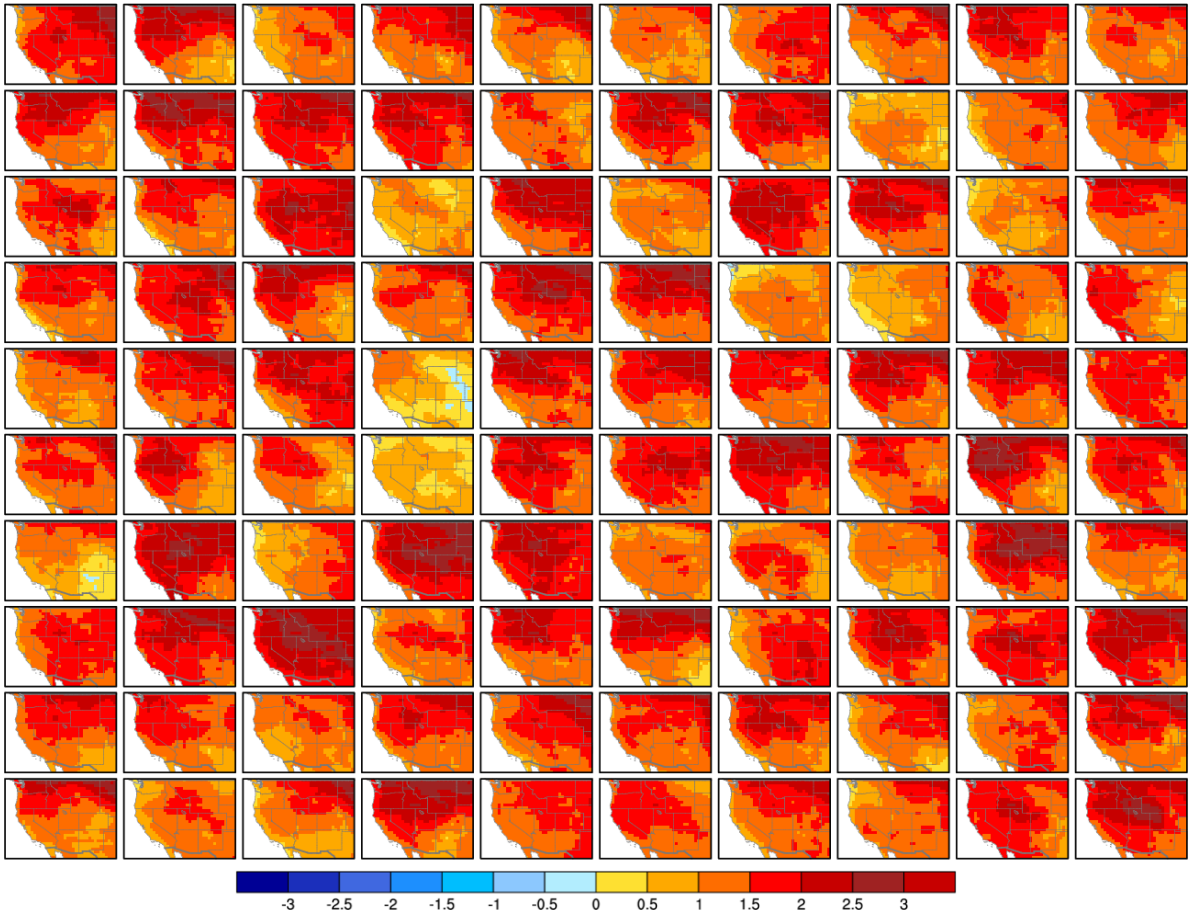
Climate Information in the AI Era

Dr. Elizabeth A. Barnes
Dalton Family Chair in Environmental Data Science & Sustainability
Professor of Computing & Data Sciences
Professor of Earth & Environment
Boston University

EPRI Energy and Climate Research Seminar
May 6, 2026

100 plausible futures for the same emissions scenario

Temperature trends for
2021–2050

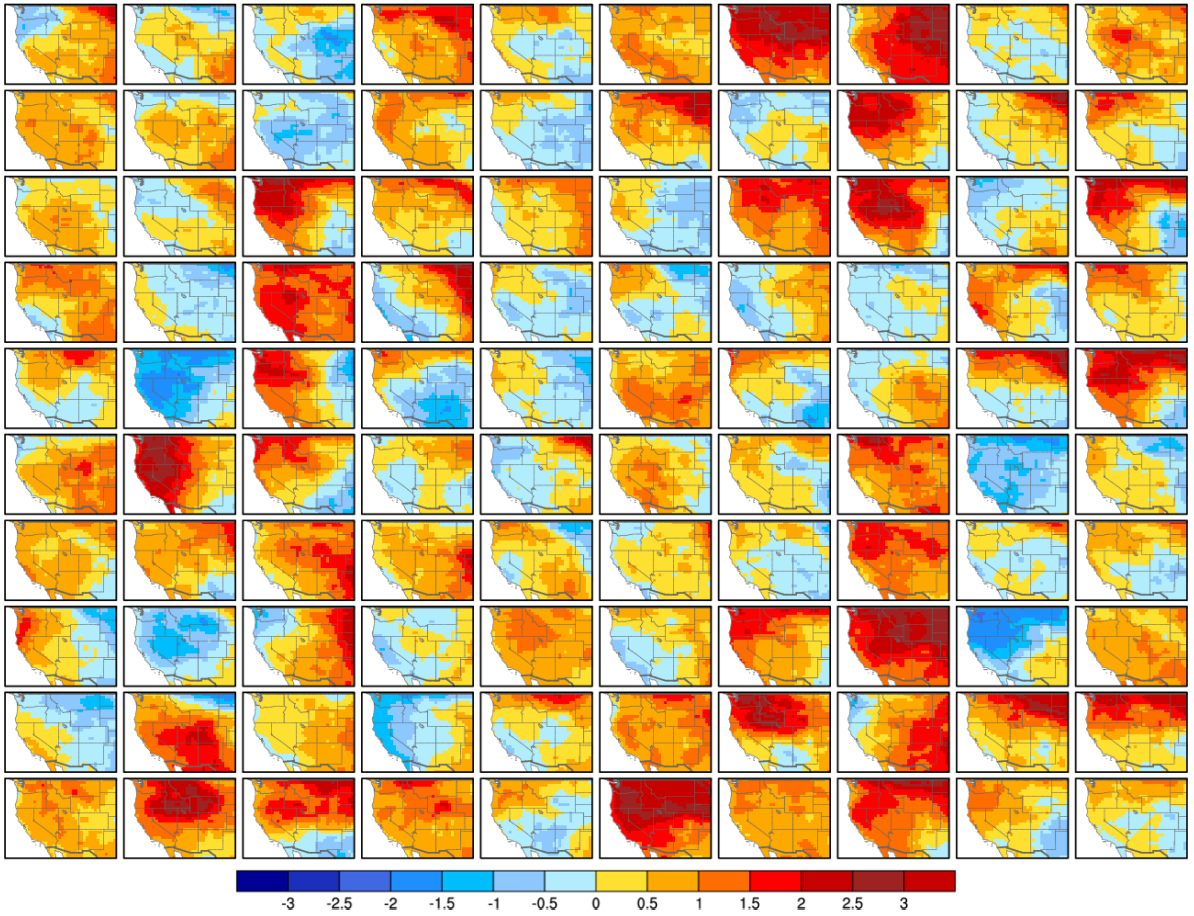


Annual mean surface temperature as simulated by the CESM2 climate model over 100 different realizations; figure by Dr. Lantao Sun

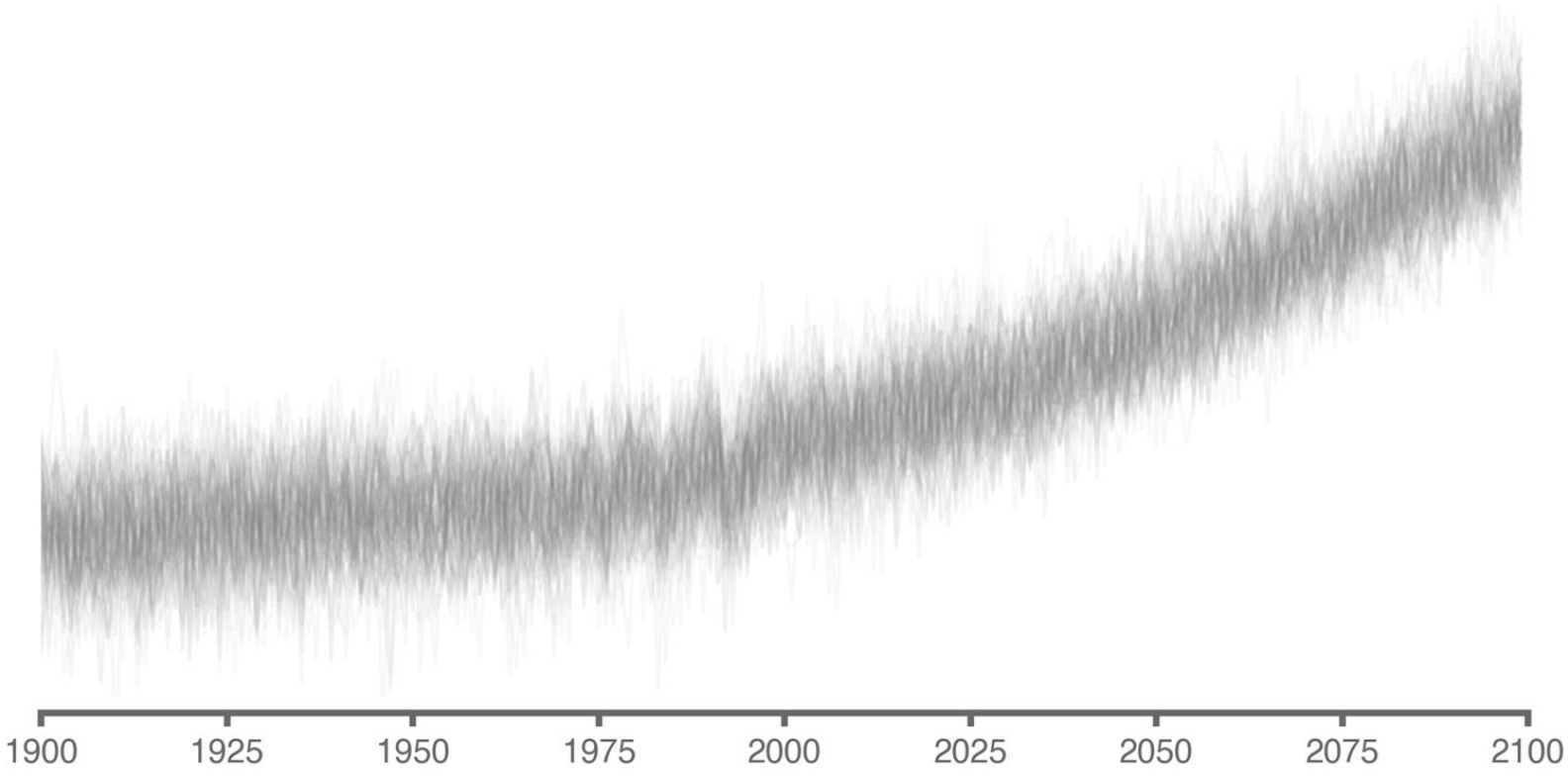
100 plausible futures for the same emissions scenario

Temperature trends for
2021-2030

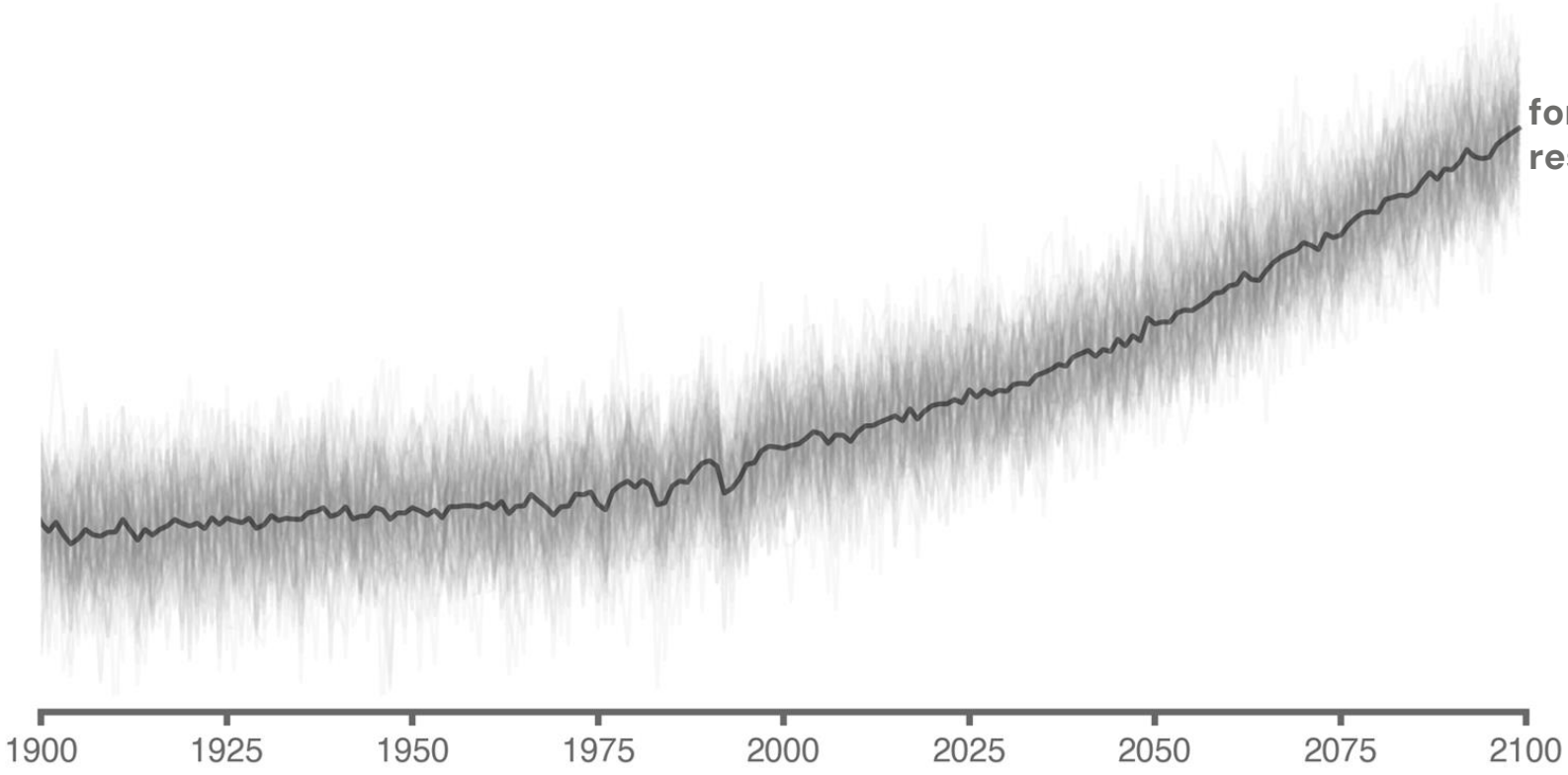
On regional scales, natural variability dominates the signal you're planning for.



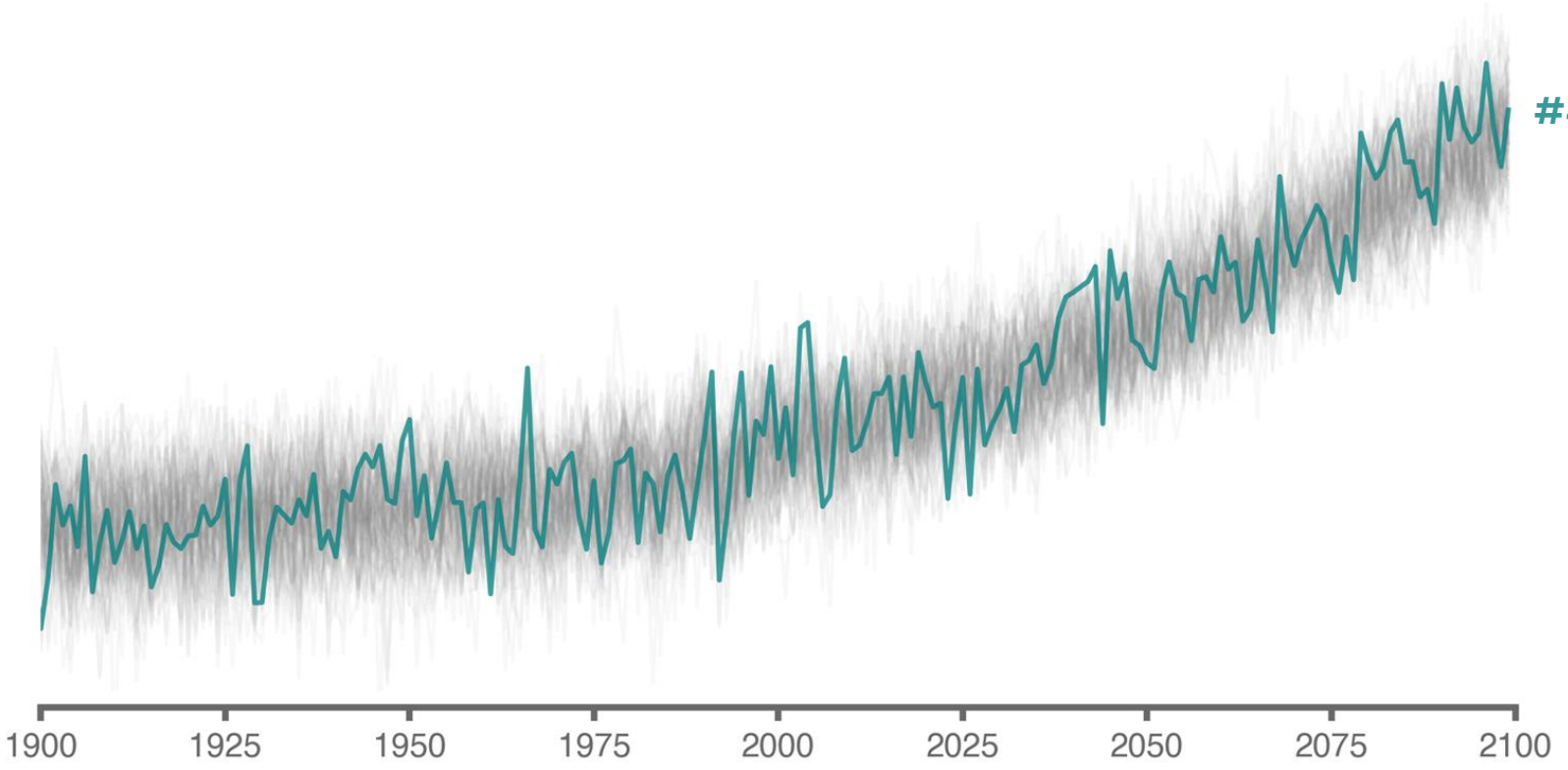
Annual mean surface temperature as simulated by the CESM2 climate model over 100 different realizations; figure by Dr. Lantao Sun



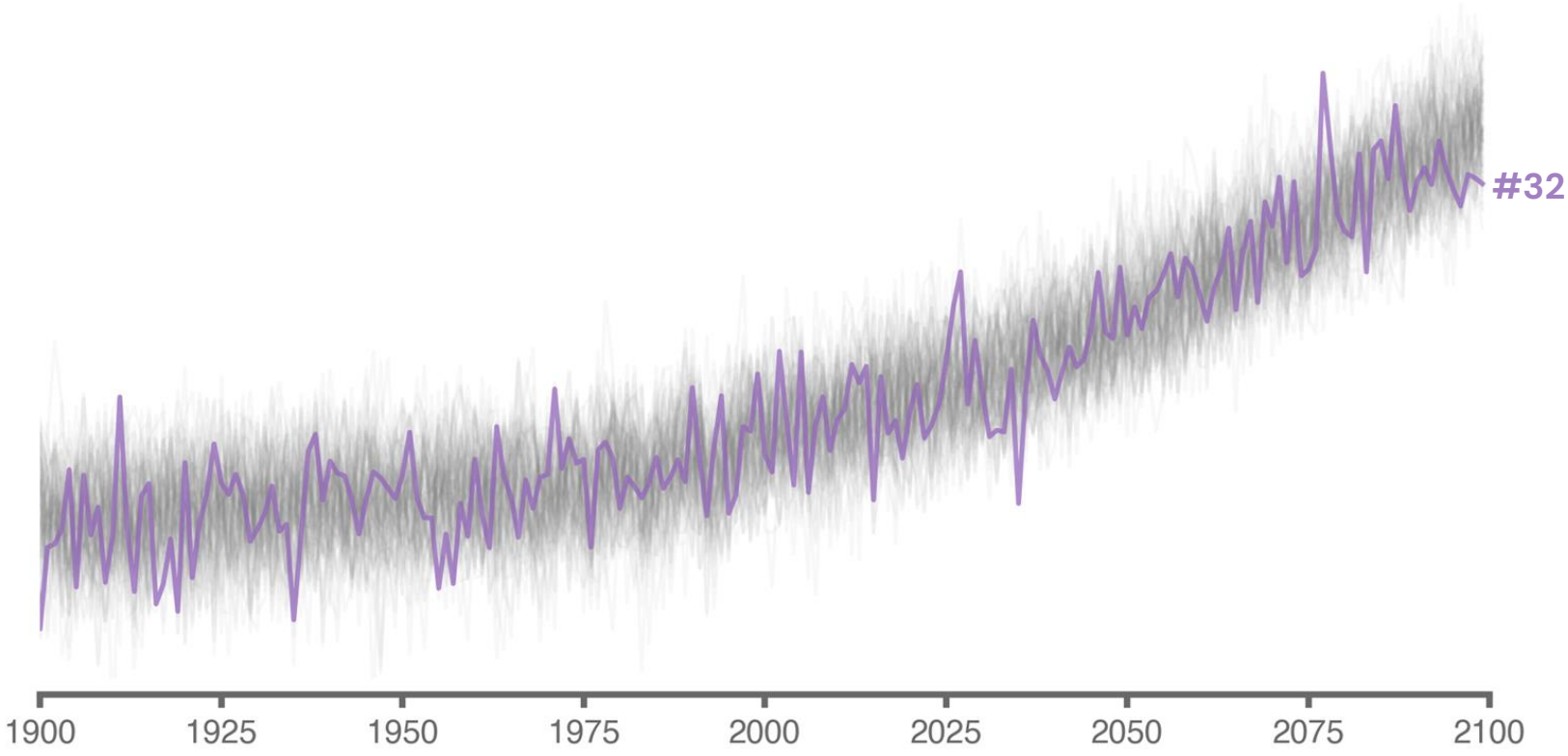
Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5



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



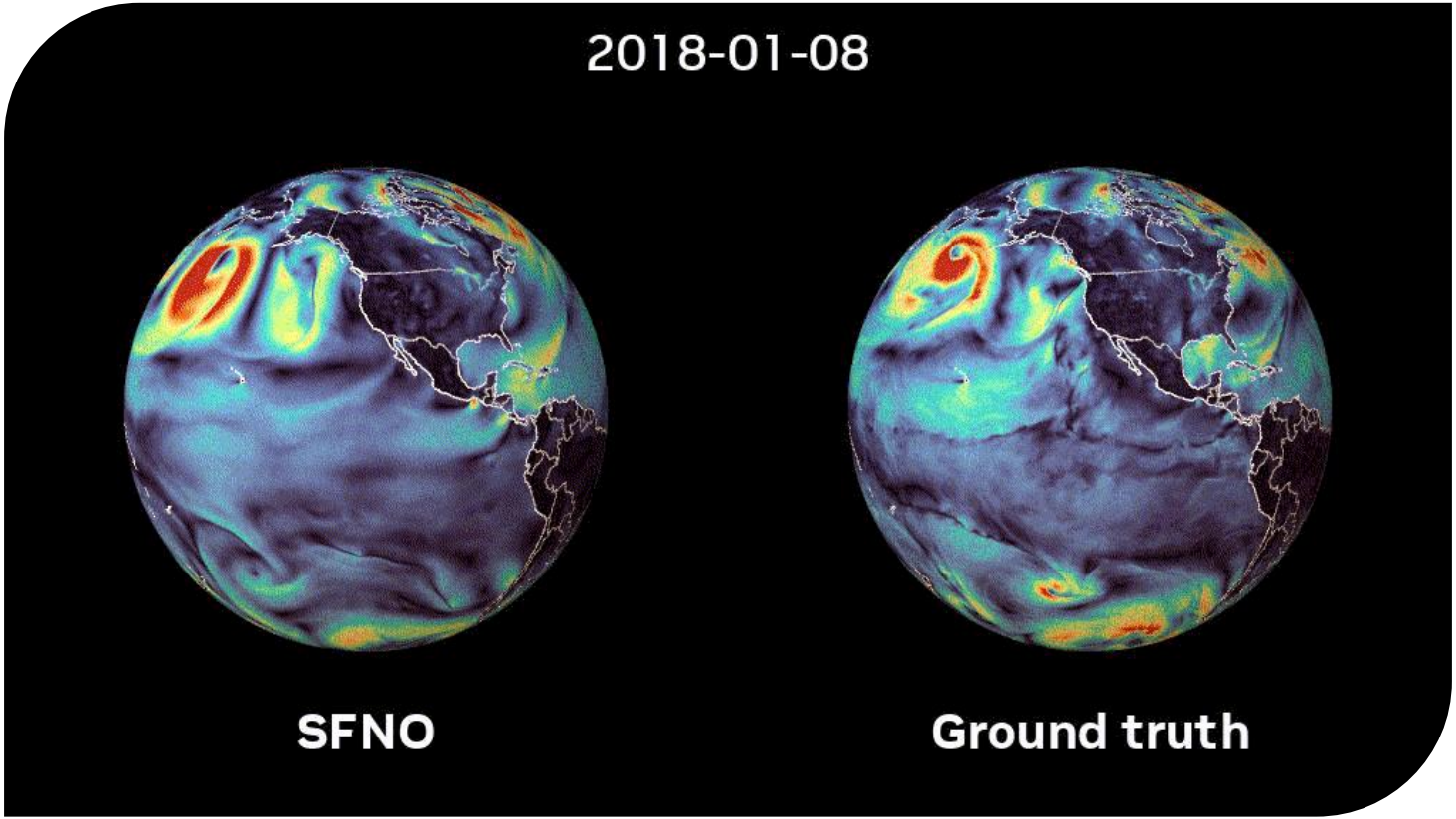
Surface temperature over Chicago, IL
MPI-ESM Large Ensemble; historical + RCP8.5



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

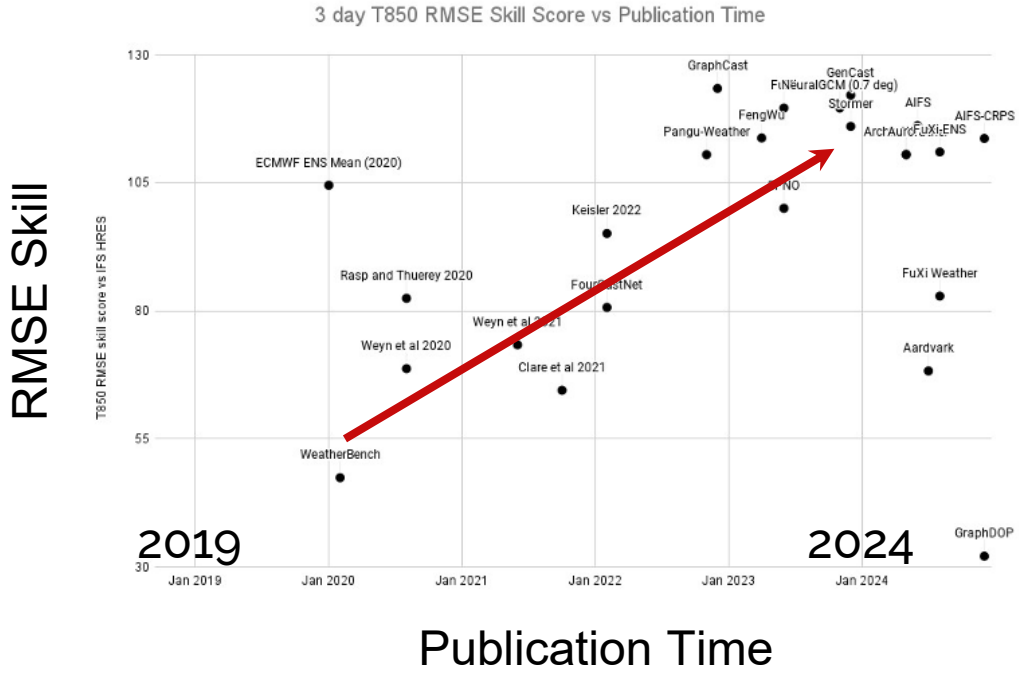
How do we sample the full range of possible futures efficiently?

AI weather & climate models are here



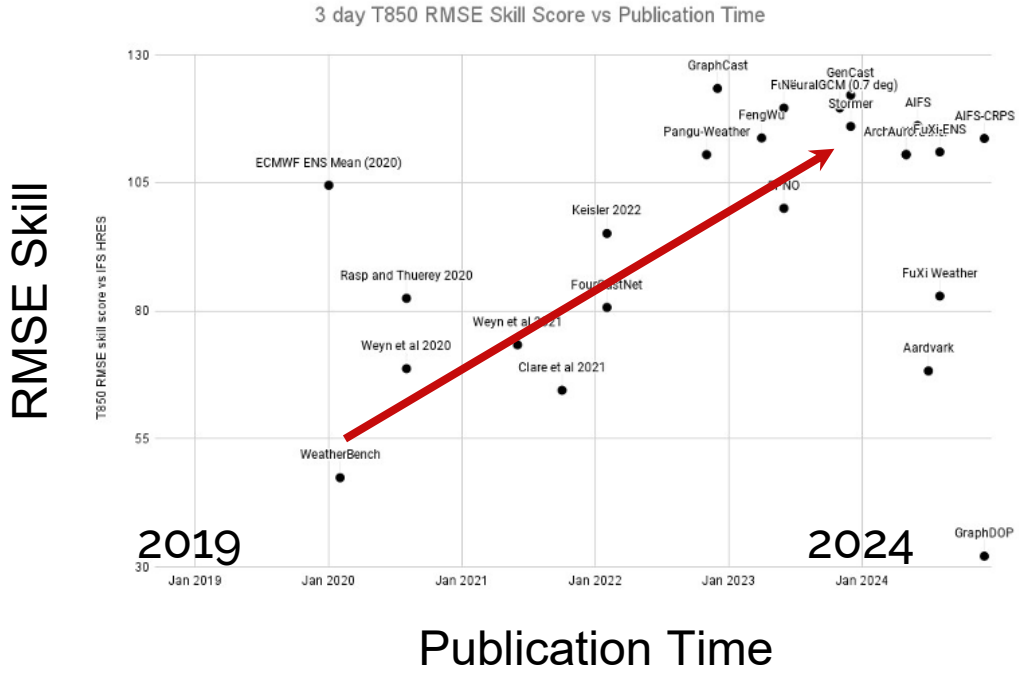
Created by NVIDIA [link]

…and they have improved dramatically



Source: Stephan Rasp

…and they have improved dramatically



We can now run *thousands* of simulations — enough to capture even the rarest events.

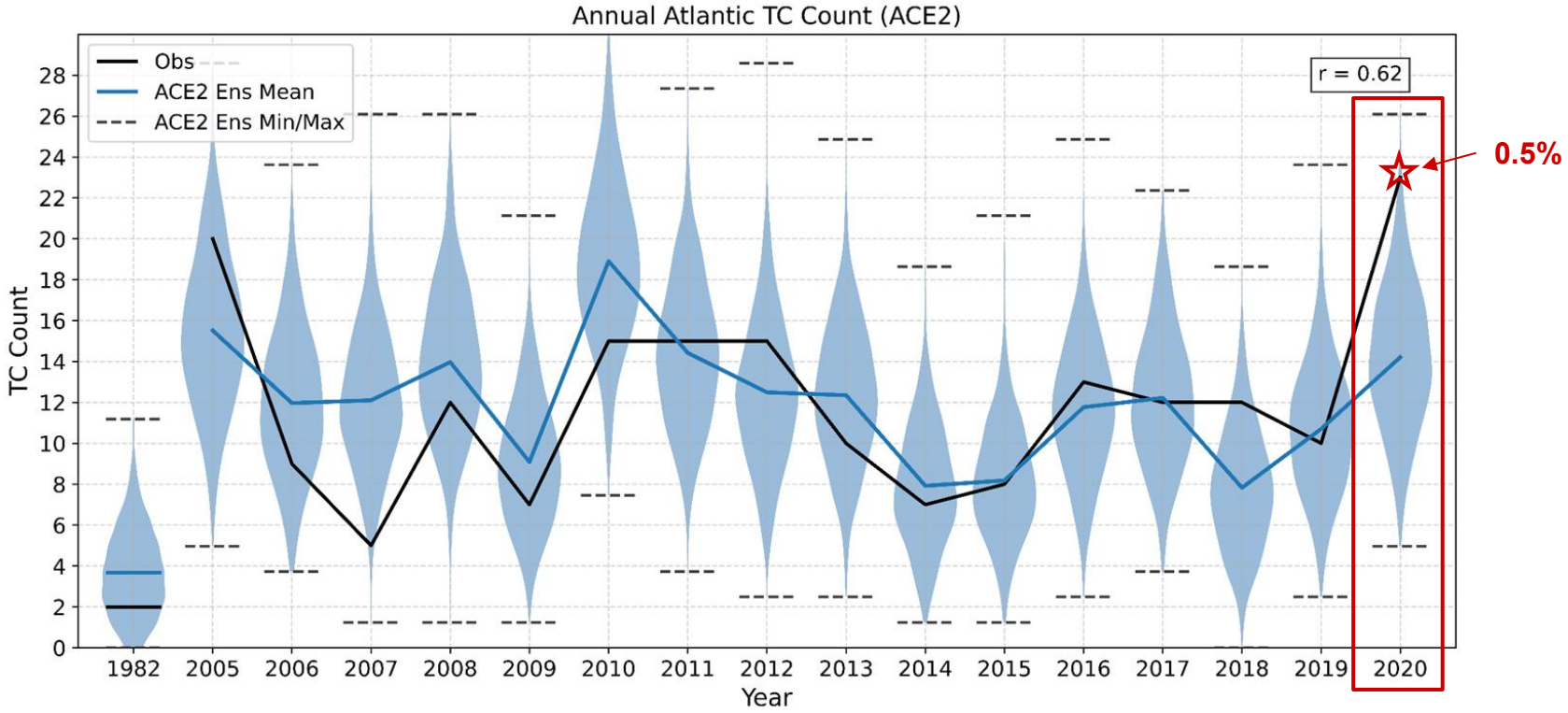
Source: Stephan Rasp

AI models are trained on average performance. Risks live in the tails.

Most AI weather models are trained to minimize average error across all conditions and learn to hedge toward the middle. For day-to-day forecasting, that's often fine.

For planning around the events that stress your grid it's a fundamental limitation. **That's why testing the tails explicitly is essential.**

Capturing the full range of TC seasons with AI

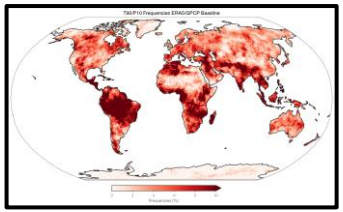


1000 AI ensemble members per season. Black = observed.
2020 was a 0.5% event.


Levin, Chien et al. (2026)

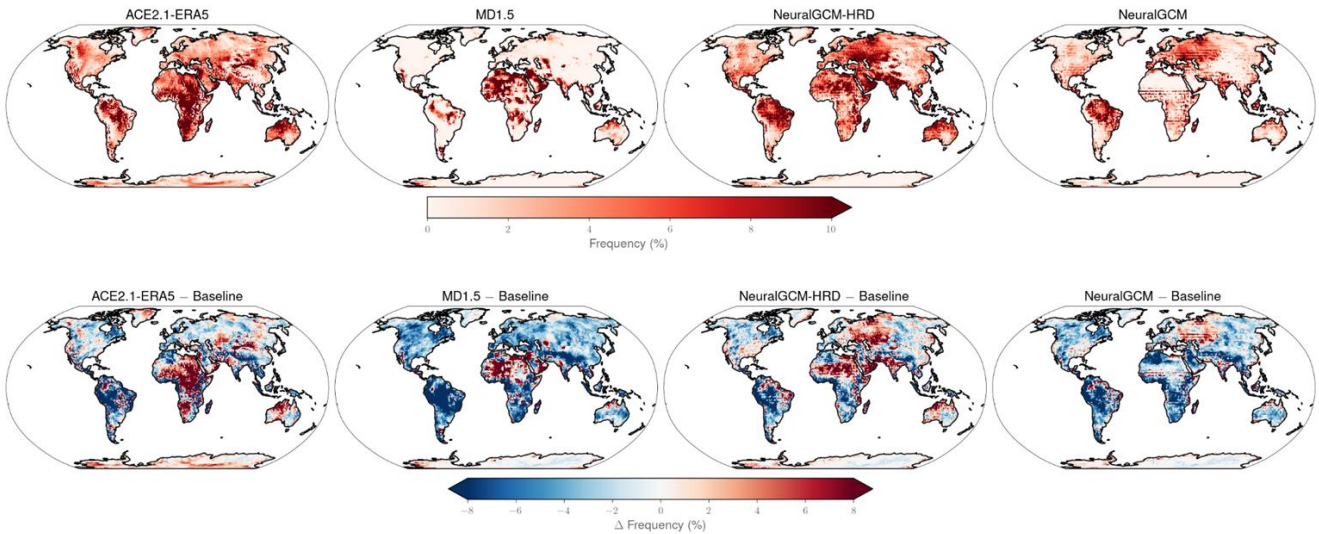
AI can sample rare events — but are the joint statistics right?

Observations



Hot-Dry Compound Event Frequency

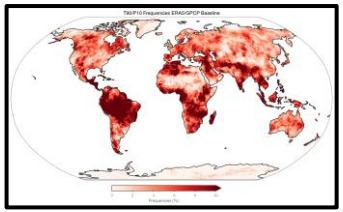
[observational percentiles]



Dwyer et al. (in prep)

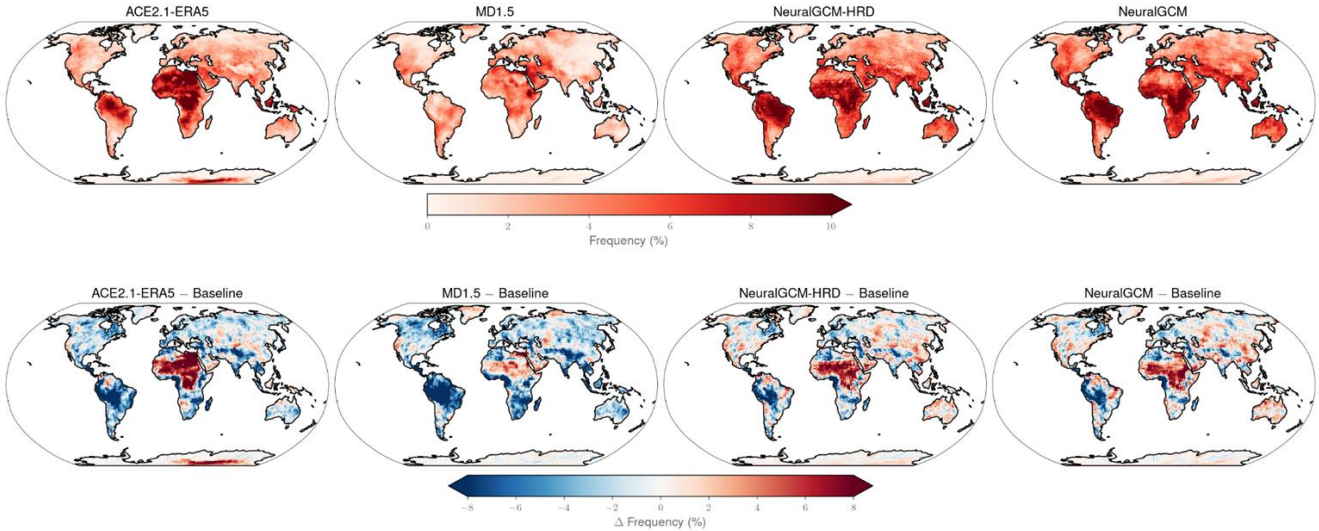
AI can sample rare events — but are the joint statistics right?

Observations



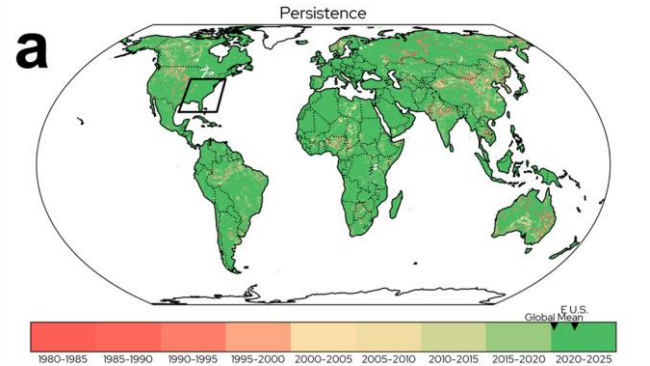
Hot-Dry Compound Event Frequency

[bias corrected via AI-model-relative percentiles]

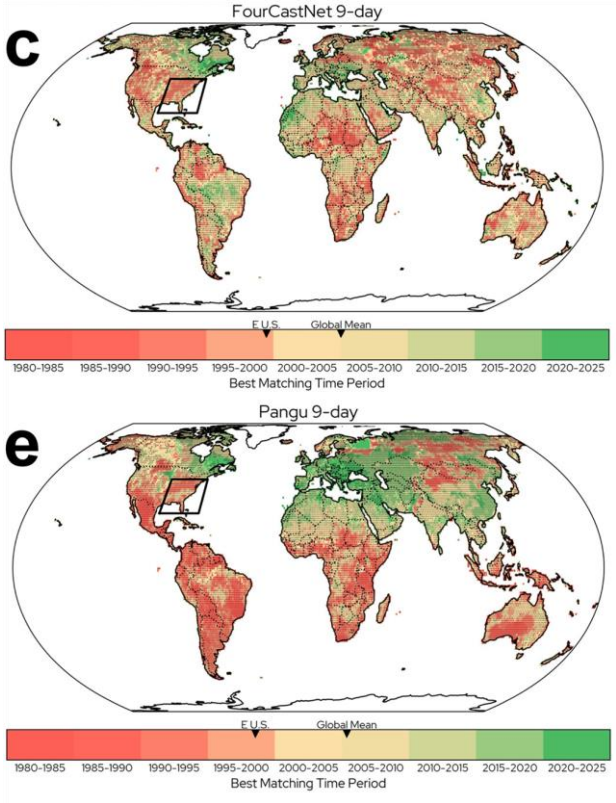


Dwyer et al. (in prep)

AI models forecast the future with yesterday's climate

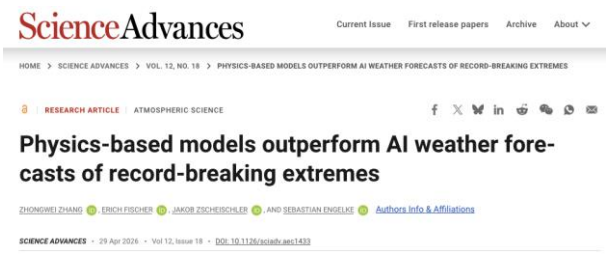


Colors show which historical period AI models drift toward when run forward 9 days.



Landsberg & Barnes (2026)

AI weather models struggle with extremes



Science Advances

Current Issue First release papers Archive About

HOME > SCIENCE ADVANCES > VOL. 12, NO. 18 > PHYSICS-BASED MODELS OUTPERFORM AI WEATHER FORECASTS OF RECORD-BREAKING EXTREMES

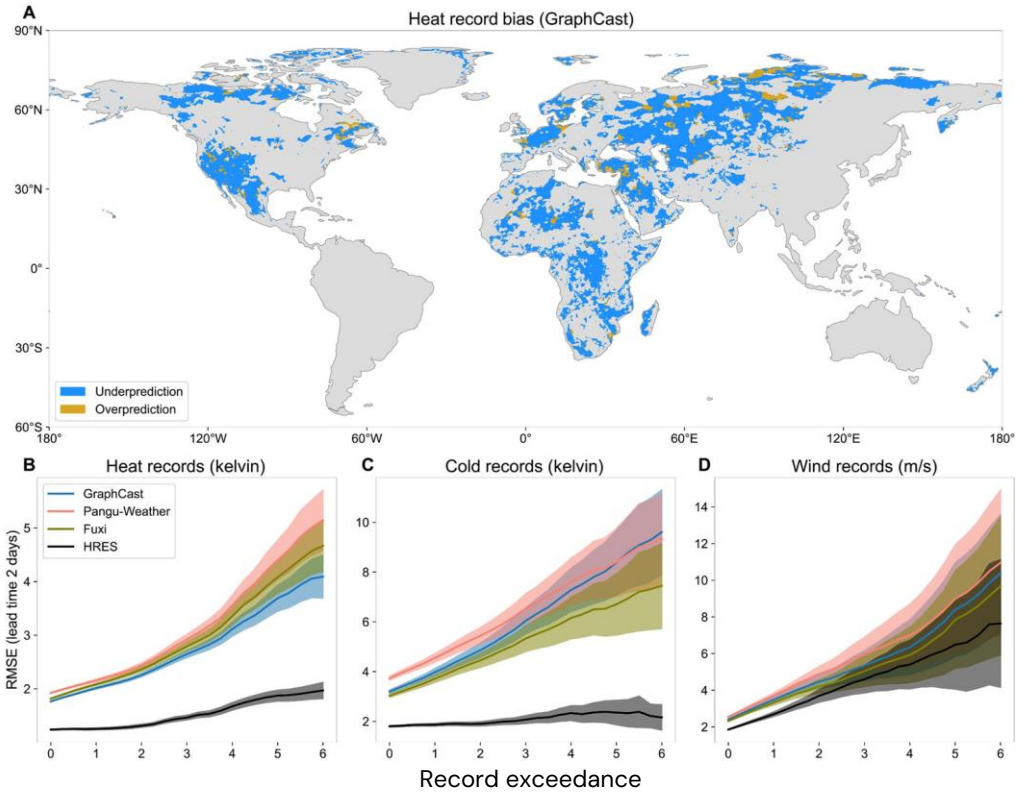
RESEARCH ARTICLE | ATMOSPHERIC SCIENCE

Physics-based models outperform AI weather forecasts of record-breaking extremes

ZHONGWEI ZHANG, ERICH FISCHER, JAKOB ZSCHEISCHLER, AND SEBASTIAN ENGELKE Authors Info & Affiliations

SCIENCE ADVANCES • 29 Apr 2026 • Vol 12, Issue 18 • DOI: 10.1126/sciadv.aec1433

The more extreme the event, the larger the gap between AI and physics-based models.

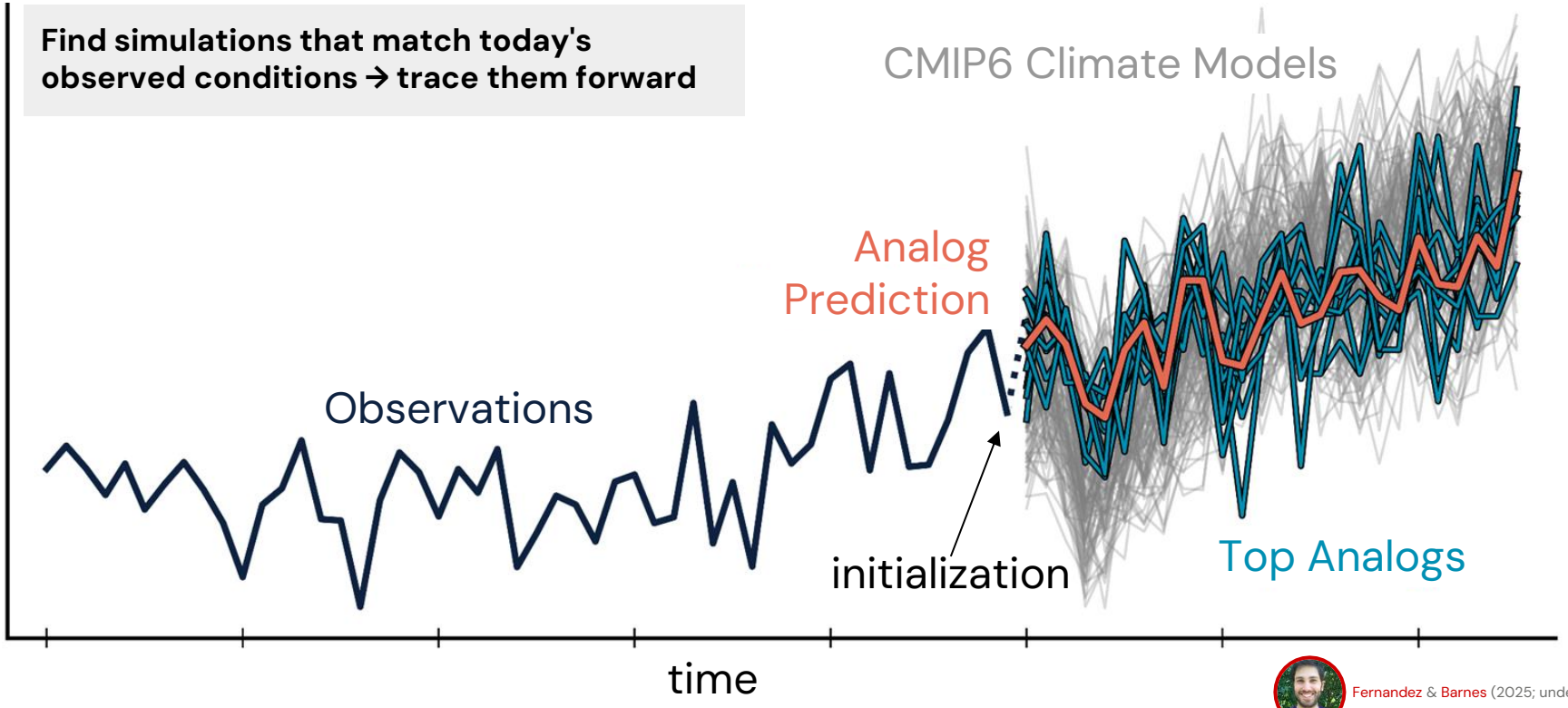


Zhang et al. (2026)

Can we predict the multi-year trajectory?

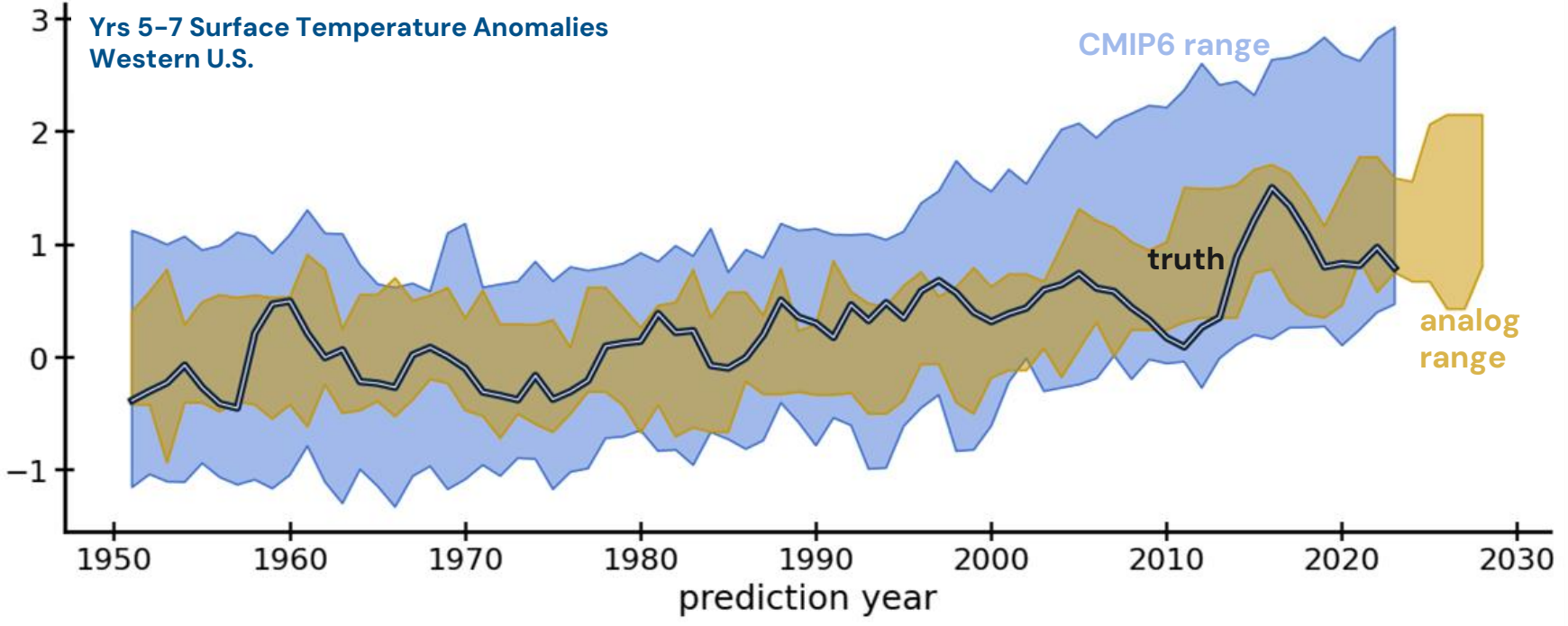
What does the future look like, starting from where we are now?

Find simulations that match today's observed conditions → trace them forward



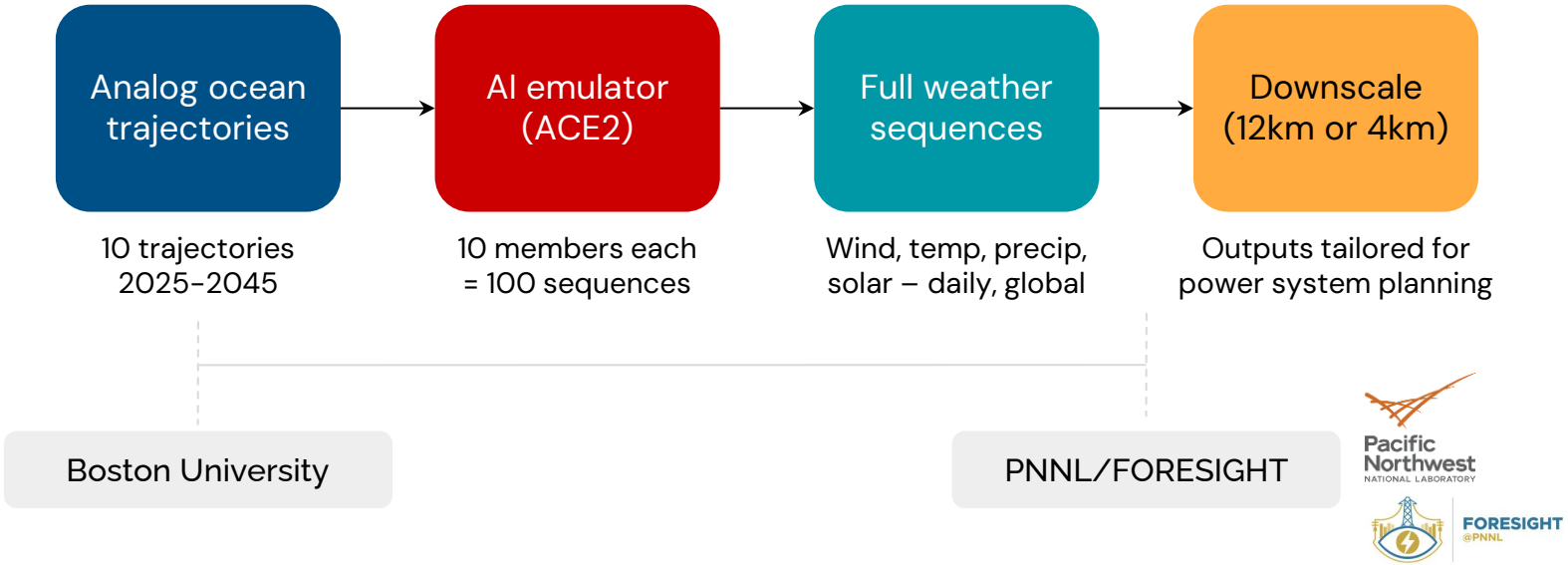
Fernandez & Barnes (2025; under review)

Starting from today's climate narrows the range of what's ahead



Fernandez & Barnes (2025; under review)

From prediction to planning

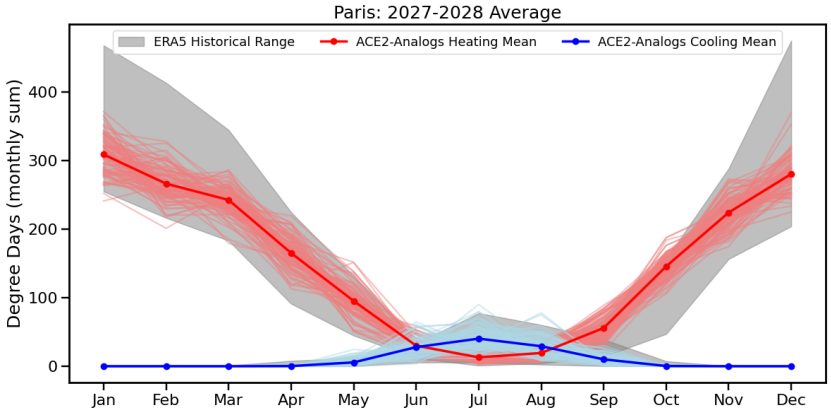


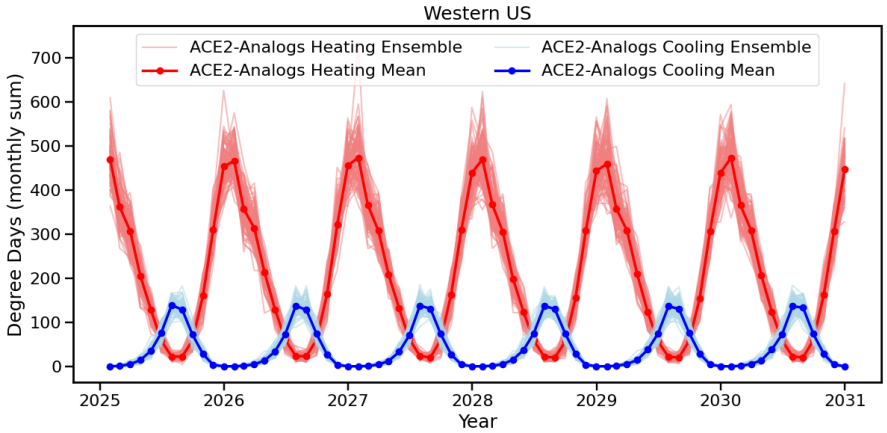
100 plausible 20-year weather futures at power-system resolution

Predicting future energy-relevant extremes

Analog Ocean Trajectories → ACE2 AI Emulator → Full Weather Trajectories

Predicted Heating & Cooling Degree Days





Fernandez et al. (in prep)

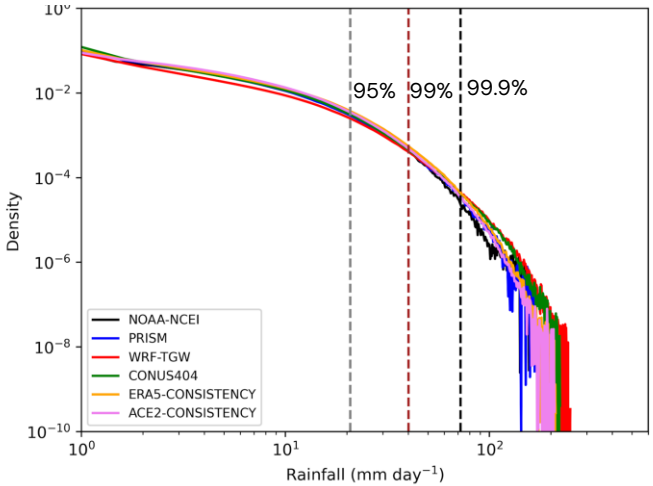
Downscaling for energy system applications

- Outputs tailored for power system planning (at specific locations, or aggregated to watershed level for example)
- Initial results are at 12km scale, also have an emulator to 4km
- Built an ensemble of 10 'historical' trajectories and 100 'near future' (2025-2045) trajectories (Historical shown)
- Evaluate against **Observations** (Blue, Black), **Training data**, **4km Regional Model**
- Emulator reproduces test data for Solar, Precipitation, Wind, and Temperature (precip and max temp shown)
- Precipitation extreme frequency well simulated to 99.9% level

Pacific Northwest NATIONAL LABORATORY

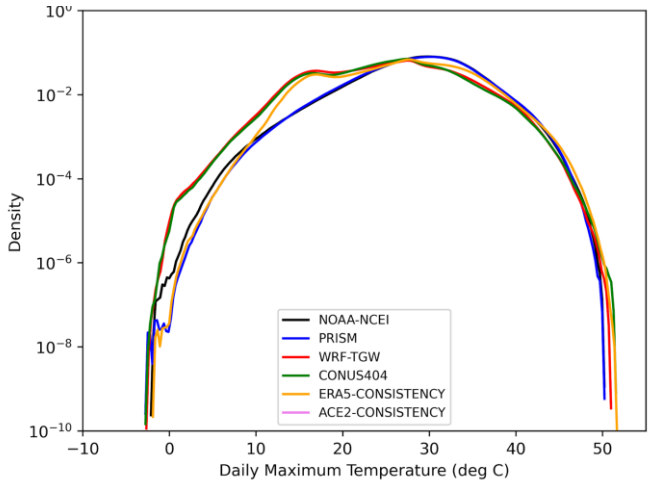
FORESIGHT @PNNL

Historical Validation of PDF of Annual Rainfall over Pacific NW



Obs 1
Obs 2
12km Training Model Eval
4km Model Eval
Downscaled Reanalysis
Downscaled AI Weather

Historical Validation of PDF of Daily Max Temp over CONUS

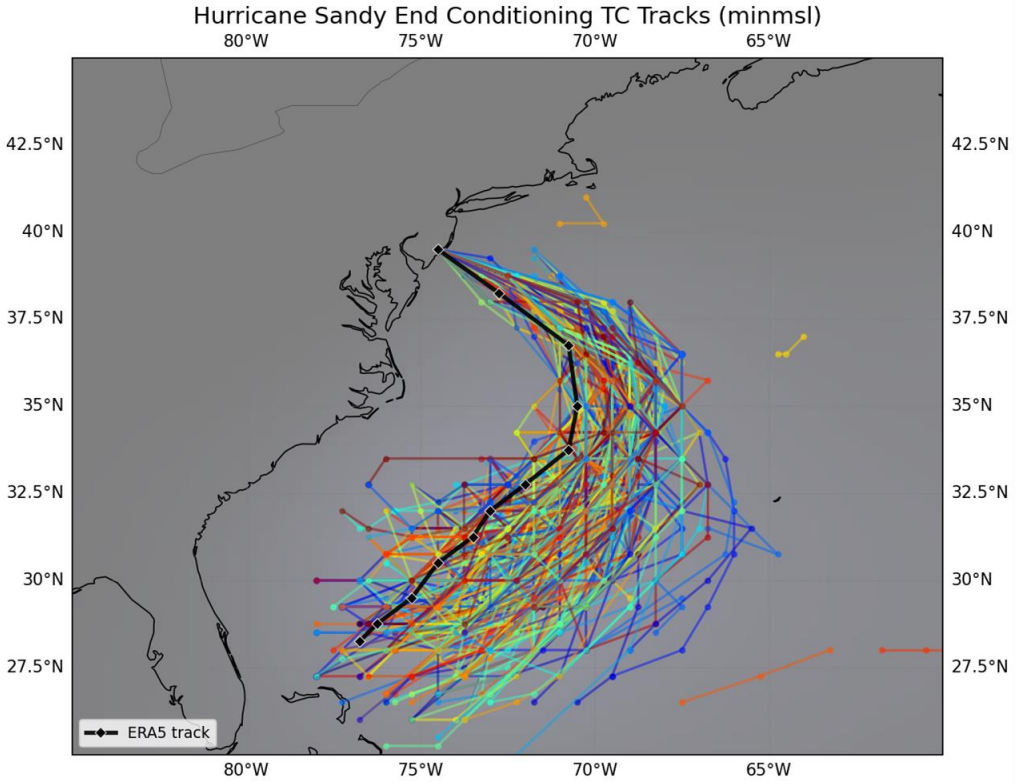


PNNL FORESIGHT Team: A. Gettelman, C. Bracken, H. Tran, A. Geiss, C. Tebaldi, S. Taraphdar

AI for “what if” scenario planning

1000 ways to make landfall

1000 AI-generated atmospheric histories that all lead to Hurricane Sandy making landfall



Lin et al. (in prep)

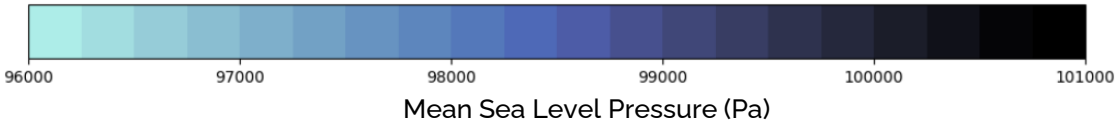
Steering trajectories of extremes for planning

Trajectory 1

Trajectory 2

Trajectory 3

*additional smoothing provided using cubic interpolation



Lin et al. (in prep)

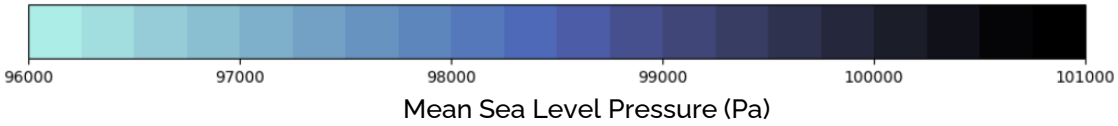
Steering trajectories of extremes for planning

Trajectory 1

Trajectory 2

Trajectory 3

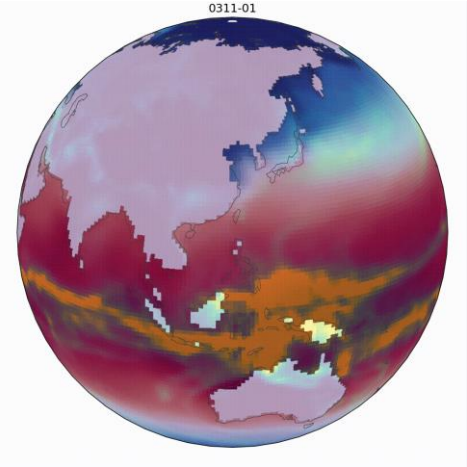
*additional smoothing provided using cubic interpolation



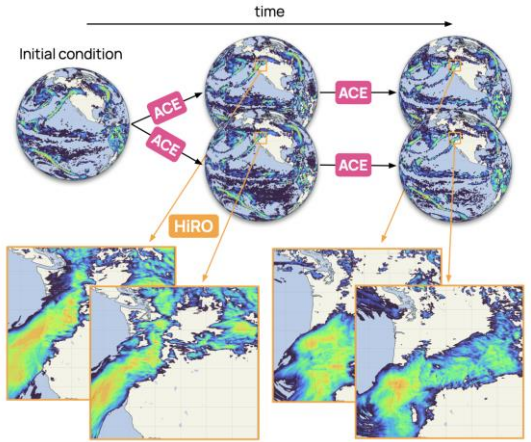
Lin et al. (in prep)

What's coming next from the community

- Fully coupled AI climate models (atmosphere + ocean) are on there way; so there will be less need to prescribe ocean temperatures in the future
- AI-native downscaling to km-scale wind, solar, precipitation are on their way
- The pace of development is outrunning community evaluation — careful validation is more important than ever



Coupled AI emulator SamudrACE
M²LInES and Ai2



HiRO-ACE from Ai2
https://allenai.org/blog/hiro-ace

Things to remember:

1. AI can generate thousands of plausible futures in minutes, **enabling us to characterize risk at high resolution quickly.**

2. But they are **trained to get the average right, not the extremes**. More research is needed.

3. They are not **explicitly trained to preserve the correlations** between variables but efforts are underway.

Thank you.

Elizabeth A. Barnes, PhD
Dalton Family Chair in Environmental Data Science & Sustainability
Professor, Computing & Data Sciences
Professor, Earth & Environment
Boston University
eabarnes@bu.edu | www.barnes-research.com
https://github.com/eabarnes1010

BOSTON UNIVERSITY