LEAP

National Science Foundation

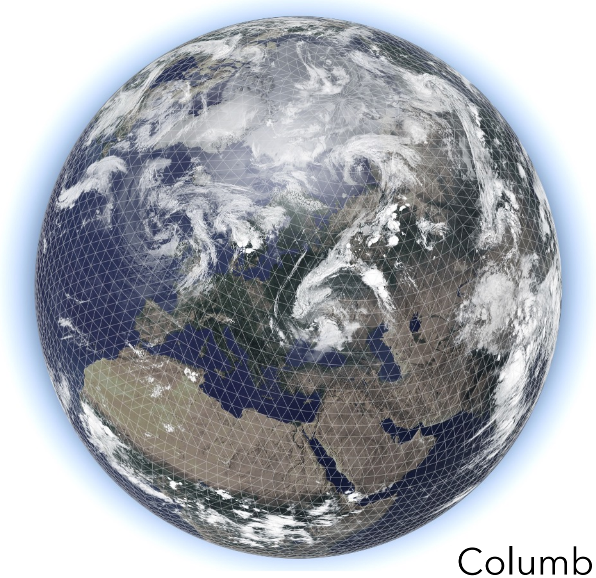
LEMONTREE
Land Ecosystem Models based On New Theory, obseRvations and ExperimEnts

Schmidt Sciences

USMILE

erc

Future(s) of Climate Simulations: Leveraging the AI revolution

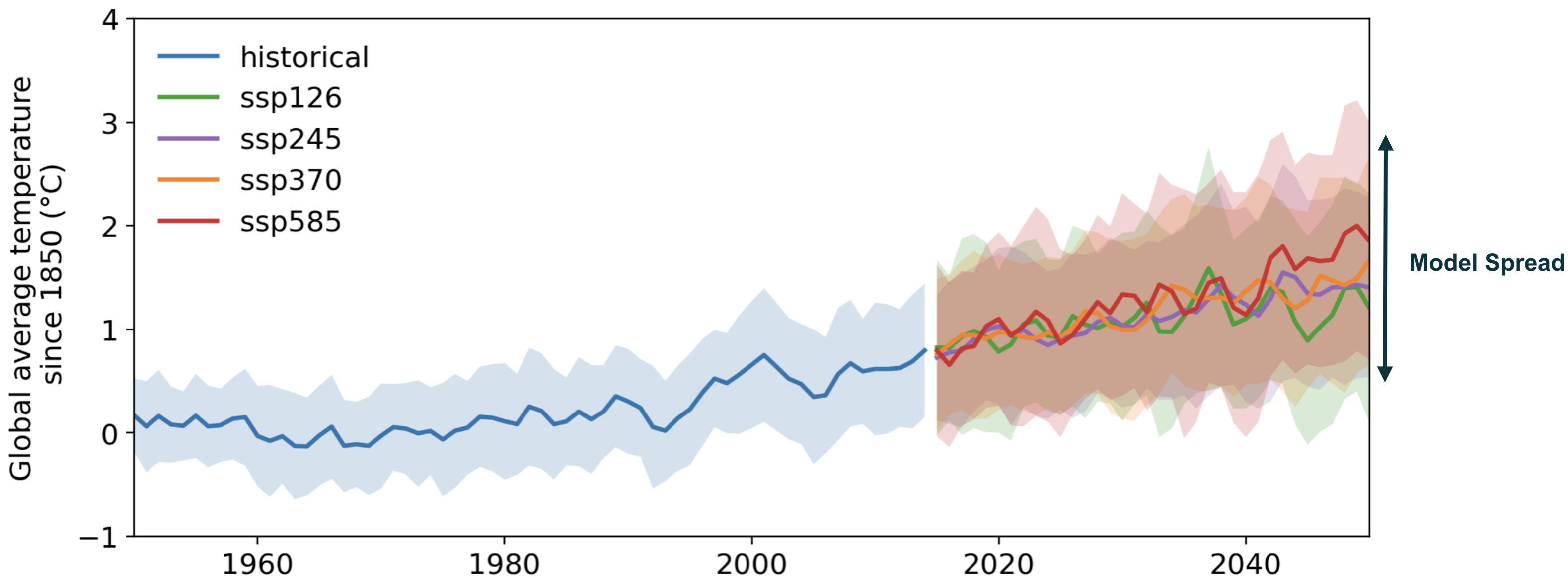


Pierre Gentine
Columbia Engineering & Climate School
Director, LEAP Center

Columbia University

Current projections are still *too* uncertain

Even for (simple) global metrics such as surface temperature

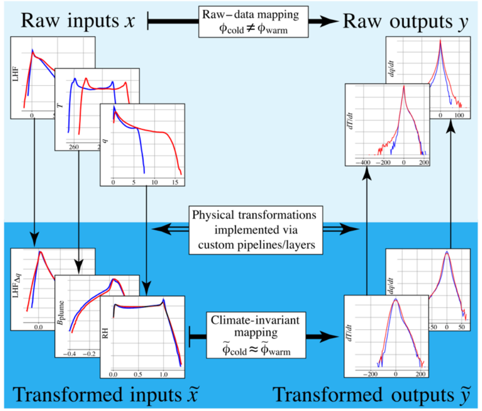


*note it takes 5 minutes to plot this with modern cloud data infrastructure (LEAP-Pangeo)

Unique challenges of climate (continuing Chris' talk)

- *Out-of-distribution* and non-stationary prediction
  → challenge for any machine learning models

Some partial fixes exist



- Need capacity to run *counterfactuals* – "what if"
  → a proper Digital Twin

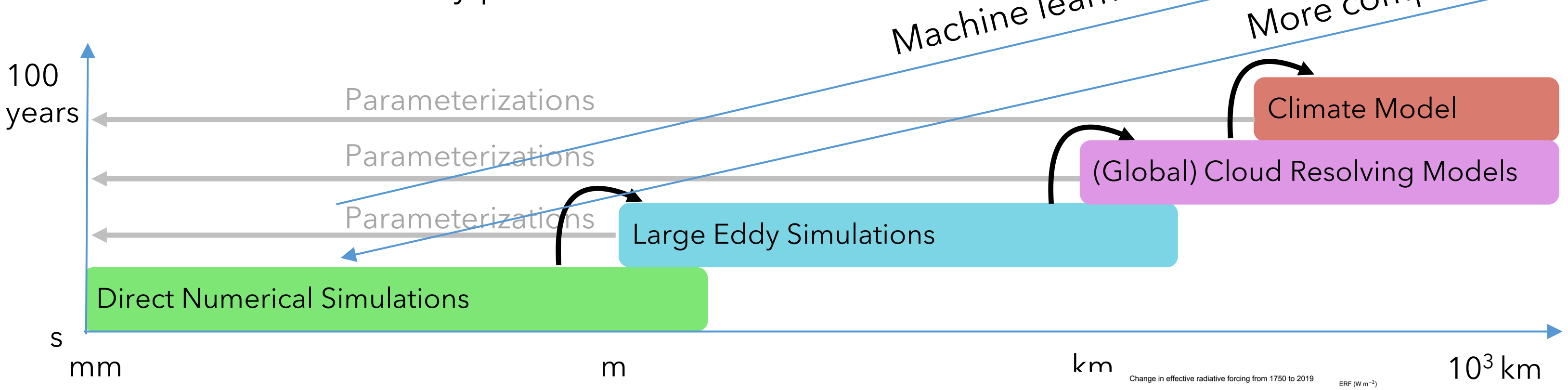
- Obviously, but often forgotten, we need *data*!

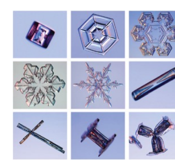
→ Many processes do not have data e.g. land carbon cycle history

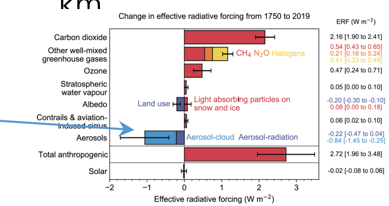
Beucler et al., *Sci Adv* 2024

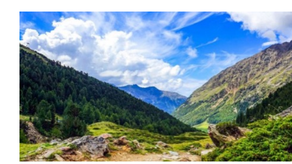
Strategies to improve climate modeling

1. Unresolved processes: parameterized, resolved or emulated:
Resolving more scales, fewer parameterizations.
We can now resolve many processes







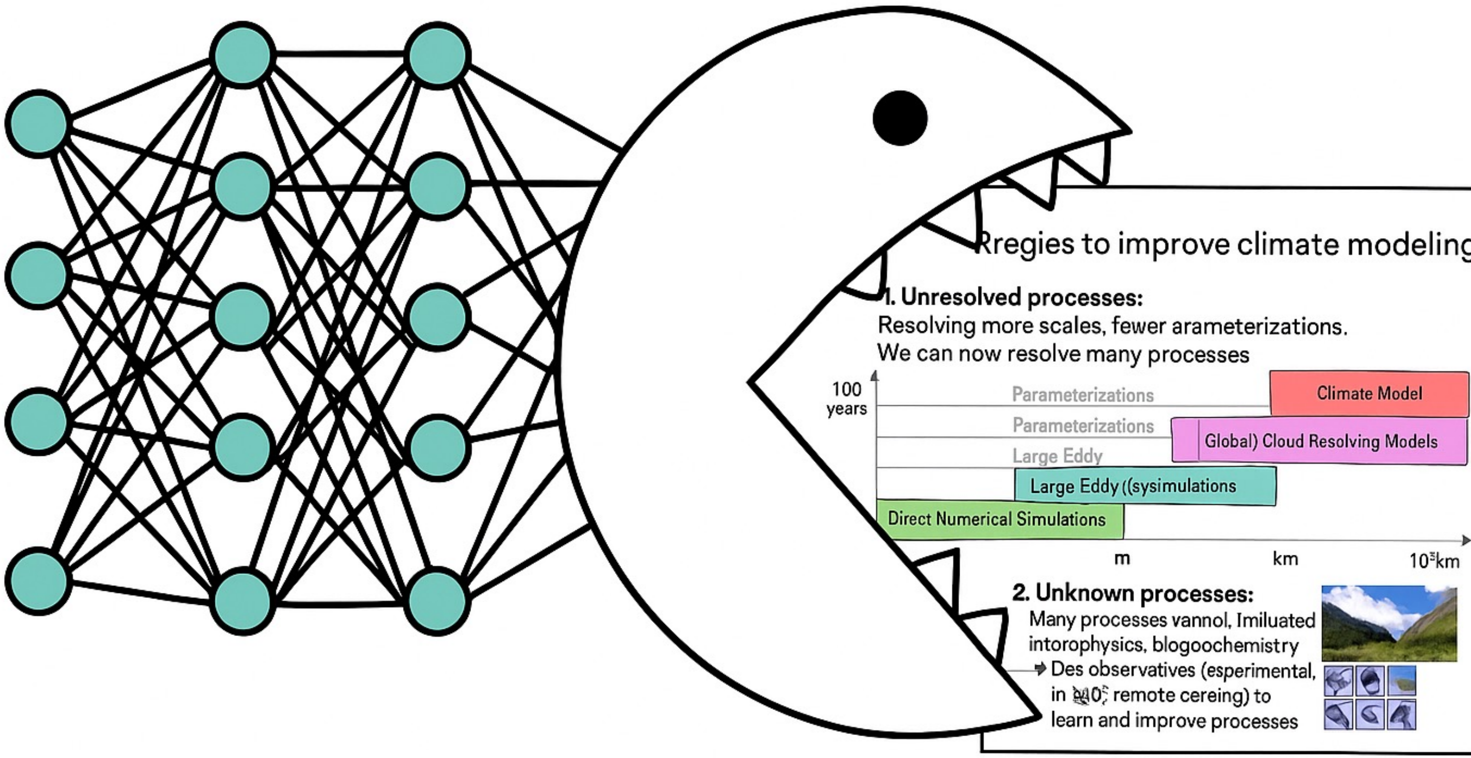


2. Unknown processes: learned or modeled:
Many processes cannot be simulated: microphysics, biogeochemistry
→ Use observations (experimental, in situ, remote sensing) to learn and improve processes

Gentine et al. 2021; Zanna et al., 2021 (Deep Learning for the Earth Sciences Linking Physics and Deep Learning Models)

Strategies to improve climate modeling

3. Machine learning the entire thing

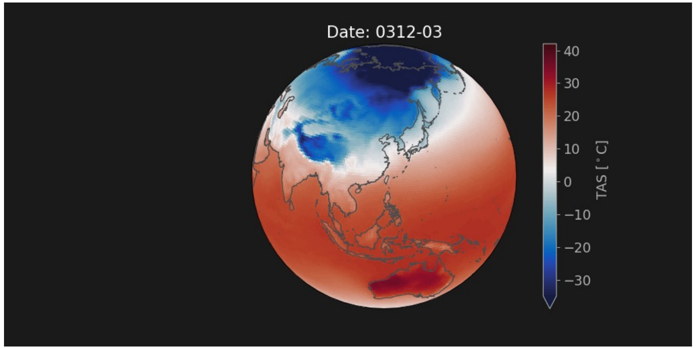


Gentine et al. 2021; Zanna et al., 2021 (Deep Learning for the Earth Sciences Linking Physics and Deep Learning Models)

Emerging strategies

Two main strategies are gaining tractions in the physical sciences

1. Machine-learning <u>emulation</u> of entire model/components. -see Chris' talk



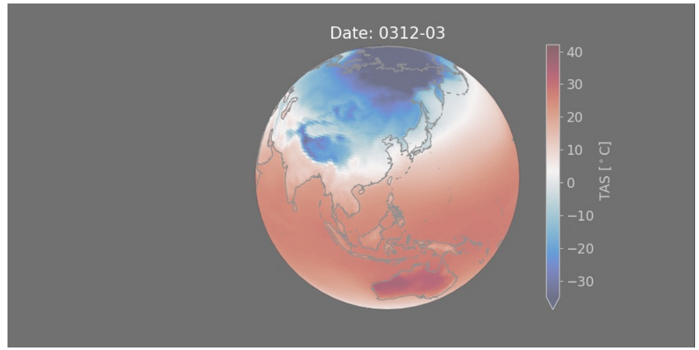
SamudrACE coupled ocean-atmosphere AI model

Duncan et al. 2025 arxiv, Fang et al., JAMES 2024, arxiv 2025, Häfner et al. 2021

Emerging strategies

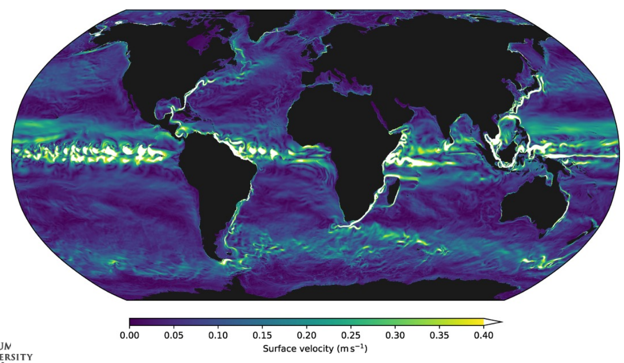
Two main strategies are gaining tractions in the physical sciences

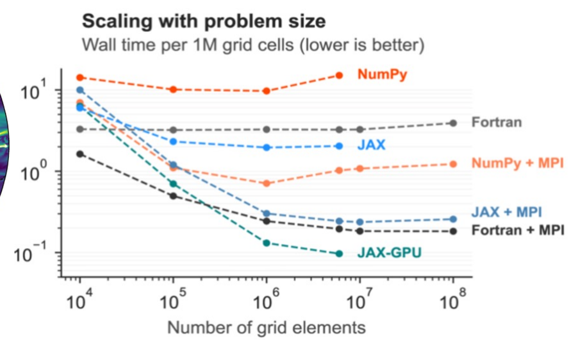
1. Machine-learning emulation of entire model/components. -see Chris' talk



SamudrACE coupled ocean-atmosphere AI model

2. Modernized physics codes (not yet in climate but working on it) that leverage the AI infrastructure

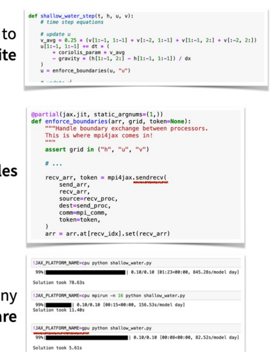




Simple to write

Scales

Runs on any hardware

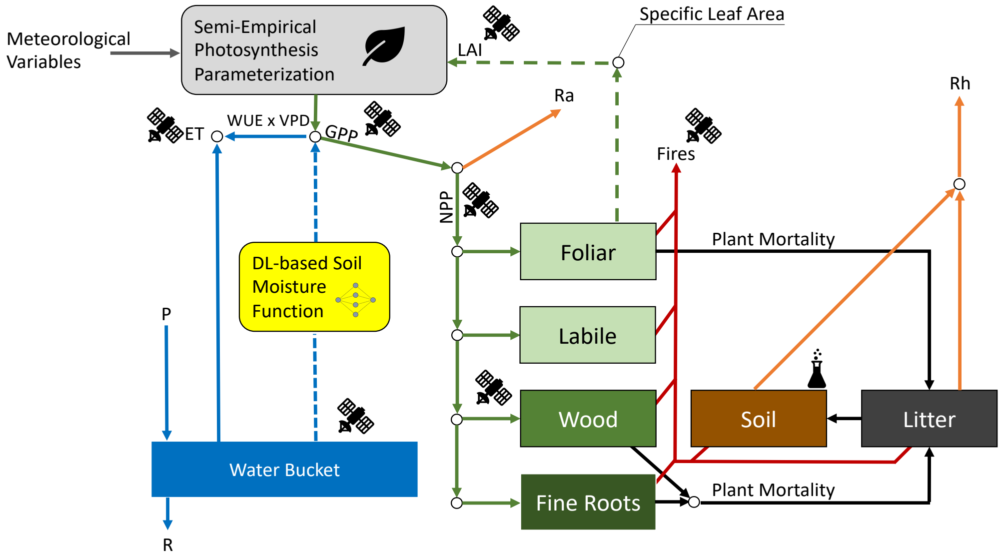


Duncan et al. 2025 arxiv, Fang et al., JAMES 2024, arxiv 2025, Häfner et al. 2021

Example differentiable model: Land Model

DifferLand: end-to-end differentiable global land model with observational operators.

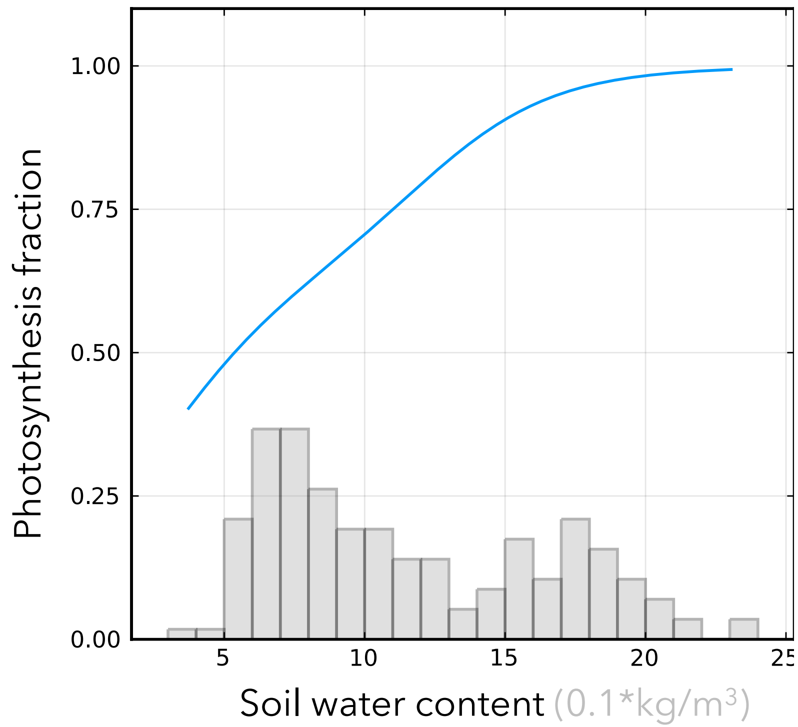
- **AI-ready** : Estimate Initial Conditions, ML or Physics Parameters
- **Agile and Fast**: Can be run on laptop, HPC (CPUs or GPUs) or on the cloud!
- **AI-ready**: Merges Data Assimilation & ML

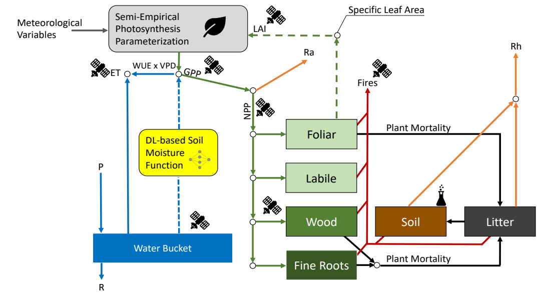


Fang et al. 2024, JAMES; Fang et al., Nature Comm in review , Qu, Bhouri and Gentine *ICLR* 2024

Example differentiable model: Land Model

Example: **parameterization** estimation: water stress

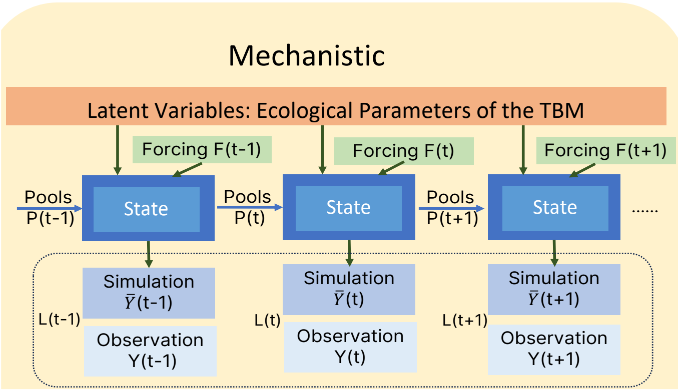




Fang et al. 2024, JAMES

Example differentiable model: Land Model

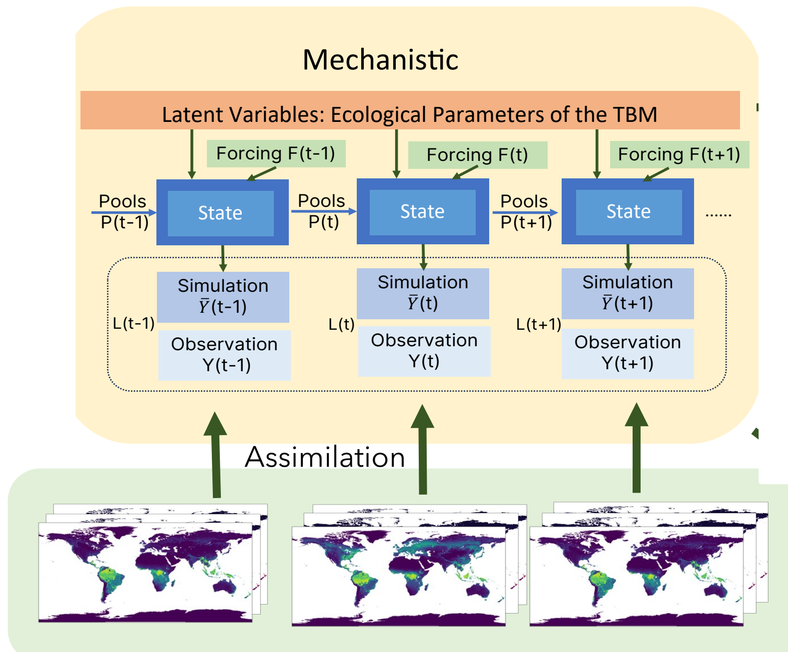
Example: spatial **parameter** estimation



Fang et al. 2024, JAMES; Fang et al., Nature Comm in review , Qu, Bhouri and Gentine *ICLR* 2024

Example differentiable model: Land Model

Example: spatial **parameter** estimation

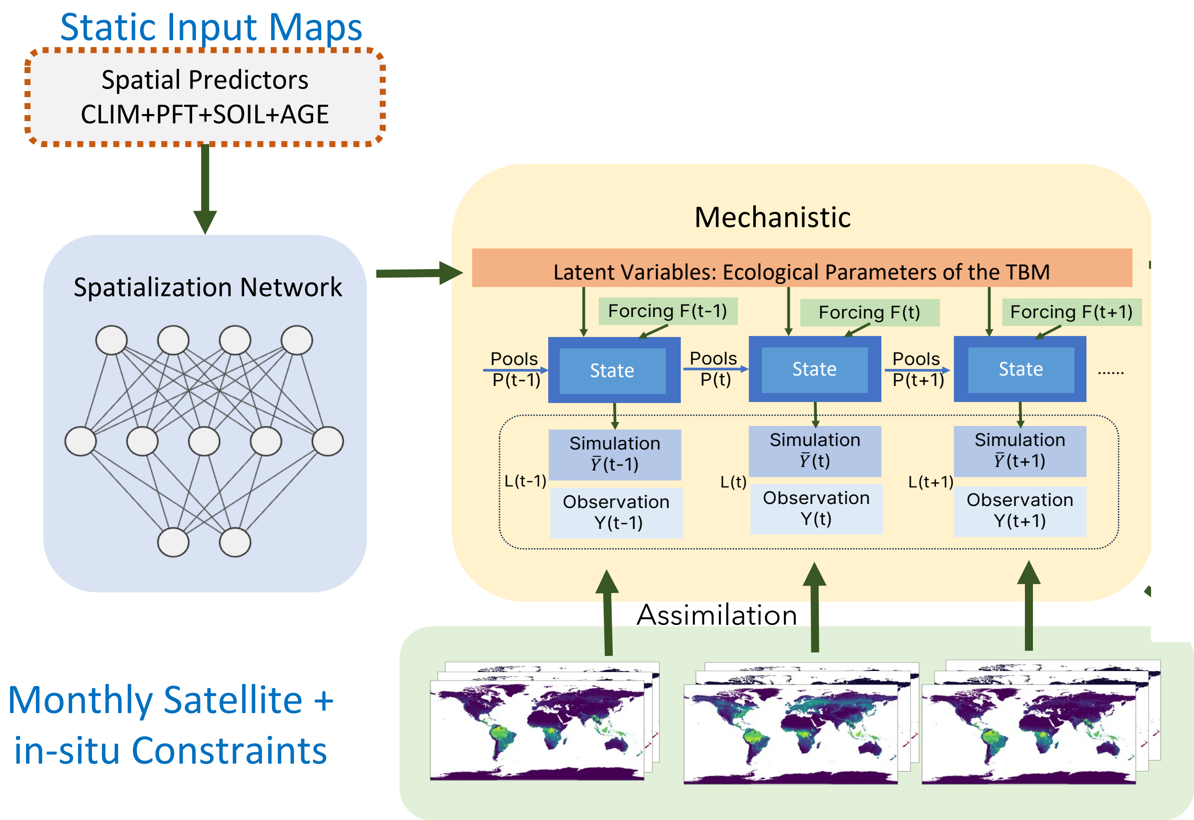


Monthly Satellite +
in-situ Constraints

Fang et al. 2024, JAMES; Fang et al., Nature Comm in review , Qu, Bhouri and Gentine *ICLR* 2024

Example differentiable model: Land Model

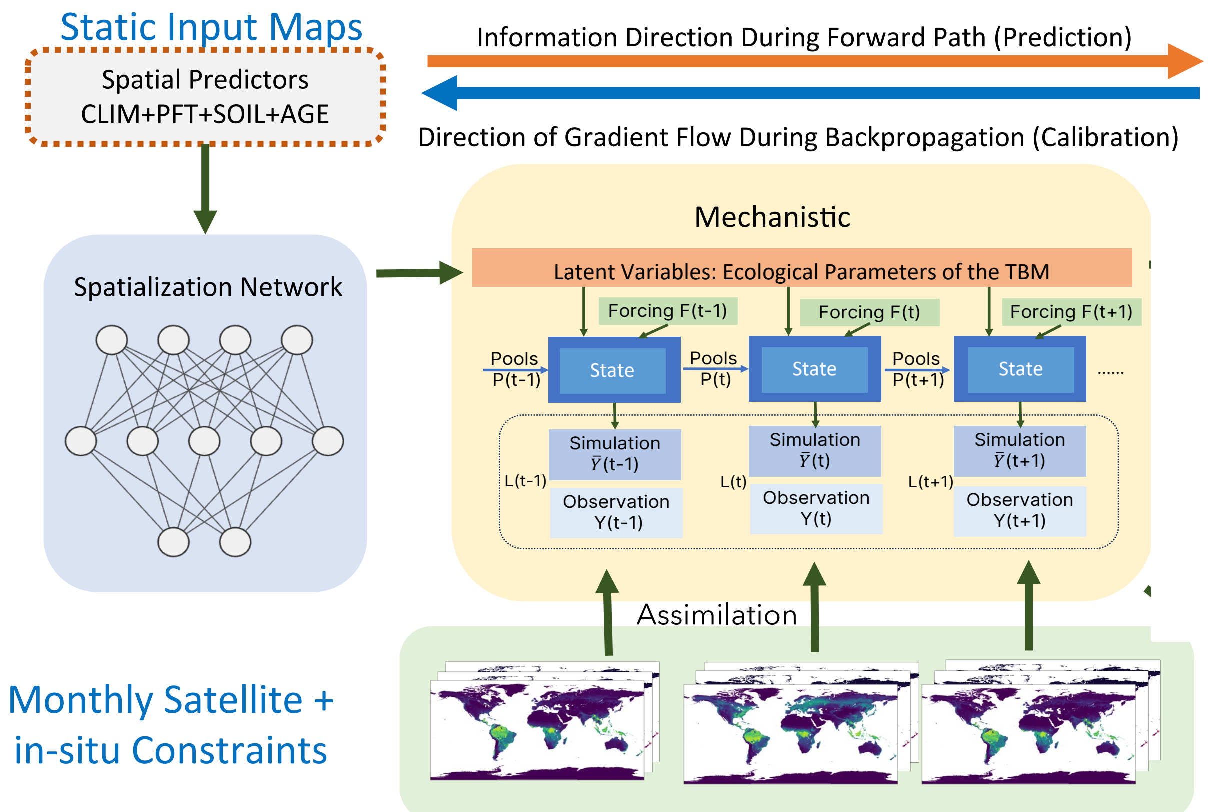
Example: spatial **parameter** estimation



Fang et al. 2024, JAMES; Fang et al., Nature Comm in review , Qu, Bhouri and Gentine *ICLR* 2024

Example differentiable model: Land Model

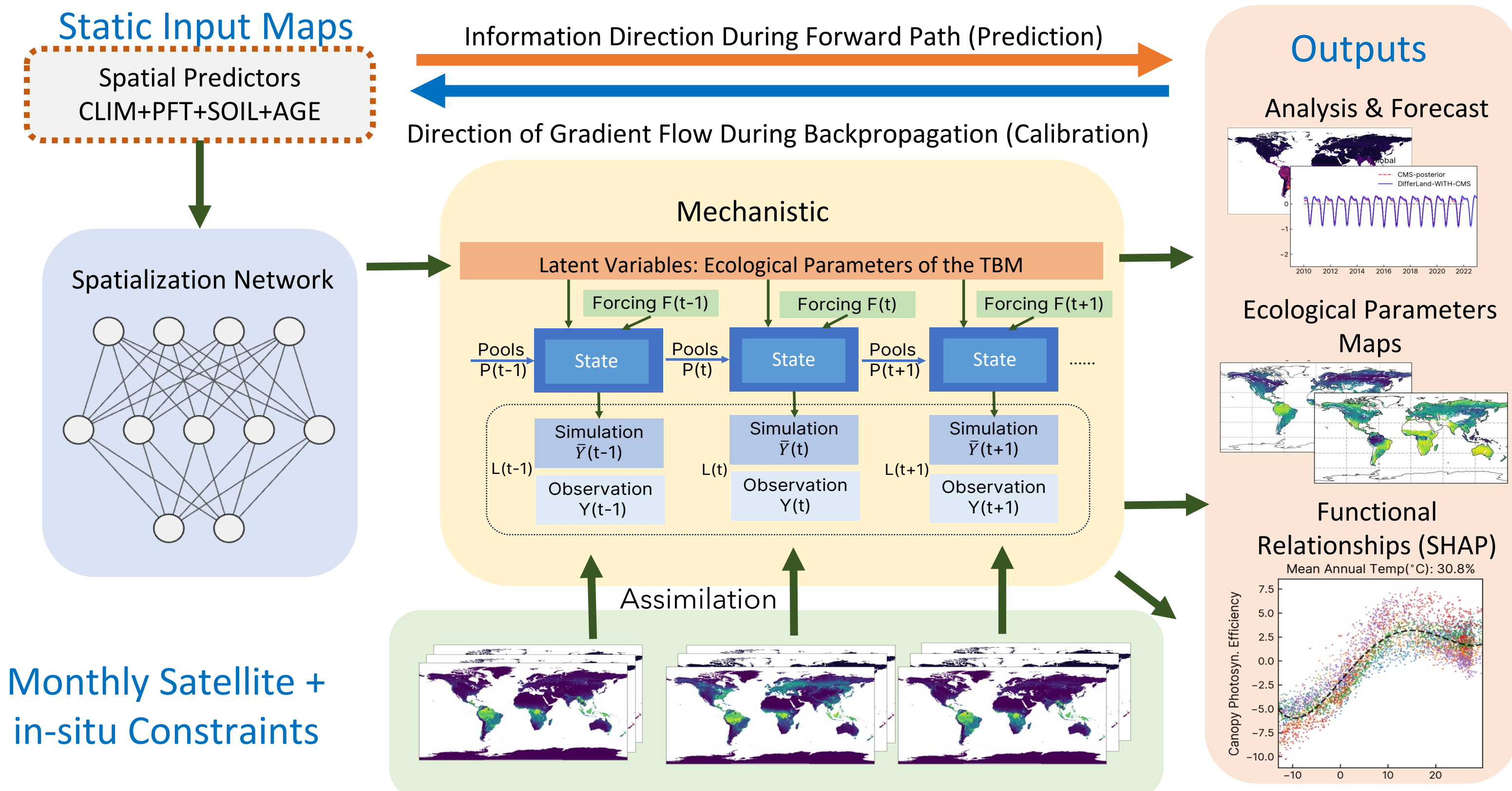
Example: spatial **parameter** estimation



Fang et al. 2024, JAMES; Fang et al., Nature Comm in review , Qu, Bhouri and Gentine *ICLR* 2024

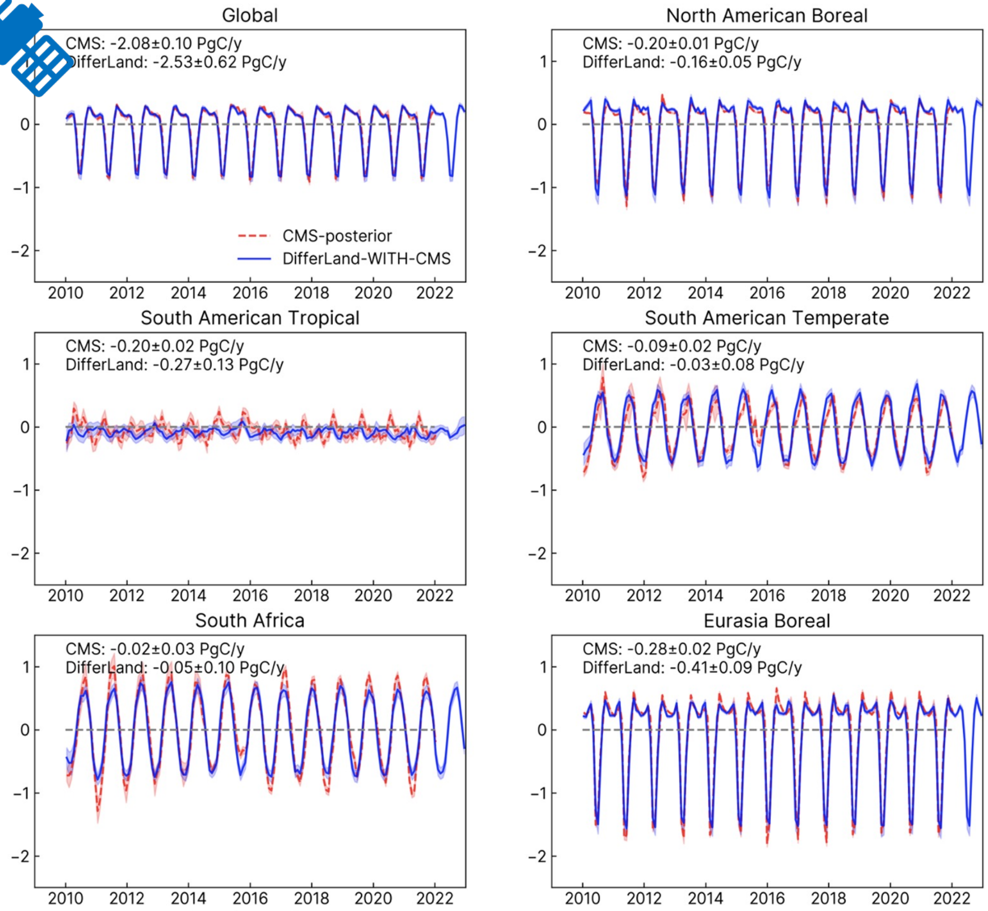
Example differentiable model: Land Model

Example: spatial **parameter** estimation

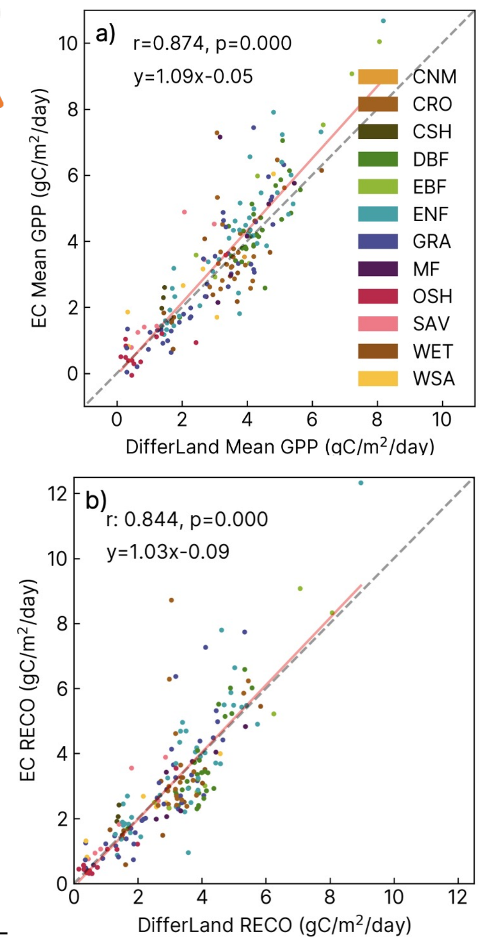


Fang et al. 2024, JAMES; Fang et al., Nature Comm in review , Qu, Bhouri and Gentine *ICLR* 2024

Top Down: Atmospheric NBE Inversion



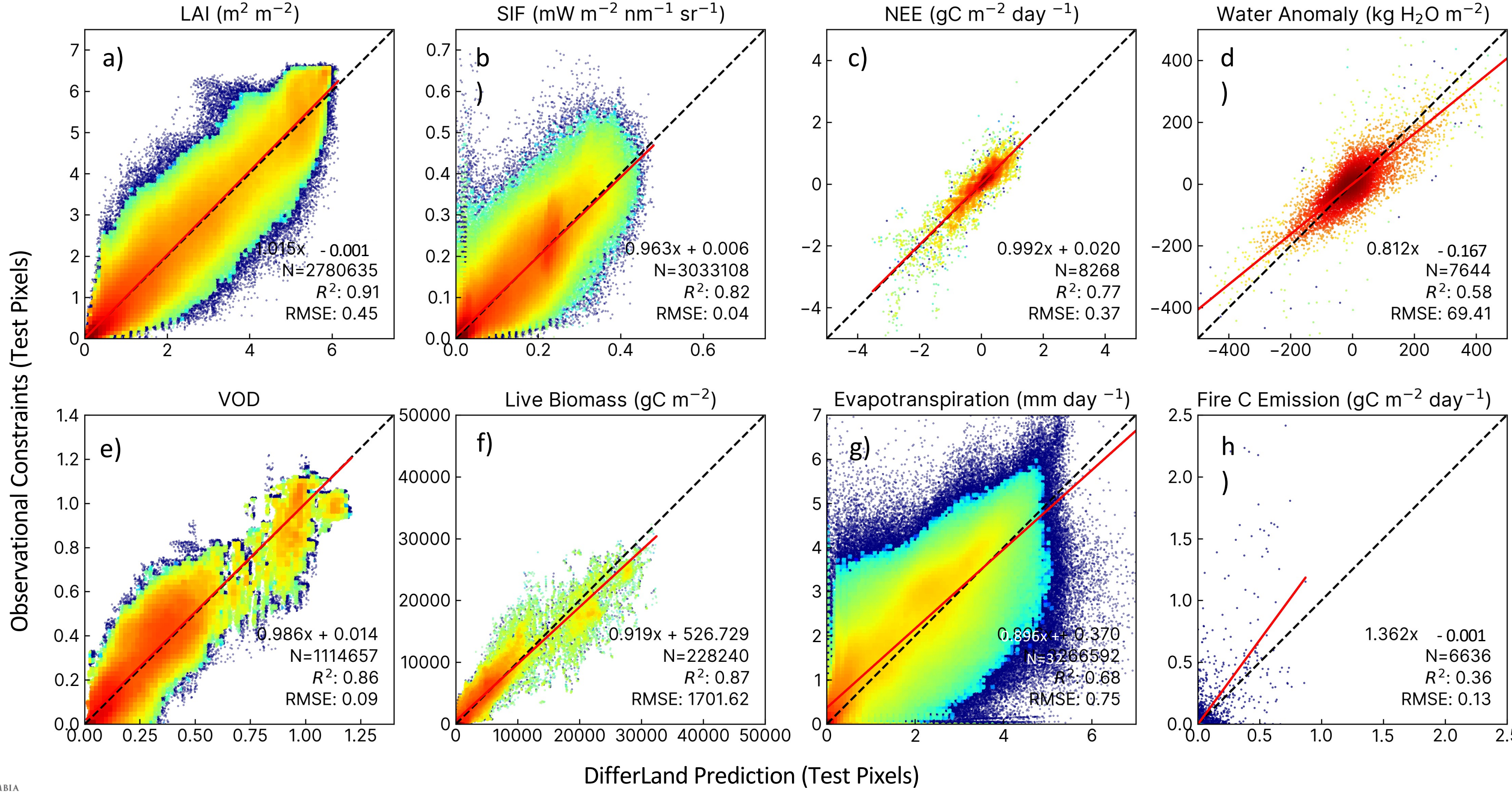
Bottom Up: Eddy Covariance



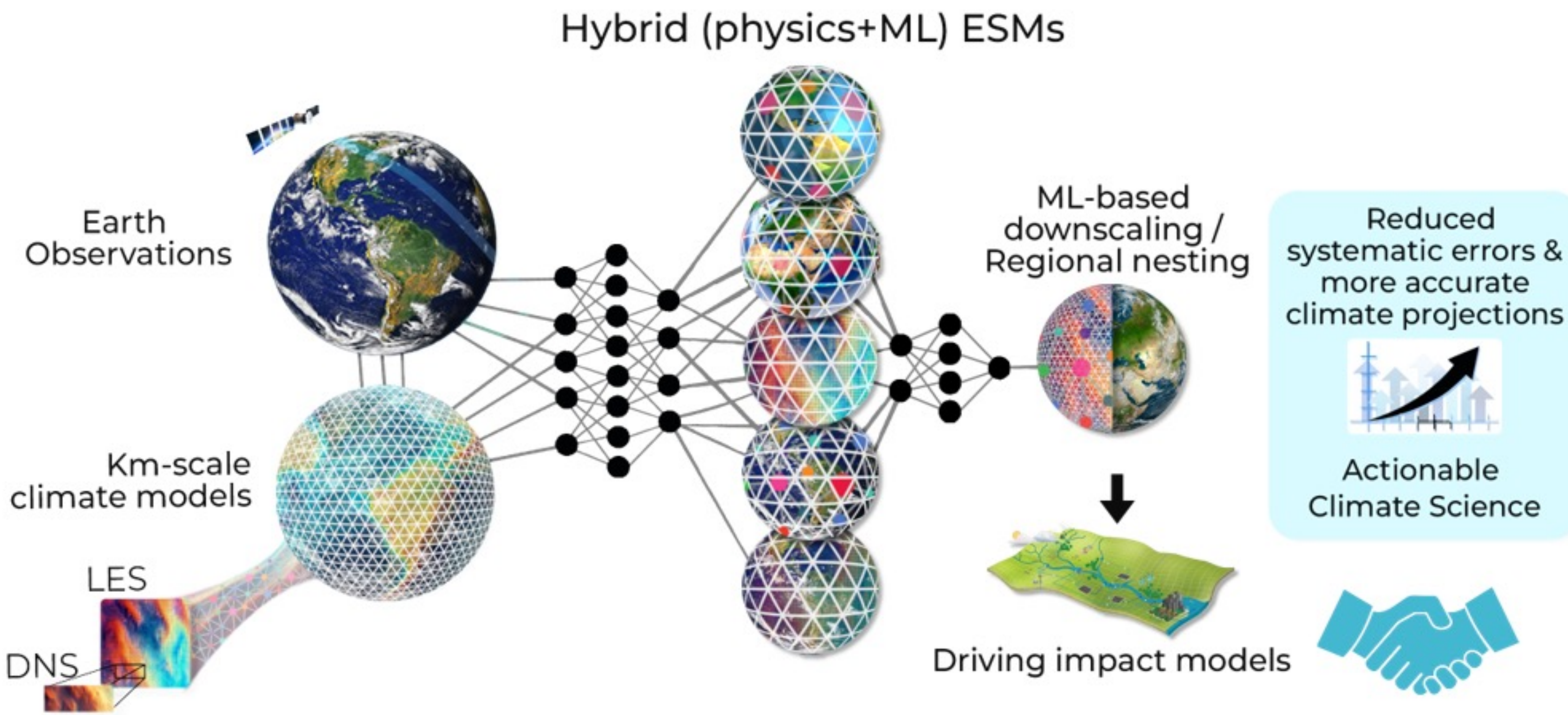
GPP

RECO

Robust Generalization at Unseen Locations

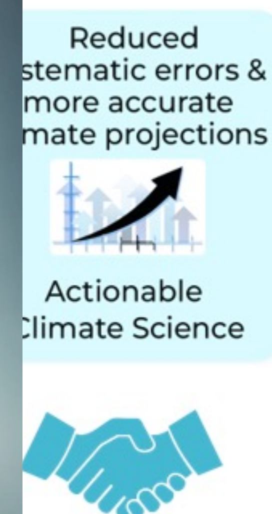


Vision for fully hybrid model



Eyring, Gentine et al *Nature Geo*, 2024

LEGO Land: swap not just parameterizations but entire modules



Eyring, Gentine et al *Nature Geo*, 2024

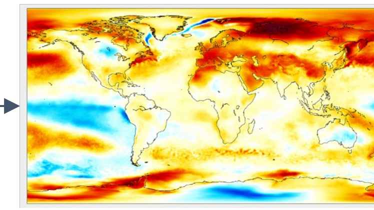
Generative AI for simulation, data assimilation and downscaling

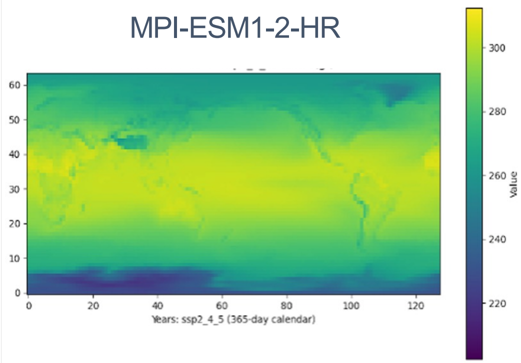
Diffusion model to emulate distribution of complex physics simulation $p_\theta(T|m, s, C, d)$

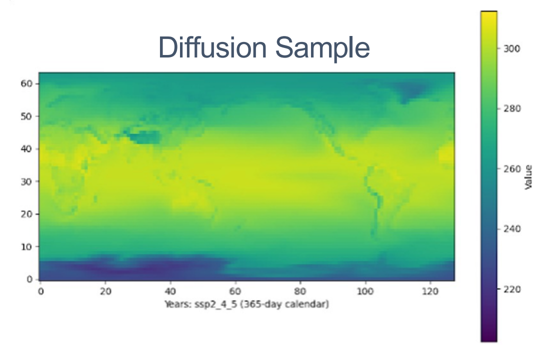
- CMIP6 model (m)
- SSP scenario (s)
- CO_2e for year y (C)
- Day of the year (d)

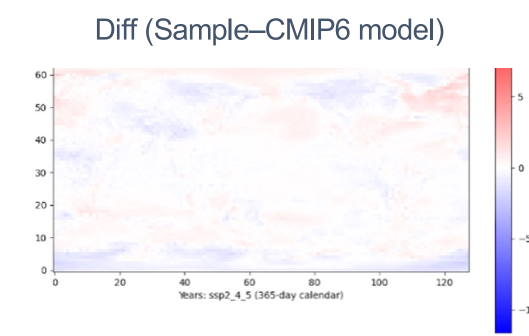
→ **Diffusion model** →

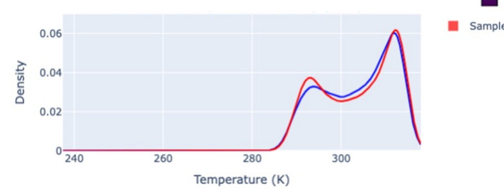
CMIP6 temperature map (T)







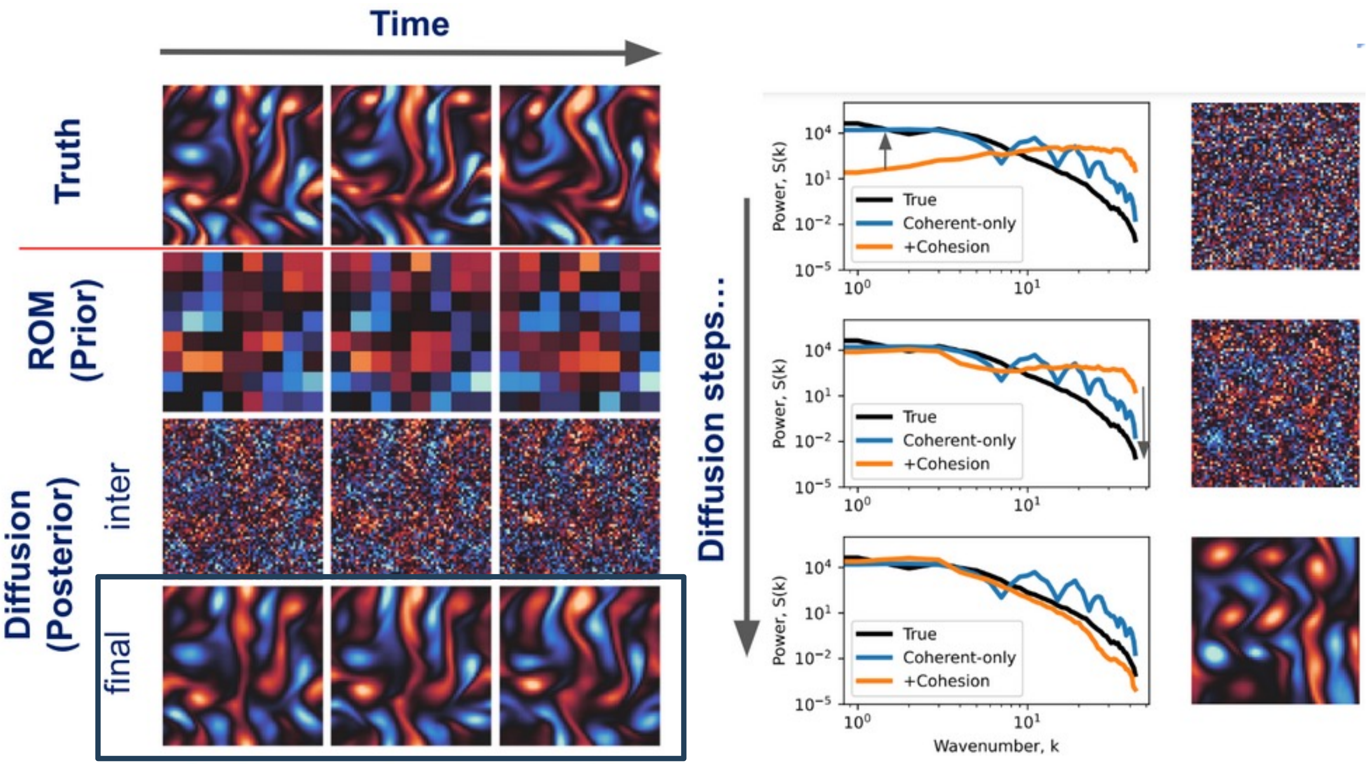




Immorlano et al., *submitted*

Generative AI for simulation, data assimilation and downscaling

Diffusion model to improve accuracy of complex multi-scale stochastic physics

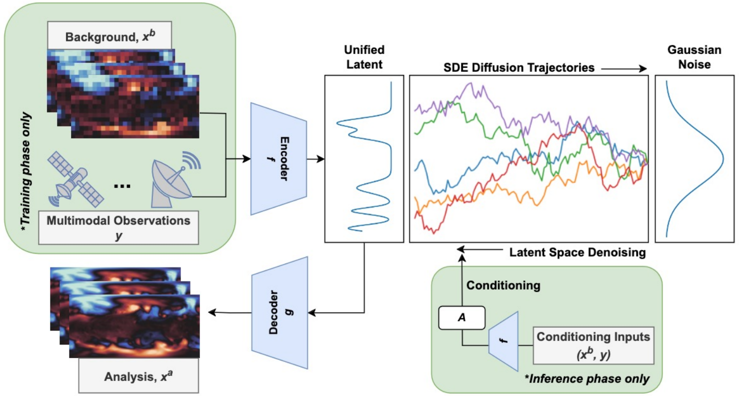


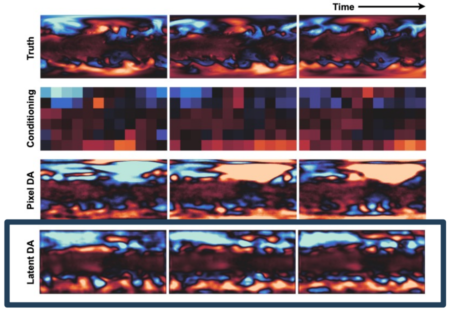
Non blurry!

Generative AI for simulation, data assimilation and downscaling

Diffusion models (Dall-E) can be used for many applications and embedded in hybrid models

Data Assimilation



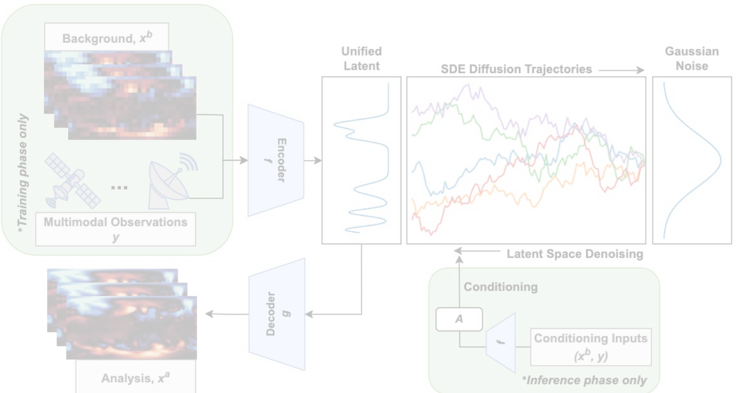


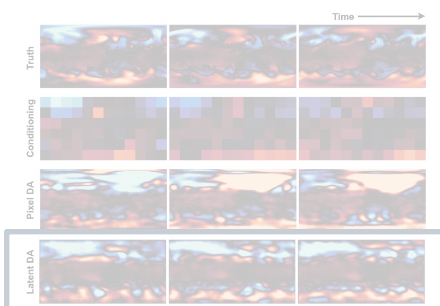
Nathaniel, Qu, Li and Gentine CVPR Best Paper 2024; Mardani et al., 2024; Lockwood et al. 2024 *JGR Atmos;* Fan et al. submitted

Generative AI for simulation, data assimilation and downscaling

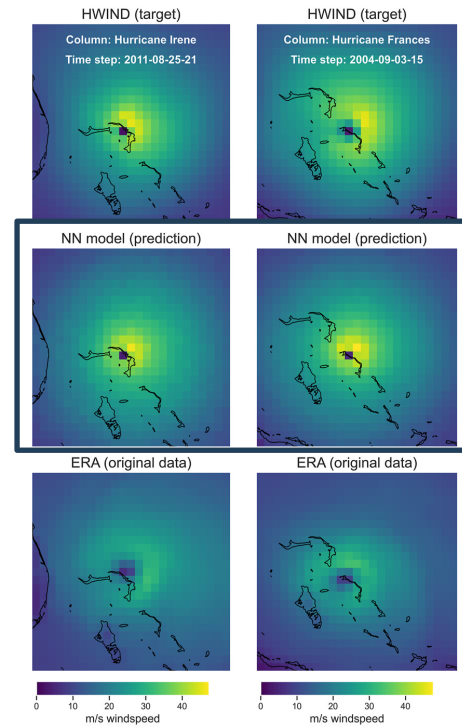
Diffusion models (Dall-E) can be used for many applications and embedded in hybrid models

Data Assimilation





Downscaling



Nathaniel, Qu, Li and Gentine CVPR Best Paper 2024; Mardani et al., 2024; Lockwood et al. 2024 *JGR Atmos;* Fan et al. submitted