From weather skill to emergent climate variability

Hannah Christensen

With thanks to

Bobby Antonio[1], Kristian Strommen[1], Pablo Ortega[2]

1. University of Oxford, 2. Barcelona Supercomputing Centre

IMSI Unpacking AI Weather Emulators, Wed 30 Sep 2026

Summary

To rigorously evaluate atmospheric emulators for *climate*, we must couple them to other Earth system components

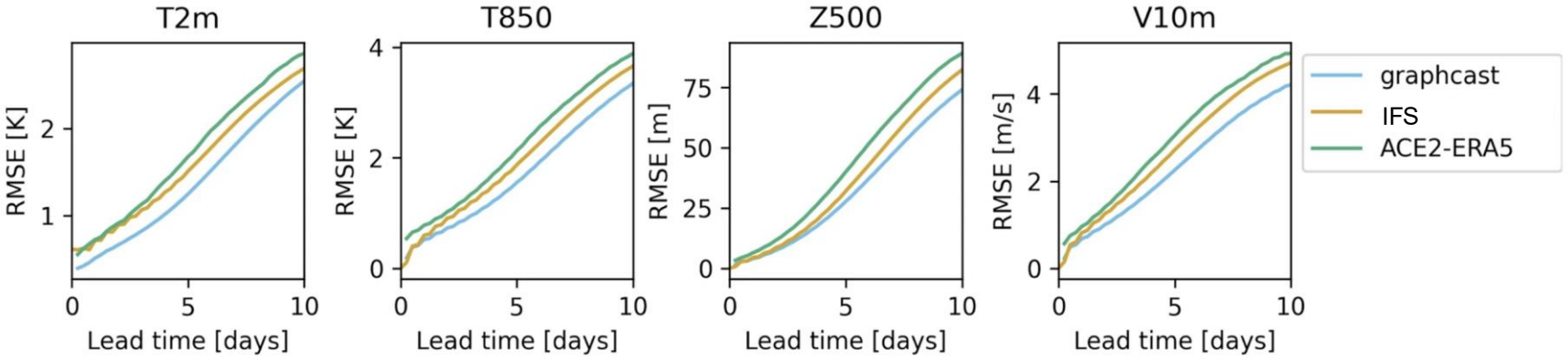
- ACE2 atmospheric emulator: from MLWP to climate
- AMIP style atmosphere only tests are too easy
- Coupling the ACE2 atmospheric emulator to the NEMO ocean model
- The good news
- The not-so-good news
- Outlook

ML atmosphere ACE2

ACE2: Developed at Ai2
Watt-Meyer et al, 2025

From weather …

- trained on ERA5, 6 hour timestep
- 1° resolution, eight terrain-following vertical layers
- SFNO architecture



ACE2 weather forecast skill – comparable to dynamical IFS and leading MLWP models

ML atmosphere ACE2

ACE2: Developed at Ai2
Watt-Meyer et al, 2025

From weather …
- trained on ERA5, 6 hour timestep
- 1° resolution, eight terrain-following vertical layers
- Spherical Fourier neural operator architecture

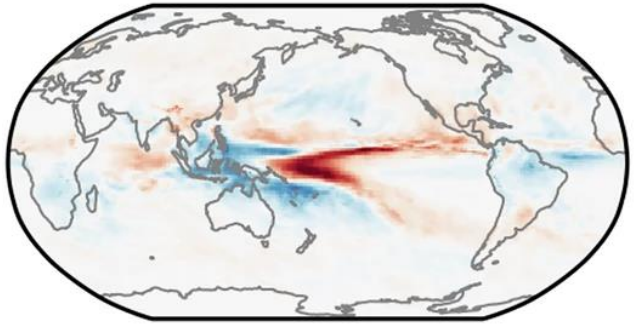
… to climate
- can stably integrate forward an arbitrary number of 6-h time steps
- take user-specified SST and CO_2 as forcing variables
- conserves mass and moisture

ACE2 is a climate emulator with high weather-prediction skill
> consider links between weather and **climate variability**

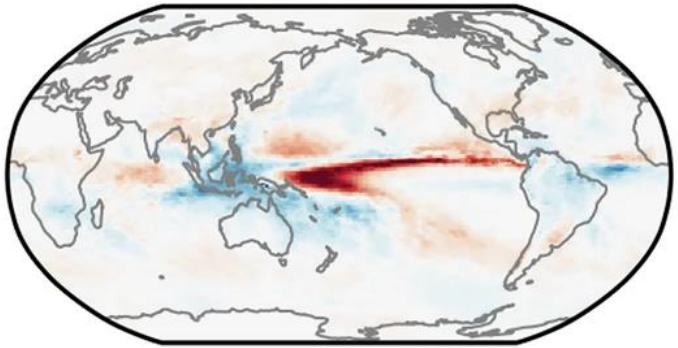
ML atmosphere ACE2

“[ACE2] accurately reproduces …
the atmospheric response to El Niño variability.”

ACE2



ERA5



Maps of regression coefficients of ACE2 and ERA5 surface precipitation against the Niño 3.4 index

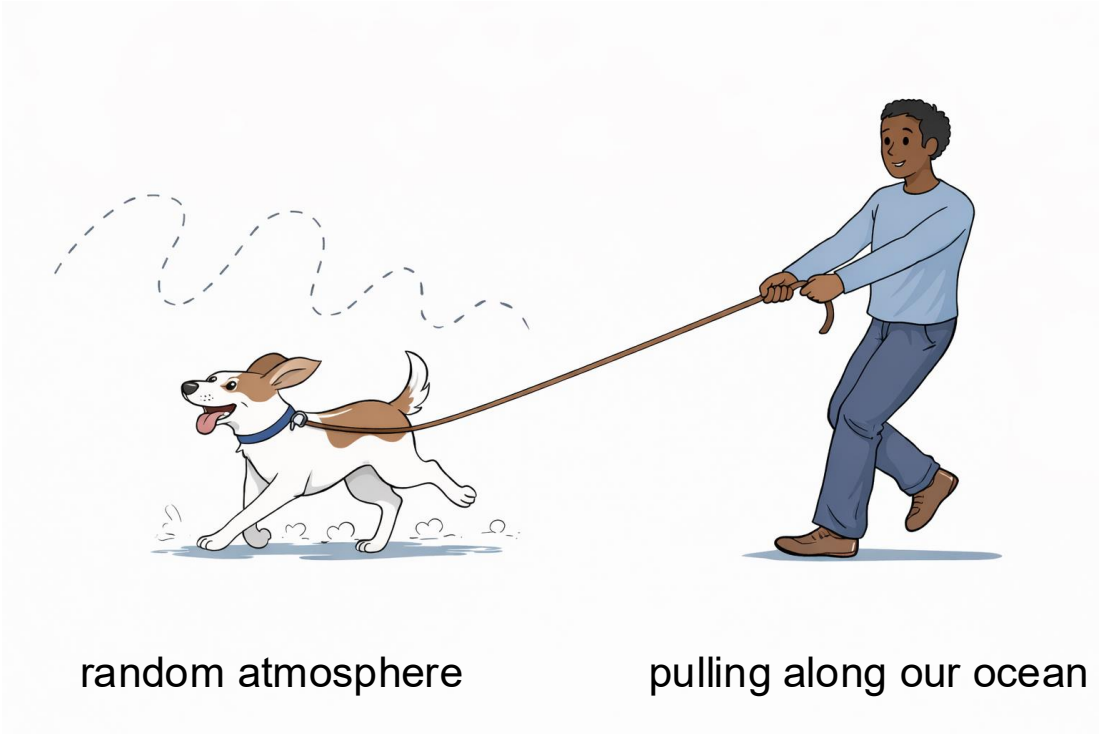
Watt-Meyer et al, 2025

But can we really say ACE2 accurately captures atmospheric variability?

- These are atmosphere only AMIP type experiments

- ACE2 supplied with observed ERA5 SST and CO_2 concentrations
- This cannot assess **feedbacks** between atmosphere and ocean (and land, ice, …) known to be **central for emergent climate variability**

A possible analogy: if this is the coupled system…



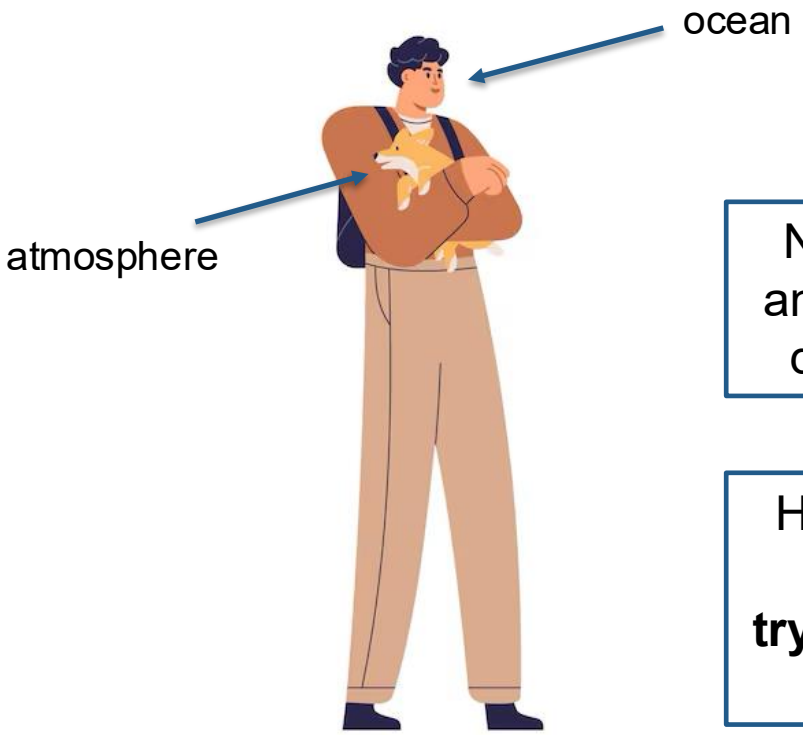
random atmosphere

pulling along our ocean …

… which is guiding the atmosphere in return

(Hasselmann:
slow climate variability
as integrated response
to continuous excitation
by short period weather
disturbances)

… then prescribing SST is like this



Not allowing our fast and slow processes to dynamically interact

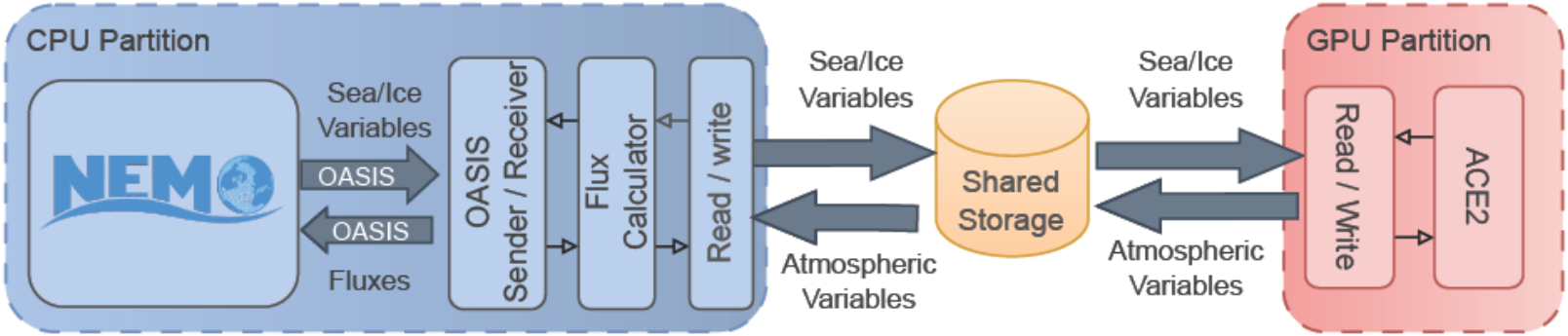
How can we tell what the atmosphere is **trying to do** if we carry it along?

Better test? Couple ACE2 to dynamical ocean

- **Dynamical ocean model:** NEMO.
  - Based on physics.
- **Atmospheric emulator:** ACE2-ERA5.
  - No retraining. Clean test of emulator
- Challenges to overcome:
  - NEMO is **CPU-intensive**, whilst ACE2 runs on **GPUs**.
  - ACE2 **does not produce all of the fluxes** that NEMO requires.
  - ACE2 has a **6hr model time step**, compared to 40mins for NEMO ORCA1.

➢ Implement a novel coupling using OASIS, shared storage, and the AirSeaFluxCode module.

Antonio, Strommen, Ortega & **HM Christensen**, submitted



ACE2-NEMO: mean state

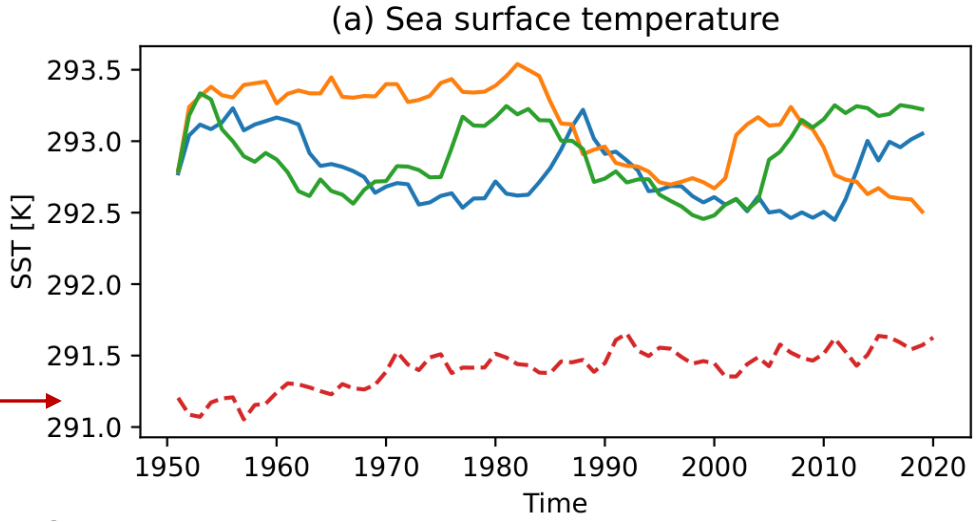
70 year run with **fixed 1950s forcings** (ACE2-NEMO-control)

☑ Stable simulations out to 100 years+

Antonio, Strommen, Ortega &
HM Christensen, submitted

ACE2-NEMO
3 ensemble members

reference: EC-Earth3
(IFS-NEMO)



ACE2-NEMO: mean state

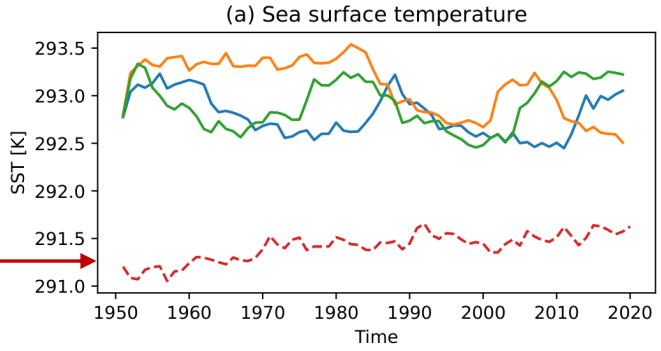
70 year run with **fixed 1950s forcings** (ACE2-NEMO-control)

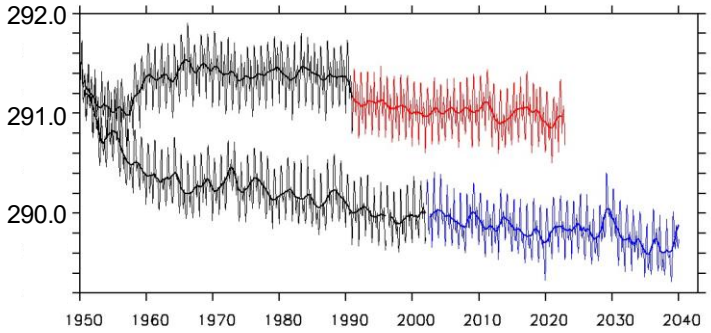
☑ Stable simulations out to 100 years+

Antonio, Strommen, Ortega & **HM Christensen**, submitted

ACE2-NEMO
3 ensemble members

reference: EC-Earth3
(IFS-NEMO)





~1.5 K difference … is this OK?

It's OK

e.g. **red** and **blue** lines are from **two versions** of the **same dynamical model** EERIE ICON ocean-atmosphere @ 10km

ACE2-NEMO: mean state

70 year run with **fixed 1950s forcings** (ACE2-NEMO-control)

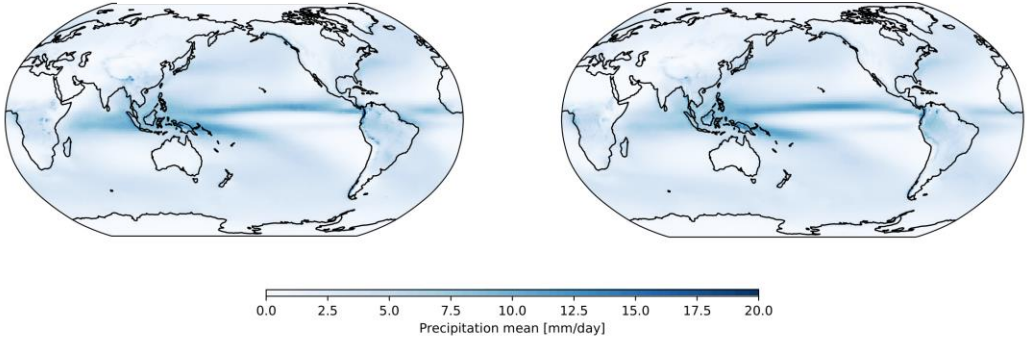
☑ Realistic mean state compared to physical model

Antonio, Strommen, Ortega & **HM Christensen**, submitted

precipitation

ACE2-NEMO

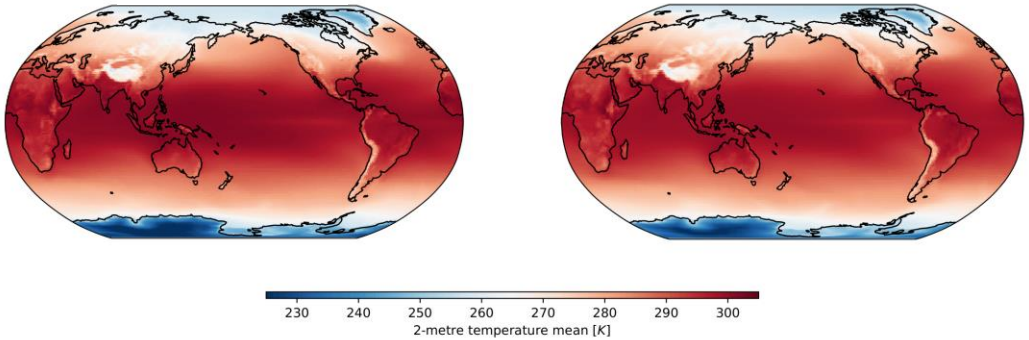
reference: EC-Earth3



2m temperature

ACE2-NEMO

reference: EC-Earth3

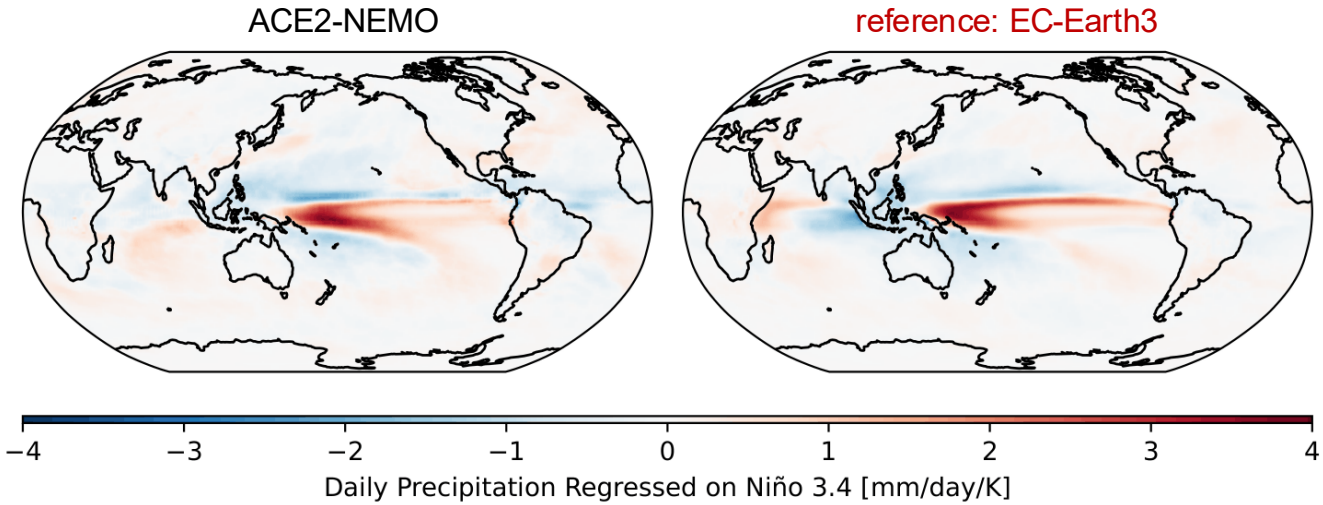


Significant as this is a realistic generalisation test:

NEMO SSTs do not look like ERA5 (!)

Coupled Variability: El Nino

70 year run with **fixed 1950s forcings** (ACE2-NEMO-control)

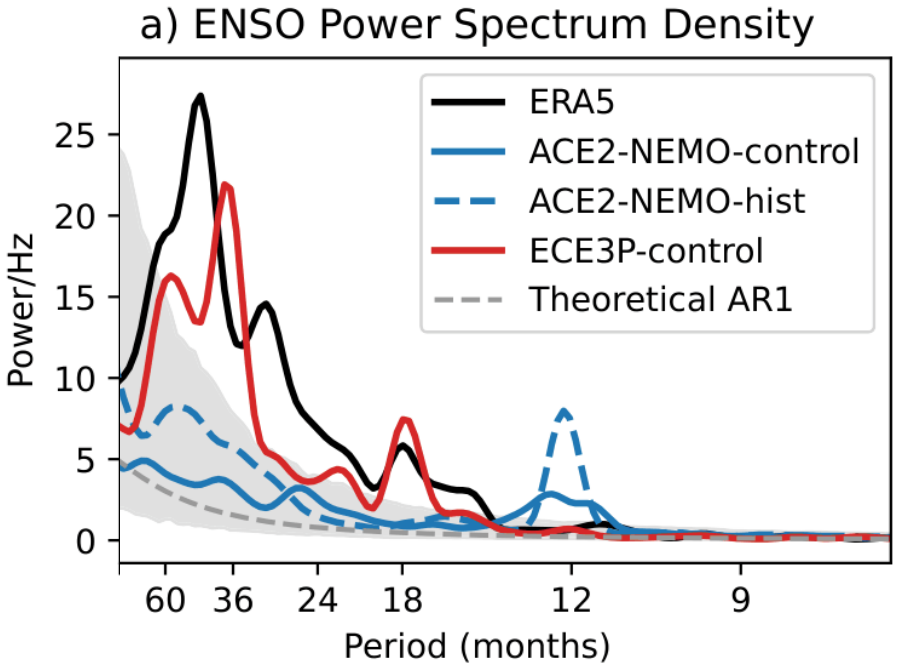


☑ Regression of precipitation onto Niño 3.4 looks reasonable (cf. Watt-Meyer et al, 2025)

Antonio, Strommen, Ortega & **HM Christensen**, submitted

Coupled Variability: El Nino

70 year run with **fixed 1950s forcings** (ACE2-NEMO-control)



⚠ But no emergent, strong, El Nino-like variability observed in power spectrum

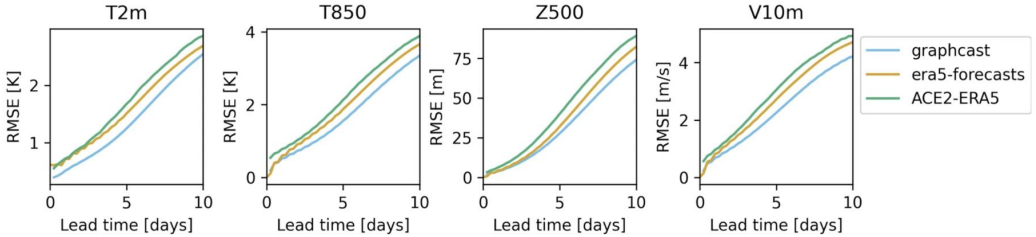
Antonio, Strommen, Ortega & **HM Christensen**, submitted

Weather skill not translating to climate. Why?

After all – climate is the integrated response to the weather

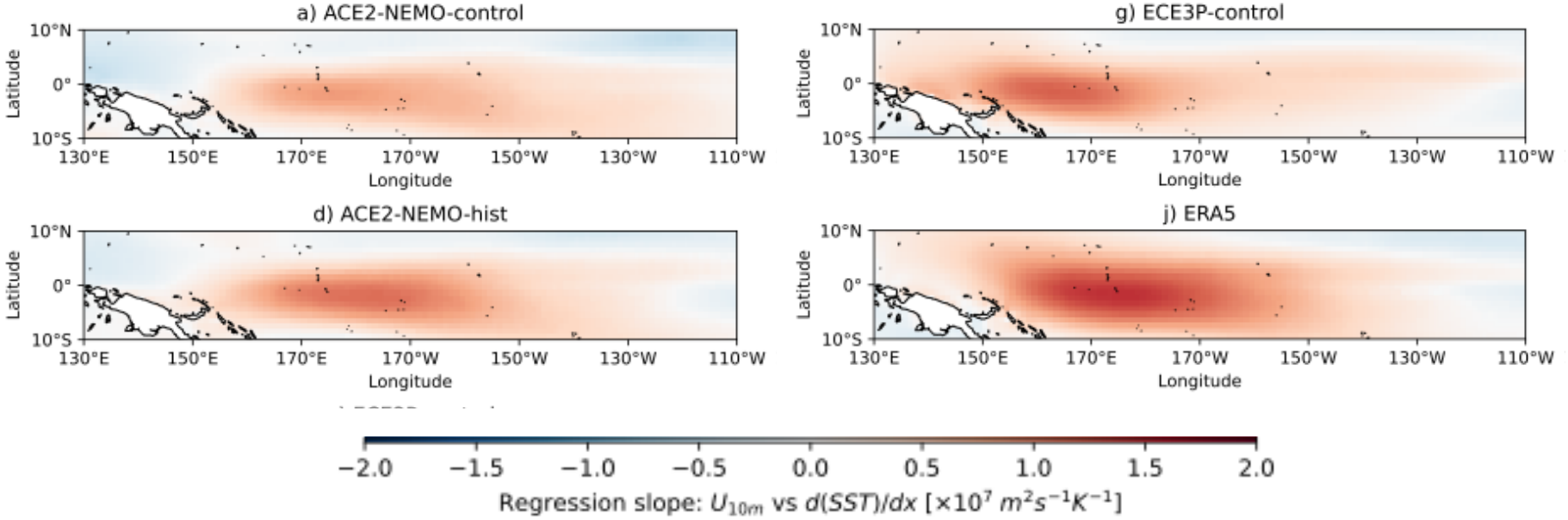


And we know we have good weather skill in ACE2



Weather skill not translating to climate. Why?

- **Too weak feedbacks?** e.g. monthly surface wind / SST gradient feedback
- Differences between ctrl and hist simulations indicate variable conflation?

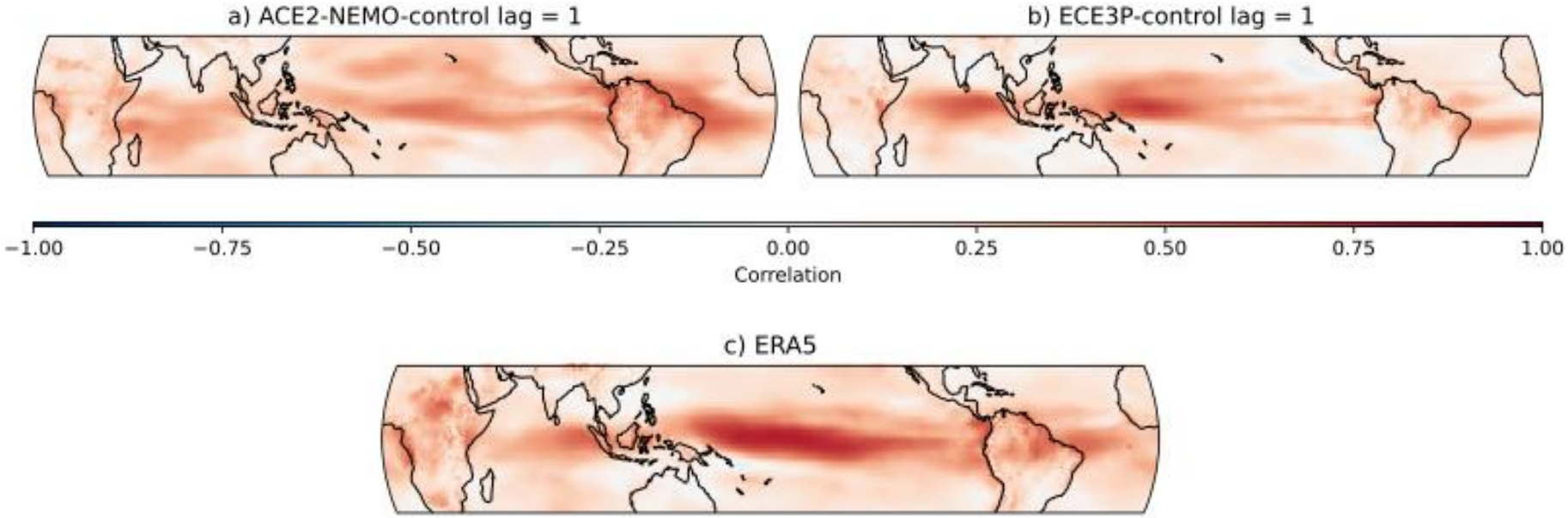


(monthly averaged fields)

Weather skill not translating to climate. Why?

- Hasselmann model: Oceanic response goes like the integrated effect of the fast atmospheric ‘noise’ **over all lags**
    - E.g. if the atmospheric forcing decorrelates quickly, its impact will cancel, while lower amplitude but correlated forcing will have a bigger impact

Too weak autocorrelation?



- 1-month lagged autocorrelation of monthly averaged U-10m

Summary

- It is possible to couple an ML atmosphere to a full-depth dynamical ocean (NEMO)

- ACE2 ML atmosphere can generate a realistic and stable ocean mean state
- But we expose problems with ACE2 that have not been observed before:
    - Weak ENSO feedback mechanisms
    - Atmospheric variability not sufficiently correlated

Funded by the European Union
UK Research and Innovation
Swiss Confederation
Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

Outlook

- Atmospheric emulators have **huge potential** for climate
- Component-wise hybrid ESM **important stepping stone** to fully ML ESM
  - Important test of atmospheric emulators designed for climate applications
  - Physical component brings interpretability
  - Assess ability of current training protocols to constrain emergent behaviour
  - Useful in their own right – e.g. for ocean spin-up problem

Check out the preprint!
B. Antonio et al, 2026

Funded by the European Union
UK Research and Innovation
Swiss Confederation
Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

Funding

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101081383

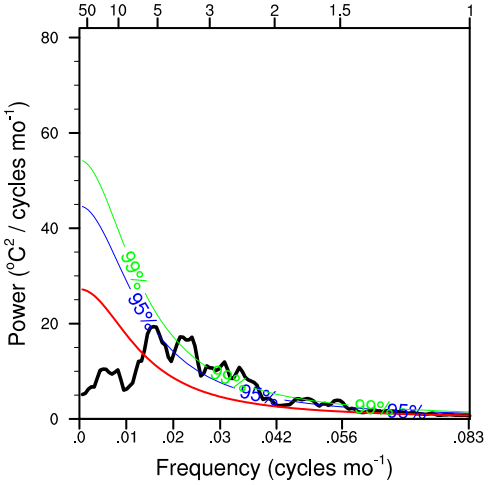
All UK Partners in EERIE are funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant numbers 10057890, 10049639, 10040510, 10040984).

ETH Zürich's contribution to EERIE is funded by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

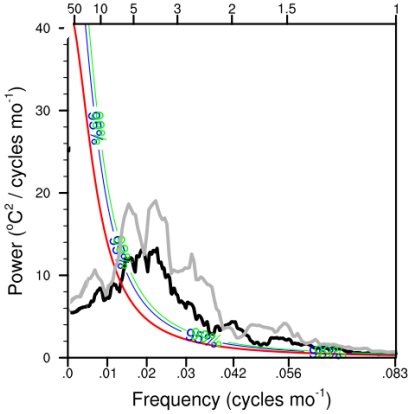


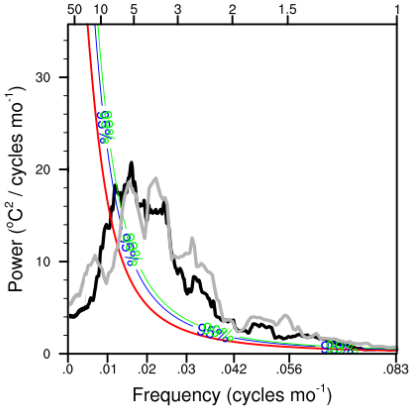
Missing stochasticity?

HadISST



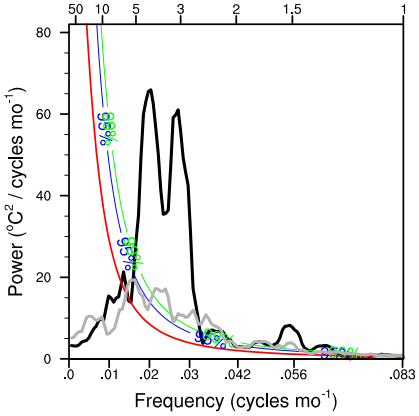
EC-Earth CTRL

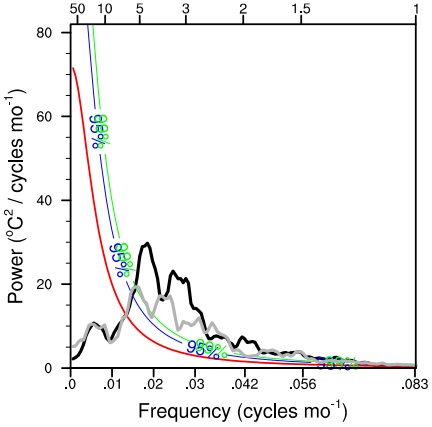




EC-Earth + State-dependent noise

CCSM CTRL

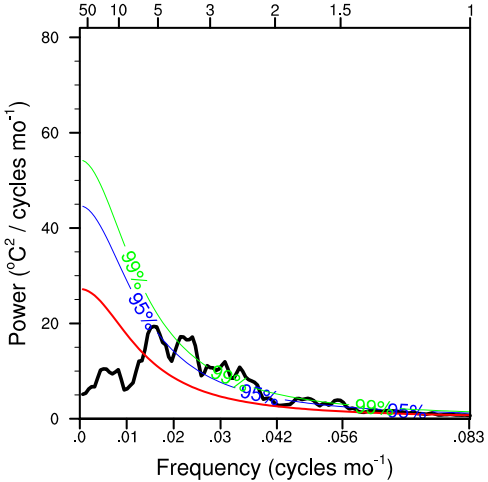




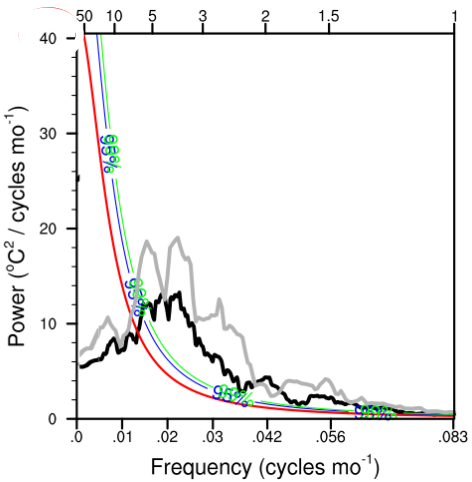
CCSM + State-dependent noise

Missing stochasticity?

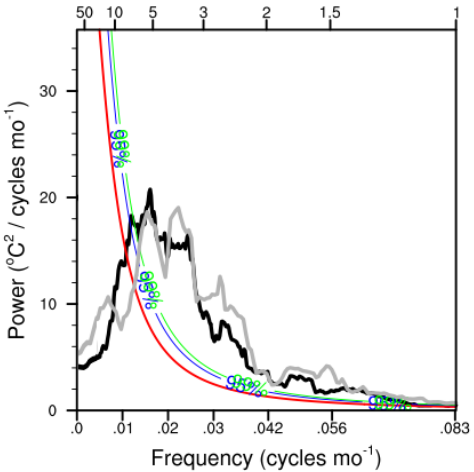
HadISST



EC-Earth CTRL

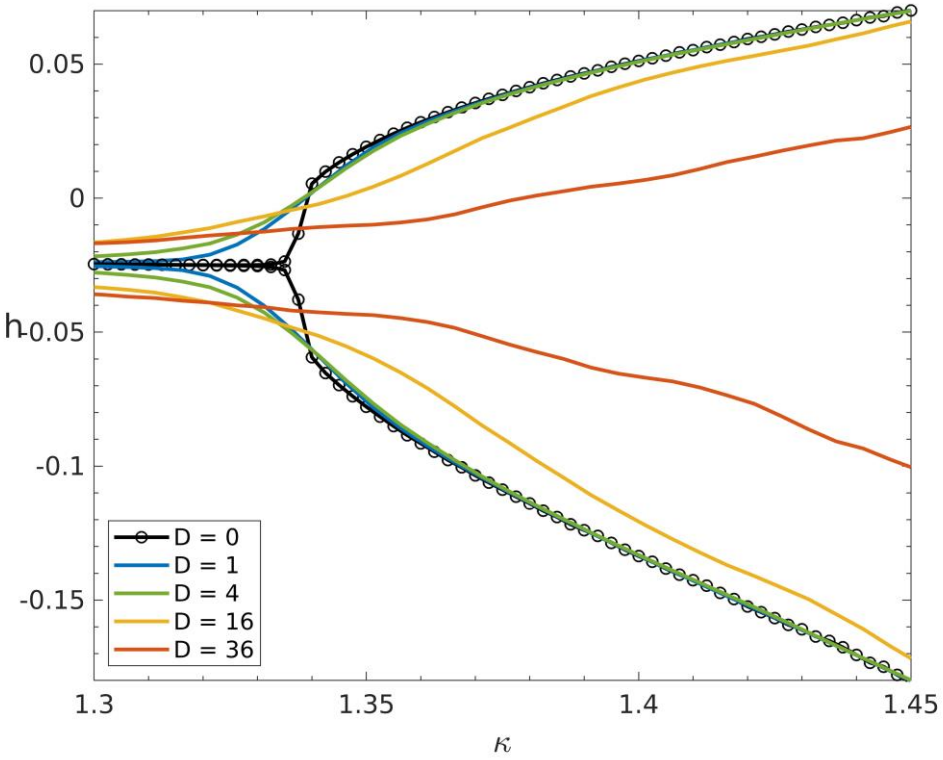


EC-Earth + State-dependent noise



Yang, Christensen et al, 2018

Missing stochasticity?



Simple delayed oscillator model of ENSO

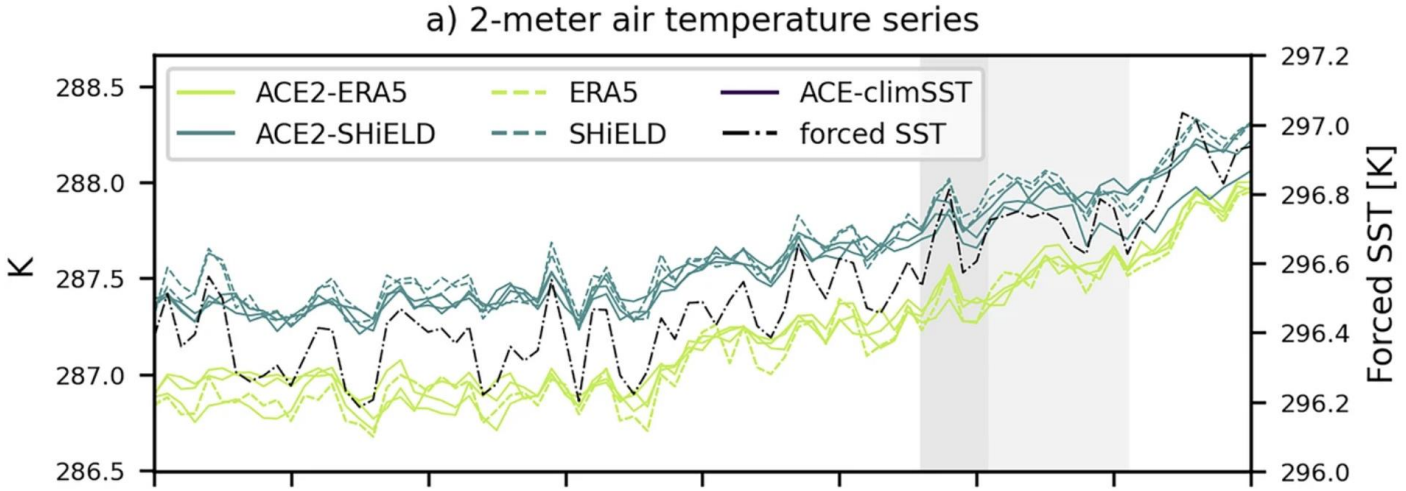
ENSO emerges above some critical coupling value (Hopf bifurcation)

Introduce noise into coupling parameter = state dependent ‘atmospheric noise’

ML atmosphere ACE2

“[ACE2] accurately reproduces …

global trends of temperature over the past 80 years.”

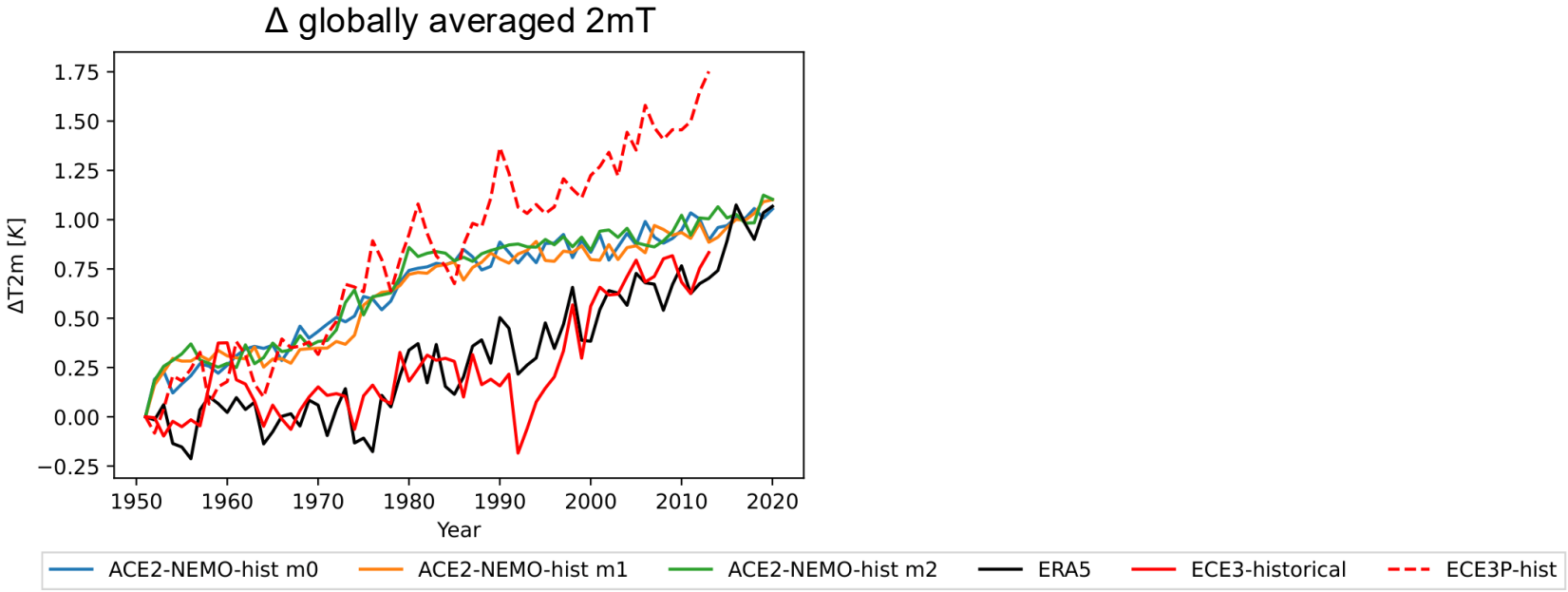


Global and annual-mean timeseries of ACE2 and reference datasets

Watt-Meyer et al, 2025

Forced response

- How does ACE2-NEMO **respond to historical forcing**?

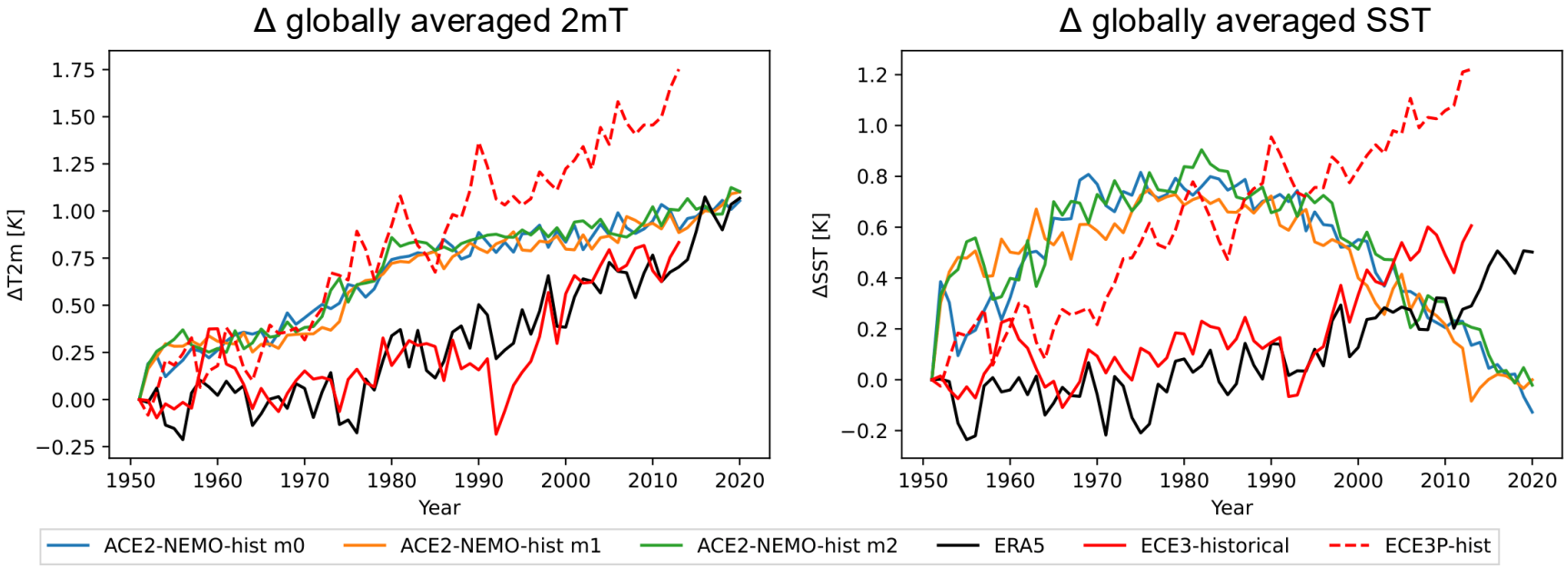


- Initial response similar to EC-Earth 3, but deviates after 1980s, with decrease in SSTs.

Antonio, Strommen, Ortega & **HM Christensen**, submitted

Forced response

- How does ACE2-NEMO **respond to historical forcing**?

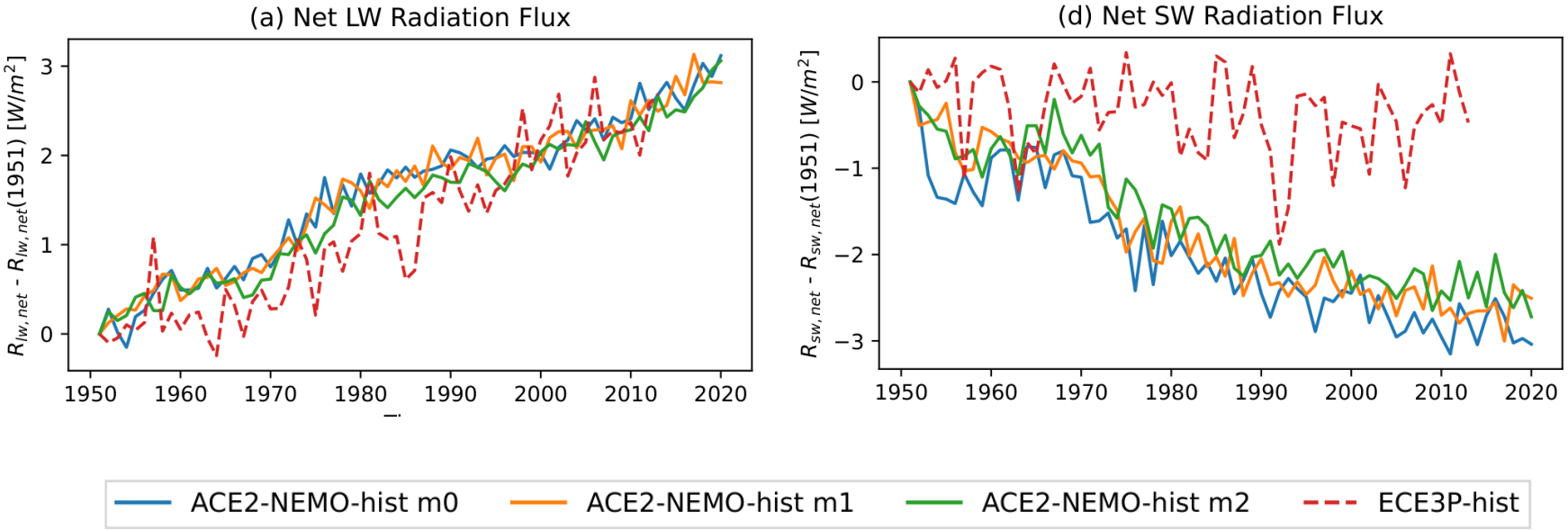


- Initial response similar to EC-Earth 3, but deviates after 1980s, with decrease in SSTs.

Antonio, Strommen, Ortega & **HM Christensen**, submitted

ACE2 historical

- Forces response appears to be due to ACE2 under-predicting downward shortwave radiation.



- Implies issues with surface energy budget

Antonio, Strommen, Ortega & **HM Christensen**, submitted