

Is CESM the best Earth System Model?

Reto Knutti

Institute for Atmospheric and Climate Science

ETH Zurich, Switzerland

reto.knutti@env.ethz.ch

Is CESM
the best Earth System Model?

Yes

Is CESM
the best Earth System Model?

Yes

No

Is CESM
the best Earth System Model?

Yes

No

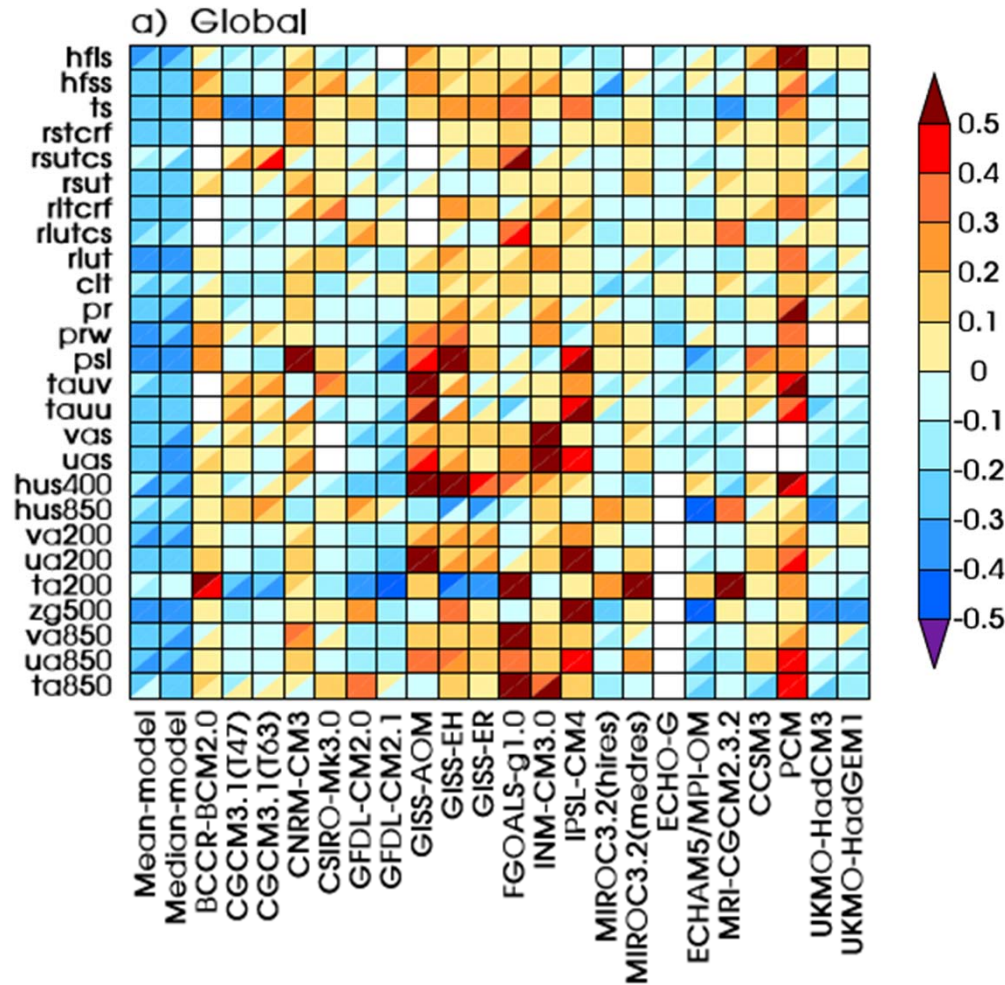
Stupid question

Is CESM

the best Earth System Model?

- What is the purpose, and is the model adequate for that purpose?
- What means “best” anyway?
- What is the evidence that a model is doing the right thing?
- Why are predictions uncertain?
- How well can we quantify uncertainty?
- How do we combine evidence from different models and observations?
- Why is it so hard, and are we making progress?

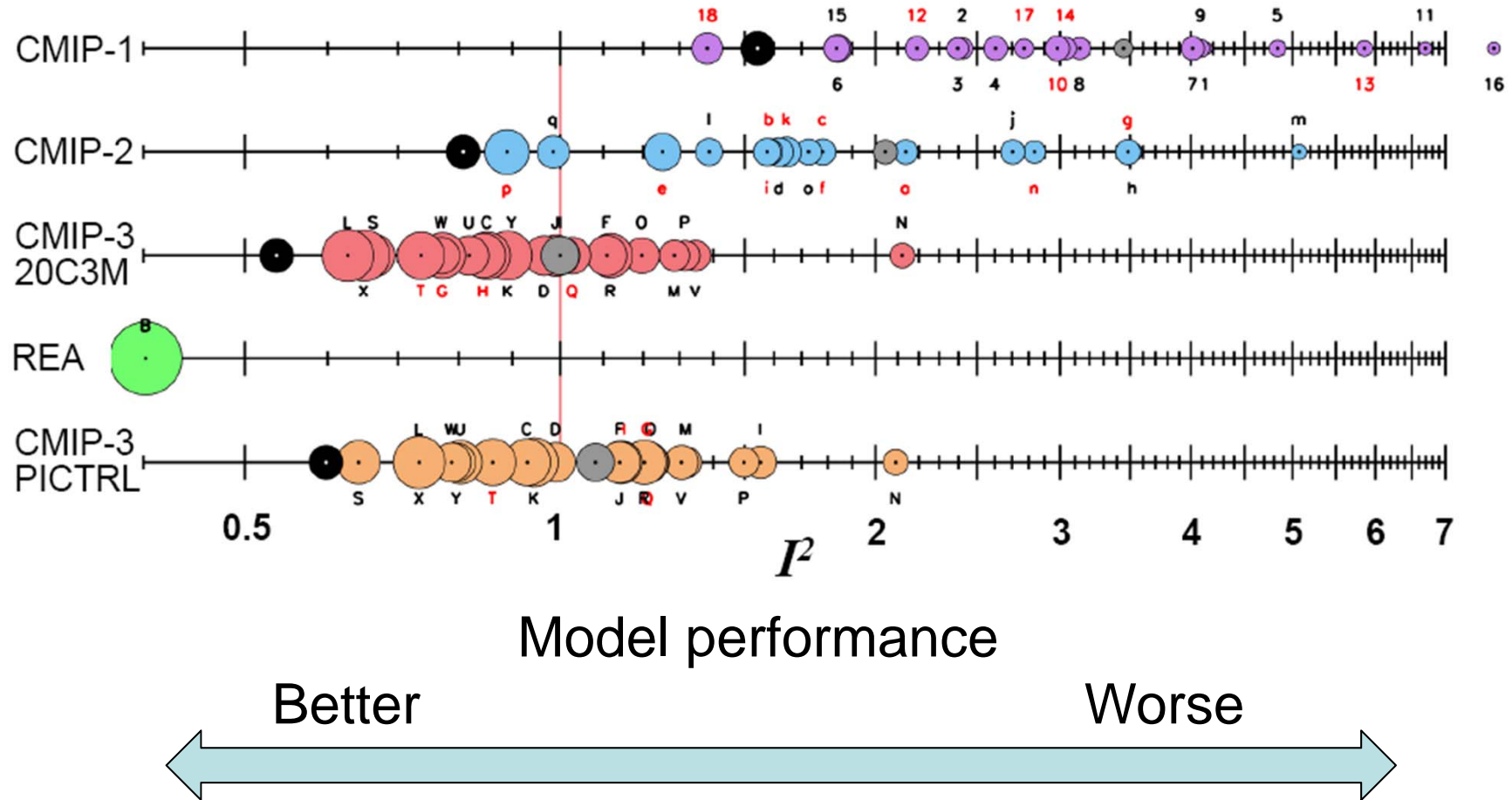
How do we measure model performance?



Performance metric
 Measure of agreement between model and observation

Model quality metric
 Measure designed to infer the skill of a model for a specific purpose

My model is better than your model



Confidence in climate models

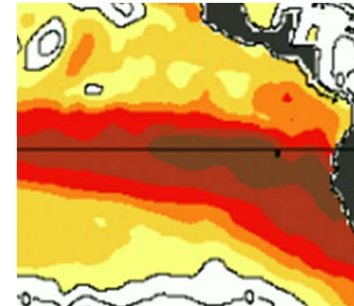
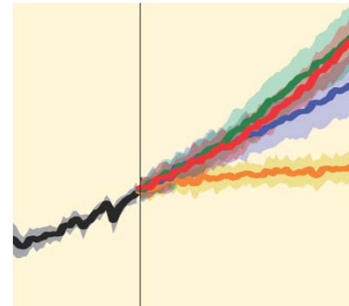
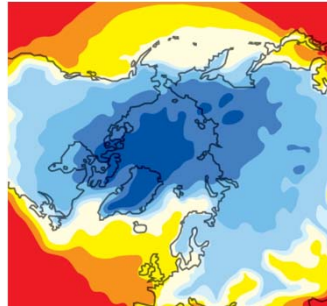
What is the evidence?

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0$$

$$\frac{\partial(\rho u)}{\partial t} + \frac{\partial(\rho u^2)}{\partial x} + \frac{\partial(\rho uv)}{\partial y} + \frac{\partial(\rho uw)}{\partial z}$$

$$\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho uv)}{\partial x} + \frac{\partial(\rho v^2)}{\partial y} + \frac{\partial(\rho vw)}{\partial z}$$

$$\frac{\partial(\rho w)}{\partial t} + \frac{\partial(\rho uw)}{\partial x} + \frac{\partial(\rho vw)}{\partial y} + \frac{\partial(\rho w^2)}{\partial z}$$





Physical principles

Reproduce climate

Reproduce trends

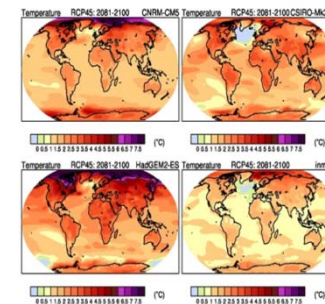
Processes

Sat	Sun
	
Clear	Rain
43°/32° Precip 10%	45°/39° Precip 30%

Weather



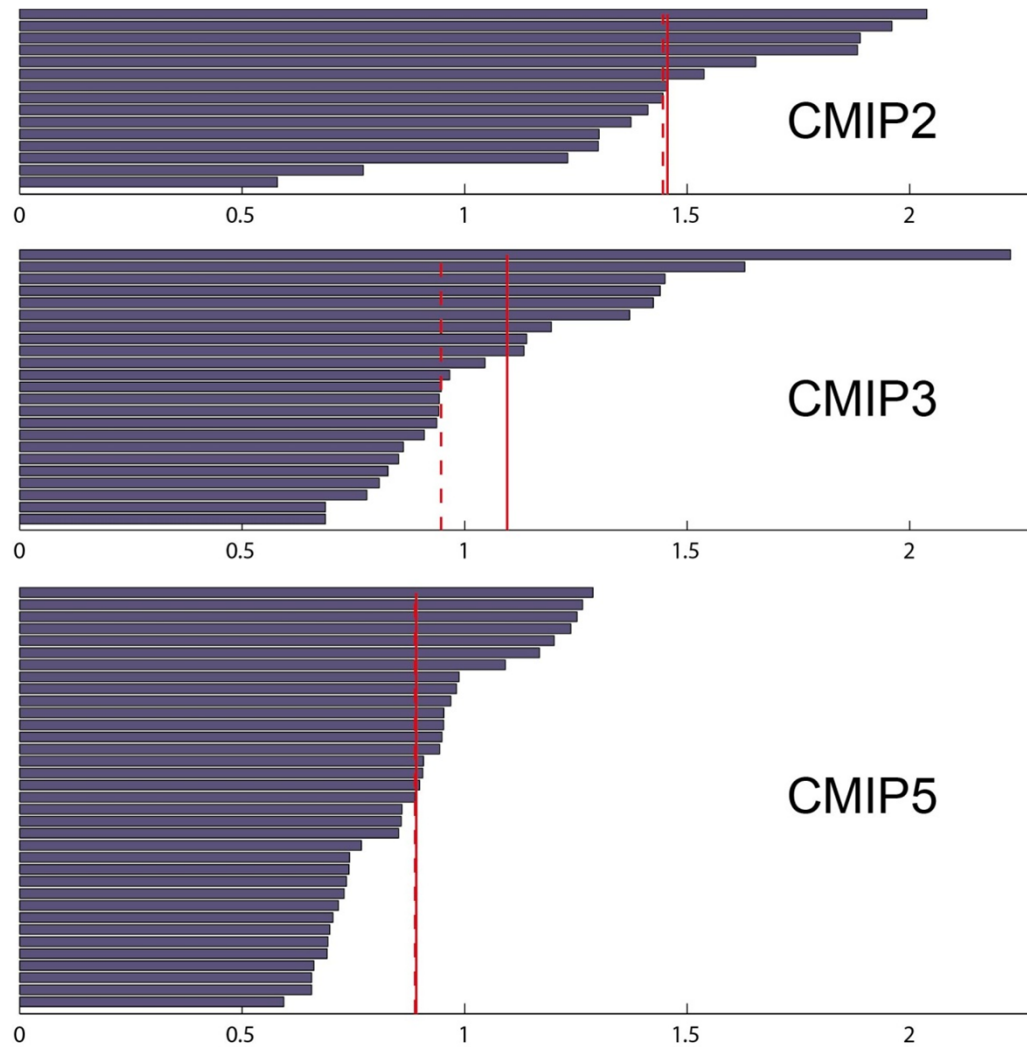
Past climate



Robustness

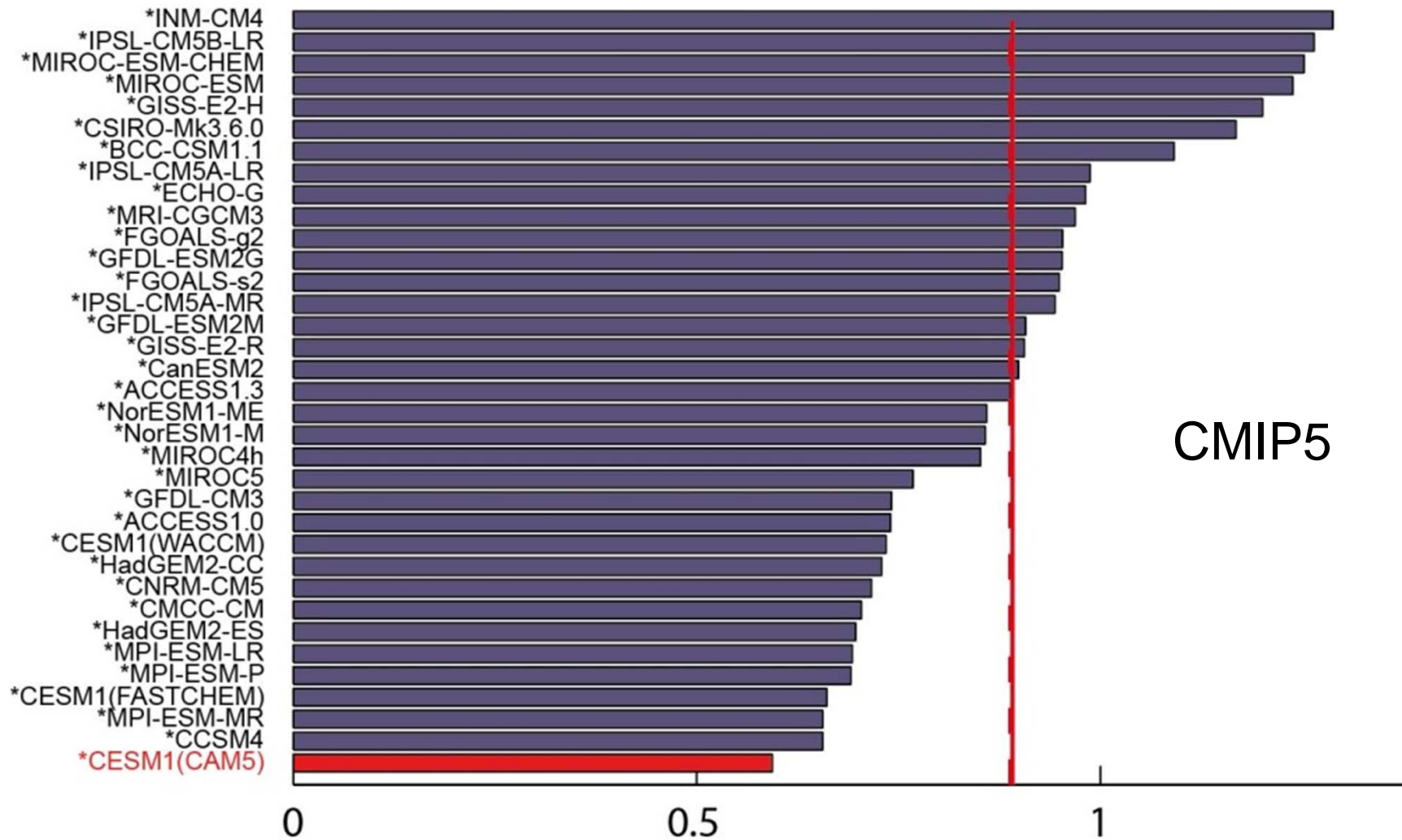
Distance to observations (tas/pr)

Is CESM the best model?



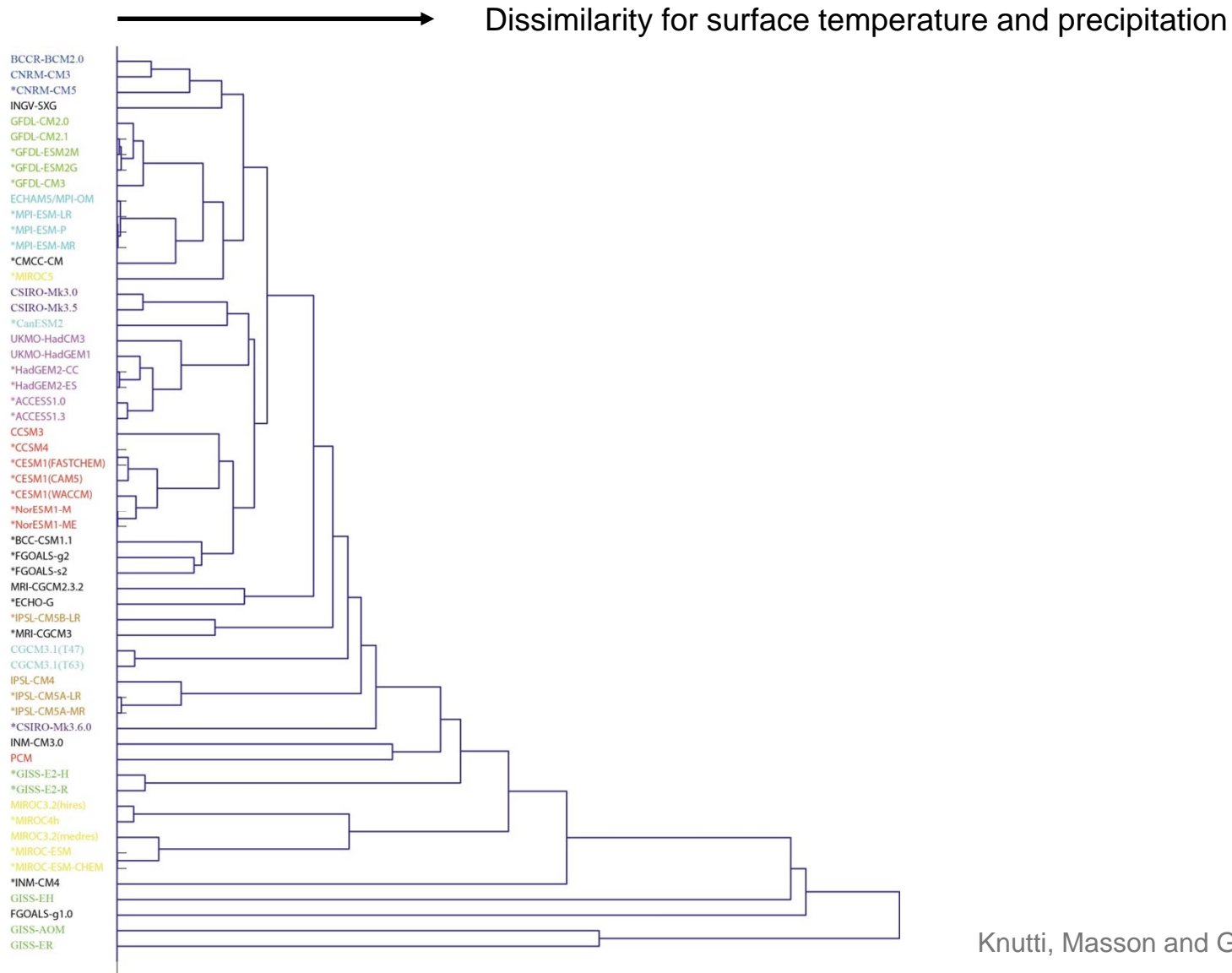
Distance to observations (tas/pr)

Is CESM the best model?



Climate model genealogy

Models are not independent

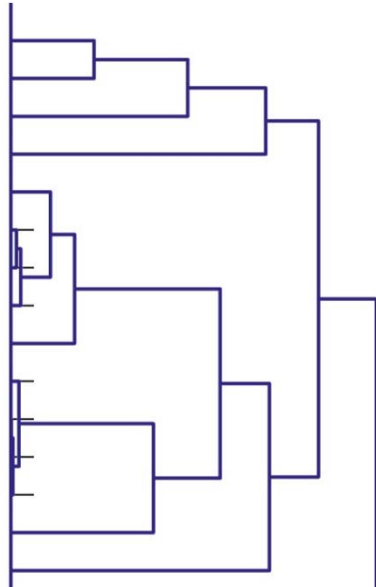


Knutti, Masson and Gettelman, in preparation

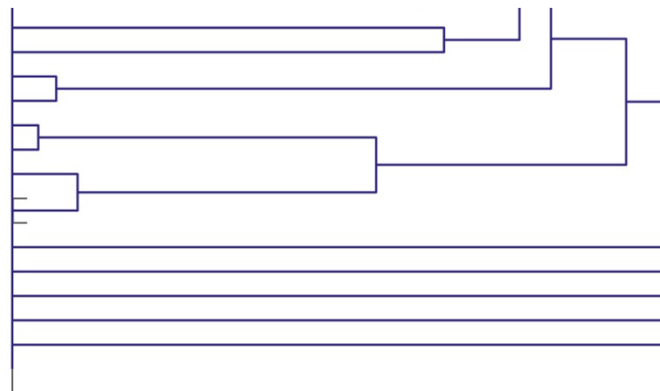
Climate model genealogy

Models are not independent

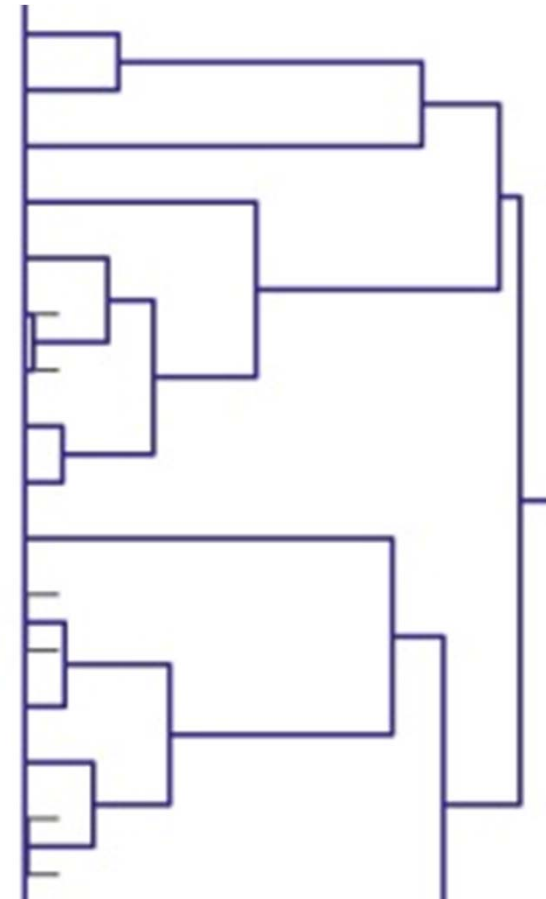
BCCR-BCM2.0
 CNRM-CM3
 *CNRM-CM5
 INGV-SXG
 GFDL-CM2.0
 GFDL-CM2.1
 *GFDL-ESM2M
 *GFDL-ESM2G
 *GFDL-CM3
 ECHAM5/MPI-OM
 *MPI-ESM-LR
 *MPI-ESM-P
 *MPI-ESM-MR
 *CMCC-CM
 *MIROC5



INM-CM3.0
 PCM
 *GISS-E2-H
 *GISS-E2-R
 MIROC3.2(hires)
 *MIROC4h
 MIROC3.2(medres)
 *MIROC-ESM
 *MIROC-ESM-CHEM
 *INM-CM4
 GISS-EH
 FGOALS-g1.0
 GISS-AOM
 GISS-ER

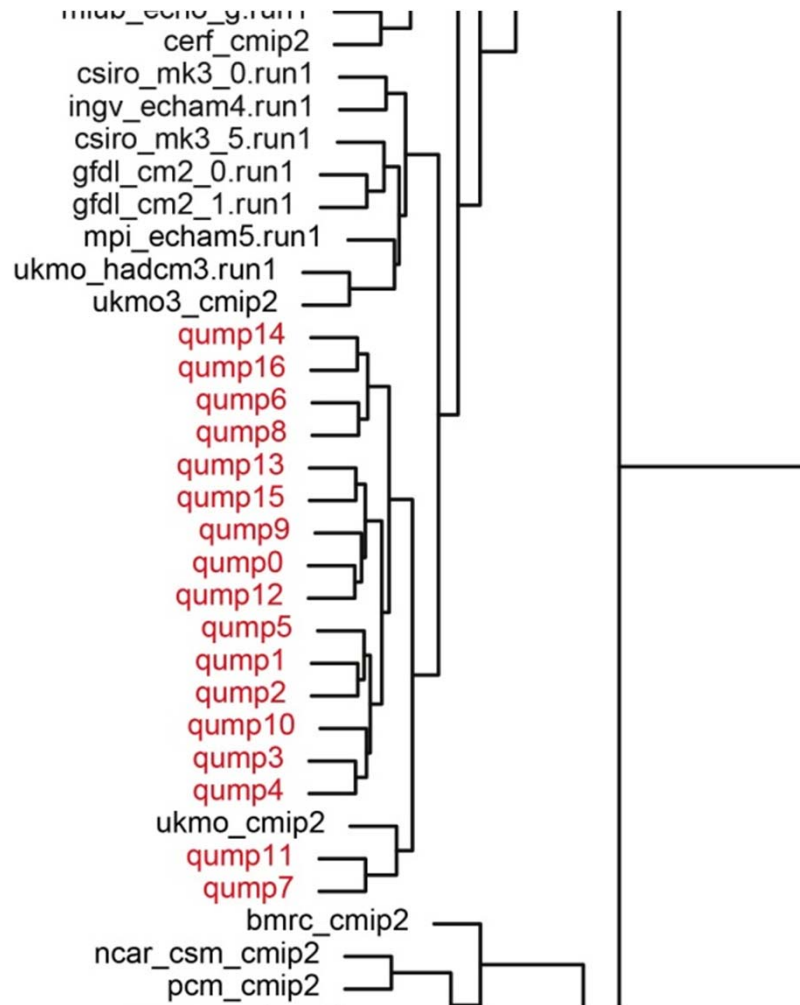


CSIRO-Mk3.0
 CSIRO-Mk3.5
 *CanESM2
 UKMO-HadCM3
 UKMO-HadGEM1
 *HadGEM2-CC
 *HadGEM2-ES
 *ACCESS1.0
 *ACCESS1.3
 CCSM3
 *CCSM4
 *CESM1(FASTCHEM)
 *CESM1(CAM5)
 *CESM1(WACCM)
 *NorESM1-M
 *NorESM1-ME



Climate model genealogy

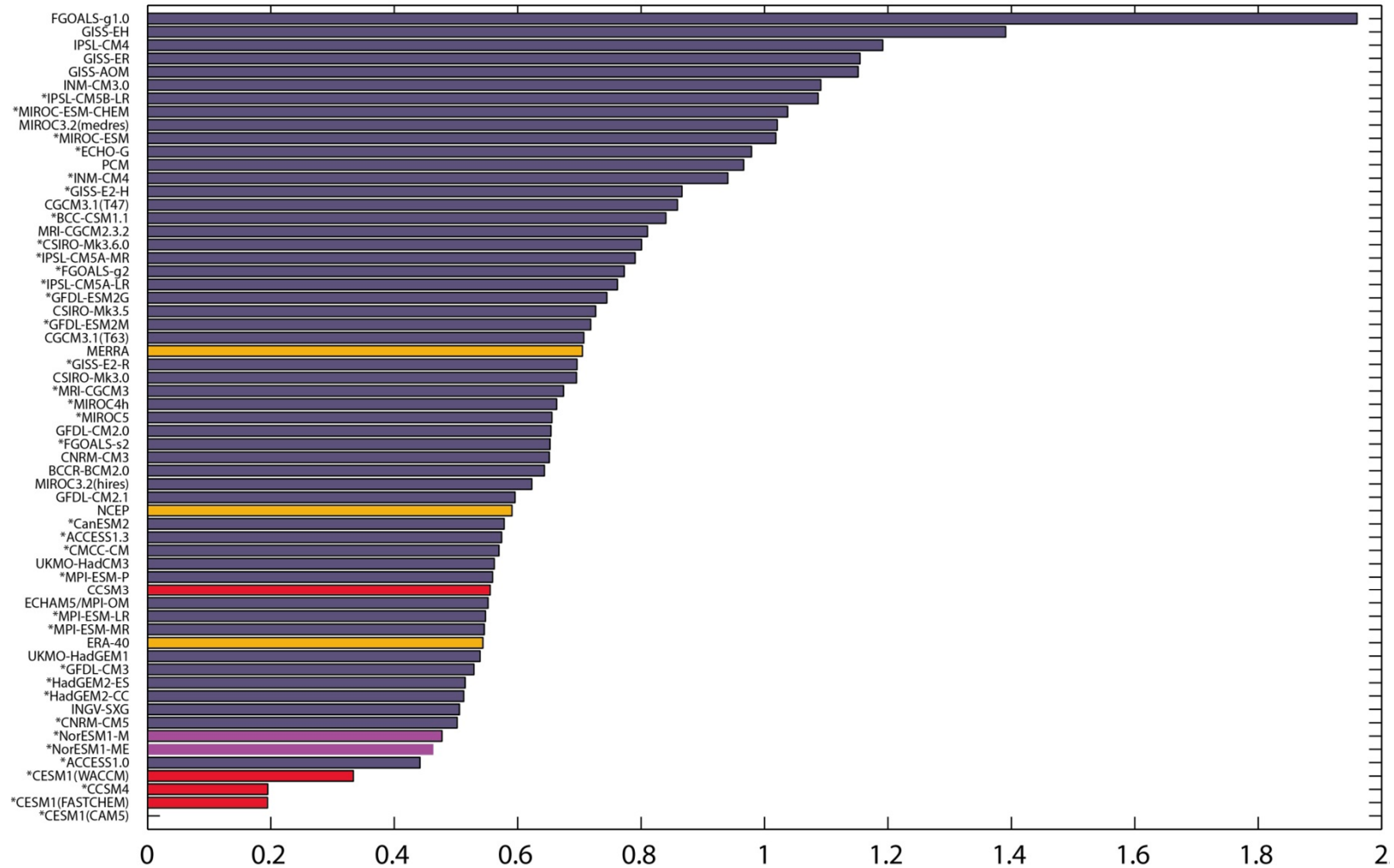
Models are not independent



Climate model genealogy

Distance from CESM1(CAM5)

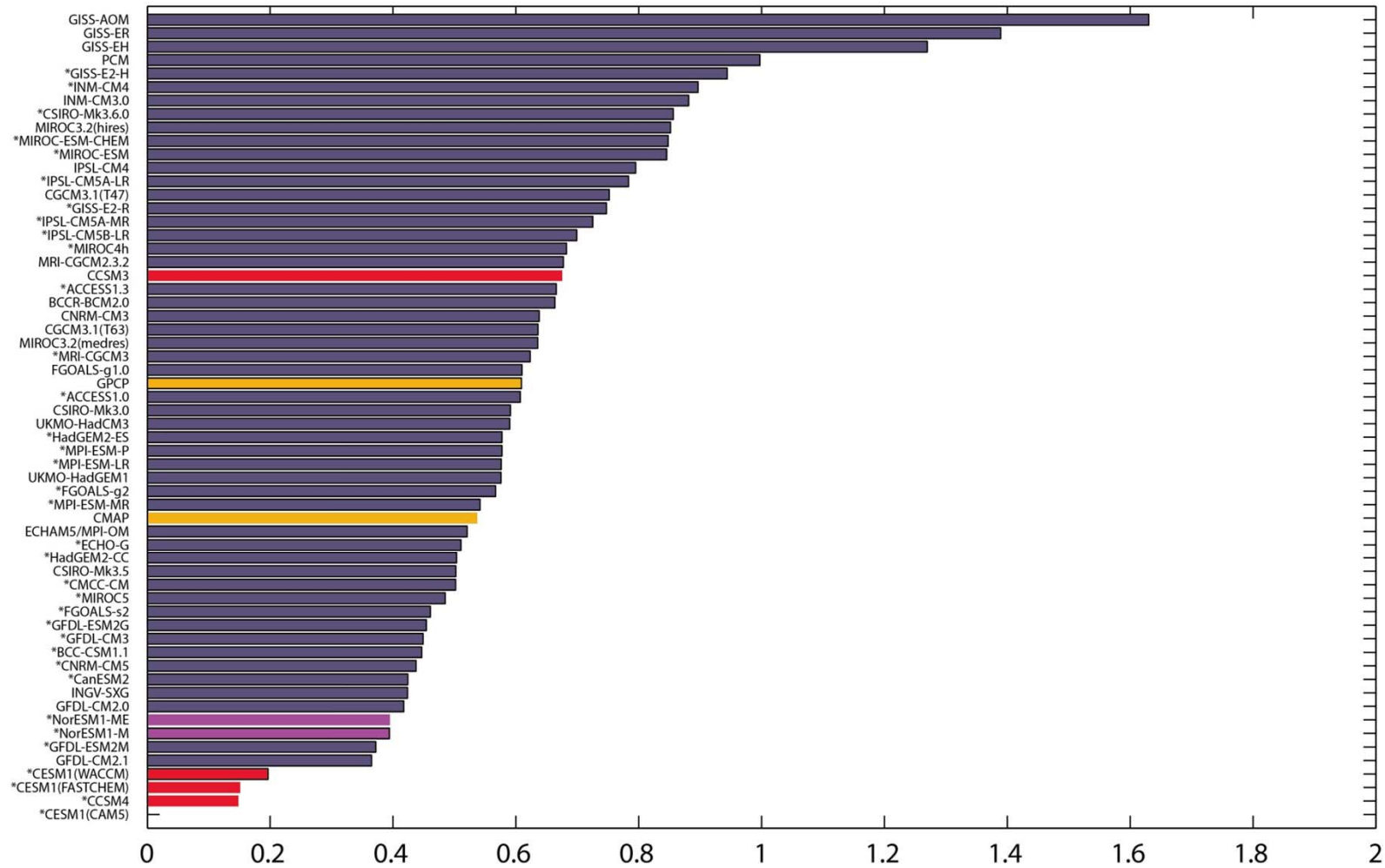
Surface temperature distance from CESM1(CAM5)



Climate model genealogy

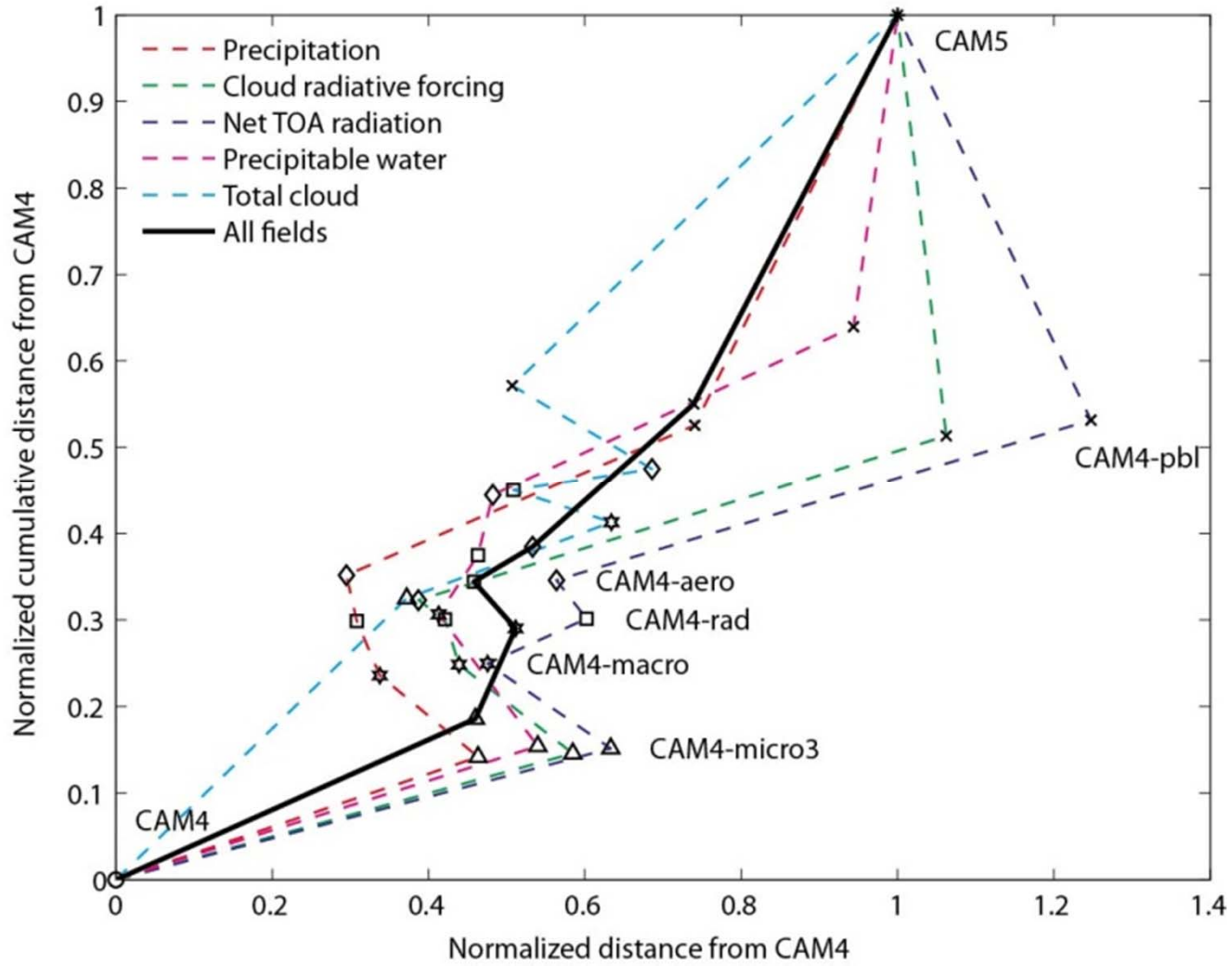
Distance from CESM1(CAM5)

Precipitation distance from CESM1(CAM5)



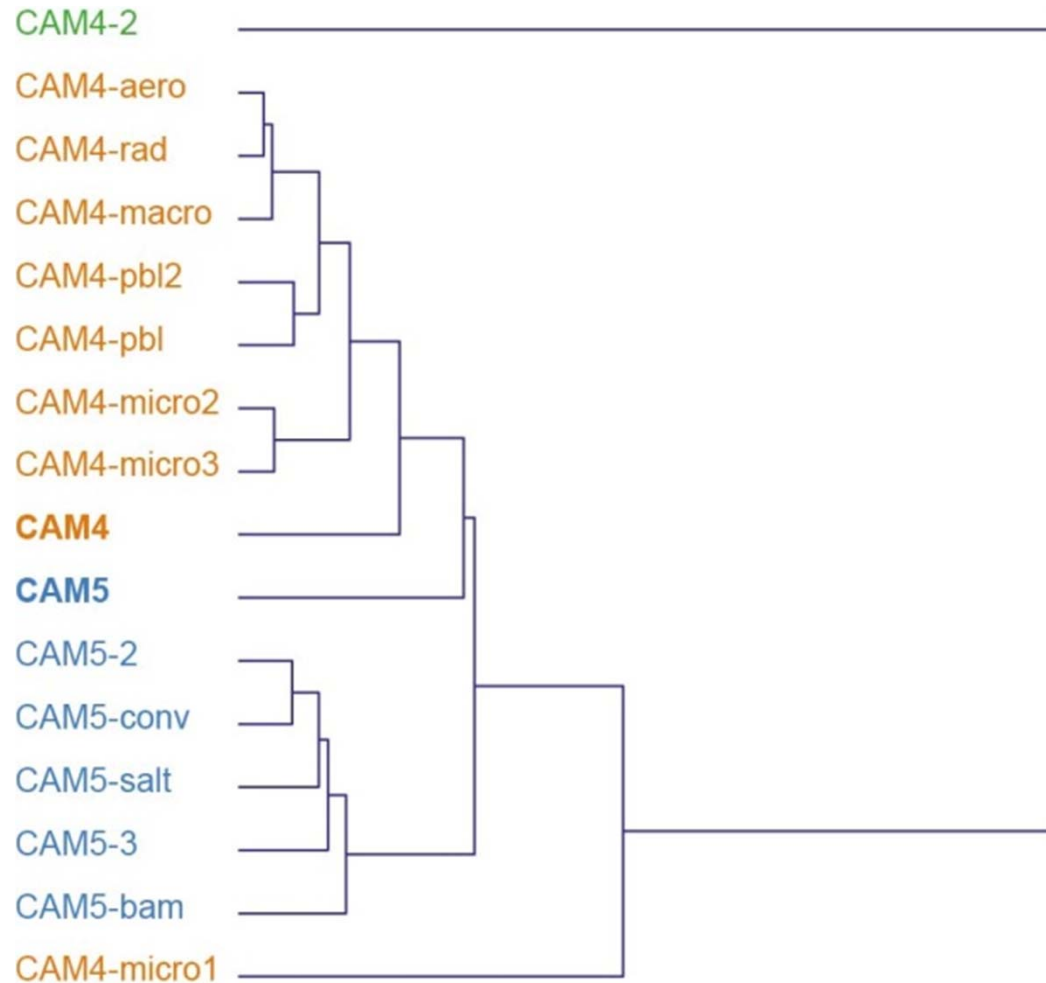
CCSM4 to CESM1(CAM5)

Tracing model development



CCSM4 to CESM1(CAM5)

Tracing model development



Conclusions

Metrics and model quality

- An infinite number of metrics can be defined.
- Many metrics are dependent.
- Observation datasets and uncertainty matters.
- The concept of a “best model” is ill-defined.
- There may be a best model for a particular purpose, where “best” measured in a specific way. But determining that is hard.



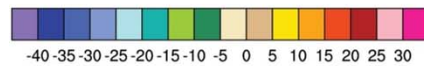
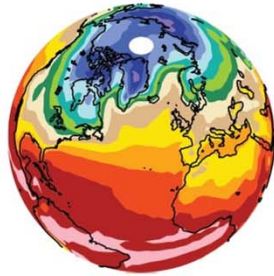
Conclusions

Where CESM certainly stands out

- Free availability
- Documentation
- Portability
- Support
- Community involvement
- Tutorials

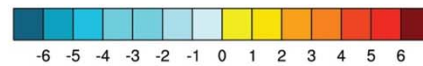
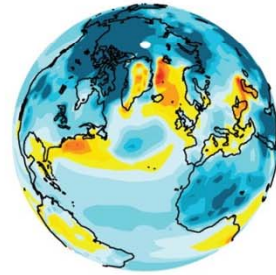
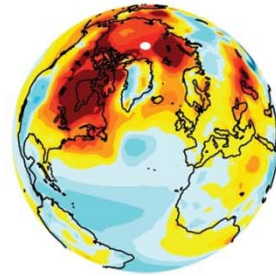
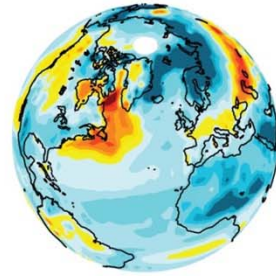
Relating model performance to projections

Temperature



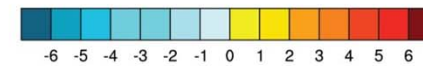
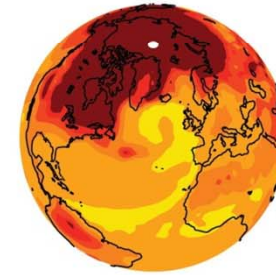
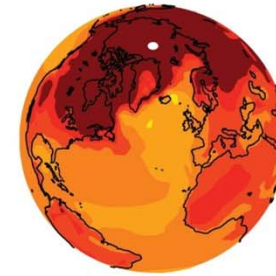
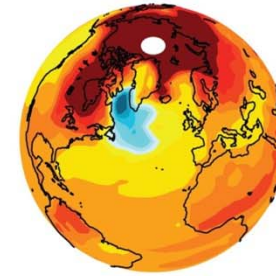
Dec-Feb surface temperature (°C)

Temperature bias



Temperature bias (°C)

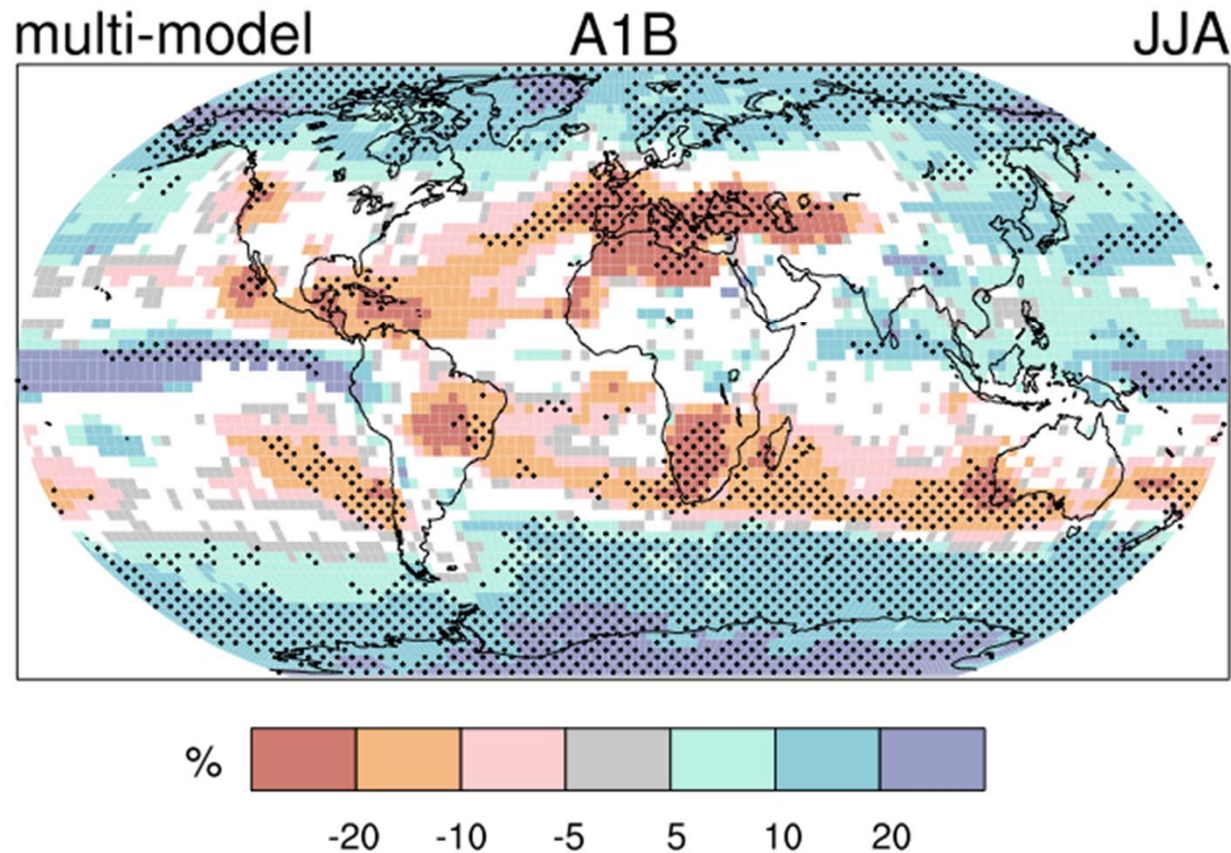
Warming



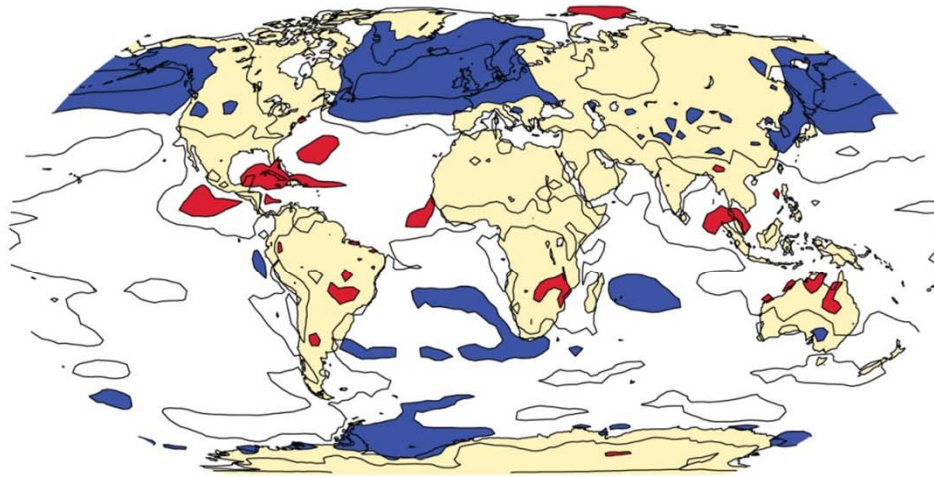
Dec-Feb warming 2080-2100 (°C)

How do we relate model performance to projections?

We don't.



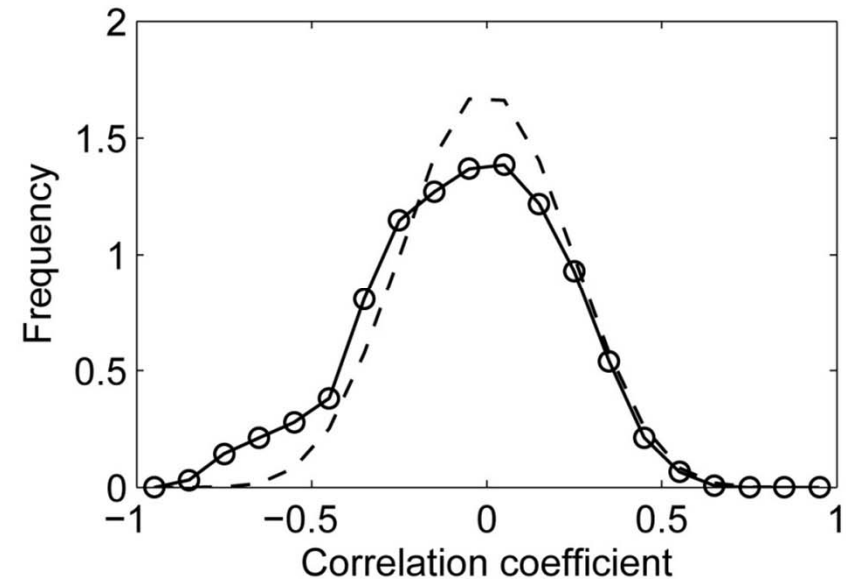
Present day temperature and warming do not correlate across models



Correlation of Dec-Feb temperature
to projected warming

$R < -0.4$

$R > 0.4$

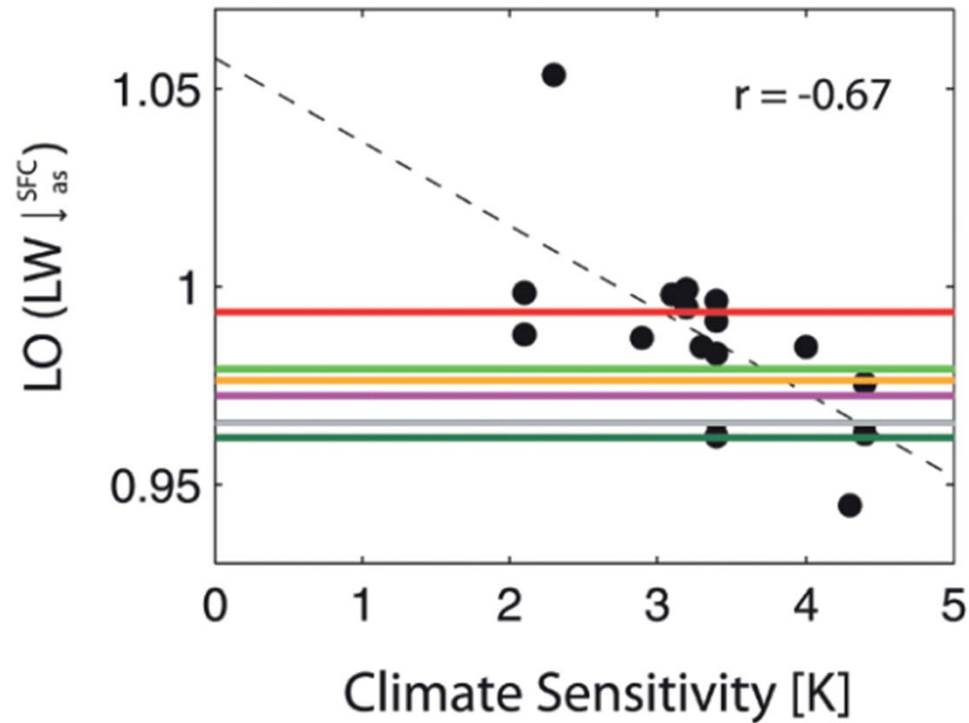


o-o-o Model

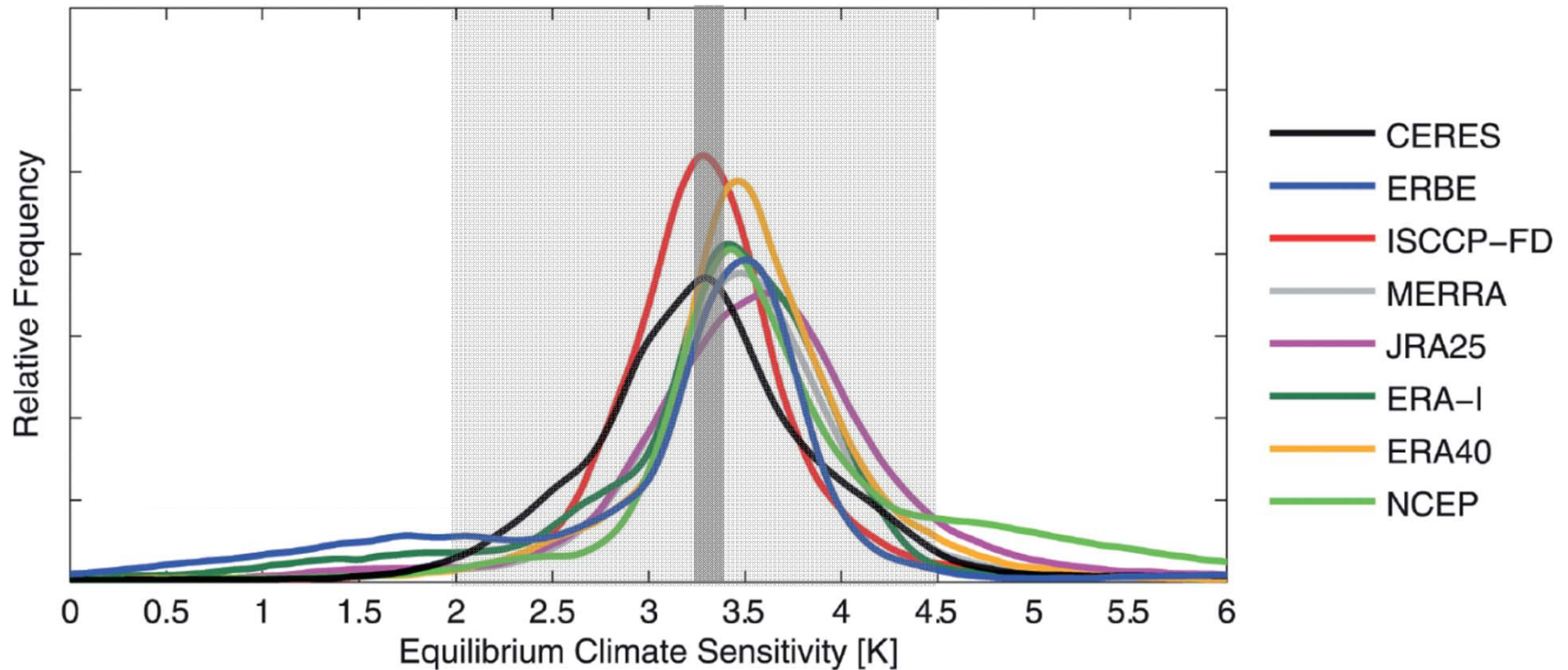
----- Random

Some observable climate indices do correlate with future warming

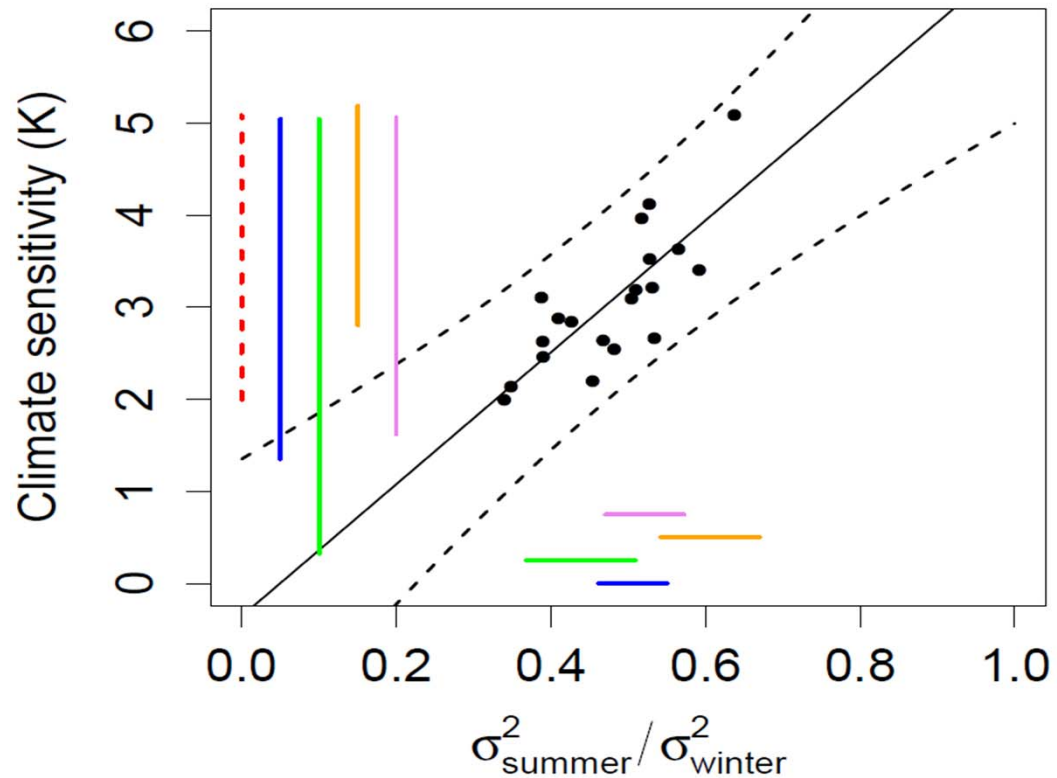
Land ocean contrast
in surface longwave
downward all sky radiation



Some observable climate indices do correlate with future warming

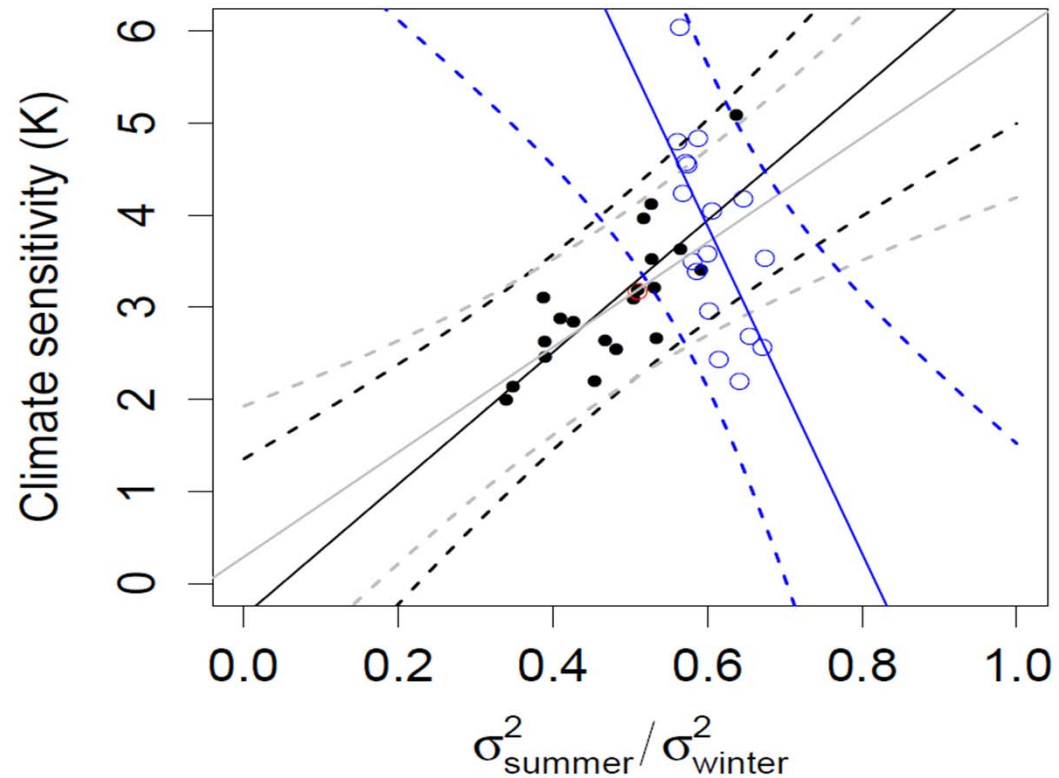


Some relationships are not stable in different ensembles



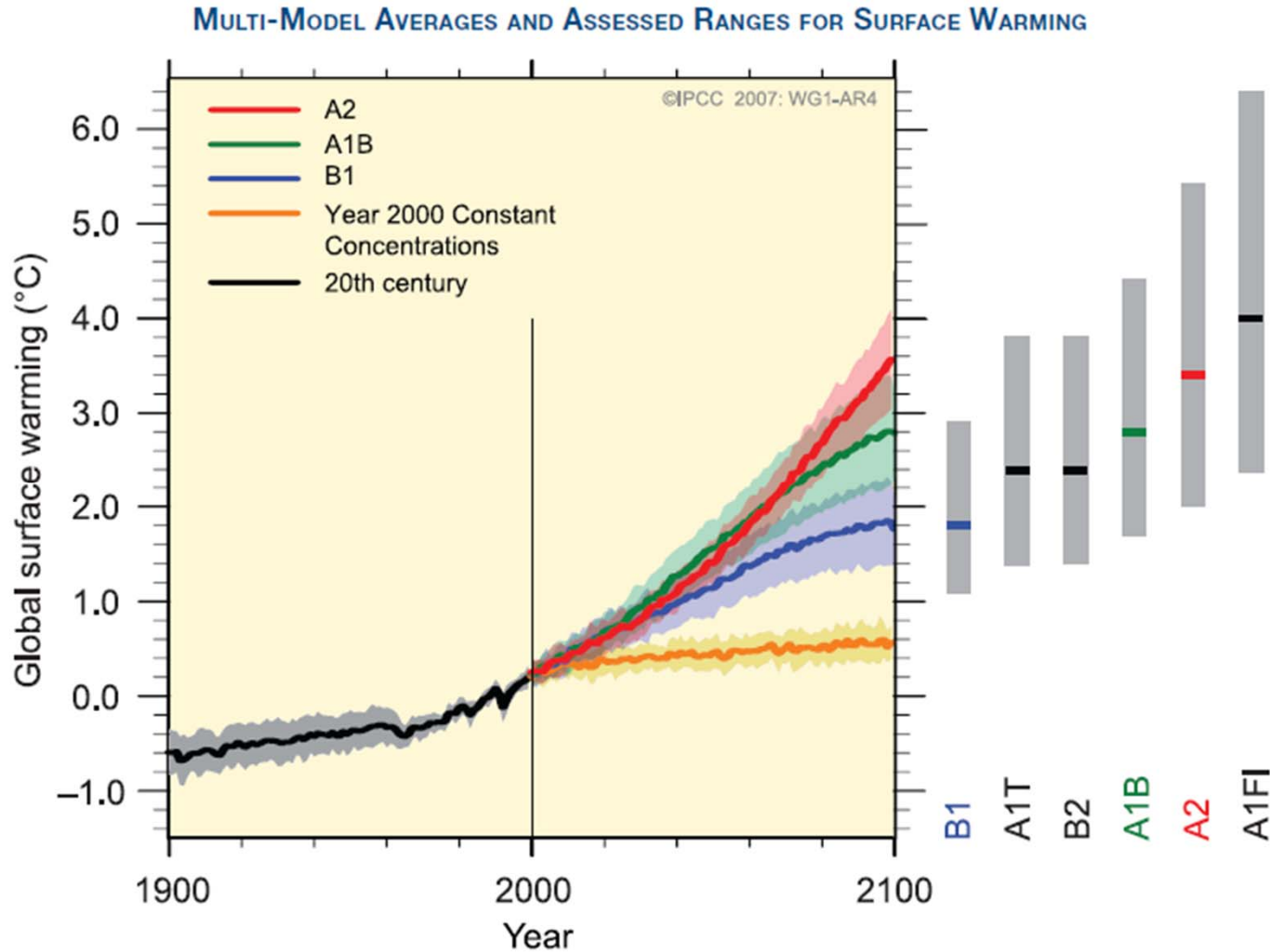
Summer to winter interannual variability

Some relationships are not stable in different ensembles



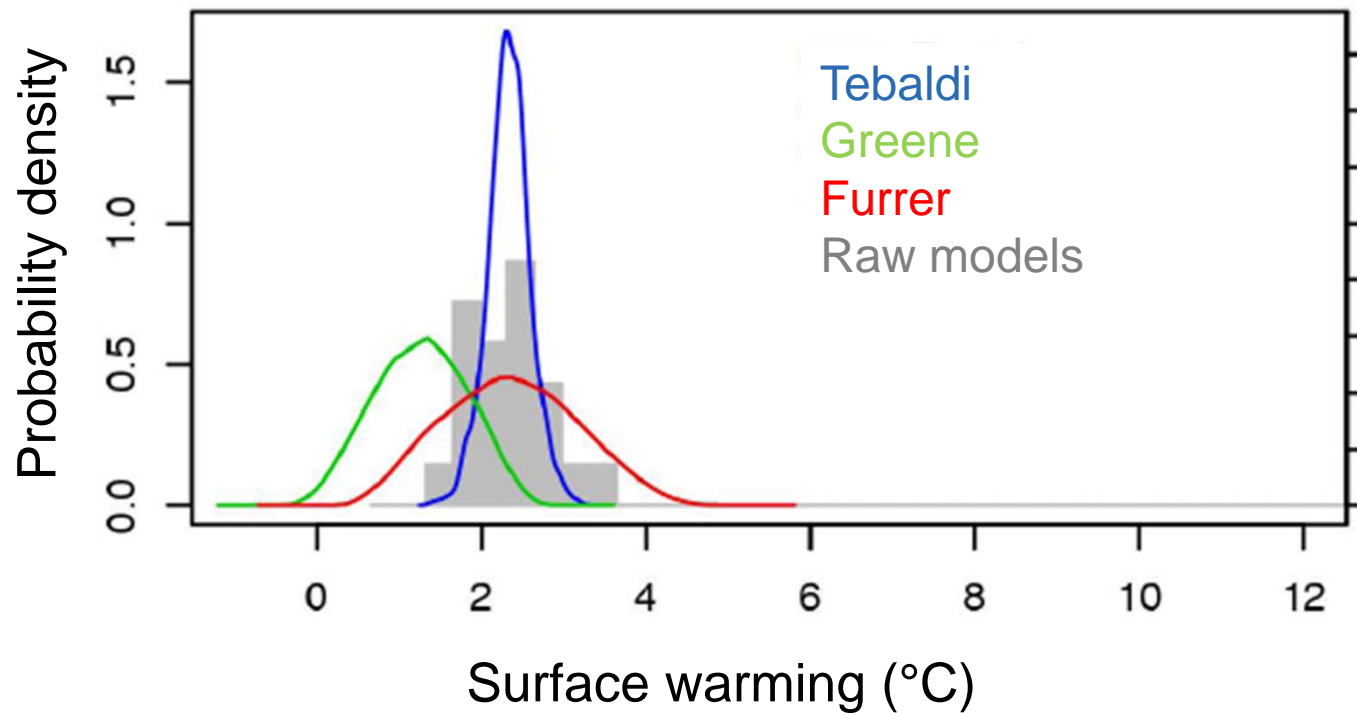
Summer to winter interannual variability

Model spread in an ensemble of opportunity not always useful as a measure of uncertainty



Assumptions in the weighting matter for the result

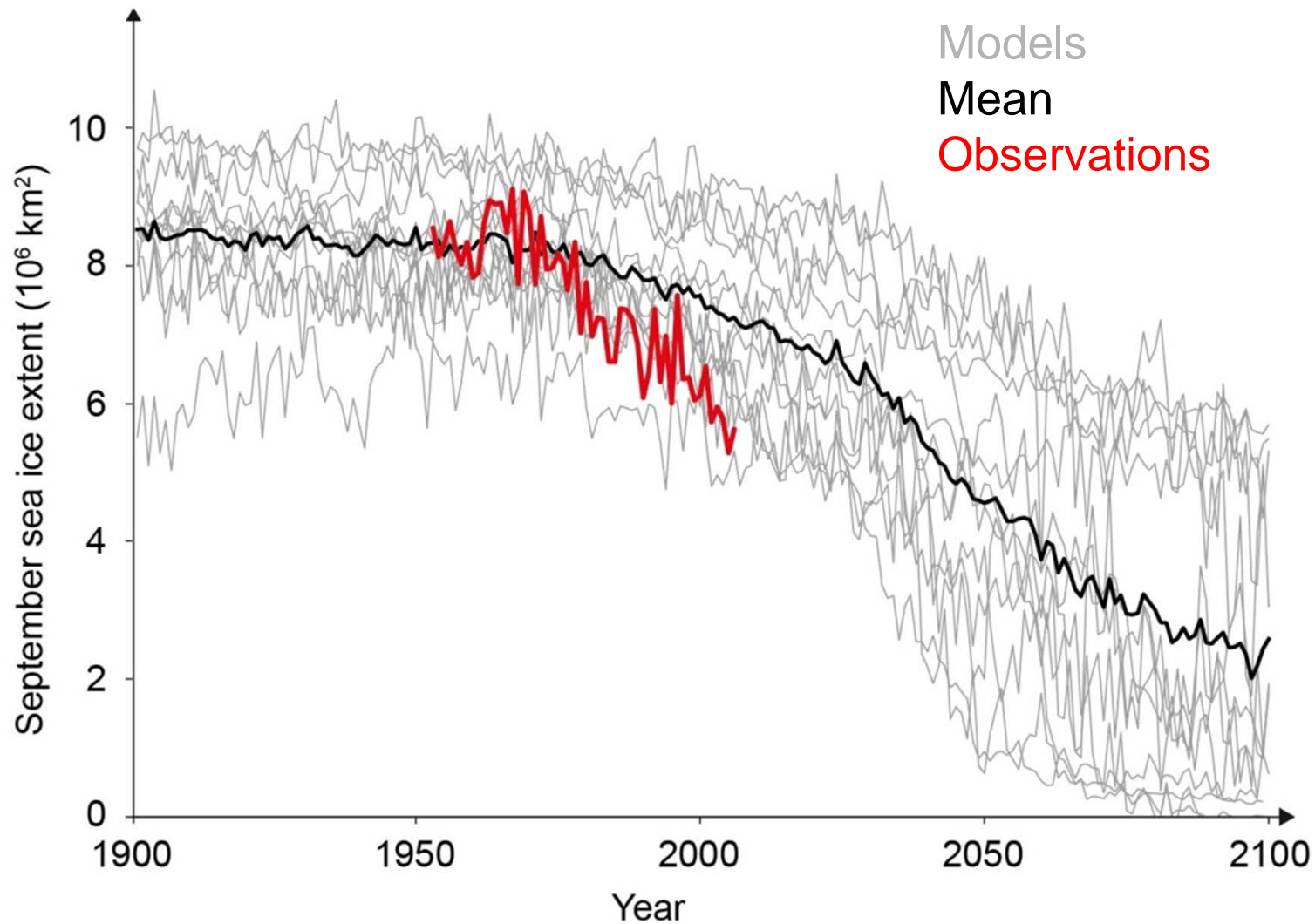
Surface warming South East Asia
Dec-Feb 2080-2099, A1B scenario



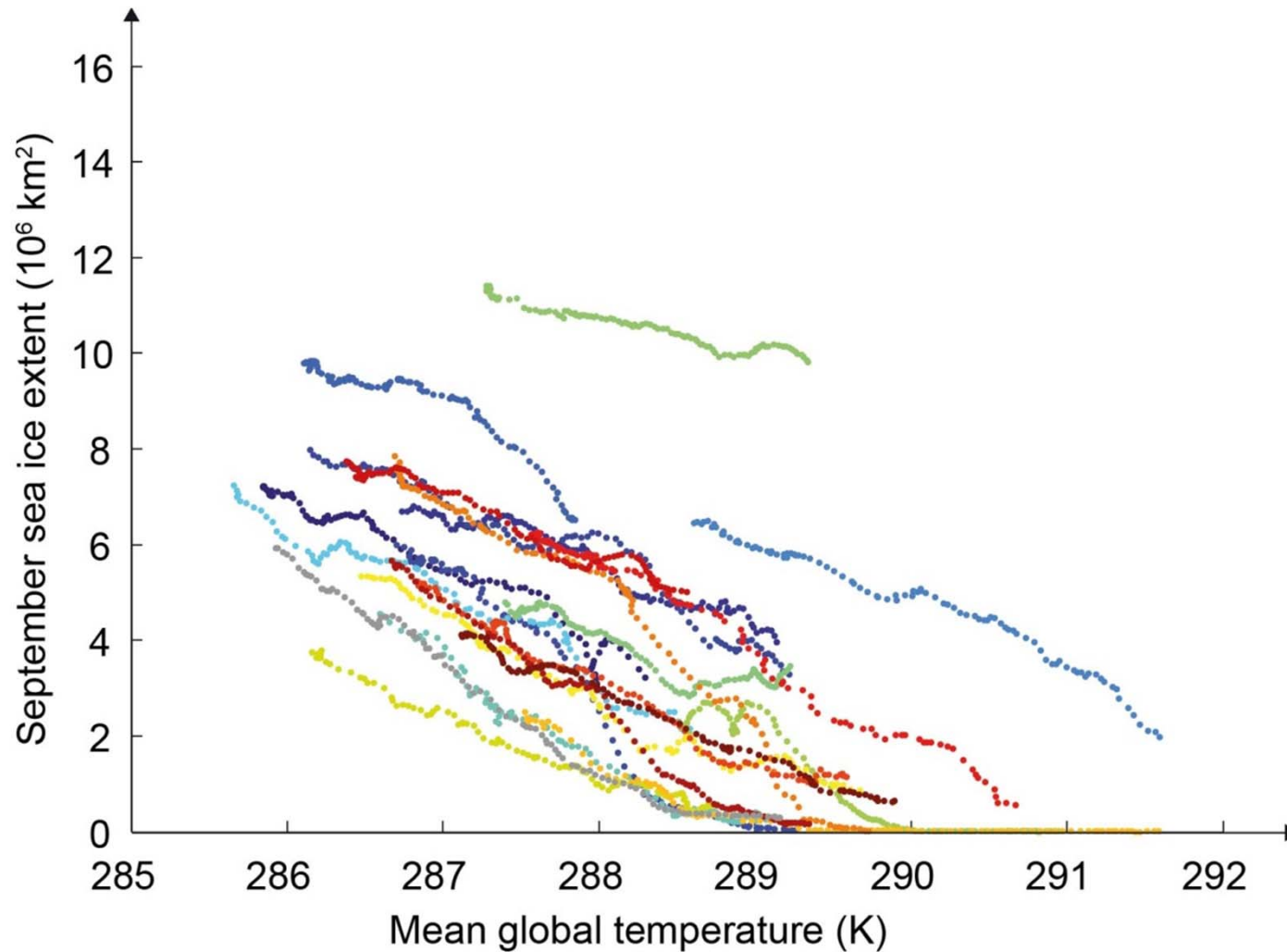
Arctic sea ice decline

Faster than forecast?

Arctic sea ice decreases in all models, but less than observed

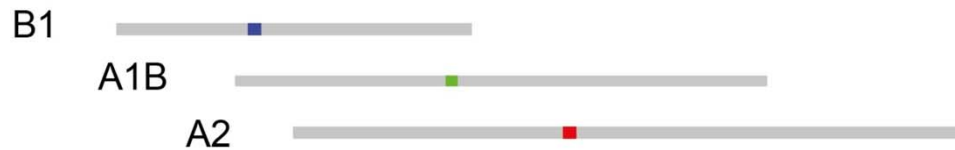
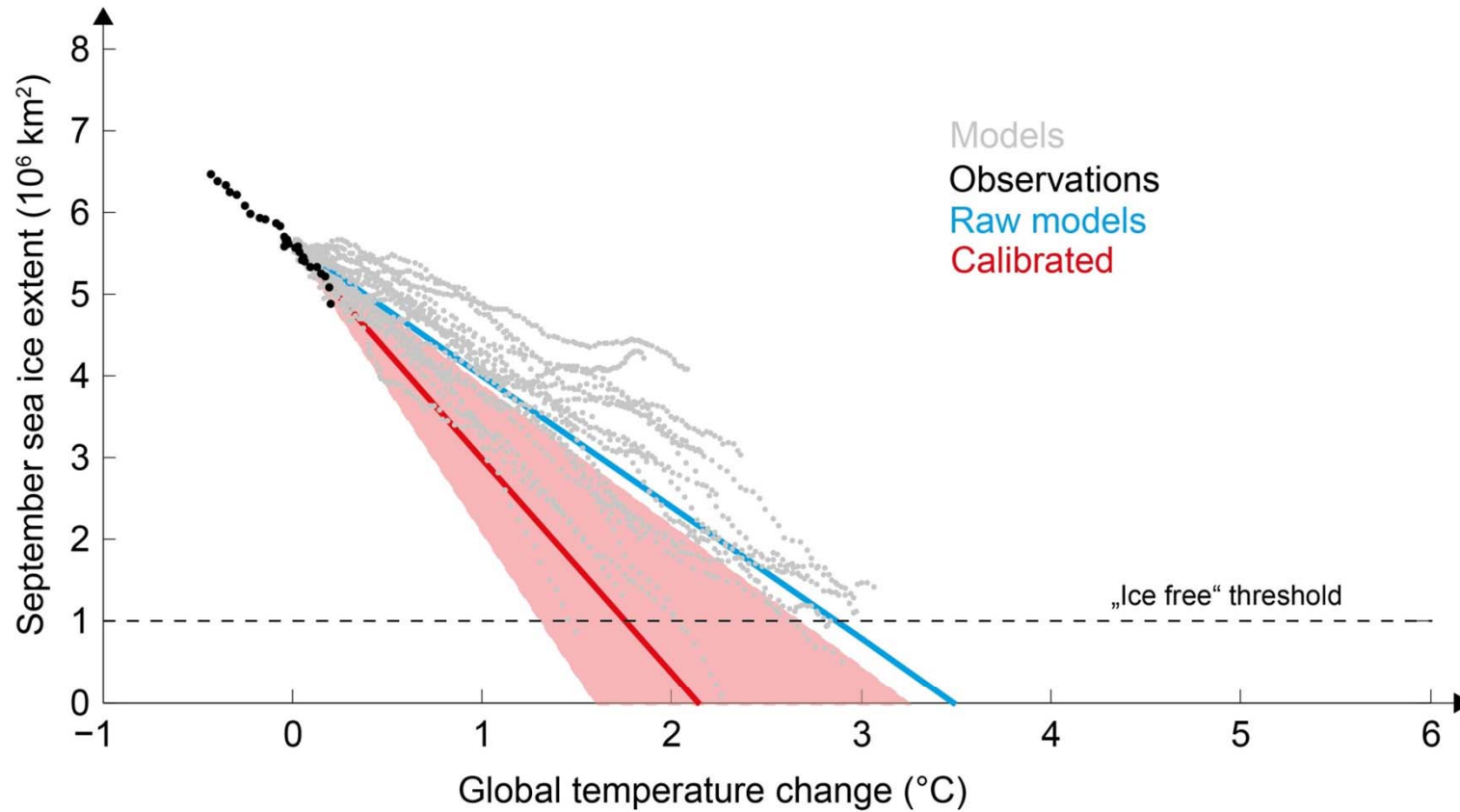


Arctic sea ice is approximately linearly related to temperature



Arctic sea ice

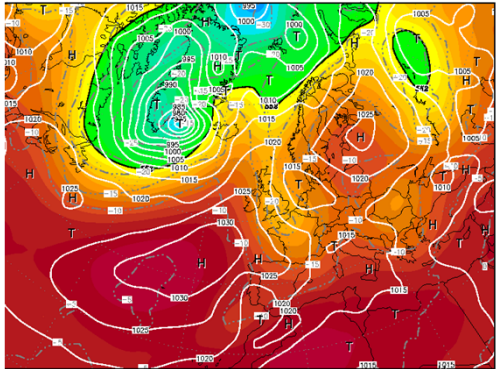
A calibrated prediction



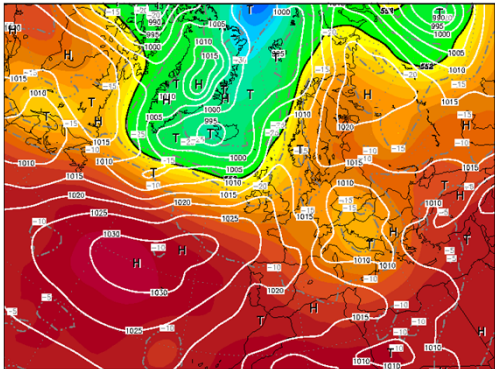
How is the weather tomorrow?

Combining lines of evidence

Init : Sat,20AUG2005 00Z Valid: Sat,20AUG2005 00Z
500 hPa Ceopot.(gpm), T (C) und Bodendr. (hPa)

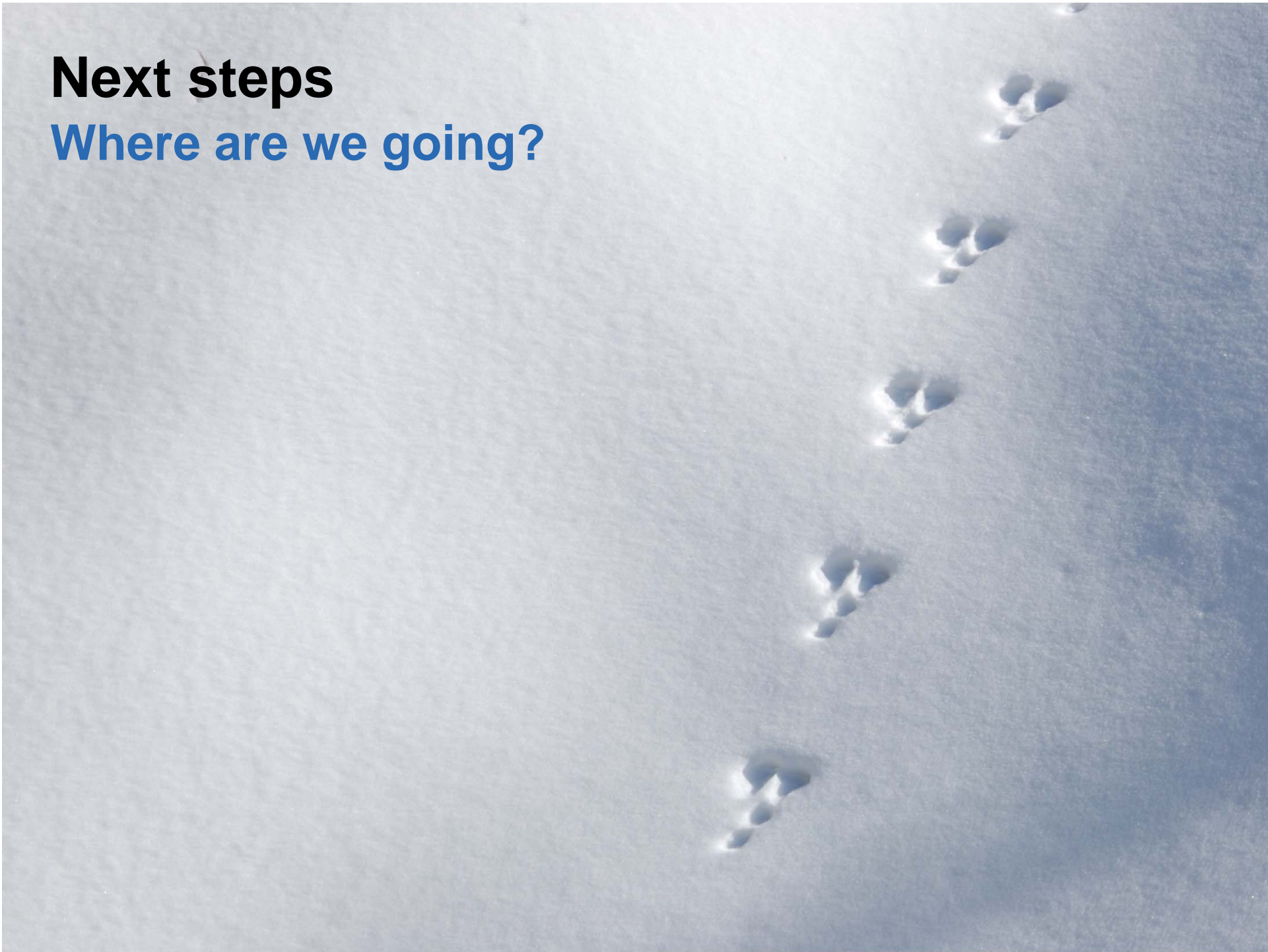


Init : Tue,23AUG2005 00Z Valid: Tue,23AUG2005 00Z
500 hPa Ceopot.(gpm), T (C) und Bodendr. (hPa)



Next steps

Where are we going?



Are we making progress?

The end of model democracy?

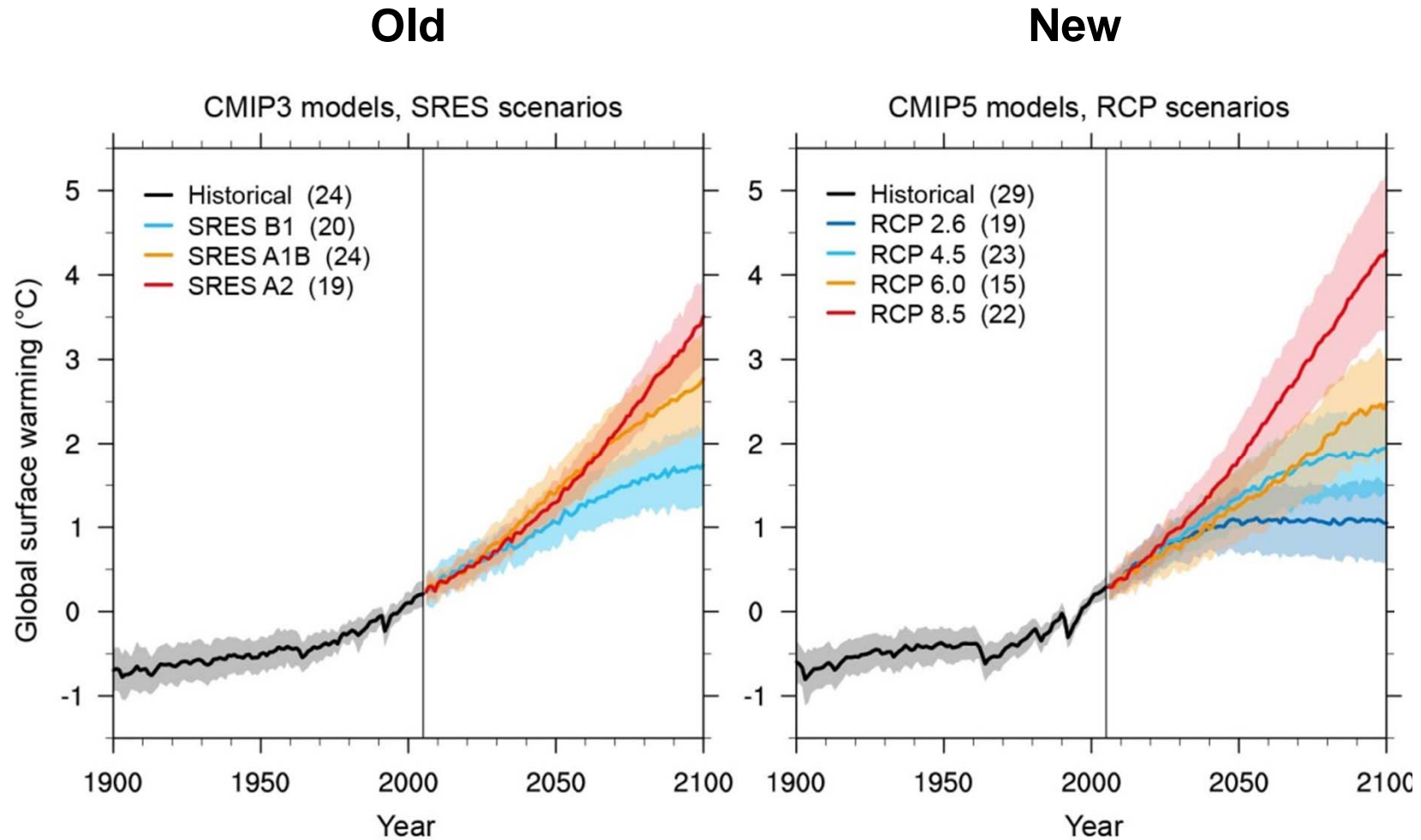
- “There should be no minimum performance criteria for entry into the CMIP multi-model database.”
- “Researchers may select a subset of models for a particular analysis but should document the reasons why.”
- “IPCC assessments should consider the large amount of scientific work on CMIP3, in particular in cases where lack of time prevents an in depth analysis of CMIP5.”

Good Practice Guidance Paper on
Assessing and Combining Multi Model Climate Projections

Core Writing Team:
Reto Knutti (Switzerland), Gabriel Abramowitz (Australia), Matthew Collins (United Kingdom),
Veronika Eyring (Germany), Peter J. Gleckler (USA), Bruce Hewitson (South Africa), Linda Mearns (USA)

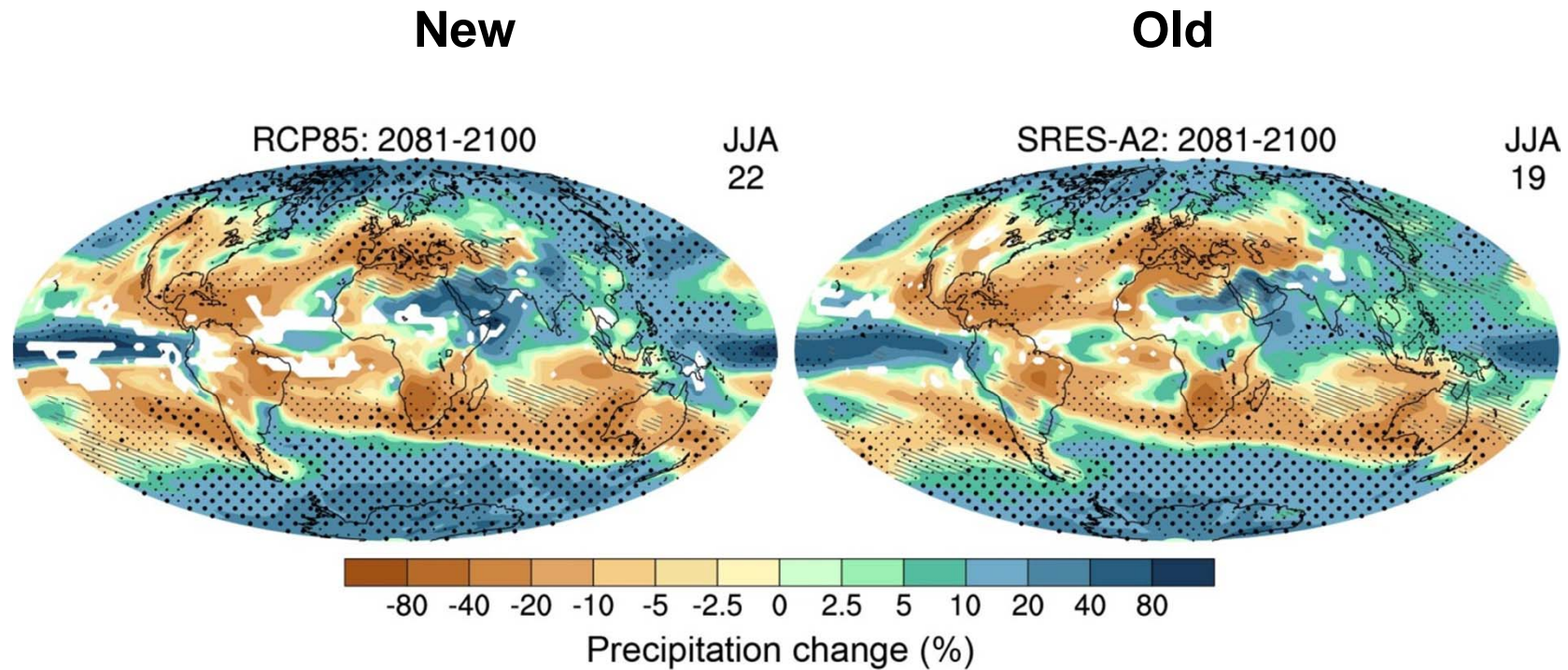
Are we making progress?

Projections are robust, uncertainties remain



Are we making progress?

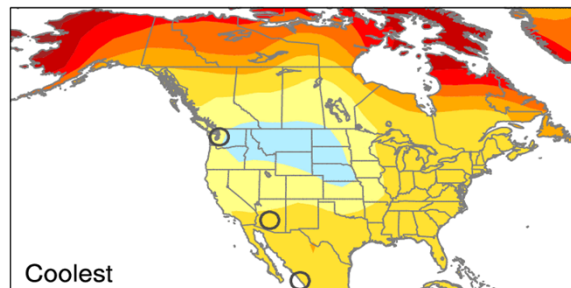
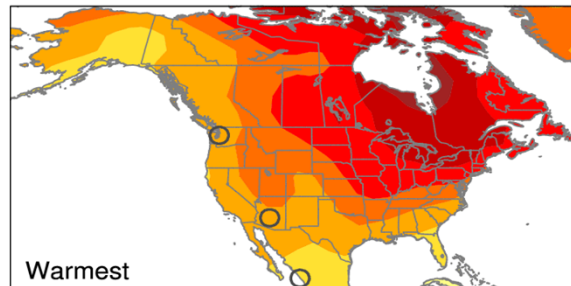
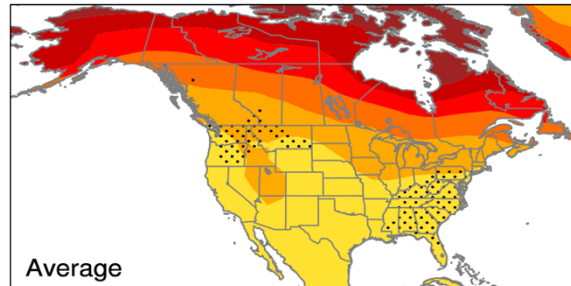
June-August precipitation projections



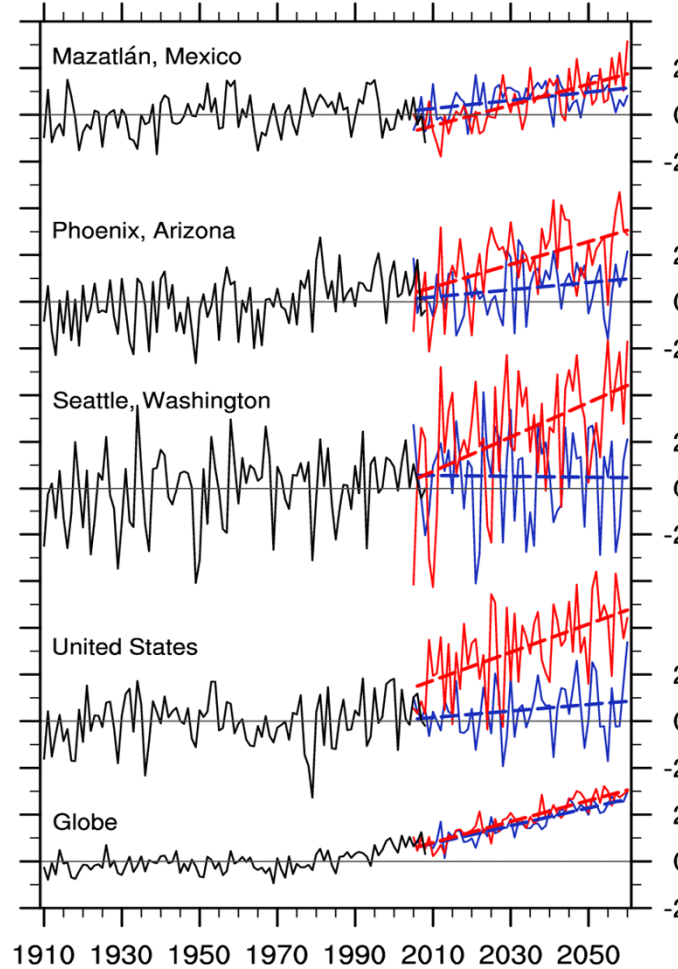
Limits of predictability

Warmest and coolest of 40 realizations

DJF Temperature Trend 2005-2060

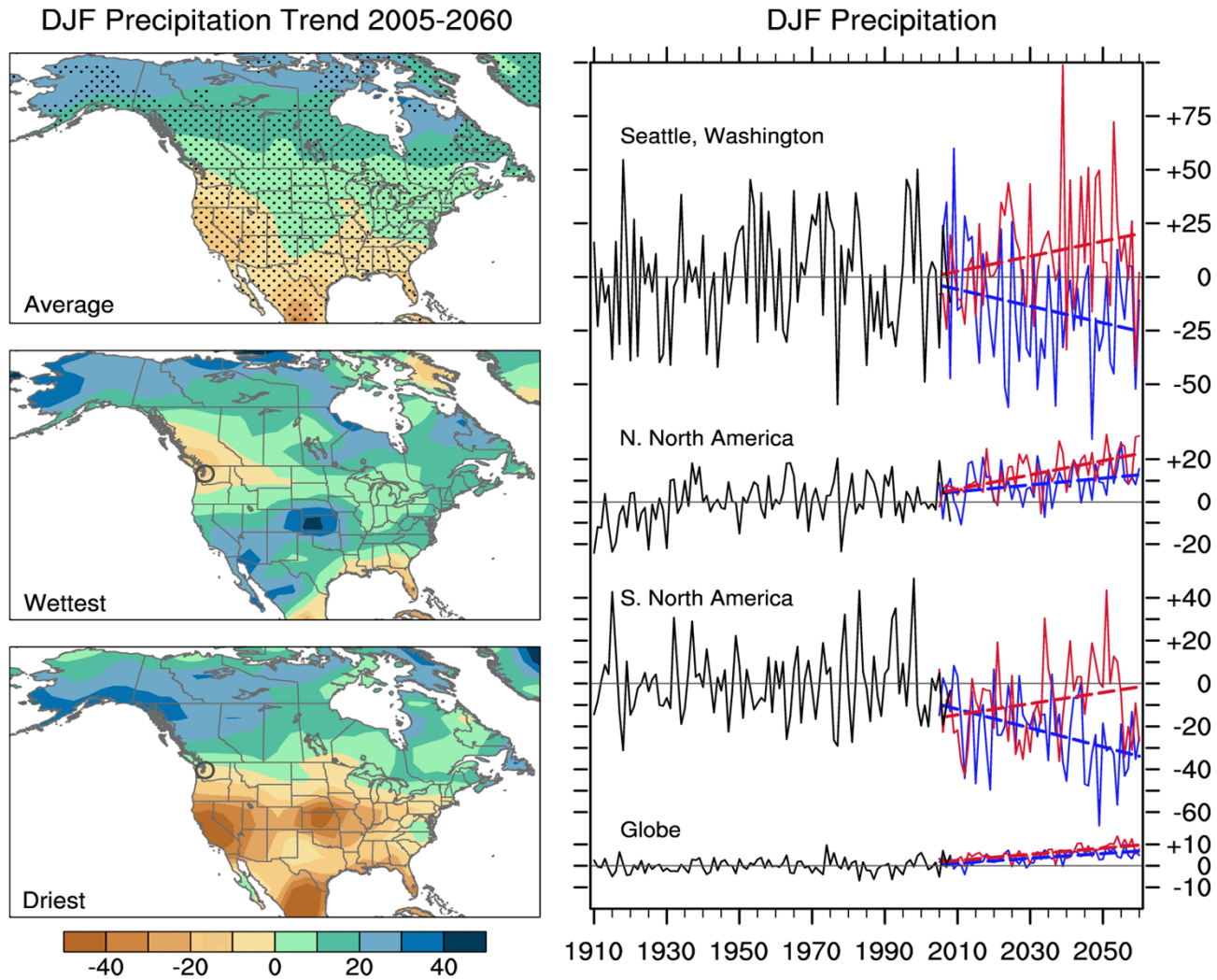


DJF Temperature



Limits of predictability

Wettest and driest of 40 realizations



Are we making progress?

Yes and no. And why it is hard.

- Open system, complex, non-linear
- Interaction from seconds to millennia, from micrometers to thousands of kilometers
- Dimensionality
- Natural variability
- Limited and uncertain observations
- Model discrepancy
- Expensive models, petabytes of data
- Calibration problem in a high-dimensional space
- Out of sample prediction, no proper verification

Establishing confidence in a prediction

Why do you believe the weather forecast?

BRECKENRIDGE CO



Fair
61°F
16°C

Humidity 7%
Wind Speed SW 17 G 21 mph
Barometer 30.40 in
Dewpoint -2°F (-19°C)
Visibility 10.00 mi

Last Update on 18 Jun 10:12 am MDT

TODAY



Breezy

High: 74 °F

TONIGHT



Breezy

Low: 45 °F

TUESDAY



Breezy

High: 70 °F

TUESDAY
NIGHT



Breezy

Low: 41 °F

WEDNESDAY



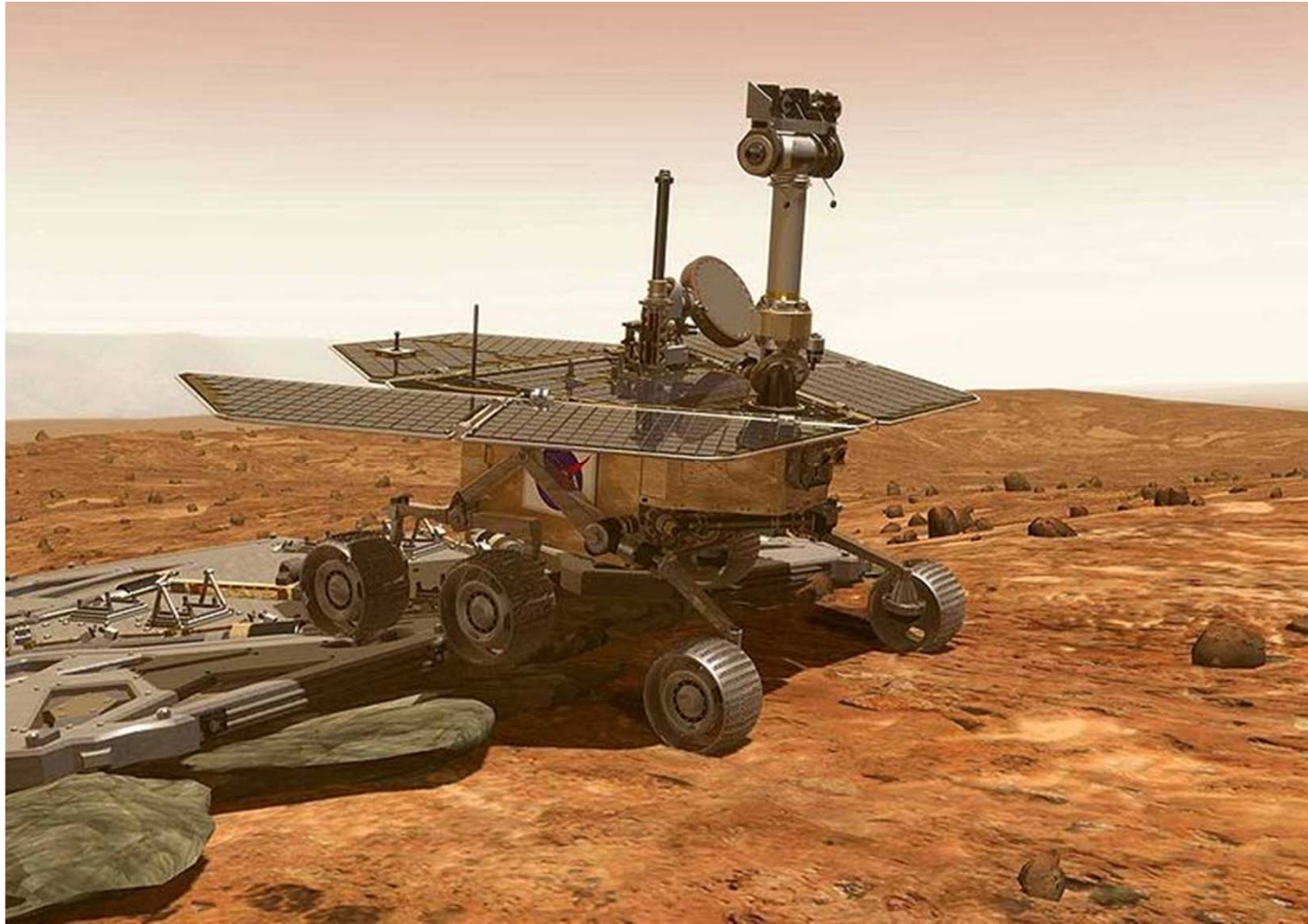
Sunny

High: 71 °F

Tuesday: Sunny, with a high near 70. Breezy, with a west southwest wind 9 to 18 mph, with gusts as high as 29 mph

Establishing confidence in a prediction

Lack of verification



Establishing confidence in a prediction

Lack of verification



Conclusions and challenges

Climate models have reached a remarkable level of maturity, but:

- Model sampling is neither systematic nor random.
- We use a collection of ‘best guesses’ not designed to span any uncertainty range, and not independent.
- Models are getting “better” at things we observe, but model spread is not reduced for projections.
- Model performance varies but we don’t know how to translate into weights. But we discard old models.
- Combining model may help but can create unphysical results.
- No independent verification of the prediction.
- There is no best model without defining a purpose.