Advancing Deep-Time Climate Reconstruction with a New Online Paleoclimate Data Assimilation Approach in CESM

An Ongoing Research of the *paleoWeather* Project

***Feng Zhu*[1]**, *Dan Amrhein*[1], *Jiang Zhu*[1], *Jess Tierney*[2],
Anta-Clarisse Sarr[3], *Sophia Macarewich*[1], *Bette Otto-Bliesner*[1], *Chris Poulsen*[3],
Elke Zeller[2], *Maya Tessler*[2], *Dervla Meegan-Kumar*[4], *Jane W. Baldwin*[4], *and Ran Feng*[5]

1. NSF NCAR 2. U Arizona 3. U Oregon 4. UC Irvine 5. U Connecticut

Jun 11, 2025
CESM Workshop 2025

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

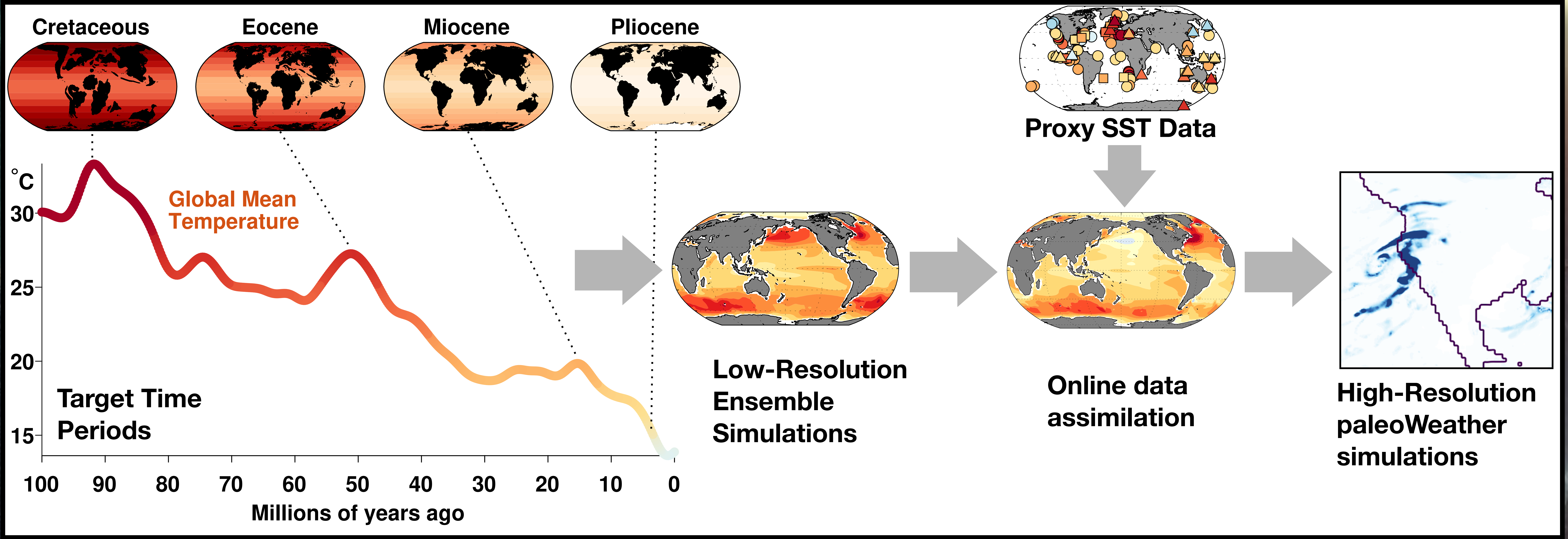
The *paleoWeather* Project

How do weather extremes like tropical cyclones and atmospheric rivers behave in warm climates?

What are the limits of habitability?

image by Dan Amrhein

The *paleoWeather* Project



How do weather extremes like tropical cyclones and atmospheric rivers behave in warm climates?

What are the limits of habitability?

image by Dan Amrhein

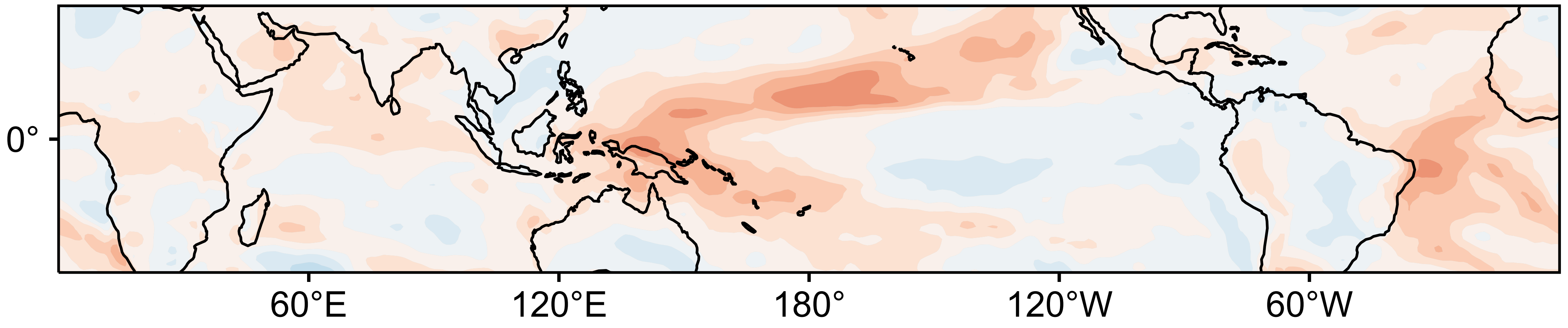
A Primer on Data Assimilation

A Bayesian Framework

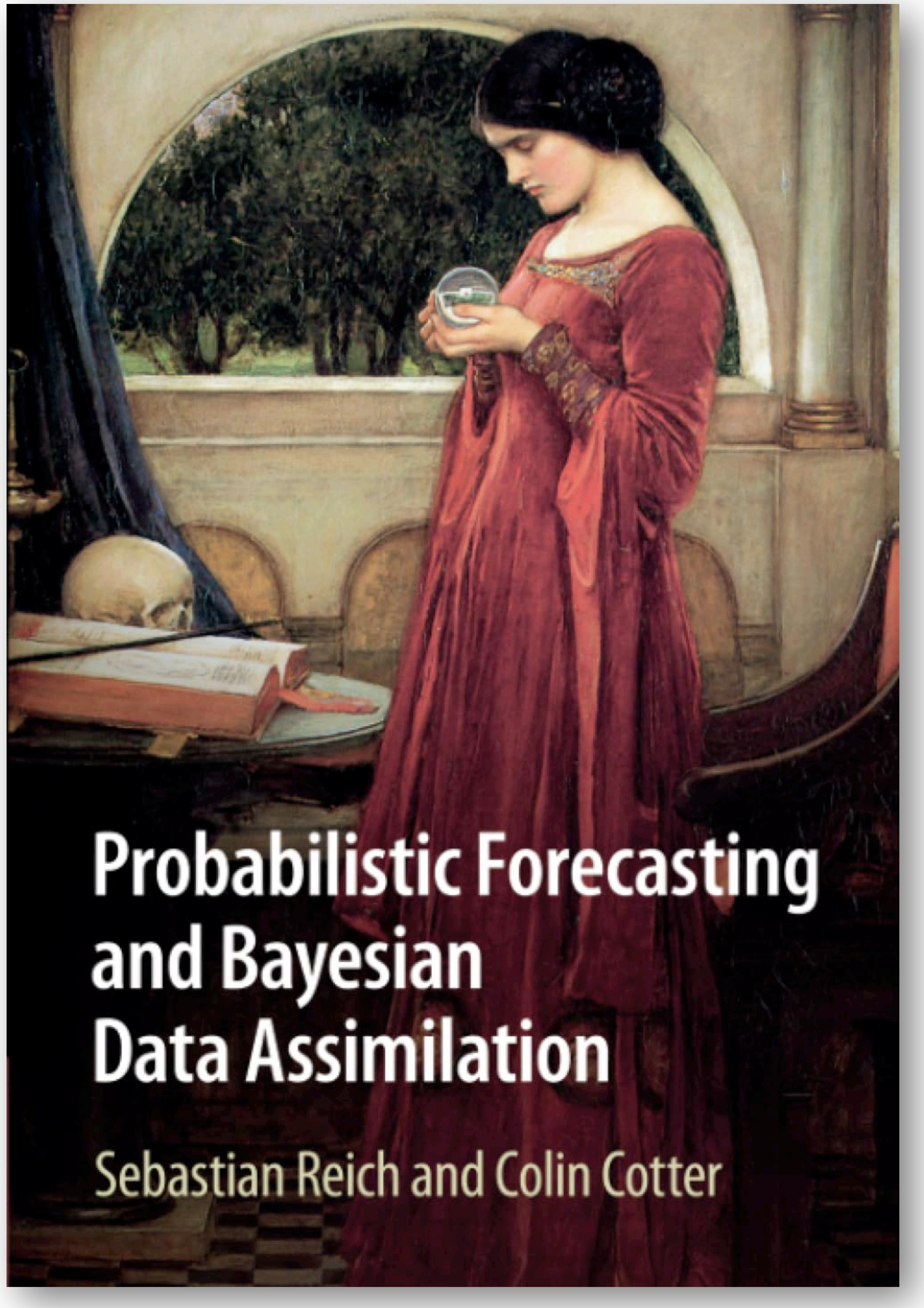
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

prior
P(A)

corr(prior, truth), mean=0.05



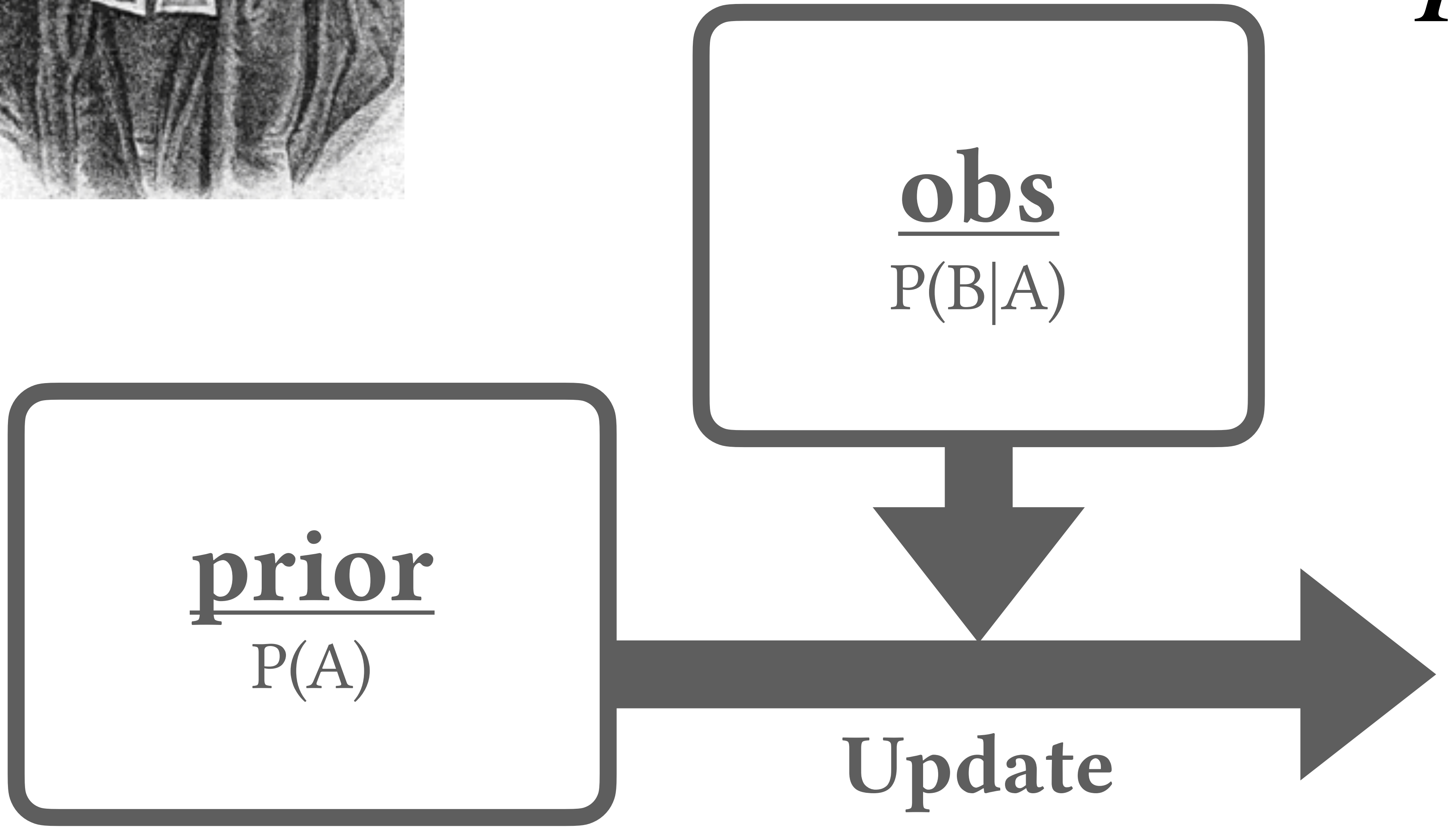
Zhu et al. (2024, *GMD*)

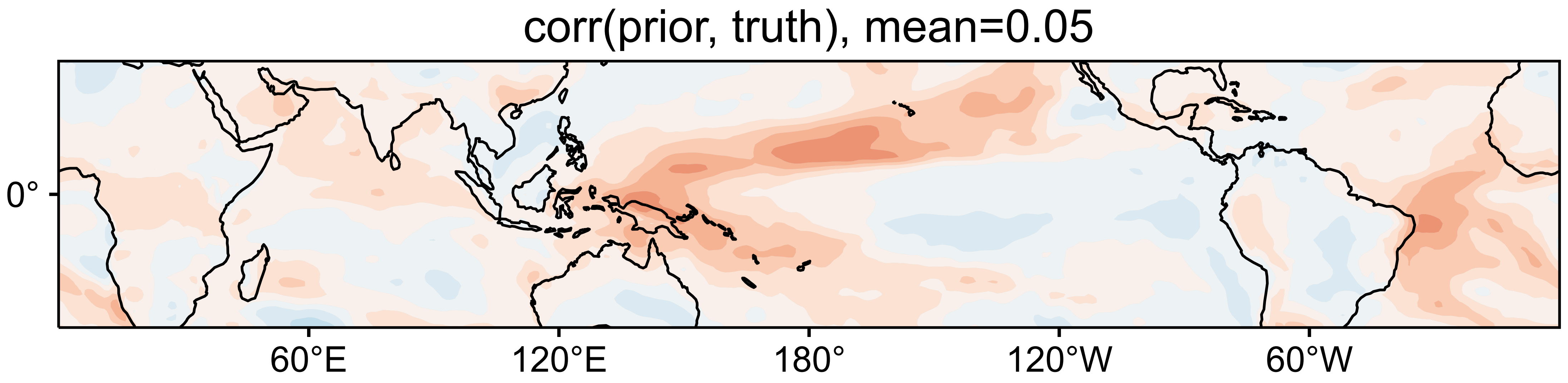


A Primer on Data Assimilation

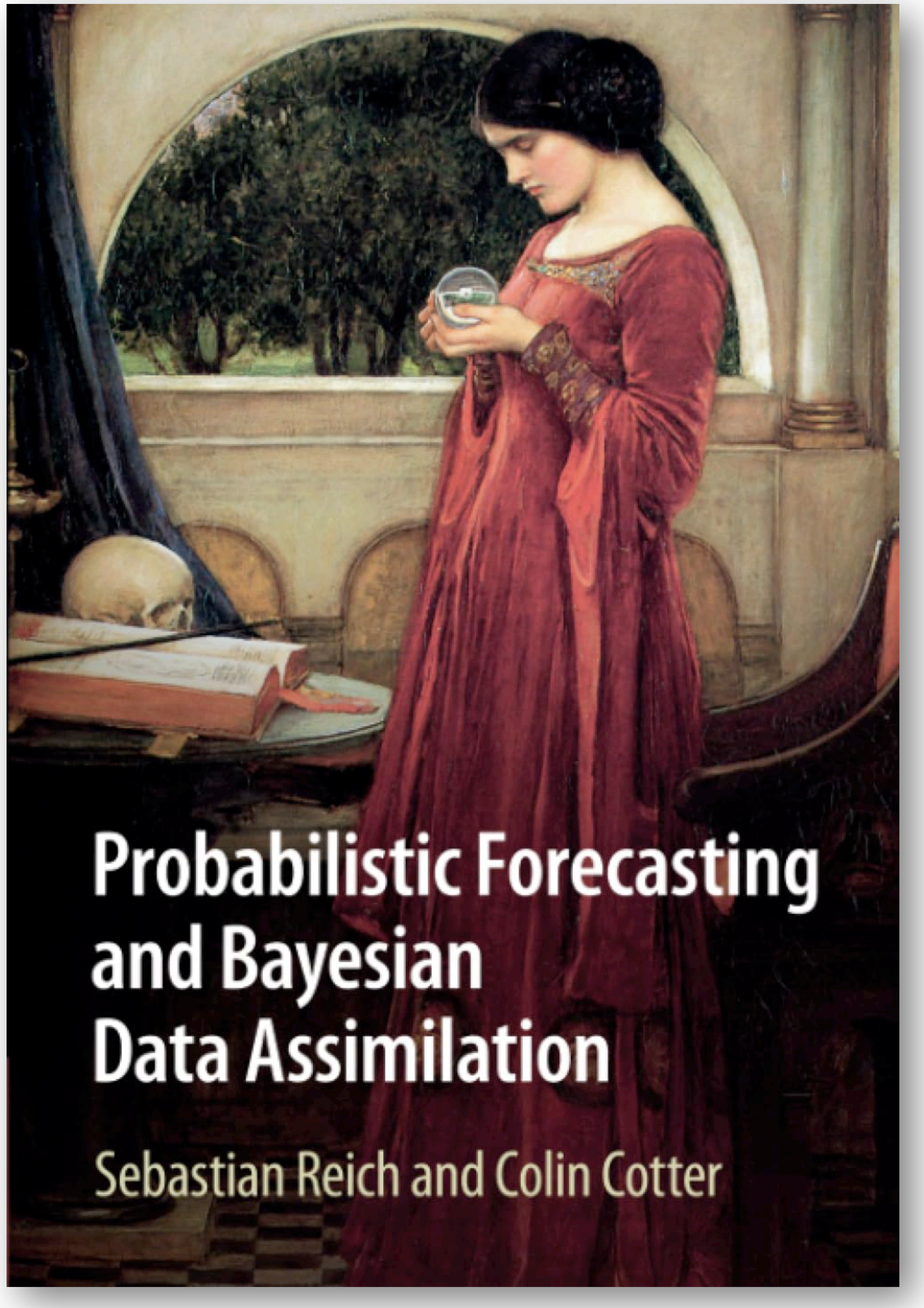
A Bayesian Framework

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$





Zhu et al. (2024, *GMD*)



A Primer on Data Assimilation

A Bayesian Framework

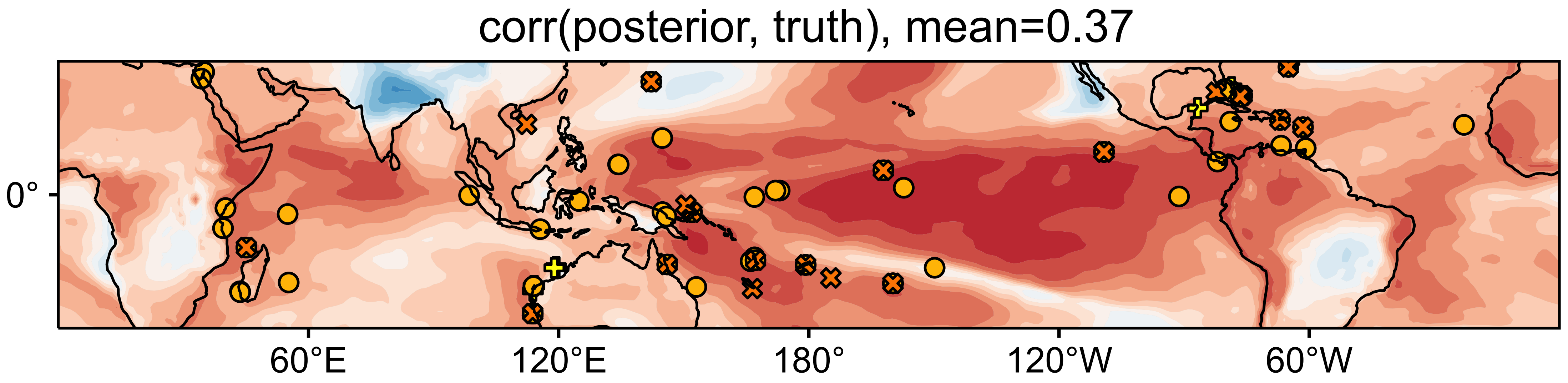
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

obs
P(B|A)

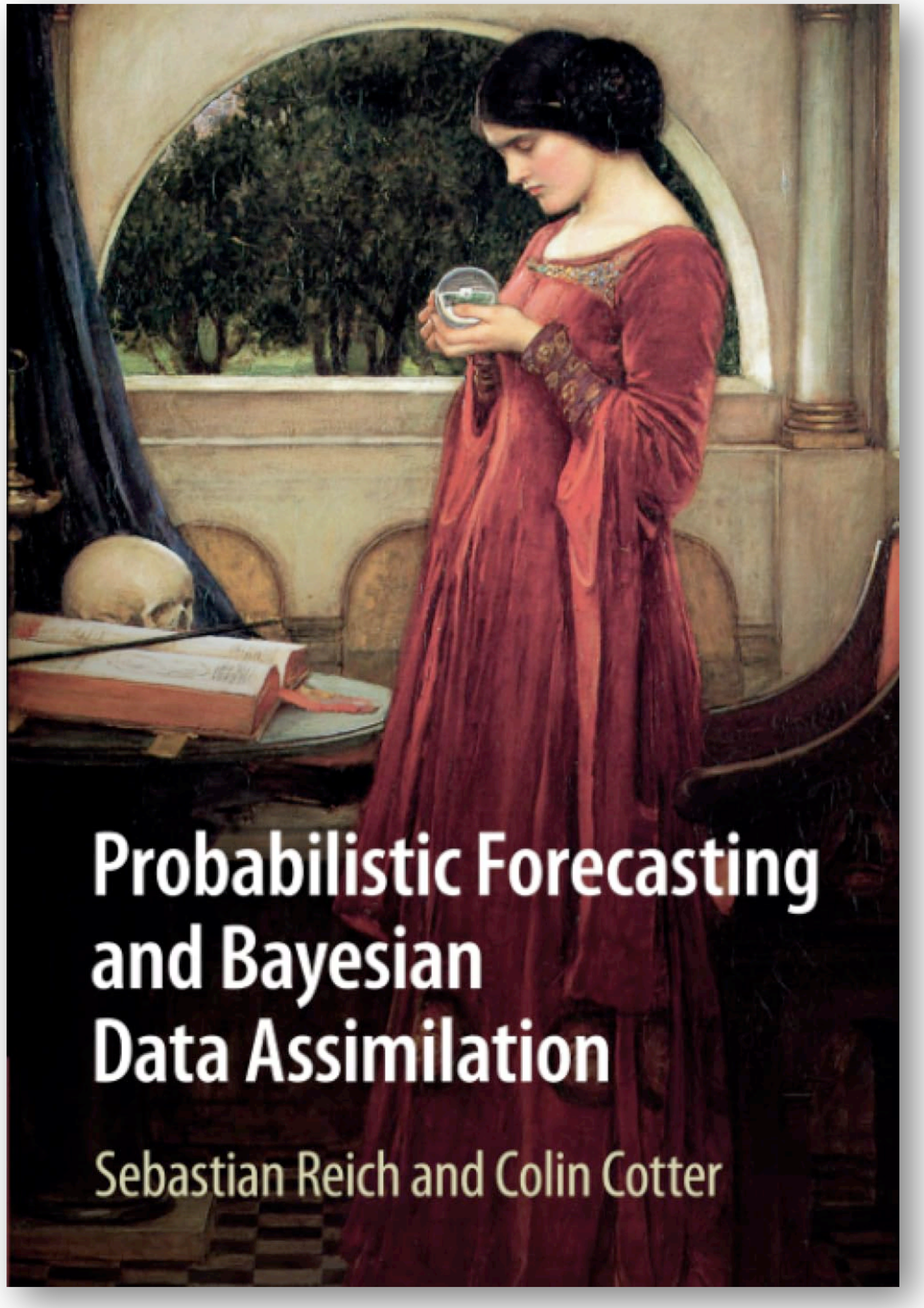
prior
P(A)

Update

posterior
P(A|B)



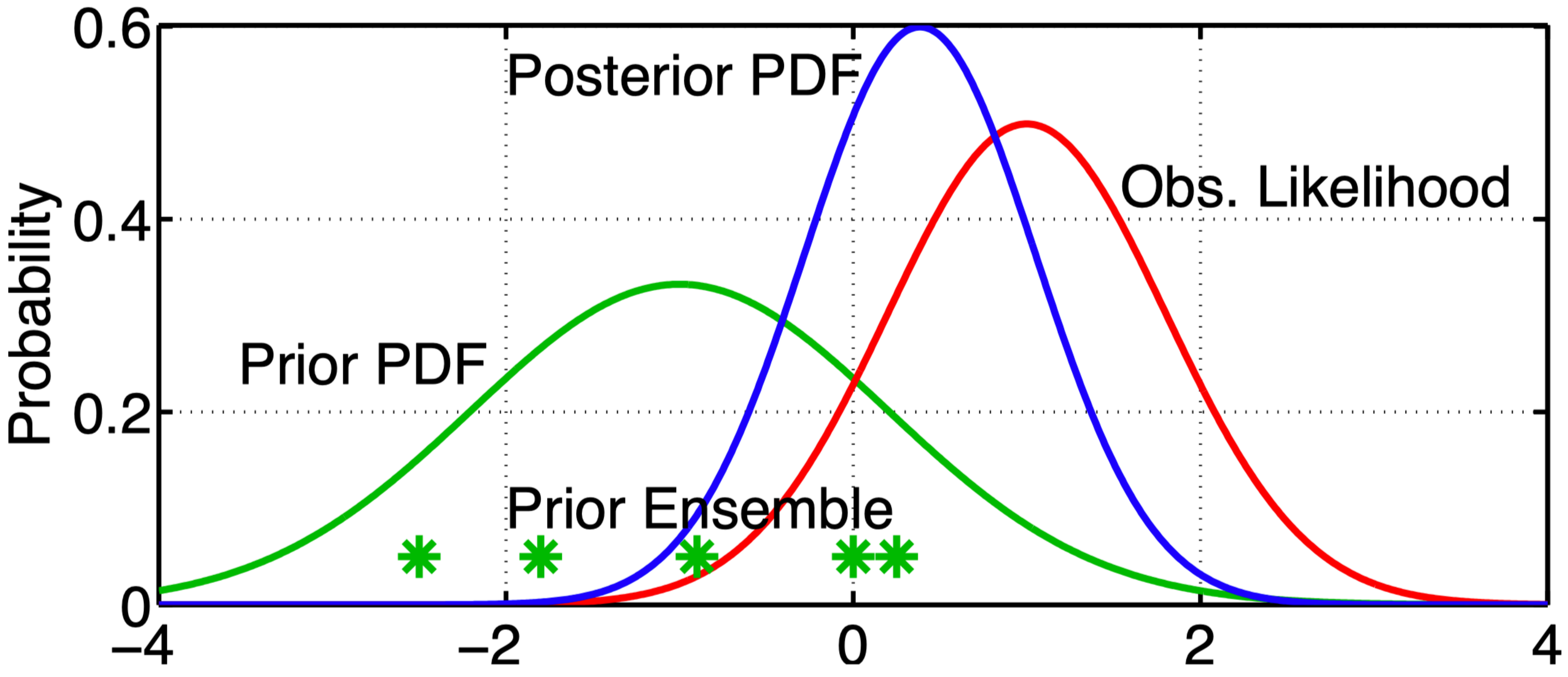
Zhu et al. (2024, *GMD*)



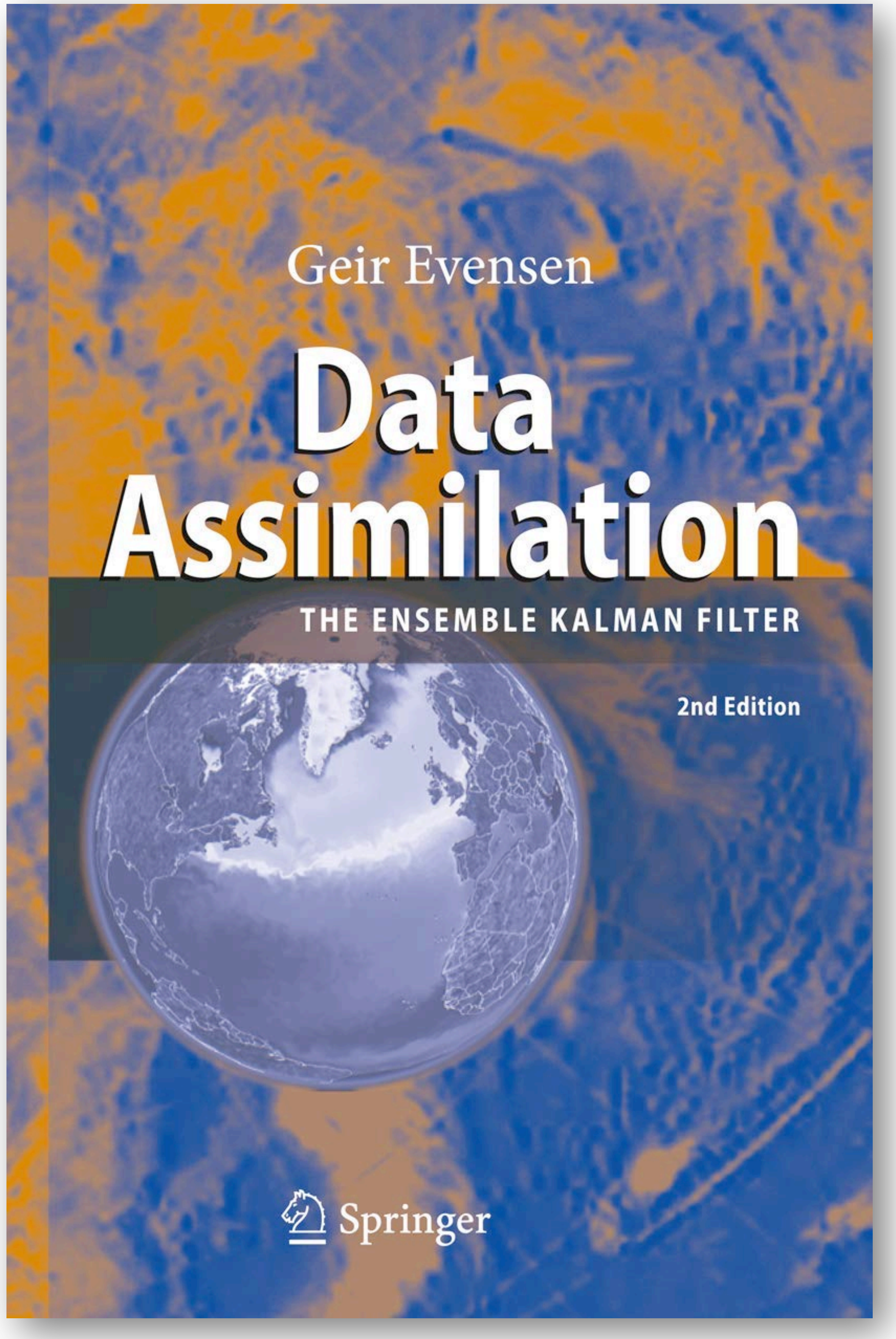
A Primer on Data Assimilation

A Bayesian Framework

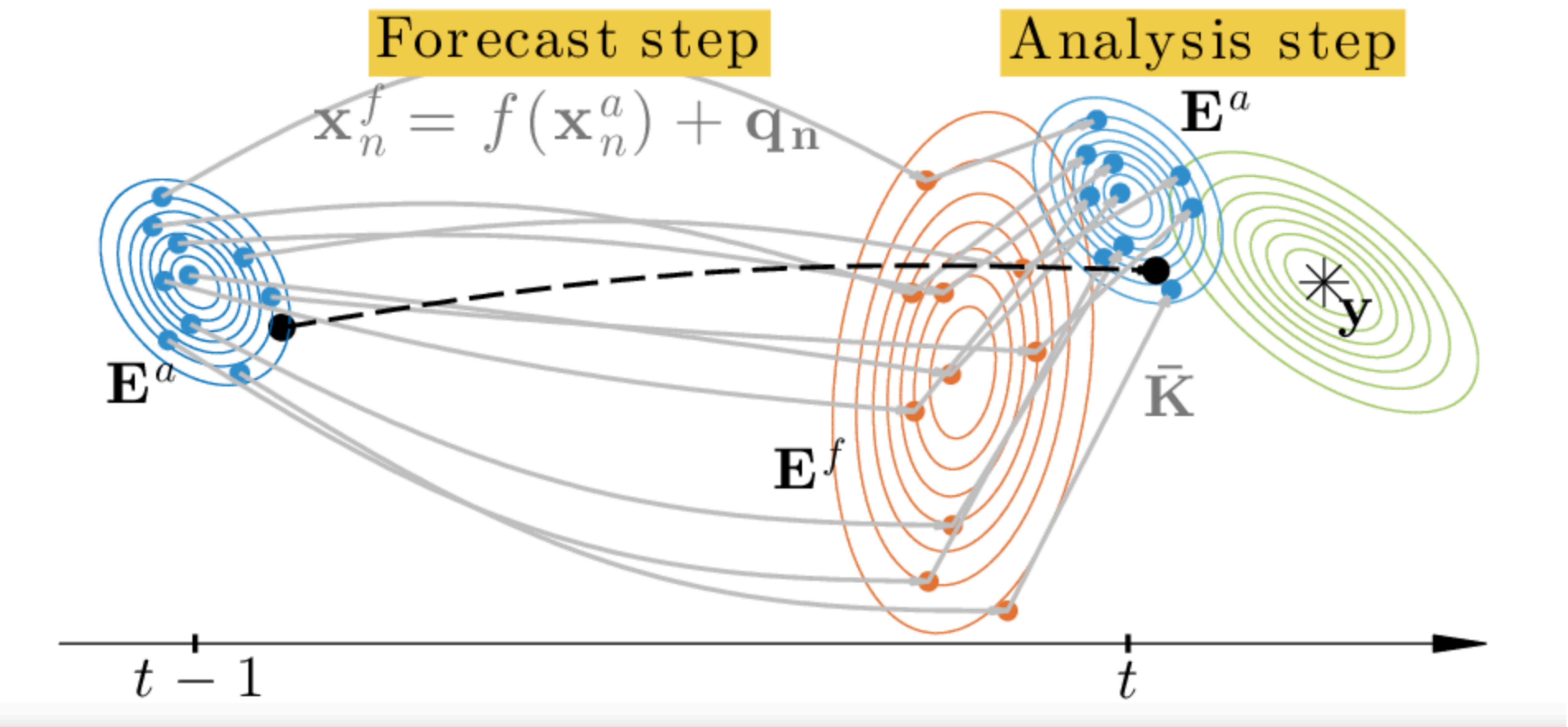
$$P(A\,|\,B) = \frac{P(B\,|\,A)P(A)}{P(B)}$$



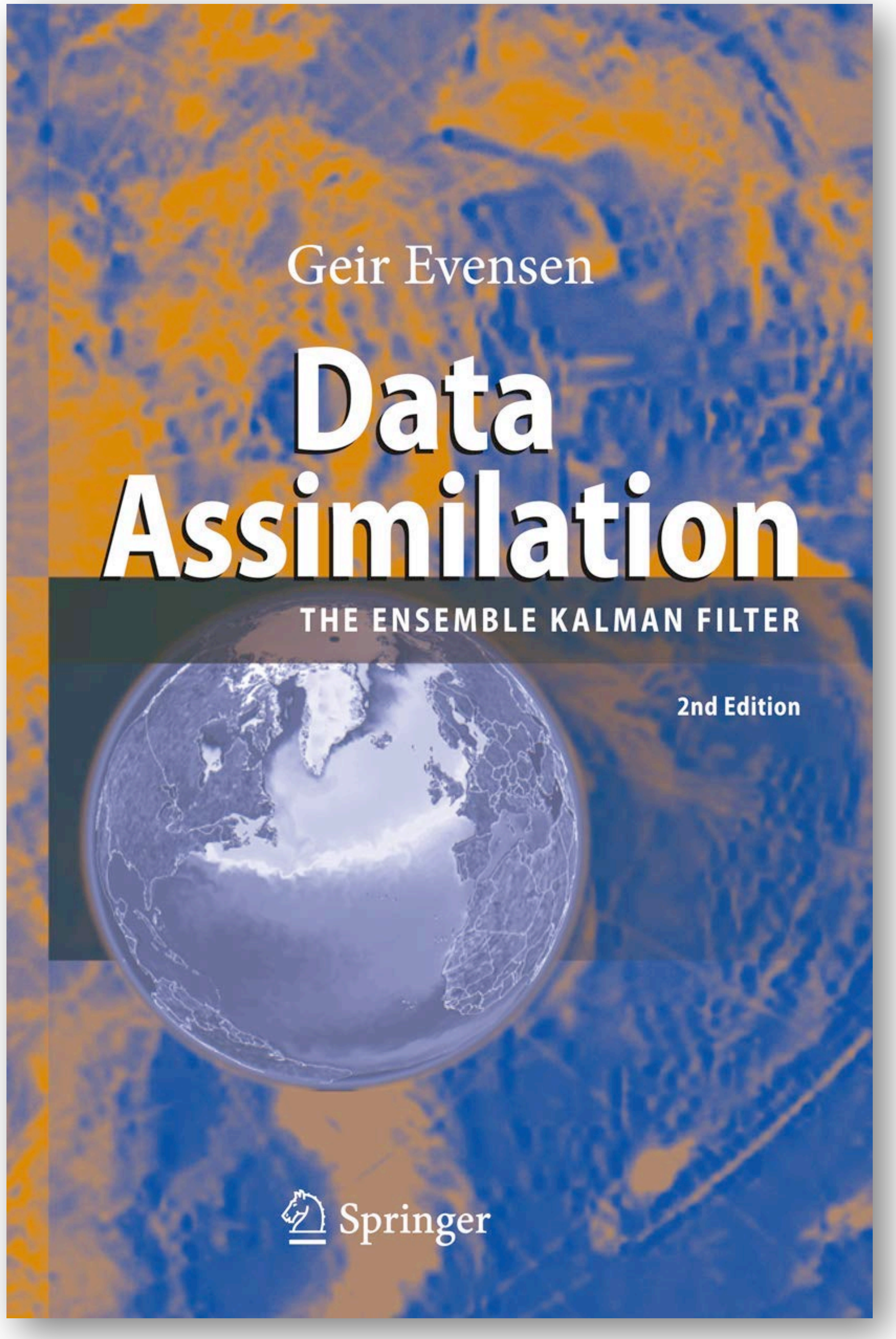
(DART Tutorial @NCAR)



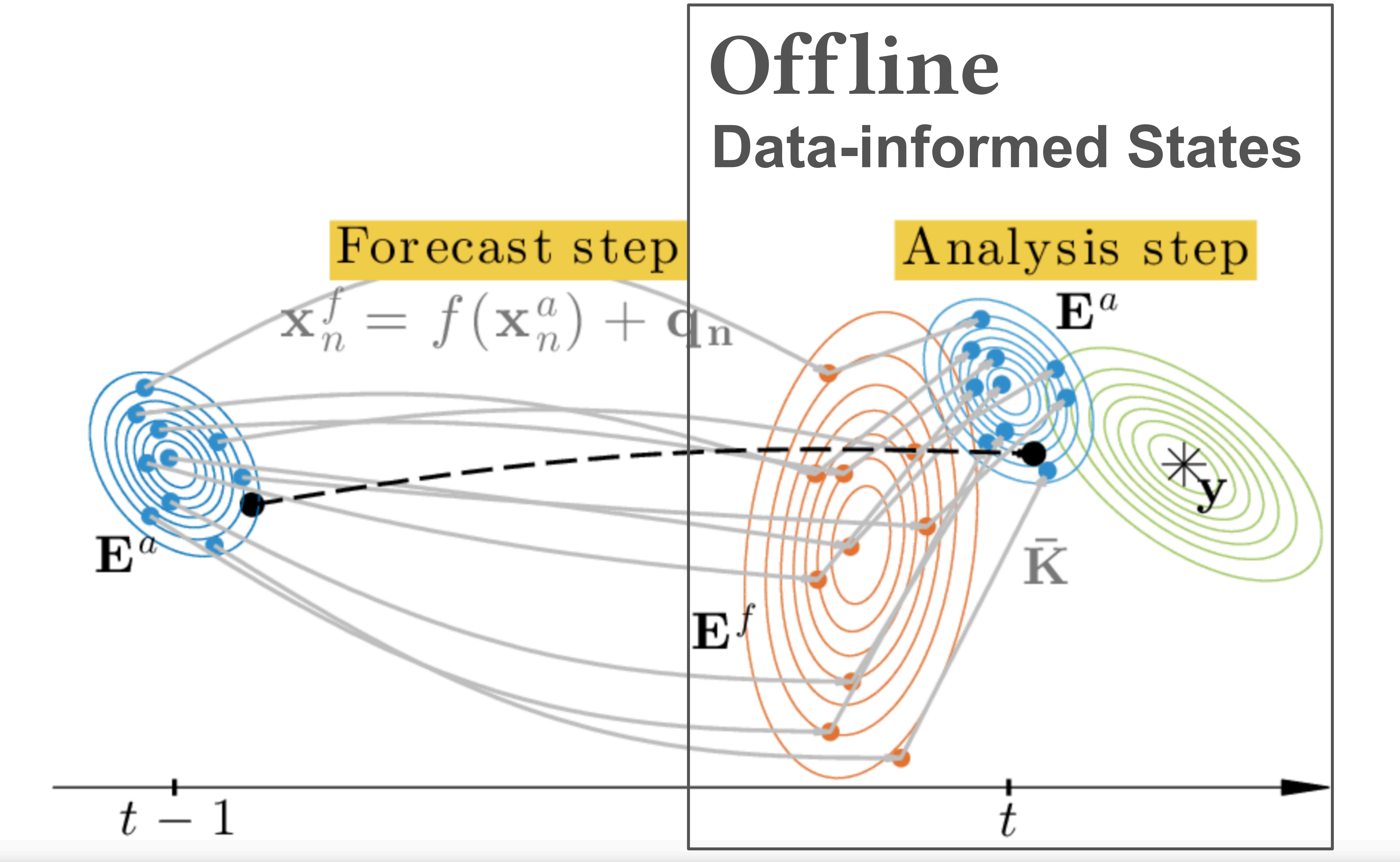
A Primer on Data Assimilation



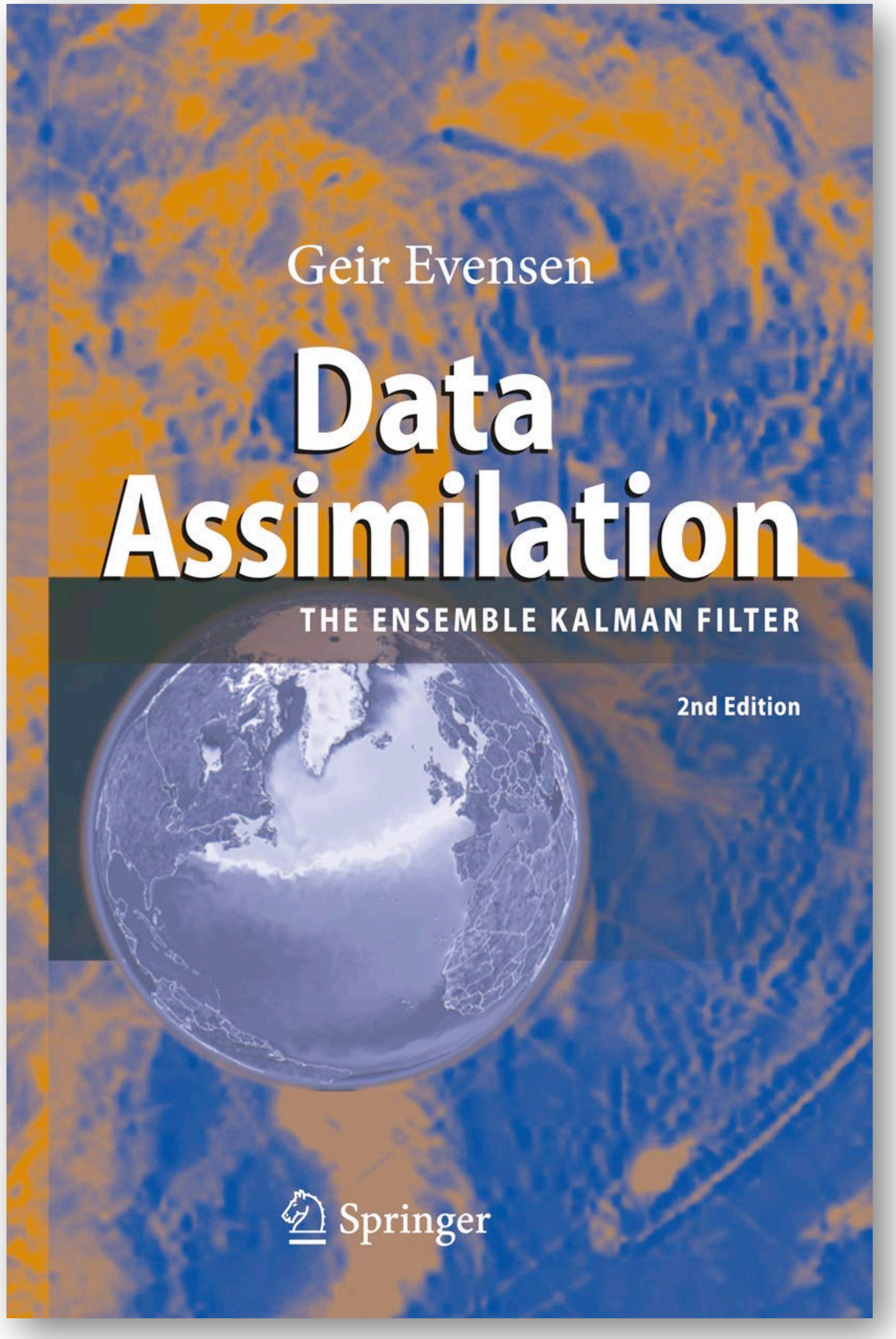
Raanes, Patrick N.. "Introduction to Data Assimilation and the Ensemble Kalman Filter." (2016).



A Primer on Data Assimilation



Raanes, Patrick N.. "Introduction to Data Assimilation and the Ensemble Kalman Filter." (2016).

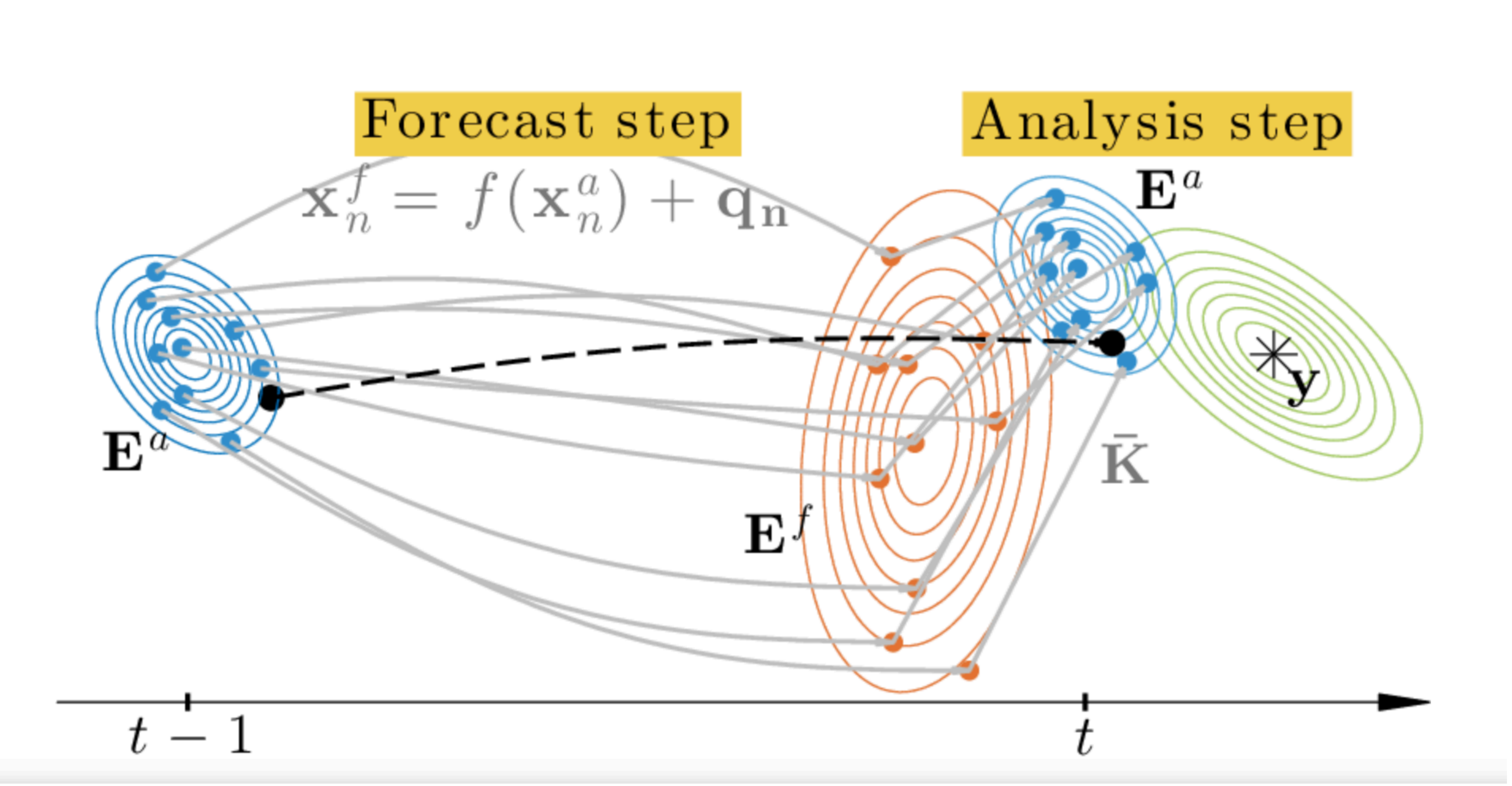


Assumption:
A linear-Gaussian system

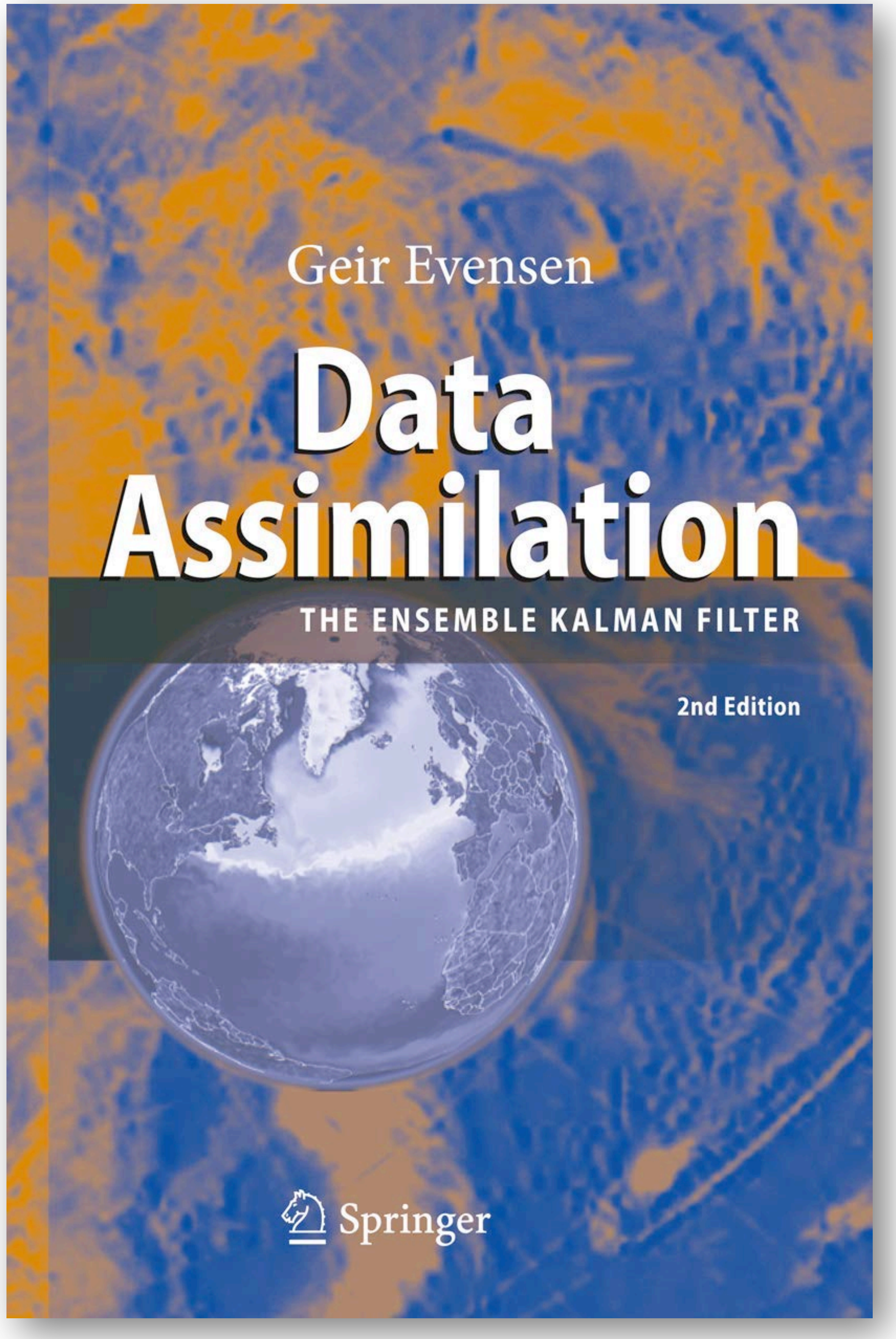
$$E^a = E^f + \mathbf{K}(y - HE^f)$$

$$\mathbf{K} = \mathrm{Cov}(E^f, HE^f)[\mathrm{Cov}(HE^f) + R]^{-1}$$

A Primer on Data Assimilation



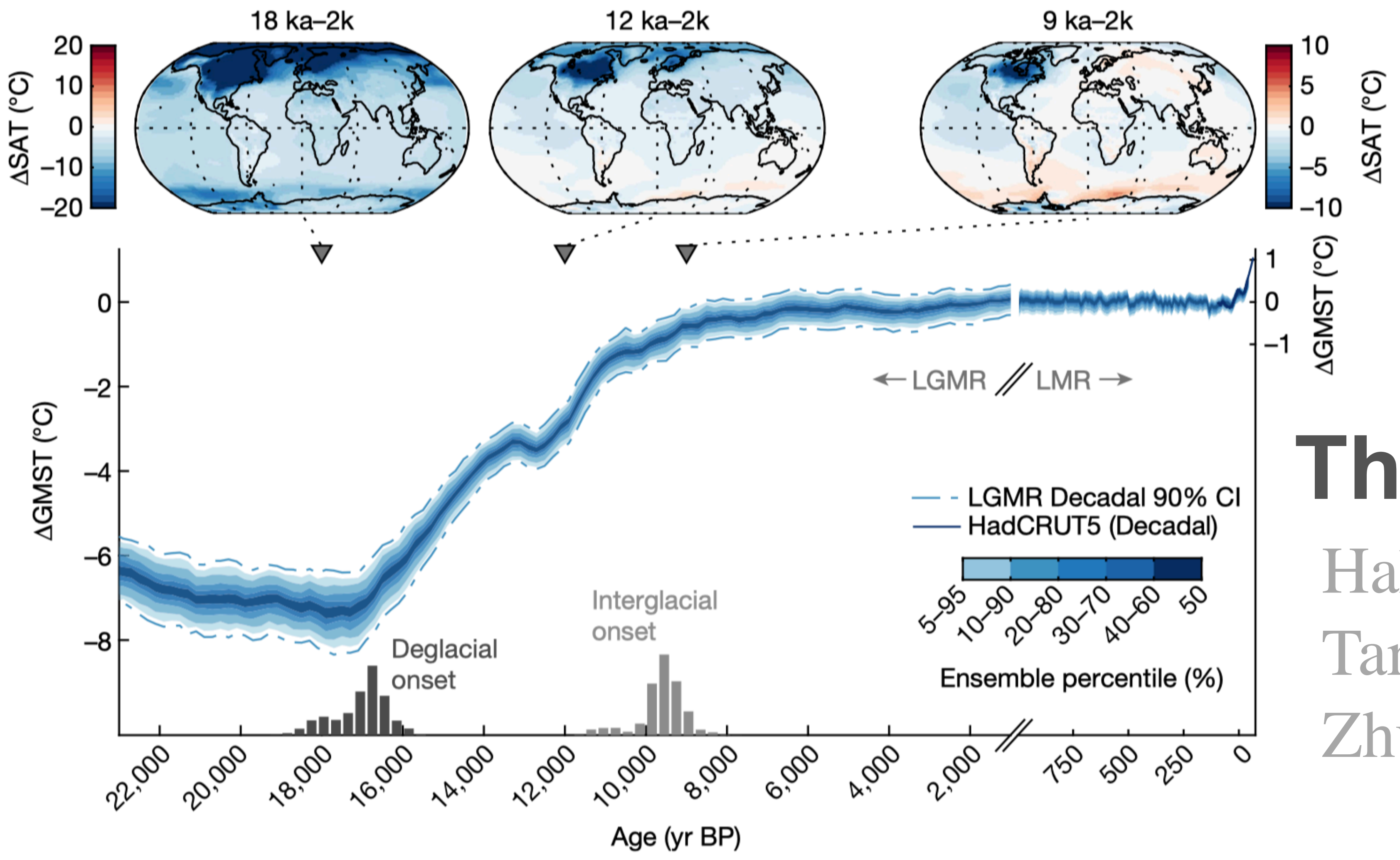
Raanes, Patrick N.. "Introduction to Data Assimilation and the Ensemble Kalman Filter." (2016).



State-of-the-art: offline paleoDA

The Last Glacial Maximum

Osman et al. (2021, *Nature*)



The Last Millennium

Hakim et al. (2016, *JGR-A*)

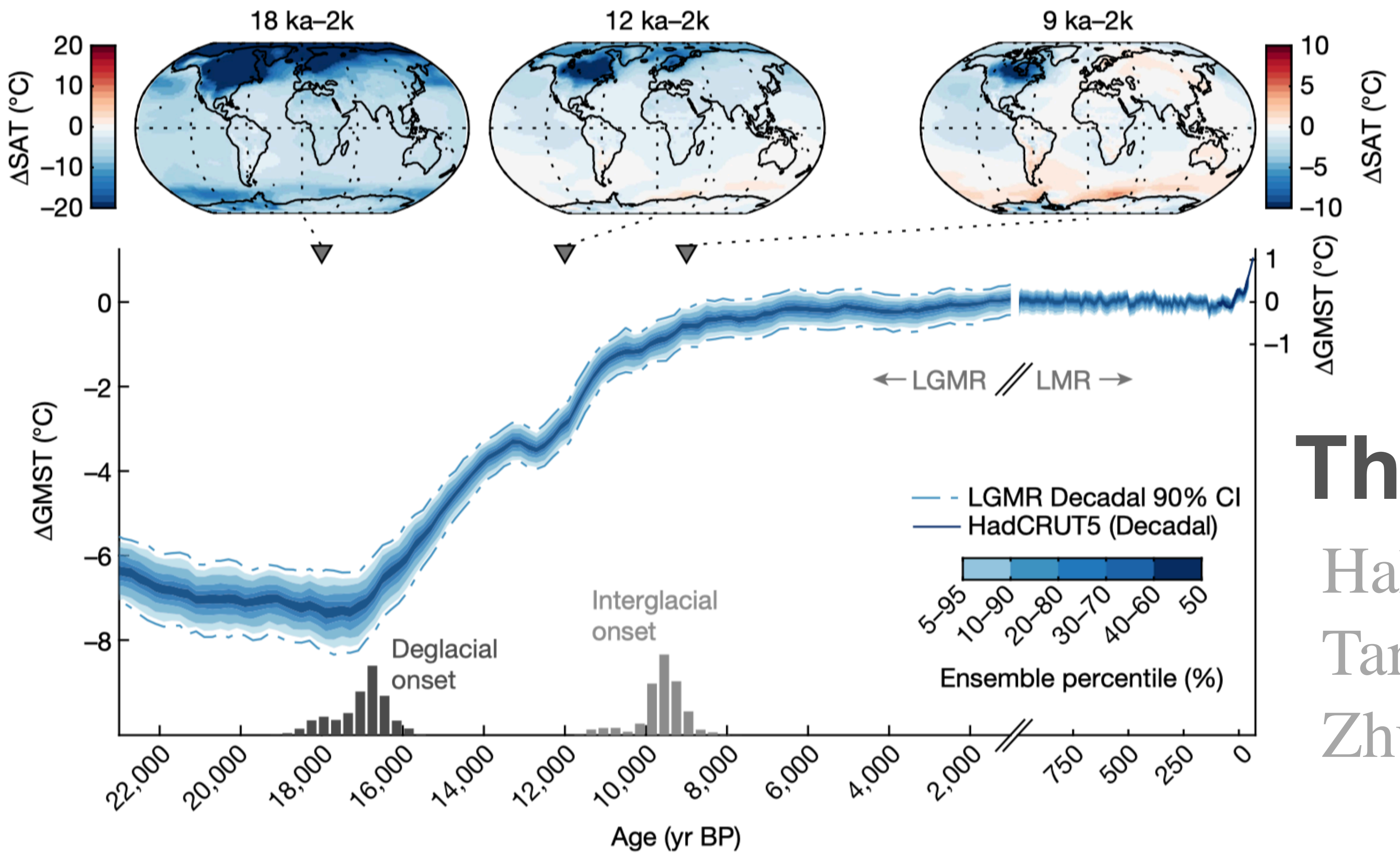
Tardif et al. (2019, *CP*)

Zhu et al. (2022, *Nat. Comms.*)

State-of-the-art: offline paleoDA

The Last Glacial Maximum

Osman et al. (2021, *Nature*)



The Last Millennium

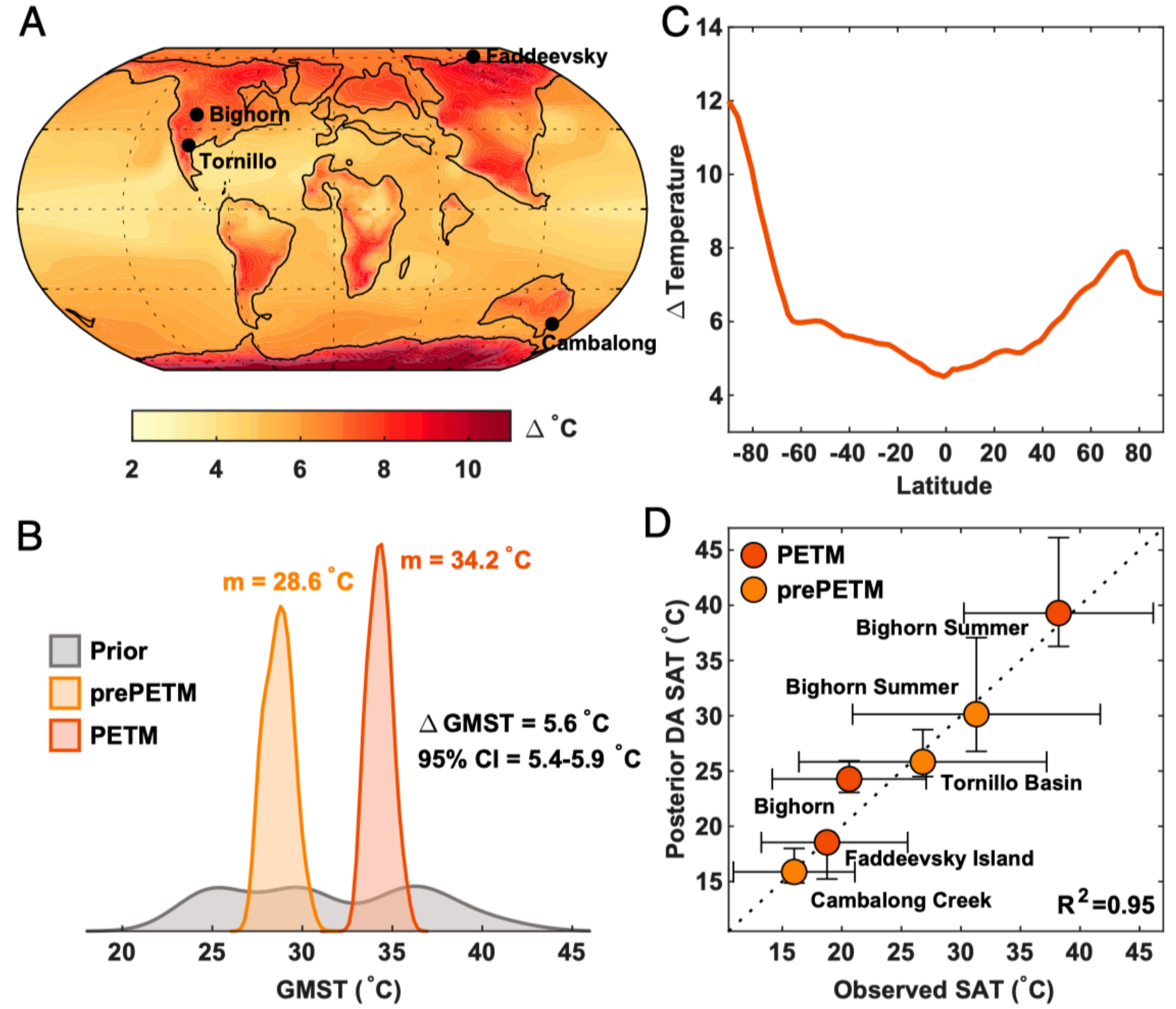
Hakim et al. (2016, *JGR-A*)

Tardif et al. (2019, *CP*)

Zhu et al. (2022, *Nat. Comms.*)

Tierney et al. (2022, *PNAS*)

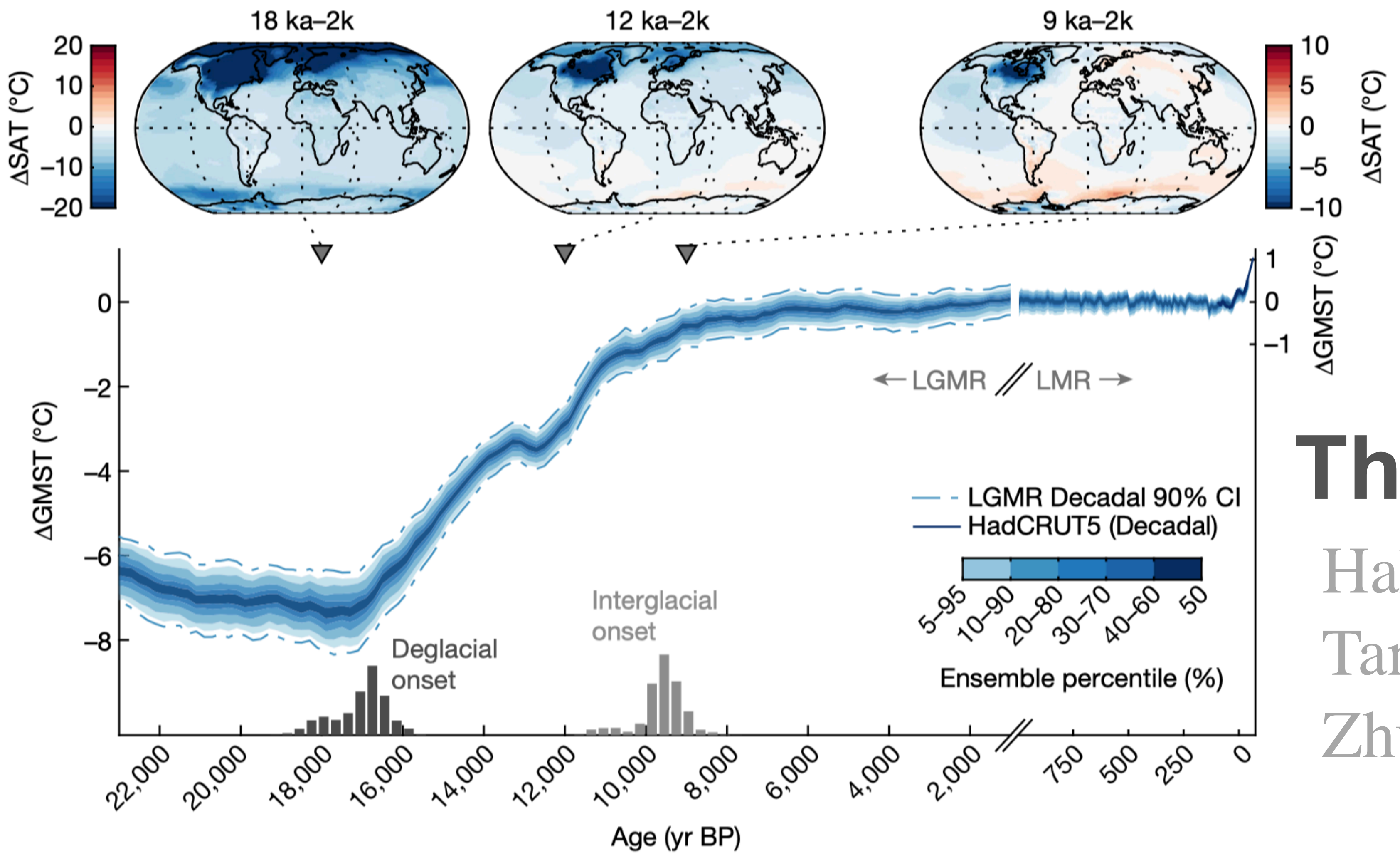
The Paleocene-Eocene Thermal Maximum



State-of-the-art: offline paleoDA

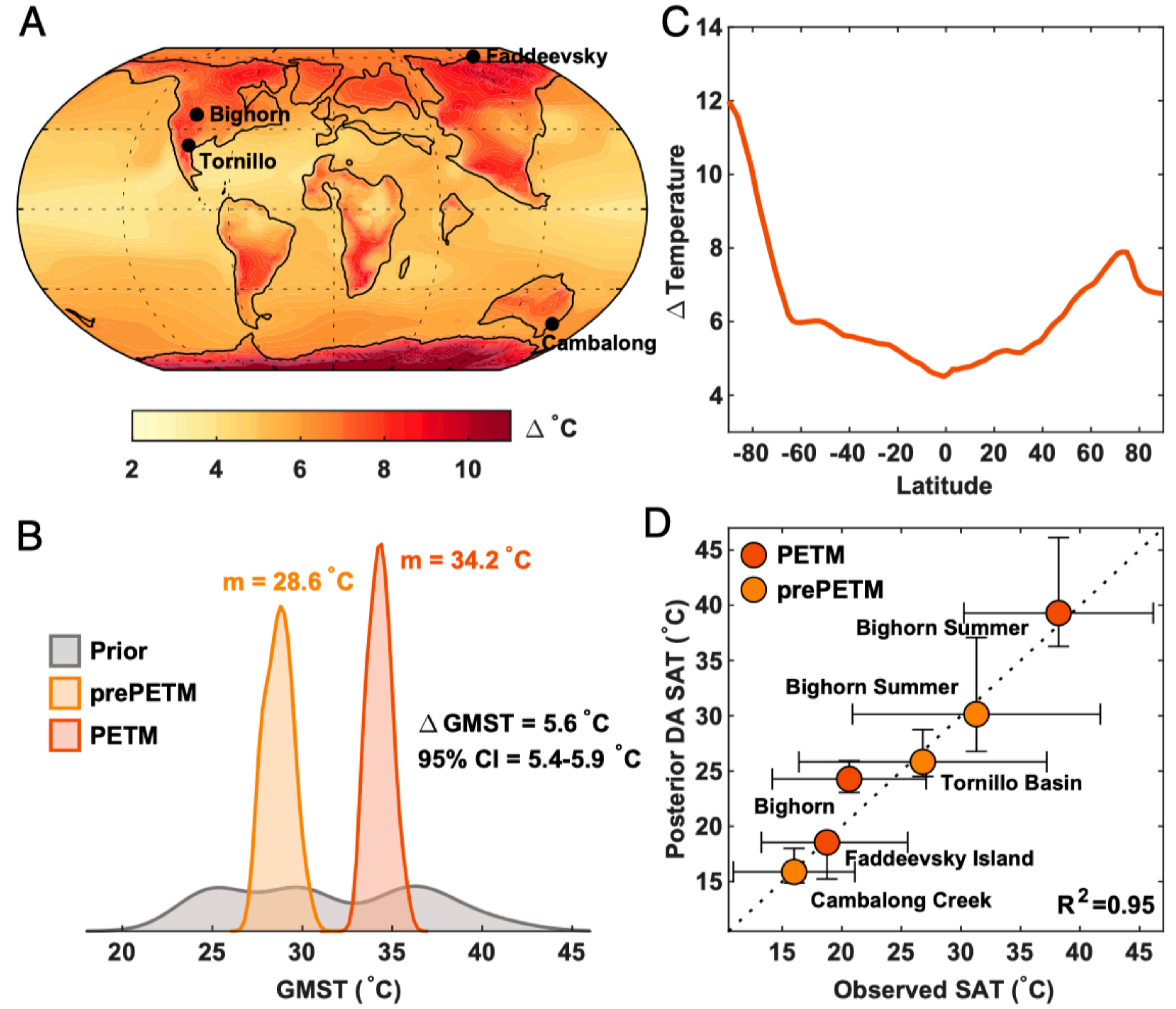
The Last Glacial Maximum

Osman et al. (2021, *Nature*)



The Paleocene-Eocene Thermal Maximum

Tierney et al. (2022, *PNAS*)



The Last Millennium

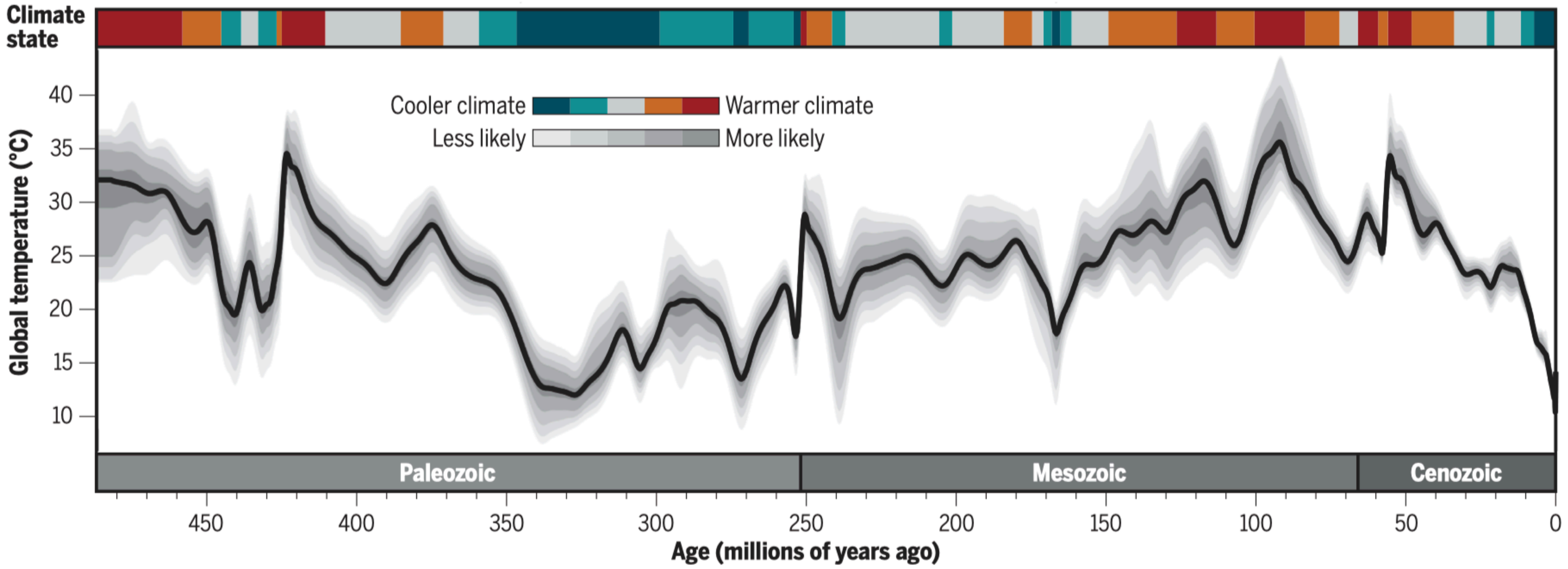
Hakim et al. (2016, *JGR-A*)

Tardif et al. (2019, *CP*)

Zhu et al. (2022, *Nat. Comms.*)

A 485-million-year history

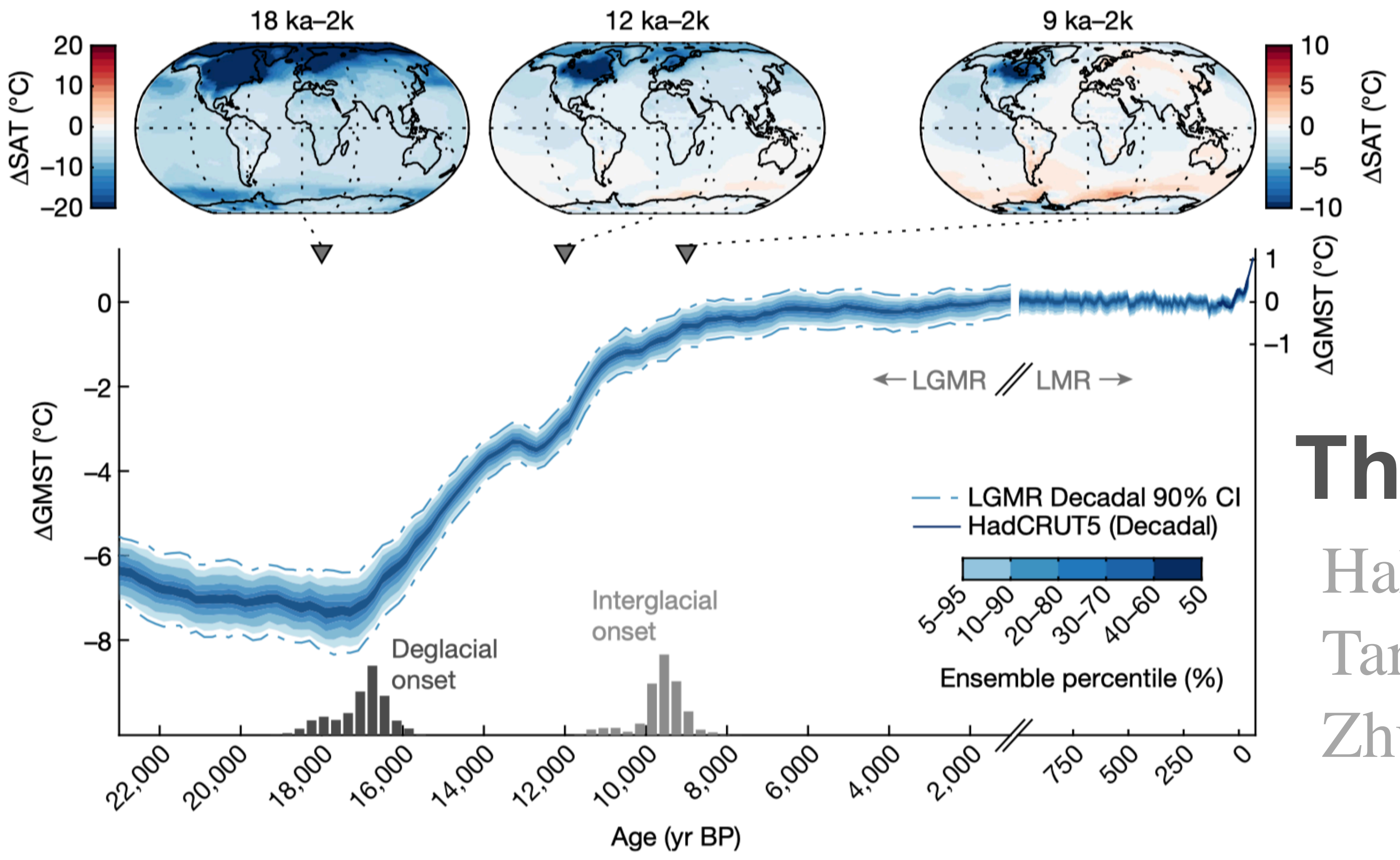
Judd et al. (2024, *Science*)



State-of-the-art: offline paleoDA

The Last Glacial Maximum

Osman et al. (2021, *Nature*)

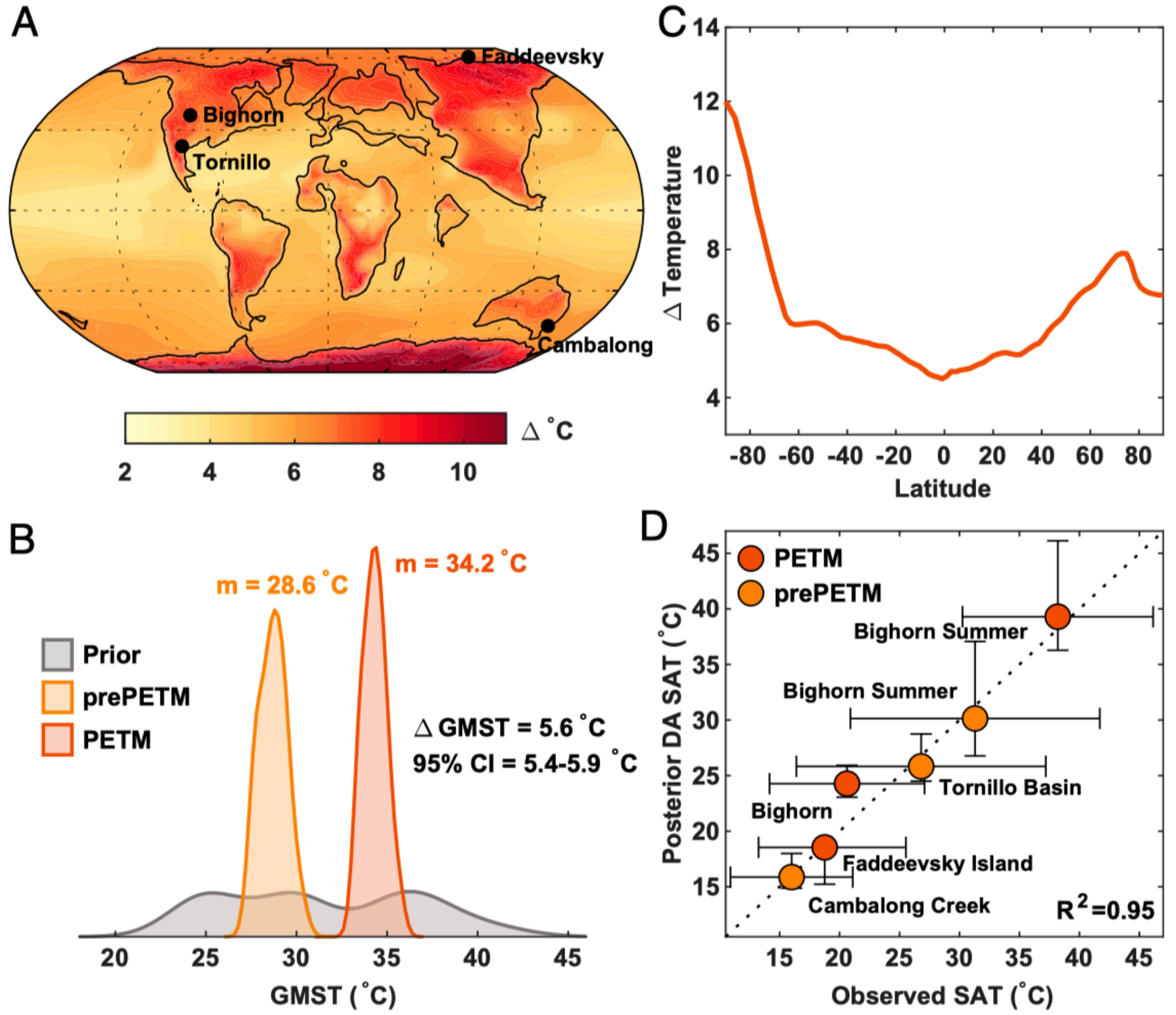


The Last Millennium

Hakim et al. (2016, *JGR-A*)
Tardif et al. (2019, *CP*)
Zhu et al. (2022, *Nat. Comms.*)

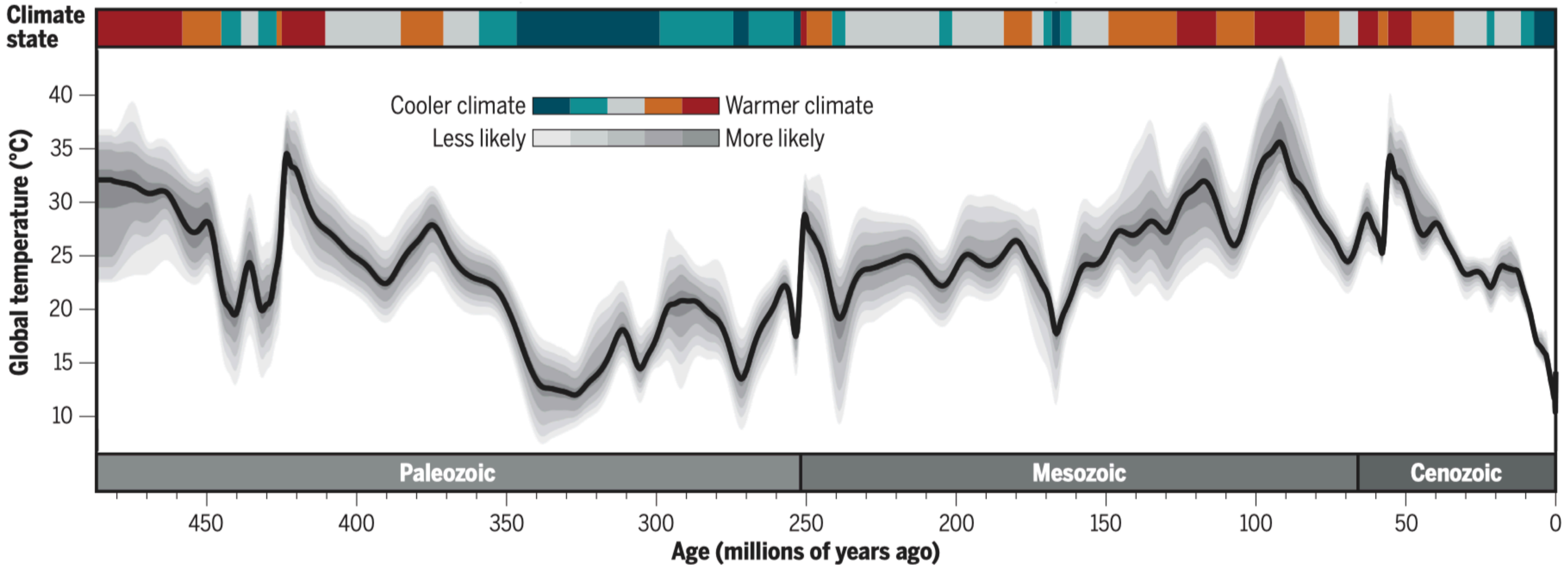
Tierney et al. (2022, *PNAS*)

The Paleocene-Eocene Thermal Maximum



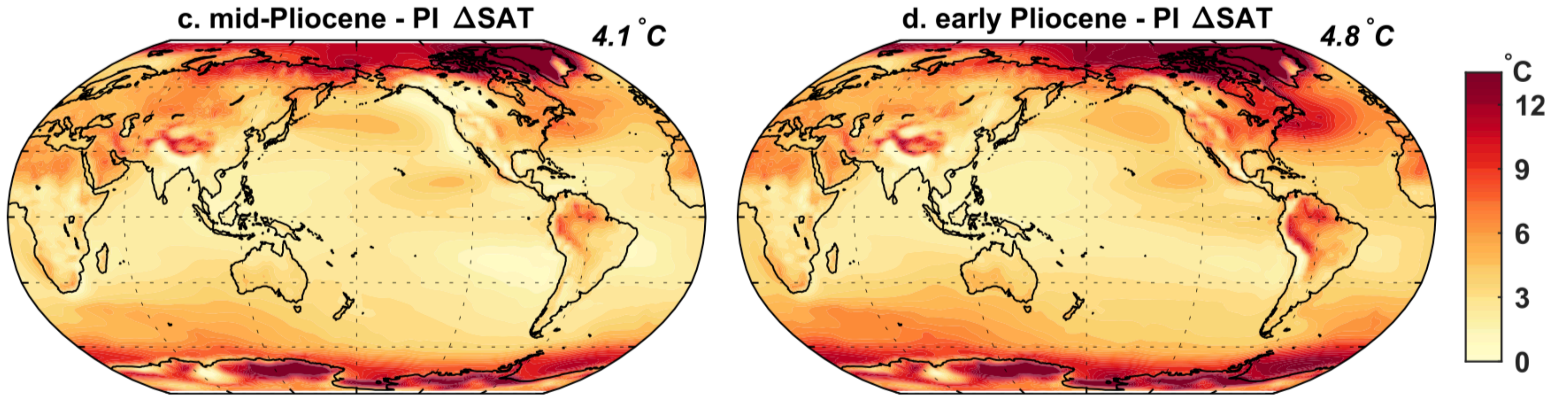
A 485-million-year history

Judd et al. (2024, *Science*)

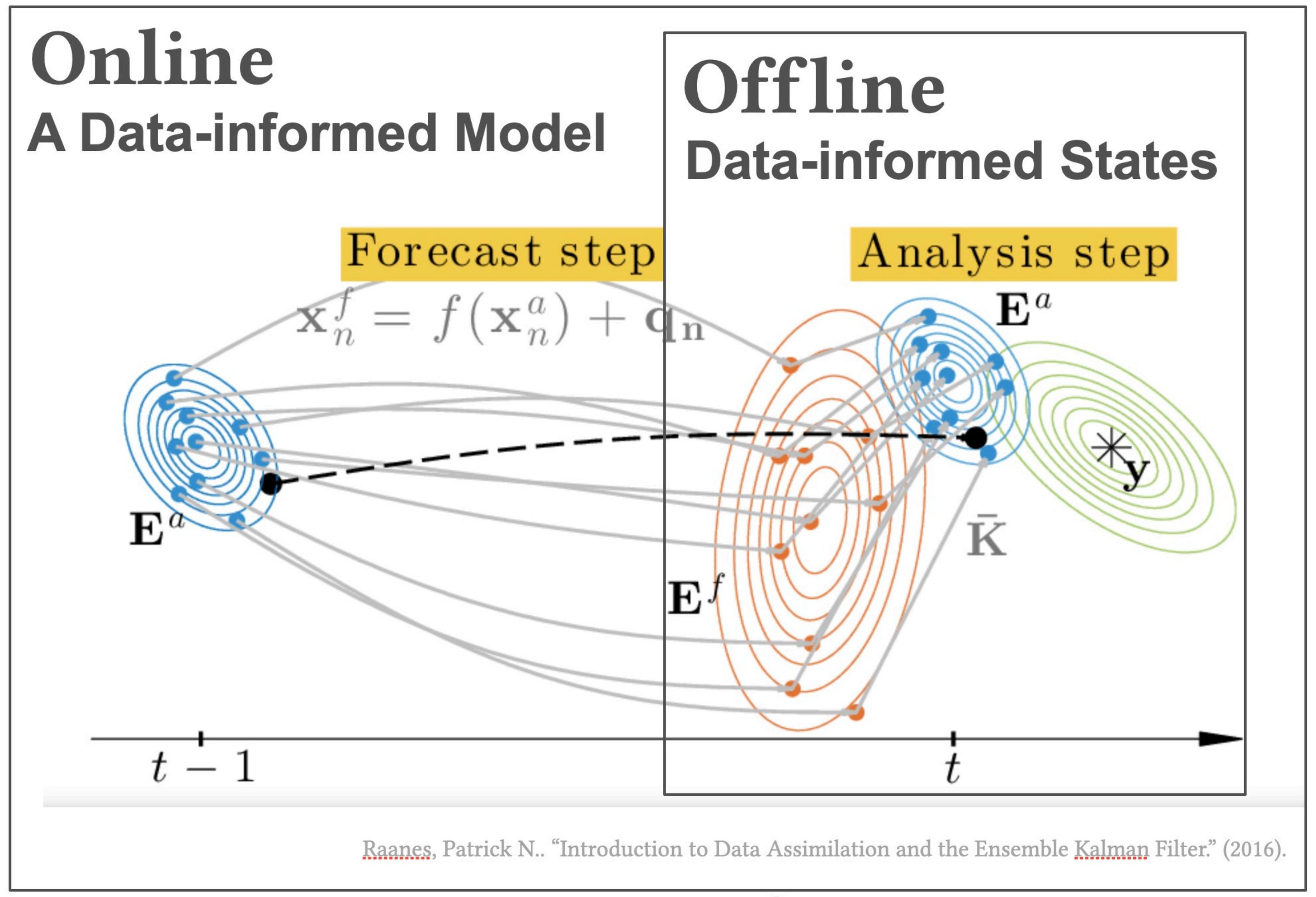


The Pliocene

Tierney et al. (2025, *AGU Adv.*)



Motivation for pursuing an online paleoDA solution



Assumption:
A linear-Gaussian system

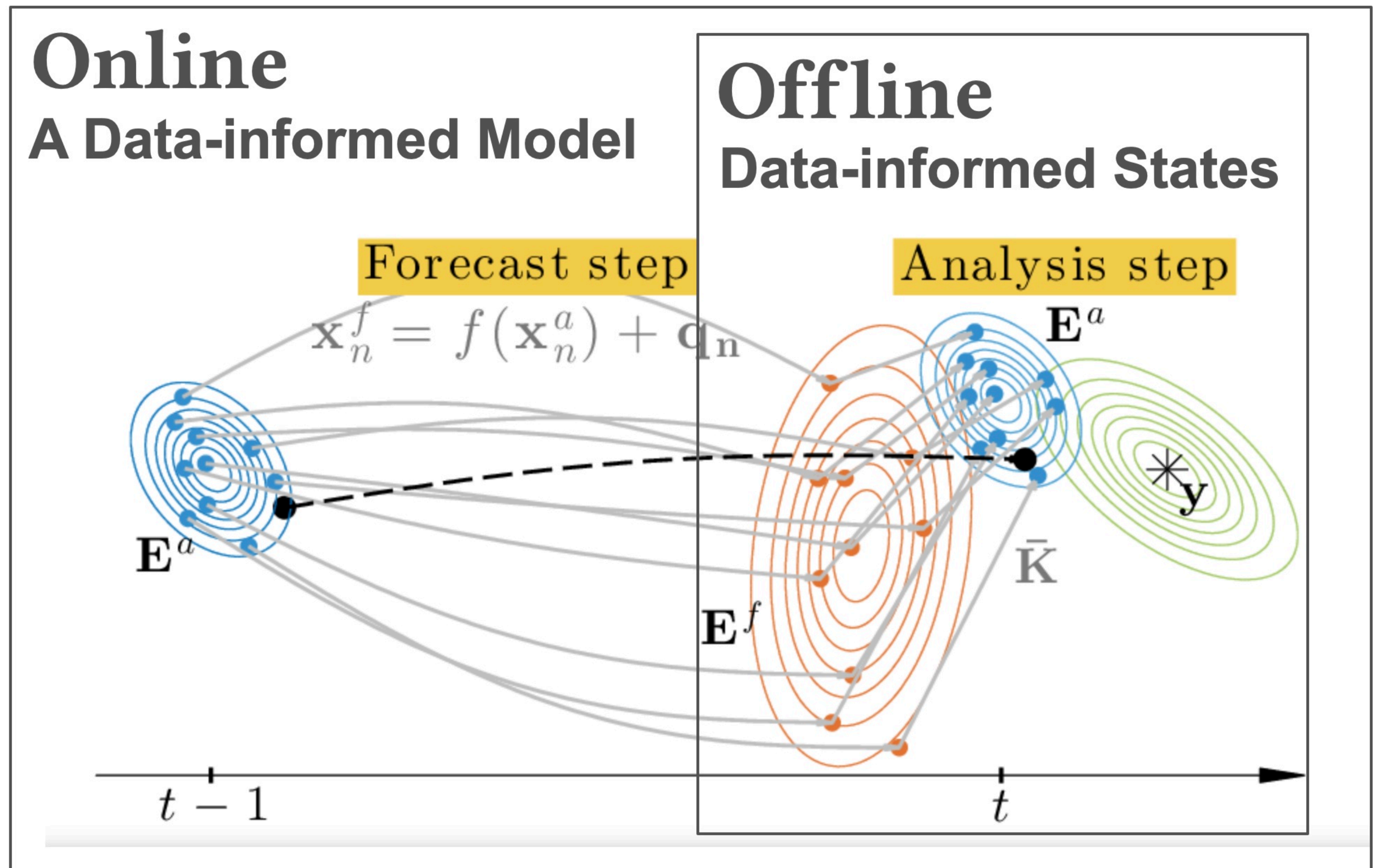
$$E^a = E^f + \mathbf{K}(y - HE^f)$$

$$\mathbf{K} = \mathrm{Cov}(E^f, HE^f)[\mathrm{Cov}(HE^f) + R]^{-1}$$

Online paleoDA: combining model and data using nonlinear-nonGaussian dynamics.

→ **more physical** multivariate & spatial relationships

Motivation for pursuing an online paleoDA solution



Raanes, Patrick N.. "Introduction to Data Assimilation and the Ensemble Kalman Filter." (2016).

Assumption:
A linear-Gaussian system

$$E^a = E^f + \mathbf{K}(y - HE^f)$$

$$\mathbf{K} = \mathrm{Cov}(E^f, HE^f)[\mathrm{Cov}(HE^f) + R]^{-1}$$

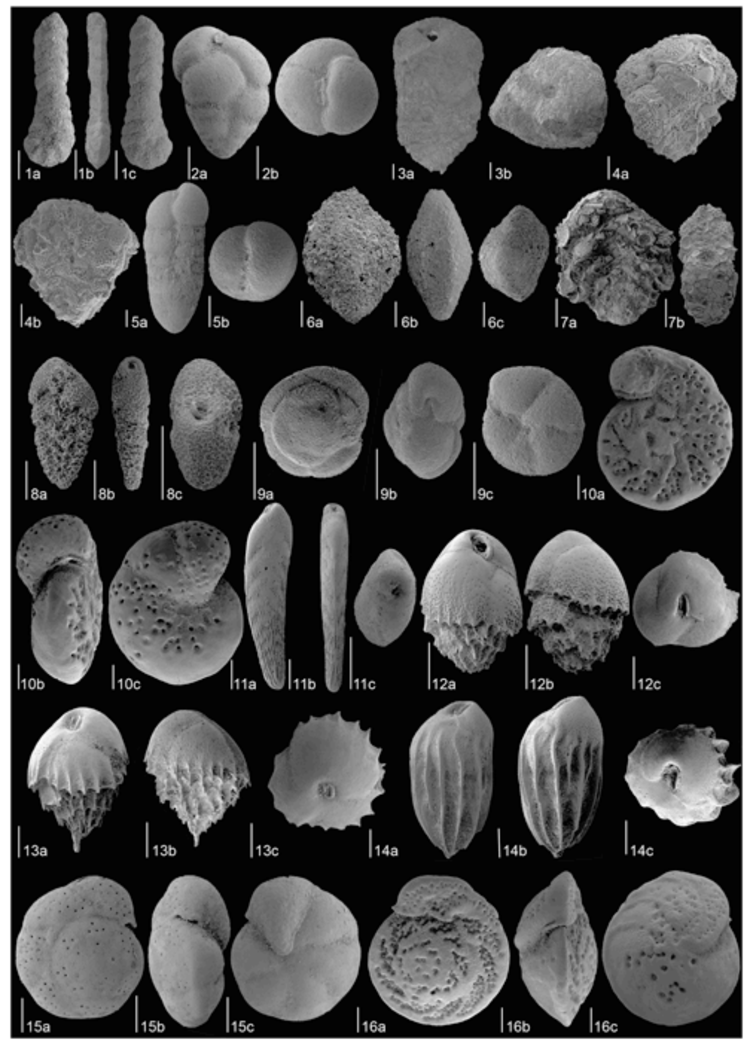
Online paleoDA: combining model and data using nonlinear-nonGaussian dynamics.

→ **more physical** multivariate & spatial relationships

A bonus for deep-time climate: assessing **internal variability** under different climatological patterns.

For example, the ENSO variability.

(deep-time paleo records are proxies for long-term mean climatology)

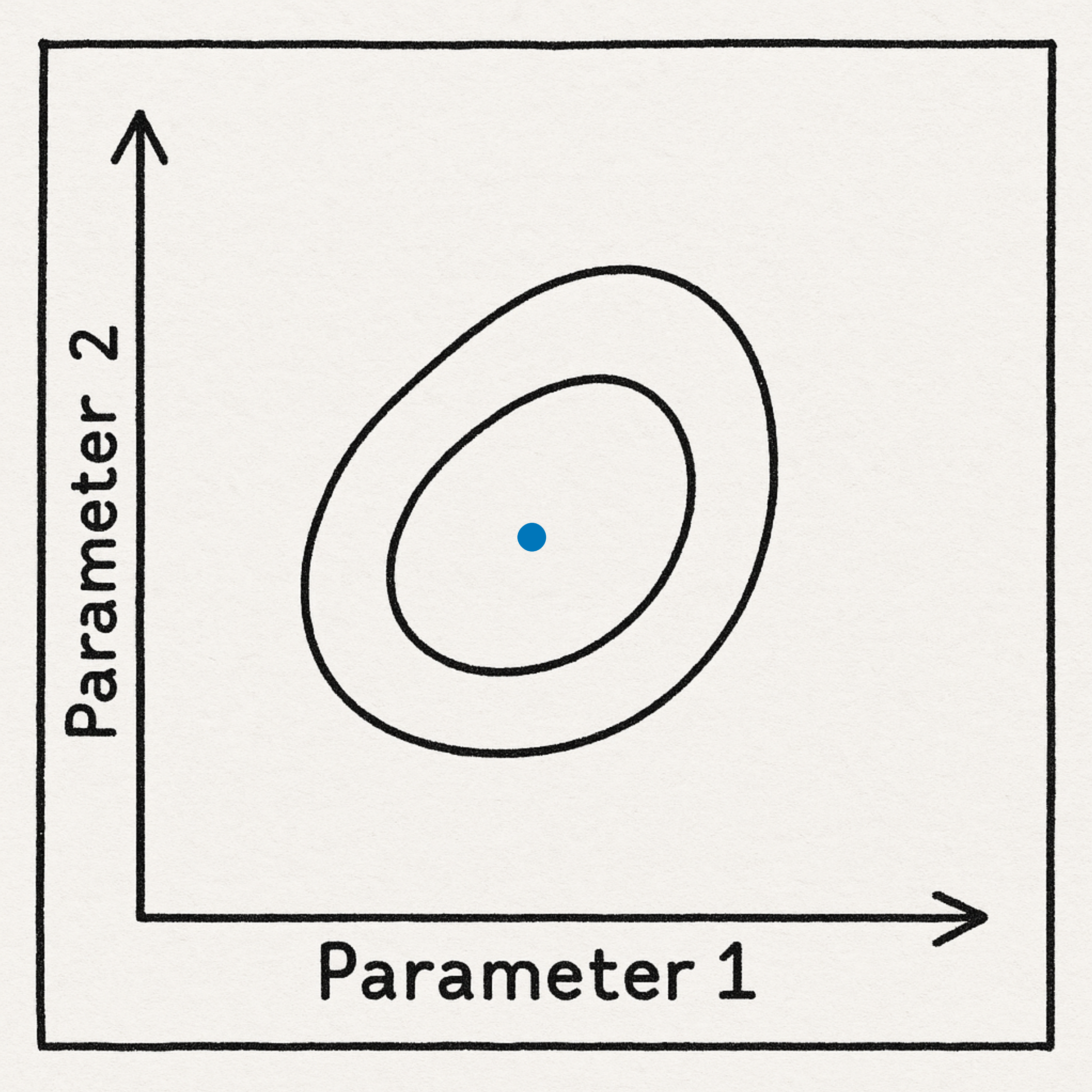


Gastaldello, M. E., Agnini, C., and Alegret, L (2024)

Deep-time climate: a boundary condition problem

An ideal approach:
Parameter Estimation

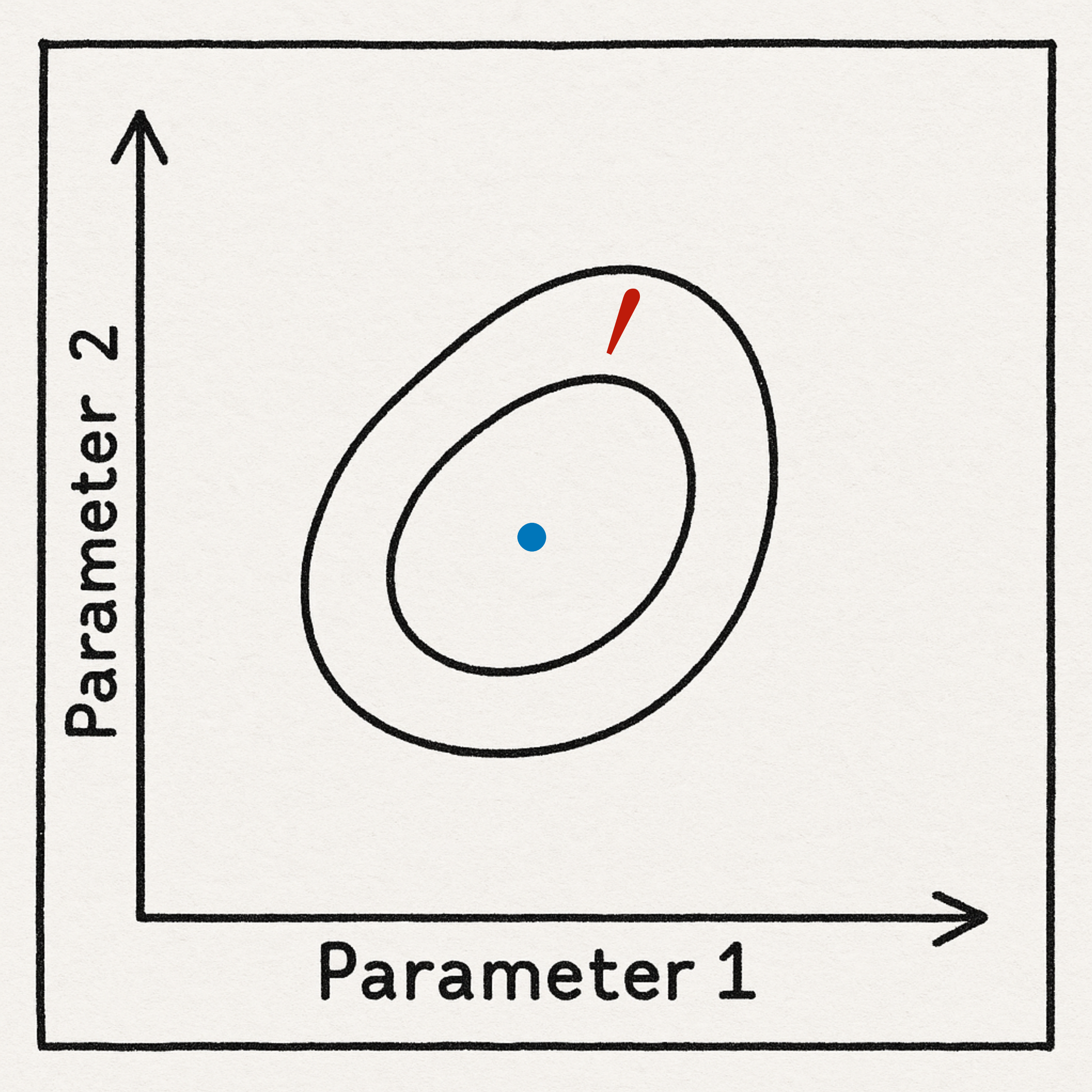
$$M(\theta) = X$$



Deep-time climate: a boundary condition problem

An ideal approach:
Parameter Estimation

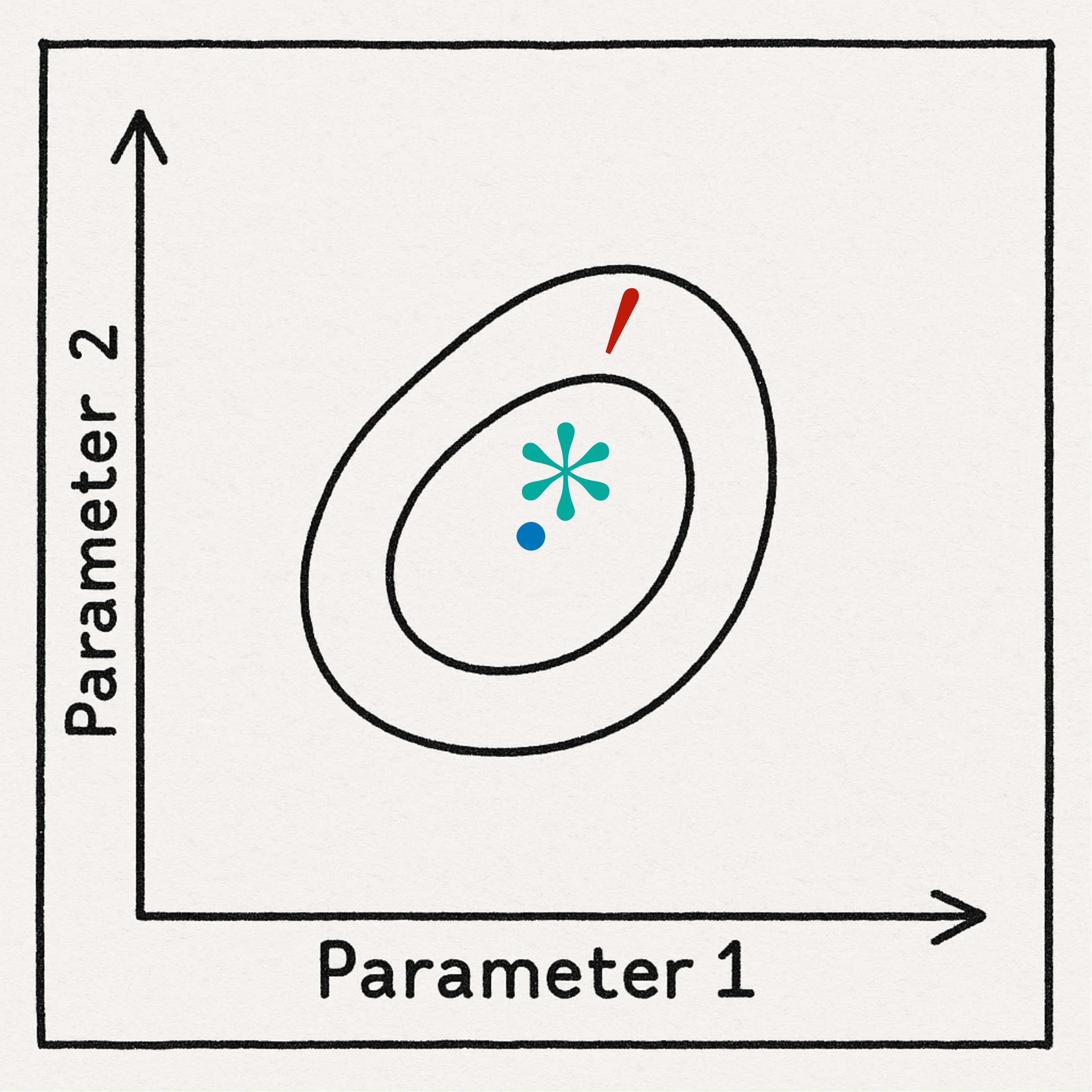
$$M(\theta) = X$$
$$M(\theta') = X'$$



Deep-time climate: a boundary condition problem

An ideal approach:
Parameter Estimation

$$M(\theta) = X$$
$$M(\theta') = X'$$
$$M(\theta^*) = X^*$$



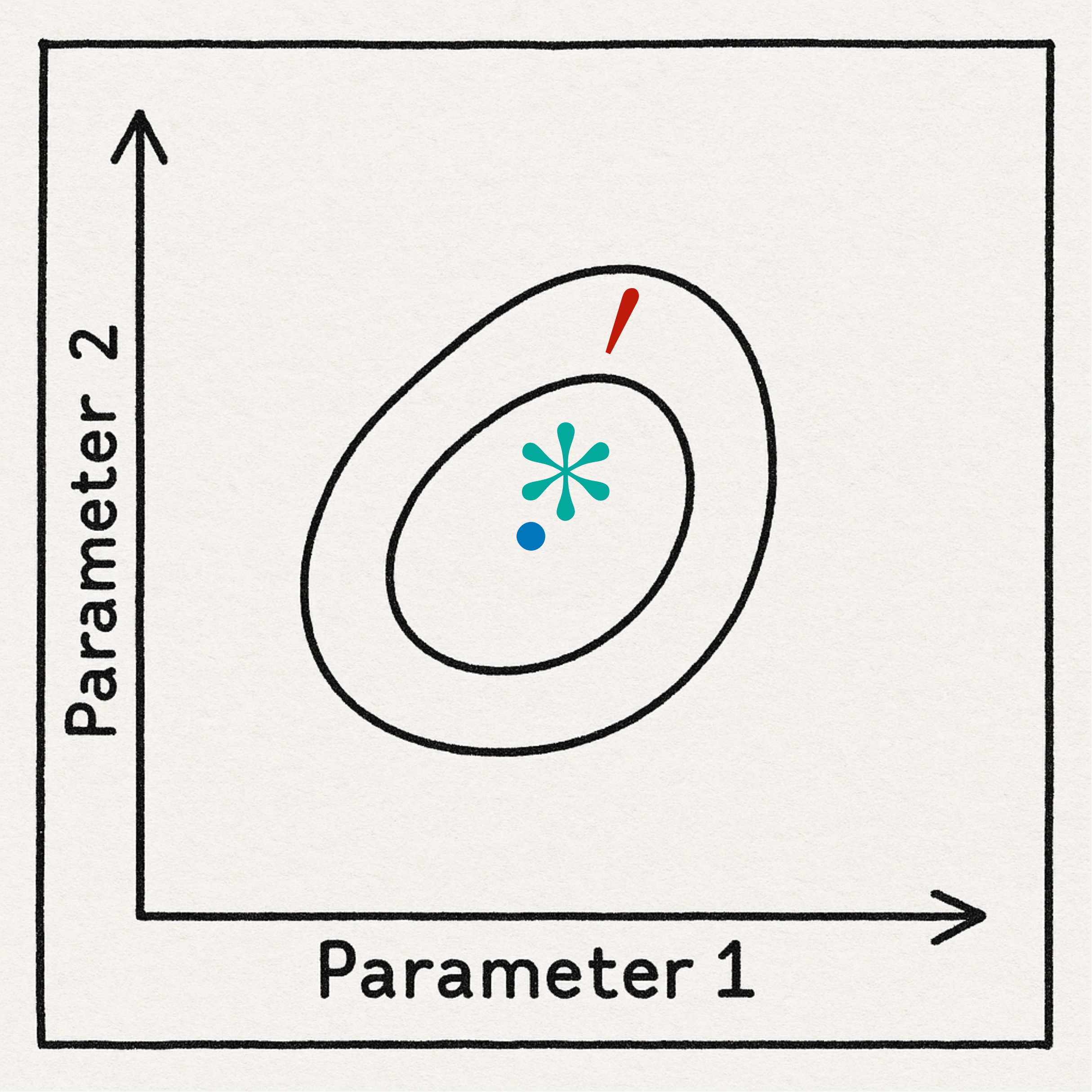
Deep-time climate: a boundary condition problem

An ideal approach: Parameter Estimation

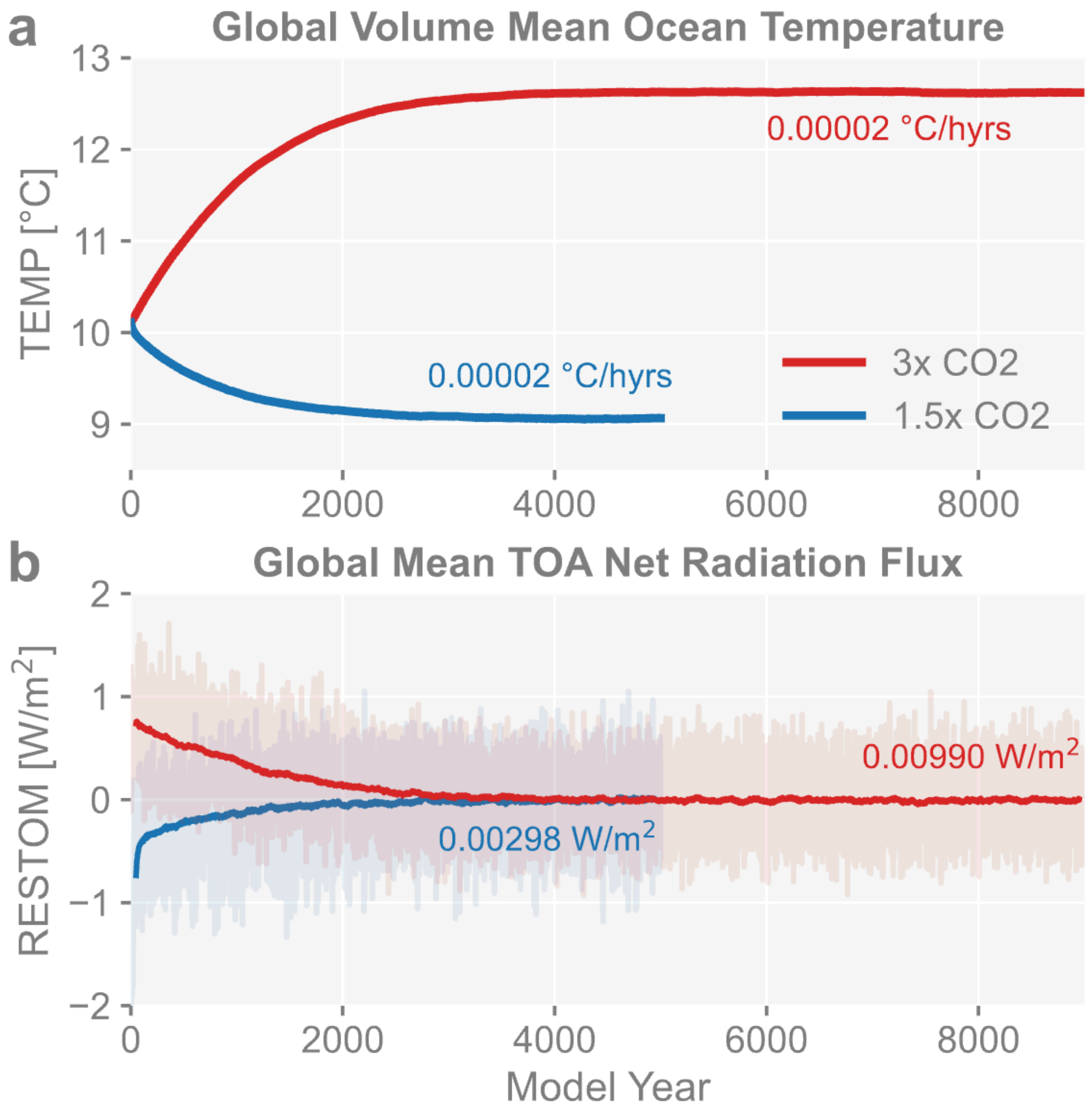
$$M(\theta) = X$$

$$M(\theta') = X'$$

$$M(\theta^*) = X^*$$



However, M is slow and expensive.



Zhu et al. (submitted to *Nat. Geosci.*)

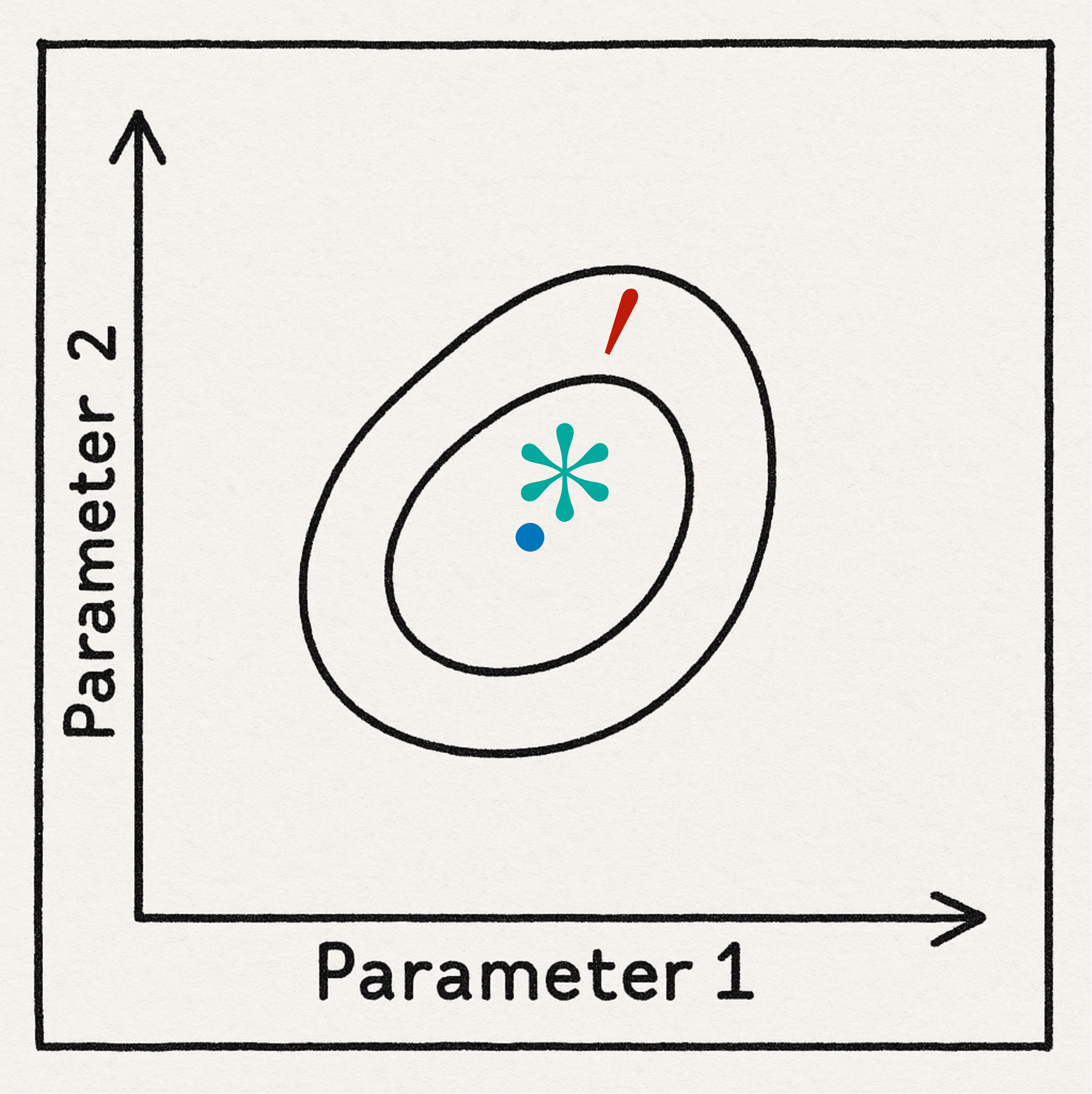
Deep-time climate: a boundary condition problem

An ideal approach: Parameter Estimation

$$M(\theta) = X$$

$$M(\theta') = X'$$

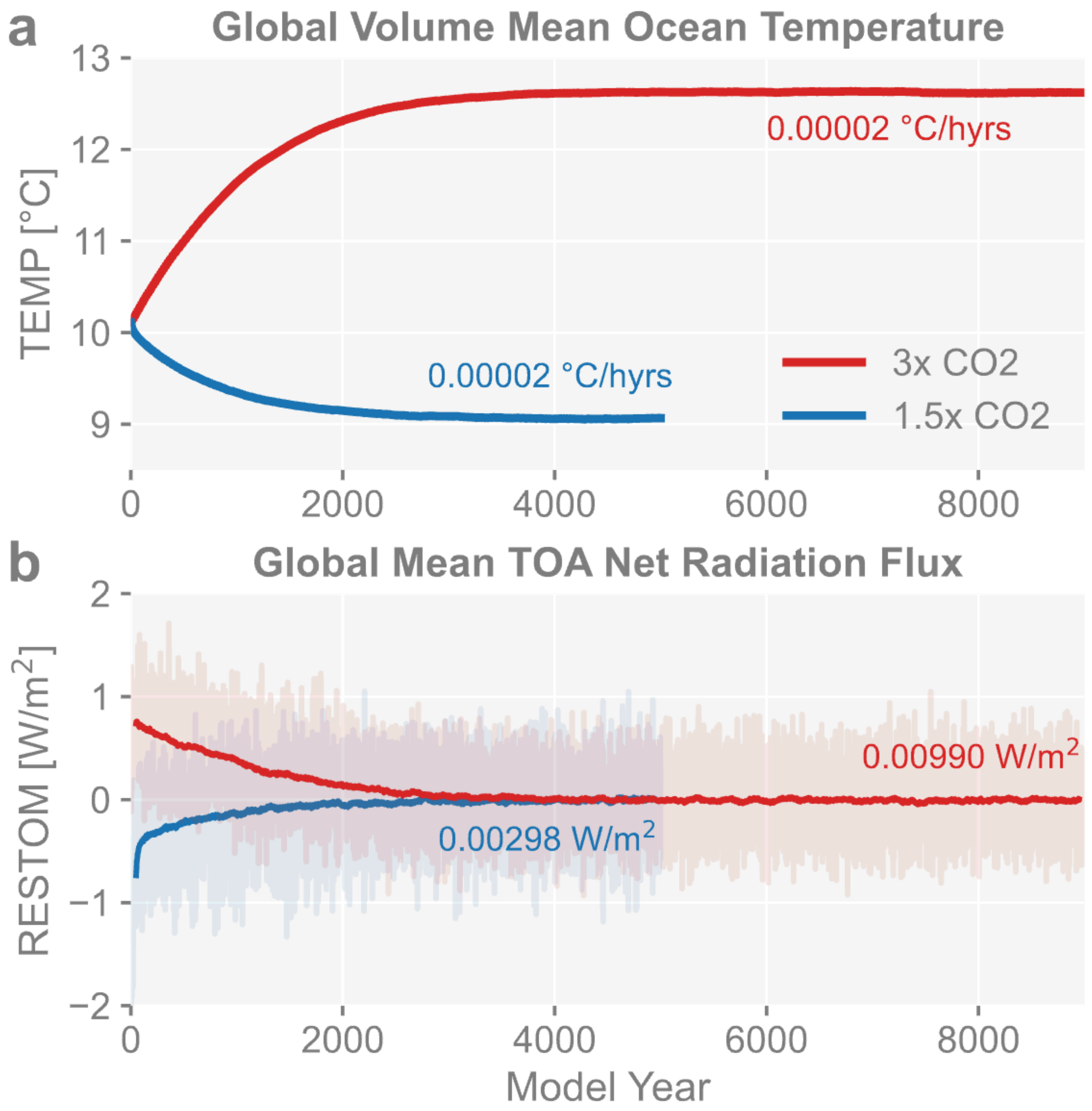
$$M(\theta^*) = X^*$$



A more feasible approach:

$$M(\theta', f_X) = X^*$$

However, M is slow and expensive.



Zhu et al. (submitted to *Nat. Geosci.*)

A new online paleoDA approach in CESM

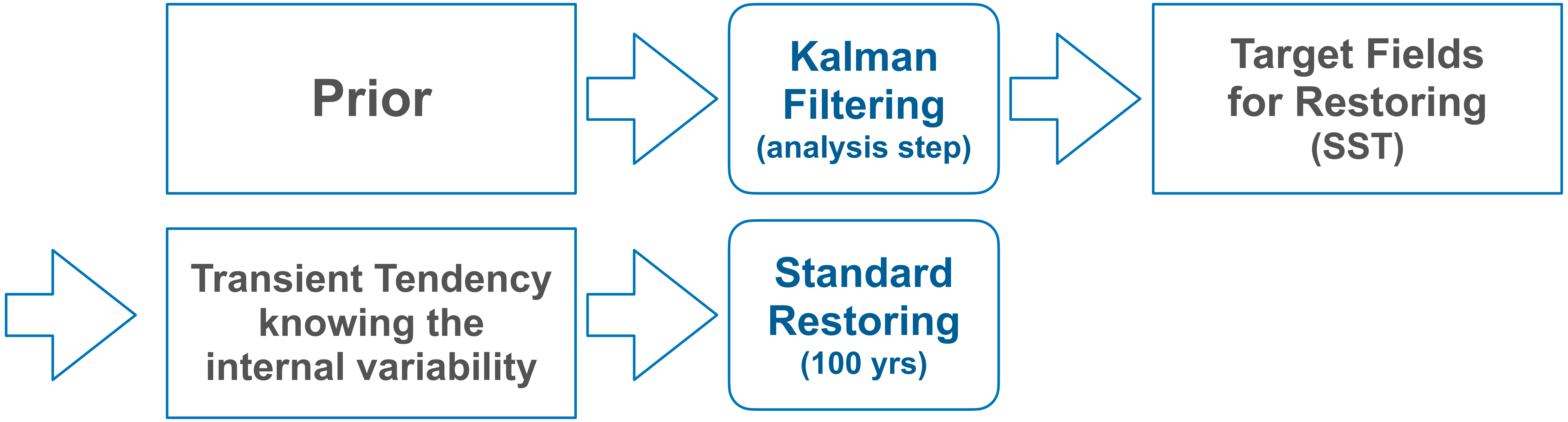
Goal:

- fit the SST proxy observations
- a running data-informed model

A new online paleoDA approach in CESM

Goal:

- fit the SST proxy observations
- a running data-informed model



μ

$$\frac{dX(t)}{dt} = -\frac{1}{\tau}(X(t) - \mu)$$

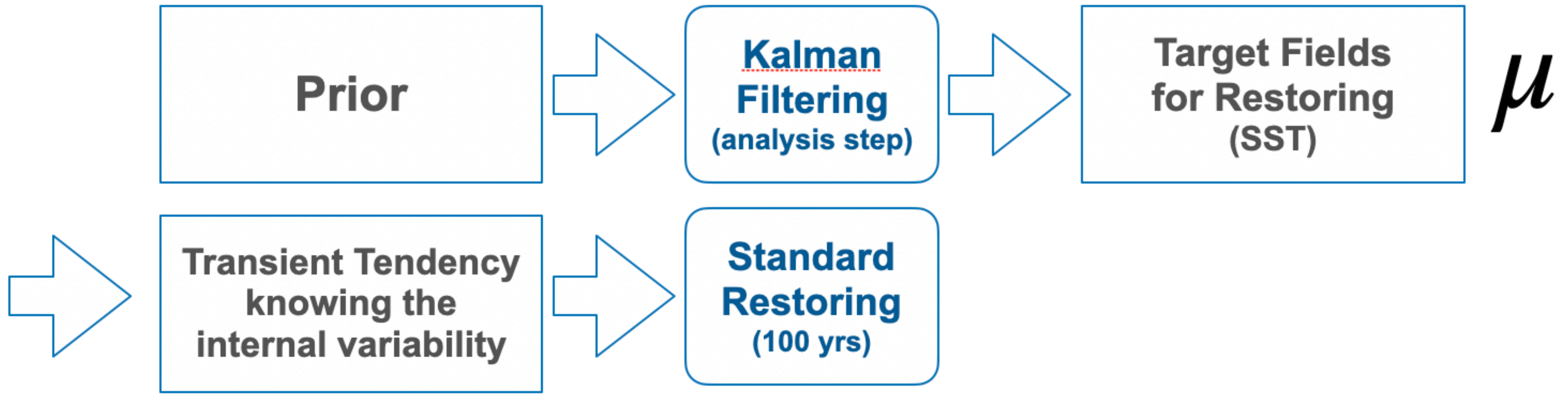
$$M(\theta', f_X) = X^*$$

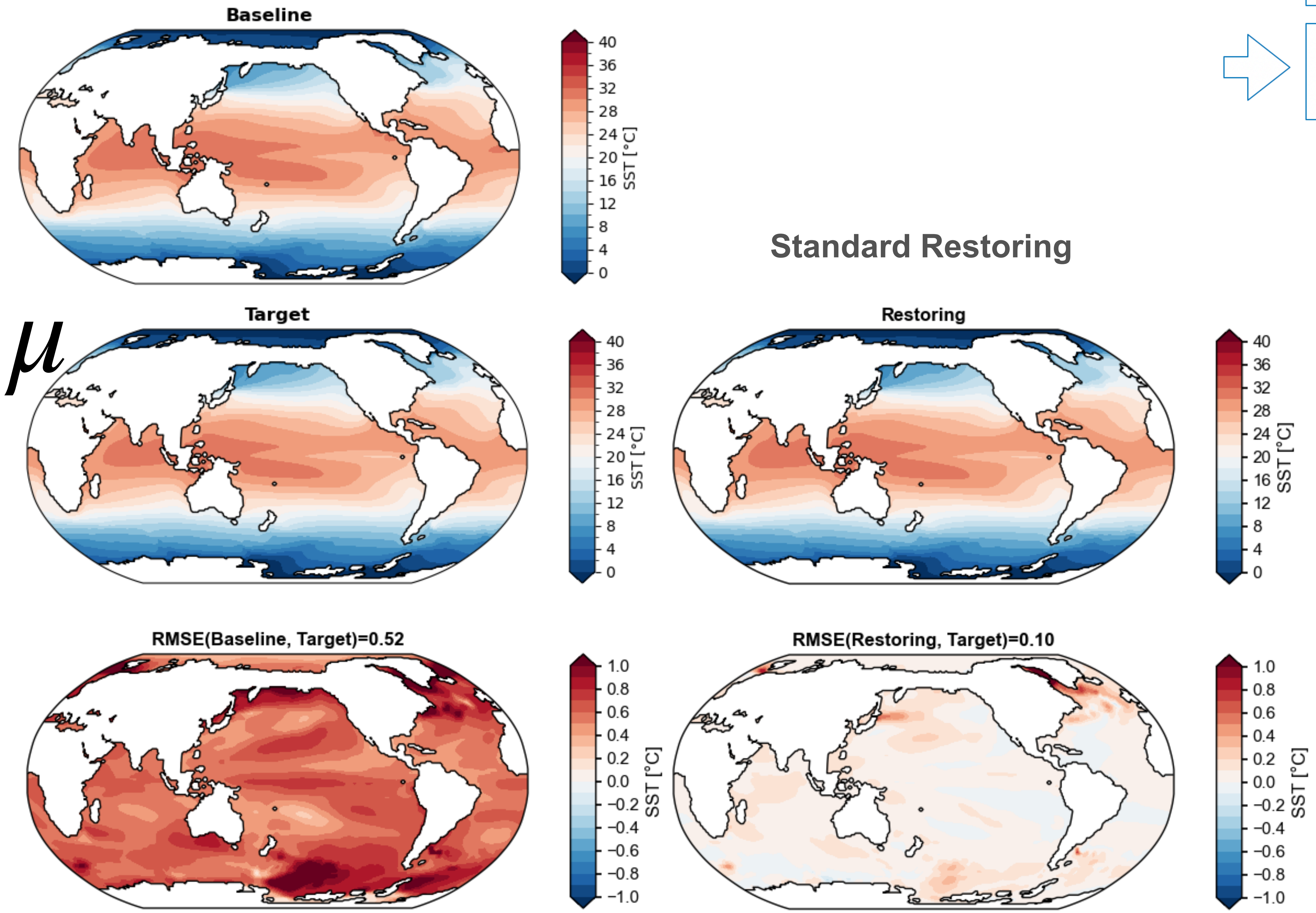
An ideal test on a Pliocene case

The Sea Surface Temperature Field

Baseline: 400ppm (Feng et al., 2020) **extended for another 800 yrs**

Target: 350ppm, branched from 400ppm (Feng et al., 2020) **and extended for another 400 yrs**



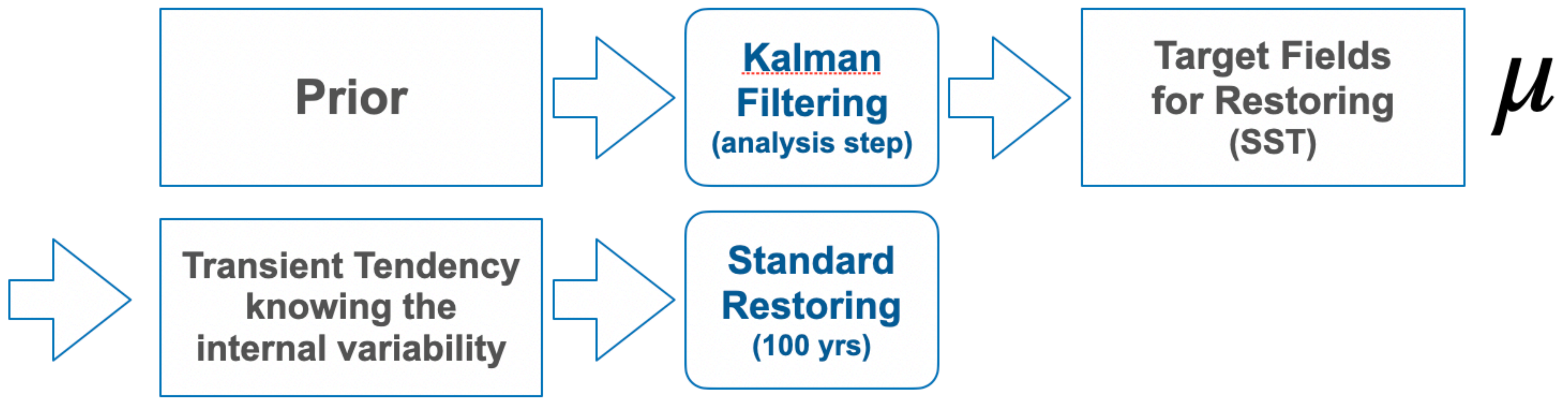


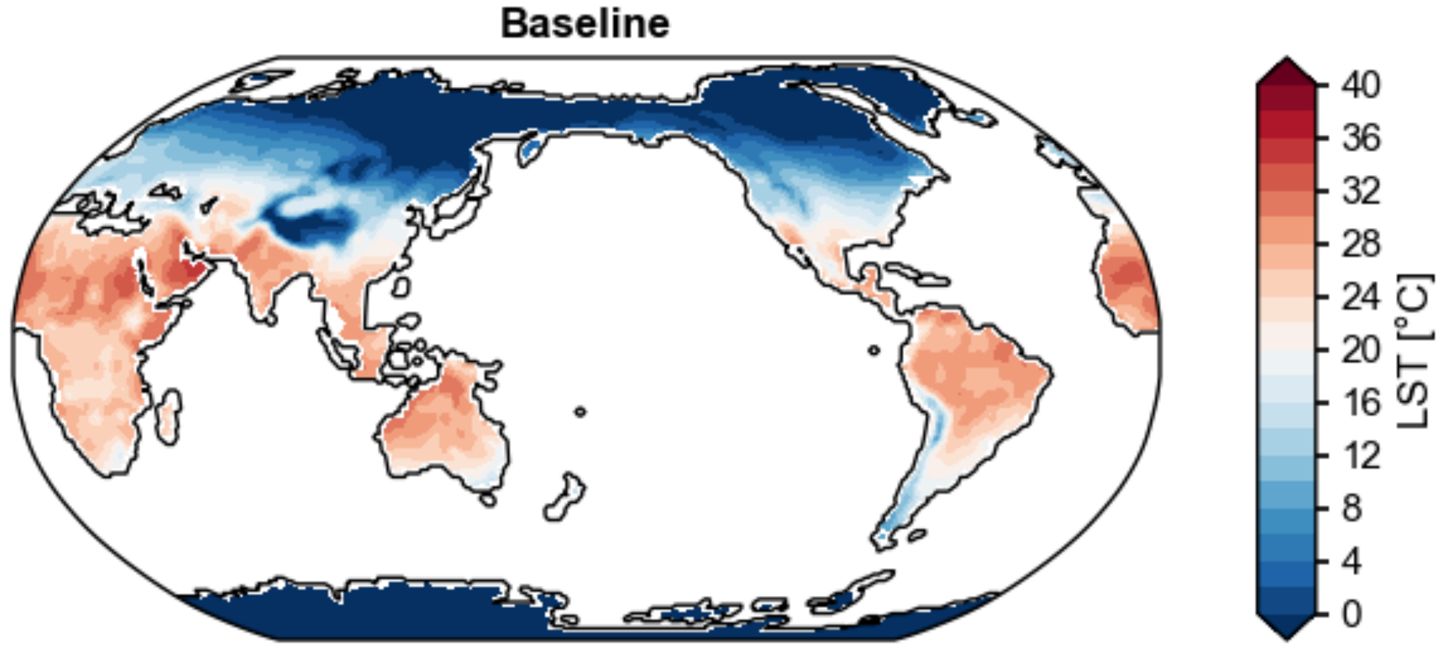
An ideal test on a Pliocene case

The Land Surface Temperature Field

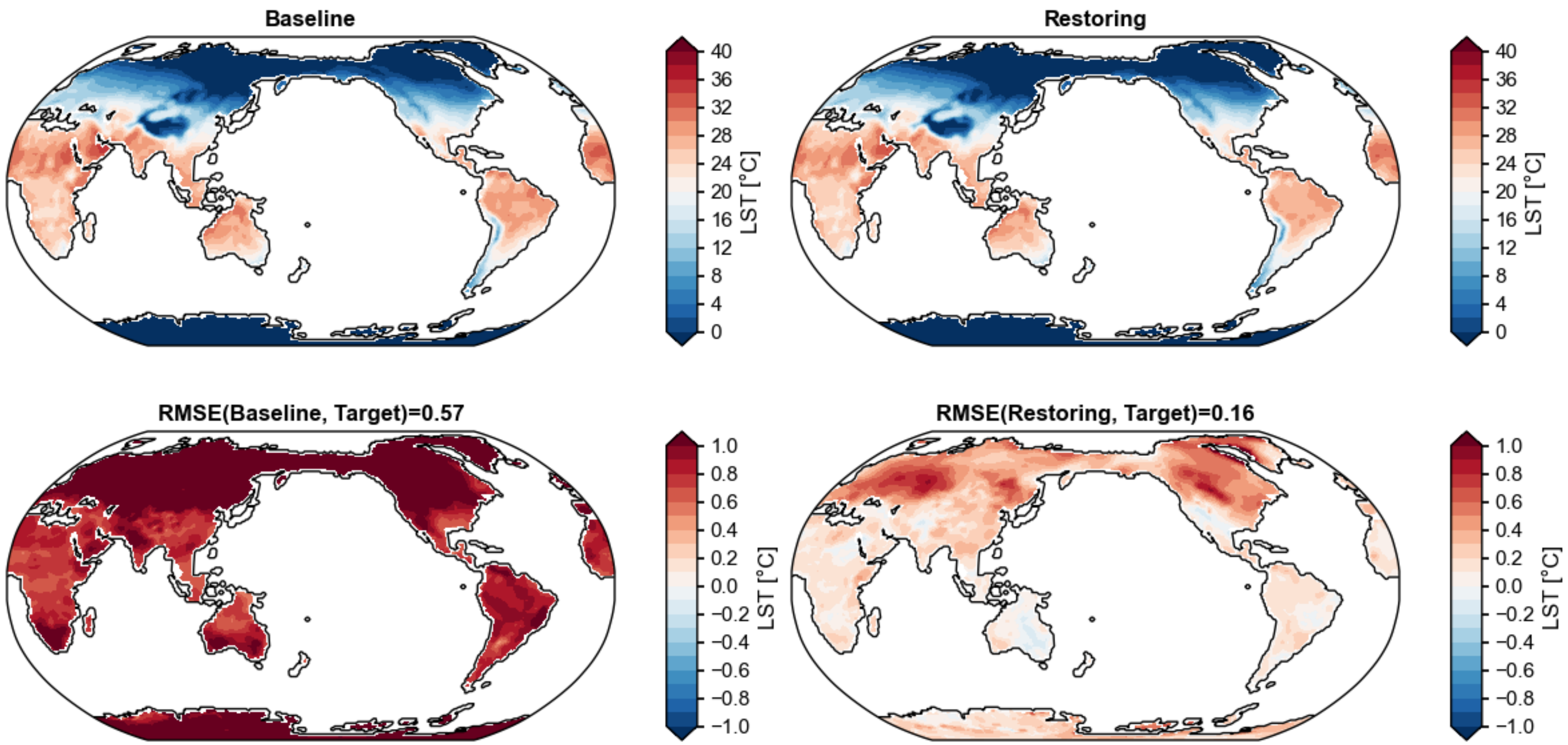
Baseline: 400ppm (Feng et al., 2020) extended for another 800 yrs

Target: 350ppm, branched from 400ppm (Feng et al., 2020) and extended for another 400 yrs





Standard Restoring

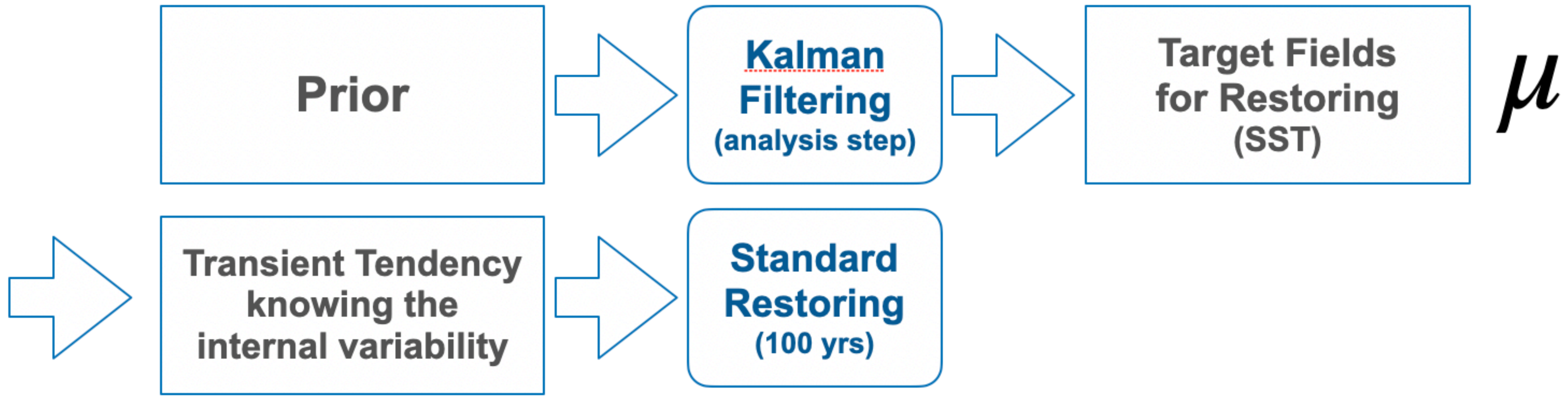


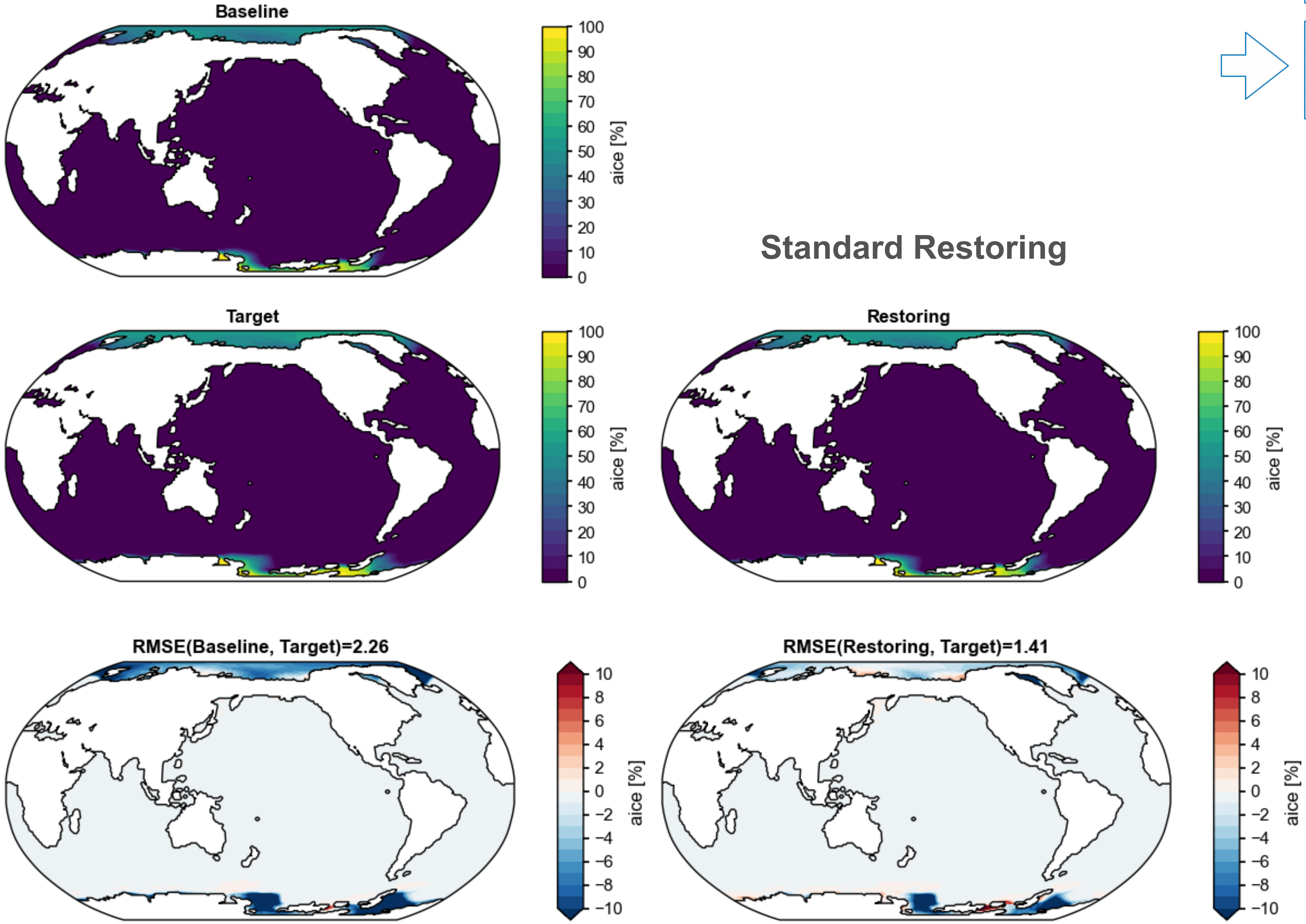
An ideal test on a Pliocene case

The Sea Ice Area Field

Baseline: 400ppm (Feng et al., 2020) **extended for another 800 yrs**

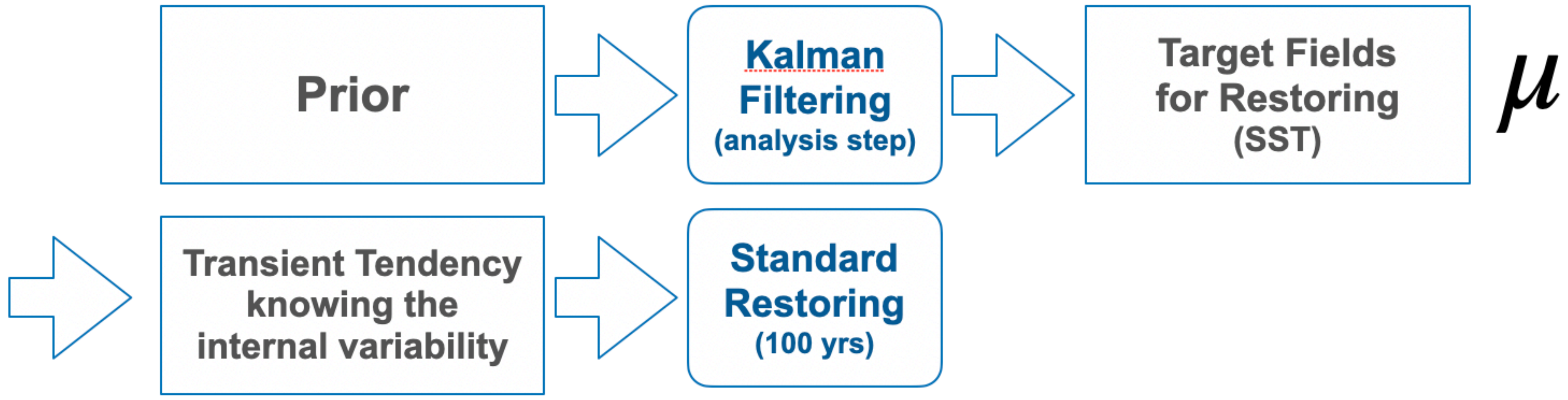
Target: 350ppm, branched from 400ppm (Feng et al., 2020) **and extended for another 400 yrs**





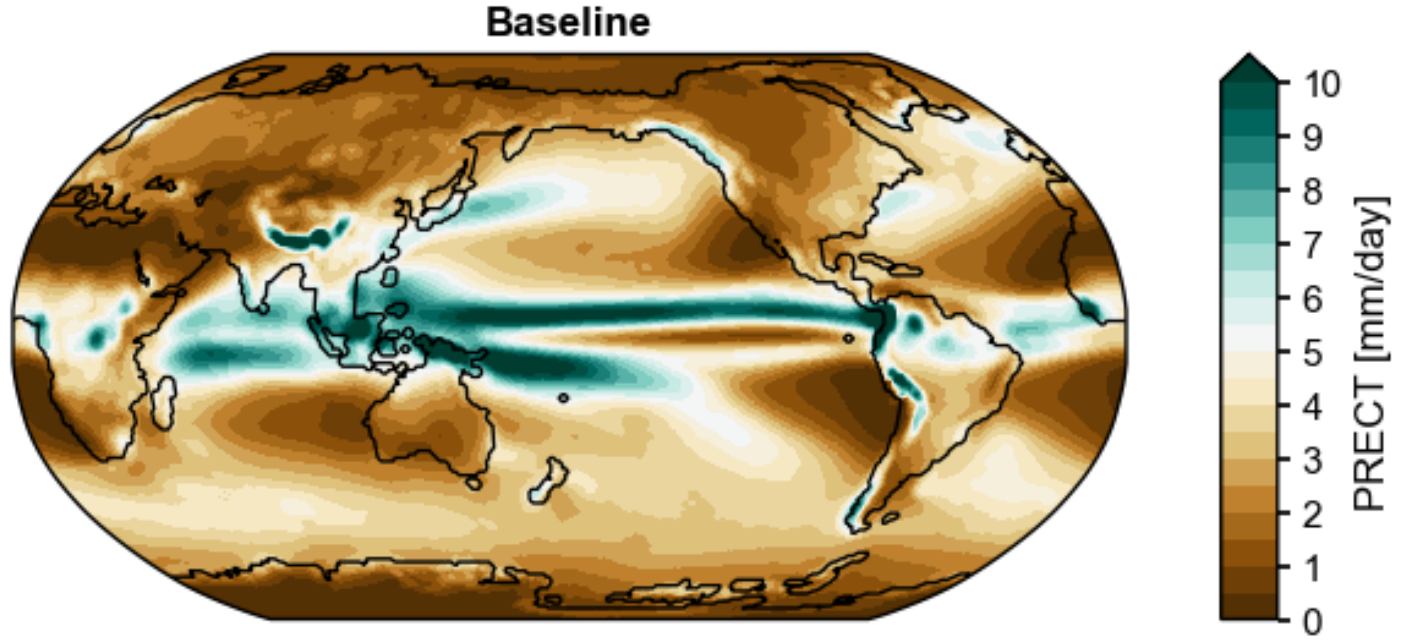
An ideal test on a Pliocene case

The Total Precipitation Rate Field

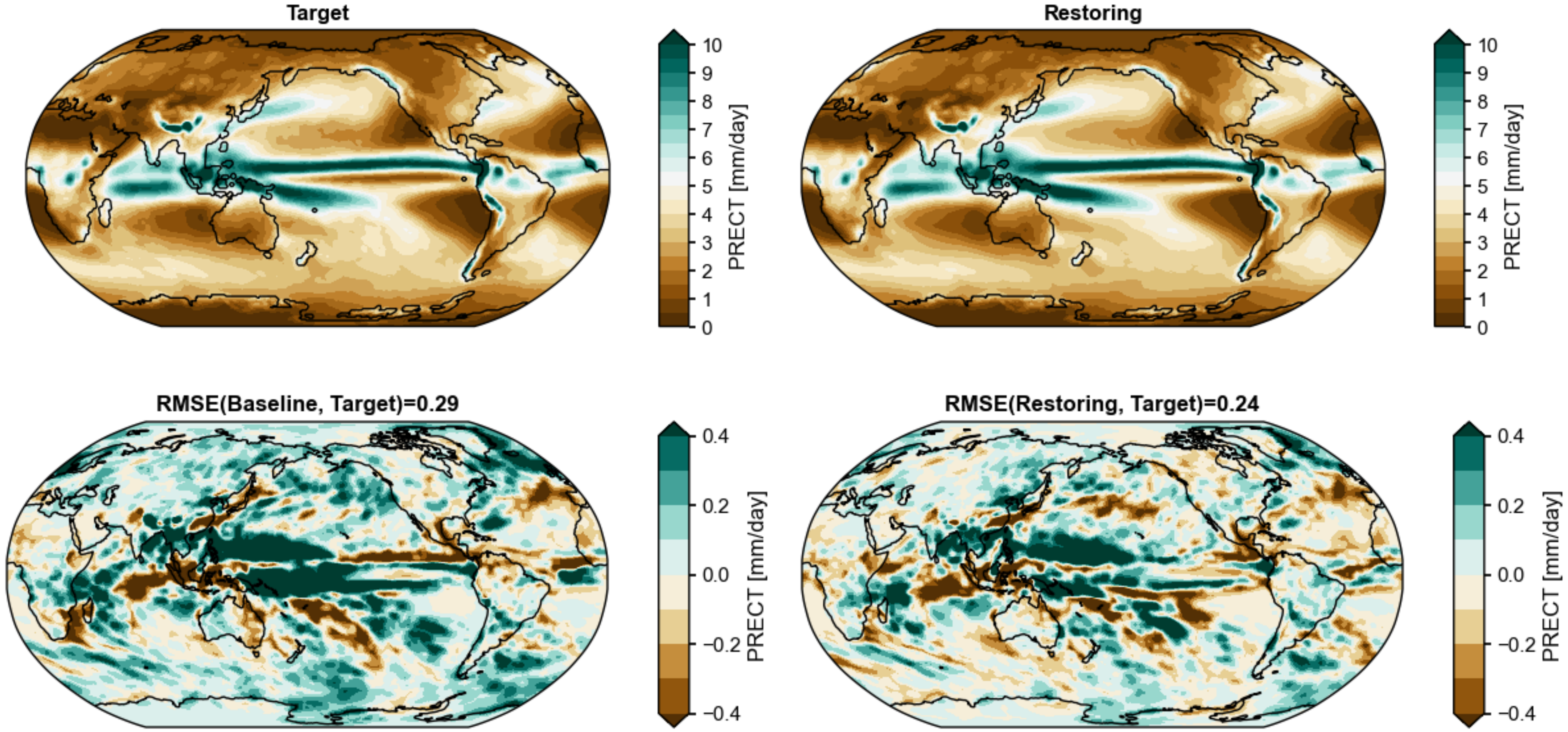


Baseline: 400ppm (Feng et al., 2020) **extended for another 800 yrs**

Target: 350ppm, branched from 400ppm (Feng et al., 2020) **and extended for another 400 yrs**



Standard Restoring

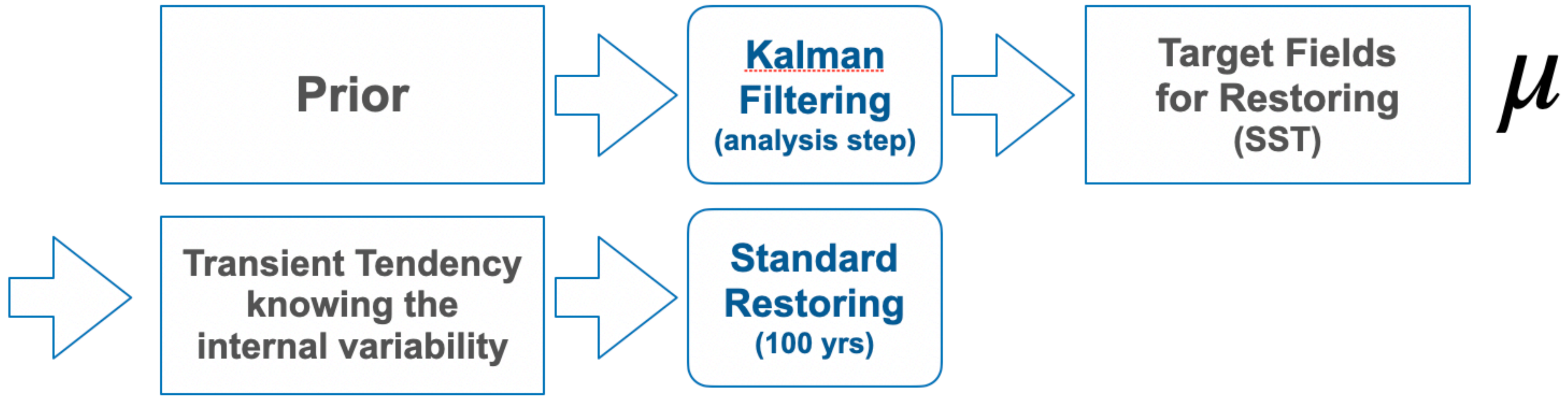


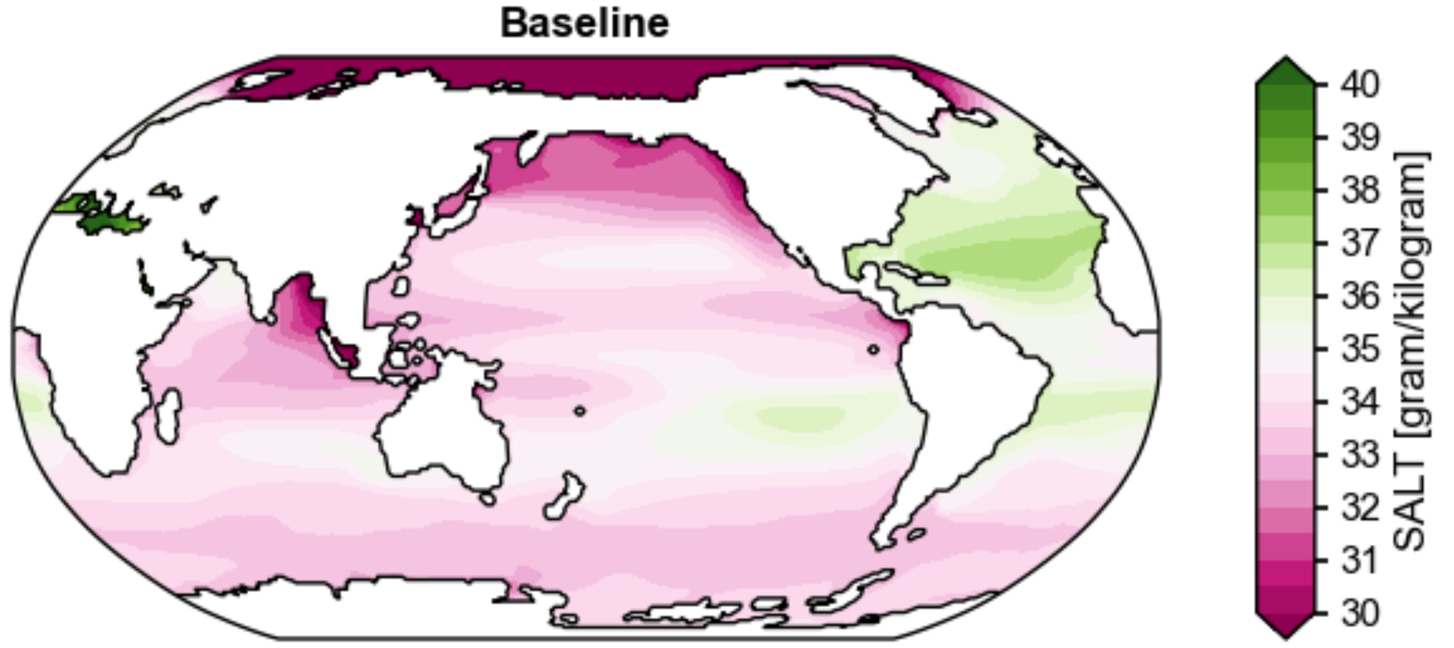
An ideal test on a Pliocene case

The Sea Surface Salinity Field

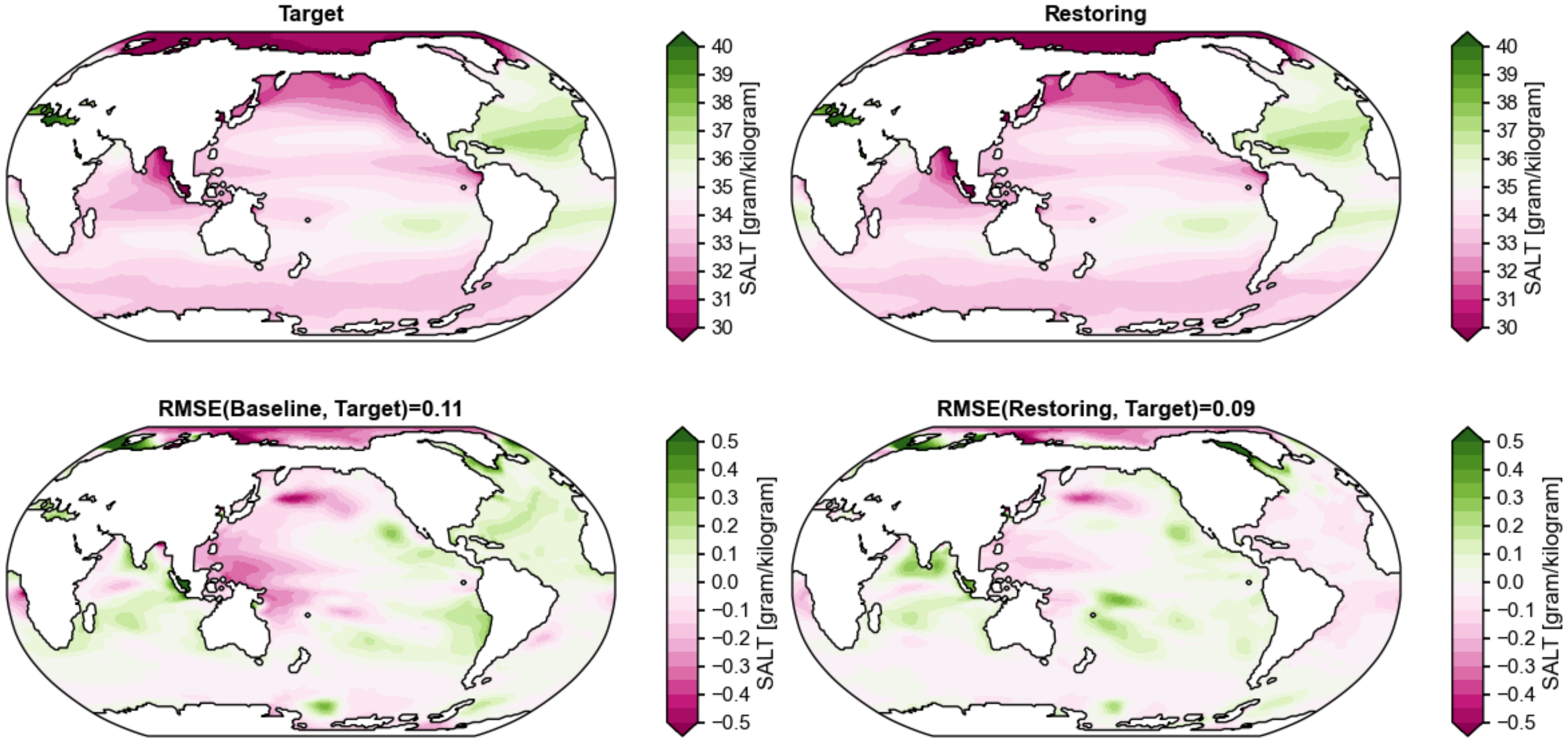
Baseline: 400ppm (Feng et al., 2020) **extended for another 800 yrs**

Target: 350ppm, branched from 400ppm (Feng et al., 2020) **and extended for another 400 yrs**





Standard Restoring

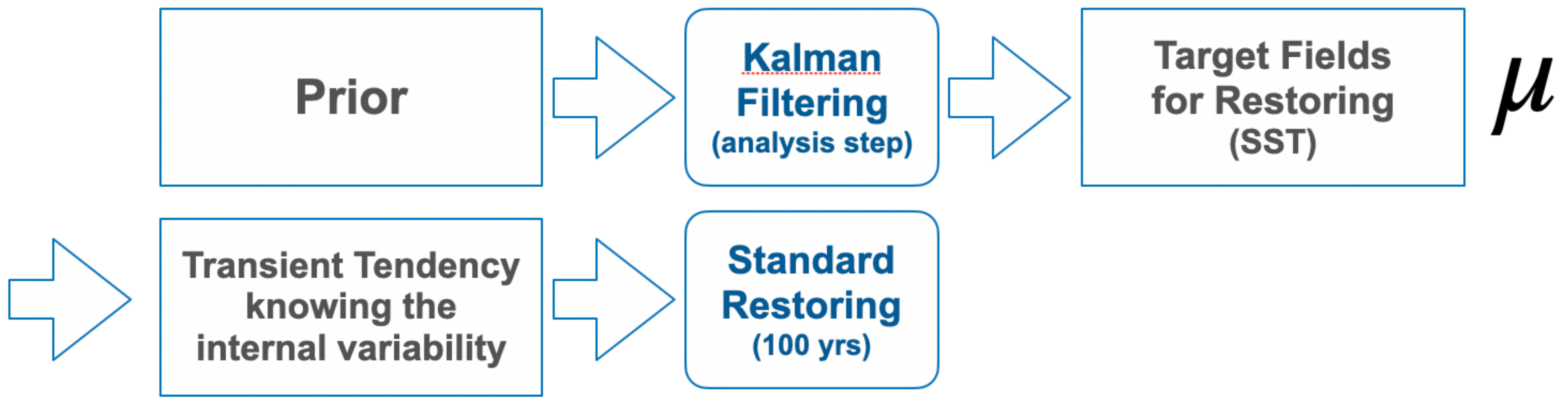


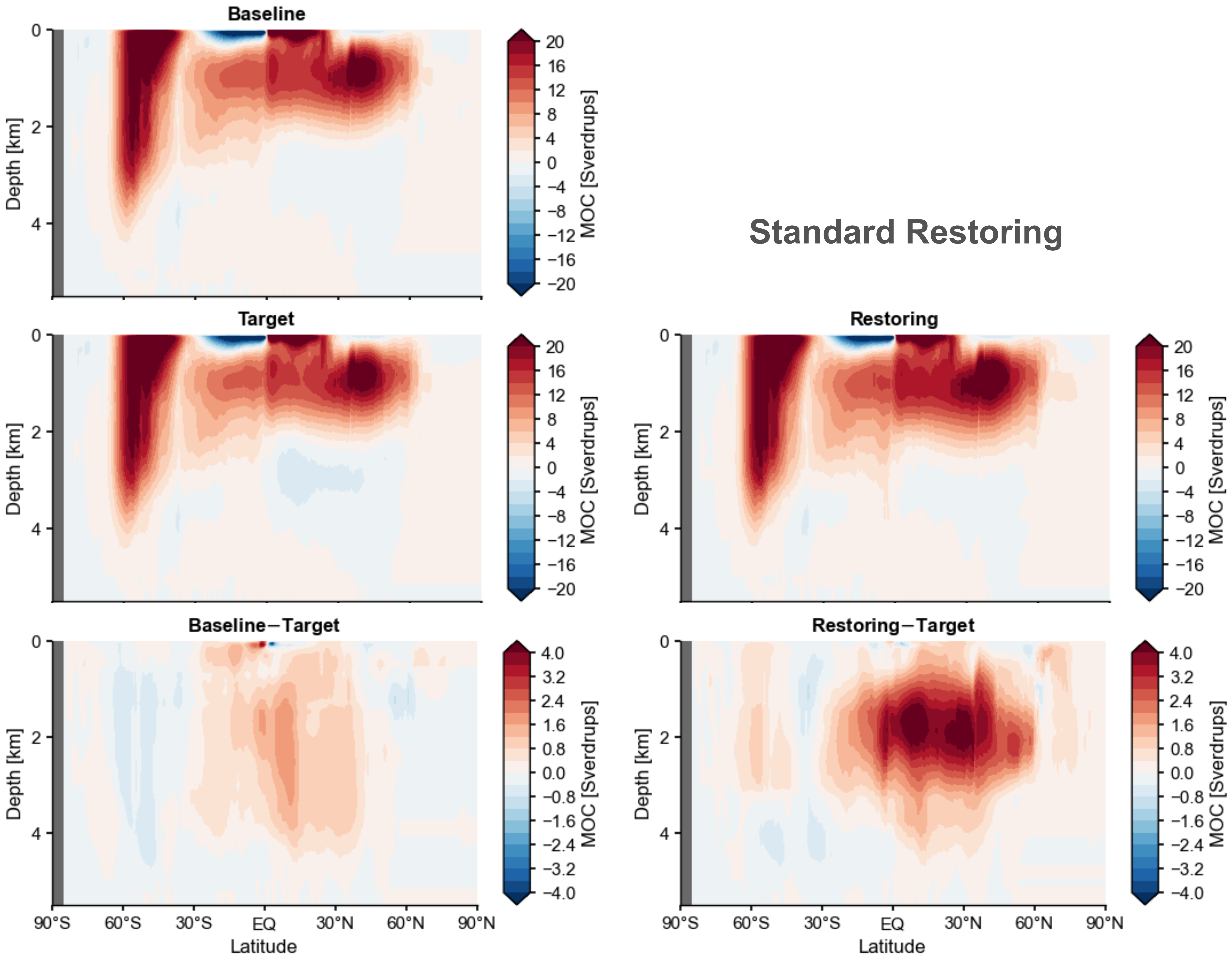
An ideal test on a Pliocene case

The Meridional Overturning Circulation (MOC)

Baseline: 400ppm (Feng et al., 2020) extended for another 800 yrs

Target: 350ppm, branched from 400ppm (Feng et al., 2020) and extended for another 400 yrs

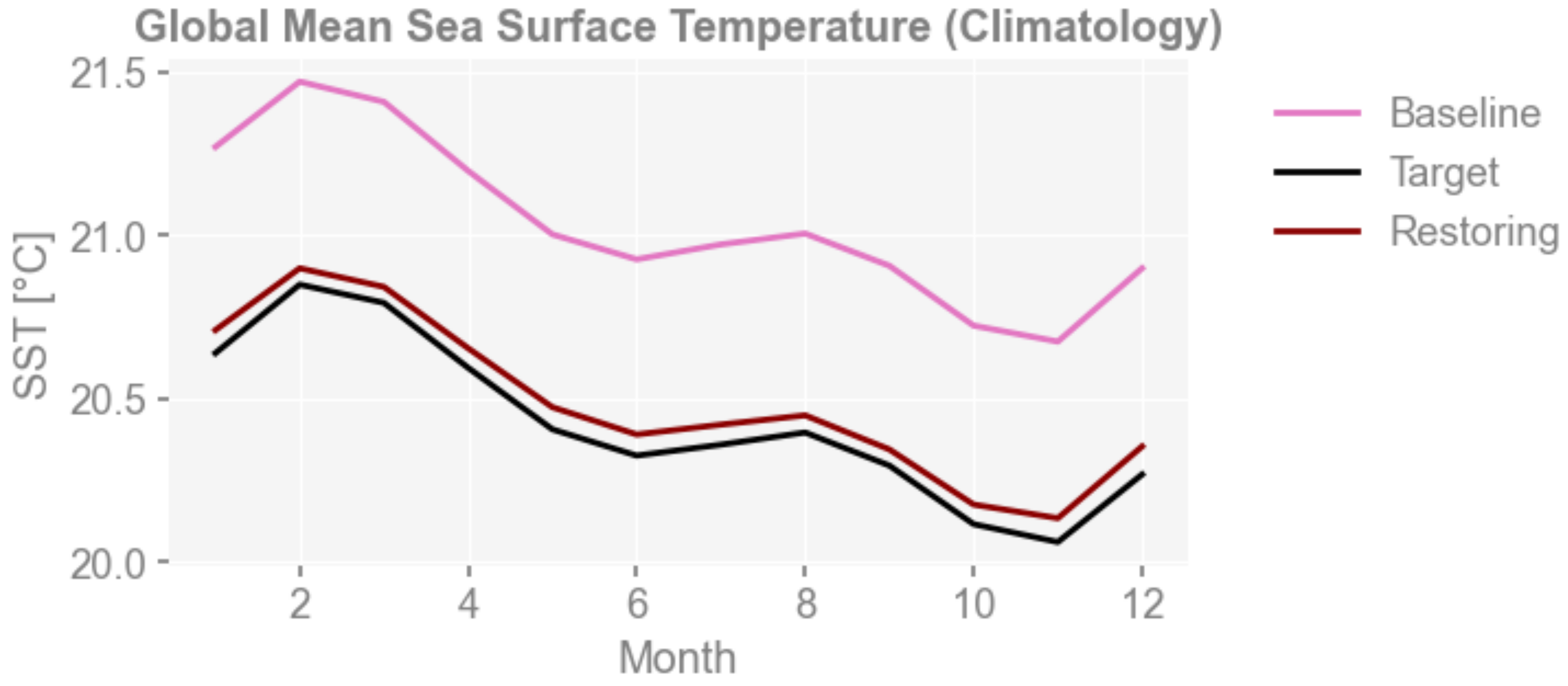


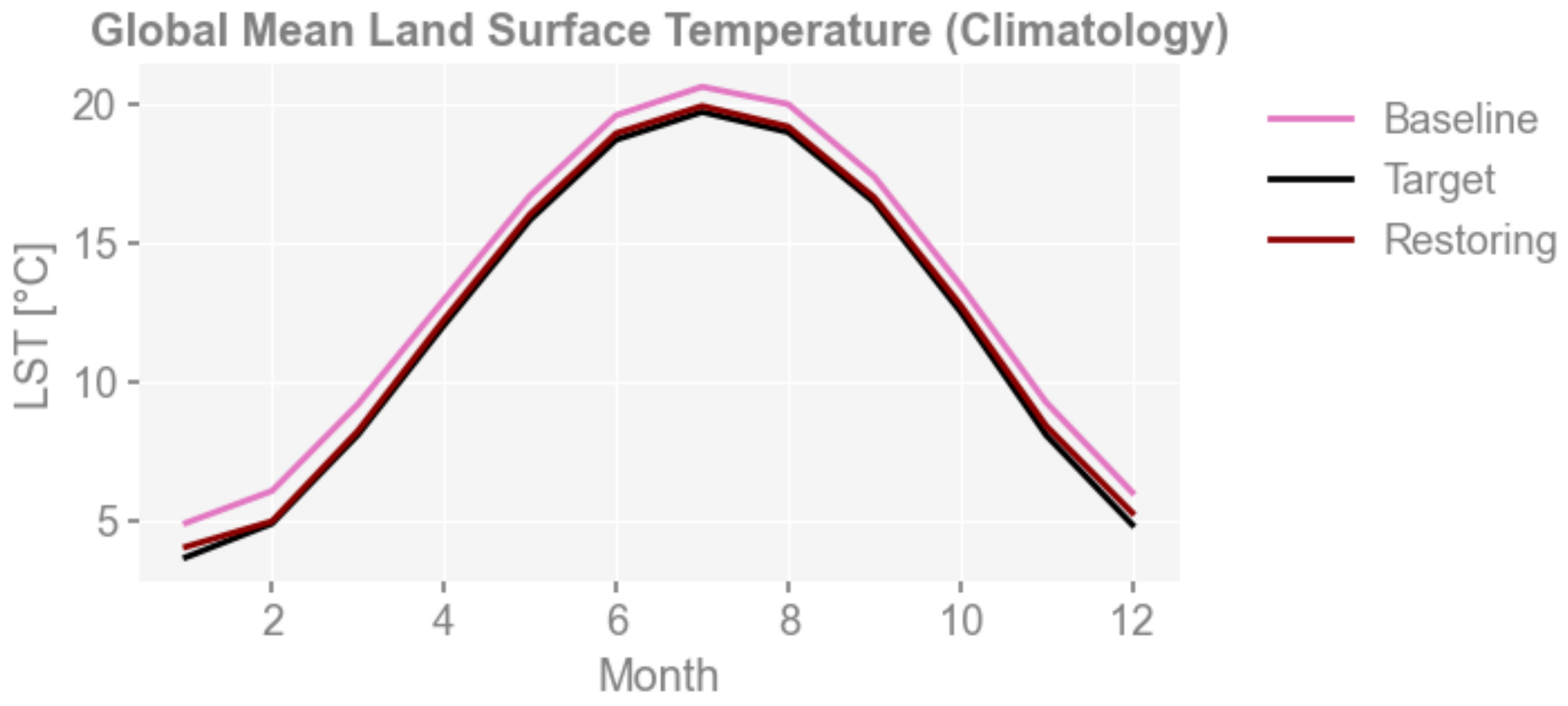


An ideal test on a Pliocene case

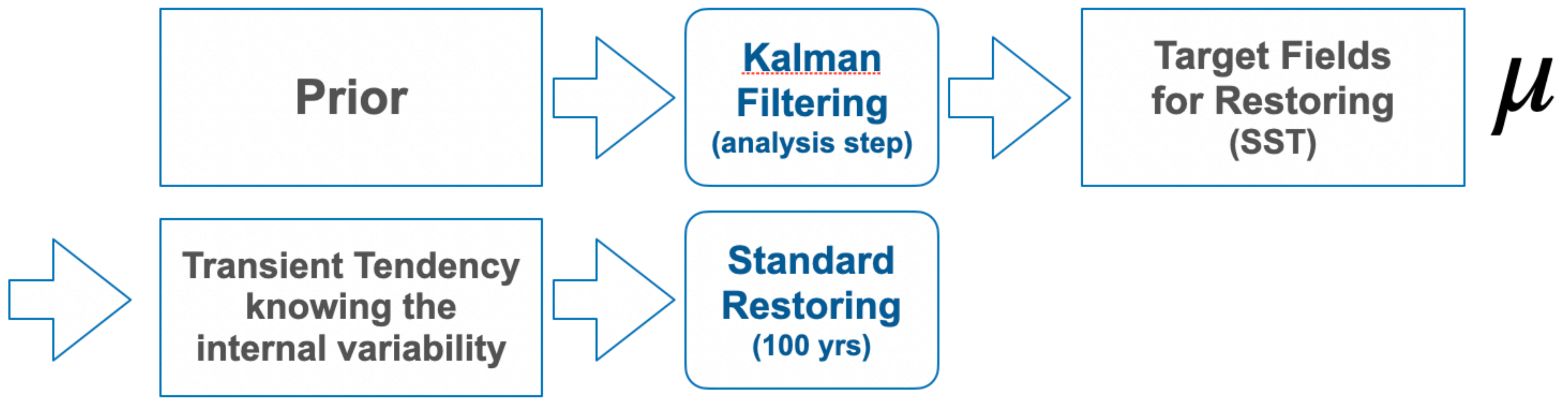
Baseline: 400ppm (Feng et al., 2020) **extended for another 800 yrs**

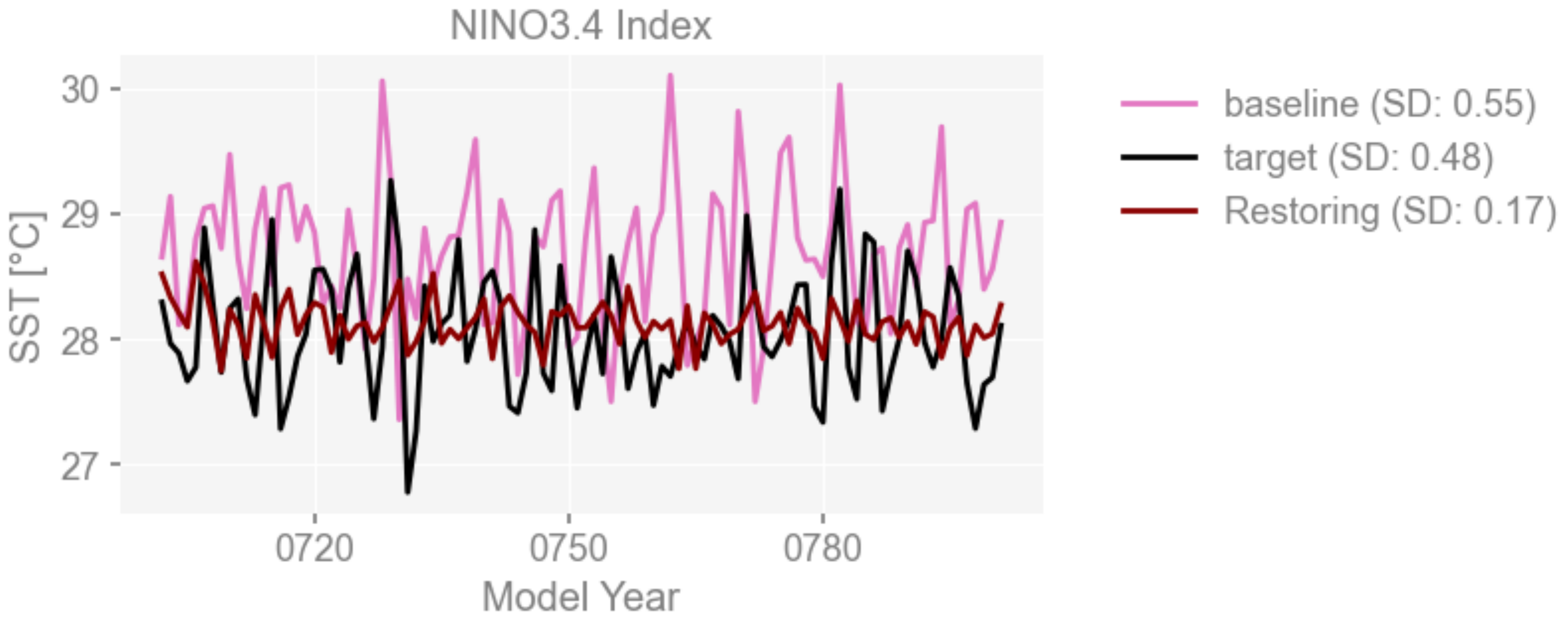
Target: 350ppm, branched from 400ppm (Feng et al., 2020) **and extended for another 400 yrs**





The Internal Variability





A new online paleoDA approach in CESM

Goal:

- fit the SST proxy observations
- a running data-informed model
- avoid unphysical disruptions (e.g., internal variability)

Prior → **Kalman Filtering** (analysis step) → Target Fields for Restoring (SST) μ

→ Transient Tendency knowing the internal variability → **Standard Restoring** (100 yrs) → Fixed Tendency *not* knowing the internal variability → **Fixed-Tendency Restoring** (100+ yrs)

$$\frac{dX(t)}{dt} = -\frac{1}{\tau}(X(t) - \mu)$$

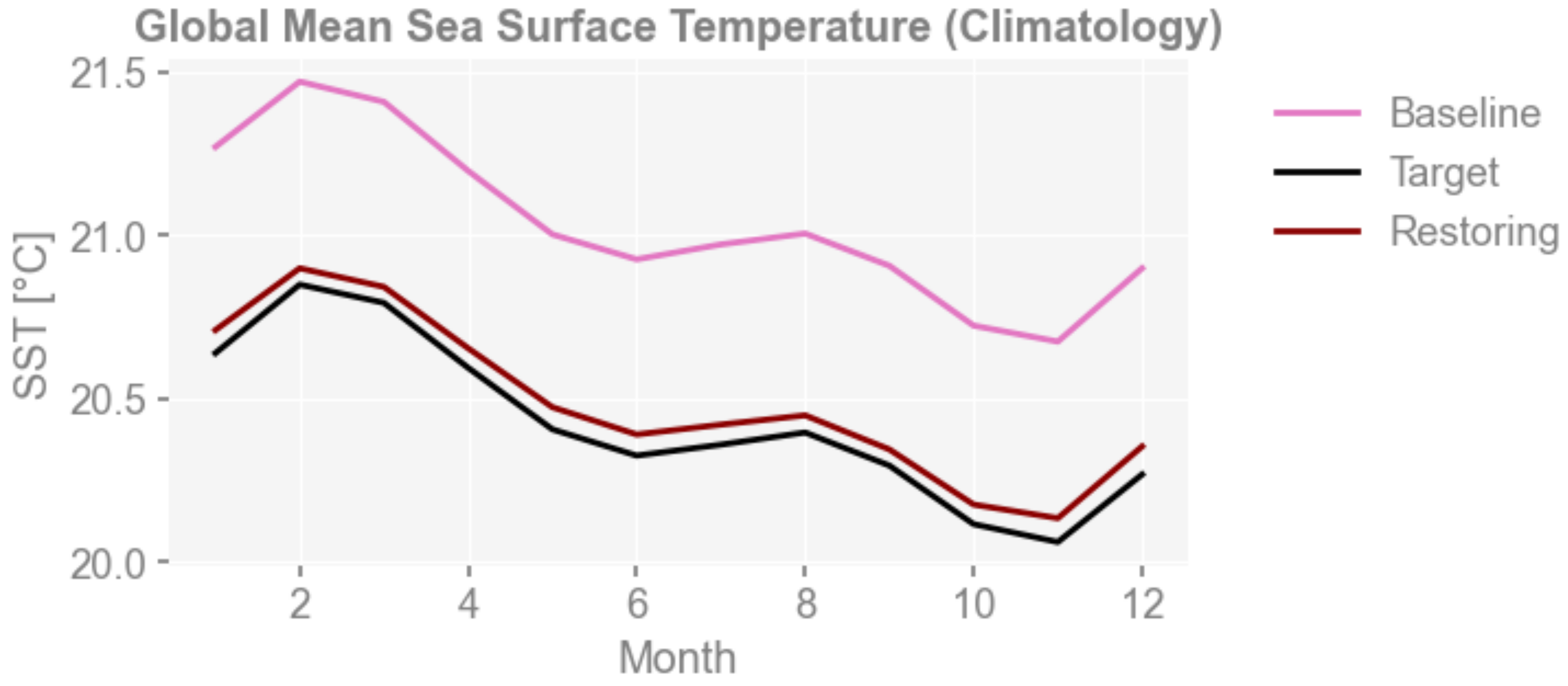
$$\frac{dX(t)}{dt} = -\frac{1}{\tau}(X_{eq} - \mu)$$

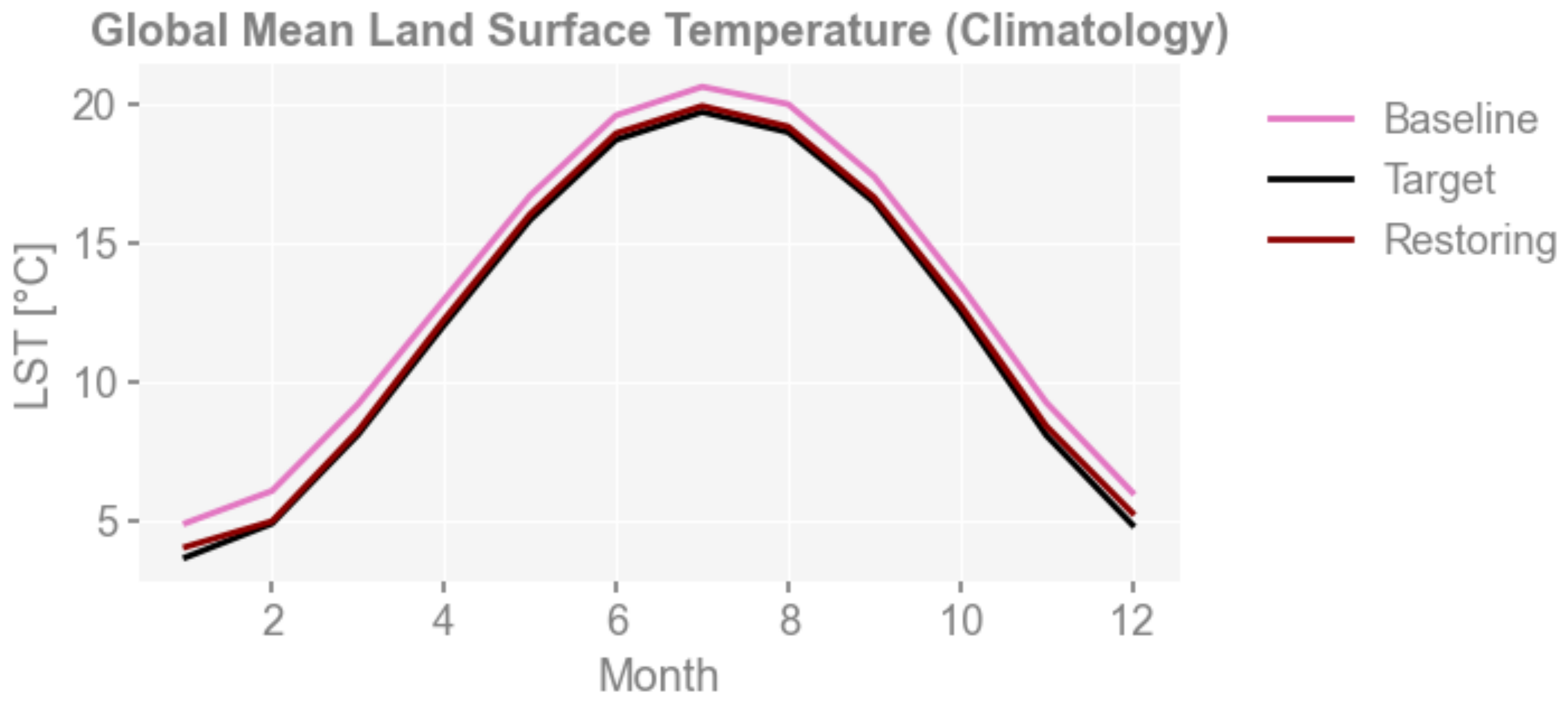
$$M(\theta', f_X) = X^*$$

An ideal test on a Pliocene case

Baseline: 400ppm (Feng et al., 2020) **extended for another 800 yrs**

Target: 350ppm, branched from 400ppm (Feng et al., 2020) **and extended for another 400 yrs**

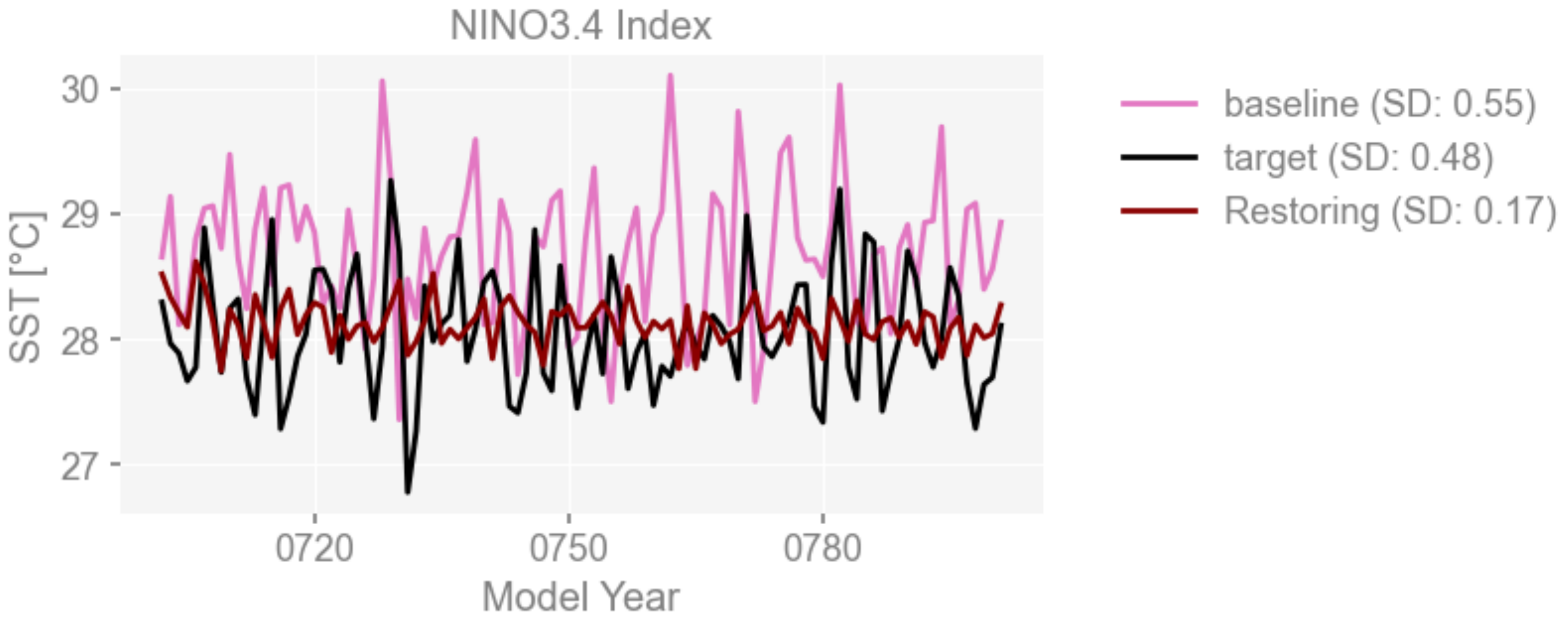




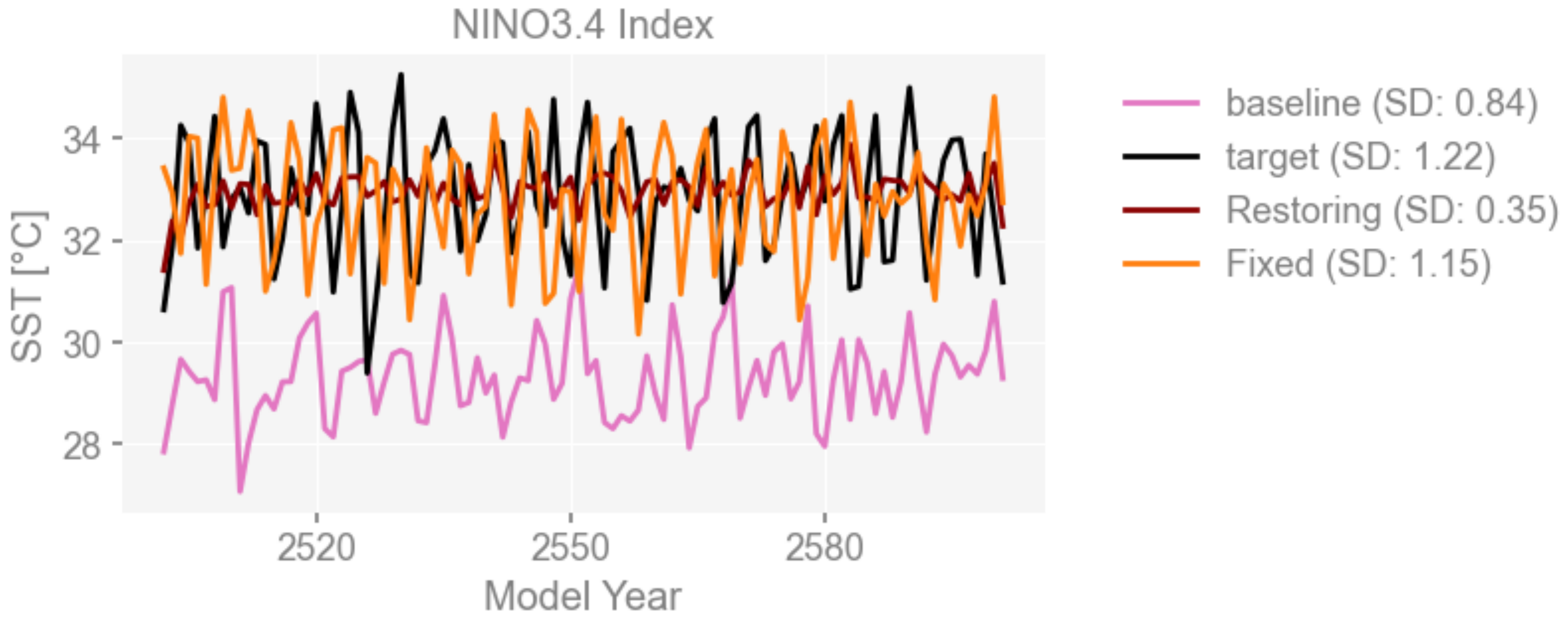
The Internal Variability

Prior → **Kalman Filtering** (analysis step) → Target Fields for Restoring (SST) μ

→ Transient Tendency knowing the internal variability → **Standard Restoring** (100 yrs) → Fixed Tendency *not* knowing the internal variability → **Fixed-Tendency Restoring** (100+ yrs)



Another test case on Miocene Climatic Optimum



Summary

- We propose **a new online paleoDA approach in CESM** for deep-time climate reconstruction.
- We perform an ideal test on Pliocene climate reconstruction and get **a data-informed model** that:
  - fits the **SST proxy observations**,
  - yields **dynamically consistent multivariate & spatial relationships**,
  - allows assessing **internal variability** under different climatological patterns,
  - much more **expensive** to run compared to offline paleoDA (wall-time: days vs minutes).
- **Preliminary** results. Welcome comments & suggestions!

Advancing Deep-Time Climate Reconstruction with a New Online Paleoclimate Data Assimilation Approach in CESM

An Ongoing Research of the *paleoWeather* Project

Thank you!

(fengzhu@ucar.edu)

Feng Zhu[1], *Dan Amrhein*[1], *Jiang Zhu*[1], *Jess Tierney*[2], *Anta-Clarisse Sarr*[3], *Sophia Macarewich*[1], *Bette Otto-Bliesner*[1], *Chris Poulsen*[3], *Elke Zeller*[2], *Maya Tessler*[2], *Dervla Meegan-Kumar*[4], *Jane W. Baldwin*[4], *and Ran Feng*[5]

1. NSF NCAR 2. U Arizona 3. U Oregon 4. UC Irvine 5. U Connecticut

Jun 11, 2025
CESM Workshop 2025

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.