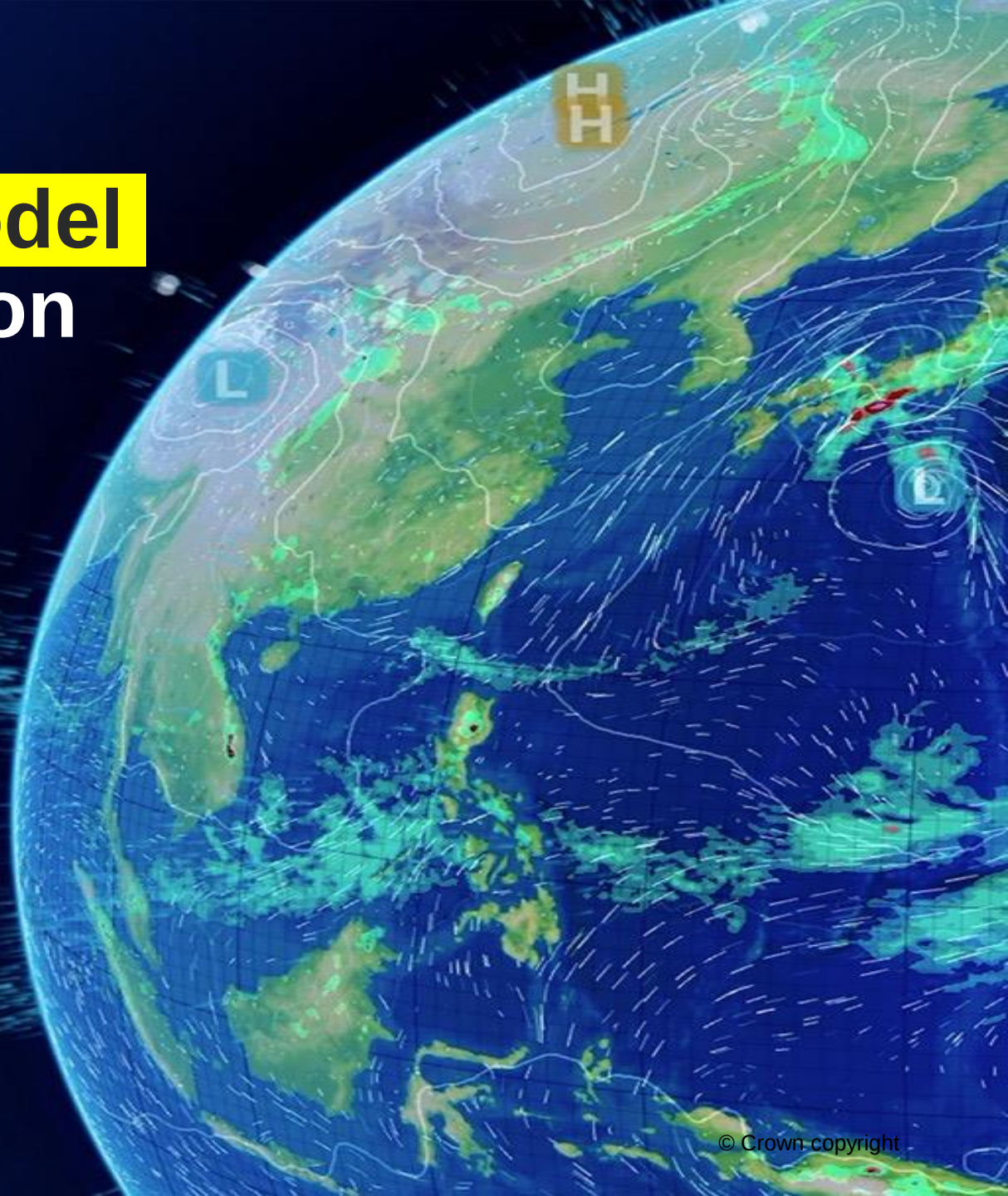
# The need to account for model error in prediction, projection and attribution




Doug Smith

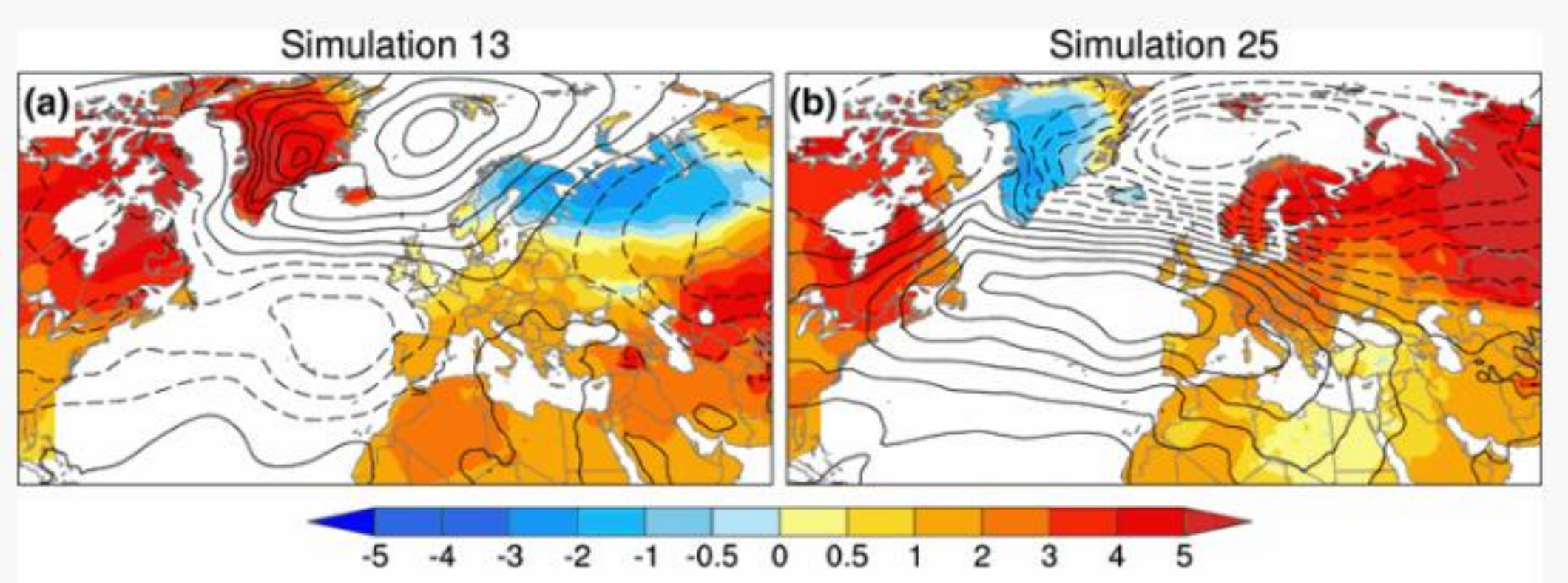
# Key messages

Climate models can have **opposite** responses to the same forcings → **they can't all be right!**

Climate models may **underestimate** the true **forced response**

Not accounting for **model errors** could leave **society unprepared** for impending **extremes**

# Irreducible internal variability?




**Opposite sign of NAO trends for 2016-2045**

**Projections use the same climate model differing by tiny perturbations to initial state**

**Irreducible uncertainty due to unpredictable internal variability (?)**

Deser et al 2017

# Models fail to capture North Atlantic trends

**Trends 1951-2020**

**Obs outside model range (303 members)**

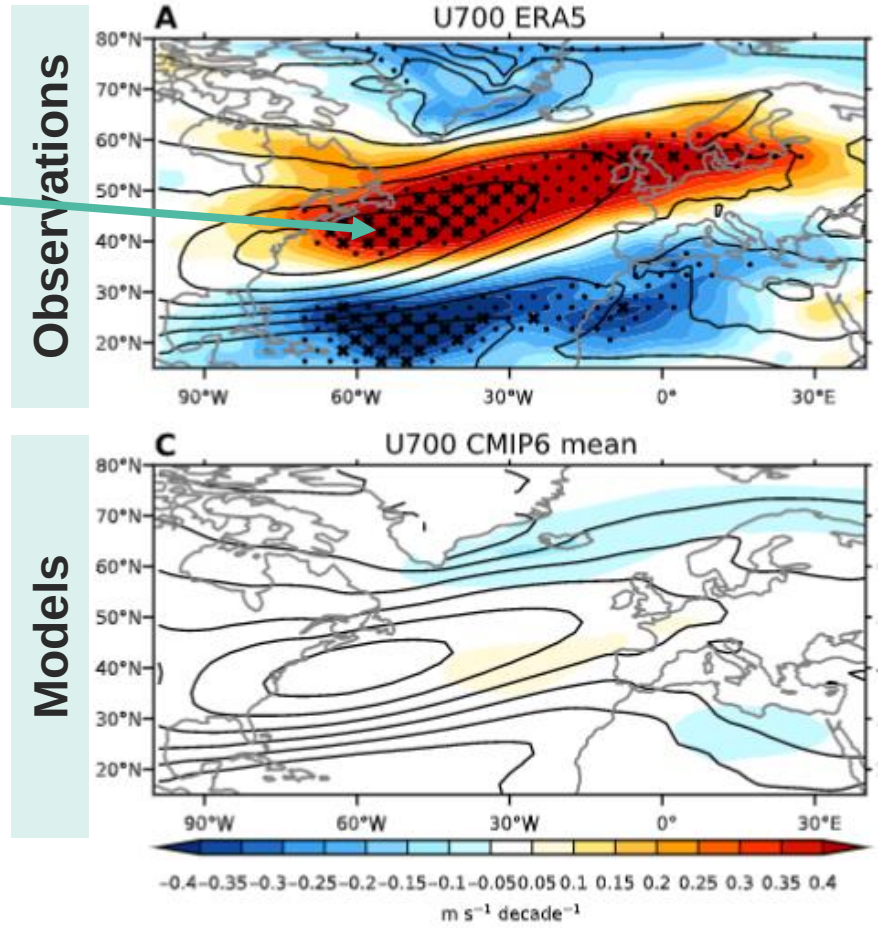
**Similar patterns but obs much stronger**

**Models may have errors!**

**→ possible underestimation of forced response?**

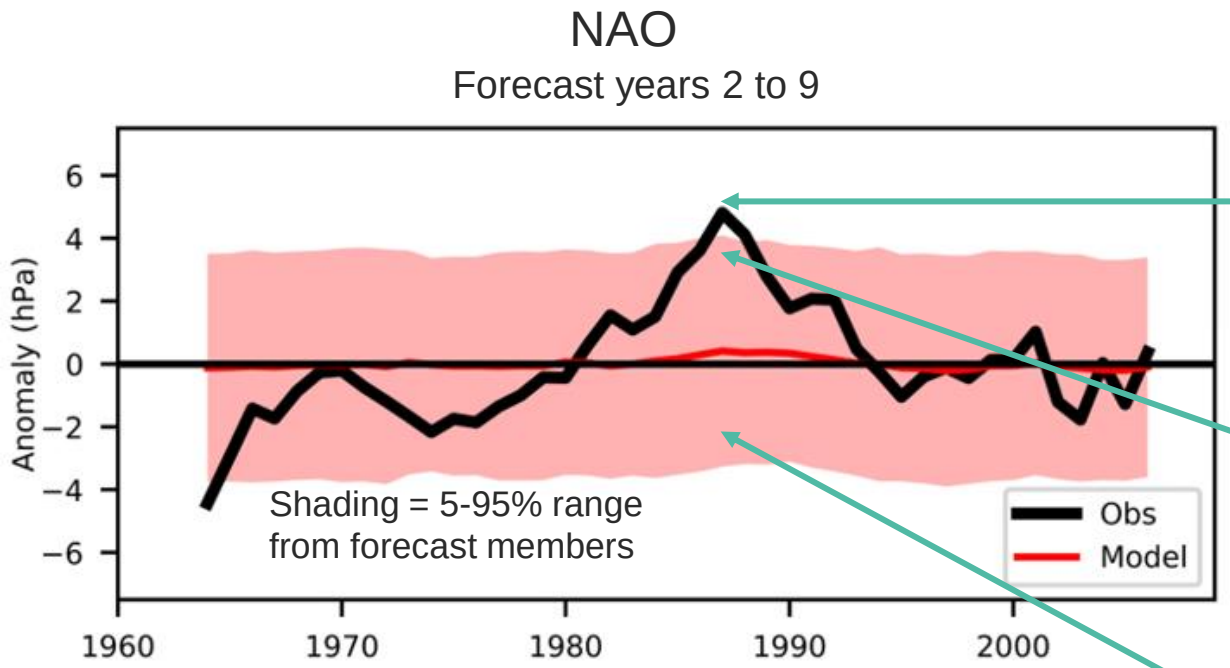
**Taking models at face value might not be the best approach…**

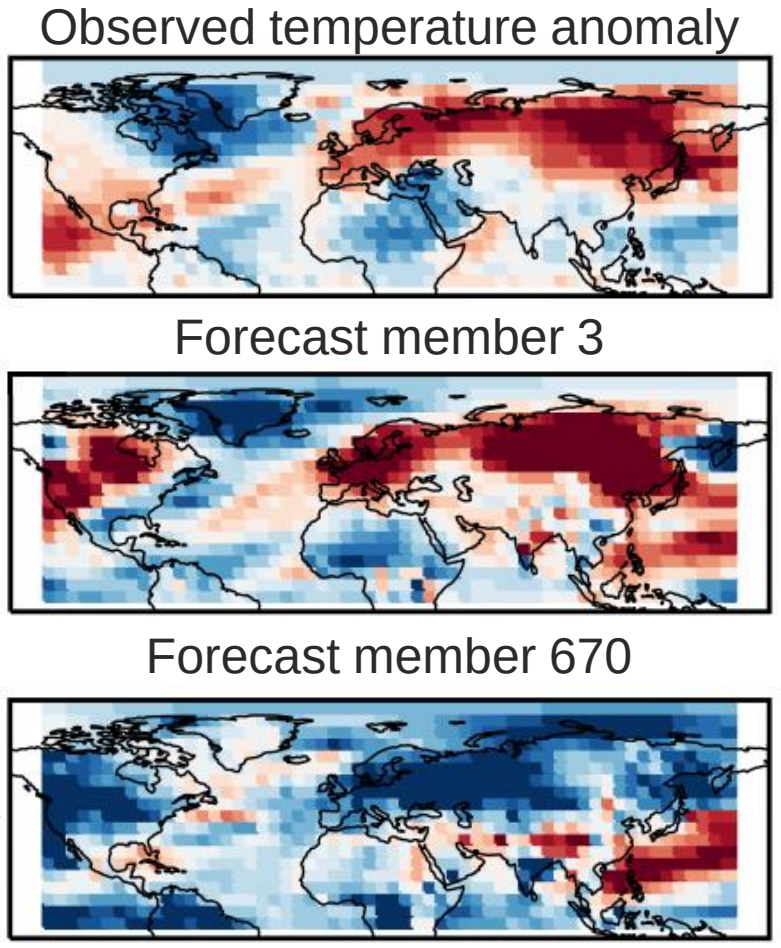
**Zonal wind**




Blackport and Fyfe 2022, Bracegirdle et al 2018, 2022, Eade et al 2022, 2024

# Decadal predictions: irreducible internal variability?




**CMIP5 + CMIP6 decadal predictions 8 year means**

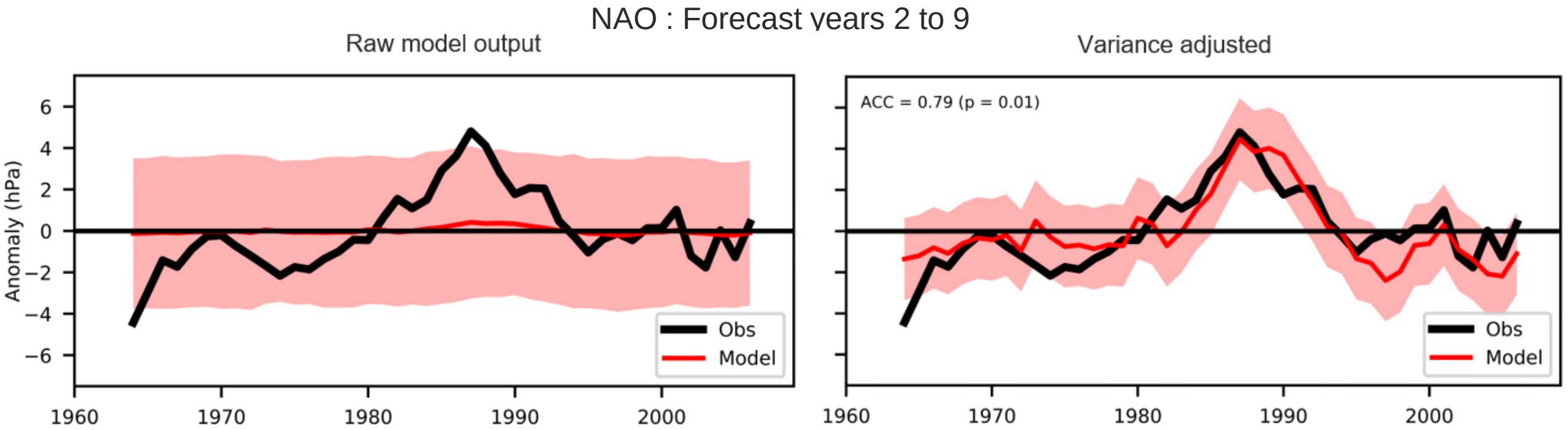
**Almost no signal in ensemble mean (red curve)**

**Irreducible internal variability if models taken at face value**

**BUT this can be tested…**

Smith et al 2020

# Forecast signal is much too weak



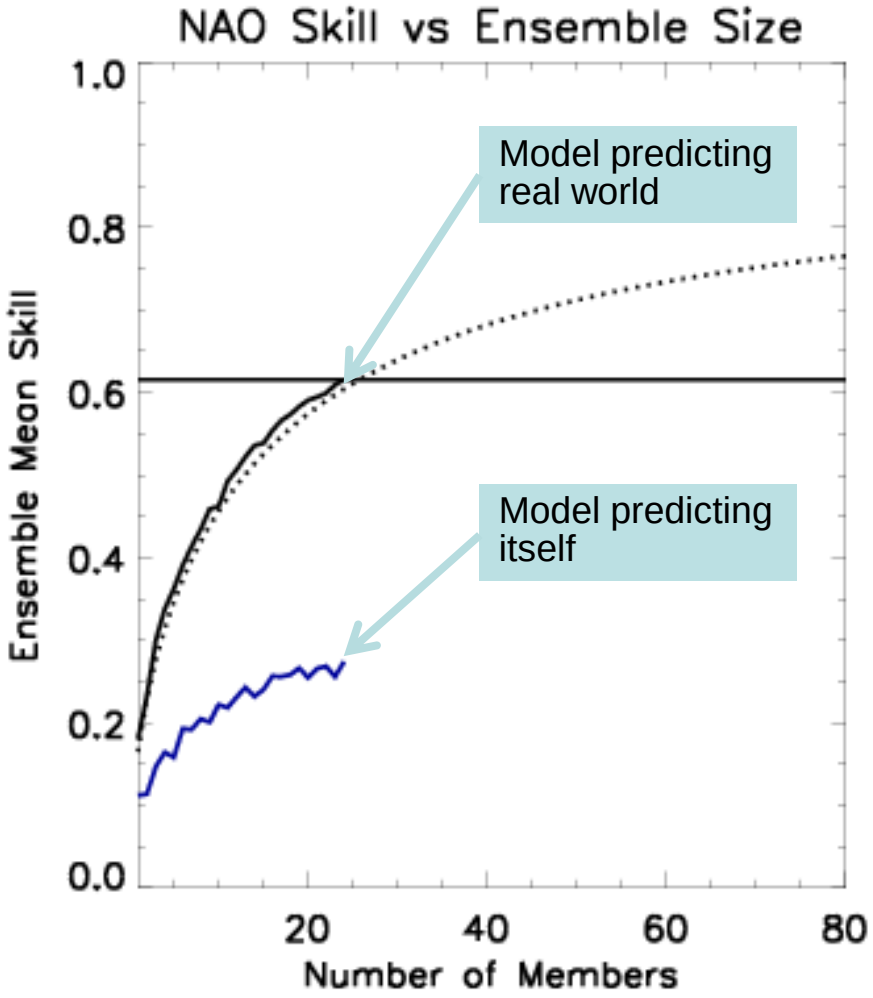

**Ensemble mean is highly correlated with obs (r = 0.79)**

**Should explain 62% of observed variability**

**Magnitude of ensemble mean variability is inconsistent with correlation**
**⇨ Signal to noise error**

Smith et al 2020

# Signal to noise paradox




**Paradox: models predict the real world better than themselves despite perfectly representing themselves**

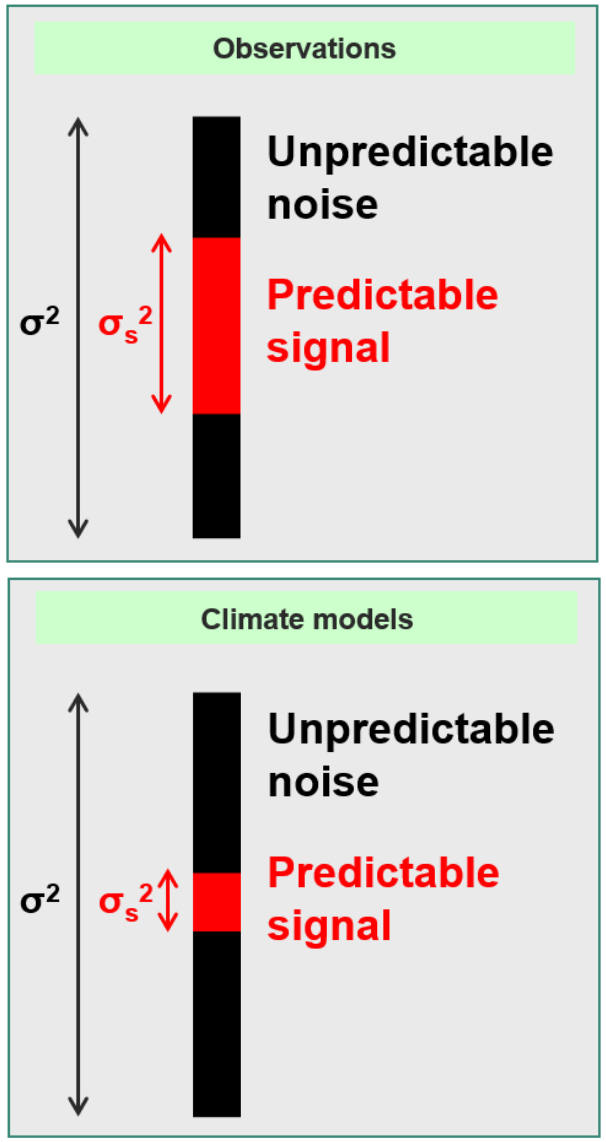
**Members NOT alternate realisations of obs**

**Need a very large ensemble to extract the predictable signal**

**Undermines the basis of ensembles**

Eade et al 2014, Scaife et al 2014, Dunstone et al 2016, 2018, Siegert et al 2016, Baker et al 2018, Scaife and Smith 2018, Smith et al 2019, 2020

# Quantifying the error




## Ratio of predictable components (RPC)

**Observations: Predictable Component PC >= r (anomaly correlation)**

**Models: PC =** $\frac{\sigma_{ensemble\ mean}}{\sigma_{ensemble\ members}}$

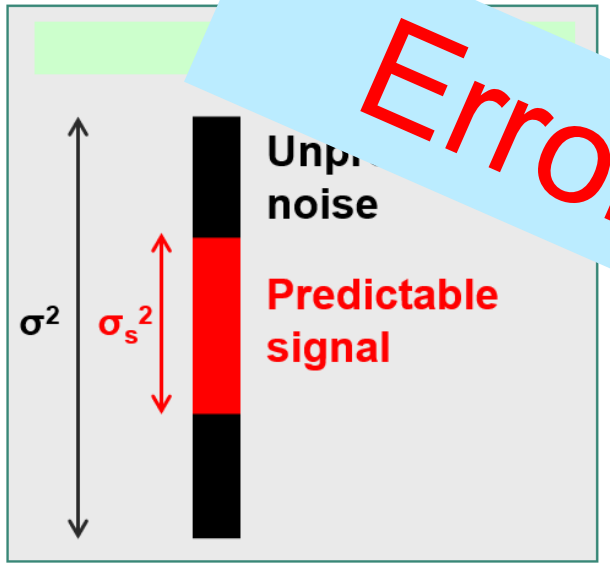
**Ratio of predictable components (RPC) >=** $\frac{r}{\sigma_{ensemble\ mean}/\sigma_{ensemble\ members}}$

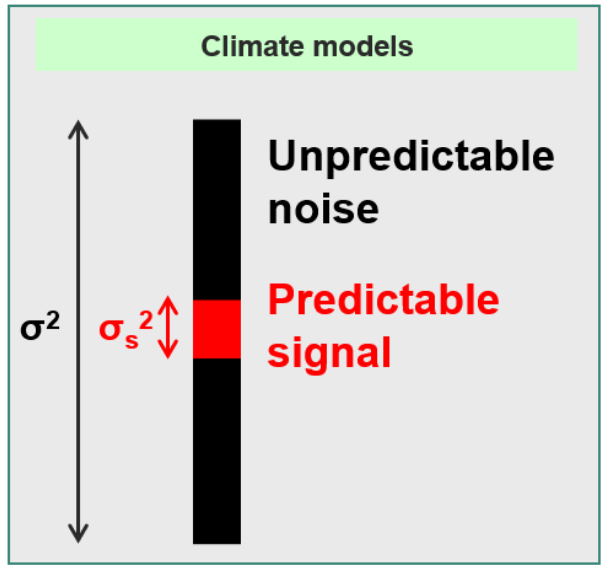
**RPC should be one**

**RPC > 1 shows the signal to noise error**

Eade et al 2014, Scaife and Smith 2018

# Quantifying the error




Error not seen in total variability

## Ratio of predictable components (RPC)

Predictable Component PC >= r (anomaly correlation)

$$\frac{\sigma_{ensemble\ mean}}{\text{ble members}}$$

Ratio of predictable components (R

$$\frac{r}{\text{members}}$$

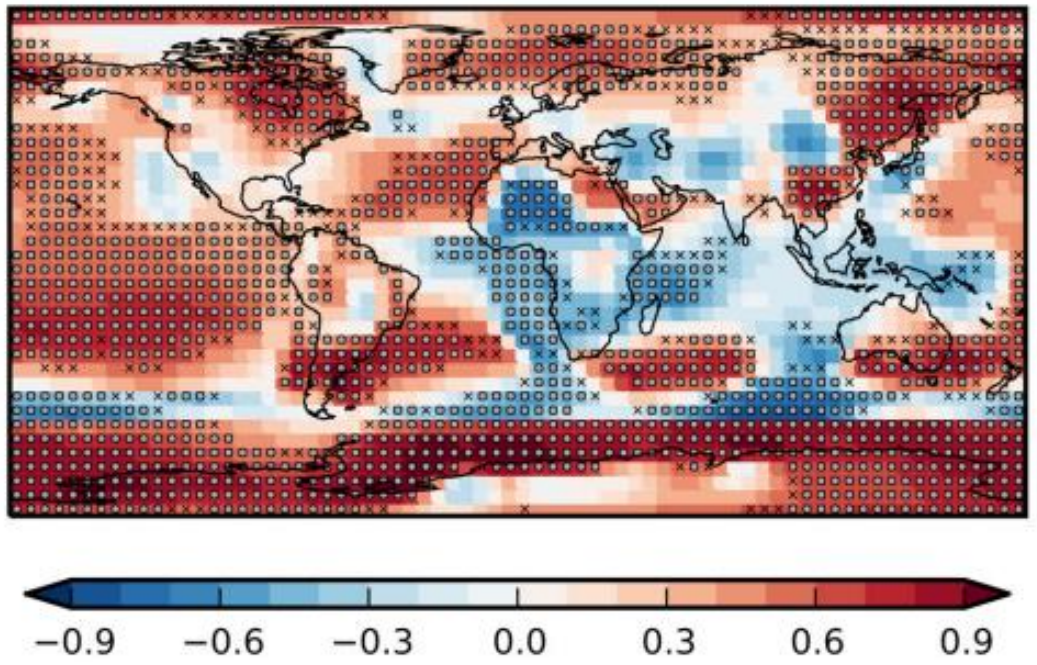
RPC should be one

RPC > 1 shows the signal to noise error

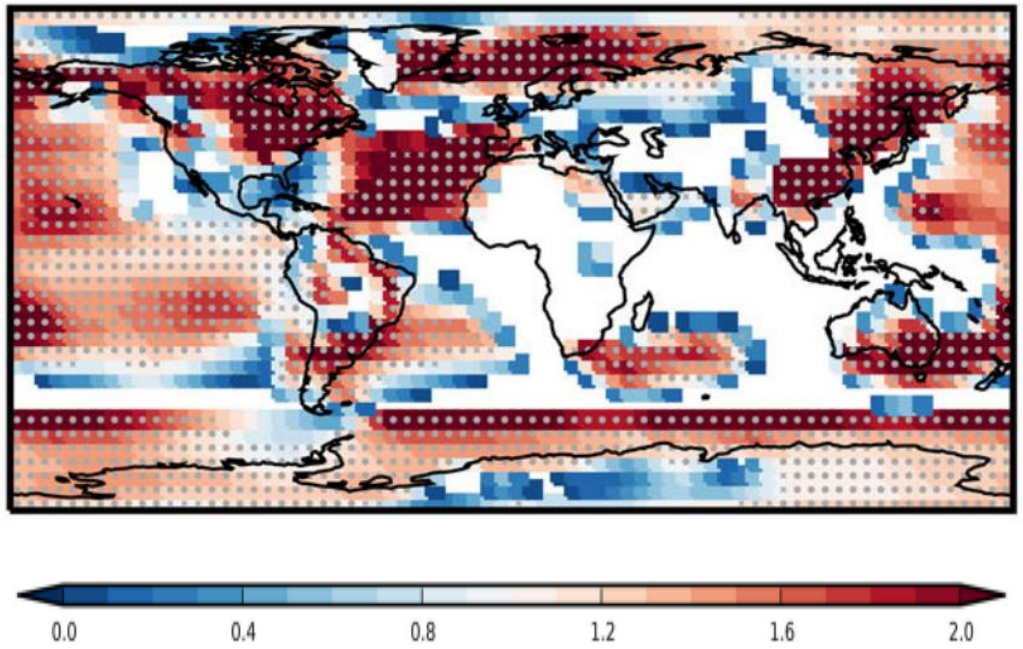
Eade et al 2014, Scaife and Smith 2018

# A key issue

**MSLP skill (years 2-9)**




**Error in magnitude of signal (RPC)**




**Wherever there is skill the modelled signals are too small!**

**Taking models at face value underestimates observed changes**

Smith et al 2019, 2020

# Models can have opposite responses!!

**NAO response to natural forcings (solar + volcanic)**

**31-year rolling means**

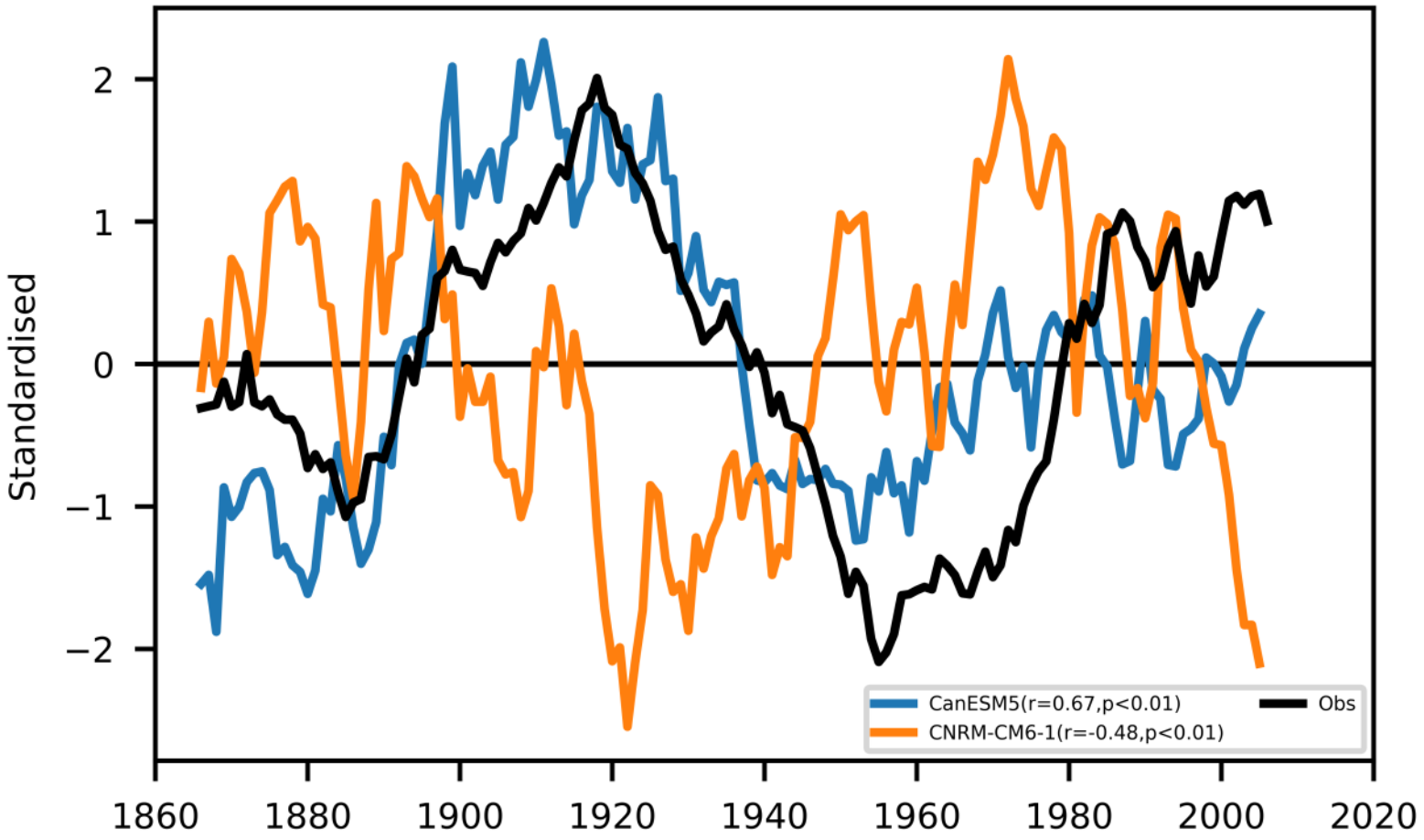
**50 members CanESM5**
**10 members CNRM-CM6-1**

**→ potential role for solar and volcanic forcings?**

**BUT opposite responses :**

**CanESM5** r = 0.67 p < 0.01

**CNRM-CM6-1** r = -0.48 p < 0.01



# Opposite model responses!!

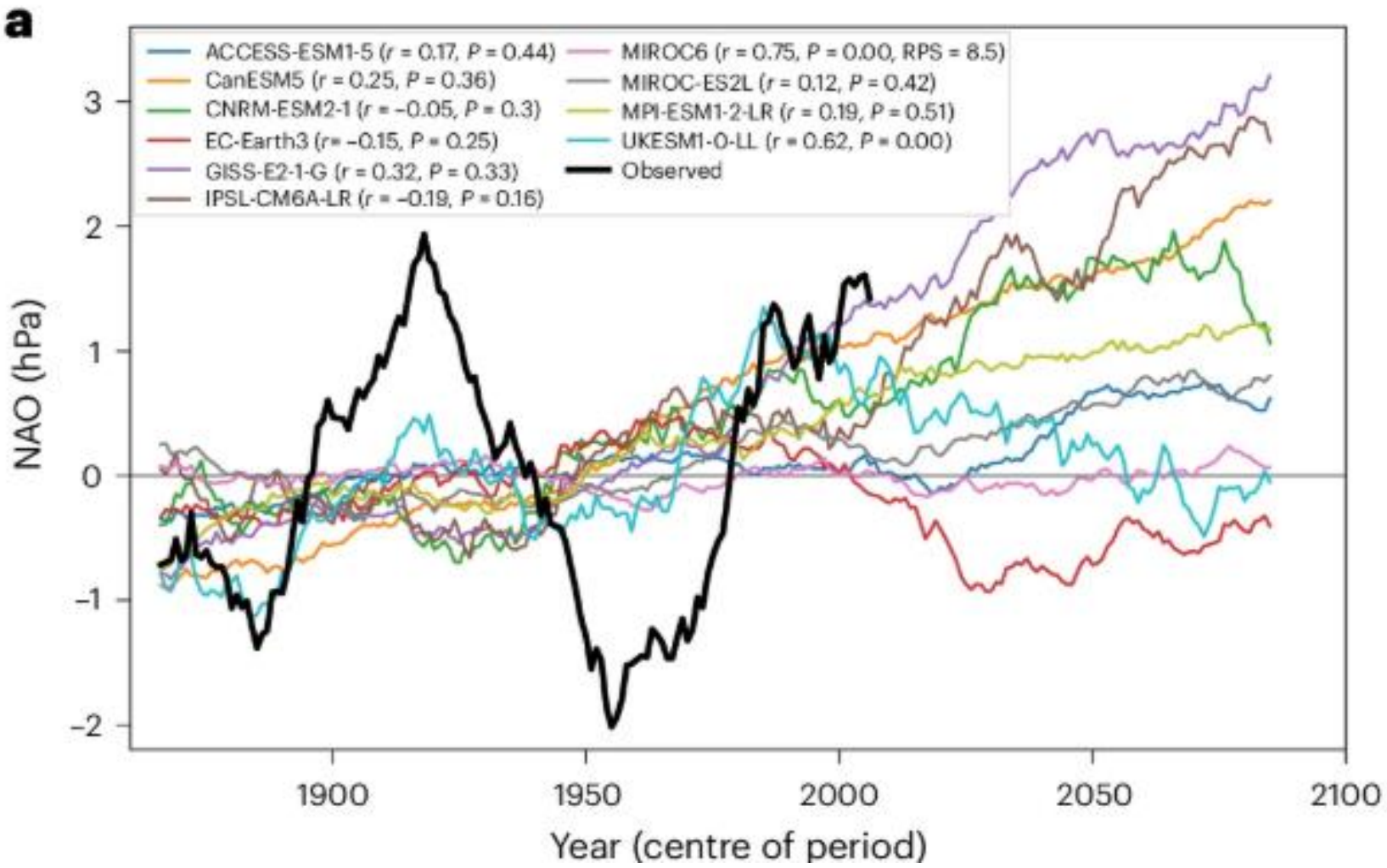
**31-year NAO**

**Historical + ssp245**

**Full range of uncertainty!**

**MIROC6, UKESM1 significantly correlated with obs**

**But very different projections**




Smith et al 2025

# Regression between EEI and U (31 year)

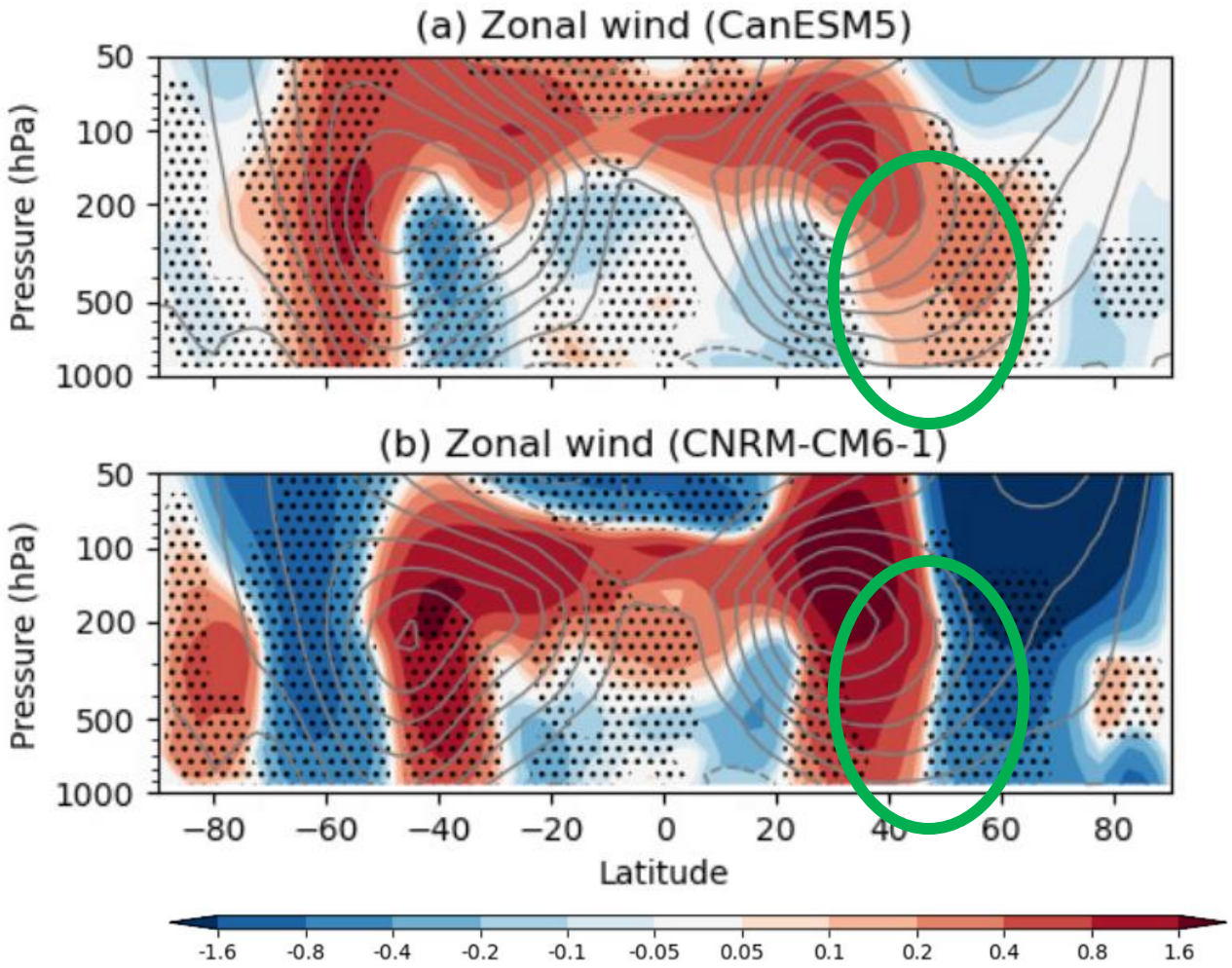
**Climate responds to bring Earth's Energy Imbalance (EEI) back to equilibrium**

**Similar horseshoe pattern increase**

**Stippled where significantly opposite**

**CanESM5 → poleward shift**

**CNRM-CM6-1 → equatorward shift**




Smith et al 2025

# Regression between EEI and T (31 year)

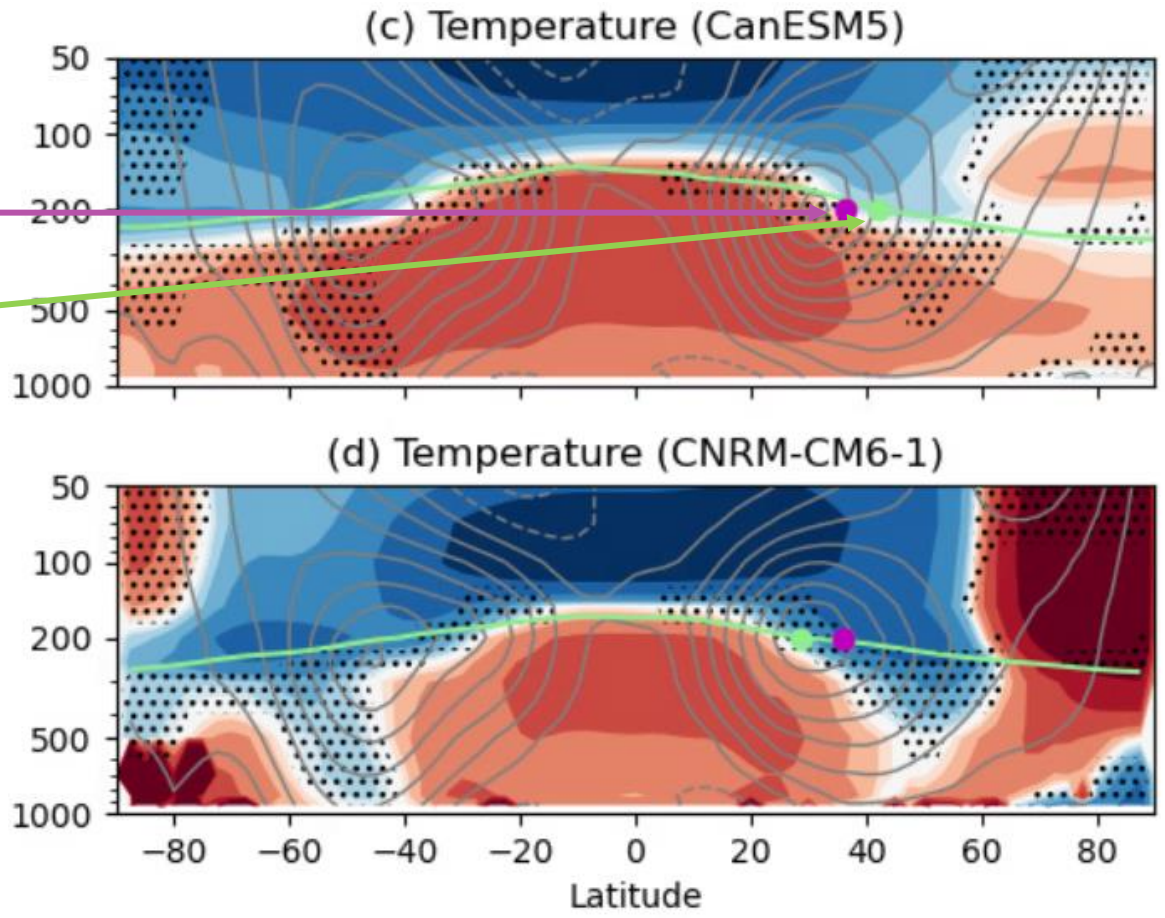
Positive energy imbalance → troposphere warming, thermal gradient at 200 hPa

- Magenta dot = jet centroid at 200 hPa
- Green dot = hygropause latitude at 200 hPa (diagnosed by water vapour contour)

CanESM5 → hygropause latitude **poleward of jet** → poleward shift

CNRM-CM6-1 → hygropause latitude **equatorward of jet** → equatorward shift

Constraint → hygropause latitude relative to jet



# Explaining model differences

16 hist-nat models (with at least 3 ensemble members)

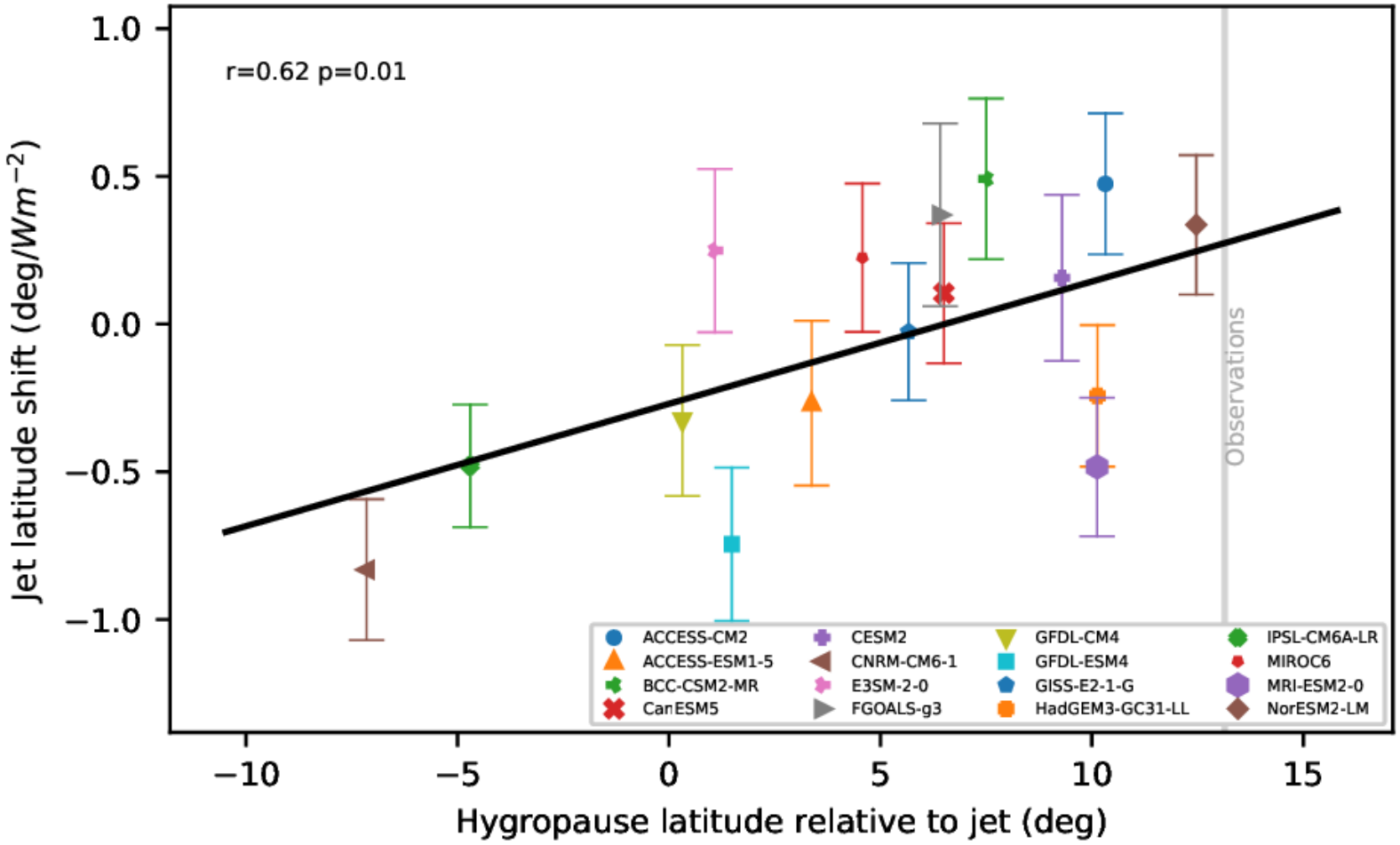
Jet shift at 200 hPa related to EEI

Significant correlation across models with hygropause latitude relative to jet (r = 0.62 p = 0.01)

Models underestimate hygropause latitude relative to jet

→ real world poleward shift

→ greater than models




Smith et al 2025

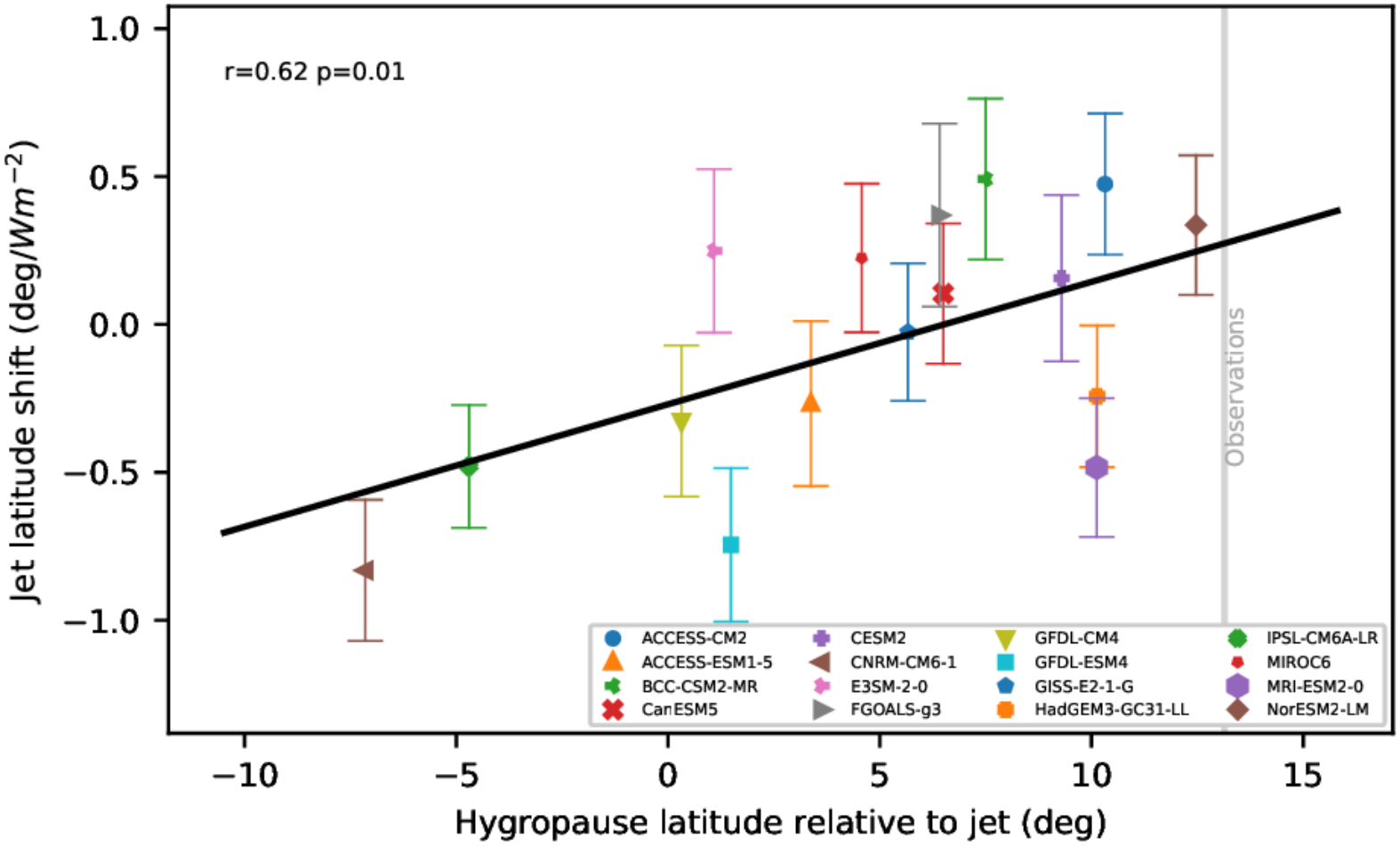
# Explaining model differences

**Emergent constraint opportunity:**

**hygropause latitude relative to jet provides weights for each model**

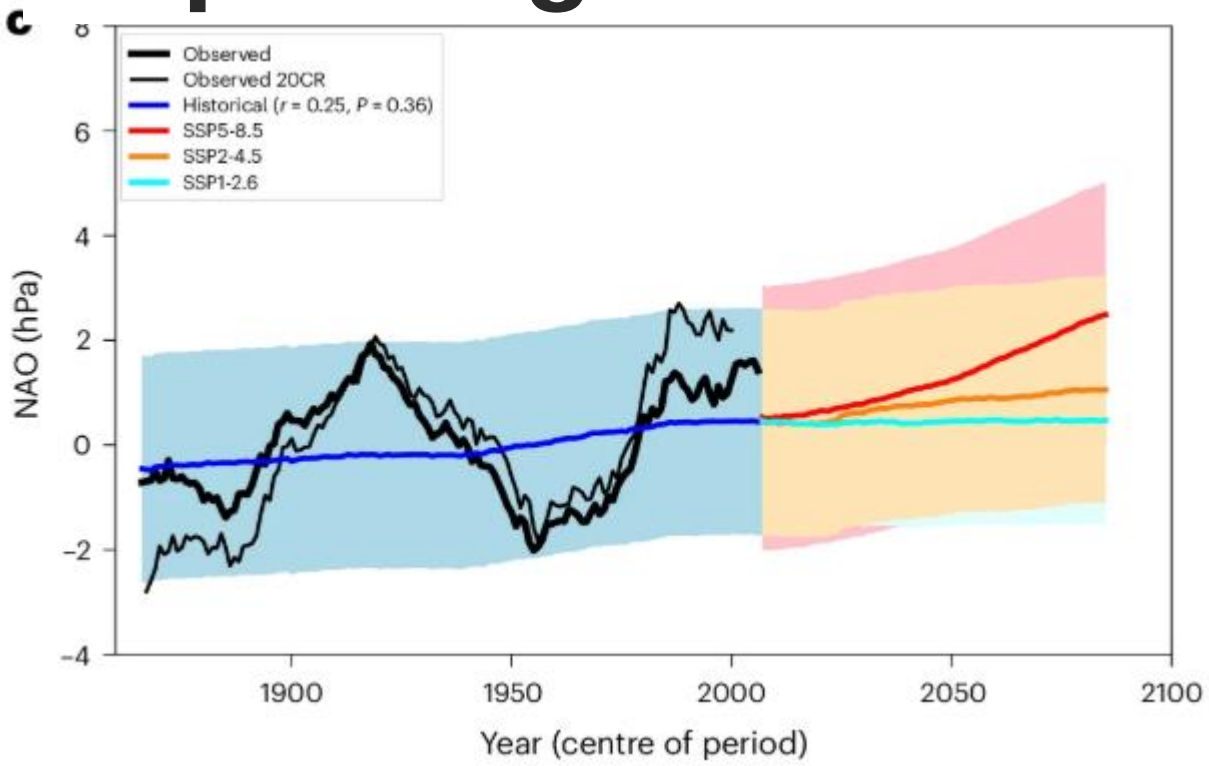
**Generally higher weights for models closer to obs**

**Negative weights for models that shift the jet the wrong way**




Ensemble regression, Bracegirdle and Stephenson 2012

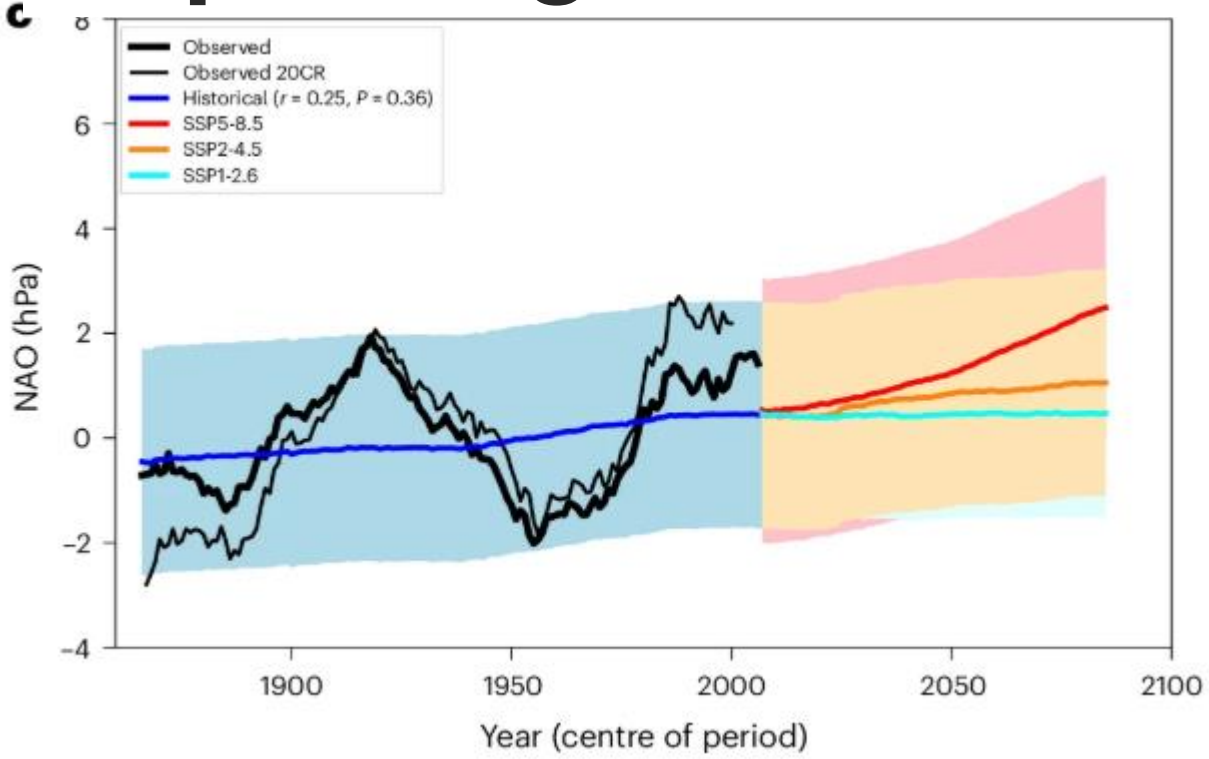
# Exploiting model differences

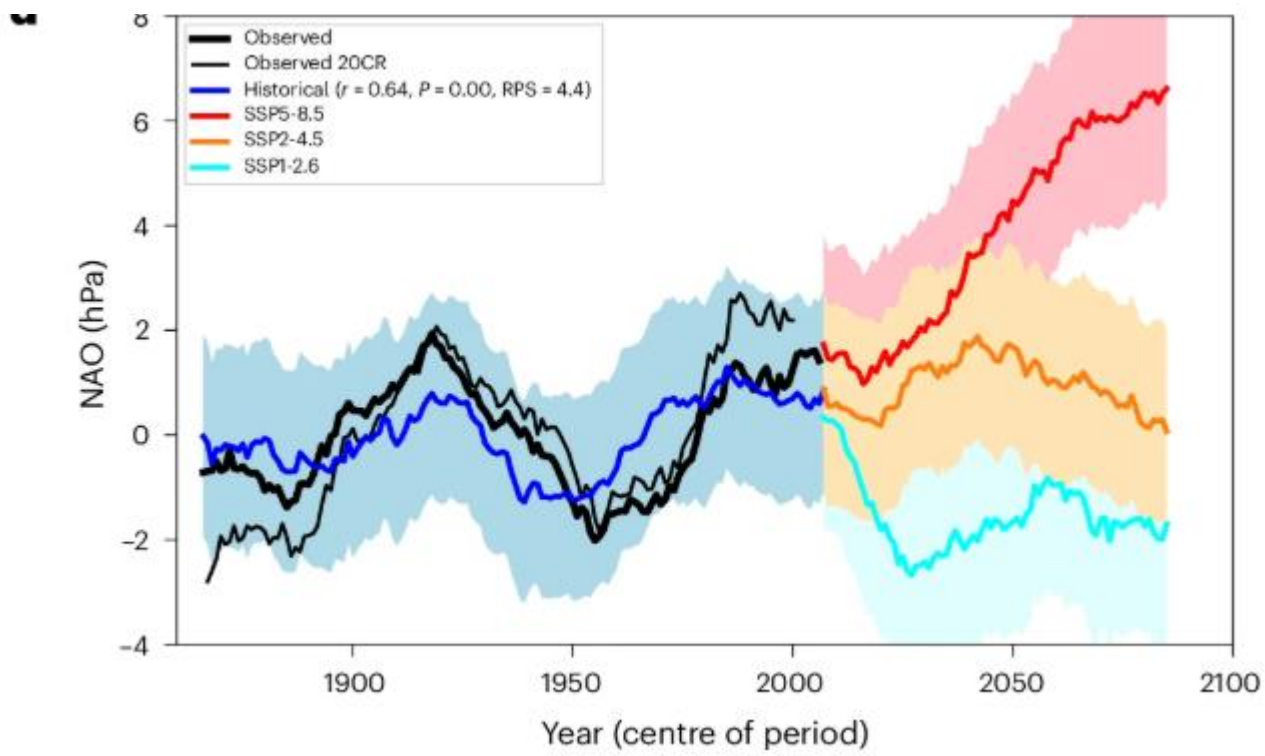



**Standard multi-model mean**
**→ no correlation with obs**
**→ decadal variability not externally forced**

Smith et al 2025

# Exploiting model differences




**Standard multi-model mean**
**→ no correlation with obs**
**→ decadal variability not externally forced**

**Constrained**
**→ high correlation with obs**
**→ decadal variability is externally forced (at least partly)**

**But subject to signal to noise paradox → scale by ~4 x**

**Projected to increase to unprecedented levels ⇨ missed when taking models at face value!**

Smith et al 2025

# Summary

Wherever decadal predictions of atmospheric circulation are skilful the signals are too weak

→ need very large ensembles to extract predictable signals

→ need to boost the forecast signal

→ also true for long term NAO projections

Models show very different (even opposite!) responses to the same forcings

– they can't all be right!

Accounting for model differences and errors reveals externally forced NAO variability and projections to unprecedented levels

→ missed if models are taken at face value

→ we should not be surprised by this – it would be a miracle if models simulated a perfect signal to noise ratio!

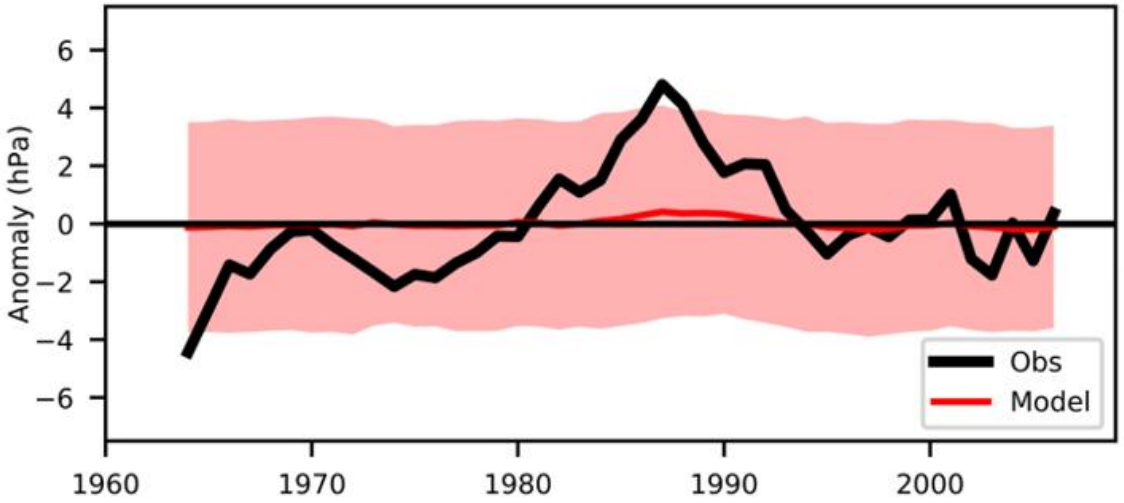
Not accounting for model errors could leave society unprepared for impending extremes

# Extra slides

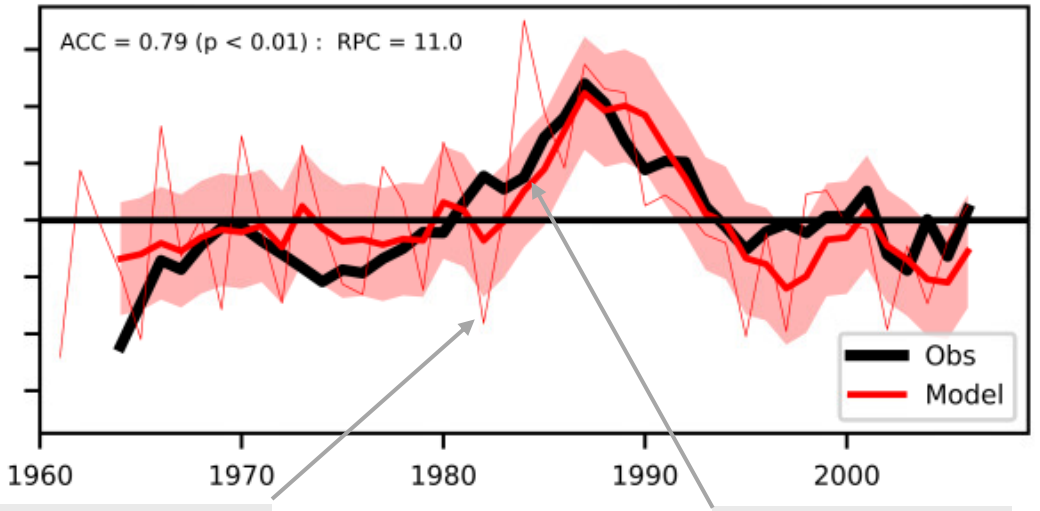
# Forecast signal is MUCH too weak

NAO : Forecast years 2 to 9

Raw model output




Variance adjusted




169 member mean

676 member mean

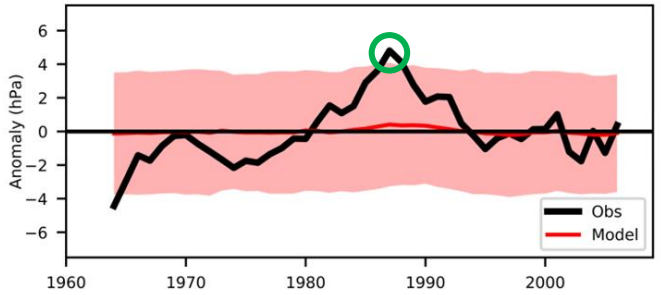
**Ratio of predictable components RPC = 11**

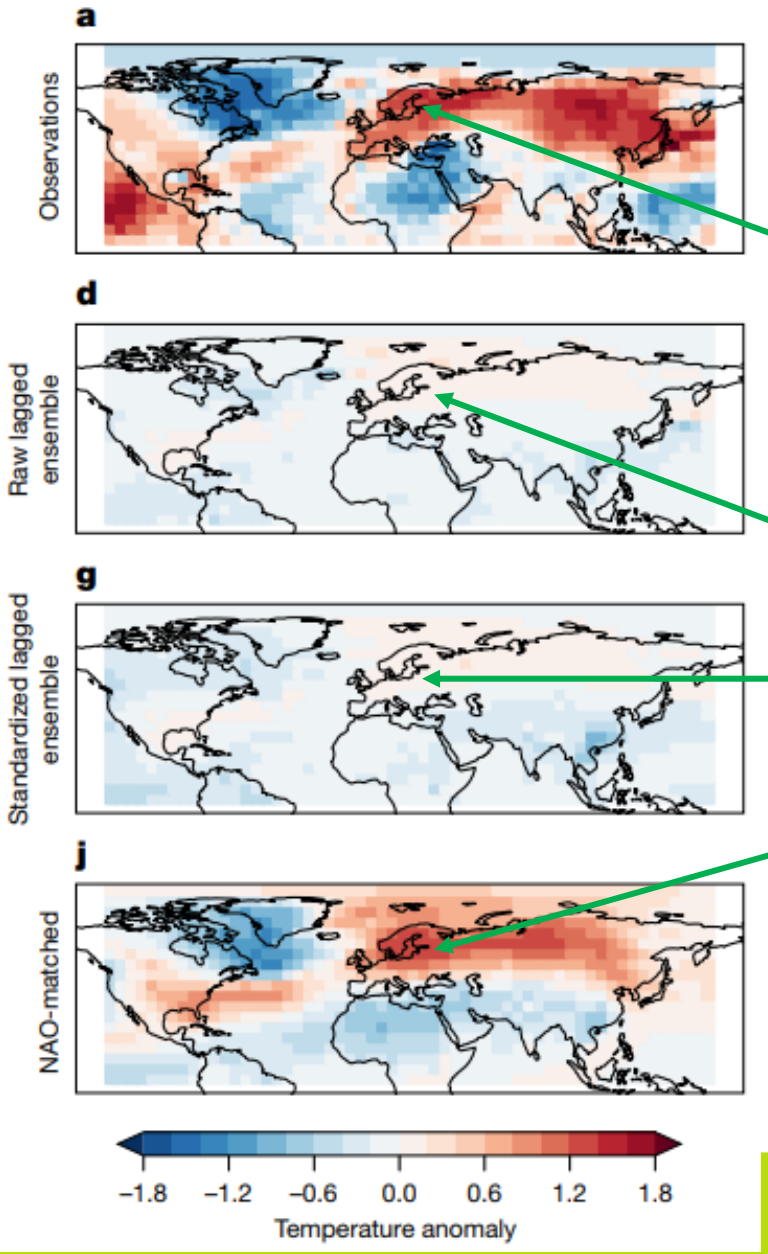
**Signal is an order of magnitude too weak in climate model ensemble**

**Need 100 times the number of ensemble members to extract the signal**

**Smith et al 2020**

# Not overcome by scaling




$T = T_{DYN} + T_{THERMO} + \varepsilon$

Real world: $T_{DYN} >> T_{THERMO}$

Ensemble mean: $T_{DYN} << T_{THERMO}$ because NAO signal too small

Scaling retains the incorrect ratio $T_{DYN}/T_{THERMO}$

Can be overcome by selecting ensemble members with correct magnitude of NAO

**Standard D&A approach will not work**

**Need to look at models in new ways**

Smith et al 2020, 2022

# Mechanism

**Hovmuller plots of rolling temperature at 200hPa anomalies from preceding 30-year mean (multi-model mean)**

**Tropical cooling following volcanic eruptions**
**→ minimum ~1990 (Agung+El Chichon+Pinatubo)**

**Greenhouse gas warming**
**→ minimum ~1960**

**$\frac{\partial \overline{T}}{\partial \emptyset}$ at jet latitude (35N) increases under SSP5-8.5**

**But reduces with mitigation**