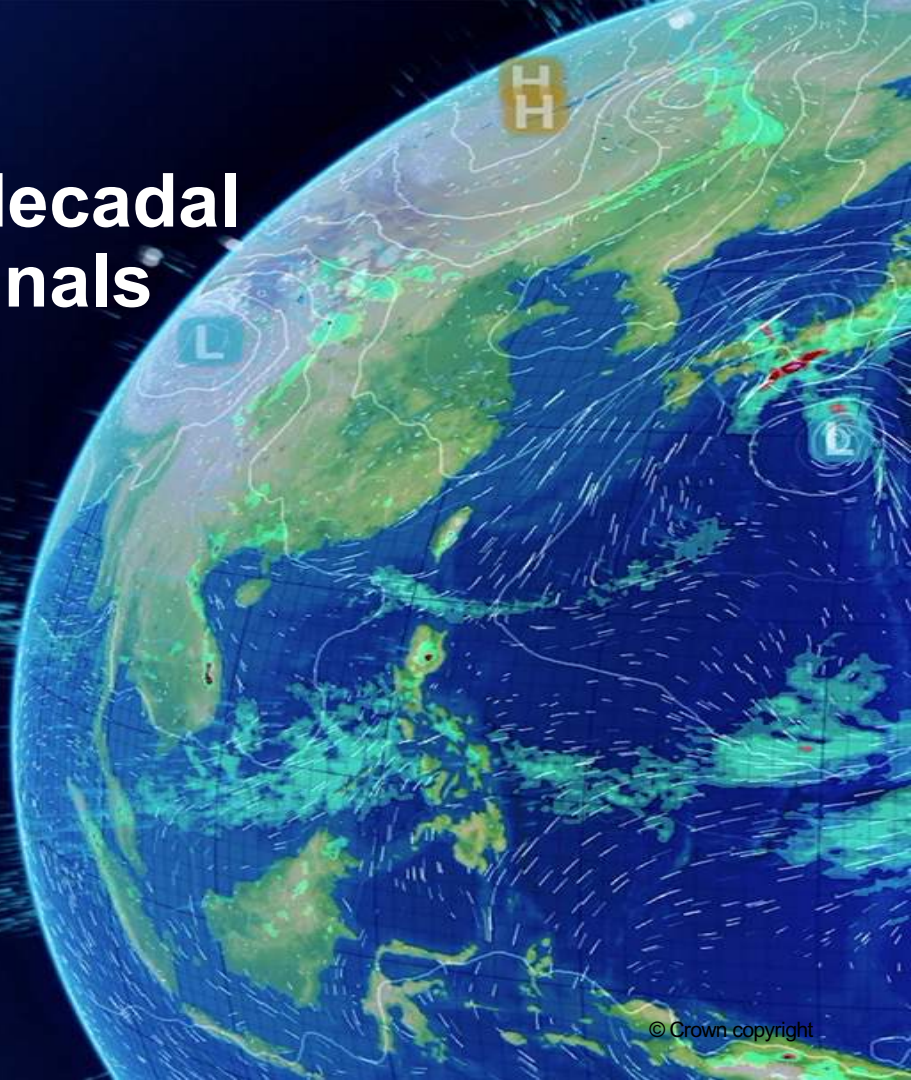
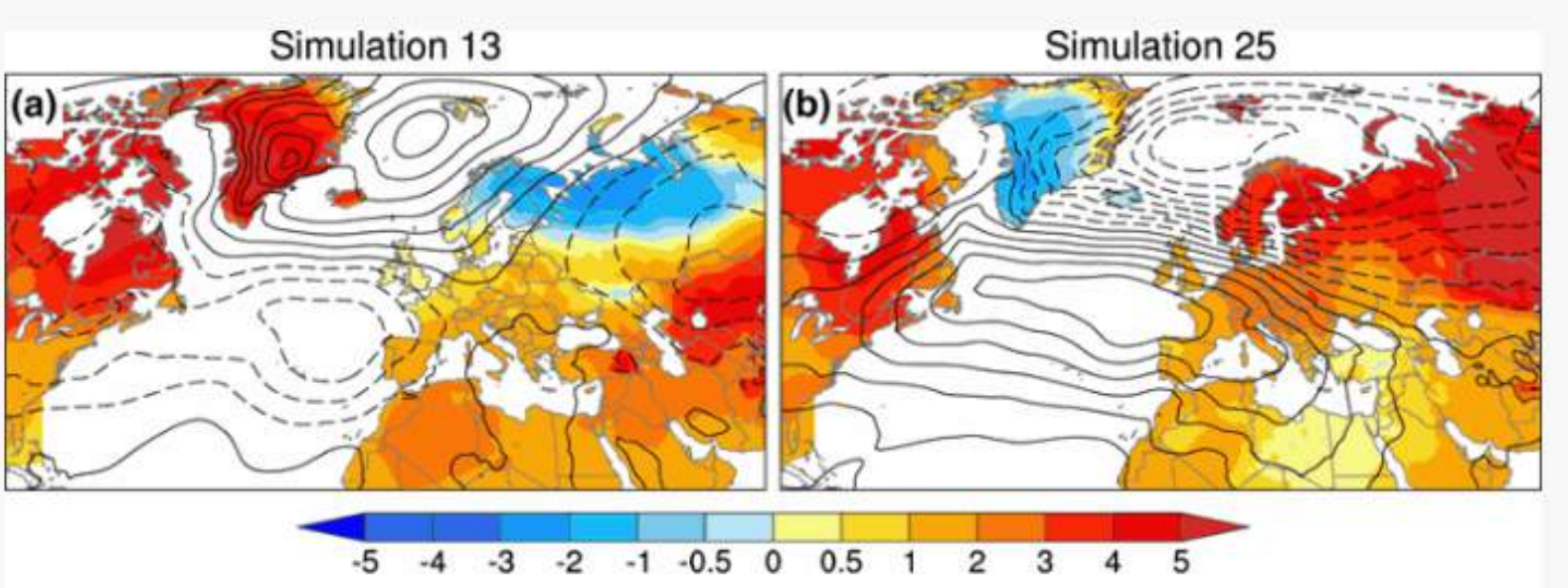
# Gross underestimation of decadal atmospheric circulation signals in climate models

D. M. Smith A. A. Scaife, R. Eade, P. Athanasiadis, A. Bellucci, I. Bethke, R. Bilbao, L. F. Borchert, L.-P. Caron, F. Counillon, G. Danabasoglu, T. Delworth, F. J. Doblas-Reyes, N. J. Dunstone, V. Estella-Perez, S. Flavoni, L. Hermanson, N. Keenlyside, V. Kharin, M. Kimoto, W. J. Merryfield, J. Mignot, T. Mochizuki, K. Modali, P.-A. Monerie, W. A. Muller , D. Nicoli, P. Ortega, K. Pankatz, H. Pohlmann, J. Robson, P. Ruggieri, R. Sospedra-Alfonso, D. Swingedouw, Y. Wang, S. Wild, S. Yeager, X. Yang and L. Zhang



# Huge uncertainty in projected 30 year trends



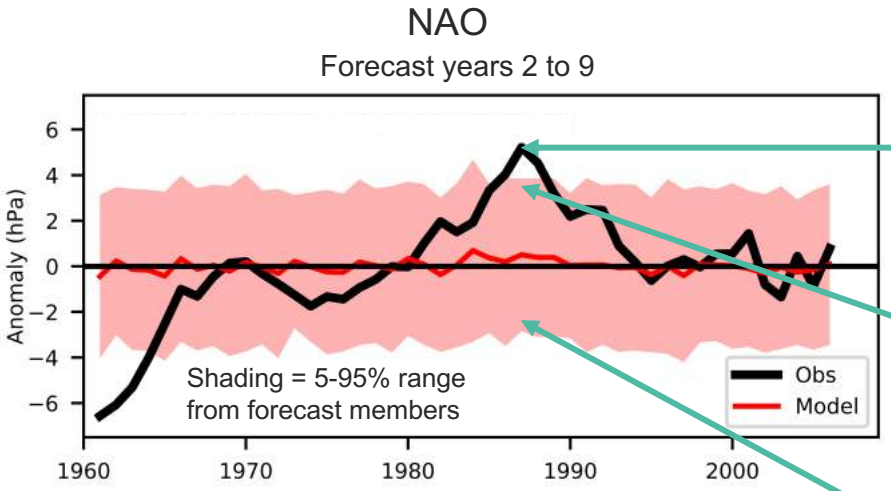

**Opposite sign of NAO trends for 2016-2045**

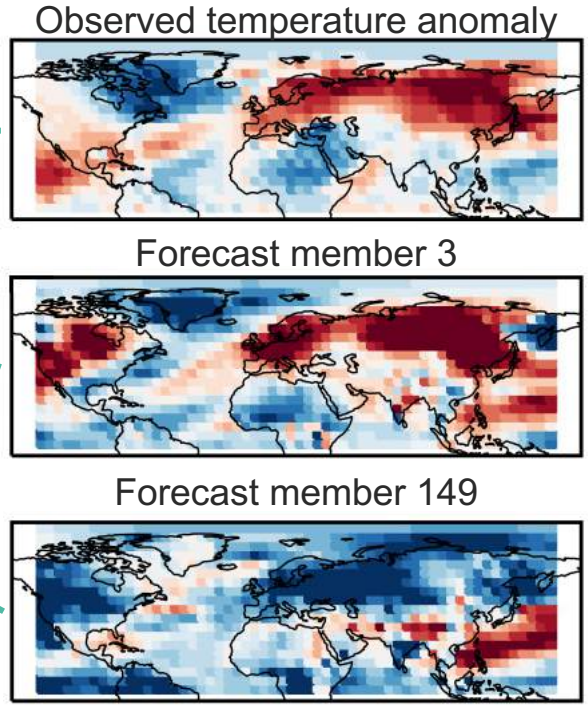
**Projections use the same climate model differing by tiny perturbations to initial state**

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

Deser et al 2017

# Huge uncertainty in decadal predictions?




**CMIP5 and CMIP6 decadal predictions**

**14 models, 169 ensemble members (>77,000 years!)**

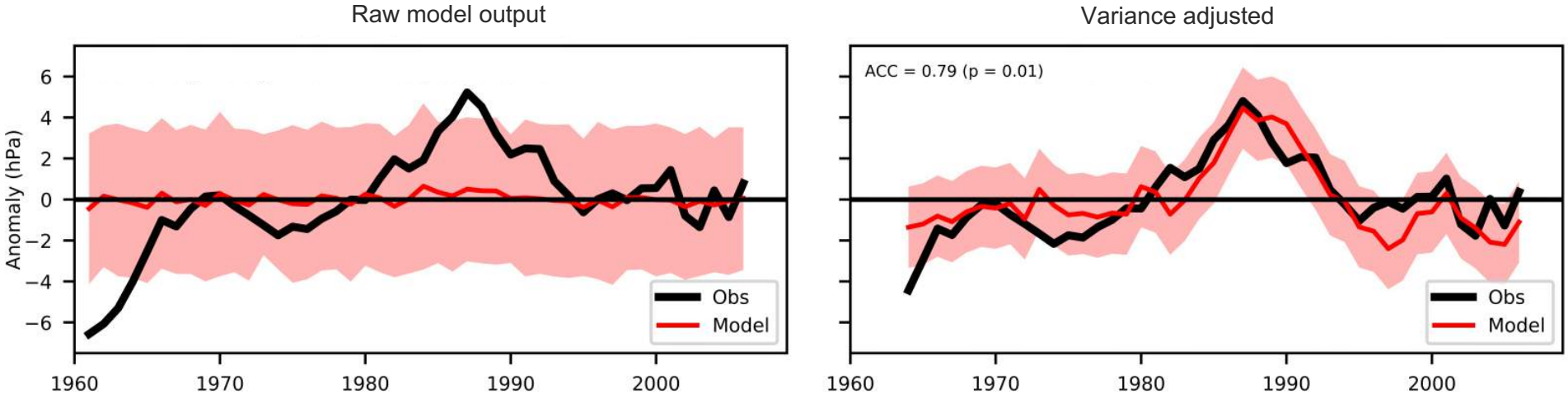
**Huge uncertainty if models taken at face value**

**BUT this can be tested…**

**Smith et al (submitted)**

# Forecast signal is much too weak

NAO : Forecast years 2 to 9



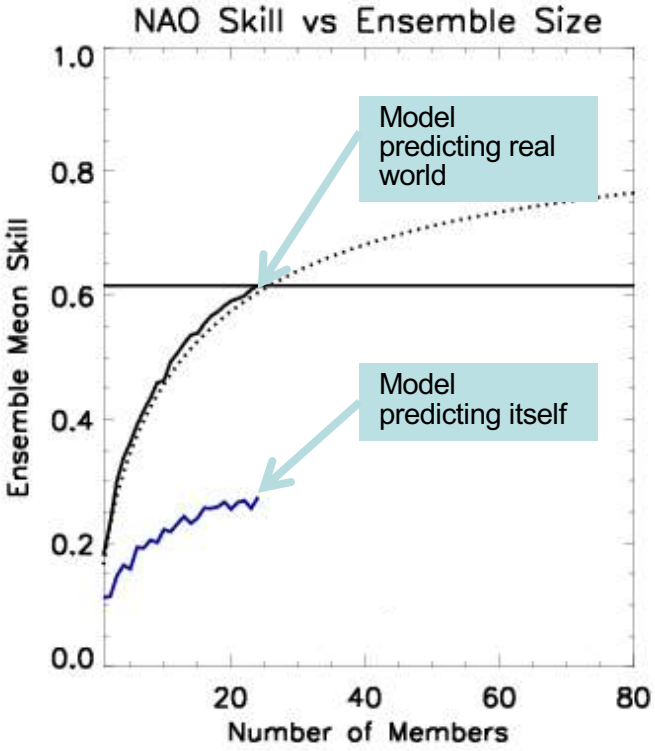

**Ensemble mean is highly correlated with obs (r = 0.79)**

**Should explain 62% of observed variability**

**Magnitude of ensemble mean variability is inconsistent with correlation**

**Smith et al (submitted)**

# Signal to noise paradox




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

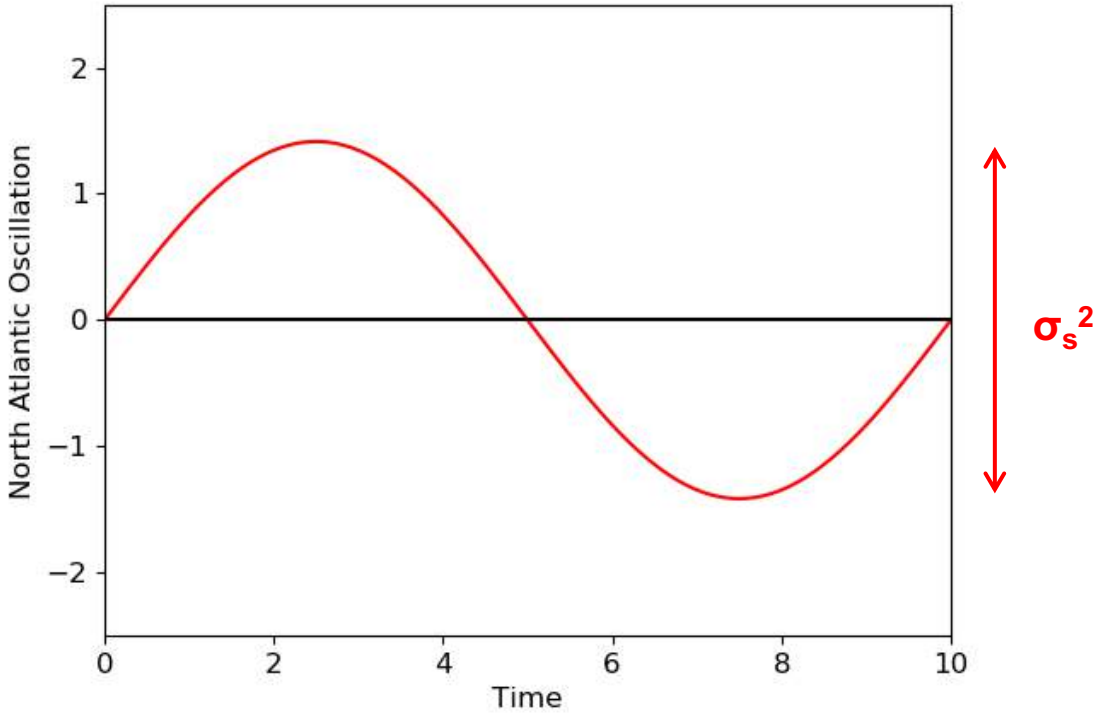
**Members NOT alternate realisations of obs**

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

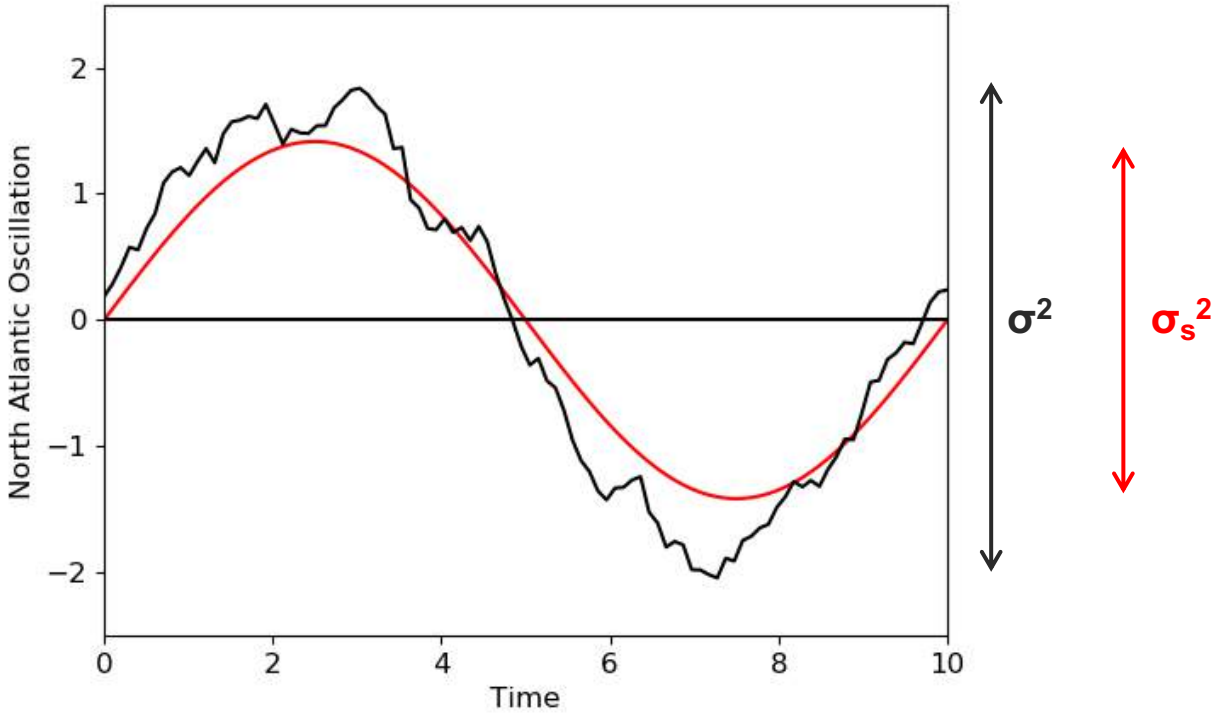
**Undermines the basis of ensemble prediction**

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

# Signal



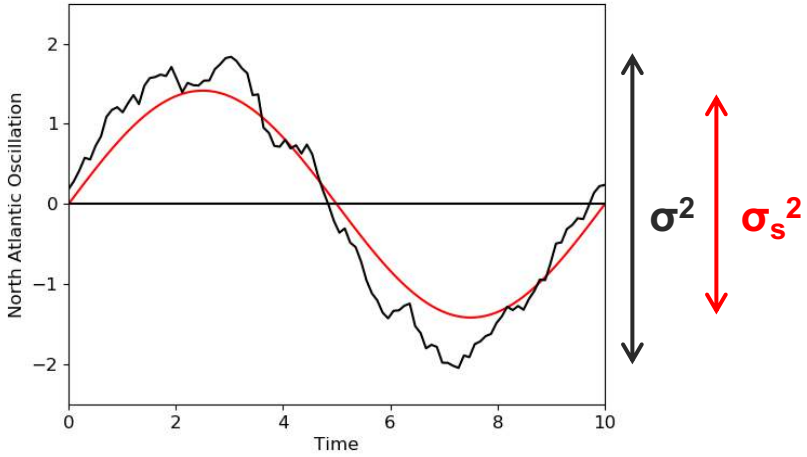
# Signal + noise



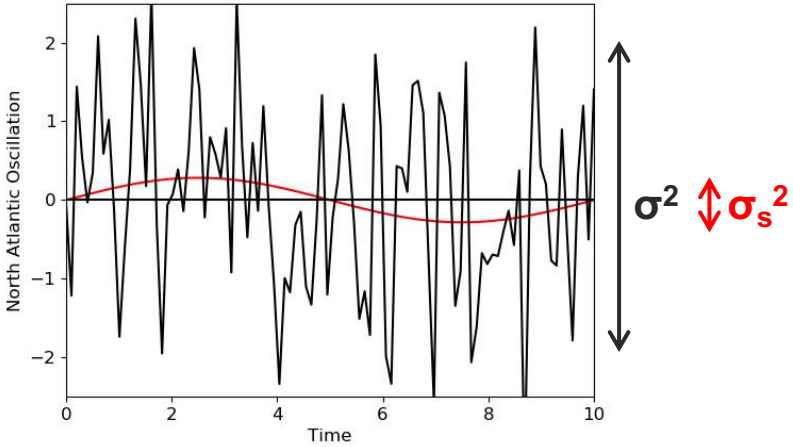

Eade et al 2014, Scaife and Smith 2018

# A simple interpretation

**Observations**




**Climate models**



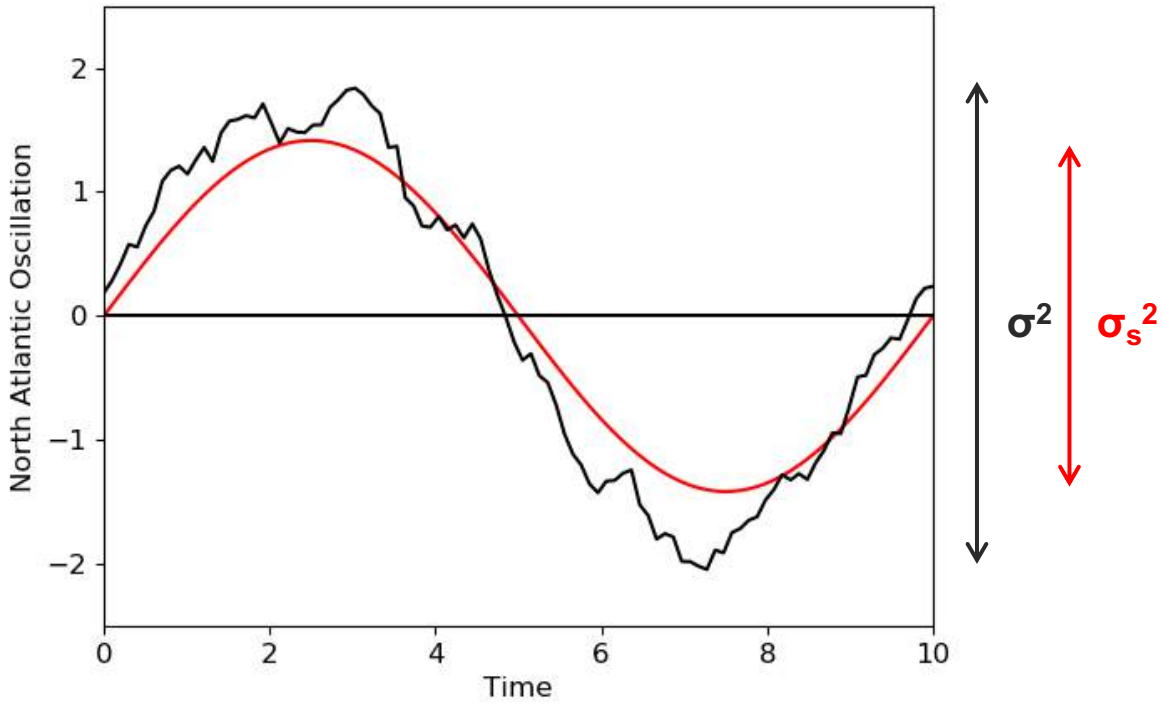

**Climate models have the right amount of variability**

**BUT**

**The *proportion* of variability that is predictable is too small**

Eade et al 2014, Scaife and Smith 2018

# Predictable component



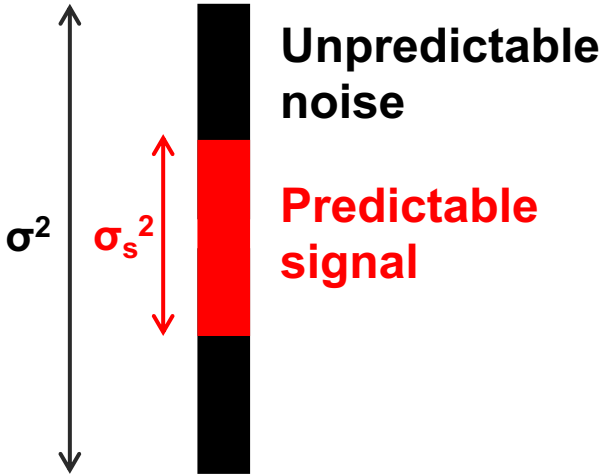

**Predictable component (PC)**

$$= \frac{signal}{total}$$

$$= \frac{\sigma_s}{\sigma}$$

Eade et al 2014, Scaife and Smith 2018

# Ratio of predictable components




Observations: PC >= r (anomaly correlation)

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

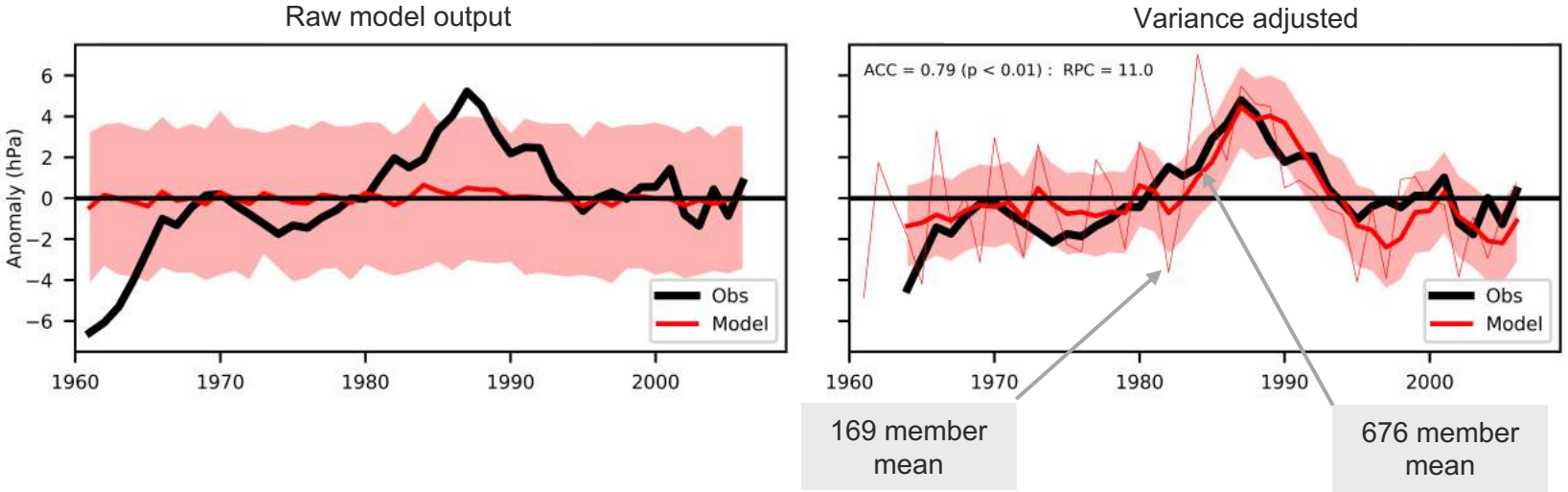
Ratio of predictable components (RPC) >= $\frac{r}{\sigma_{ensemble\ mean}/\sigma_{ensemble\ members}}$

Ratio of predictable signals = $\frac{\sigma_s^{obs}}{\sigma_s^{models}} = RPC\ \frac{\sigma^{obs}}{\sigma^{models}} \approx$ RPC

Eade et al 2014, Scaife and Smith 2018

# Forecast signal is MUCH too weak

NAO : Forecast years 2 to 9




**Ratio of predictable components RPC = 11**

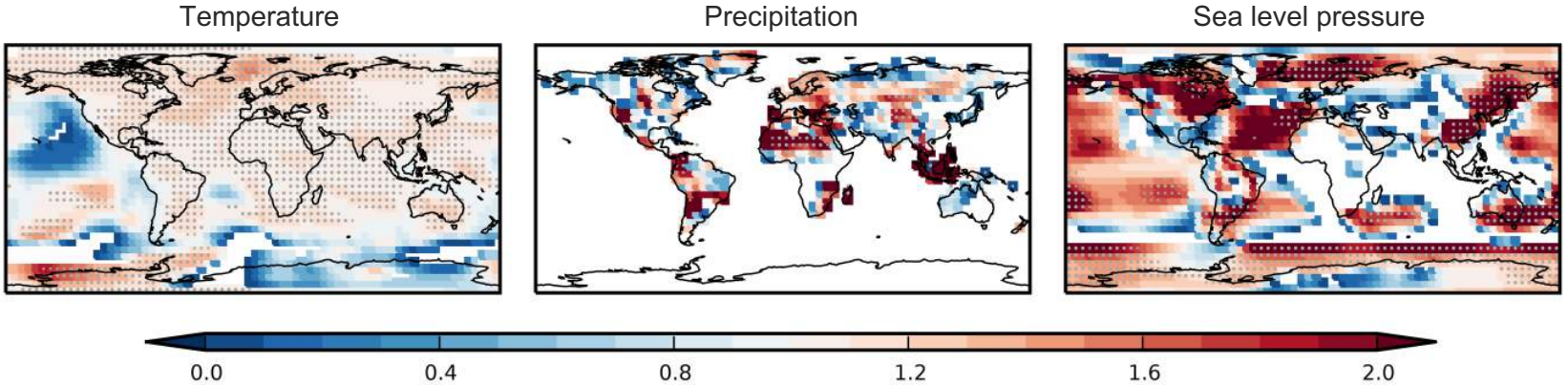
**Signal is an order of magnitude too weak in climate model ensemble**

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

Smith et al (submitted)

# A widespread issue

RPC : Forecast years 2 to 9




**RPC > 1 in many regions**

**Especially for precipitation and pressure**

**Atmospheric circulation signals too weak**

Smith et al 2019

# Why worry about this?

**Probabilistic and error based skill measures will give inaccurate estimates of the forecast skill that is potentially available**

**Ensemble forecasts of are sometimes overconfident and their statistical properties may be improved by techniques such as stochastic physics which often increase ensemble spread. However, such techniques could potentially exacerbate this problem where the signal-to-noise ratio is too small and models are under-confident**

**Upper limits on predictability are sometimes estimated from model ensembles but this is not the case here**

**Event attribution will give inaccurate estimates of the probability of extremes, especially in the North Atlantic sector, where the signal-to-noise ratio is too small**

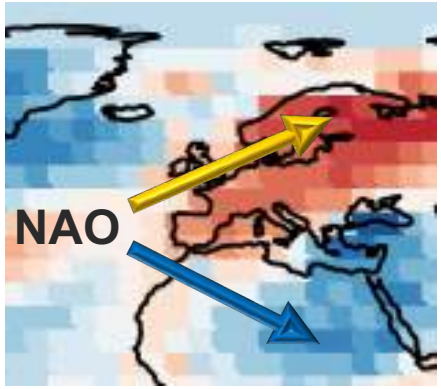
**Estimates of unpredictable internal variability in regional climate change over the coming decades may be too large and forced signals may be too weak**

**Resolving the paradox would allow reduced ensemble sizes and increase prediction skill**

# Impacts of the NAO

**Real world**

GHGs ↓↓↓↓↓↓




**Ensemble mean**

GHGs ↓↓↓↓↓↓


**In reality NAO dominates over other factors (GHGs)**

**NAO too weak in ensemble mean → GHGs dominate (but small impact)**

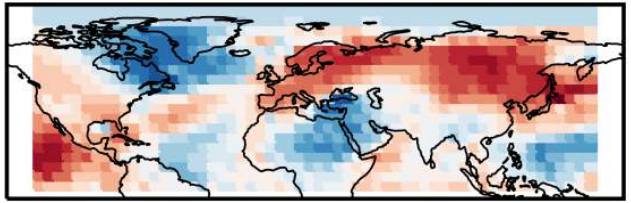
Smith et al (submitted)

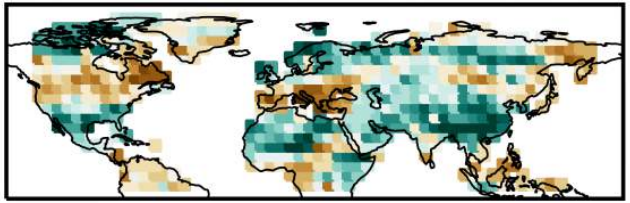
# Impacts of extreme NAO

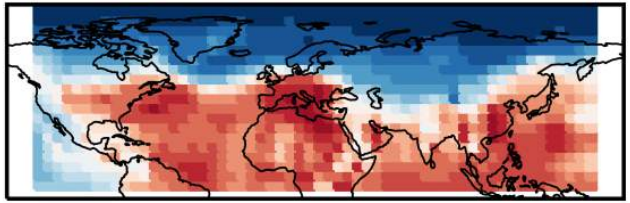
Temperature
(a) Observations

Precipitation
(b) Observations

Pressure
(c) Observations



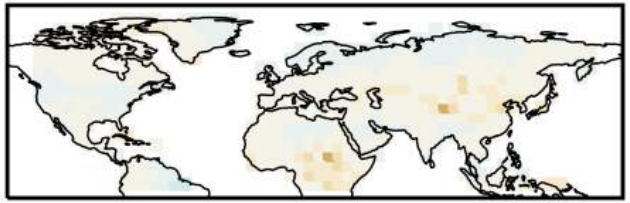


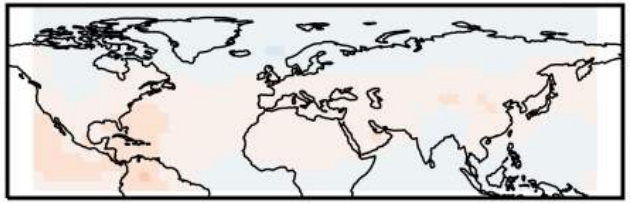


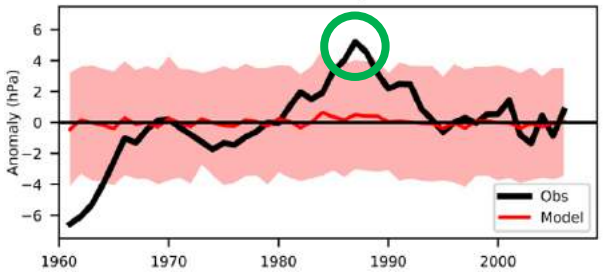
(d) Raw lagged ensemble

(e) Raw lagged ensemble

(f) Raw lagged ensemble








**Extreme positive NAO in late 1980s to 1990s**

**Clear quadrupole impact seen in observations**

**Not captured in raw model data**

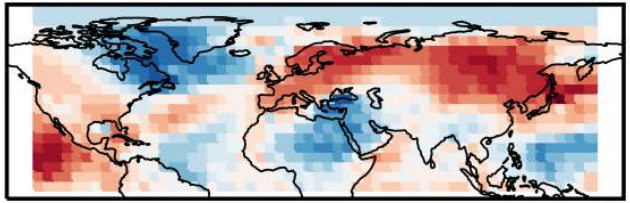
Smith et al (submitted)

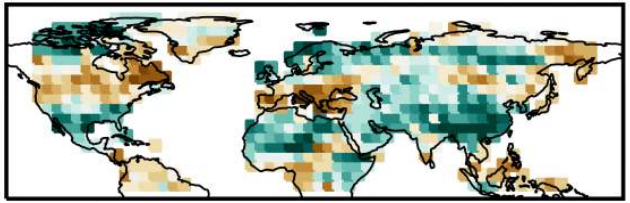
# Impacts of extreme NAO

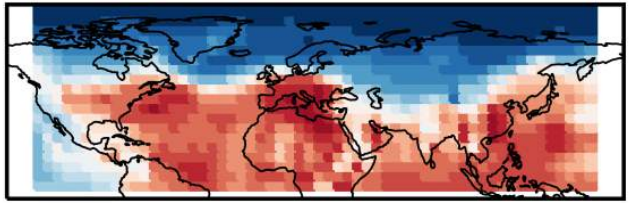
Temperature
(a) Observations

Precipitation
(b) Observations

Pressure
(c) Observations



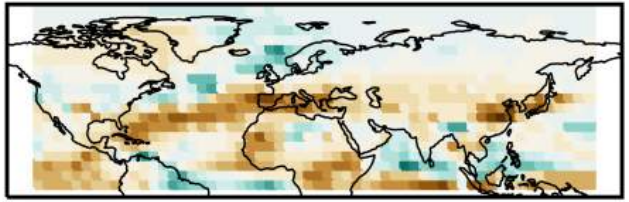


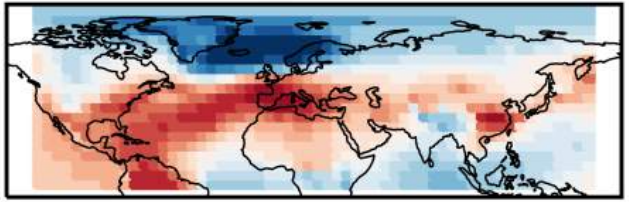


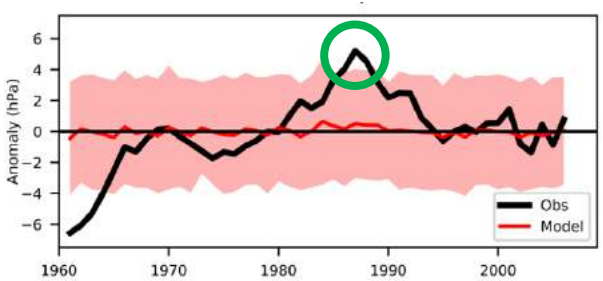
(g) Standardised lagged ensemble

(h) Standardised lagged ensemble

(i) Standardised lagged ensemble








**NAO captured by ensemble mean (standardised)**

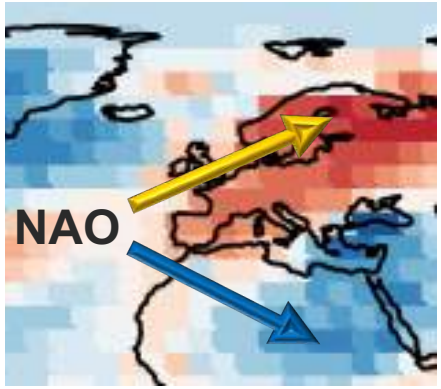
**Other impacts not captured (especially temperature)**

**Smith et al (submitted)**

# Impacts of the NAO

**Real world**

GHGs ↓↓↓↓↓↓

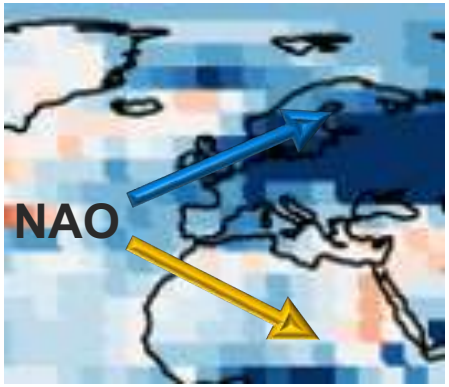



**Ensemble mean**

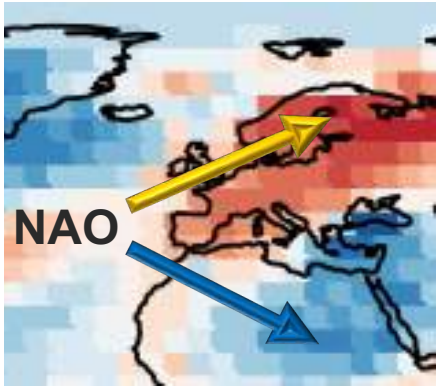
GHGs ↓↓↓↓↓↓


**Ensemble members**

GHGs ↓↓↓↓↓↓




GHGs ↓↓↓↓↓↓




**In reality NAO dominates over other factors (GHGs)**

**NAO too weak in ensemble mean → GHGs dominate**

**Chose ensemble members with the correct NAO magnitude**

Smith et al (submitted)

# NAO-matching

**In reality NAO dominates over other factors (GHGs)**

**NAO too weak in ensemble mean → GHGs dominate**

**Chose ensemble members with the correct NAO magnitude:**

**Compute ensemble mean NAO adjusted for underestimated signal**

**Compute NAO for each ensemble member**

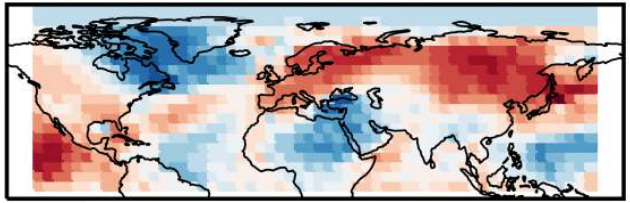
**Select ensemble members whose NAO is close to adjusted ensemble mean NAO**

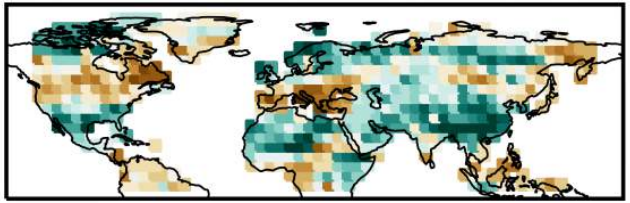
# Impacts of extreme NAO

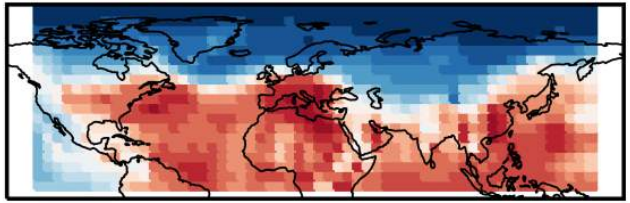
Temperature
(a) Observations

Precipitation
(b) Observations

Pressure
(c) Observations



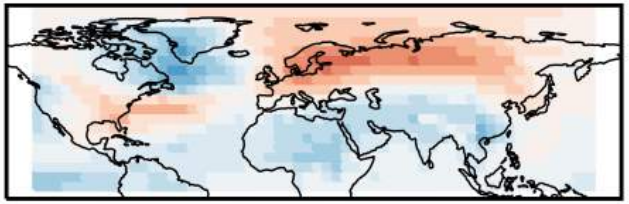


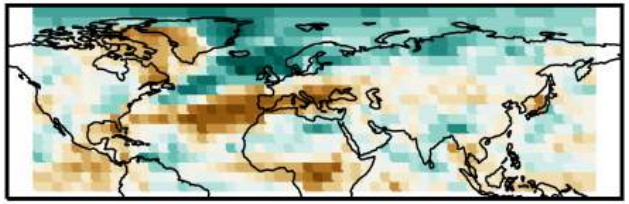


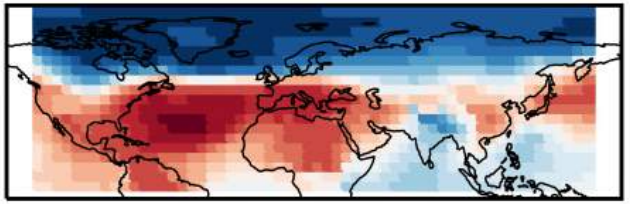
(j) NAO-matched

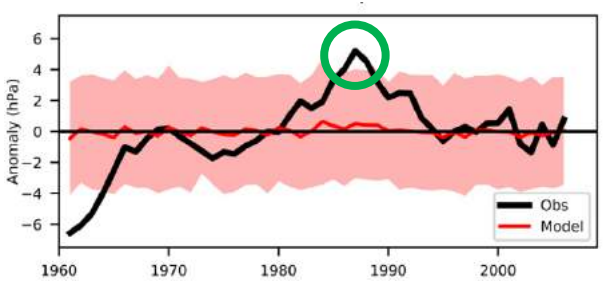
(k) NAO-matched

(l) NAO-matched







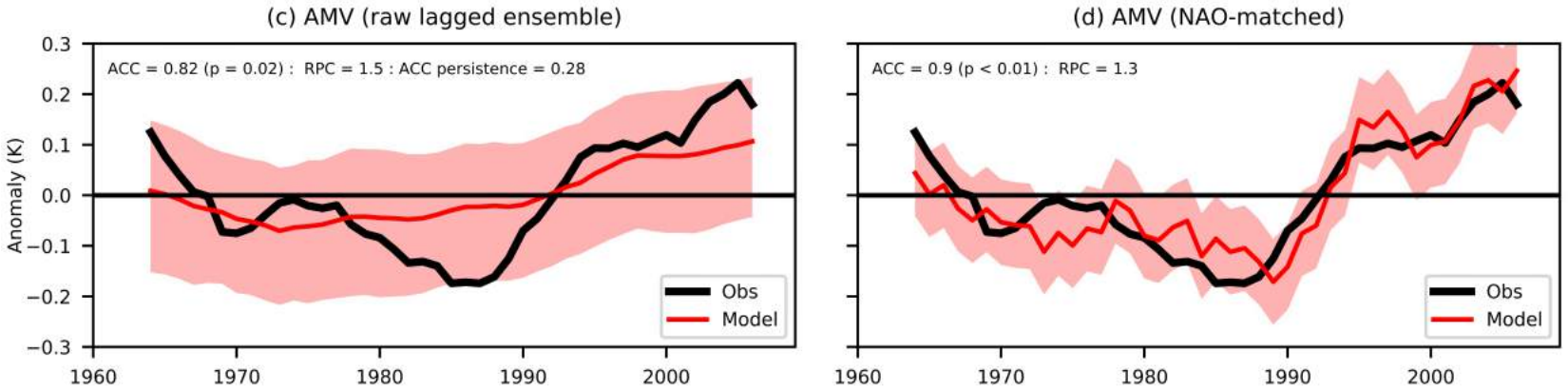



**Impacts captured very well by 20 ensemble members with correct NAO magnitude**

**“NAO-matching”**

**Smith et al (submitted)**

# Impacts of the NAO on Atlantic ocean




**NAO-matching clearly improves AMV predictions**

**AMV not the sole driver of the NAO**

Smith et al (submitted)

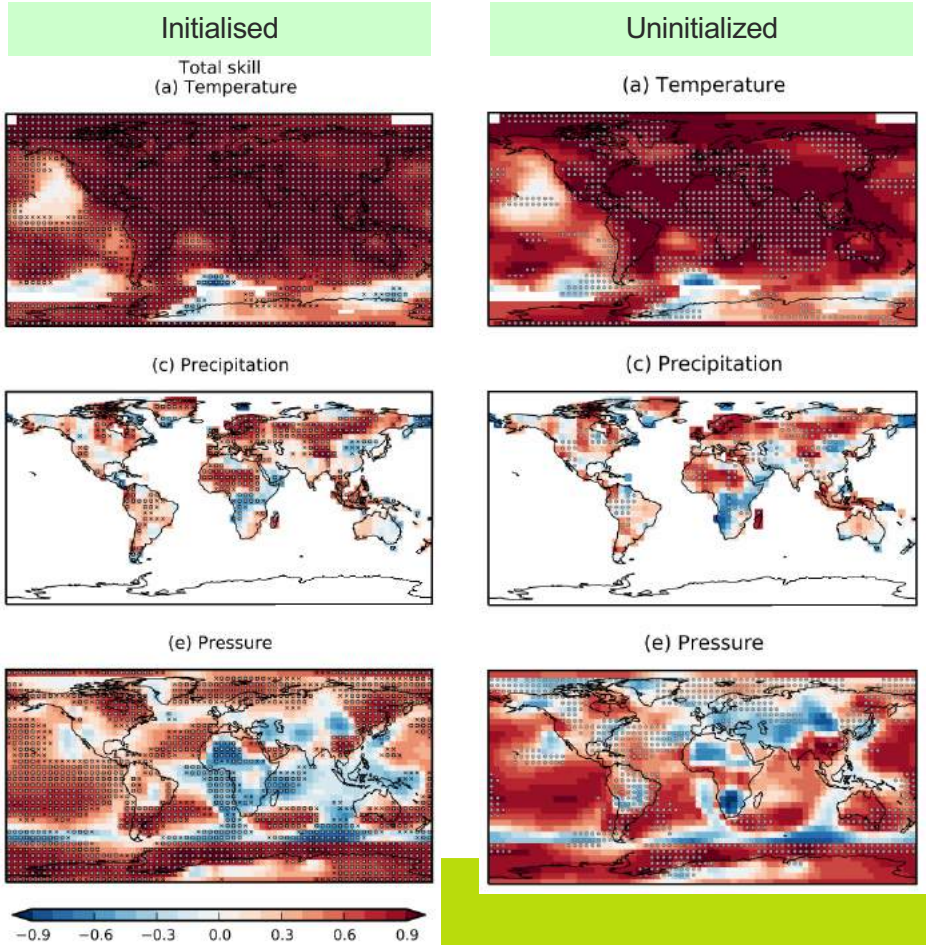
# Internal variability or external forcing?

**Significant skill using a large ensemble**

**Precipitation and pressure as well as temperature**

**Patterns of skill are captured by uninitialized simulations**

**Initialisation mainly improving the response to external forcings?**




Smith et al, 2019

# Summary

## The good news:

Climate is *much* more predictable than previously thought

NAO correlation skill = 0.8 in decadal predictions

## The bad news:

There is a serious deficiency in climate models

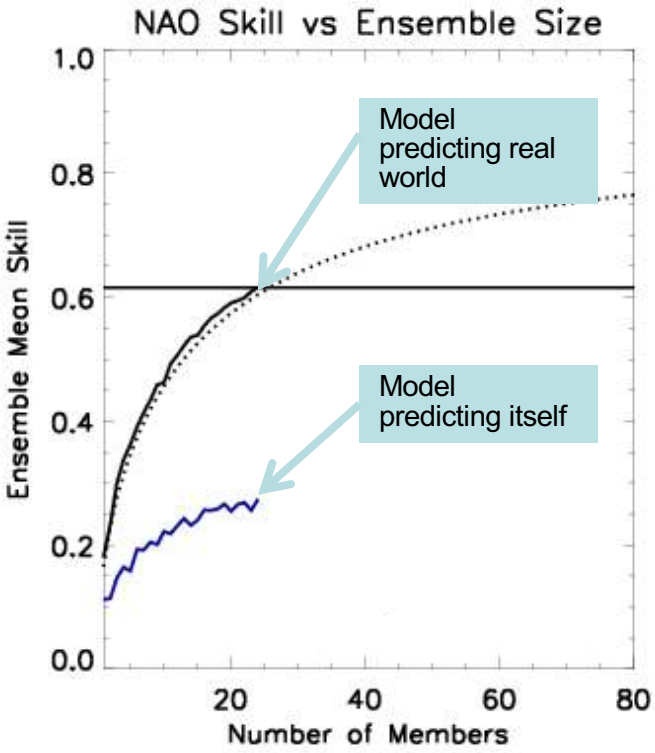
Predictable signal is an order of magnitude too small

What is the cause?

Does it affect projections beyond 10 years?

Need to understand the role of external forcings

# Signal to noise paradox: hypotheses



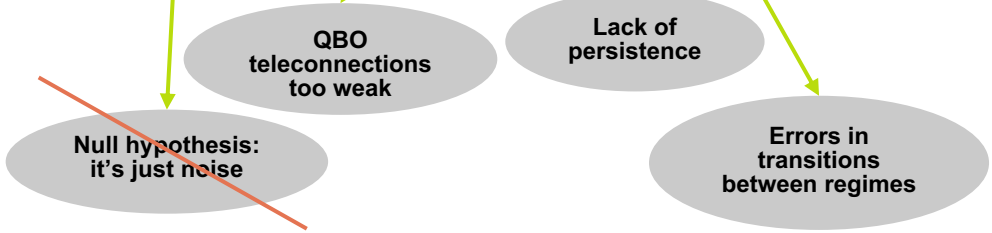

**Paradox: models predict the real world better than themselves despite perfectly representing themselves**

**Members NOT alternate realisations of obs**

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

**Undermines the basis of ensemble prediction**

**Scaife et al 2014, Eade et al 2014, Dunstone et al 2016, Siegert et al 2016, Weisheimer et al 2018, O'Reilly et al 2018, Strommen and Palmer 2018, Scaife and Smith 2018, Zhang and Kirtman 2019**



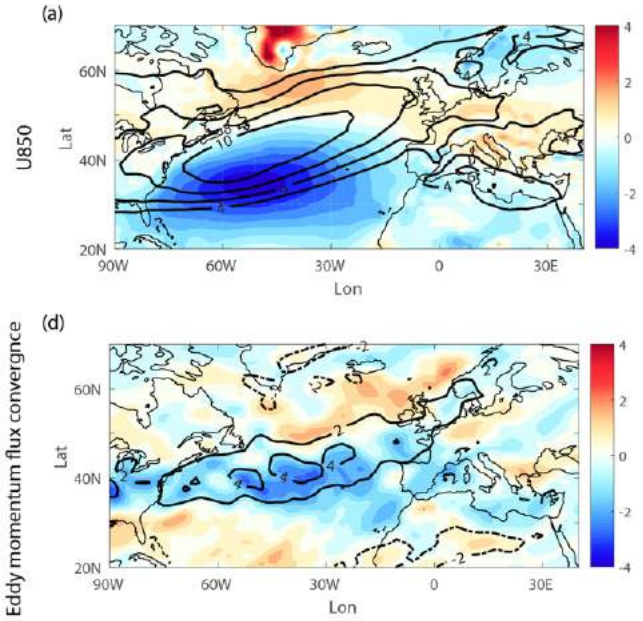
# Eddy feedback too weak?

**U wind**

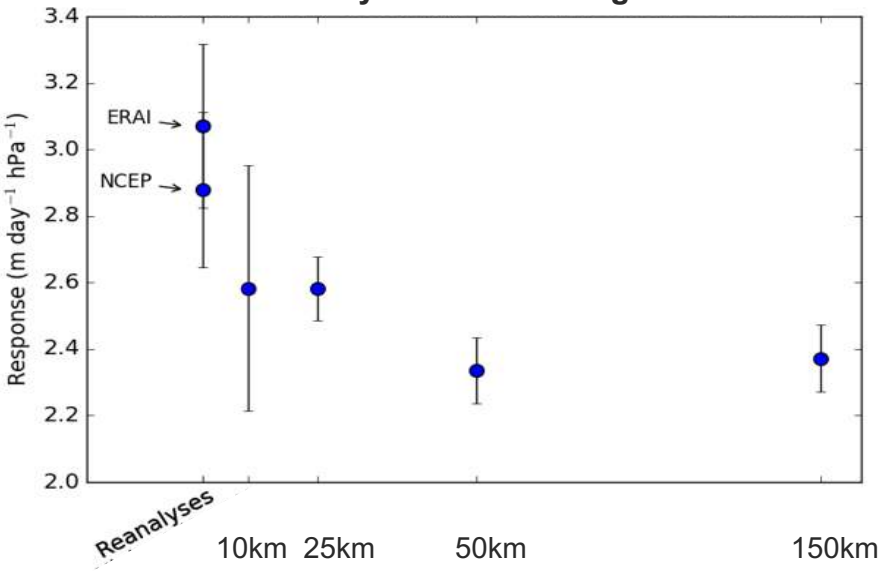
**Eddy momentum flux convergence**

$$F_y = -\frac{\partial(\overline{u'v'})}{\partial y}$$

**Model bias**




**Eddy feedback strength**




**← increasing model resolution**

**Eddy driven jet too weak**

**Eddy momentum flux convergence too weak**

**Eddy feedback weaker in models than reanalysis**

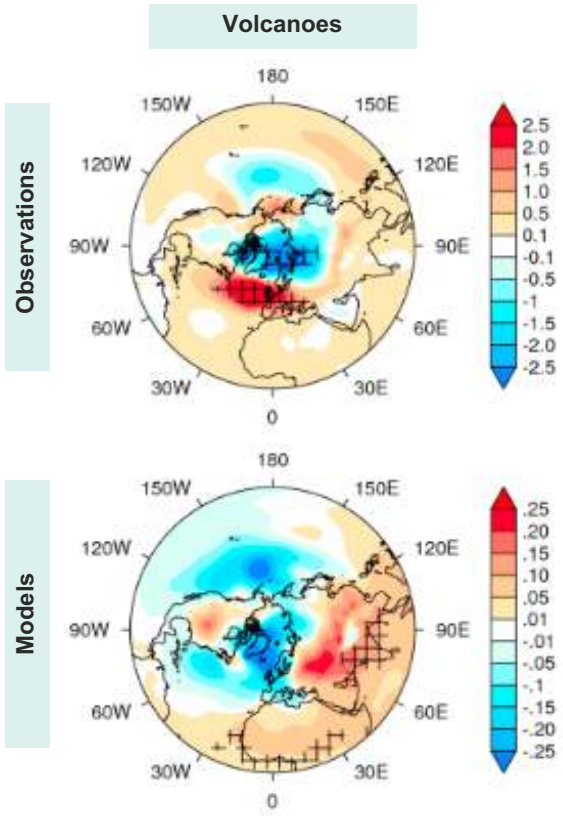
**Strengthens at higher resolution**

**May need 10km or higher resolution**

**Scaife et al, ASL, 2019**

# Signal to noise paradox in external drivers?

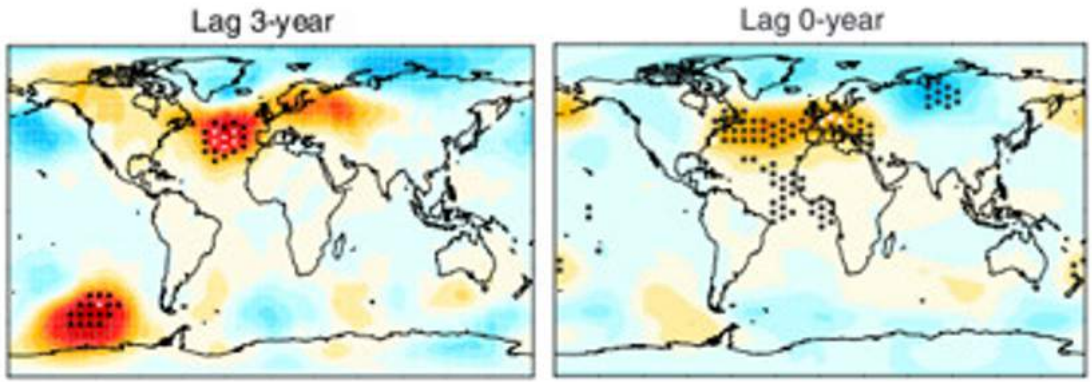
**Volcanoes**




Note different colour scales!

**Solar**

**Observations** | **Models**

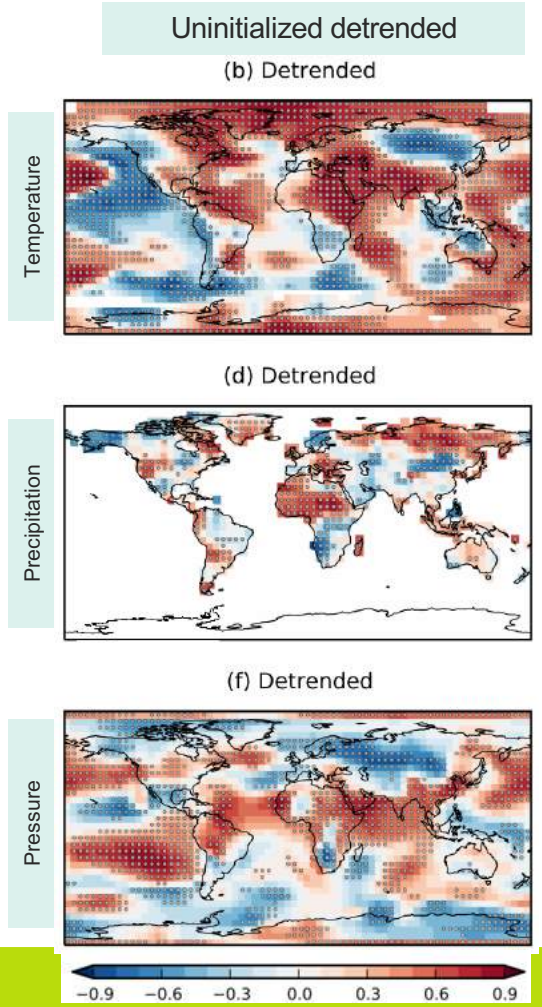



Model response is too weak, and not lagged

Driscoll et al 2013; Gray et al 2013

# Not just a trend

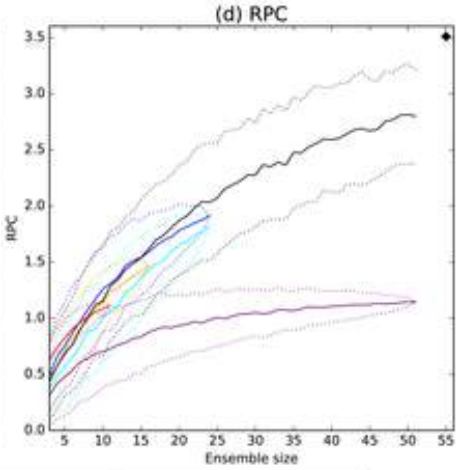
- Significant skill in many regions after detrending
- Highlights importance of non-GHG forcings
  - Anthropogenic aerosols
  - Volcanoes
  - Solar
  - Ozone




Smith et al, 2019

# A widespread issue

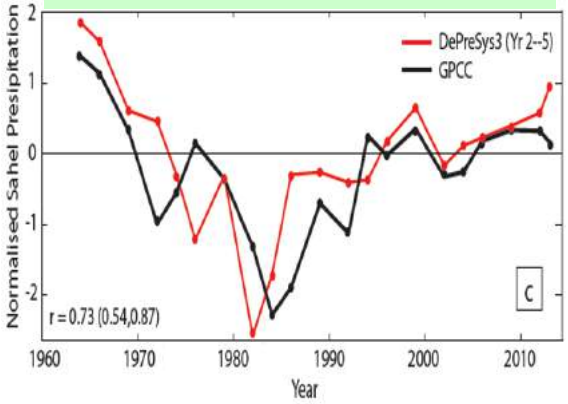
## Seasonal NAO




RPC > 1 in multiple seasonal forecasting systems

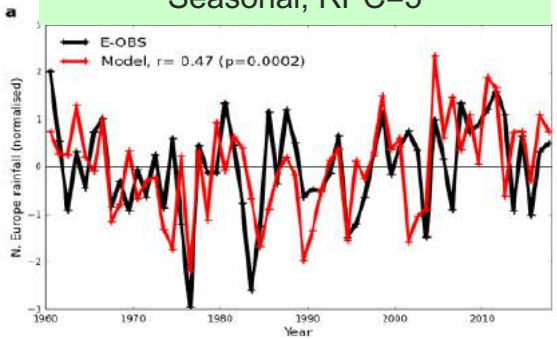
## Sahel rainfall

Years 2-5, RPC=1.2

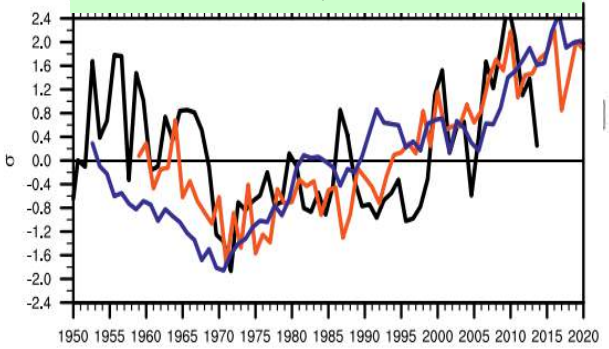



## European summer rainfall

Seasonal, RPC=5




Years 3-7, RPC=1.6



Baker et al 2018, Yeager et al 2018, Sheen et al 2017, Dunstone et al 2018