Towards a machine learning enhanced version of the Community Earth System Model (CESM3-MLe)

David Lawrence
CESM Chief Scientist
LEAP Model Development Liaison

NSF | NCAR

LEAP

Next-generation Earth System modeling

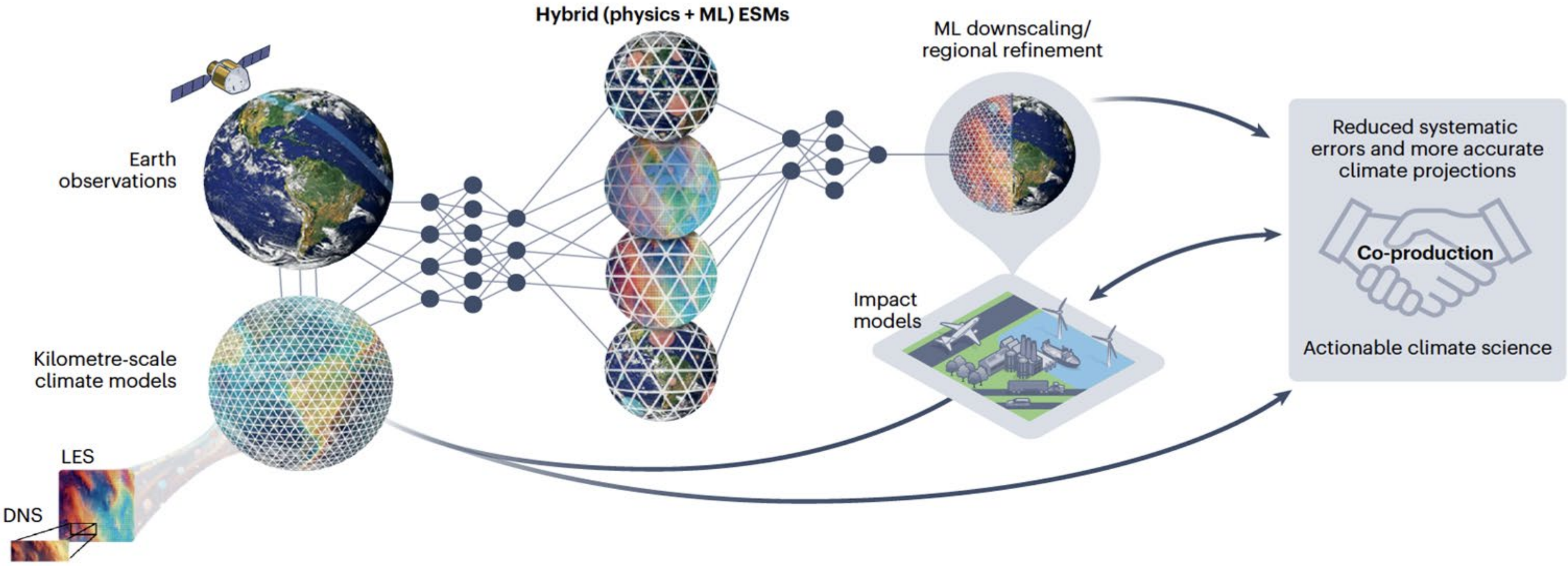
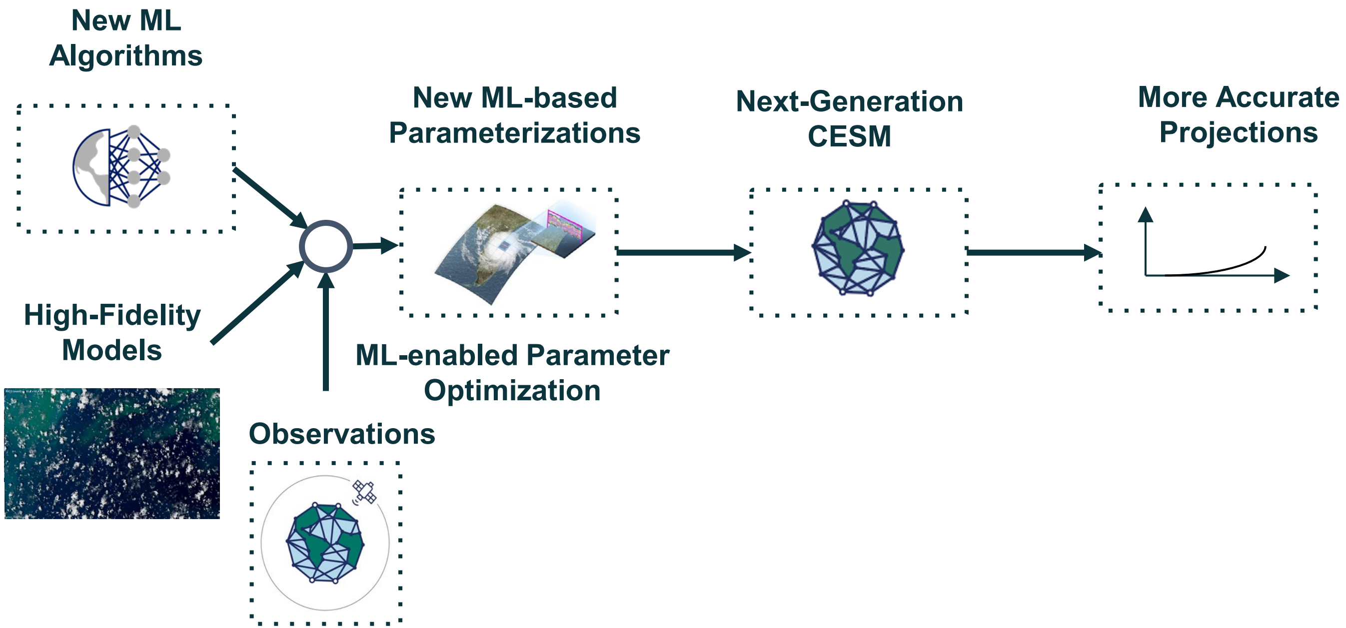


Figure from Eyring, Gentine, Camps-Valls, Lawrence, Reichstein (Nature Climate Change, 2024)

Towards a machine learning enhanced version of CESM (CESM3MLe)

Learning the Earth with Artificial intelligence and Physics
NSF Science and Technology Center

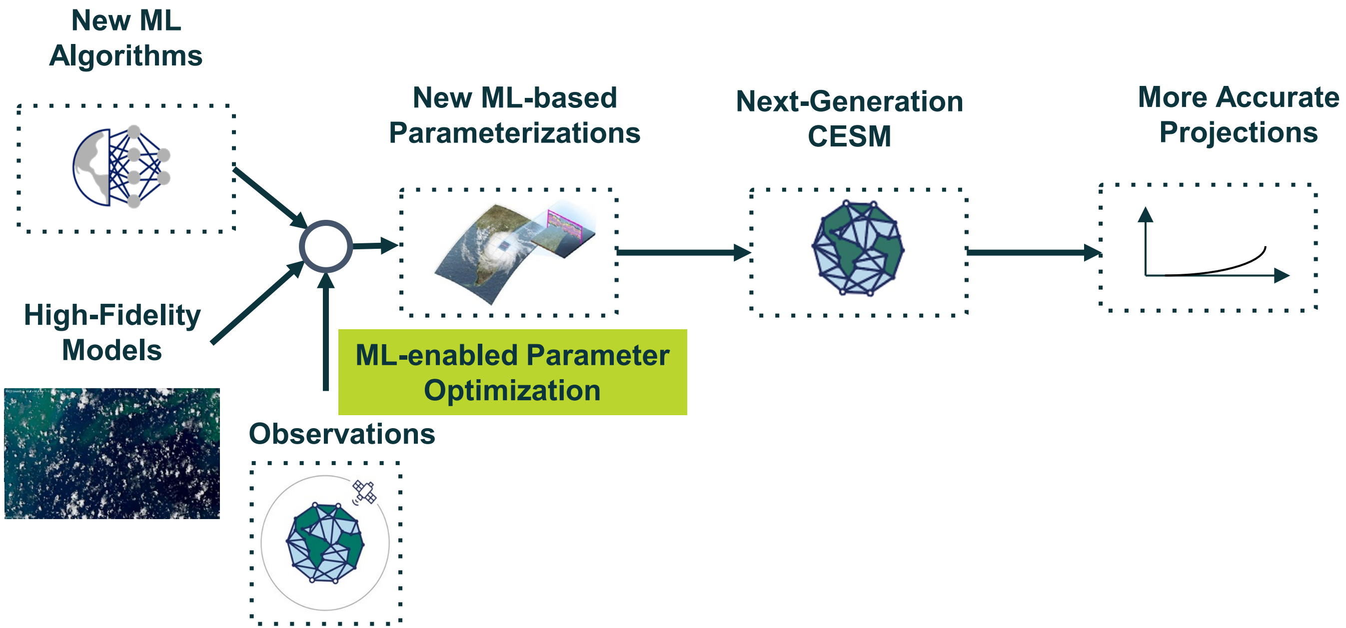
M²LInES
Schmidt Sciences



Towards a machine learning enhanced version of CESM (CESM3MLe)

Learning the Earth with Artificial intelligence and Physics
NSF Science and Technology Center

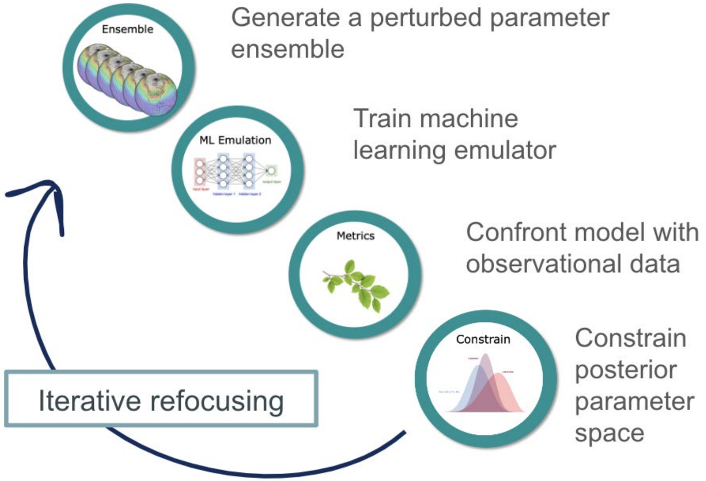
M²LInES
Schmidt Sciences



Towards a machine learning enhanced version of CESM (CESM3MLe)

LEAP Phase 1: Parameter estimation methodologies have been developed for land (CLM) and atmosphere (CAM) model components

CLM



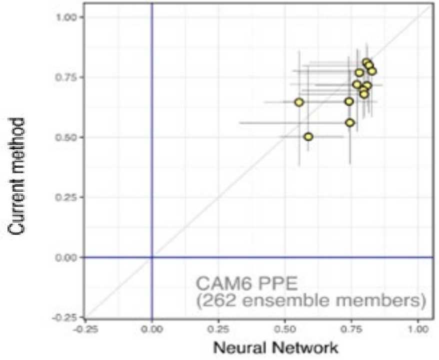
Linnia Hawkins, Daniel Kennedy, Katie Dagon, Dave Lawrence

CAM

Additive Gaussian Process Emulator

Designed for sparse state spaces:

- Additive and simple
- Parameter interaction considered
- Less likely to overfit



Qingyuan Yang, Greg Elsaesser, Brian Medeiros, Addisu Semie, Marcus van Lier-Walqui

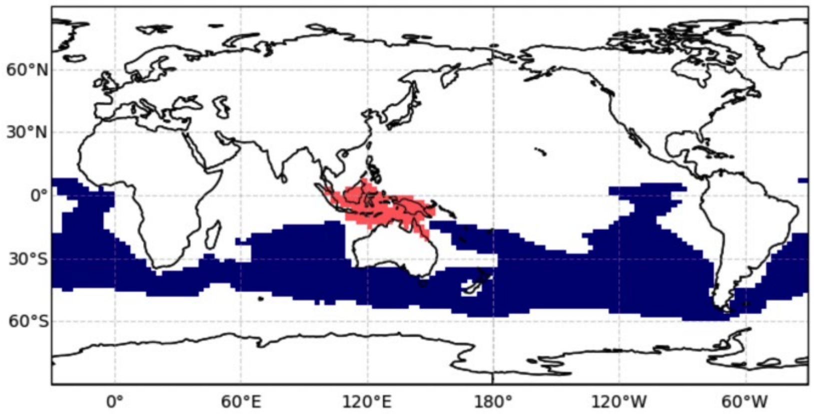
CAM: Additive Gaussian Process Emulator (Yang et al., JAMES, accepted)

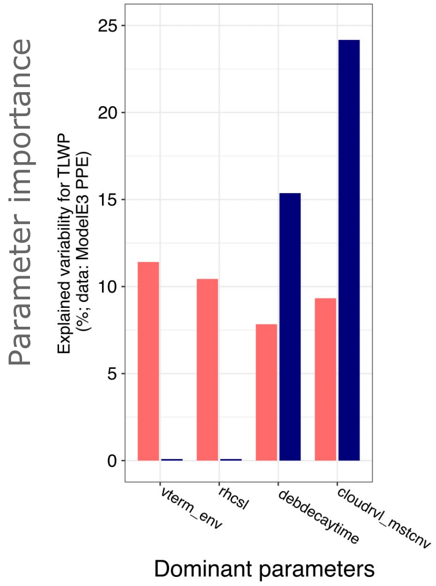
Insight: Individually insensitive parameters can be cumulatively important

Why? Some parameters are only locally/regionally sensitive

Therefore: Emulating only global climatologies may not be sufficient for all problems

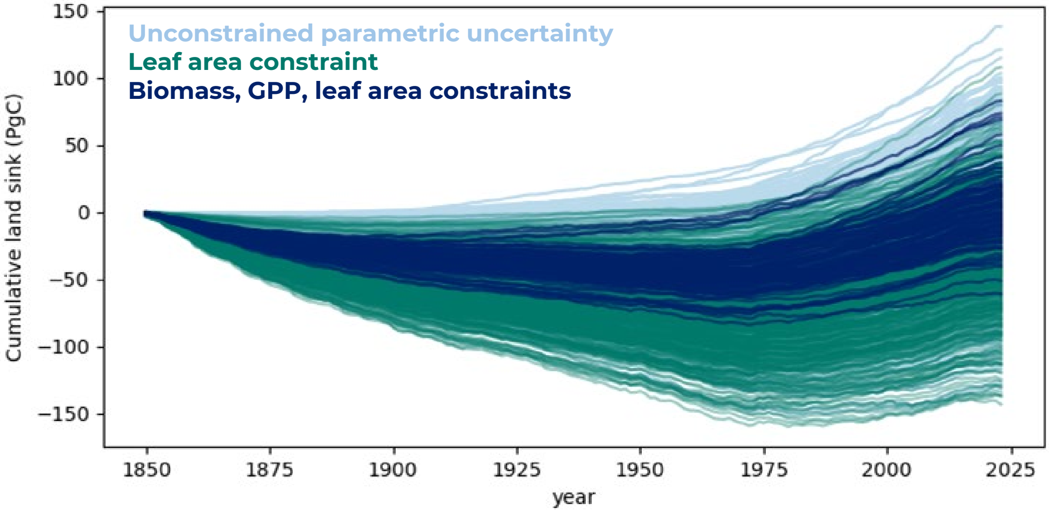
Implication: May be able to both decrease local biases while still calibrating globally

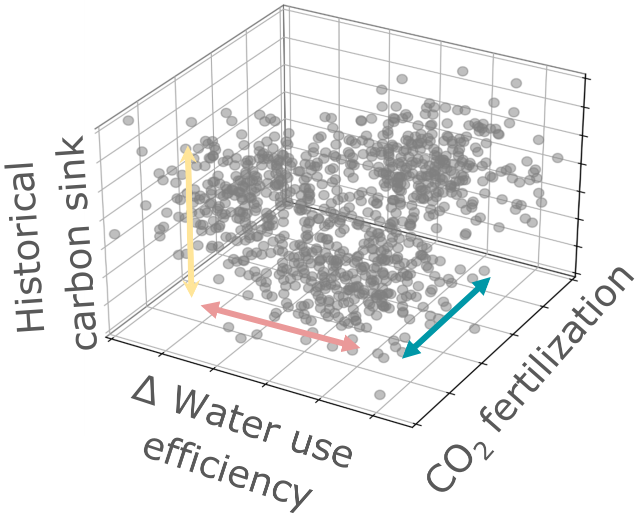




CLM PPE activities

- Calibrate CLM6 for CESM3
- Publish an observationally constrained perturbed parameter ensemble
- Constrain / understand land carbon cycle uncertainty
- For a possible future CESM3 large ensemble (emissions-driven mode), utilize carefully selected parameter sets that span unconstrained emergent behavior in land carbon cycle and ECS or TCR

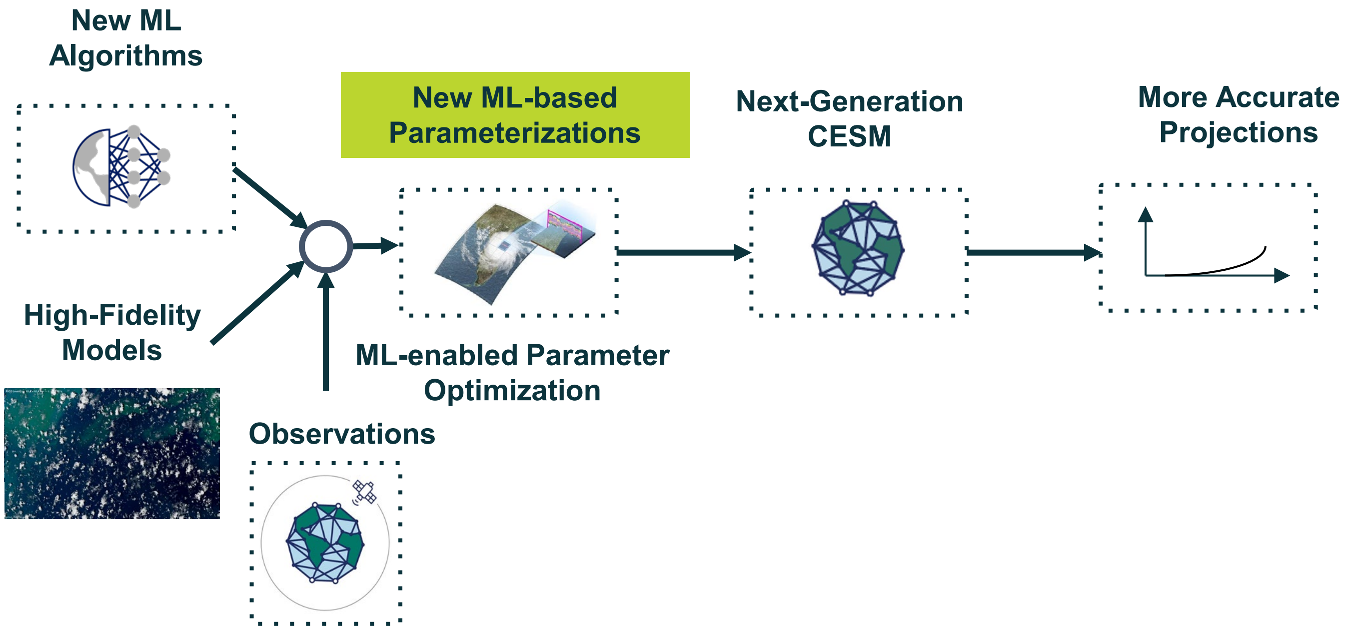




Towards a machine learning enhanced version of CESM (CESM3MLe)

Learning the Earth with Artificial intelligence and Physics
NSF Science and Technology Center

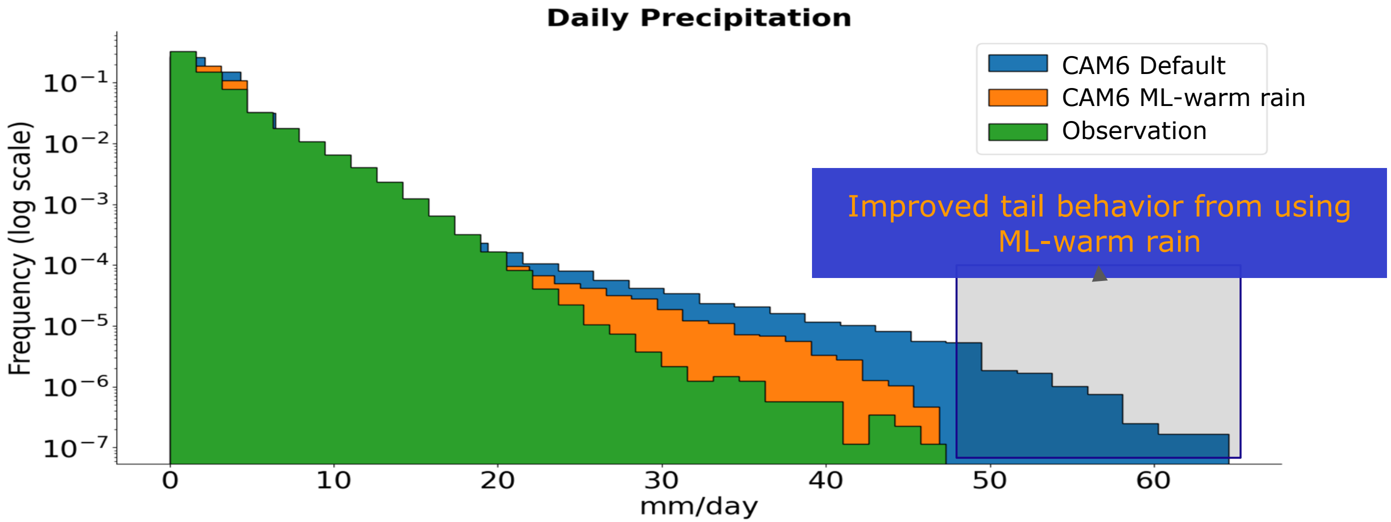
M²LInES
Schmidt Sciences



Towards a machine learning enhanced version of CESM (CESM3MLe)

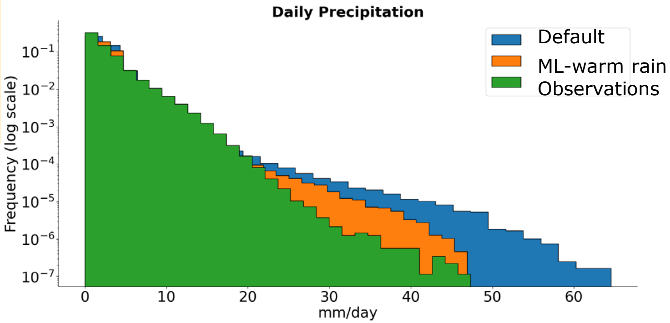
LEAP Phase 1: Demonstrate pathway and impact of ML-based parameterizations in CESM

Warm rain microphysics: Emulate cloud droplet autoconversion and accretion with NNs trained on CAM simulations with warm rain process replaced with highly resolved bin microphysics (TAU code)



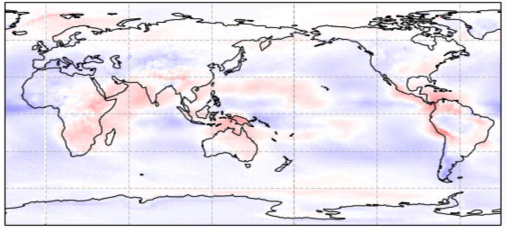
Developing workflow to recalibrate after ML paramaterizations implemented

1. ML warm rain microphysics → improvement in rainfall distribution

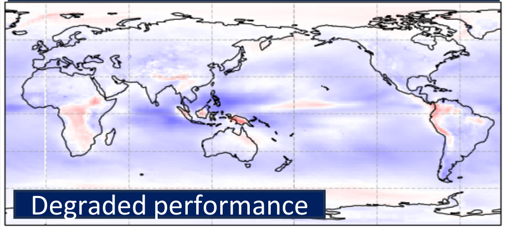


2. But, degraded performance for other fields is likely with new parameterization

CAM default - observation

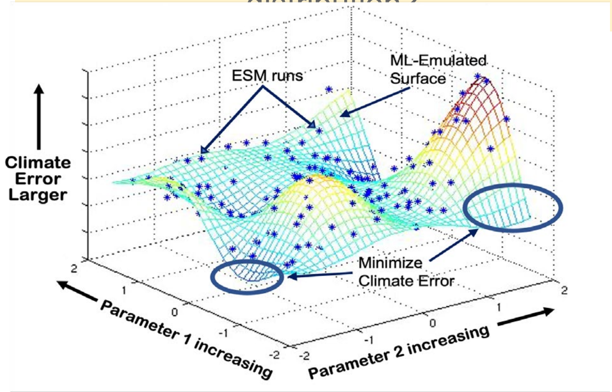


CAM with warm rain ML- observation

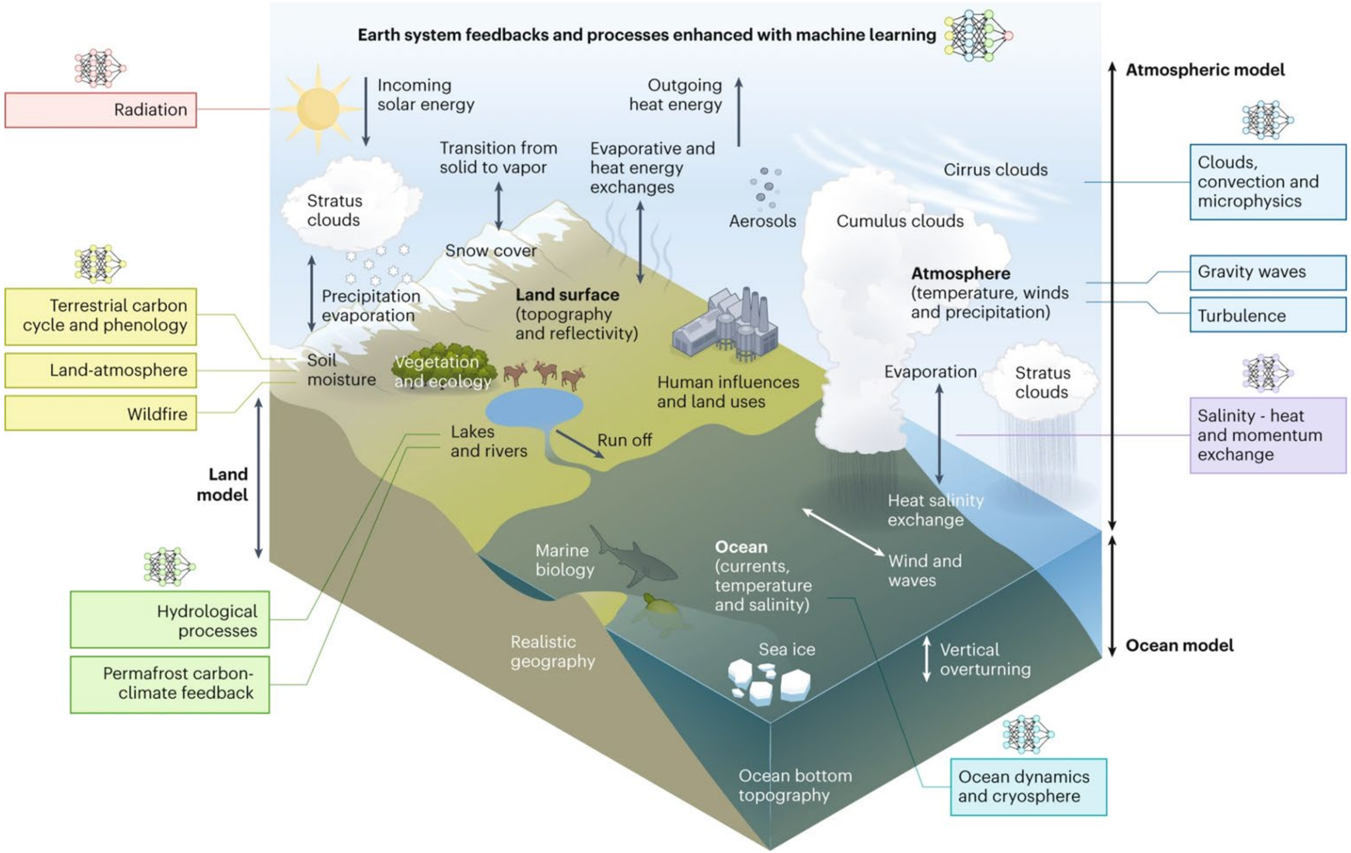


Schematic only; representative of a climatological radiation or cloud field

3. Using ML for auto-tuning, can we re-calibrate CAM to correct the degraded performance, while simultaneously retaining the improvement in rainfall distribution?



Next-generation Earth System modeling to address urgent mitigation and adaptation needs



Candidate parameterizations and tools for CESM3-MLe

Radiation

Phenology

Pedotransfer functions

Fire

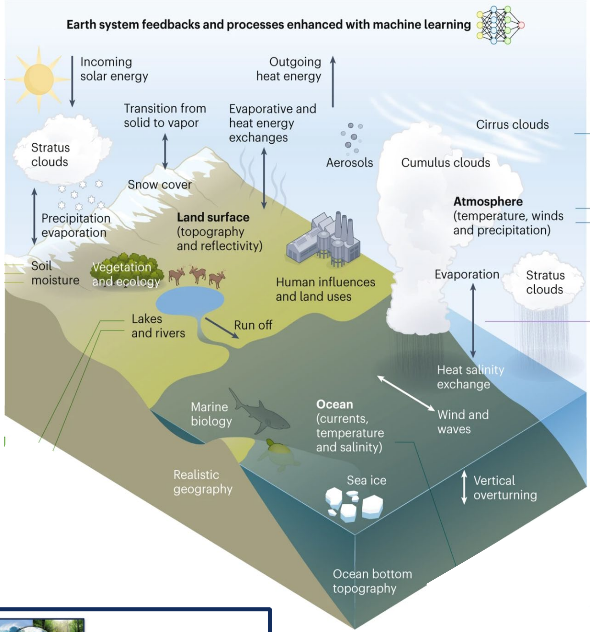
Sea ice heat conduction

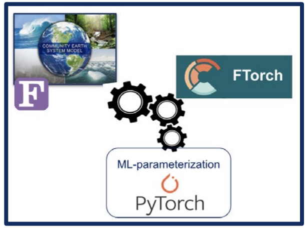
Ice sheet friction

LEAP

M²LInES

ICON, E3SM, lit.





Warm rain microphysics

Ice microphysics

Turbulence

Convection

Cloud cover

Atmospheric Chemistry

Vertical Mixing (KPP)

Ocean Mesoscales

Air-sea turbulent fluxes

Figure modified from Eyring, Gentine, Camps-Valls, Lawrence, Reichstein (Nature Climate Change, 2024)

Candidate parameterizations into CESM-MLe

→ **LEAP is enabling** the platform for flexible adoption and experimentation of **ML parametrization & calibration into CESM** from LEAP, M²LInES, and other projects

CESM-MLe Integration Team

Identified need

More productive and sustained interactions between LEAP and M^2LInES projects and CESM scientists and developers

CESM-MLe Integration Team

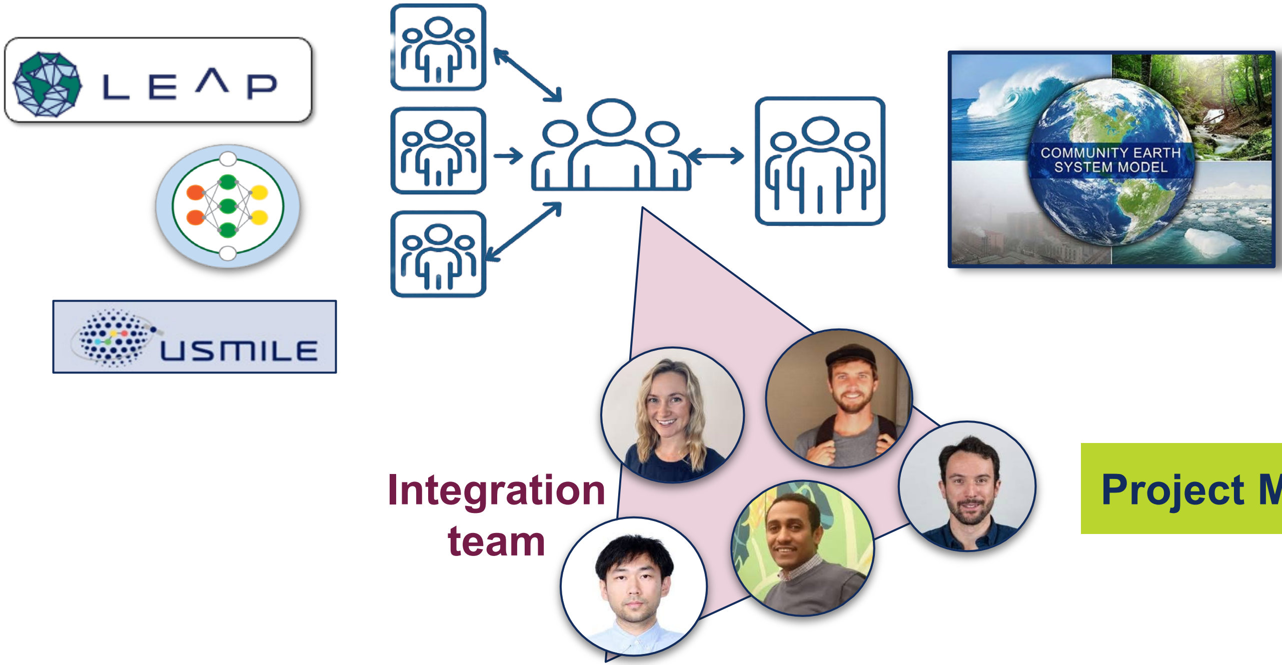
Identified need

More productive and sustained interactions between LEAP and M^{2}LInES projects and CESM scientists and developers

CESM-MLe Integration Team

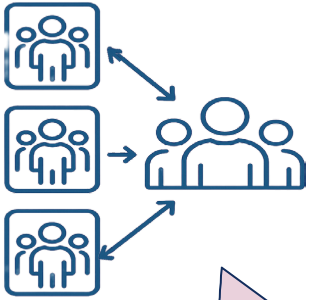
Identified need

More productive and sustained interactions between LEAP and M²LInES projects and CESM scientists and developers



Project Management

CESM-MLe Integration Team

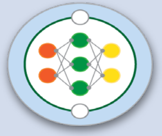


Integration team

CESM Integration Survey (selected questions)

- Which physical process is your ML model targeting?
- How well do your inputs/outputs align with those available in the CESM component model?
- What dataset(s) are being used for training?
- How are you currently testing your ML parameterization?
- How will you test in a CESM-relevant context?
- What is the planned mechanism for integrating your ML model into CESM?
- Are there intermediate calculations or dependencies that must be implemented within CESM?
- Will the ML parameterization require additional tuning once integrated into CESM?
- What additional support or collaboration would help facilitate your model's integration into CESM?
- Are there any anticipated concerns or risks regarding your ML parameterization's integration into CESM?

Towards a Machine-Learning enhanced CESM (CESM3-MLe)



Defining Success for CESM3-MLe

- Several ML-based parameterizations into CESM (1-2 atm, 1-2 ocn, 1-2 lnd, 1 sea ice/land ice)
- ML parameter calibration (lnd, atm)
- Reduced biases in critical fields, especially extremes

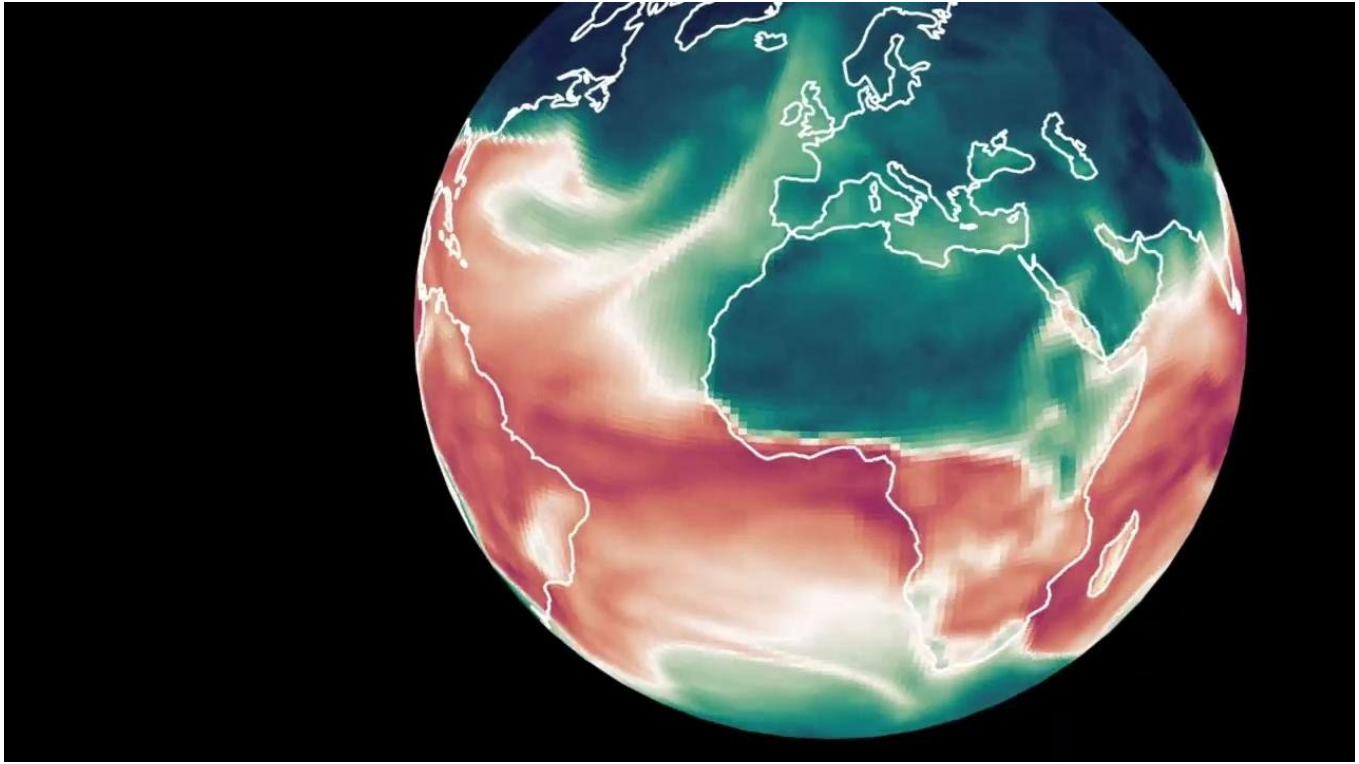
Push forward over next 18-24 months to test the hypothesis that ML can help build better and more accurate ESMs

- Sustained team interactions (e.g., PI, ML-param developer, experienced CESM developer, and SE)
- More coordination / communication (github CESM-MLe project management, development meetings)

Anticipate that there will be challenges

- ML parameterizations in out-of-training climates
- CESM model instabilities with new ML
- Unanticipated interdependencies
- Substantially new simulated climate that may degrade orthogonal aspects of simulation
- New tuning challenges with some tuning knobs removed

CAMulator (CREDIT):

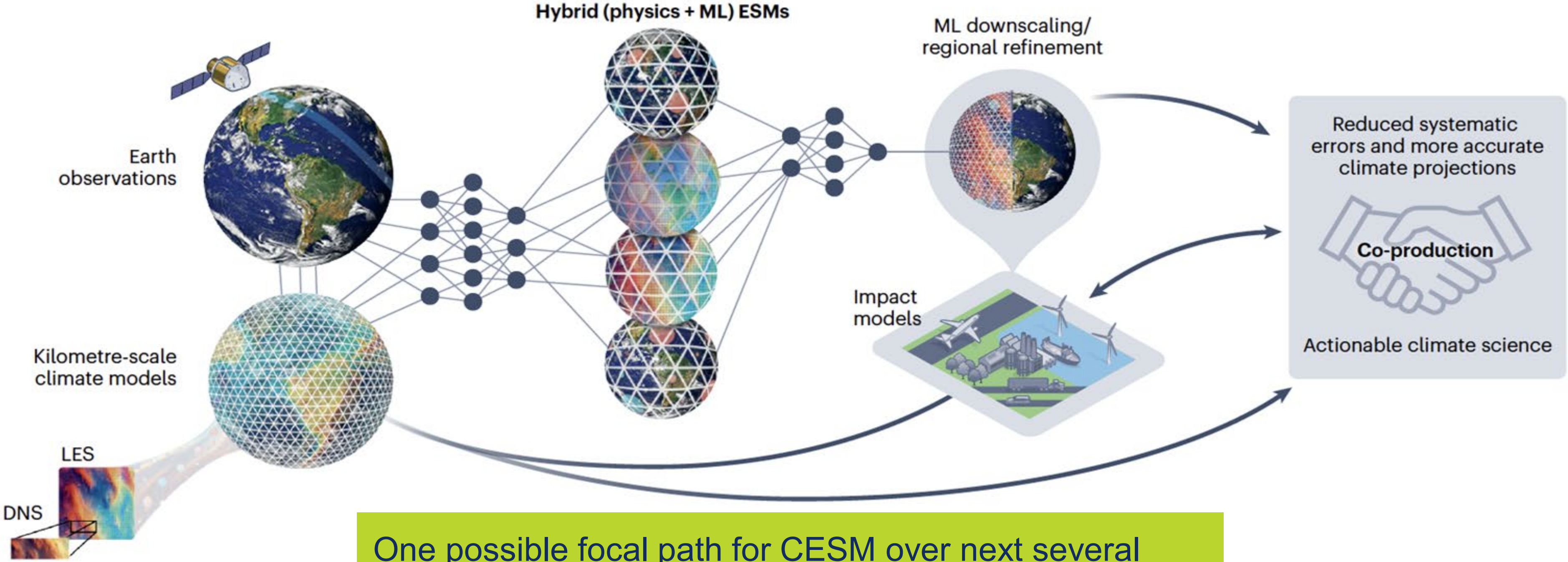


Looking Forward with CREDIT / CESMulator

- Coupling to a dynamic ocean / land (Chapman, Lauritzen)
- Radiative forcing experiments / CESM future climate (Chapman)
- CESMulator S2S initialized forecast system (Mayer, Chapman)
- Emulating CLM (Hawkins)
- Diffusion based post-processing (CISL + CGD)
- Near Real Time forecasts (CISL + CGD)
- Ensemble Capabilities (CISL + CGD)

(An) idea is to include CAMulator along with CREDIT platform within the CESM3 release as a supported research tool

Next-generation Earth System modeling



One possible focal path for CESM over next several years is to push towards a CESM3-HR-MLe configuration along with a CESMulator of it to generate a novel and impactful next generation LE