APPLICATE General Assembly

28-30 January 2019 - ECMWF

Challenges in climate model evaluation

François Massonnet

Challenges in climate model evaluation

1. Standard error metrics are often over-interpreted

2. Model error is not the only cause for mismatch with observations

3. Dealing with uncertainty

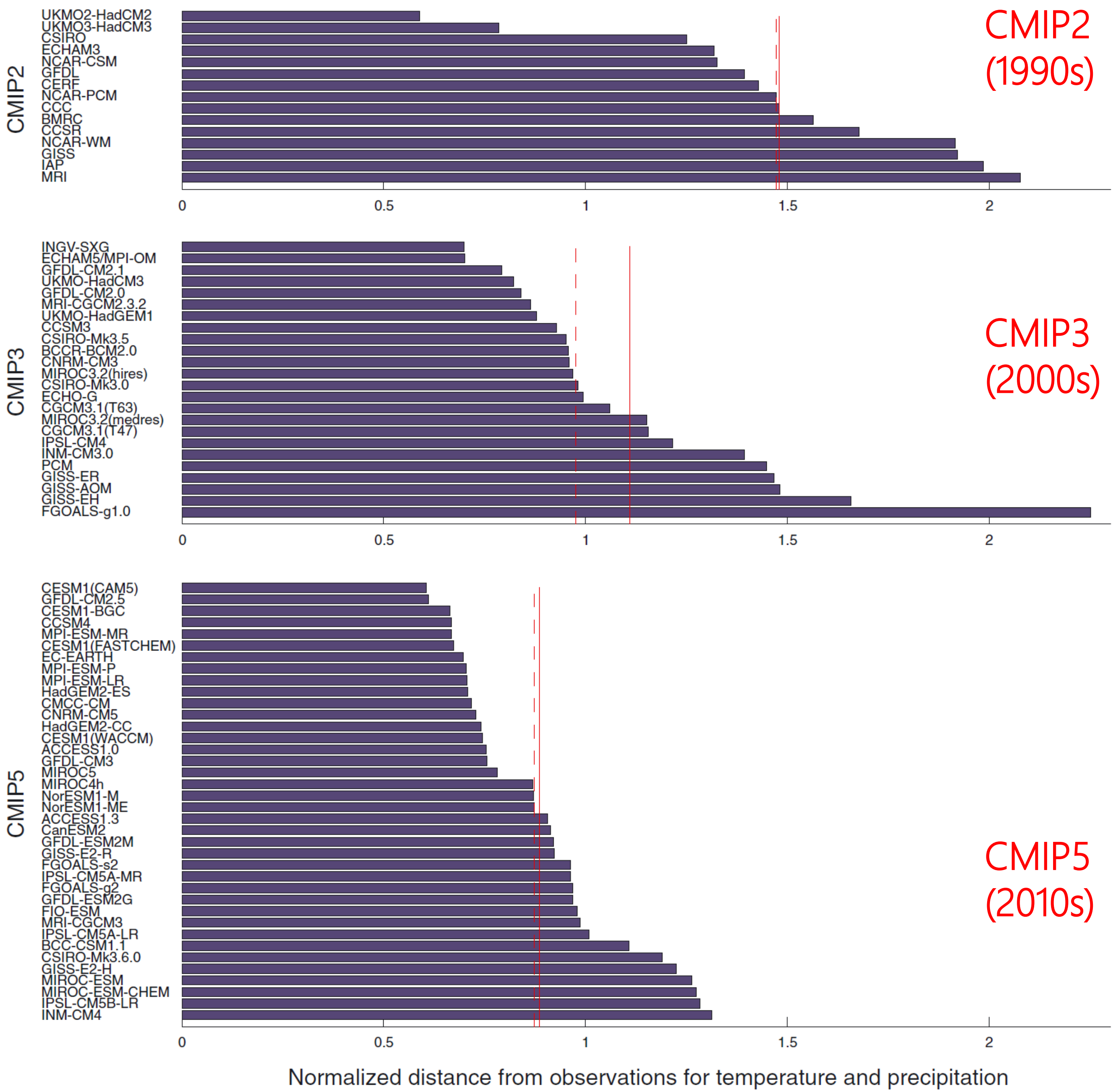
Challenges in climate model evaluation

1. Standard error metrics are often over-interpreted

2. Model error is not the only cause for mismatch with observations

3. Dealing with uncertainty

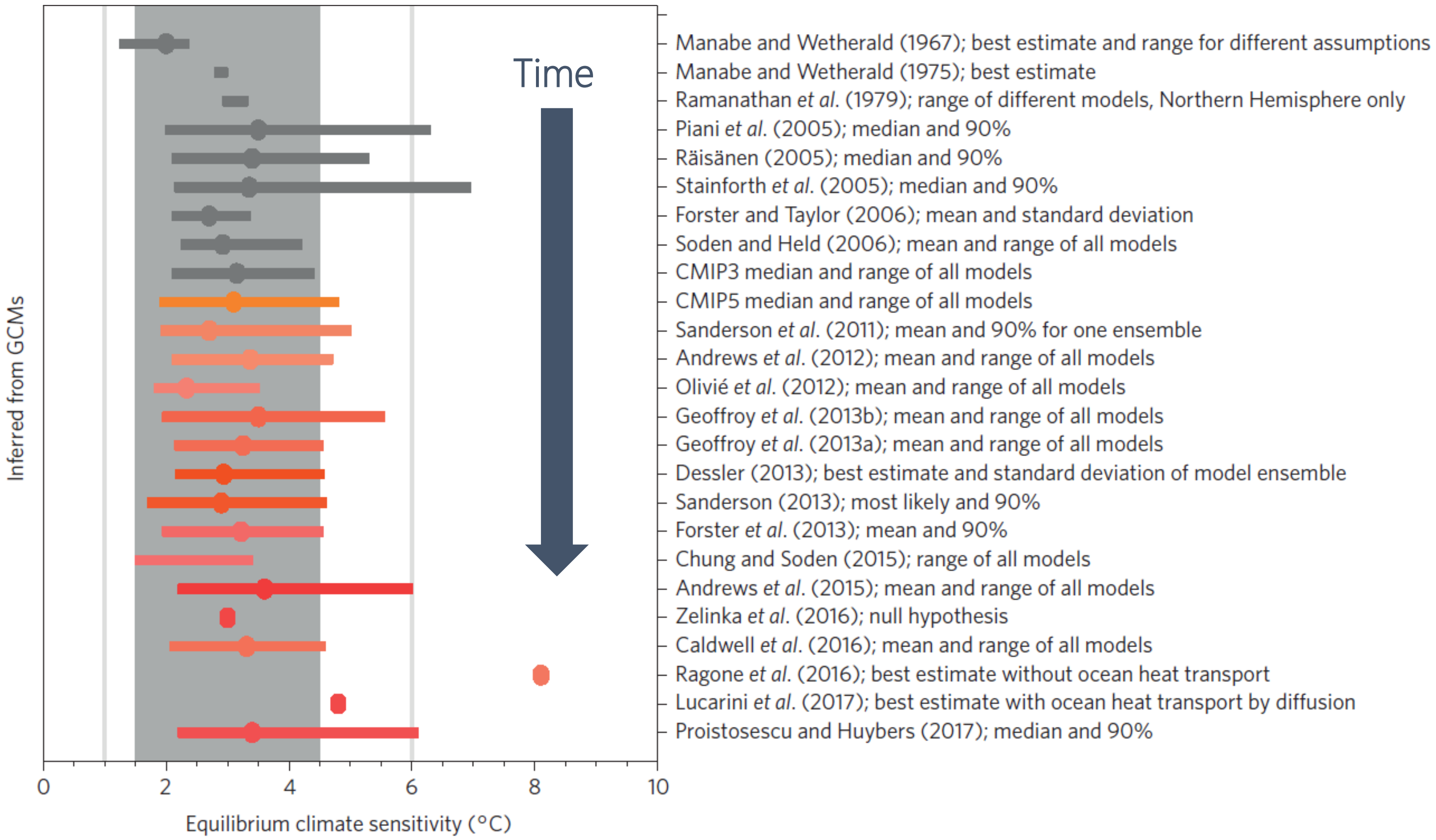
CMIP models are getting better over time…



[Knutti et al., *Geophys. Res. Lett.*, 2013]

… but are they getting more certain?

Estimated equilibrium climate sensitivity from GCMs



[Knutti et al., *Nature Geosci.*, 2017]

Climate models cannot be validated, but they can sometimes be invalidated

ARTICLE

Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences

Naomi Oreskes,* Kristin Shrader-Frechette, Kenneth Belitz

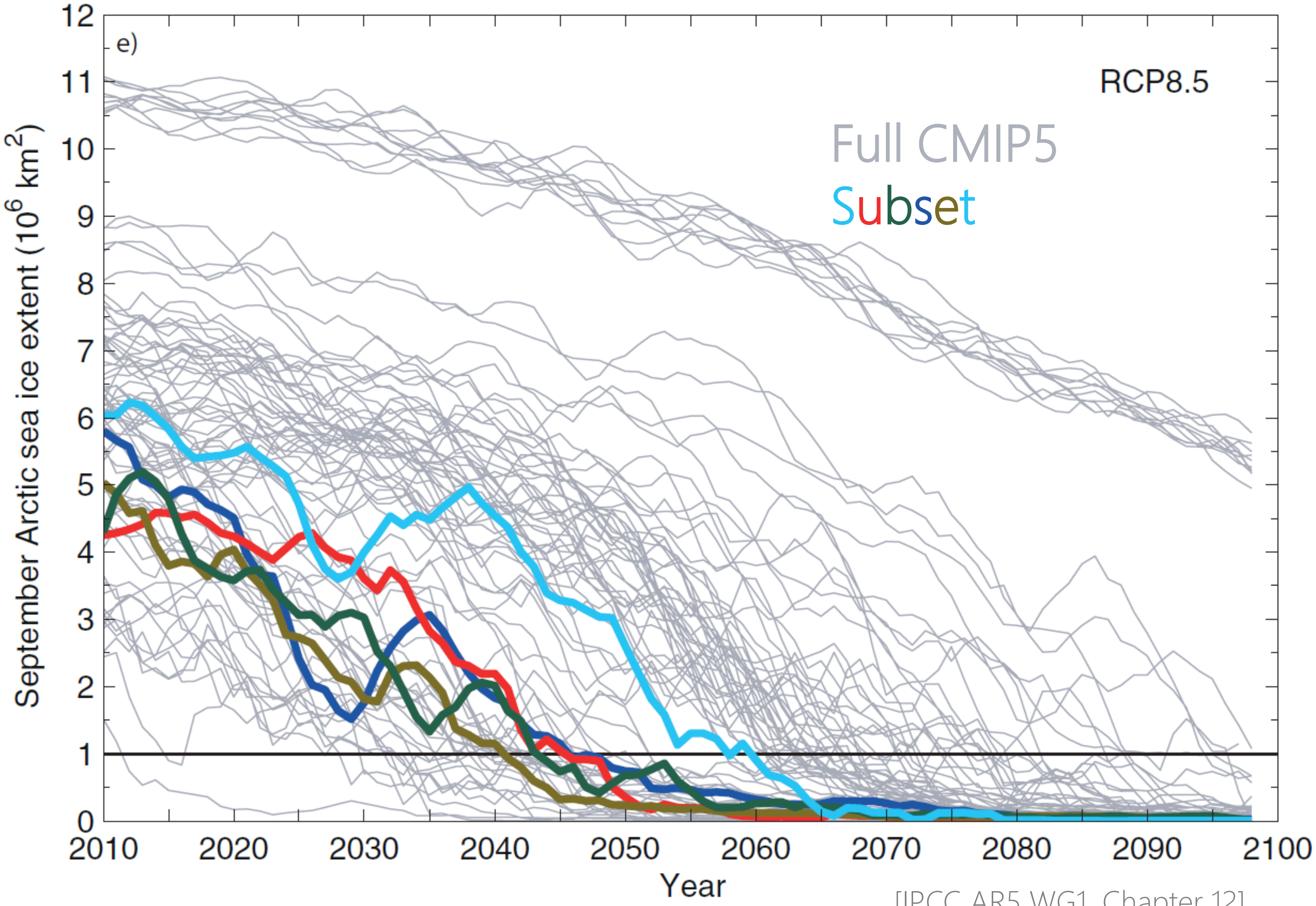
Verification and validation of numerical models of natural systems is impossible. This is because natural systems are never closed and because model results are always non-unique. Models can be confirmed by the demonstration of agreement between observation and prediction, but confirmation is inherently partial. Complete confirmation is logically precluded by the fallacy of affirming the consequent and by incomplete access to natural phenomena. Models can only be evaluated in relative terms, and their predictive value is always open to question. The primary value of models is heuristic.

puter program may be verifiable (*12*). Mathematical components are subject to verification because they are part of closed systems that include claims that are always true as a function of the meanings assigned to the specific symbols used to express them (*13*). However, the models that use these components are never closed systems. One

[Oreskes et al., *Science*, 1994]

Constraining summer Arctic sea ice projections

Projections of September Arctic sea ice extent



[IPCC AR5 WG1, Chapter 12]

Challenges in climate model evaluation

1. Standard error metrics are often over-interpreted

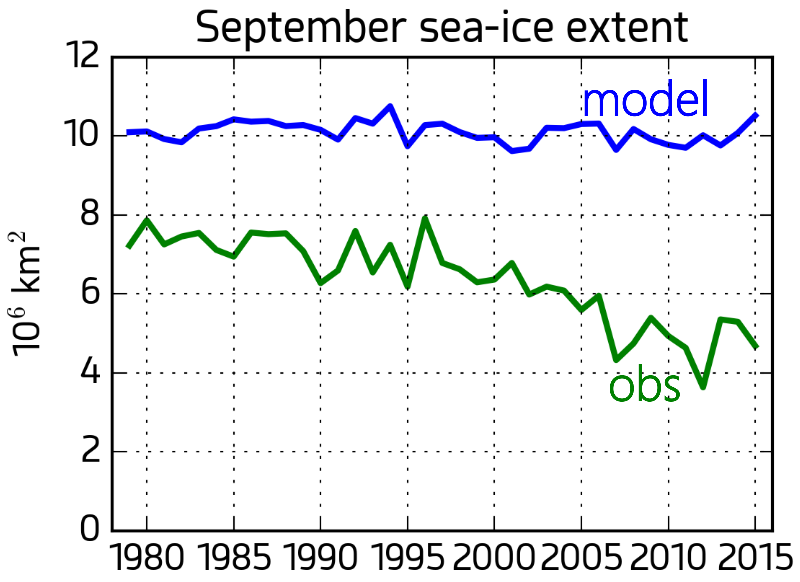
2. Model error is not the only cause for mismatch with observations

3. Dealing with uncertainty

Challenges in climate model evaluation

1. Standard error metrics are often over-interpreted
2. Model error is not the only cause for mismatch with observations
3. Dealing with uncertainty

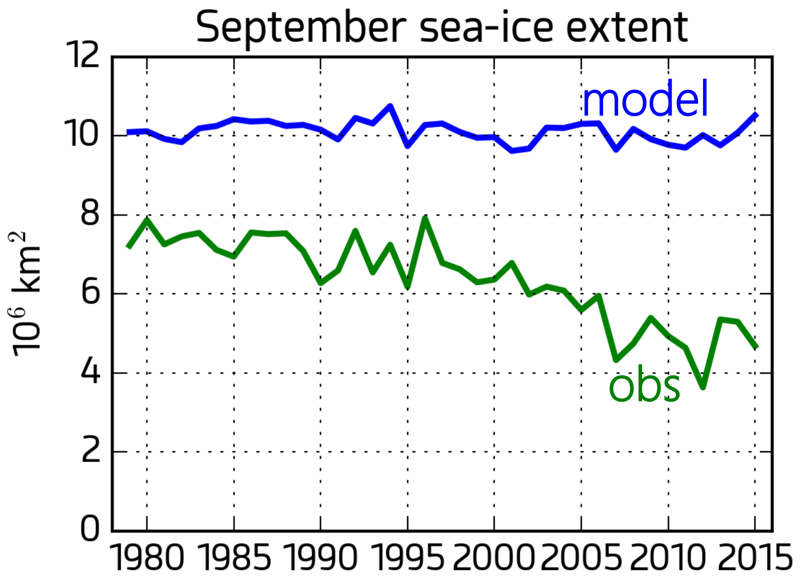
Why don't models and observations match each other?



Why don't models and observations match each other?

It's the modelers fault

-Physical equations are wrong
-Equations are discretized
-Forcing is not correct
-Initial conditions are not correct
-Processes are parameterized
-There are computational errors

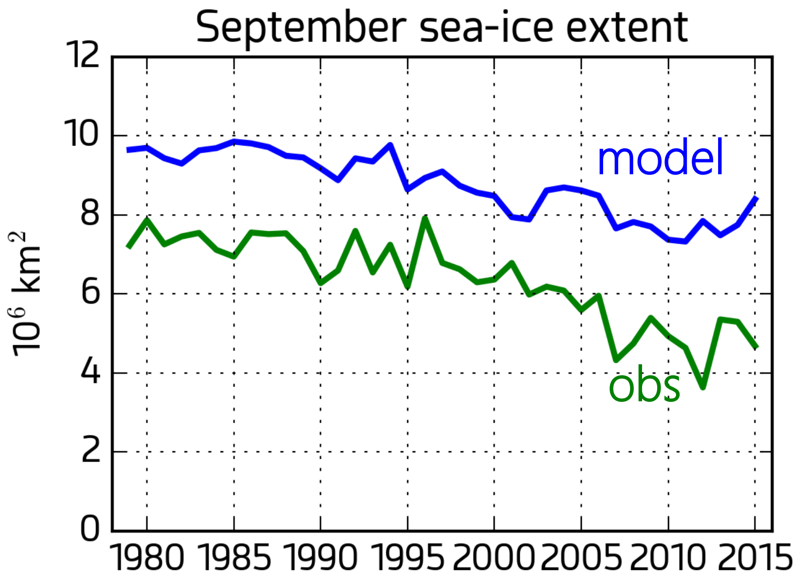


[Orrell et al., *Nonlin. Proc. Geophys.*, 2001]

Why don't models and observations match each other?

It's the modelers fault

-Physical equations are wrong
-Equations are discretized
-Forcing is not correct
-Initial conditions are not correct
-Processes are parameterized
-There are computational errors



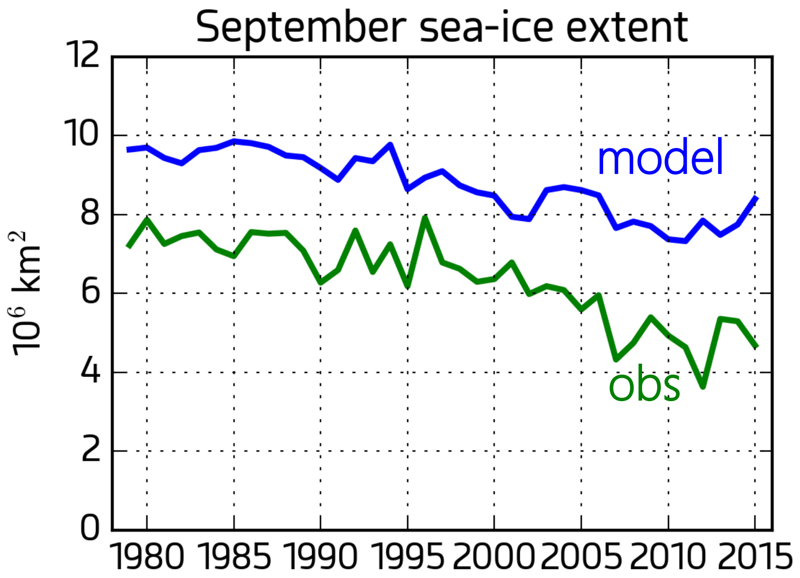
[Orrell et al., *Nonlin. Proc. Geophys.*, 2001]

Why don't models and observations match each other?

It's the modellers fault

It's the observers fault

Instrumental errors
Algorithm errors
Assumptions (e.g. hydrostatic)
Sampling errors



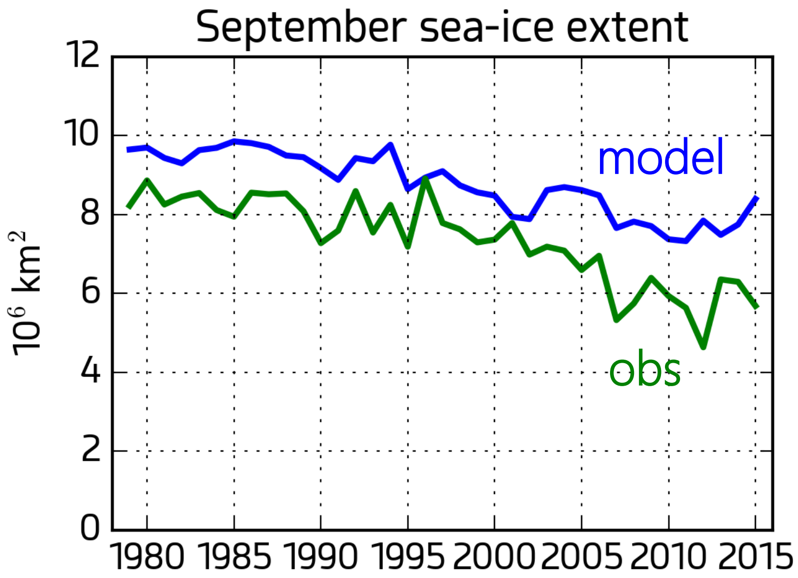
[Ivanova et al., *Cryosphere*, 2014; Zygmuntowska et al., *Cryosphere*, 2014; Worby et al., *J. Geophys. Res.*, 2008]

Why don't models and observations match each other?

It's the modellers fault

It's the observers fault

Instrumental errors
Algorithm errors
Assumptions (e.g. hydrostatic)
Sampling errors

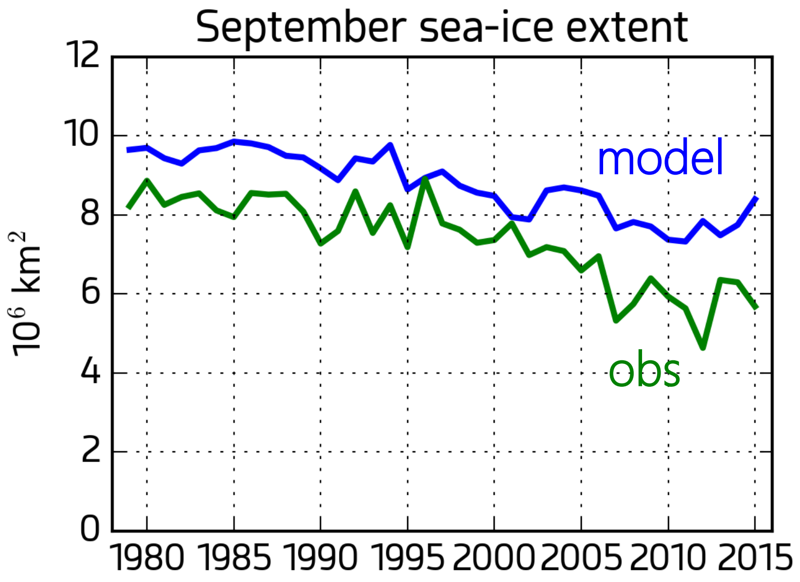


[Ivanova et al., *Cryosphere*, 2014; Zygmuntowska et al., *Cryosphere*, 2014; Worby et al., *J. Geophys. Res.*, 2008]

Why don't models and observations match each other?

It's the modellers fault

It's the observers fault



It's my fault

No scale-awareness

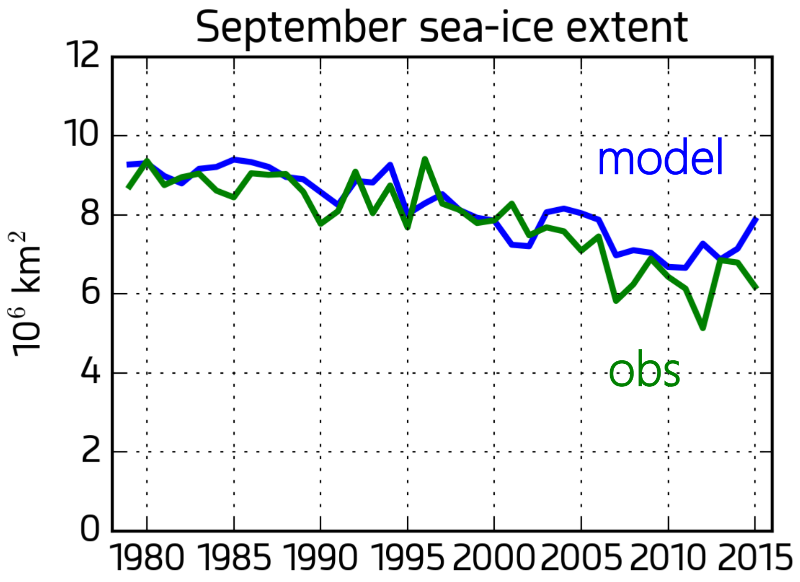
No definition-awareness

[Kay et al., *J. Geophys. Res.*, 2016]

Why don't models and observations match each other?

It's the modellers fault

It's the observers fault



It's my fault

No scale-awareness

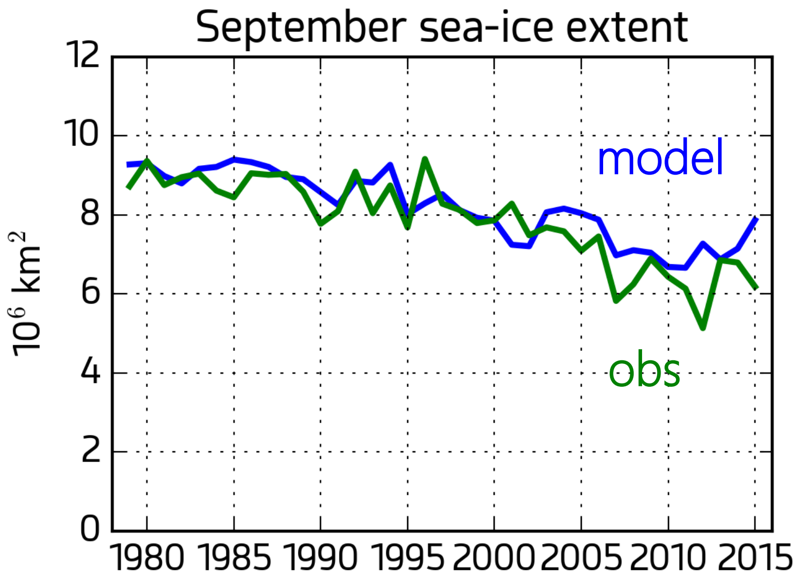
No definition-awareness

[Kay et al., *J. Geophys. Res.*, 2016]

Why don't models and observations match each other?

It's the modellers fault

It's the observers fault



It's my fault

No scale-awareness

No definition-awareness

It's no one's fault

Internal variability

[Notz, *Phil. Trans. Roy. Soc.*, 2015]

Challenges in climate model evaluation

1. Standard error metrics are often over-interpreted

2. Model error is not the only cause for mismatch with observations

3. Dealing with uncertainty

Challenges in climate model evaluation

1. Standard error metrics are often over-interpreted

2. Model error is not the only cause for mismatch with observations

3. Dealing with uncertainty

Comput. Methods Appl. Mech. Engrg. 200 (2011) 2131–2144

ELSEVIER

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing

Christopher J. Roy [a,*], William L. Oberkampf [b]

[a] *Aerospace and Ocean Engineering Department, Virginia Tech, 215 Randolph Hall, Blacksburg, Virginia 24061, USA*
[b] *Consulting Engineer, 5112 Hidden Springs Trail, Georgetown, Texas 78633, USA*

[Roy and Oberkampf, *Comput. Methods Appl. Mech. Engrg.*, 2011]

Classification of uncertainty

Uncertainty in a prediction

« Aleatoric »

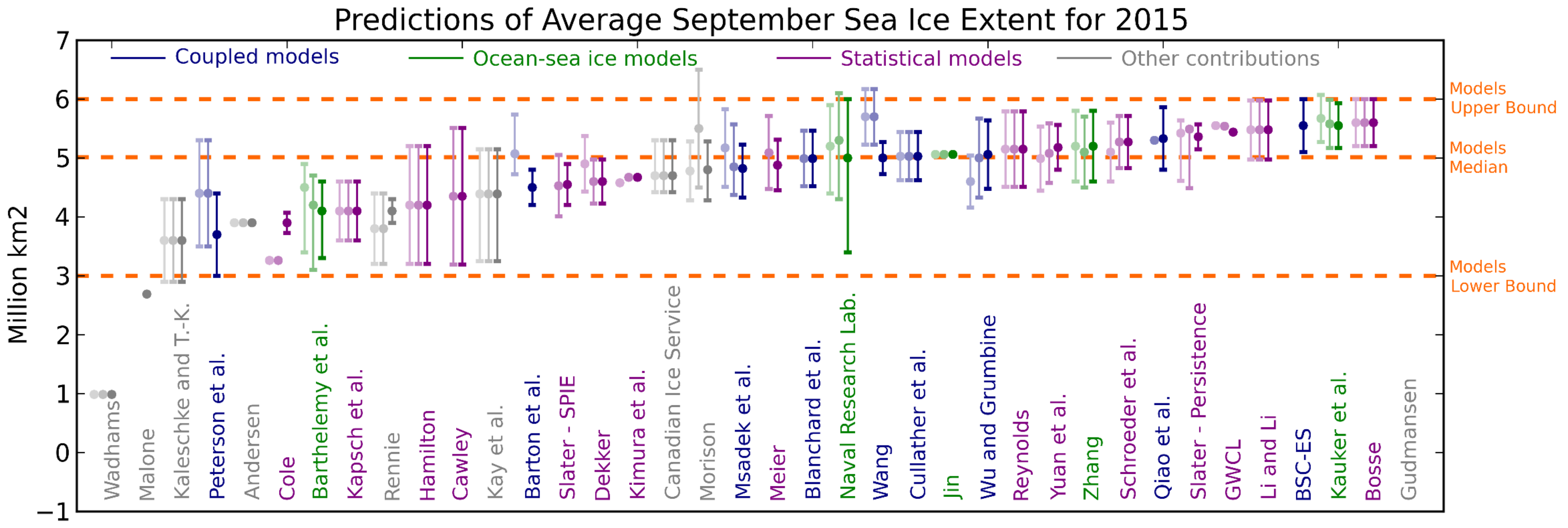
- Due to random effects
- Characterized by a PDF (frequentist interpretation)
- Irreducible

« Epistemic »

- Due to ignorance
- Characterized by an interval
- Can, in principle, be eliminated

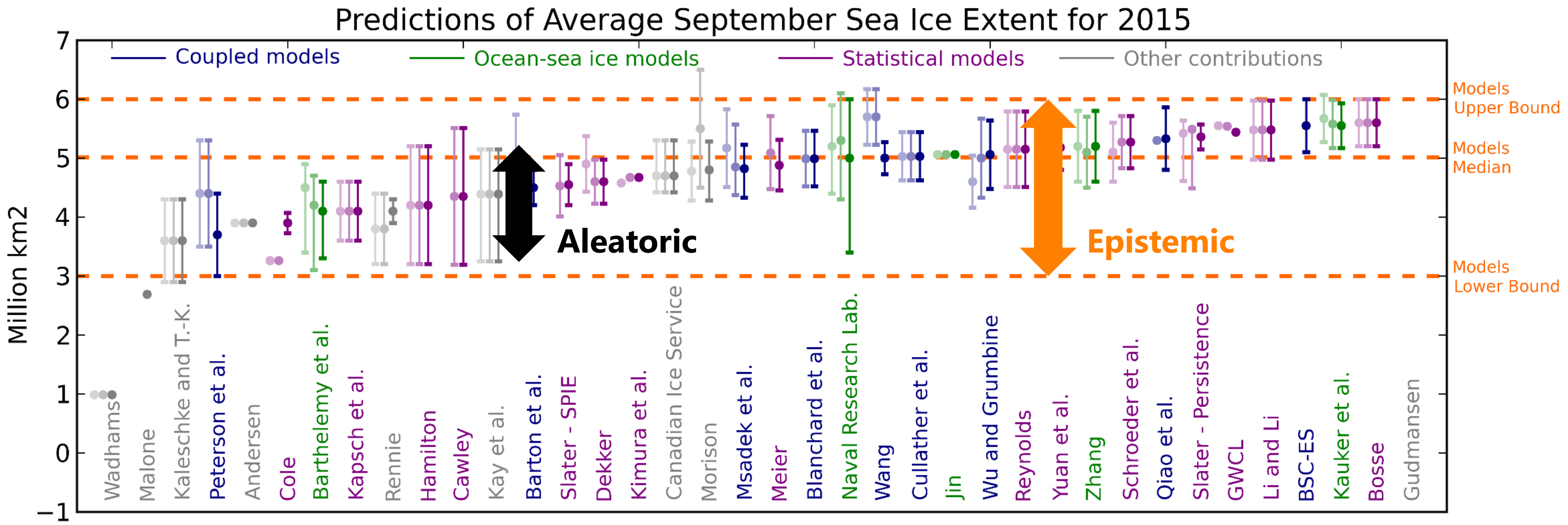
[Roy and Oberkampf, *Comput. Methods Appl. Mech. Engrg.*, 2011]

Aleatoric *vs* epistemic uncertainty in the Sea Ice Outlooks



https://www.arcus.org/sipn/sea-ice-outlook/2015/post-season

Aleatoric *vs* epistemic uncertainty in the Sea Ice Outlooks



https://www.arcus.org/sipn/sea-ice-outlook/2015/post-season

Challenges in climate model evaluation

1. Standard error metrics are often over-interpreted
2. Model error is not the only cause for mismatch with observations
3. Dealing with uncertainty

Some food for break-out group discussions

- Do APPLICATE metrics account for aleatoric uncertainty?
    - e.g.: observational error, internal variability?
- Can you list all the reasons why your simulated and observed diagnostics could differ from each other?
- Can models be used to guide the development of future Arctic observing systems? How?
- What are the model developments that should go in CMIP7? How do we decide?

"distance" in geometry. Ideally, they should be defined according to a set of axioms too (such as positivity, triangle inequality, symmetry, nullity). Several types of metrics must be distinguished from each other:

- ***Standard error metrics*** are developed in order to check the overall consistency of a model or prediction system with a reference. Standard error metrics are useful: they put pressure on centers to be responsive in addressing obvious model biases, but they also allow for tracking the first-order evolution of model development through time (Gleckler et al., 2008; Reichler and Kim, 2008; Eyring et al., 2016). Such metrics should be handled by "responsible adults" because they are easily over-interpreted. For instance, a model may simulate a realistic trend in annual-mean, global-mean near-surface air temperature, but thanks to the cancellation of major regional biases. Ideally, standard error metrics metrics should never be computed in isolation (e.g. for one specific variable) but rather be part of an overall assessment process – this would allow an instant visualization of the system's consistency with the reference(s) as a whole.

"The root mean squared error of Arctic sea ice thickness in my model is 1.2 m over 2004-2008, compared to the ICESat sea ice thickness dataset."

(Standard error metric)

From APPLICATE WP1 Model Assessment Plan