



Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada



# Physical systematic biases

**Aspen Model Evaluation Workshop**

**Greg Flato**

**Canadian Centre for Climate  
Modelling and Analysis**

**August, 2017**

# Issues

---

- Many large-scale errors/biases persist from one generation of models to another; if they were well understood and easy to fix, they wouldn't persist.
- Our ability to quantify these biases depends on observational quality and on the confounding effects of internal variability.
- Will more systematic evaluation help? How much do errors/biases affect model fitness for purpose? Can we focus on the 'essential' ones? Is weighting/selection really a good idea? Is post-processing (e.g. bias correction) helpful?

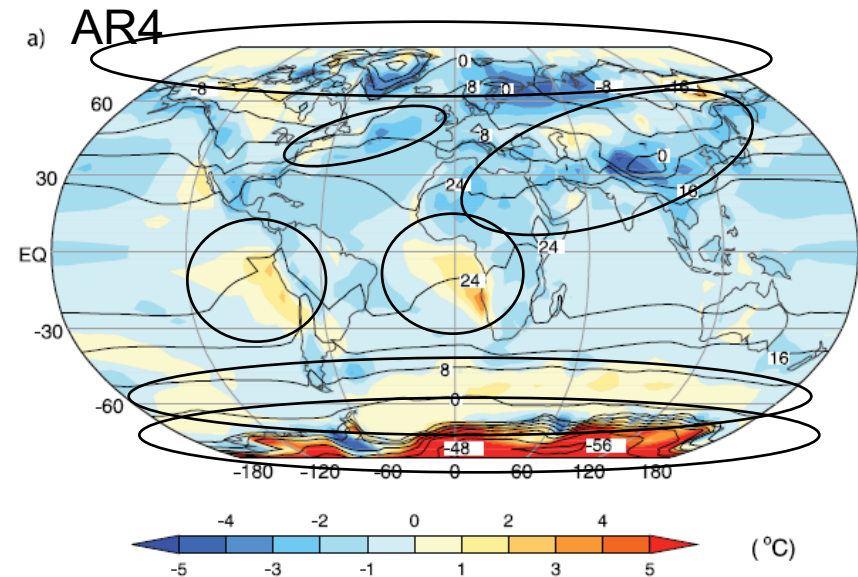
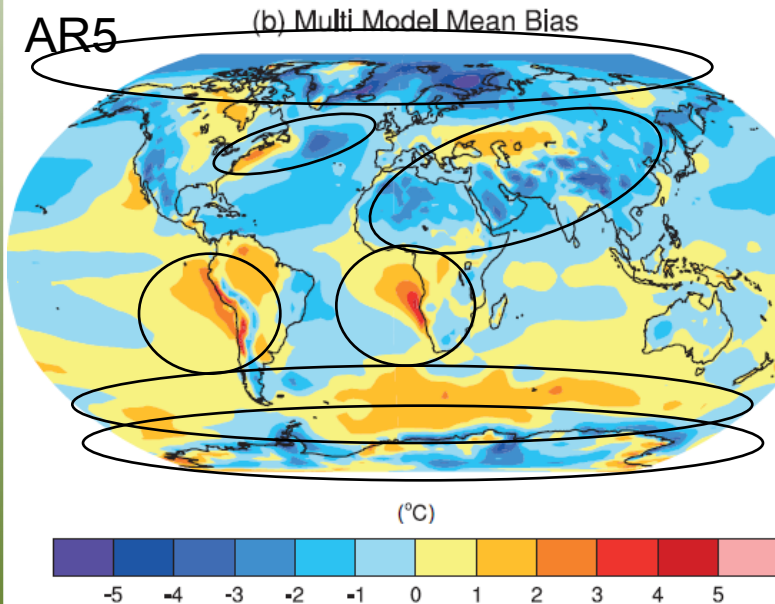


# IPCC Assessments ...

- ... have shown some aspects of persistent, systematic errors, and temperature is indicative.

Not all systematic errors are persistent

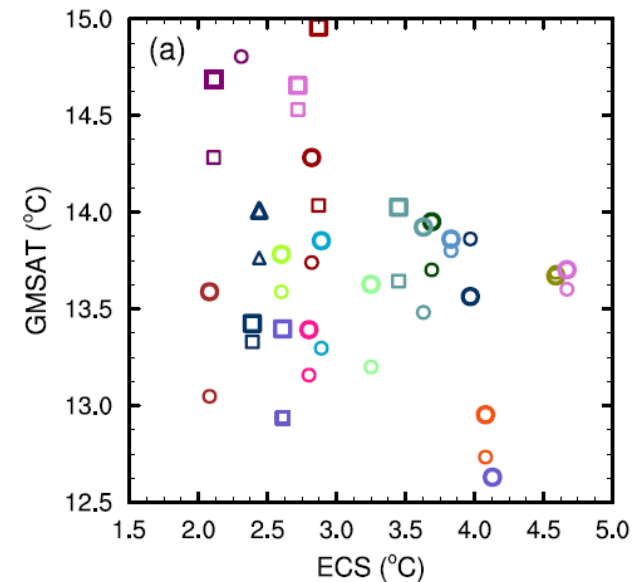
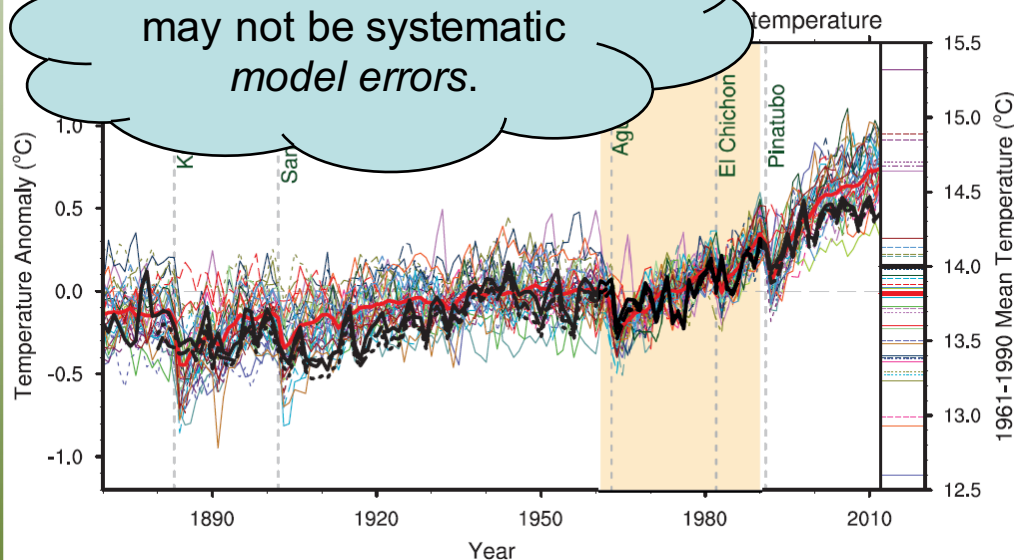
Annual mean surface air temperature



Not all errors or biases are of equal importance – different applications will place different demands on model quality.

- How important are these biases?
- Some you think would be important, don't seem to be.

Systematic differences may not be systematic model errors.



IPCC AR5, Ch. 9

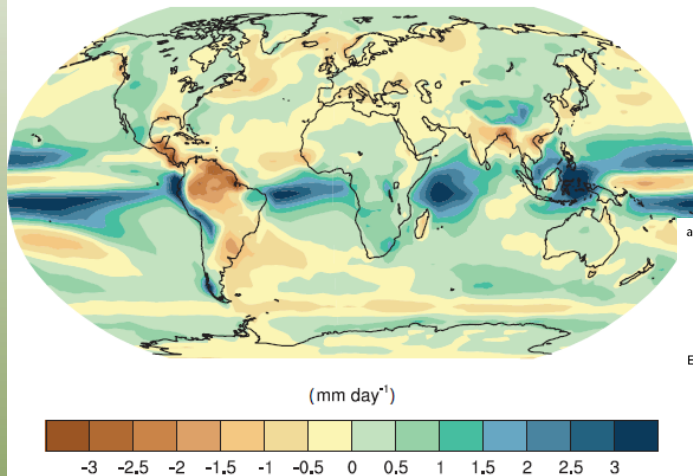


For some quantities, like precipitation, changing presentation format in IPCC reports makes tracking systematic errors difficult

## The famous 'double ITCZ'

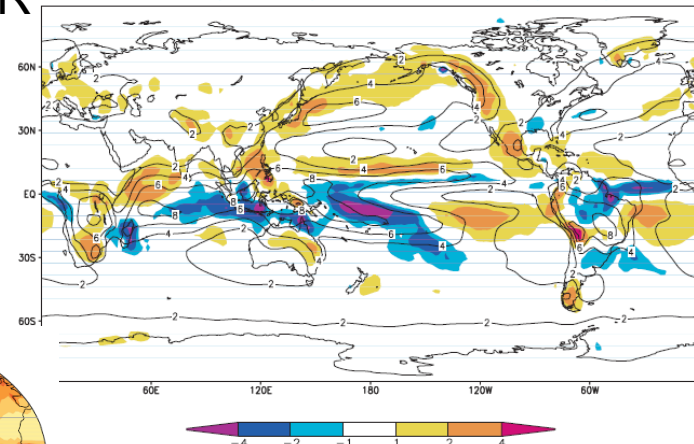
AR5

(b) Multi Model Mean Bias

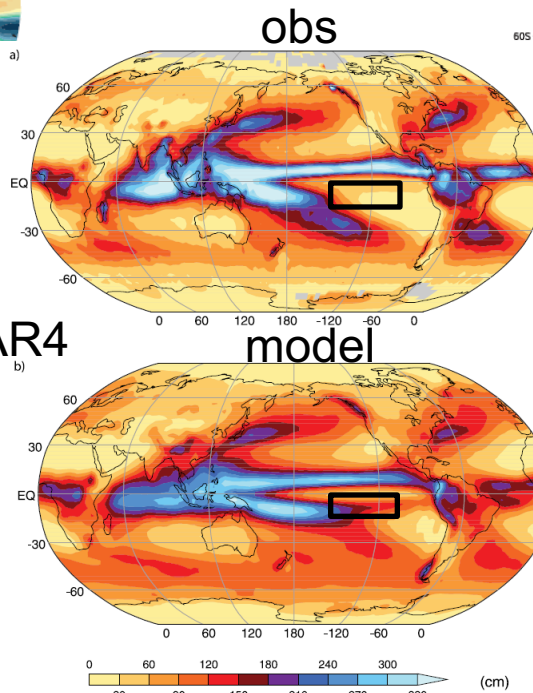


TAR

Model precipitation (DJF)  
Model mean (contoured) mean minus observed (shaded)



AR4



More systematic evaluation tools (applied retrospectively where possible) would improve this.

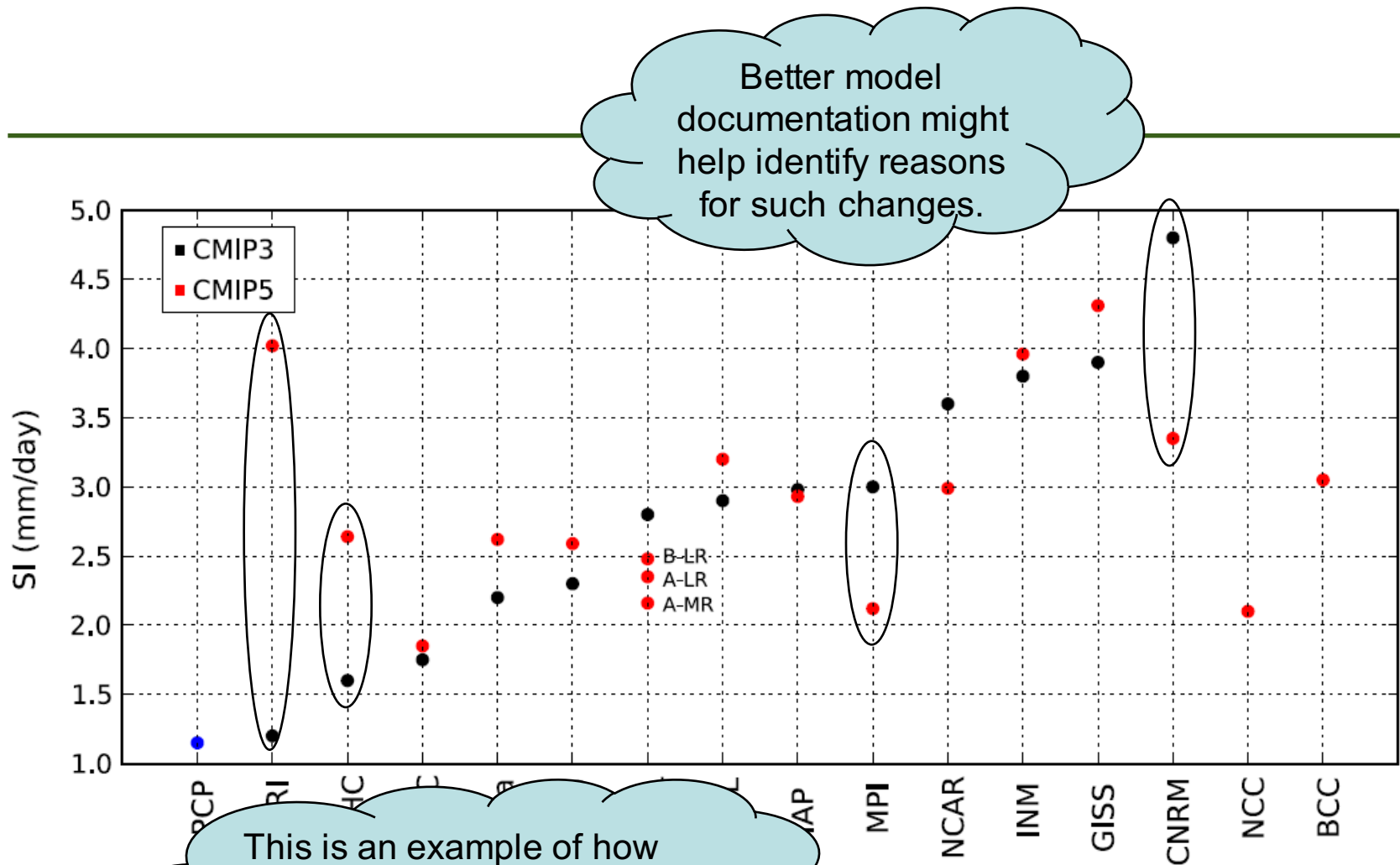


Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada

## Little evidence for improvement in this particular bias



This is an example of how more 'automated' diagnostics could benefit modelling groups

*Deslati and Bellon, Clim. Dyn., 2015*

Page 6 – October 24, 2017



Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada

## Other systematic errors (list courtesy of Ron Stouffer):

---

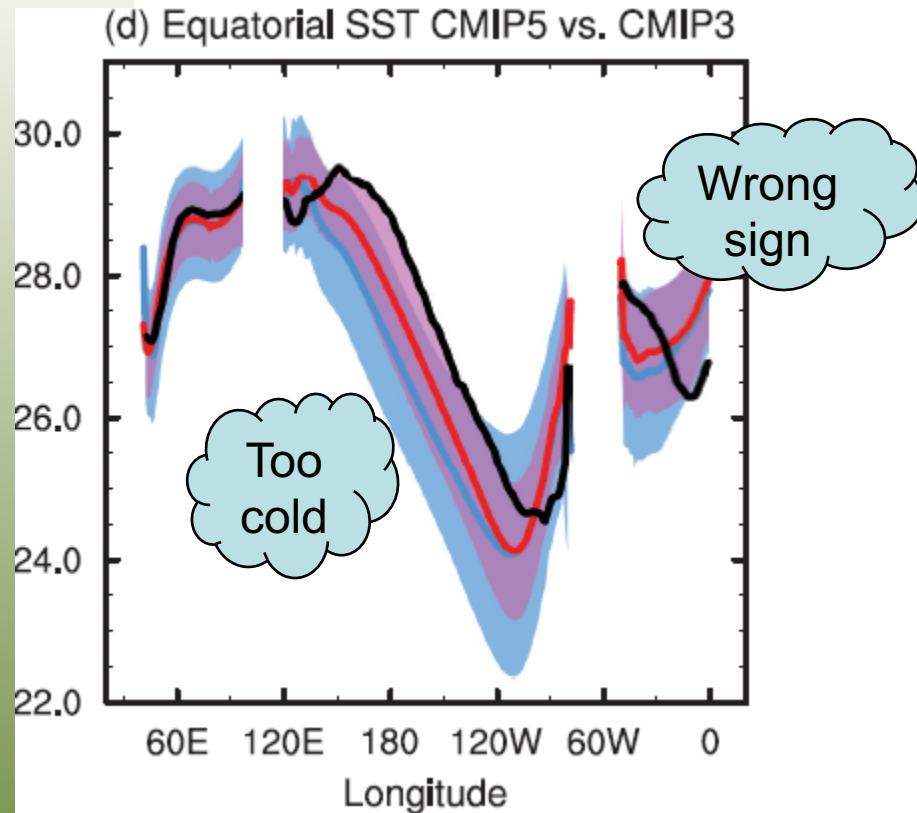
1. Double ITCZ
2. Clouds in general,
  - a. low clouds over low latitude oceans
  - b. radiative properties
3. Oceanic heat uptake – Southern Ocean in particular
4. Sea ice – both hemispheres (coverage and thickness and seasonal cycle)
5. Precipitation over land, river flow
6. Diurnal cycle – T, P
7. Oceanic mixing both near surface and in regions of topography;
  - a. horizontal and vertical => eddies, tides and etc.
8. Magnitude of oceanic boundary flows
9. AMOC strength and poleward heat transport (model mismatch with obs)
  - a. Model typically have too large volume transport and too little heat transport
  - b. High resolution may improve things



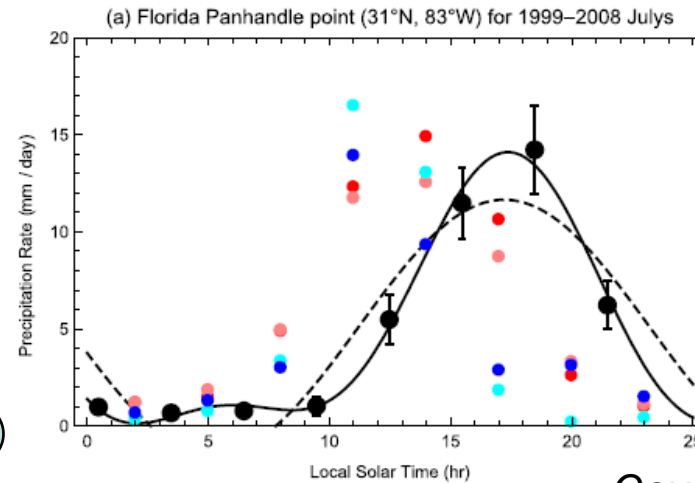


# Diurnal cycle of Precipitation

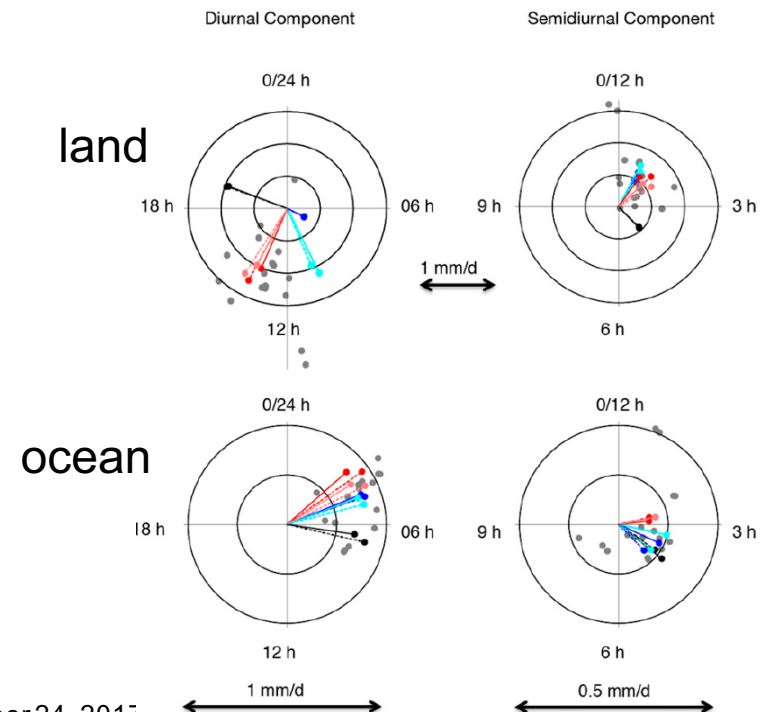
## Sea-Surface Temperature



IPCC AR5, Ch. 9



Covey et al., 2016





Lots of new diagnostic methods, new observations, new errors identified ....



The screenshot shows the official website for the 5th WGNE workshop. At the top, there's a navigation bar with links for Canada.ca, Services, Departments, and Français. Below this is a search bar and a navigation menu with links: Home, About, Scientific Committee, Keynote Speakers, Programme, Posters, Instructions, Venue, and Local Information. The main heading is "5th WGNE workshop on systematic errors in weather and climate models", dated "June 19-23, 2017, Montréal, Québec, Canada". Logos for the World Meteorological Organization, WCRP, MAPP, and OURANOS are displayed. The "Description" section explains that the workshop is organized by the WCRP-JSC/CAS Working Group on Numerical Experimentation (WGNE) and hosted by Environment and Climate Change Canada (ECCC) in Montreal during June 19-23, 2017. It outlines the goal of increasing understanding of model errors and lists five themes: Atmosphere-land-ocean-cryosphere interactions, Clouds and precipitation, Resolution issues, Teleconnections, and Metrics and diagnostics, and Model errors in ensembles.

## 5th WGNE workshop on systematic errors in weather and climate models

June 19-23, 2017, Montréal, Québec, Canada






### Description

The WCRP-JSC/CAS Working Group on Numerical Experimentation (WGNE) is organizing a workshop on systematic errors in weather and climate models, to be hosted by Environment and Climate Change Canada (ECCC) in Montreal during **19-23 June 2017**.

The principal goal of the workshop is to increase understanding of the nature and cause of errors in models used for weather and climate prediction, including intra-seasonal to inter-annual scales. Of special interest will be studies that consider errors found in multiple models and errors which are present across timescales. Diagnostics and metrics to identify and characterize systematic errors are welcomed.

Considering recent reports from WGNE members, affiliated centres and groups, WGNE has identified some processes that models currently fail to represent accurately. The workshop is therefore organized around the following themes:

- **Atmosphere-land-ocean-cryosphere interactions:** errors in the representation of surface fluxes and drag processes; stable boundary layer issues; impact of coupled modeling.
- **Clouds and precipitation:** cloud-radiative feedback problem; tropical convection issues; representation of low clouds, especially at high latitudes; excess low accumulations of precipitation; underestimation of precipitation extremes; summer continental precipitation; precipitation over orography.
- **Resolution issues:** dependence of systematic errors on model resolution; grey zones of physical parametrizations.
- **Teleconnections\*:** errors in the simulation of interactions between high-latitudes, mid-latitudes and tropics.
- **Metrics and diagnostics:** emphasis on novel techniques (e.g. process-based diagnostics; use of data assimilation or coupled modeling) to diagnose and measure systematic errors.
- **Model errors in ensembles:** characterization of ensemble spread and identification of systematic errors in multi-model ensembles and ensemble prediction systems; evaluation of stochastic representations.




We need to go beyond  
'documenting' model  
errors/biases.

In the IPCC 5<sup>th</sup> Assessment  
we struggled to make a  
connection between the  
model evaluation chapter and  
the later chapters where  
models were used (in D&A, in  
future projections, in driving  
RCMs, etc.)

Workshop on Connecting Climate Model  
Evaluation to Assessing Fitness-for-Purpose

Max-Planck Institute for Meteorology, Hamburg, Germany  
23-24 February, 2017



**Funding/Sponsorship**

This meeting was sponsored by the World Climate Research Programme (WCRP) which supported travel of developing-country scientists, the IPCC Working Group I Technical Support Unit, providing organizational assistance, the Max-Planck Institute for Meteorology, providing meeting space, coffee/tea breaks, and group dinner.

ipcc  
INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE

Max-Planck-Institut  
für Meteorologie

Page 10 – October 24, 2017



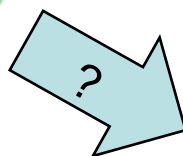
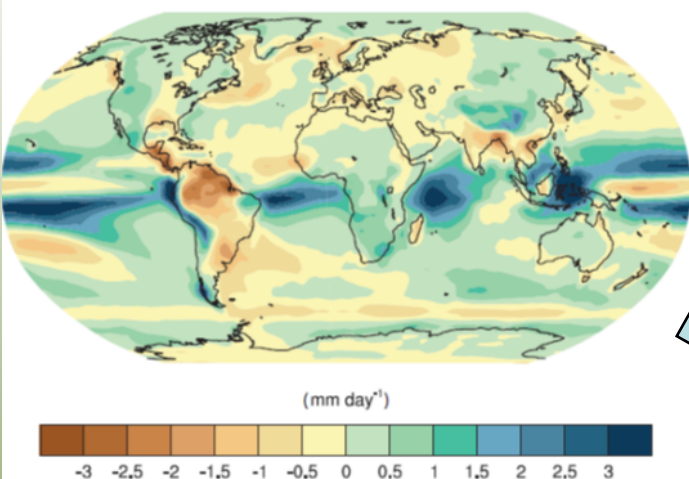
Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada

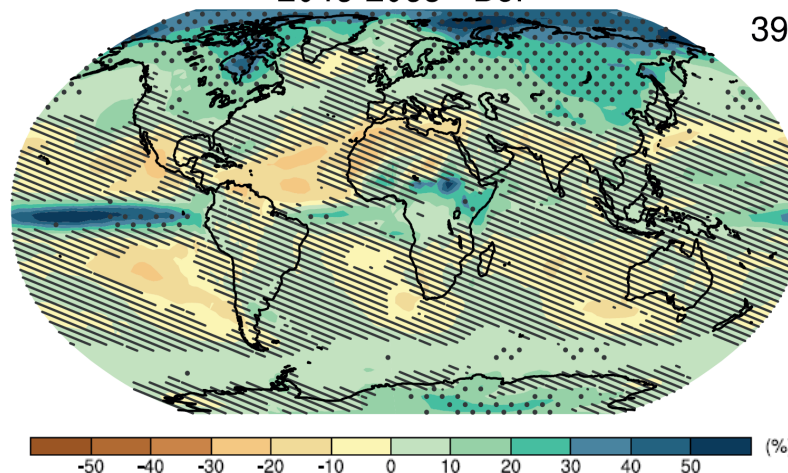
In other words ... we have a growing suite of performance measures, but what we'd like to know is how errors/biases affect, e.g., a model projection of future change.

(b) Multi Model Mean Bias



Multi-model mean precipitation change, relative to 1985-2005 (RCP8.5).

2046-2065 - DJF

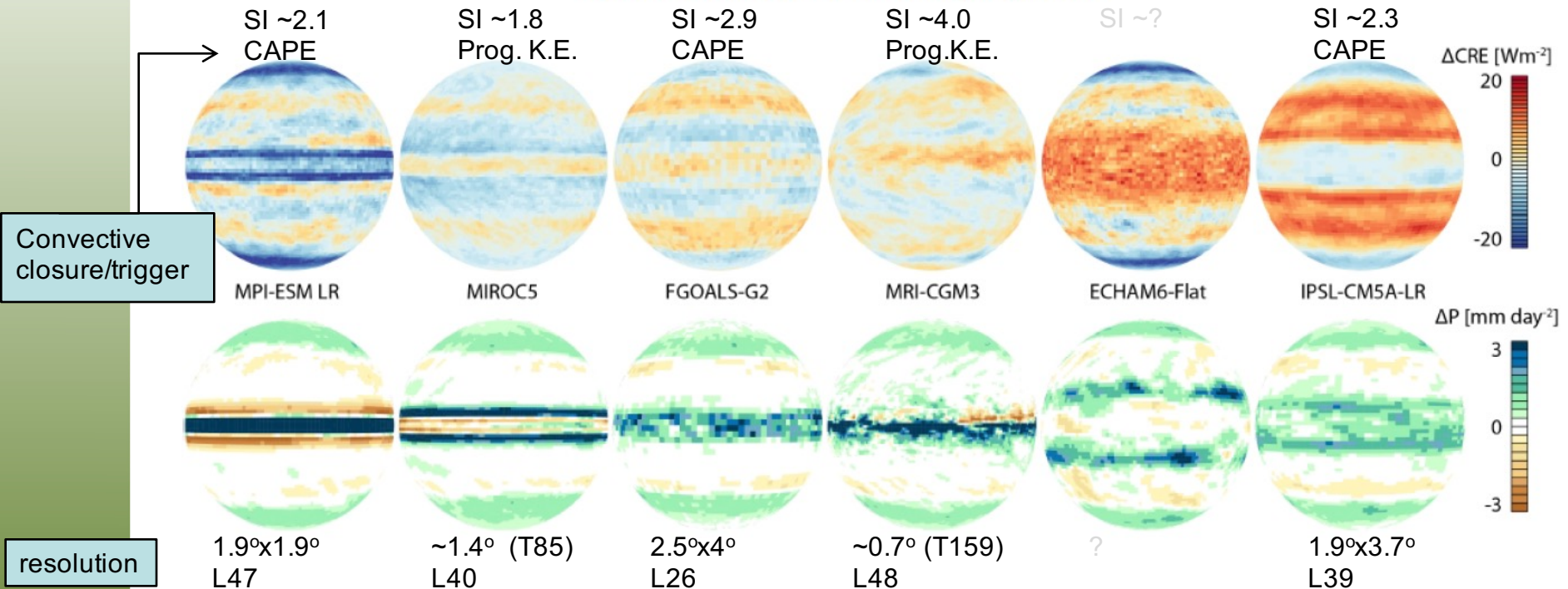


*IPCC AR5, Ch 12*



# Can we get some insight from idealized experiments?

## Response of Cloud Radiative Effects and Precipitation to a uniform +4K in **CMIP5 aqua-planets**



Presentation by Bony and Stevens, 2013.

Page 12 – October 24, 2017

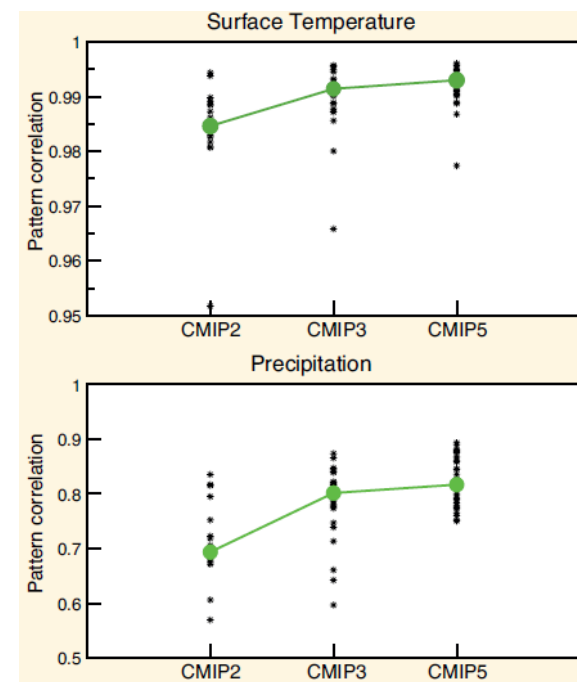


Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada

- So, we have systematic biases that have persisted from one generation of model to the next. They represent *hard* problems and compelling model development challenges.
- But, many of the diagnostics we have to not clearly identify the root cause or even the processes involved, and so model improvement is slow (though visible).
- So we have to explore some other options in parallel.

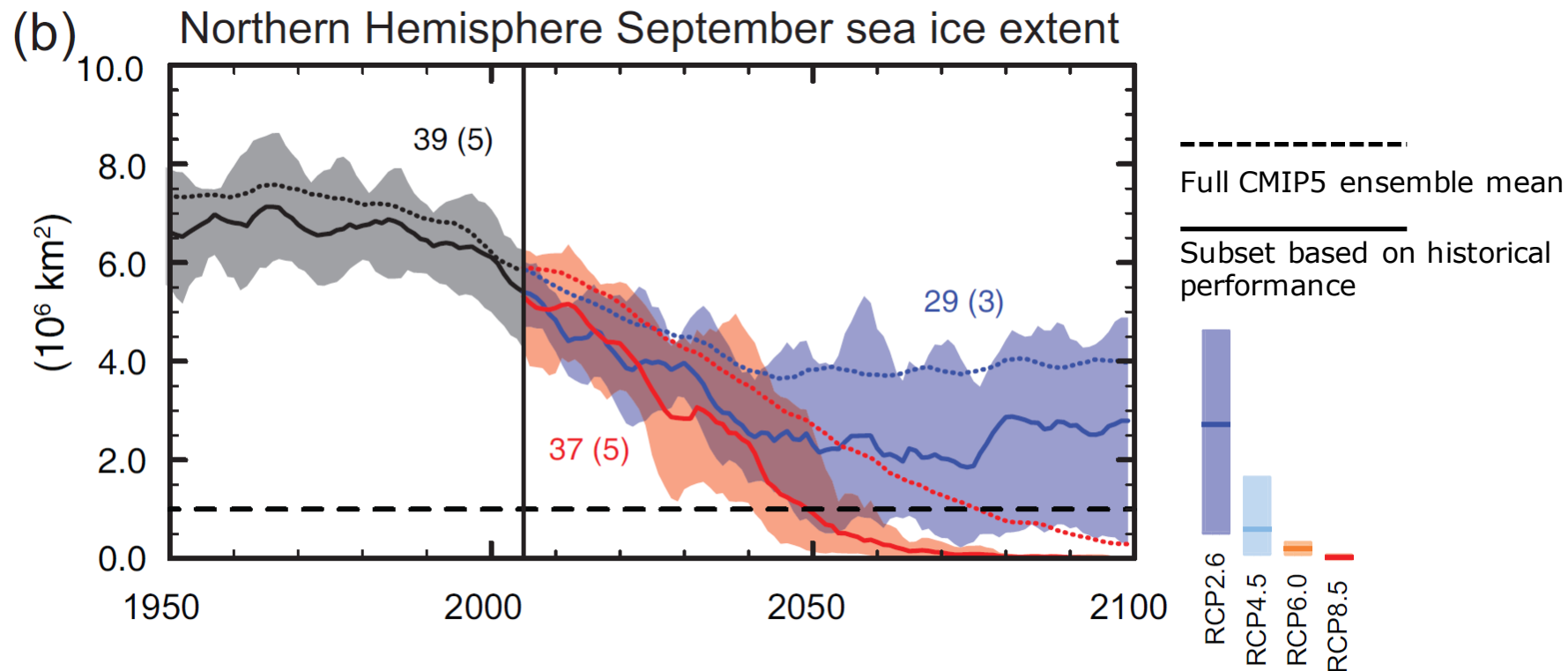


IPCC AR5, Ch. 9





Model sub-setting was used in the IPCC AR5 for Arctic sea-ice extent, based on a paper by Massonnet et al. (2012).



IPCC AR5, SPM (based on Ch.12)

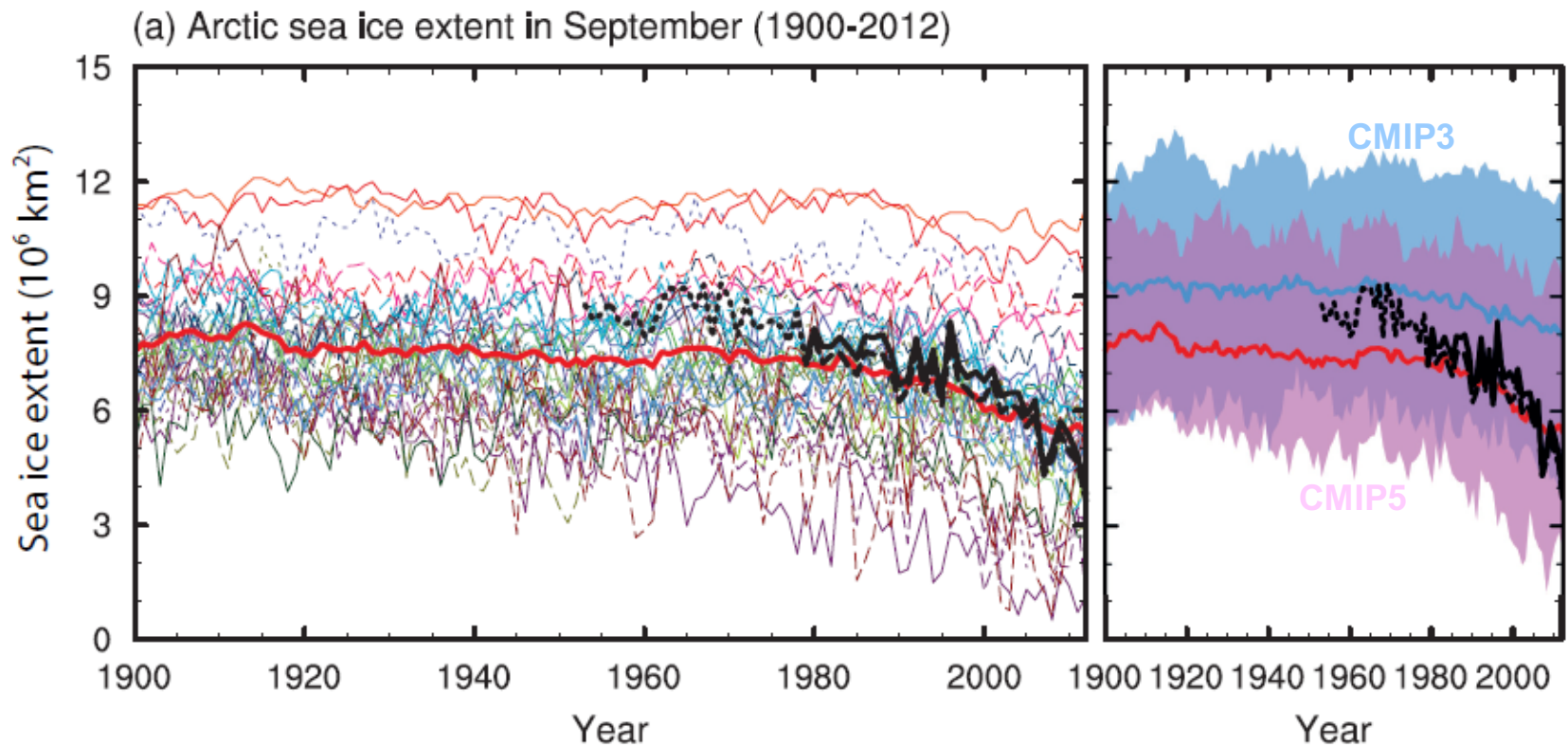
Page 14 – October 24, 2017



Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada



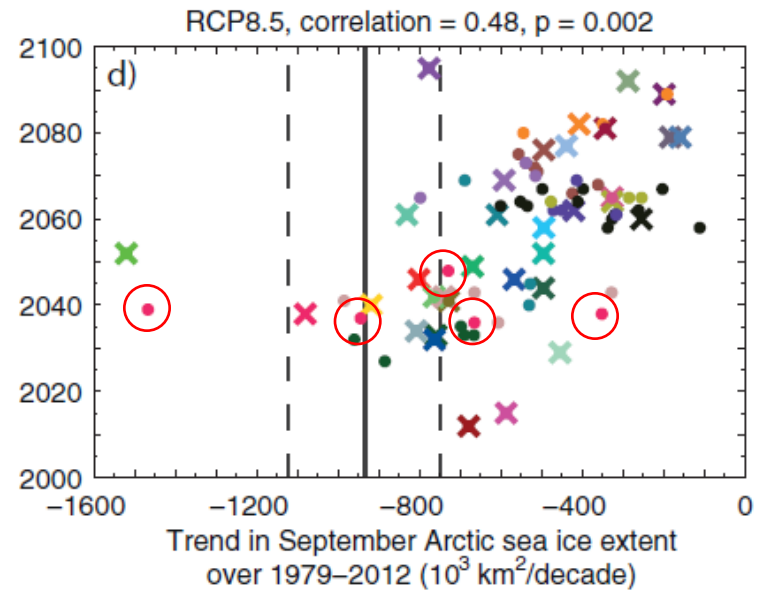
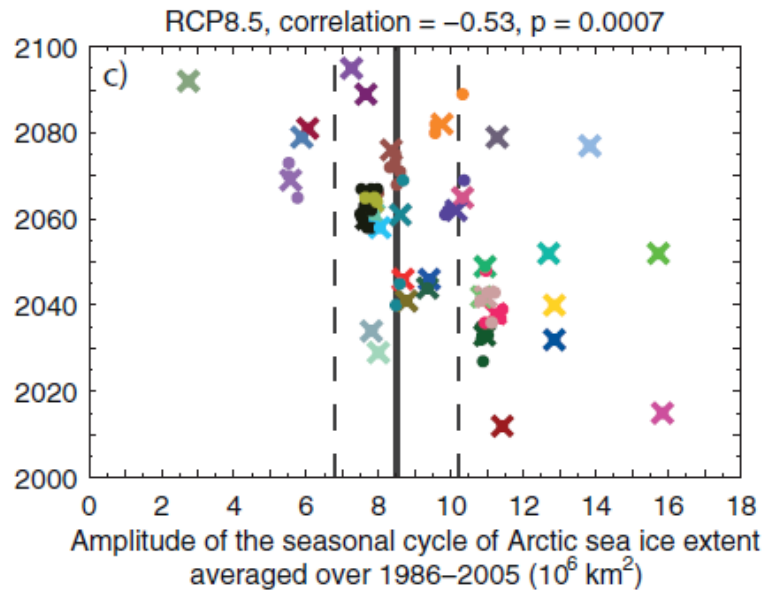
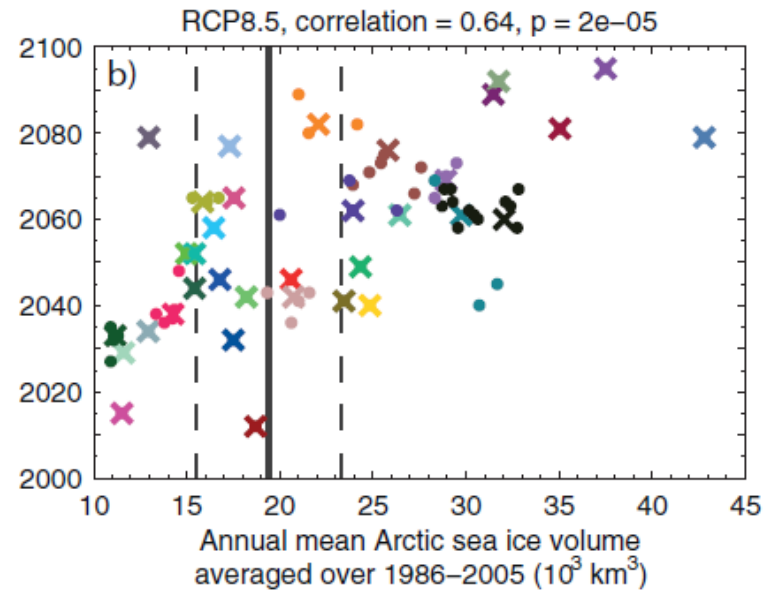
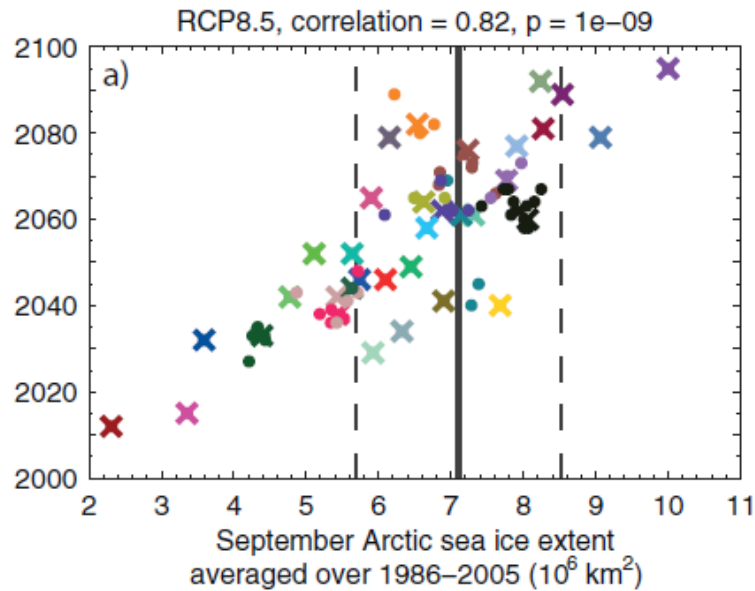
*IPCC AR5, Ch. 9*



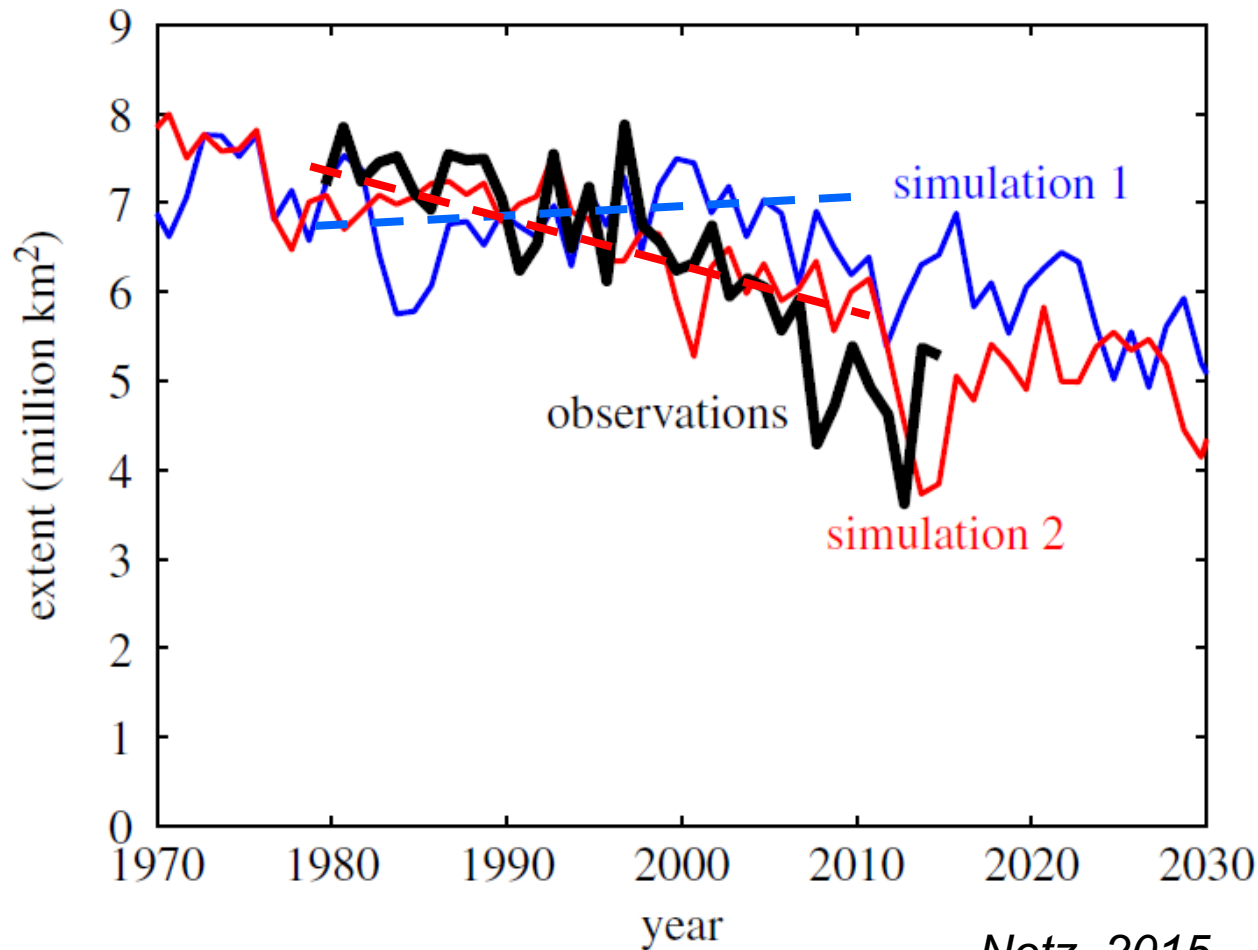


- 
- Systematic bias in *mean* and *trend* improved from CMIP3 to CMIP5.
  - But large spread remains.
  - And that directly impacts our ability to use model output to estimate time at which Arctic will become essentially 'ice free' (an important, impact and policy relevant issue).





An example of internal variability can completely confound a performance metric.



*Notz, 2015*

Page 18 – October 24, 2017

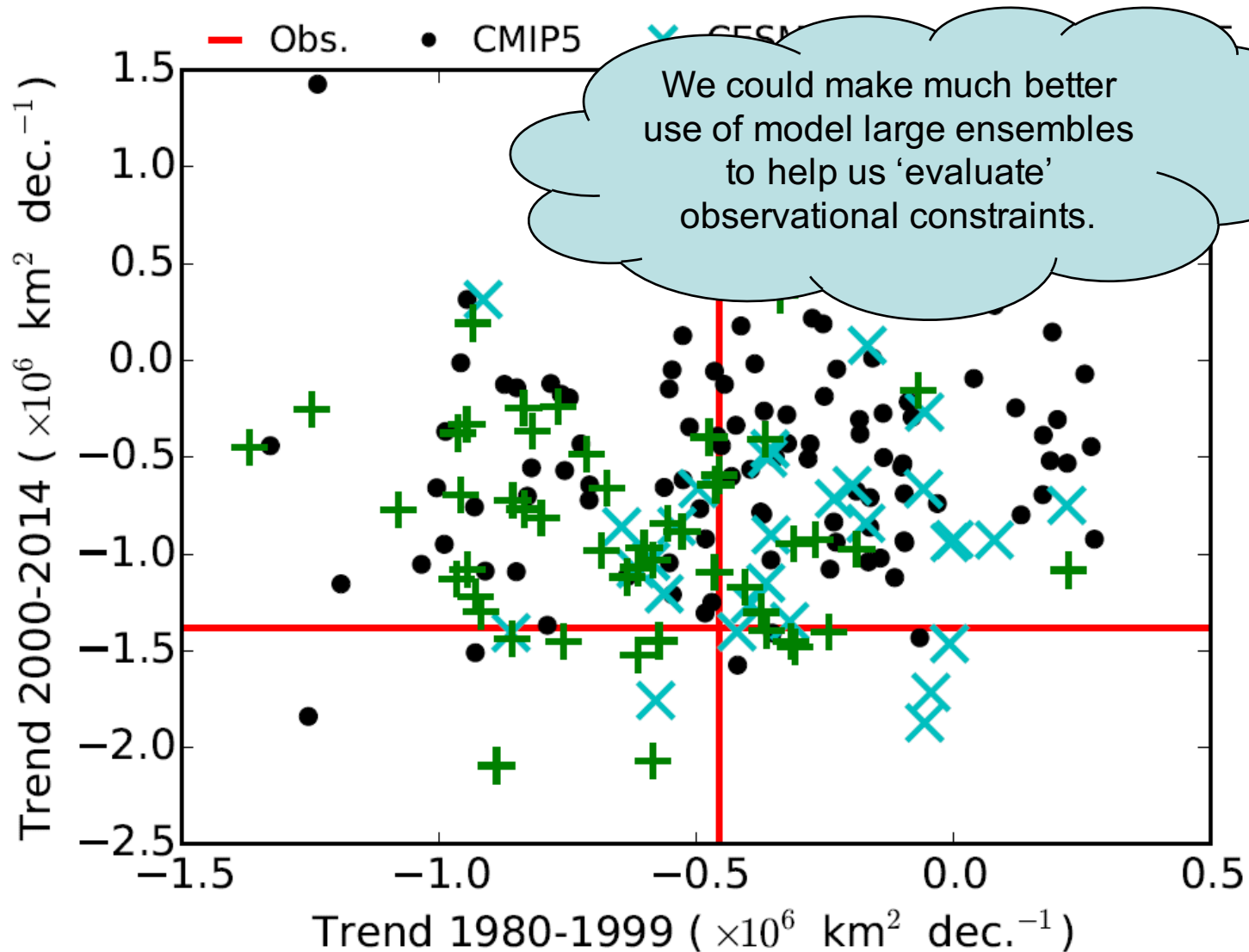


Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada

Large ensembles make this very clear ...



Page 19 – October 24, 2013 *Figure prepared by Neil Swart, CCCma*



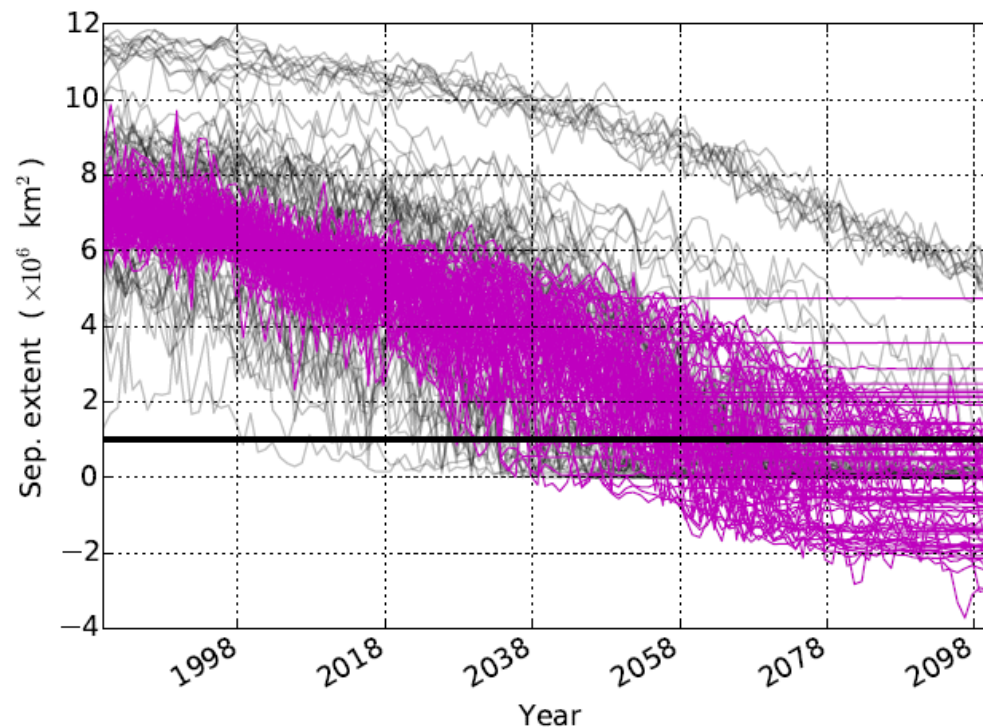
Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada

# Alternatives to selection/weighting?

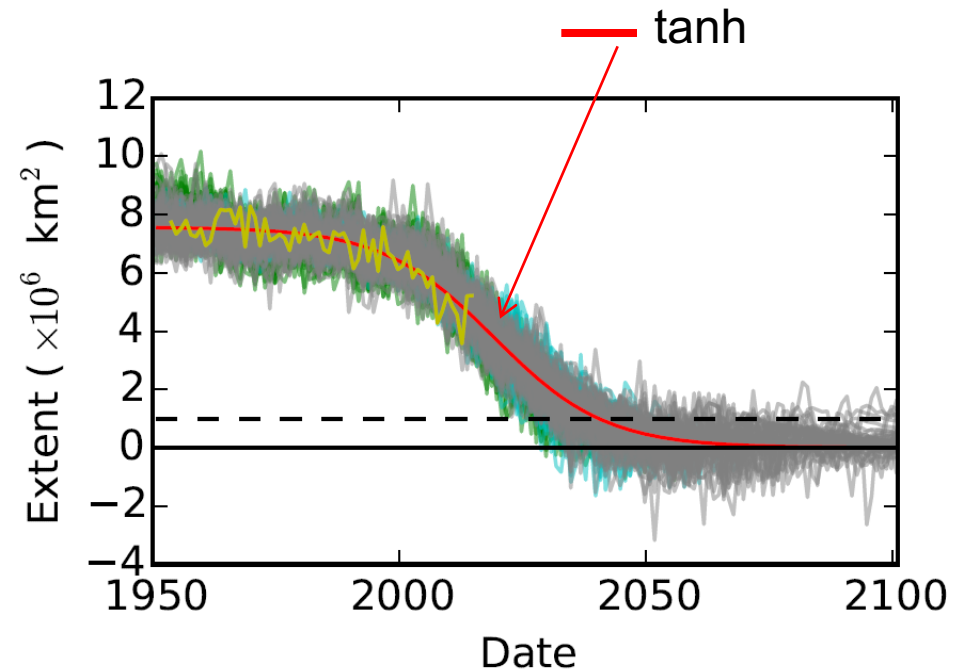
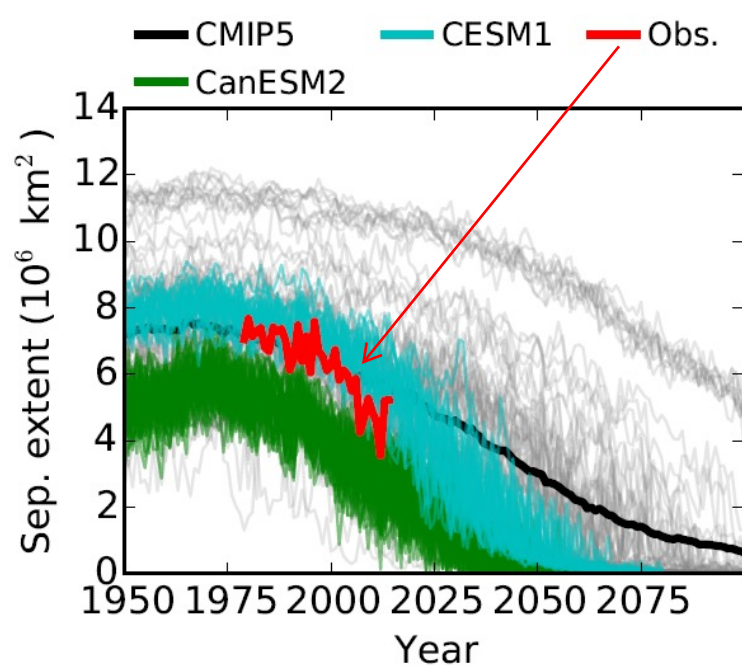
Bias correction (subtracting climatological error) can be useful in some case, but not all ...



*Flato, Swart and Fyfe, in prep.*



Another possibility would be to try and extract a signal from model output, 'calibrated' by observations ...



*Flato, Swart and Fyfe, in prep.*

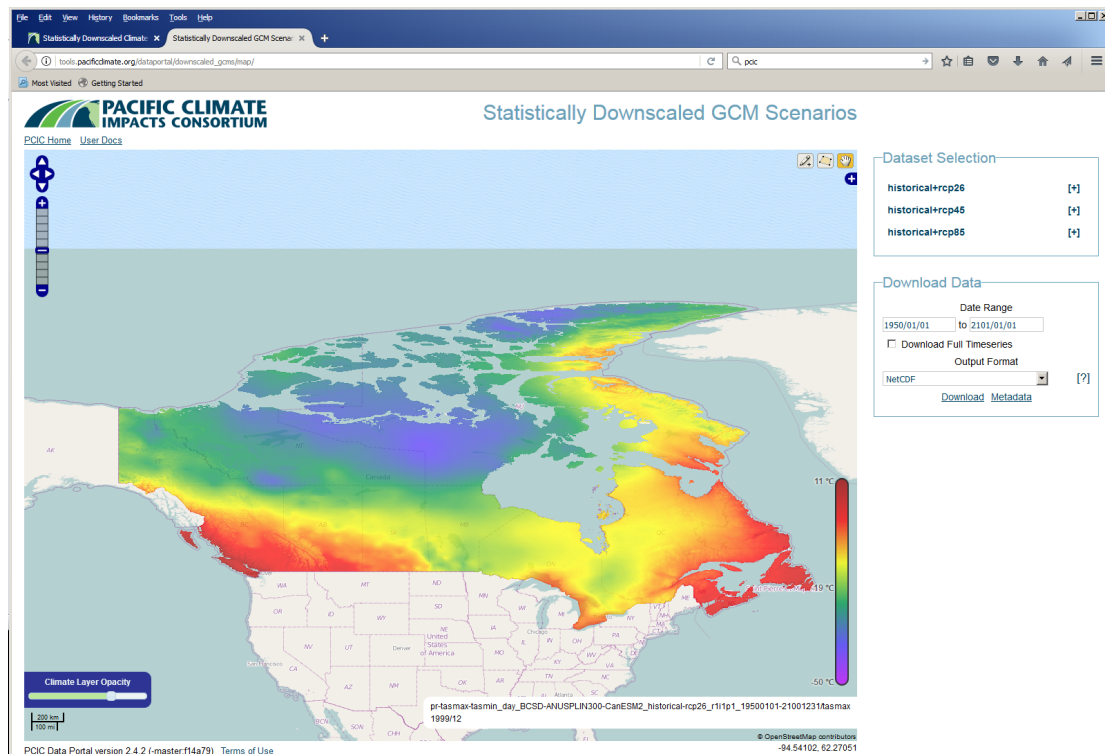


## Statistical post-processing techniques ...

Useful for estimating changes in extreme events as shape of PDF is corrected.

e.g.:

- Bias-Correction Spatial Disaggregation (BCSD)
- Bias Correction/Constructed Analogues with Quantile mapping reordering (BCCAQ)



Page 22 – October 24, 2017



Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Canada



# Concluding thoughts ...

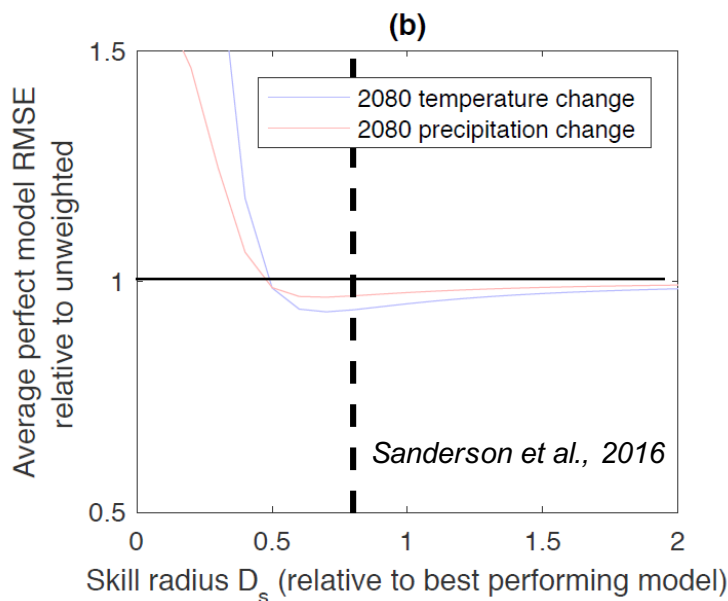
---

- Many traditional performance metrics (e.g. climatological bias), while revealing systematic errors, may not be particularly valuable in assessing model fitness-for-purpose (estimating TCR, future projection, b.c. for downscaling, ...).
- More systematic evaluation will undoubtedly help, but we need to focus on things that matter (how do we identify those things?).
- Systematic evaluation without systematic (and comprehensive) model documentation will limit its value.
- The lack of progress on many longstanding biases suggests the need for more work on value-added post-processing (bias correction, 'signal extraction', other statistical approaches ...).
- Model weighting/sub-setting may be helpful, but internal variability needs to be more carefully considered.



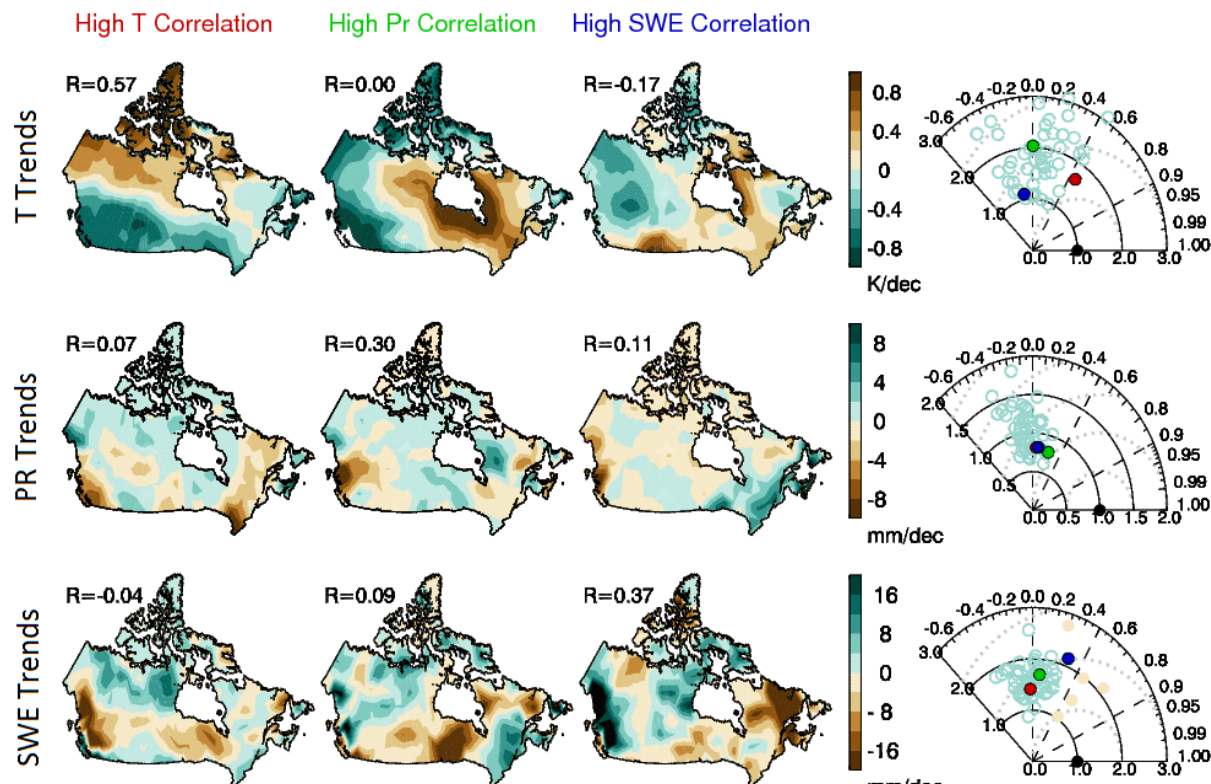
# Concluding thoughts, cont'd ...

- We should take more advantage of large ensembles and 'perfect model' studies to ask about the connection between historical performance and projection skill.
  - A recent example is that of Sanderson et al. (GMD, 2016)



# Concluding thoughts, cont'd ...

- But issues remain when we need consistent multi-variate information ...



1981-2005 trends

*L. Mudryk (pers. Comm.)  
Based on Kushner et al., in press.*

---

*Thank you*

Page 26 – October 24, 2017



Environment and  
Climate Change Canada

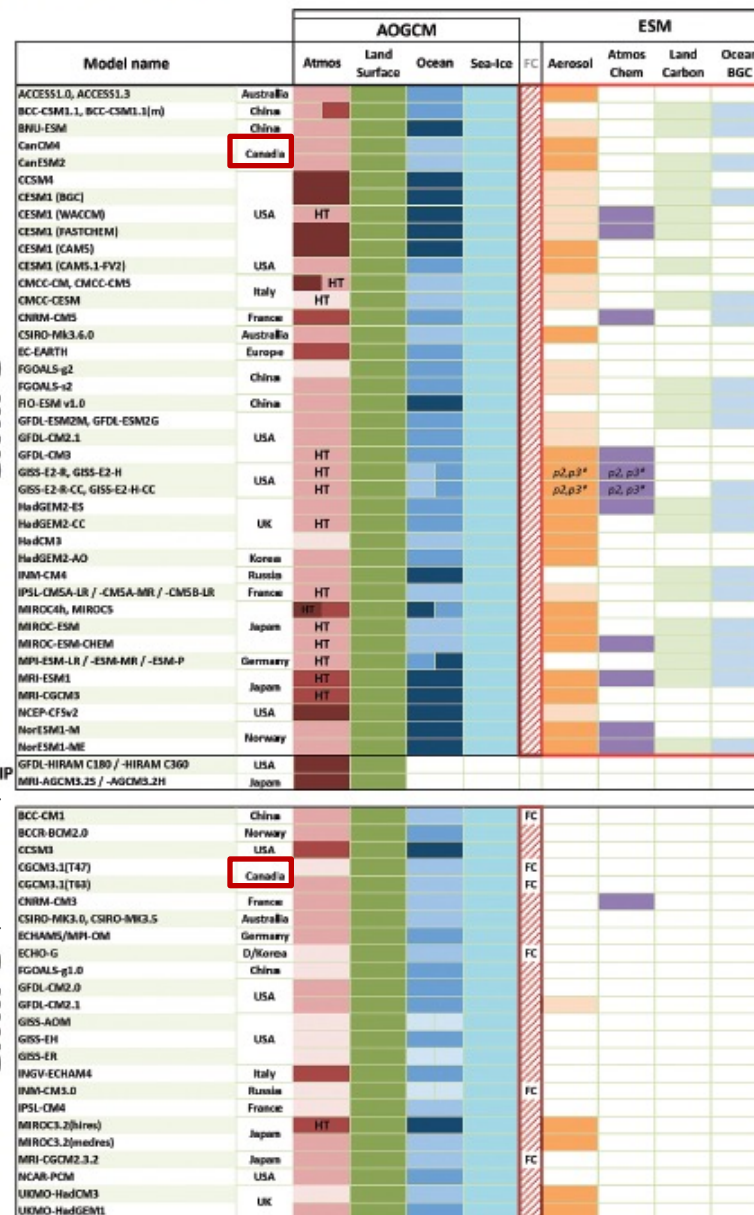
Environnement et  
Changement climatique Canada

Canada 

More models; more comprehensive ...

CMIP5 (39 models)

CMIP3 (24 models)



Increasing resolution Atmosphere / Ocean  
(total number of horizontal grid points)

8000 30000 52000 12000 50000 110000



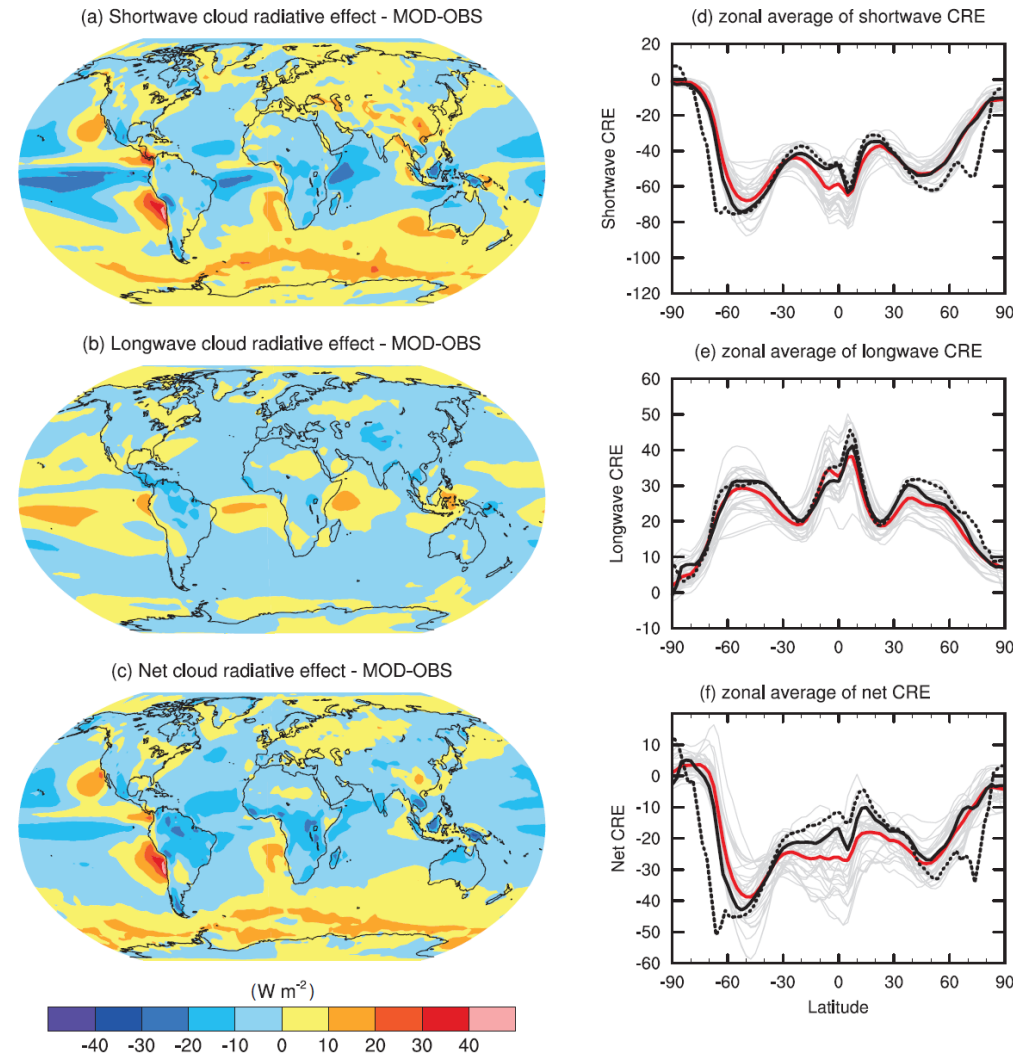
Environment and  
Climate Change Canada

Environnement et  
Changement climatique Canada

Source: Table 9.1, IPCC AR5.

Canada

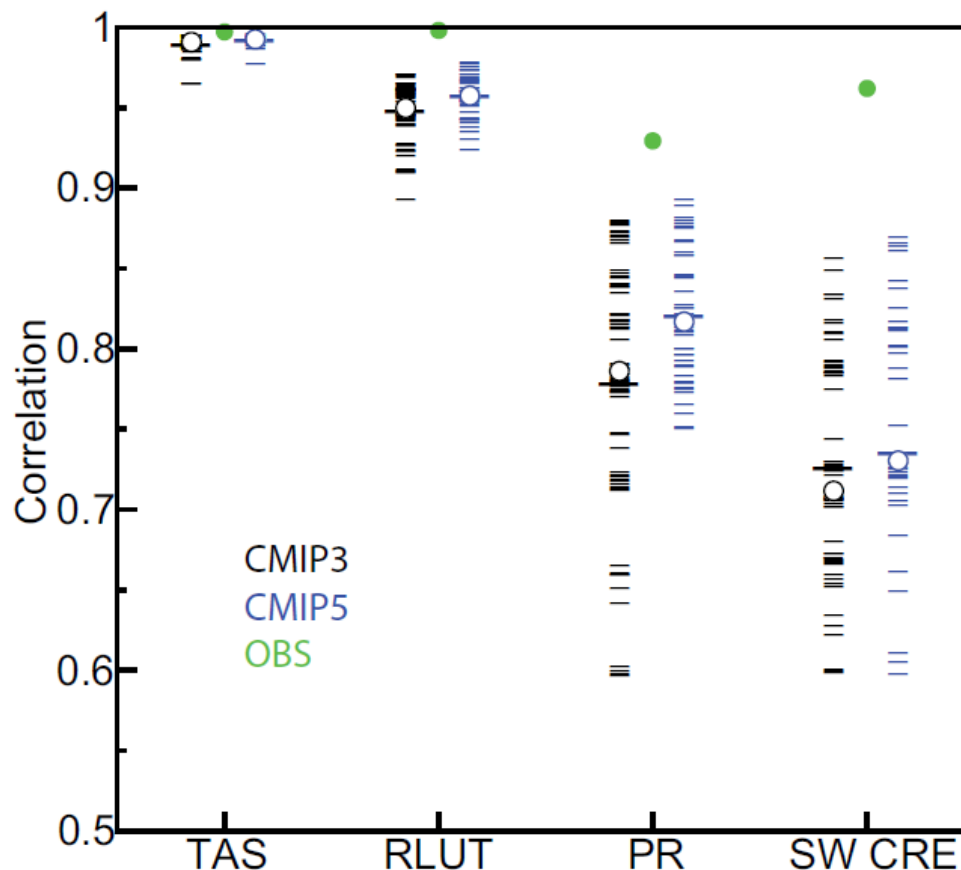
New capabilities, such as the widespread implementation of 'satellite simulators', were exploited.



**Figure 9.5** | Annual-mean cloud radiative effects of the CMIP5 models compared against the Clouds and the Earth's Radiant Energy System Energy Balanced and Filled 2.6 (CERES EBAF 2.6) data set (in W m<sup>-2</sup>; top row: shortwave effect; middle row: longwave effect; bottom row: net effect). On the left are the global distributions of the multi-model-mean biases, and on the right are the zonal averages of the cloud radiative effects from observations (solid black: CERES EBAF 2.6; dashed black: CERES ES-4), individual models (thin grey lines), and the multi-model mean (thick red line). Model results are for the period 1985–2005, while the available CERES data are for 2001–2011. For a definition and maps of cloud radiative effect, see Section 7.2.1.2 and Figure 7.7.



We also made many comparisons of recent CMIP5 results to earlier CMIP3 results (used in the Fourth Assessment).



**Figure 9.6** | Centred pattern correlations between models and observations for the annual mean climatology over the period 1980–1999. Results are shown for individual CMIP3 (black) and CMIP5 (blue) models as thin dashes, along with the corresponding ensemble average (thick dash) and median (open circle). The four variables shown are surface air temperature (TAS), top of the atmosphere (TOA) outgoing longwave radiation (RLUT), precipitation (PR) and TOA shortwave cloud radiative effect (SW CRE). The observations used for each variable are the default products and climatological periods identified in Table 9.3. The correlations between the default and alternate (Table 9.3) observations are also shown (solid green circles). To ensure a fair comparison across a range of model resolutions, the pattern correlations are computed at a resolution of 4° in longitude and 5° in latitude. Only one realization is used from each model from the CMIP3 20C3M and CMIP5 historical simulations.

Page 29 October 27, 2017

