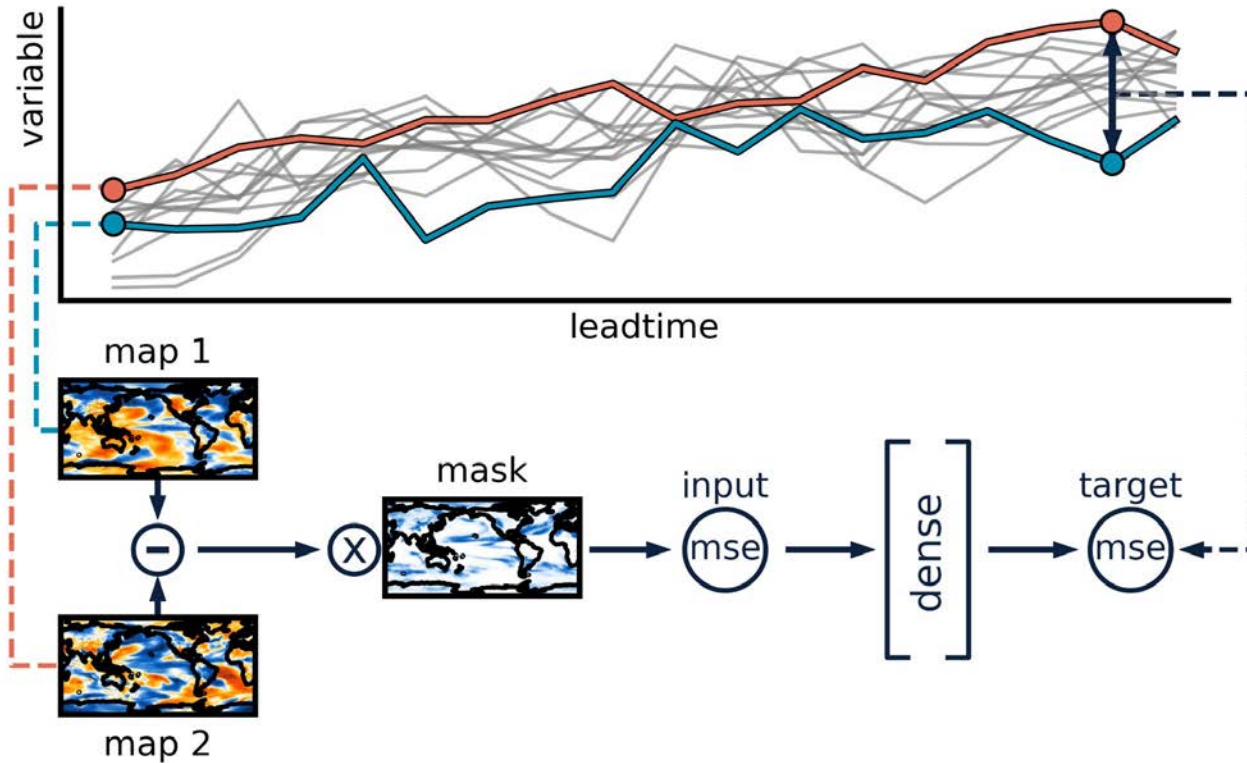
Analog Forecasts for North Atlantic Sea Surface Temperature

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

$$\text{abs}(\text{state of interest} - \text{potential analog}) = \text{difference}$$


The diagram illustrates the calculation of the absolute difference between two climate states. It shows the equation $\text{abs}(\text{state of interest} - \text{potential analog}) = \text{difference}$. Each term is accompanied by a global map with a color scale from blue (negative) to red (positive). The 'state of interest' map shows a mix of blue and red. The 'potential analog' map shows a similar mix. The 'difference' map shows the result of subtracting the two, with most areas in red, indicating positive differences.

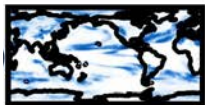
This assumes that every region is equally important for determining how the climate system will evolve. **We can do better.**



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



mask

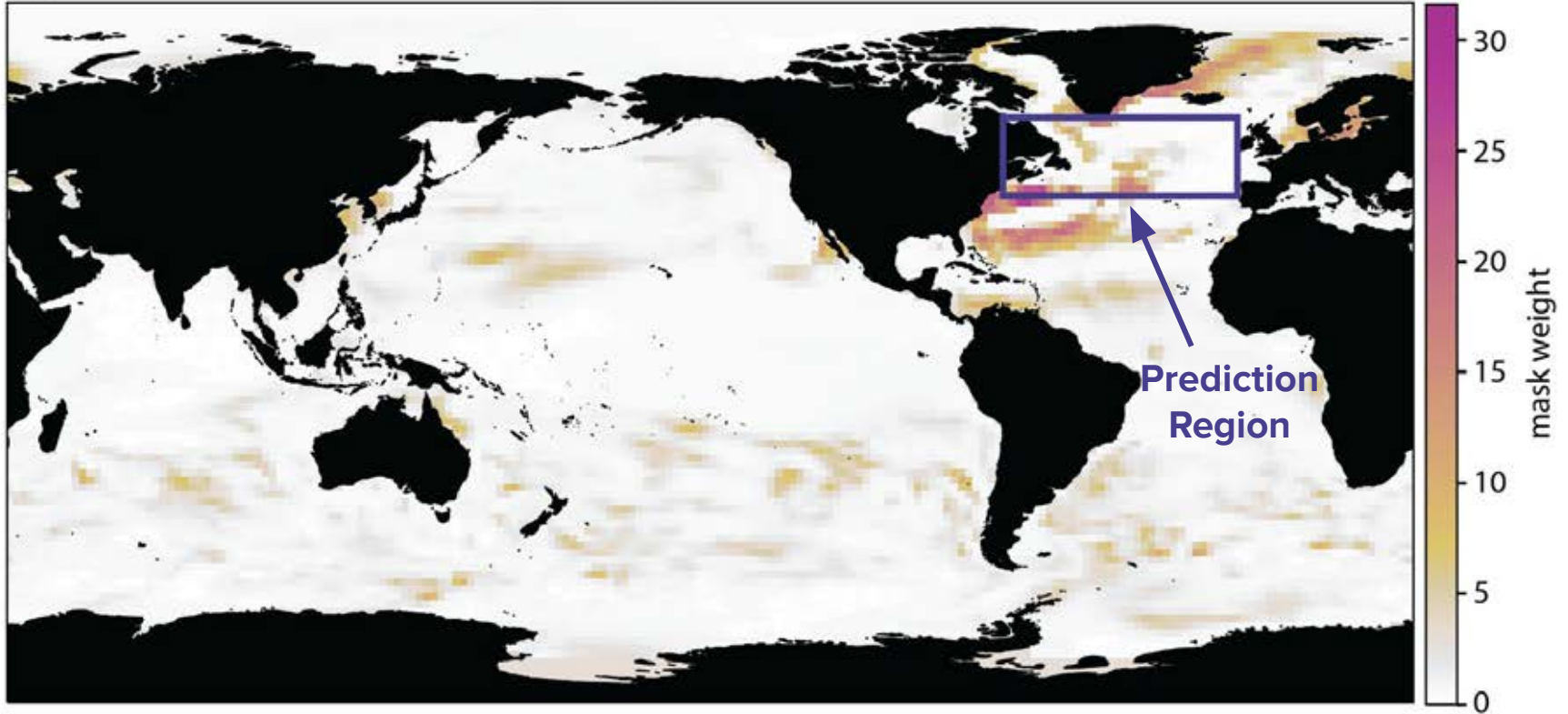


Take this mask and use it to determine the “best analogs” in the standard, non-AI, way.

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



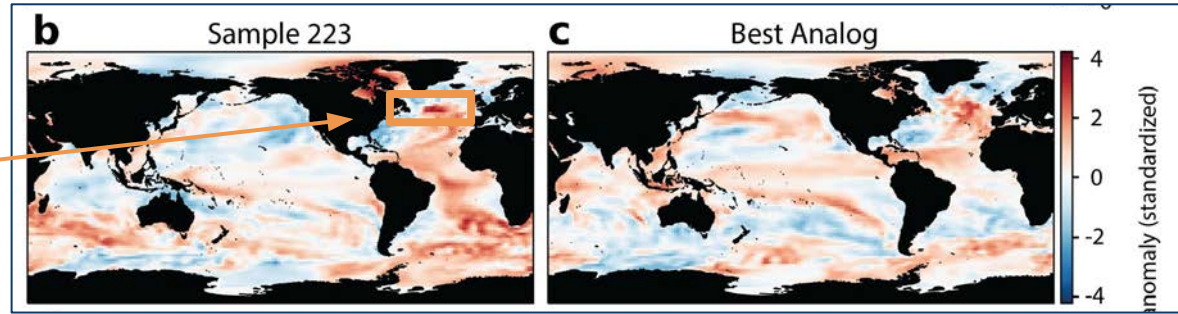
Weighted Mask



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

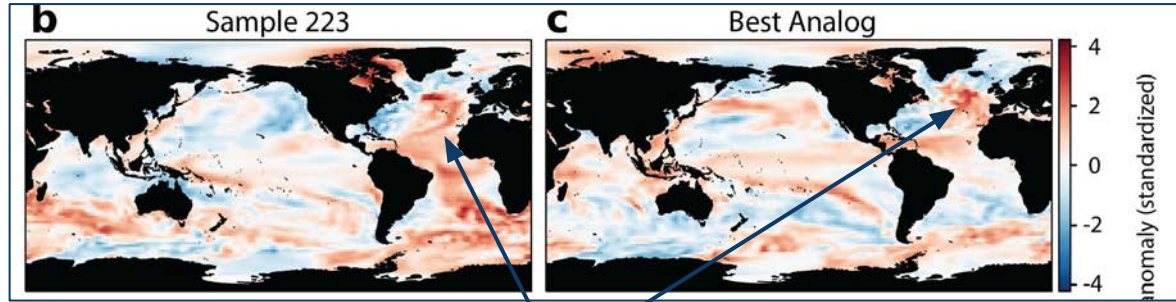


Predicting 5-yr
SST in the North
Atlantic



Identify the analogs with the weighted mask

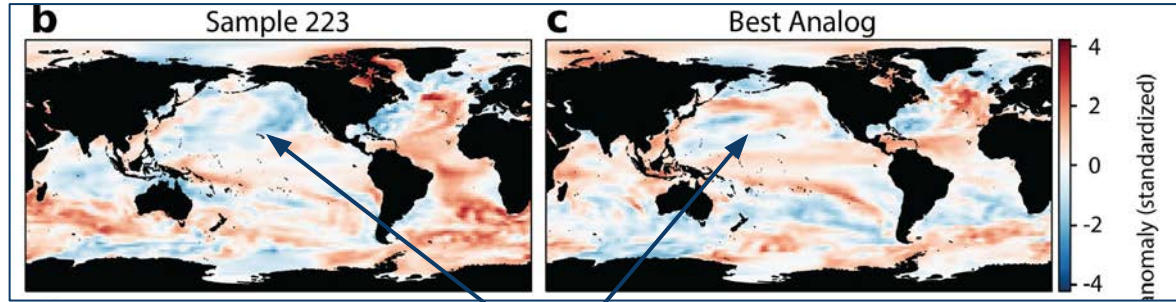




Some regions look similar

Identify the analogs with the weighted mask

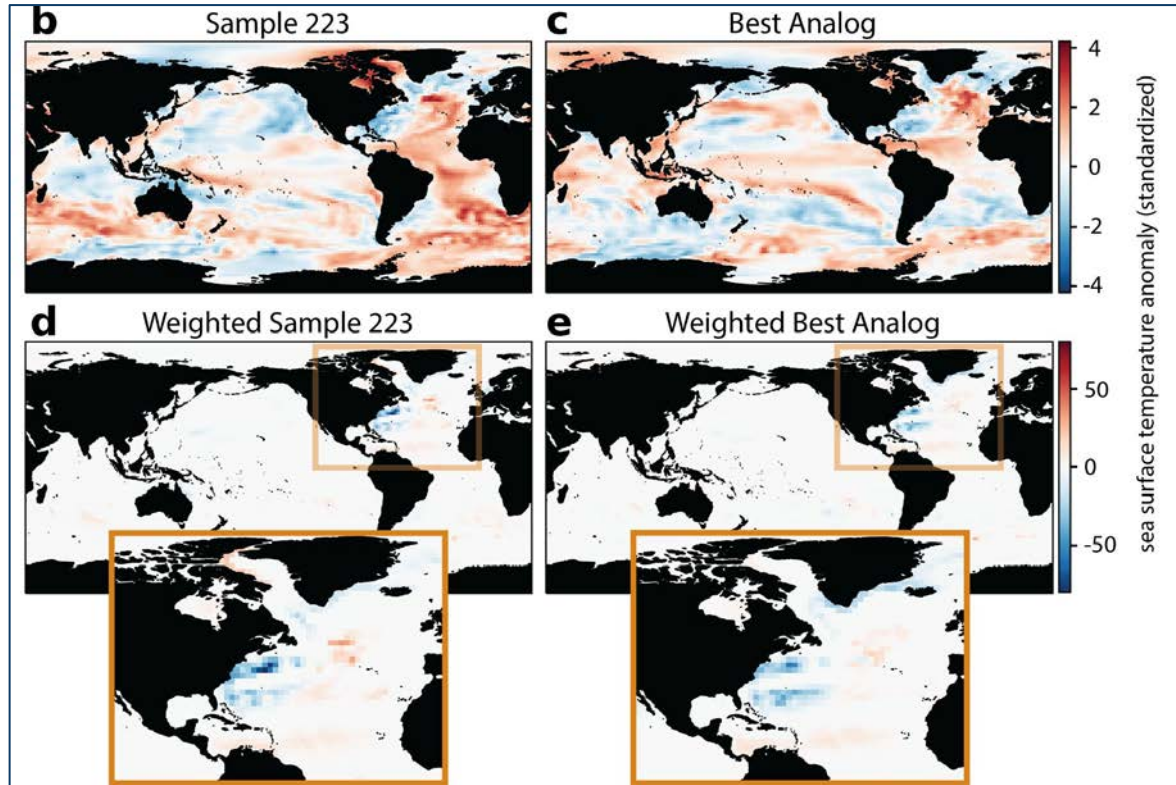




Some regions look different

Identify the analogs with the weighted mask



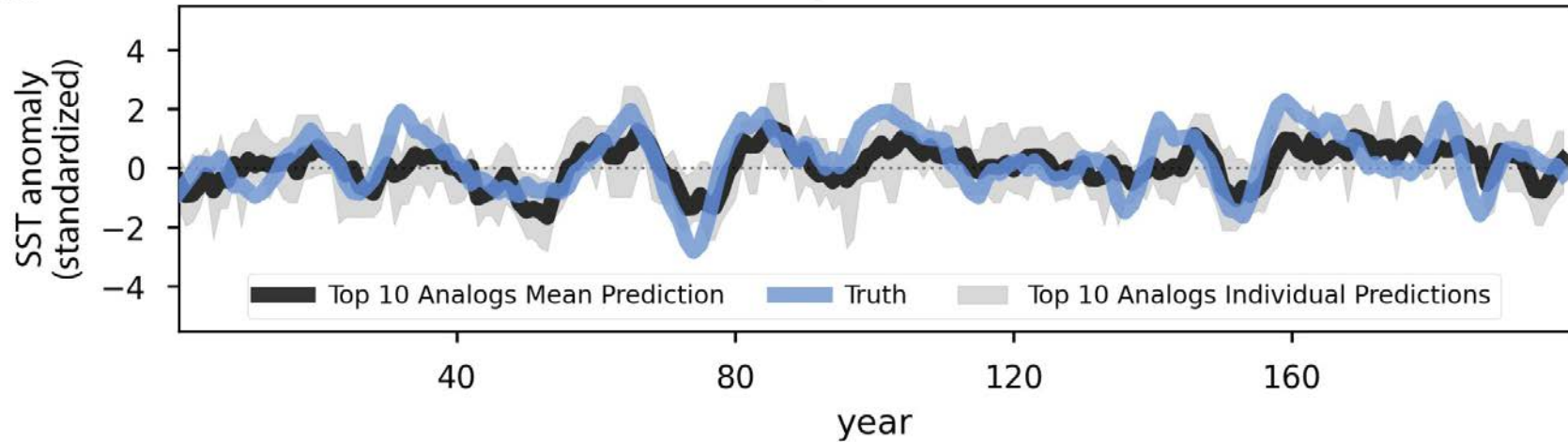


Maps look similar
in the precursor
regions

Identify the analogs



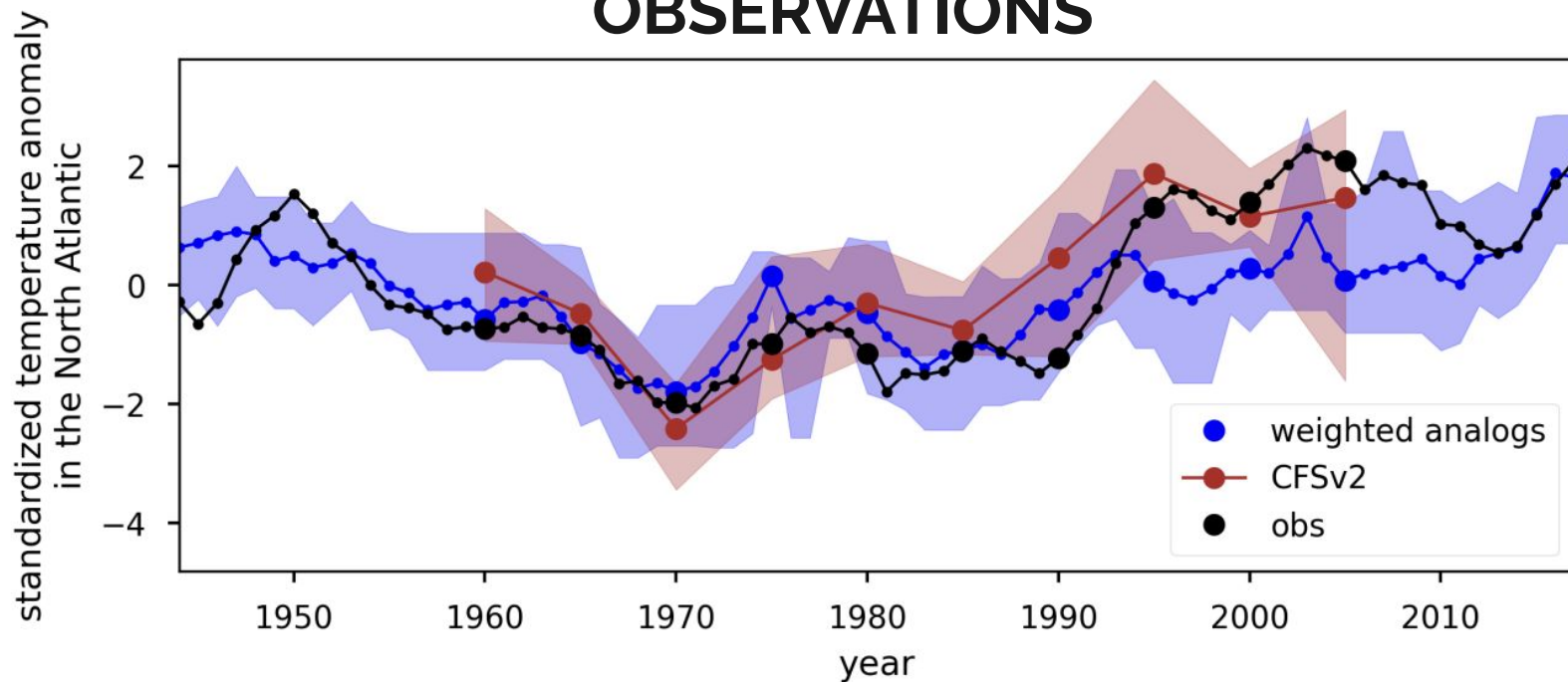
CLIMATE MODEL RESULTS



Forecasts are skillful



OBSERVATIONS



These optimized analogs rival the skill of dynamical models



This approach is interpretable AI!

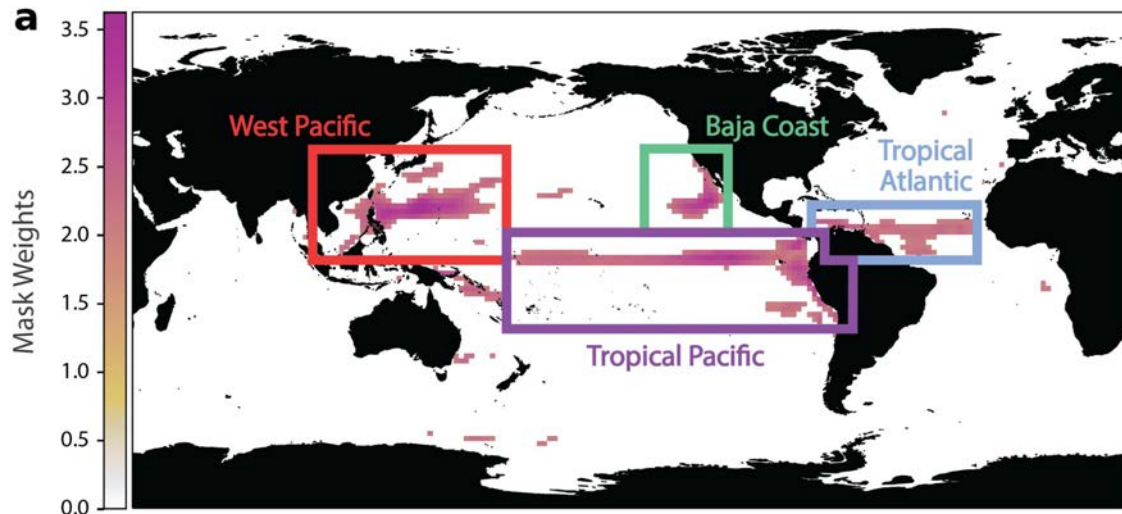
While one could train a massive black-box AI model to make these predictions, the benefit of analogs is that they are easily understood.

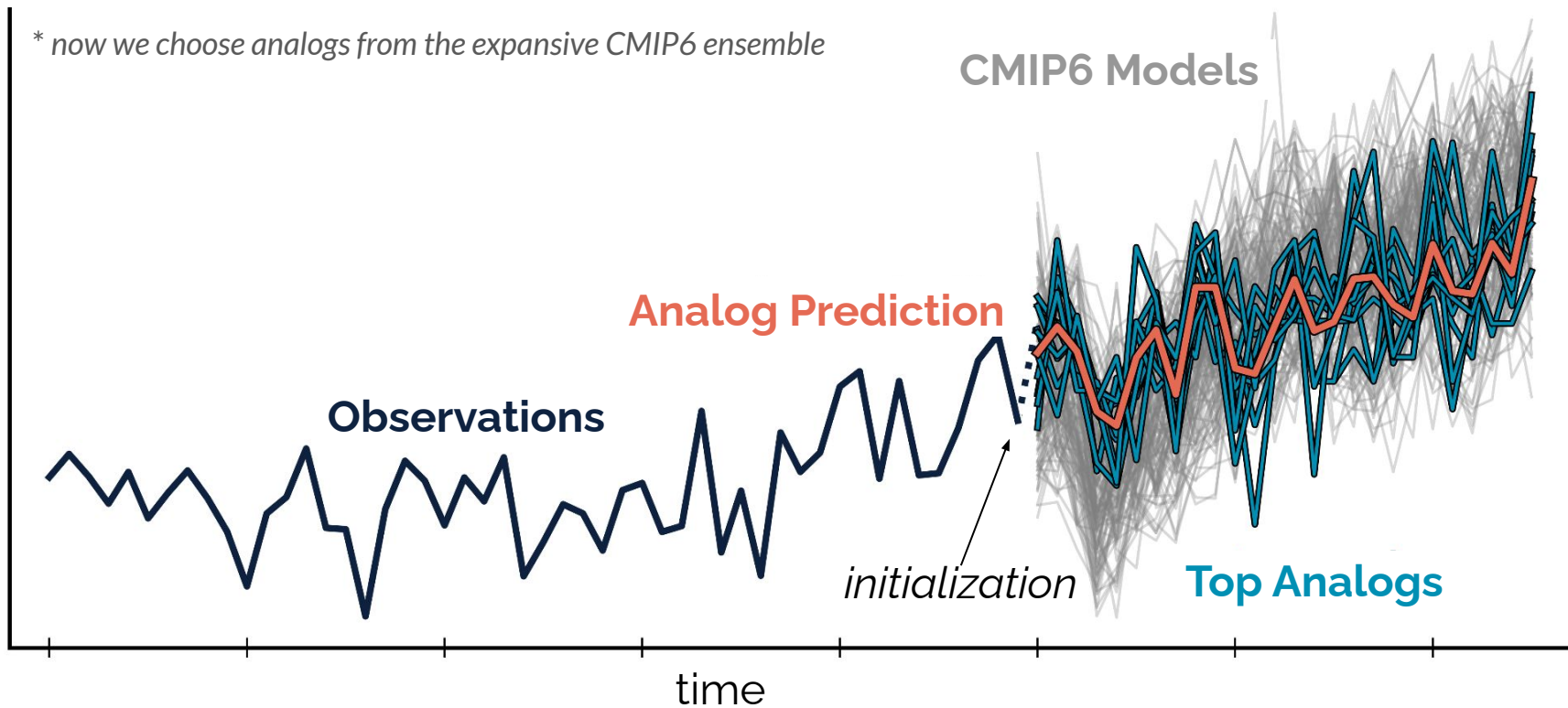
AI is merely used to learn the optimal mask for choosing the analogs. The masks are the “XAI”.

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)

Learned ENSO Prediction Mask

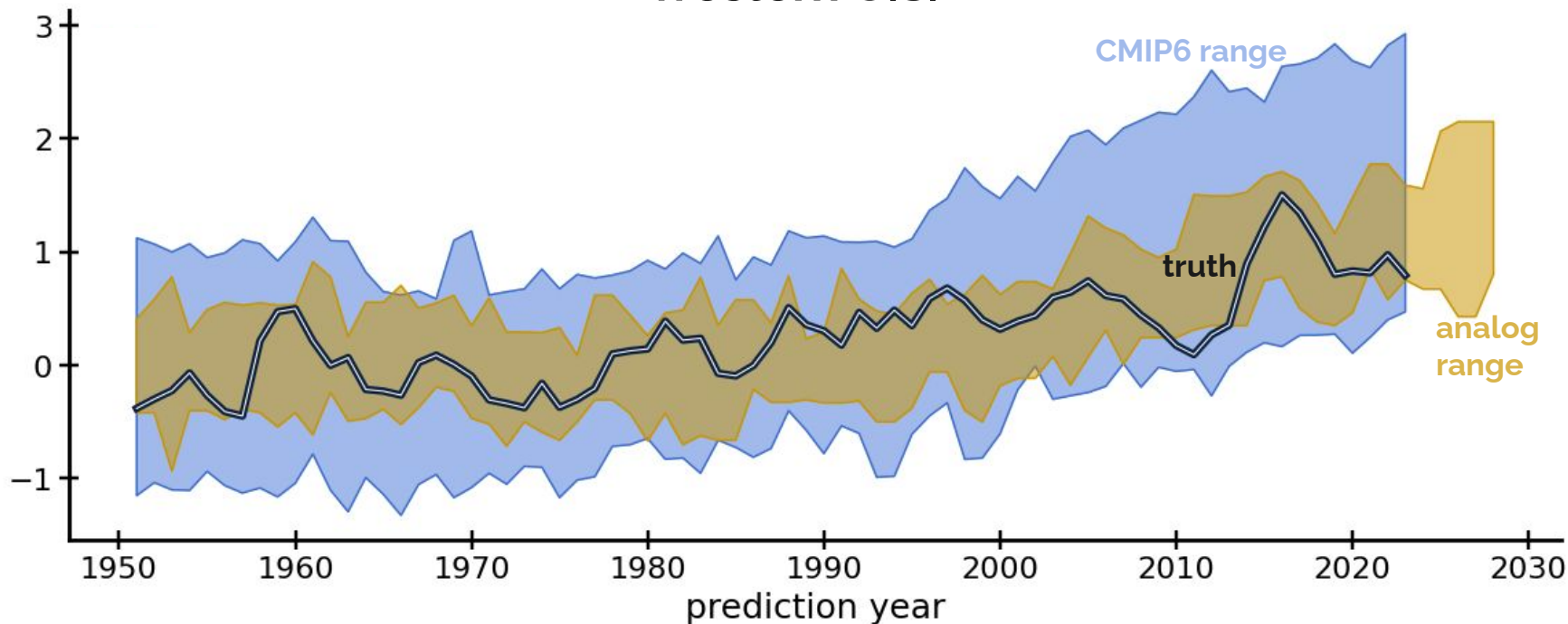




Analogs can help constrain projections → initialized predictions



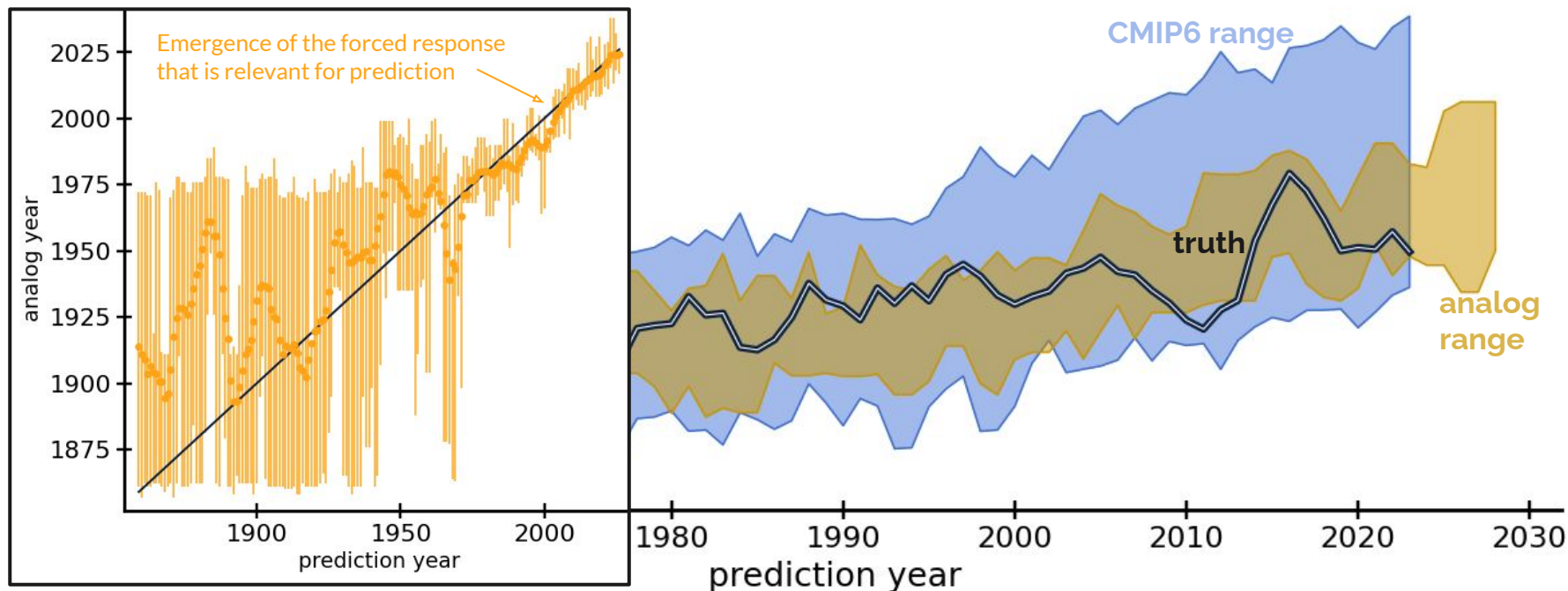
Yrs 5-7 Surface Temperature Anomalies Western U.S.



Analogs can help constrain projections → initialized predictions



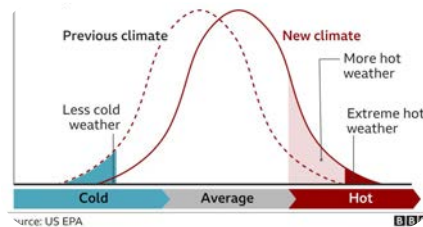
Yrs 5-7 Surface Temperature Anomalies Western U.S.



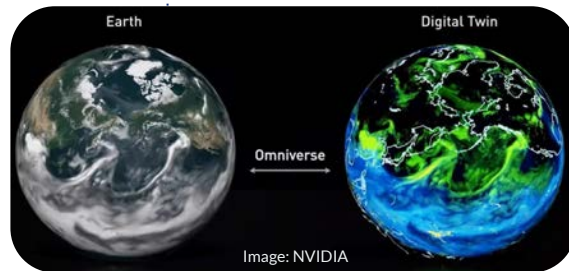
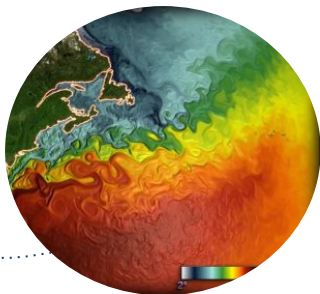
Analogs can help constrain projections → initialized predictions



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