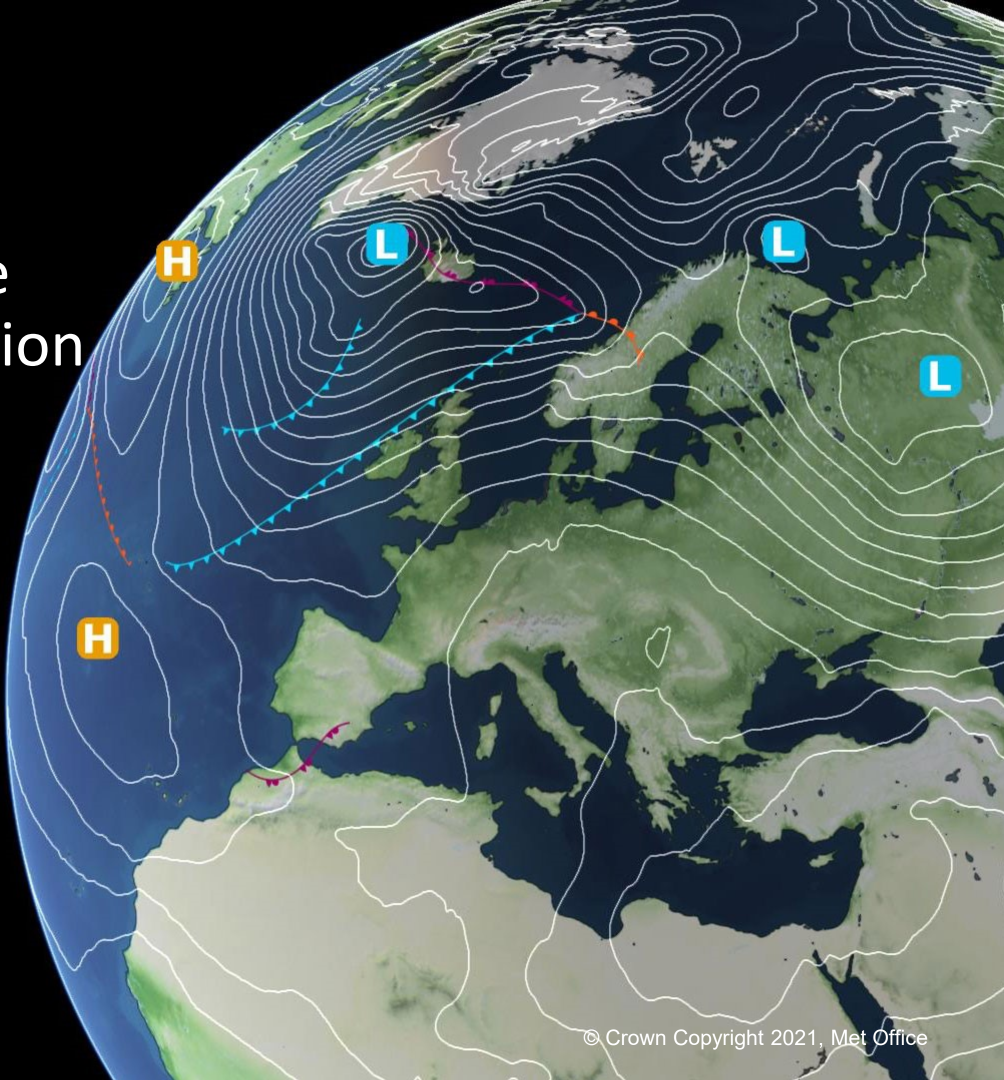
Met Office

The importance of confidence in observations to constrain climate projections used to inform adaption actions



David Sexton

Measurement for Climate Action Workshop , 13th October 2021

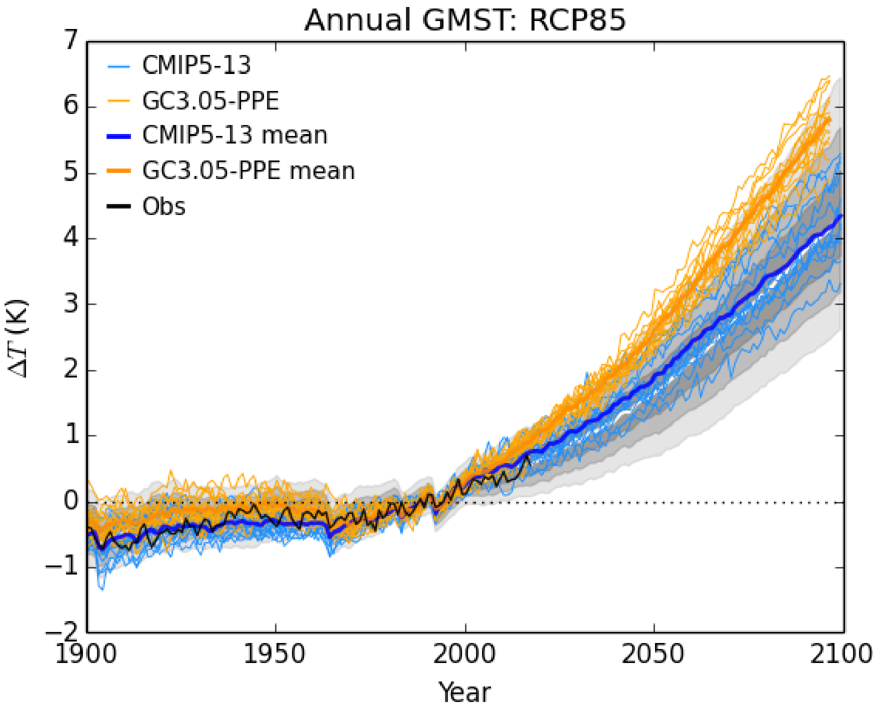
Particular thanks to Jill Johnson (now Sheffield University for work done at Leeds University)

Contents

- UK’s national climate projections
- Climate modelling and evaluation using observations
- Constraining climate projections
- Future directions for using observations for climate projections

Climate modelling and projections

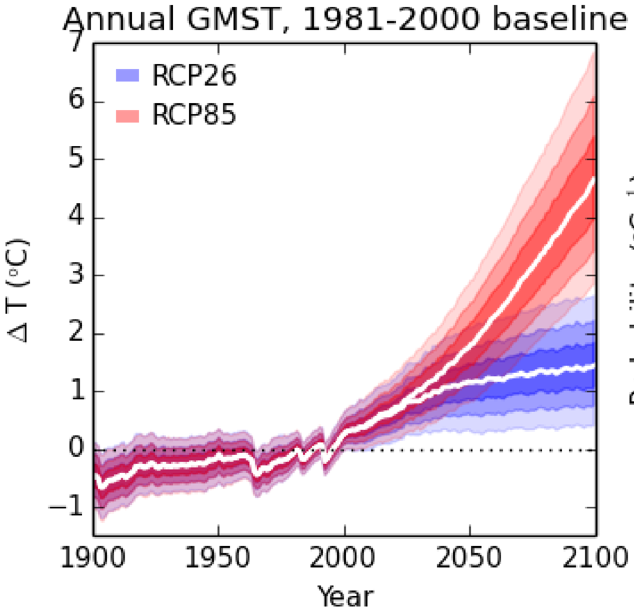
- Data like this from UK Climate Projections (UKCP18) are fed into impact studies (e.g. floods, heat extremes), which then inform the UK's climate change risk assessment, that informs the National Adaptation plan.
- Each of the thin lines are a **"plausible"** simulation of how rapidly global temperature might increase under a scenario of intensive fossil fuel emissions.
- But what is "plausible"? We **filter out** many alternative simulations based on carefully quantified comparisons with observations.

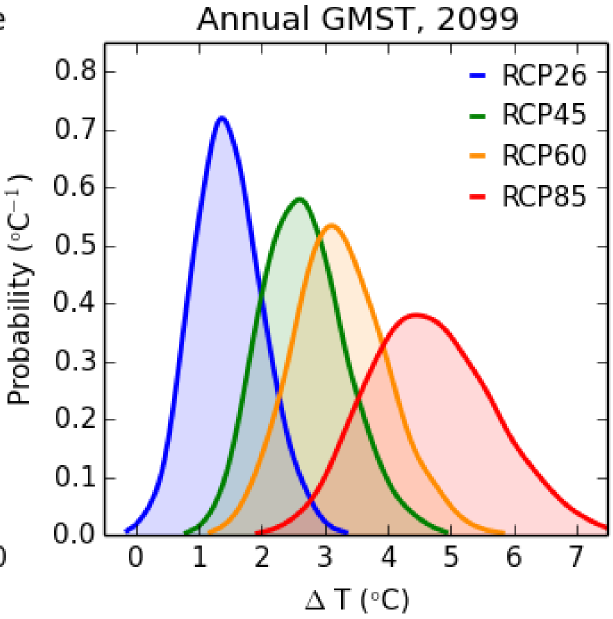


Probabilistic projections

Uses a Bayesian framework (Goldstein and Rougier 2004; Sexton et al 2012)

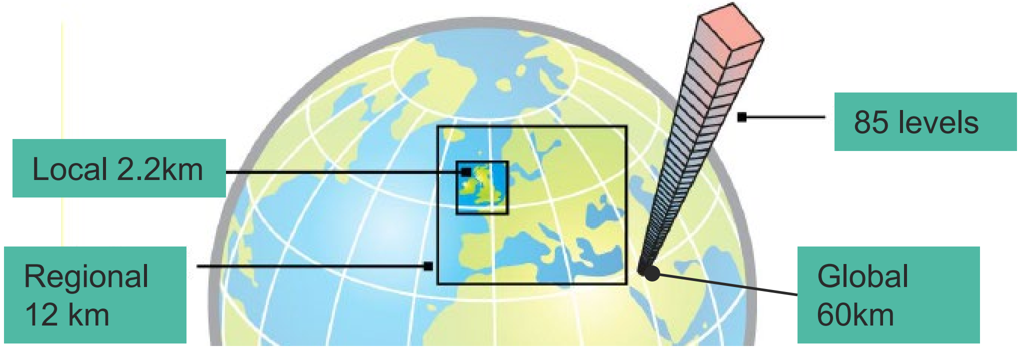
Combination of thousands of statistically generated pseudo-simulations combined with a **weighting** that measures likelihood of observable characteristics of the simulation given the observations.



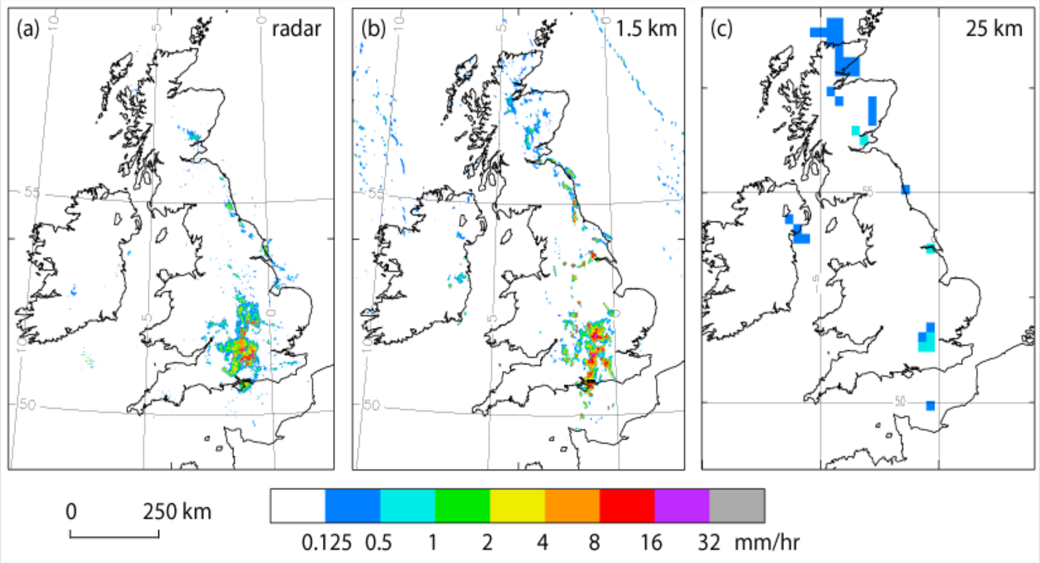


Climate modelling

- Climate models represent a range of processes, some resolved at horizontal/vertical resolution, some un-resolved and so parameterised (e.g. cloud microphysics).
- In our 60km global model, there are hundreds of variables defined at 140,000 grid boxes on each of 85 levels.
- For the fraction of variables observed, none will be observed as a grid box average.



Qualitative comparisons with observations to build confidence in climate simulations



Mesoscale convective system at 0000 UTC 14 June 2014.
(Clark et al, 2016, Meteorological Applications)

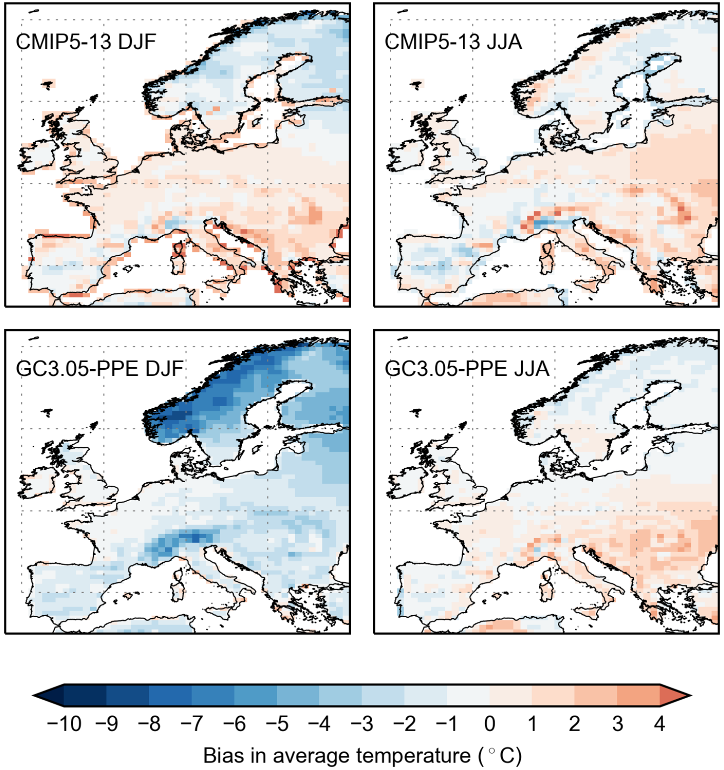
- In general, the radar data provide reliable information on the spatial patterns and temporal characteristics of rainfall.
- We have lower confidence in the absolute rainfall amounts especially for intense events and for hail events. Rain gauges are trusted more for rainfall intensity.

Harrison et al, 2000

Met Office

Biases in long-term averages

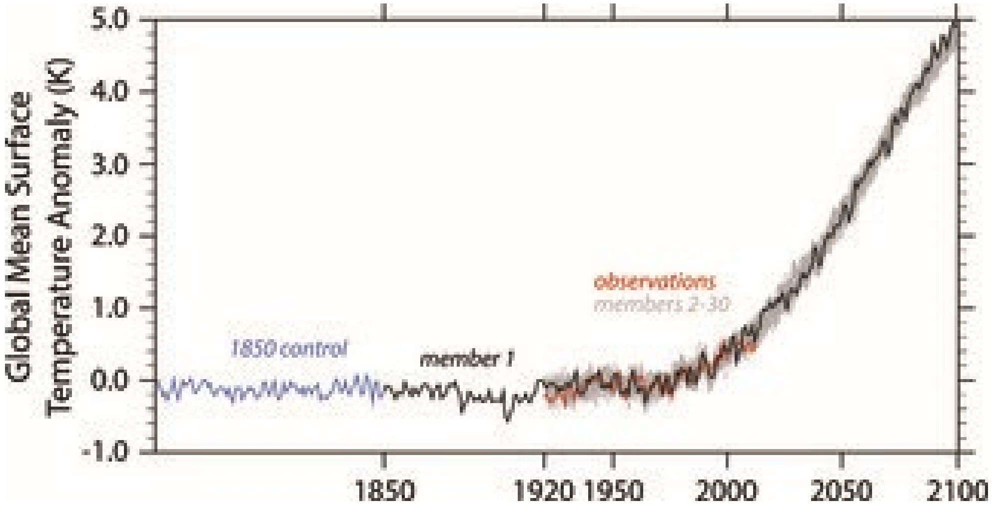
- Biases are differences in the long-term averages of modelled and observed data.
- Biases across different variables form the main basis for deciding whether or not a climate simulation is plausible. We also use historical trends.
- Here is an example using a gridded product of 1.5m temperature from EOBS.



Signal and noise

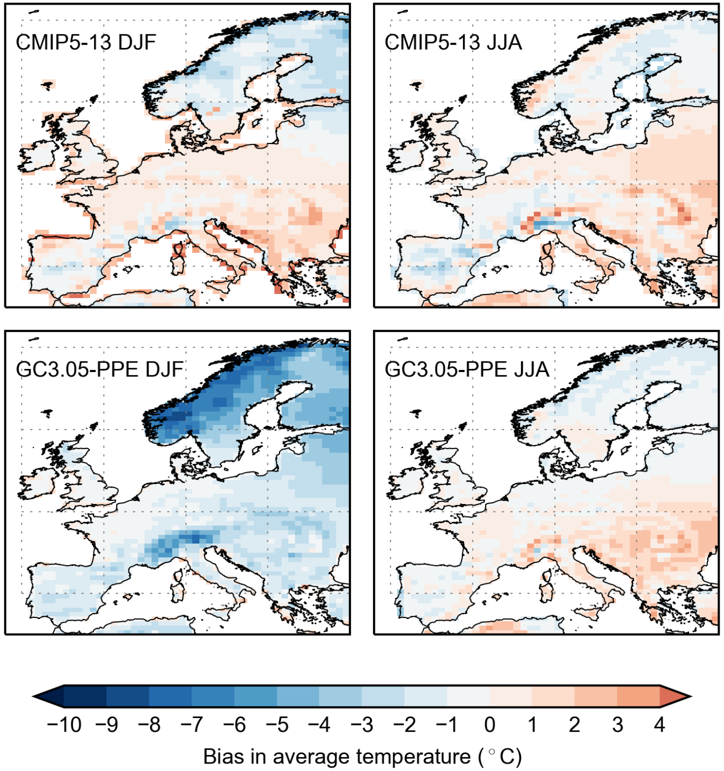
- Both real world and climate models have inherent variability due to chaotic nature of climate system.
- Simulations not designed to reproduce real-world noise, just the real-world signal
- There are other causes of error or uncertainty that we must account for:
    - In real world there is also measurement error and representation uncertainty
    - Models are imperfect and have structural differences caused by approximations and missing processes common to all our models

From LARGE ensemble of CESM projections run with multiple initial conditions (Kay et al 2015)



Impact of uncertain data on climate model evaluation

- We need our constraints to reflect that noisy data or imprecise measurements or poorly modelled climate processes **reduce** our ability to discern good models from poor ones.
- Spatial or time averaging help to reduce noise.
- We need a method which captures that improving precision of observations or our climate models helps to provide better constraints.



Met Office

The Implausibility Metric, I

Implausibility

Observations

Model output

$$I = \frac{|O - M|}{\sqrt{[Var(\epsilon) + Var(\psi) + Var(\phi) + Var(\delta)]}}$$

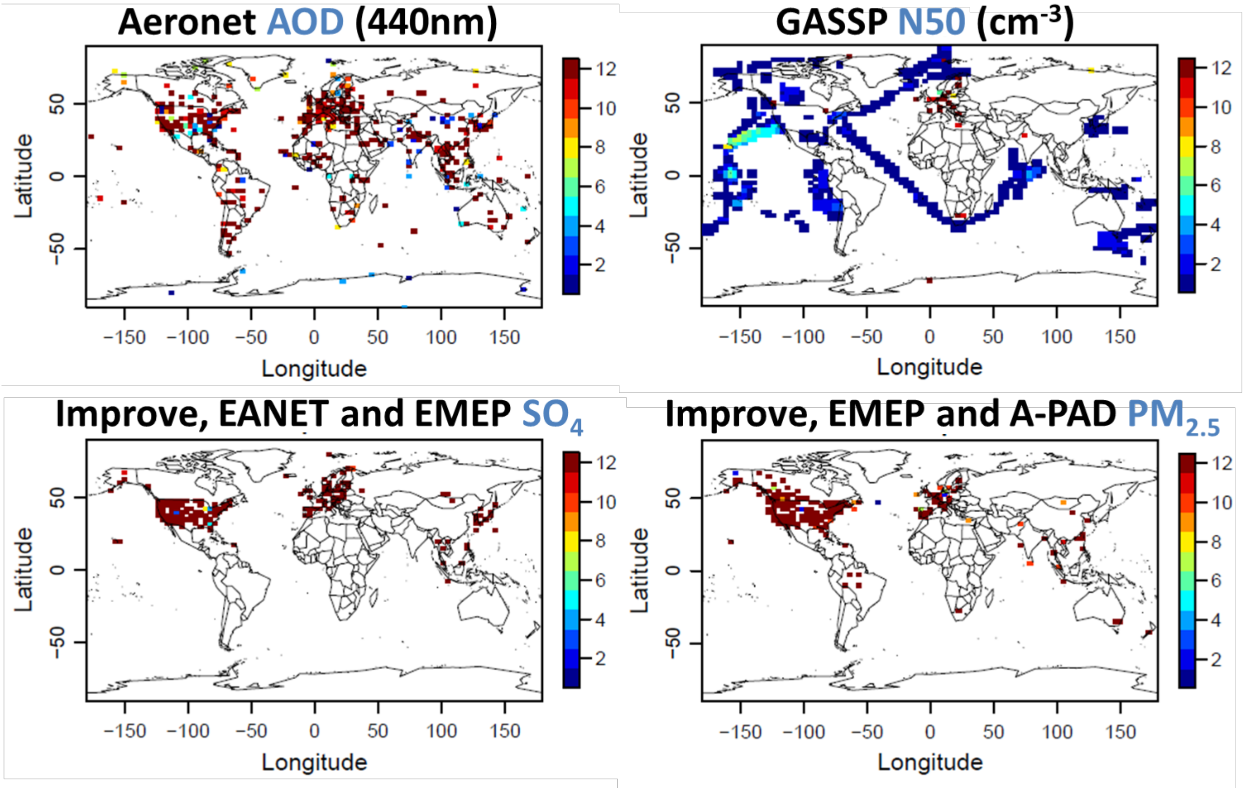
Measurement uncertainty

Representation uncertainty including real-world noise

Model noise

Structural uncertainty (or “discrepancy”)

Extensive aerosol observations to use for constraint



colour = monthly temporal coverage

An extensive set of aerosol observations, processed for the comparison:

- Varied global coverage: Spatially/temporally sparse
- Data from large networks (e.g. AERONET)
- Data from ship and Aircraft campaigns
- **9000+** observations

Johnson et al. (2020)

Components of representation uncertainty

UNIVERSITY OF LEEDS

- Spatial co-location
  - Comparing point measurements with the model grid
  - Where in the grid-box (central / edge) the observation lies

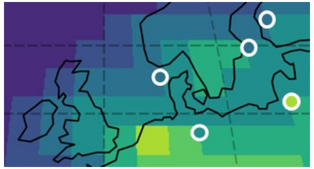
(Schutgens *et al*, 2016a)

- Temporal co-location
  - Comparing campaign data (measured over a few hours/days) to monthly mean model output

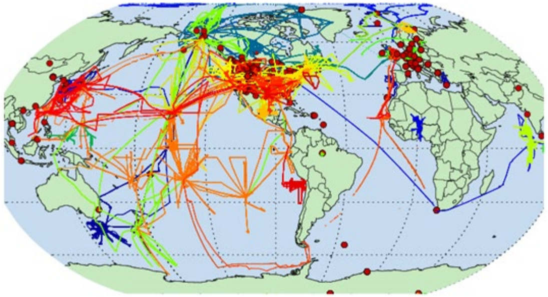
(Schutgens *et al*, 2016b)

- Inter-annual variability
  - Campaigns are 'one-off' studies
  - Comparing observations taken in a particular year to model output of a different year

Model v's Observations resolution

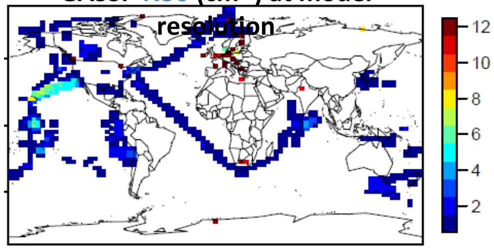


GASSP observations

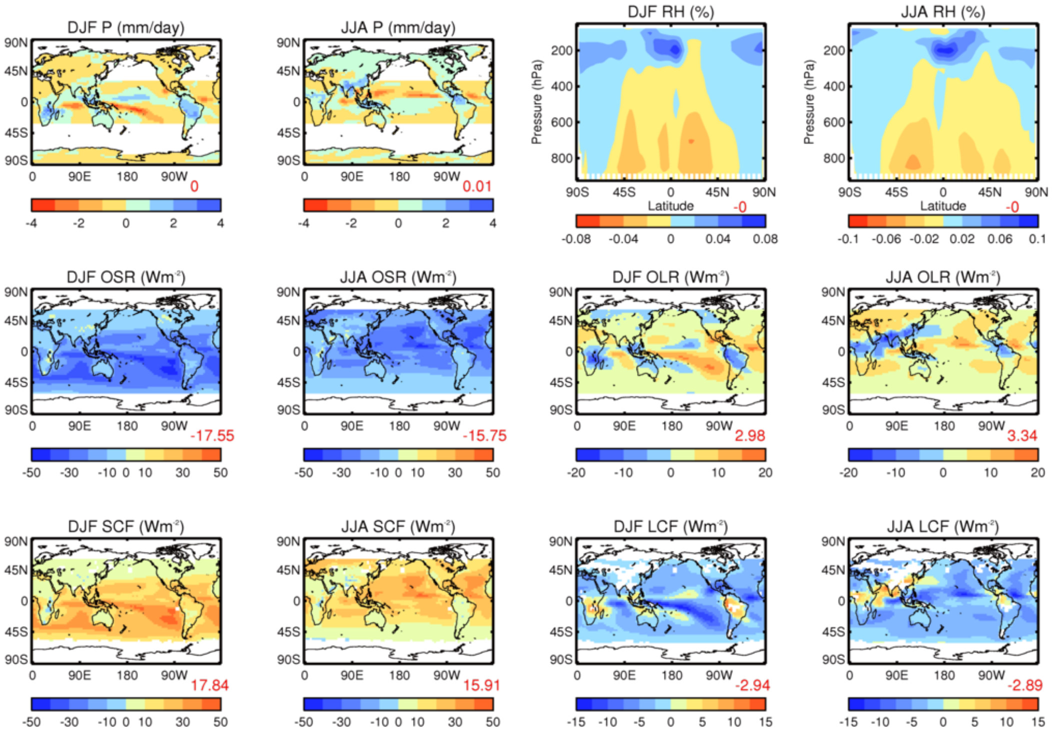


Reddington *et al*, 16

GASSP N50 (cm^{-3}) at model resolution



Can use several metrics to constrain climate projections



- First of six metrics used in Sexton et al (2012) and UKCP09
- More metrics, less chance for rewarding a poor model
- Where possible use two data sets to represent observational uncertainty from measurements.

The Implausibility Metric, I

Filtering: if I > specified threshold for certain fraction of observables, rule out simulation.

Weighting: weight is proportional to exp(- I^2)

Observ

Implausibility

$$I = \frac{|O - M|}{\sqrt{[Var(\epsilon) + Var(\psi) + Var(\phi) + Var(\delta)]}}$$

Measurement uncertainty

Representation uncertainty including real-world noise

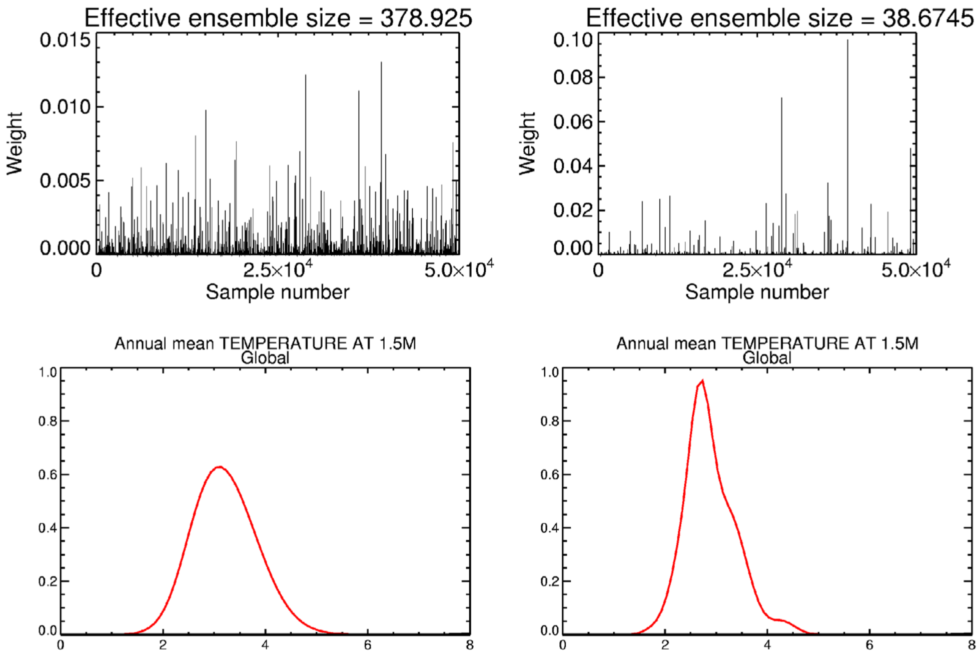
Model noise

Structural uncertainty (or “discrepancy”

Effect of discrepancy on weighting

If we do not factor in uncertainties like measurement error, we end up with over-confident projections

Discrepancy included **excluded**

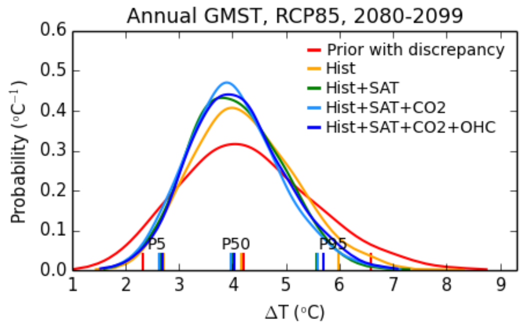


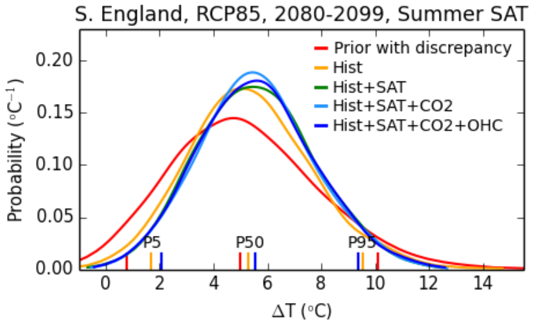
Sexton et al 2012

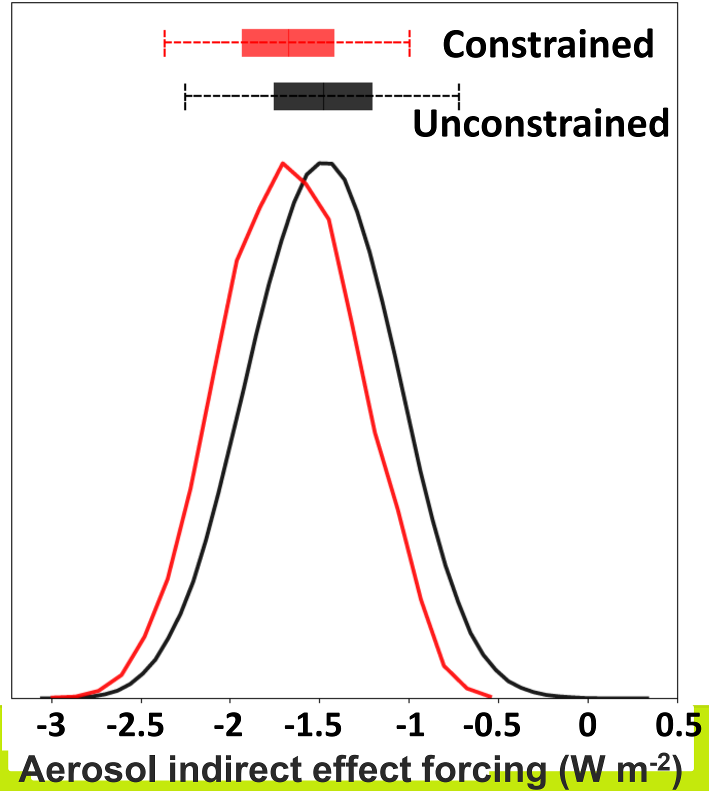
Met Office

The effect of constraining projections

The 95% CI on the aerosol forcing reduces by 11%.





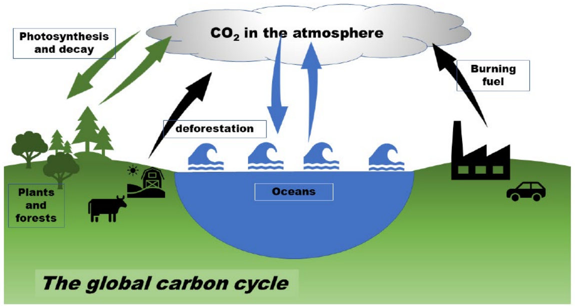


Summary

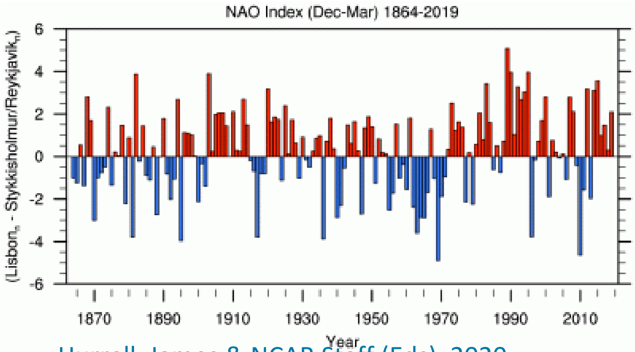
- Observations of multiple variables are used to either filter out implausible climate simulations or weight simulations according to goodness-of-fit
- Large errors in observations are one contributory factor that reduces our ability to use those observations to discern a good climate simulation from a relatively poor one.
- Our methods capture the benefit of improving precision of observations, (likely) leading to a more tightly constrained projection, more targeted adaptation.

The future of using observations for climate projections

- As well as biases in average climate, more emphasis on coupled Earth System and process evaluation.

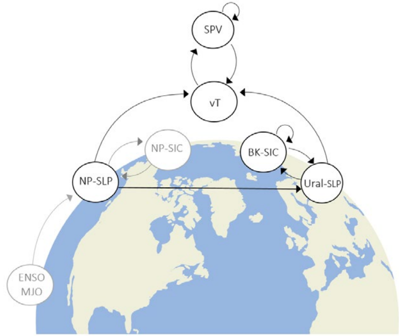


Earth system processes



Hurrell, James & NCAR Staff (Eds). 2020

Variability in dynamic drivers



Kretschmer et al 2020

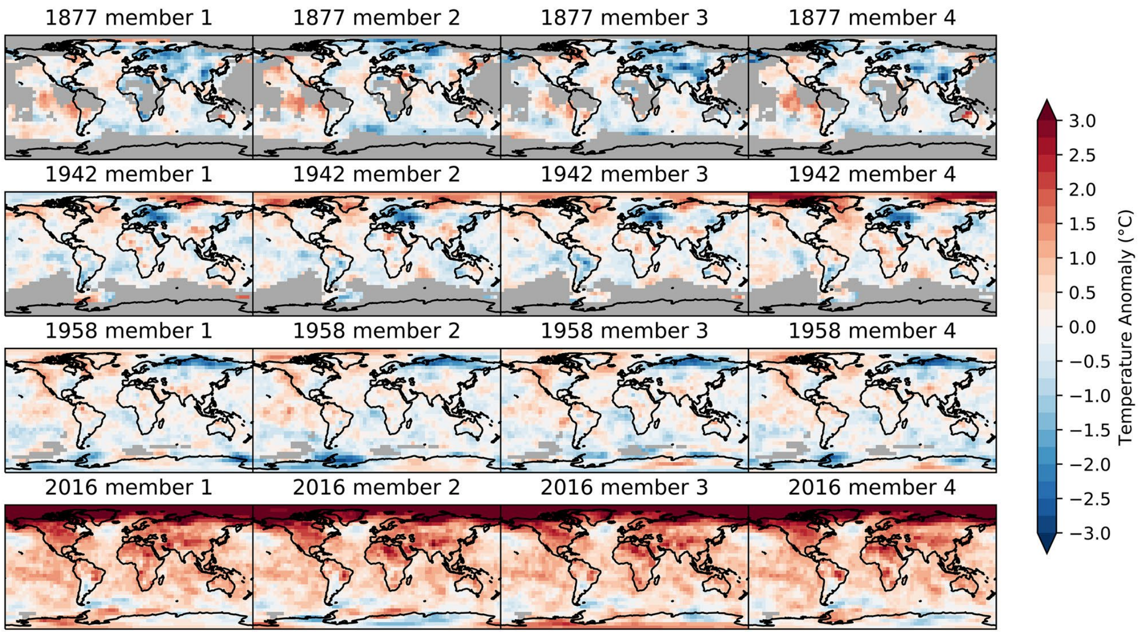
Connections between drivers

- This implies need for wider range of variables, consistency across time and space, greater granularity, preferably with quantified observational errors...

Met Office

HadCRUT5 - A useful way to provide observational data and associated uncertainties

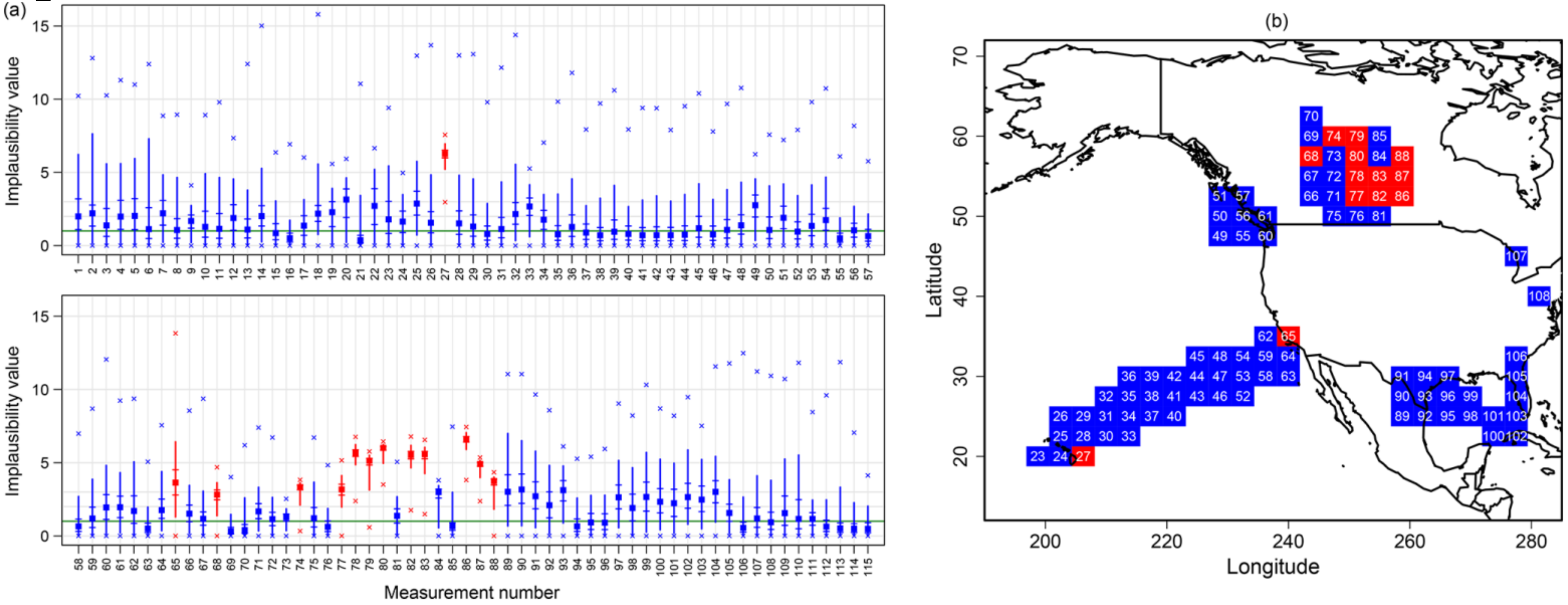
- HadCRUT5 (Morice et al 202) combines sea surface temperature with near surface temperature over land.
- Consists of a 200-member ensemble of realisations sampling measurement and representation uncertainty plus other uncertainty information provided.



Backup slides

Identifying observations that do not compare well

We remove all observations if the lower 95% credible interval bound on *I* (across variants) is >1



It can be difficult to pin-point the cause: Are the mis-matches due to **representation errors**?

Or, are they **indicators of structural errors** in the model?

Fig 4, Johnson et al. (2020)

Met Office

Bayes Linear synthesis of multiple lines of evidence

- UKCP09 based on a Bayesian methodology of Goldstein and Rougier 2004
- Large multivariate problem
    - Model parameters (X)
    - Historical and future model output (m_h,m_f)
    - True climate (y_h,y_f)
    - Observations (o)
    - Model imperfections = discrepancy (d)
- Best input assumption - Model not perfect so there are processes in real system but not in our model that could alter model response by an uncertain amount. We assume that one choice of these values, x*, is better than all others. Any point in parameter space has a probability of being x* so we need to sample parameter space

$$y = f(x^*) + d$$

True climate

Model output of best choice of parameter values x*

Discrepancy

d=0 for perfect model