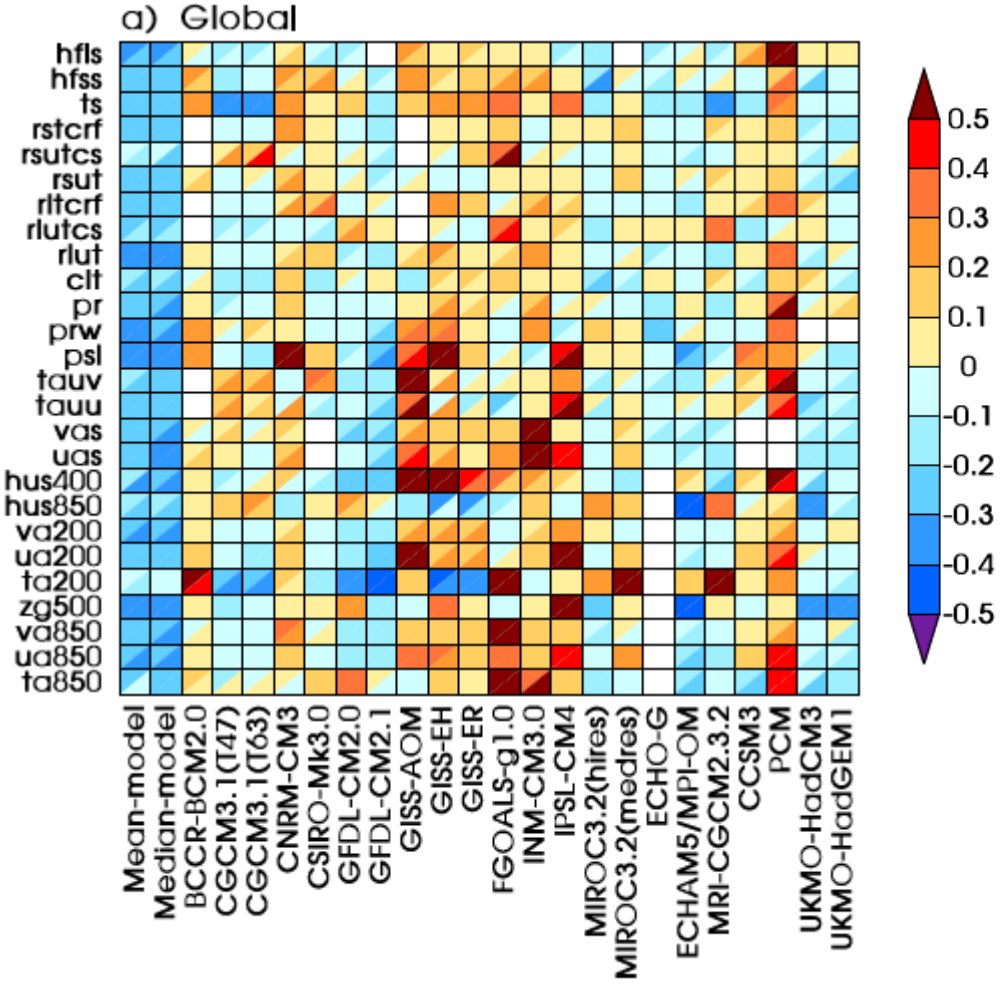
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Relating model performance to credibility of projections

Reto Knutti
Institute for Atmospheric and Climate Science
ETH Zurich, Switzerland
reto.knutti@env.ethz.ch

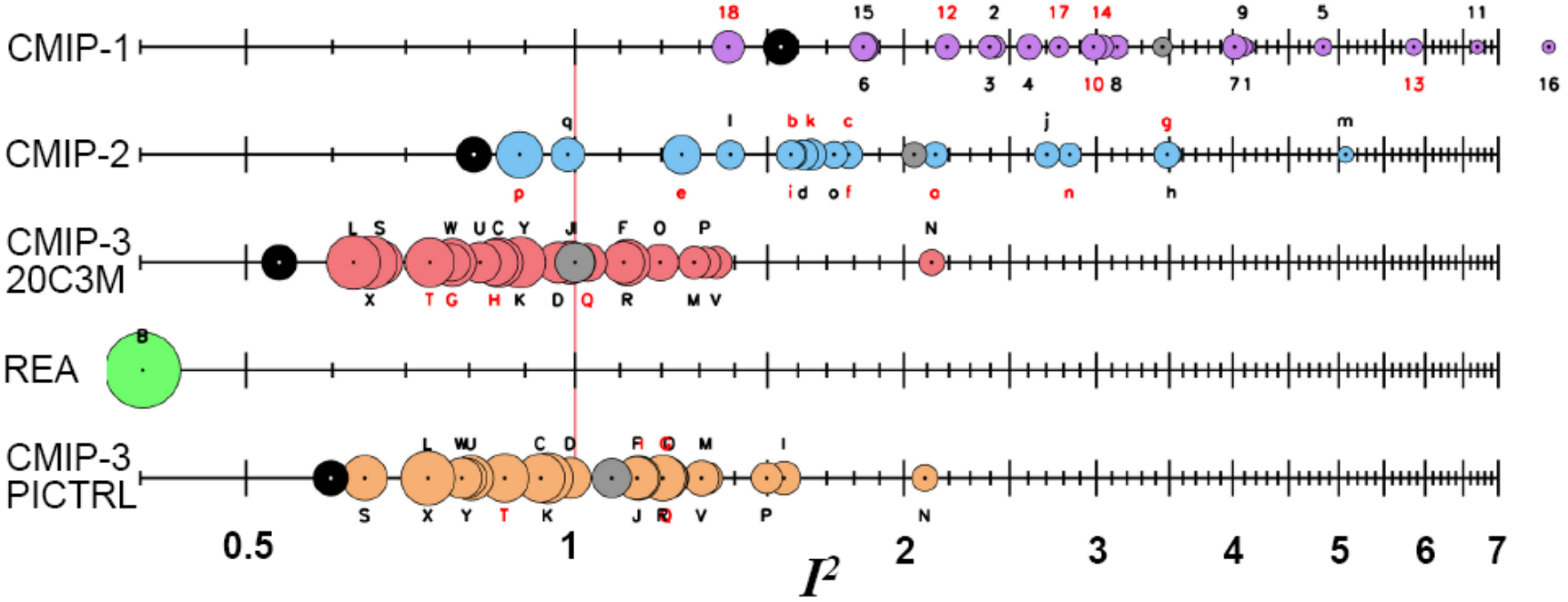
Translating model agreement with observations into skill



- 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

(Gleckler et al., JGR 2008)

Is there are best model?



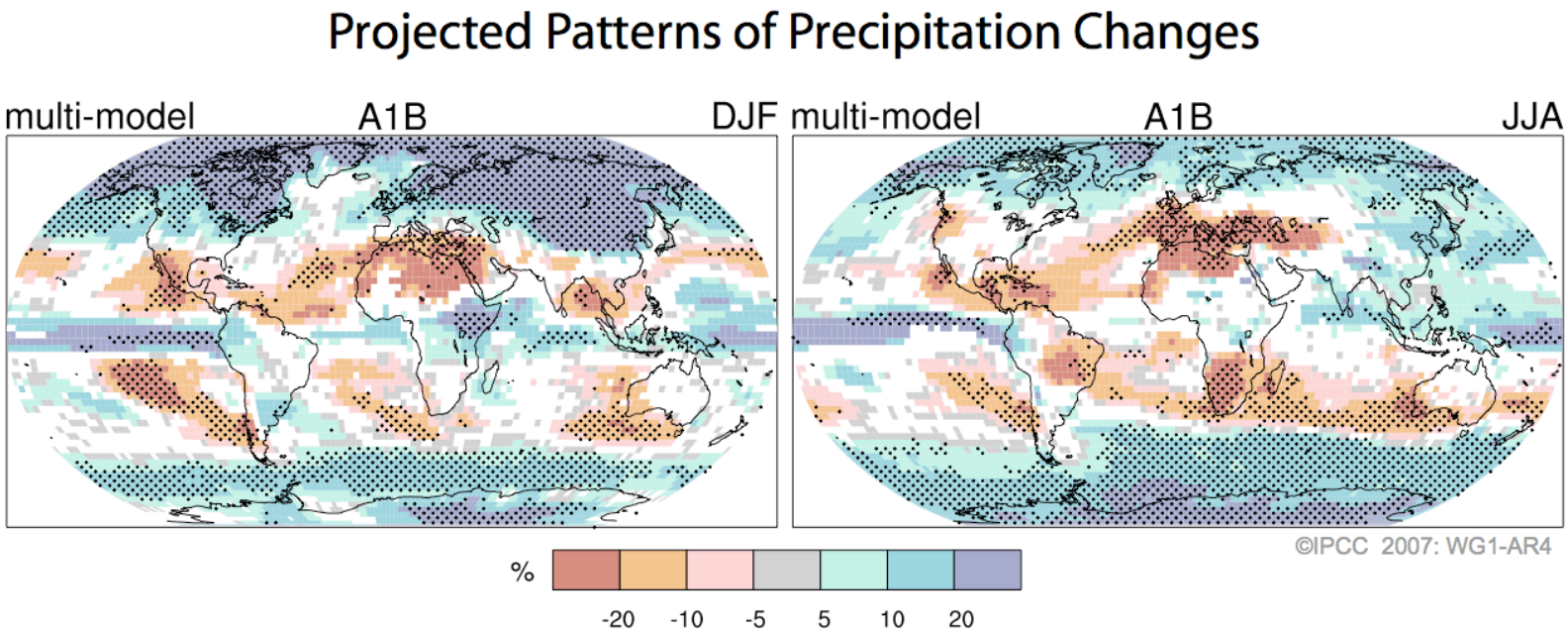
Model performance

Better ⟷ Worse

(Reichler and Kim, BAMS 2008)

How do we relate model performance to credibility of projections?

We don't.

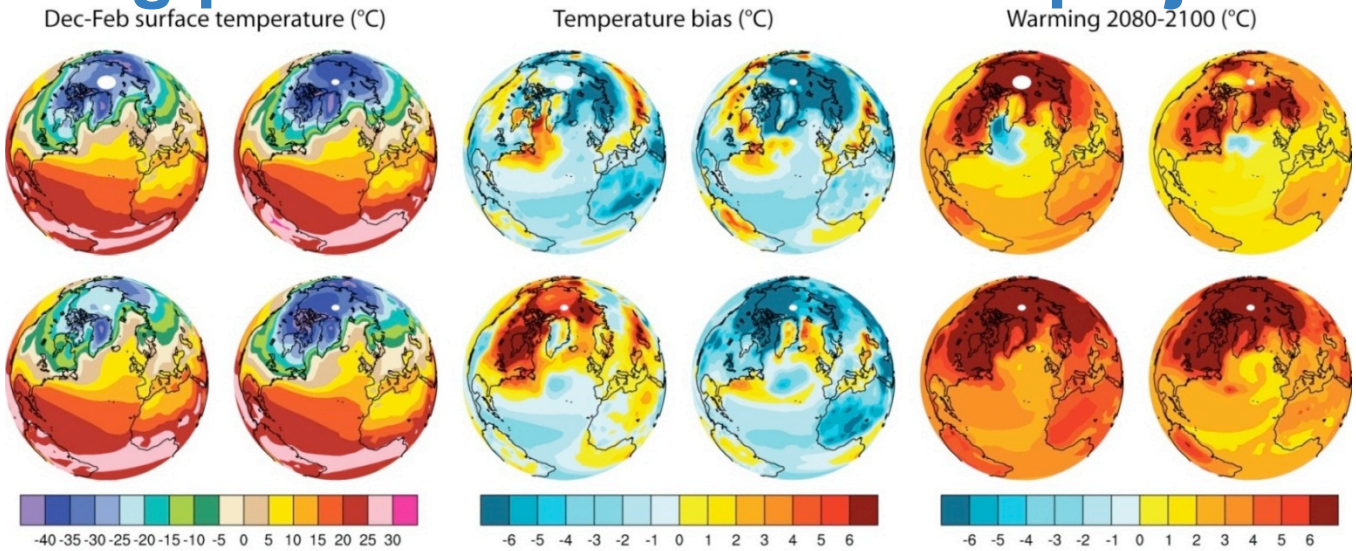


(IPCC AR4, Fig. SPM7)

Relating performance metrics to projections

(Knutti, Phil Trans Roy Soc 2008)

Relating performance metrics to projections



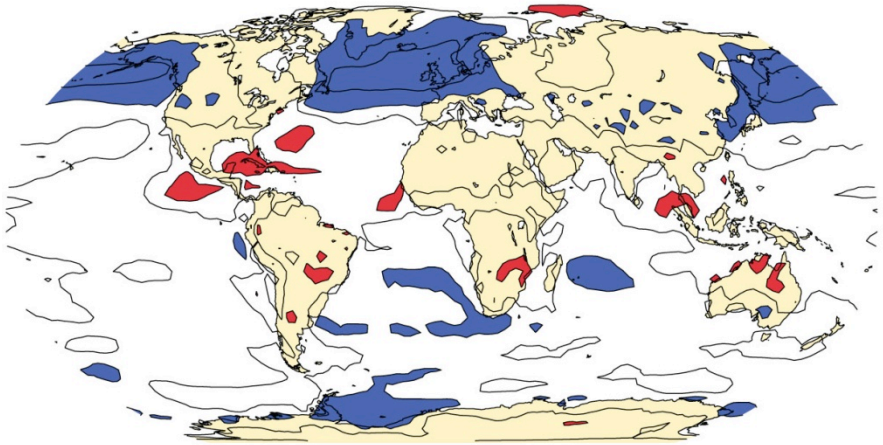
- Which metrics should we use?
- How might a metric translate into weight?
- Structural model uncertainty
- Observation uncertainty
- Range of the ensemble and model dependence

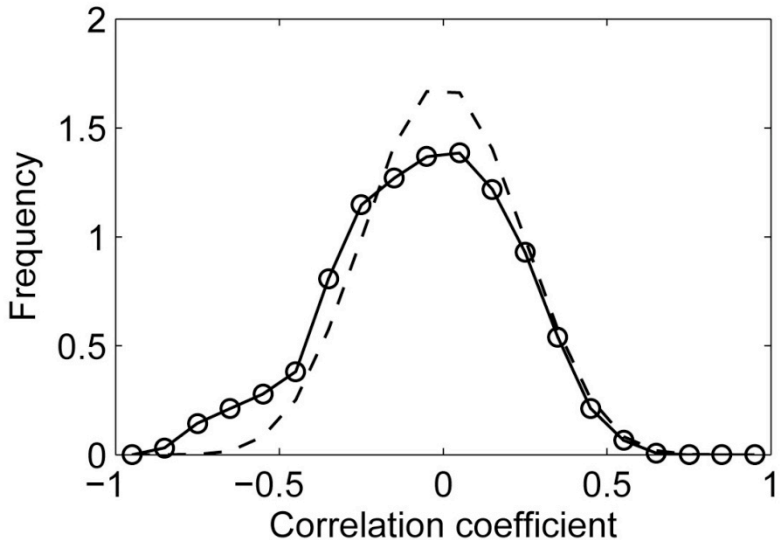
(Knutti, Phil Trans Roy Soc 2008)

Relating present day climate to future changes

- Correlation between present day temperature and future warming across models is small.
- How useful are metrics based on present day climate?

a) DJF correlations between current temperature and predicted warming (blue < -0.4, red > 0.4)





(Knutti et al., J. Climate 2010)

How should we evaluate climate models?
What is a good model?

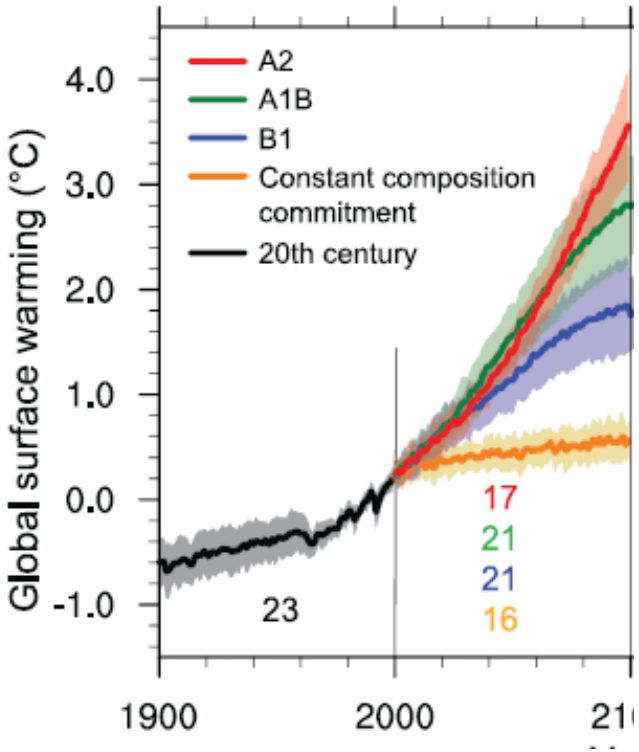
There is considerable confidence that climate models provide credible quantitative estimates of future climate change, particularly at continental scales and above. This confidence comes from the foundation of the models in accepted physical principles and from their ability to reproduce observed features of current climate and past climate changes. (IPCC AR4 FAQ 8.1)

Aspects of observed climate that must be simulated to ensure reliable future predictions are unclear. For example, models that simulate the most realistic present-day temperatures for North America may not generate the most reliable projections of future temperature changes. (US CCSP report 3.1)

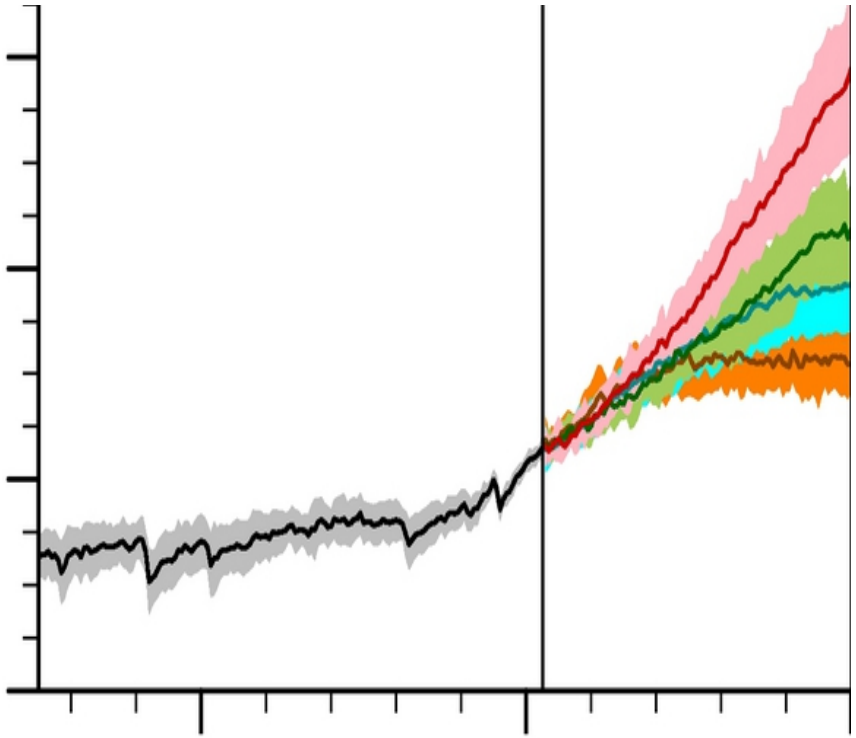
Relating present day climate to future changes

- Model agreement with observations improves, but future spread is not decreasing.

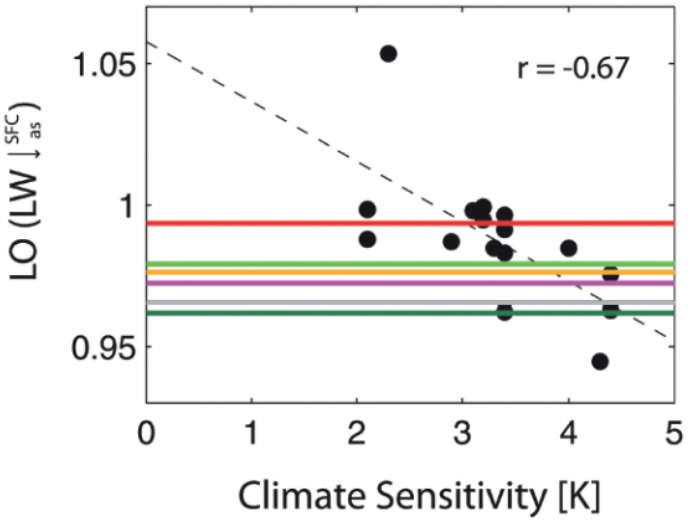
CMIP3

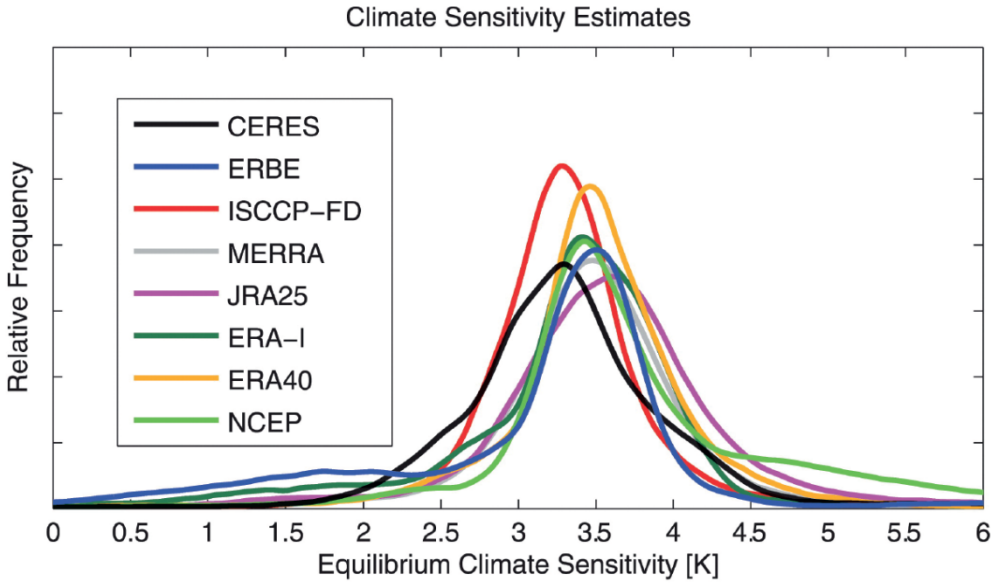


CMIP5
Preliminary, based on ~10 models!



Correlation across models



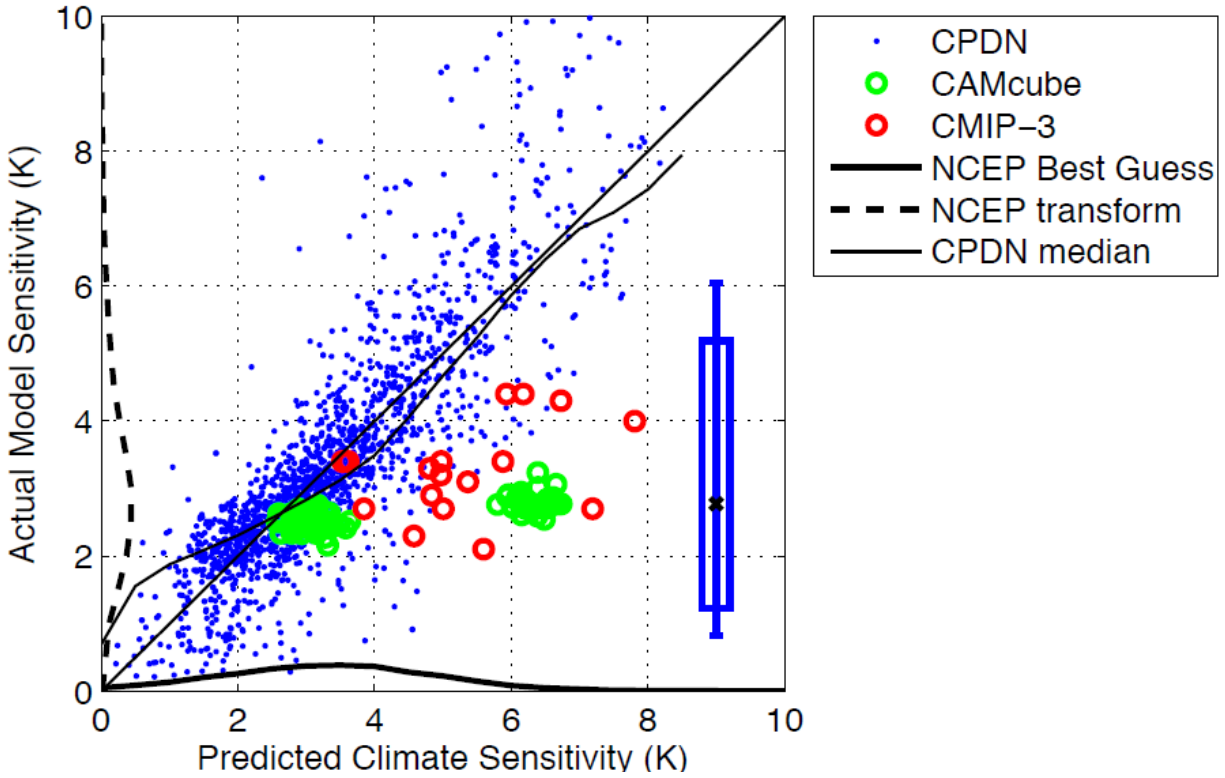


- Relative frequency based on hundreds of indices
- Remarkably similar to the original CMIP3 range

(Huber et al., J. Climate 2011)

Model structure matters

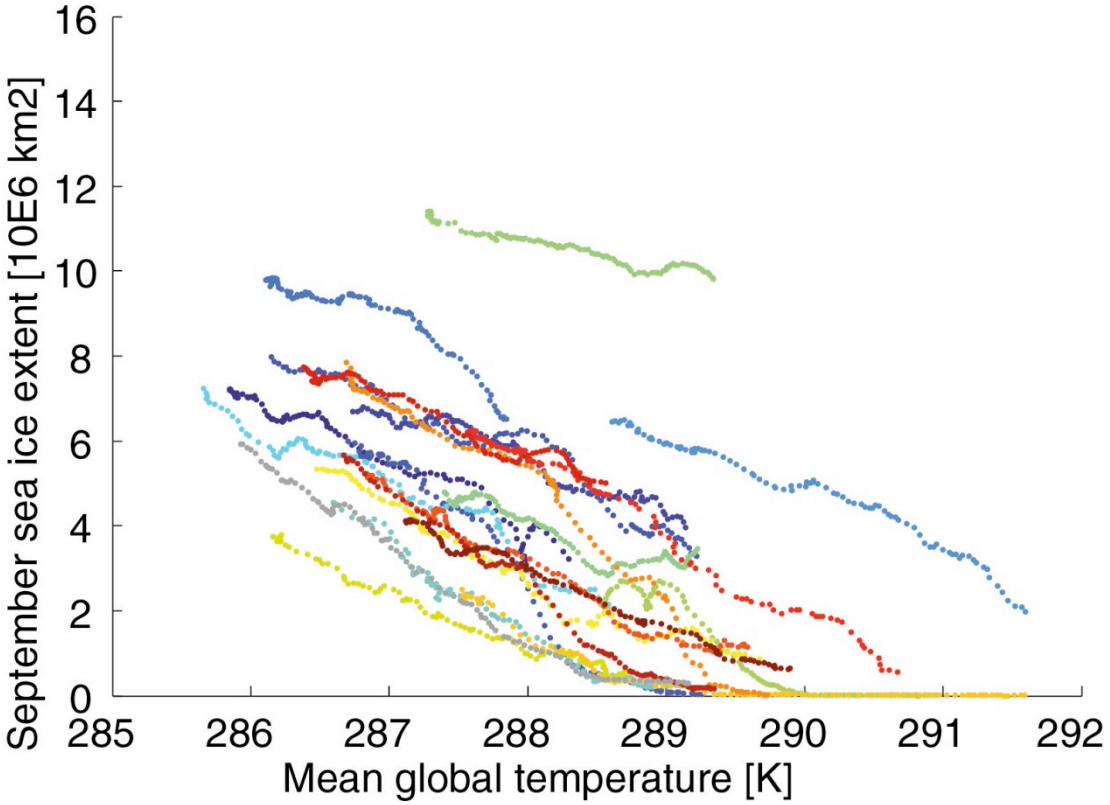
- Relationships derived in one perturbed physics ensemble may not translate easily to other ensembles.



(Sanderson, 2011, submitted)

Sea ice trends

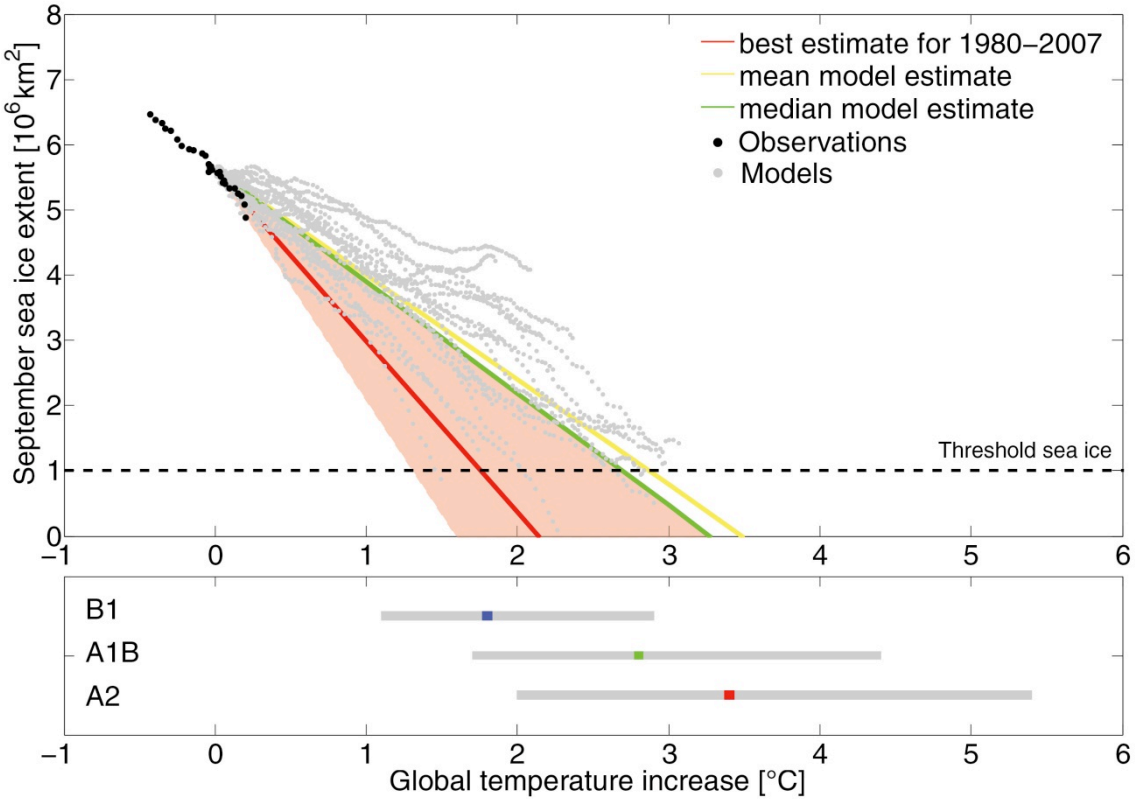
- Sea ice decreases near linearly with temperature in all models, i.e. past trends relate strongly to future trends.



(Mahlstein and Knutti, JGR, submitted)

Sea ice trends

- Observed trends constrain future. Use models and metrics to establish relationship, but no model weighting.



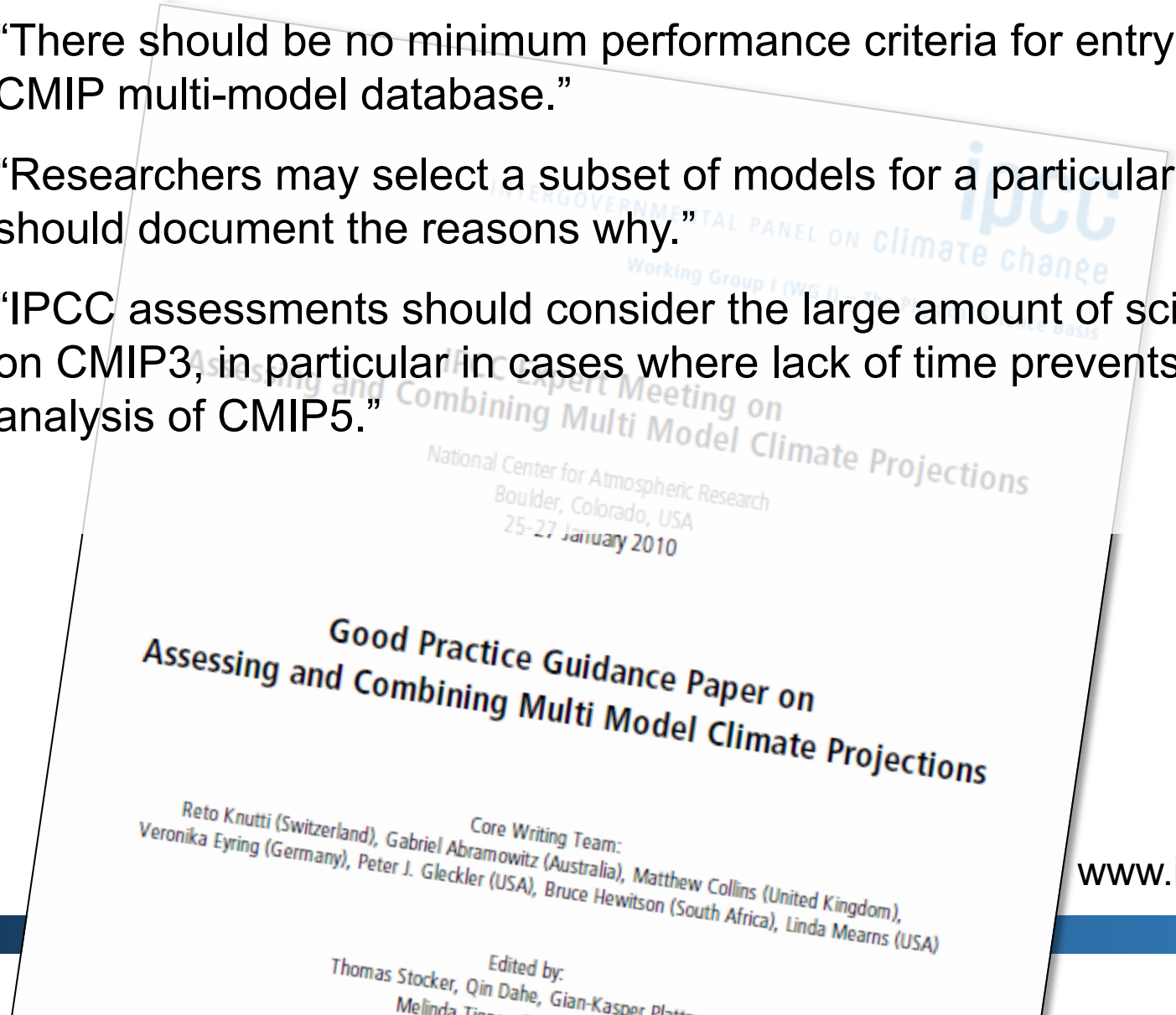
(Mahlstein and Knutti, JGR, submitted)

Metrics and weighting

- Metrics should ideally be simple.
- Metrics should demonstrably be related to the prediction. Finding suitable metrics is likely to be easier for particular purposes.
- Results should be understood in terms of known processes.
- Robust against slight variations in the definition of the metric.
- Observations available with uncertainties sufficiently small to discriminate between models.
- How do we aggregate variables? Units, correlation, observation uncertainties, temporal and spatial coverage?

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."



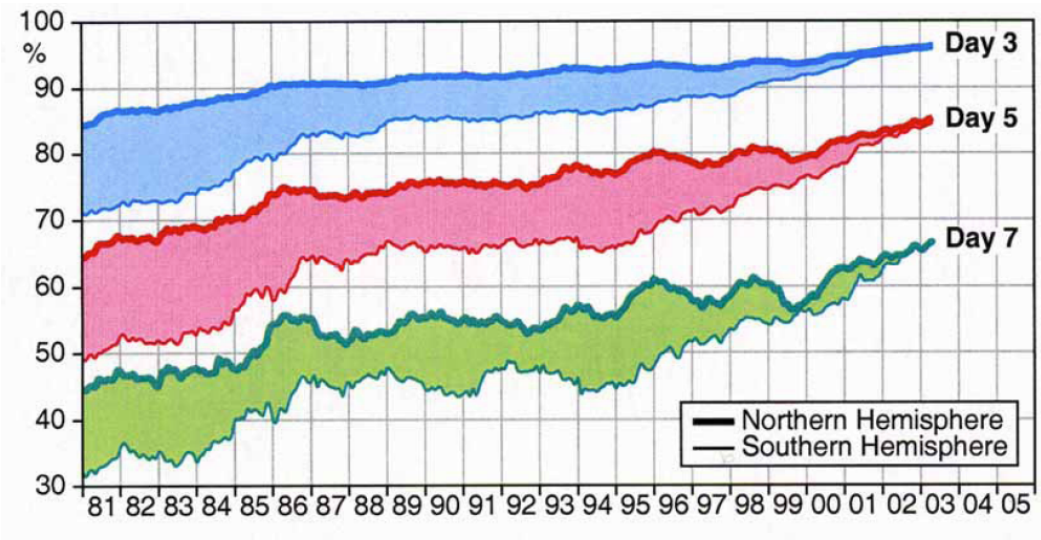
Establishing confidence in a prediction

- Why do we trust the weather forecast for tomorrow but not the forecast for three weeks?

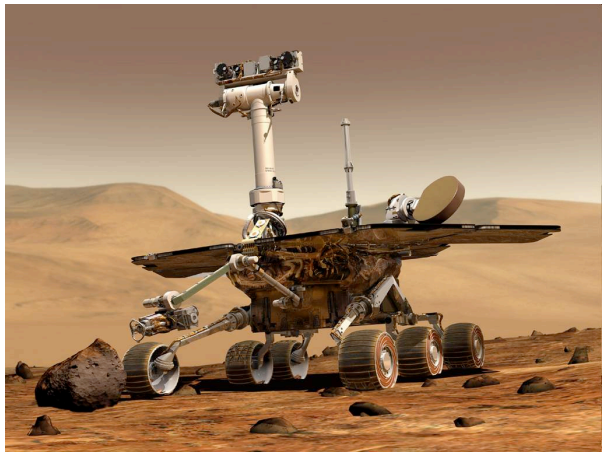
Tuesday Night: Rain and snow likely, becoming all snow after 9pm.

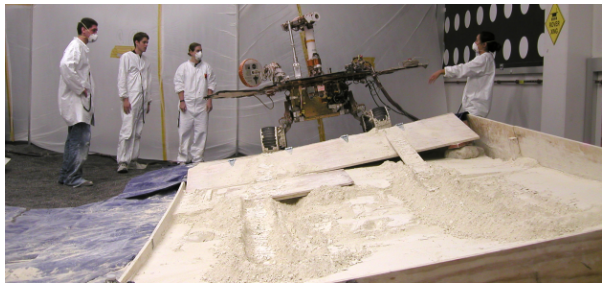
Wednesday: Snow likely

Correlation of hemispheric pressure fields at ~5km altitude



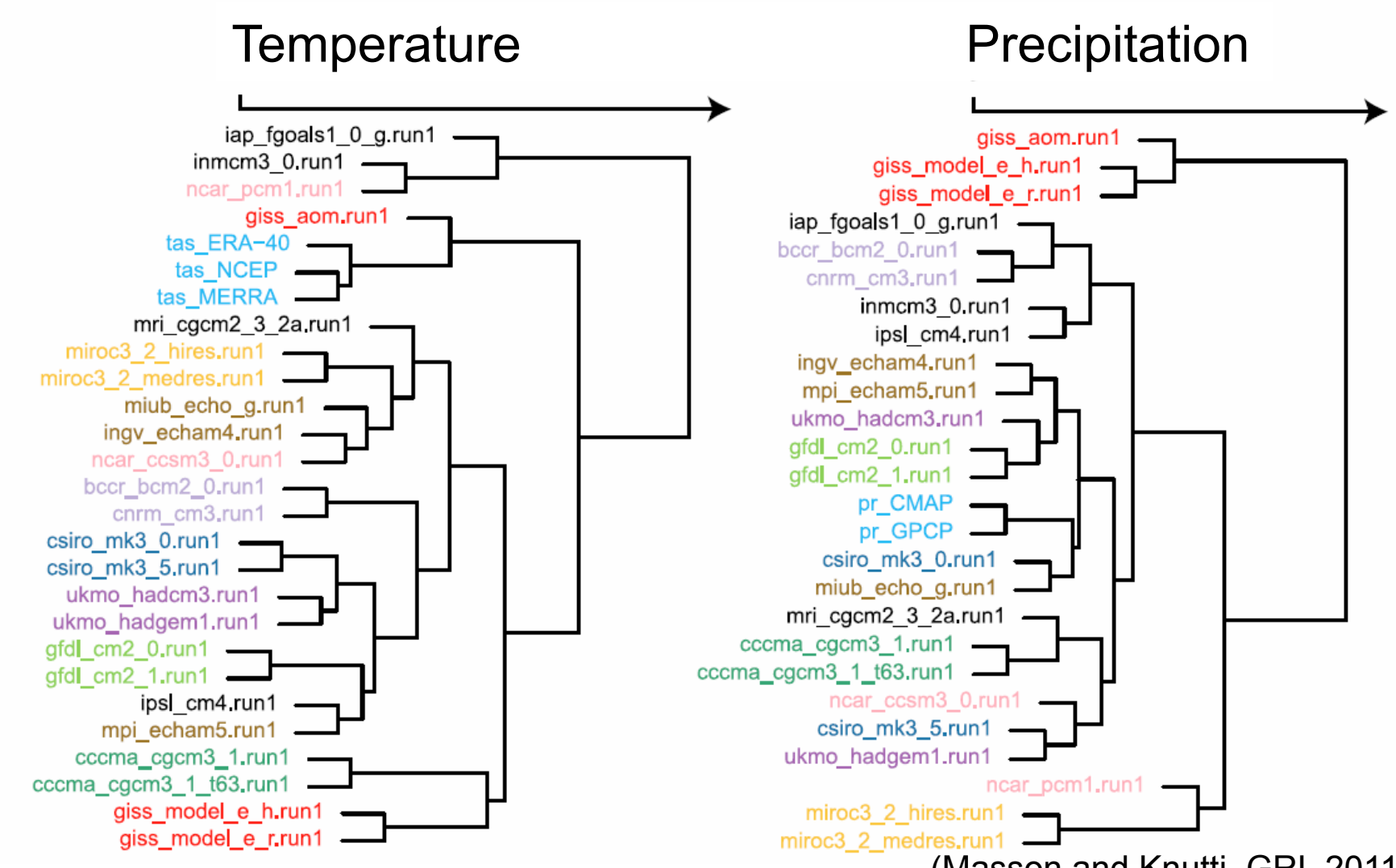
Establishing confidence in a prediction





- Unlike in weather prediction, the confidence in future climate change projections cannot be established by repeated verification.
- We can only test models indirectly. Which tests are most appropriate?

Climate model genealogy



(Masson and Knutti, GRL 2011)

Which model is the best?

- Wrong question. Which models are adequate for purpose?
  Let's not build the best model, let's build many. The idea of developing the best model is strange without defining the purpose first.
- Model performance varies, but most observable metrics provide only a weak constraint on the future. We don't really know how to weight models but implicitly do it by discarding old models.
- Projection spread doesn't decrease. Are we looking at the wrong metric? Are we starting with an sample that is too tight?
- CMIP is an ensemble of opportunity with model dependencies.
- What is a good model? Metrics are a thorny issue. Inappropriate weighting is likely to be worse than no weighting. We may not agree on how to weight, but we should at least talk about it and try to do better than we do now.

Some papers: http://www.iac.ethz.ch/people/knuttir/papers