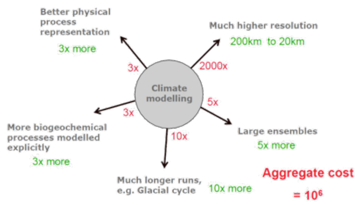
Scientific Challenges in Climate Modelling

Jochem Marotzke

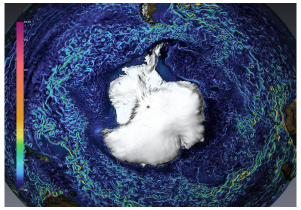
Max-Planck-Institut für Meteorologie
KlimaCampus, Hamburg

Outline

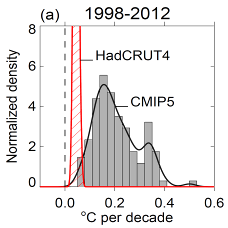
1. HPC requirements by climate modelling



2. Example of high-resolution climate modelling: STORM

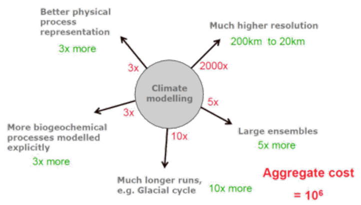


3. Examples of large-ensemble climate modelling

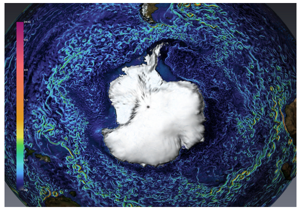


Outline

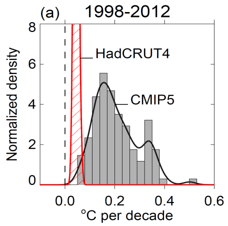
1. HPC requirements by climate science drivers



2. Example of high-resolution climate modelling: STORM



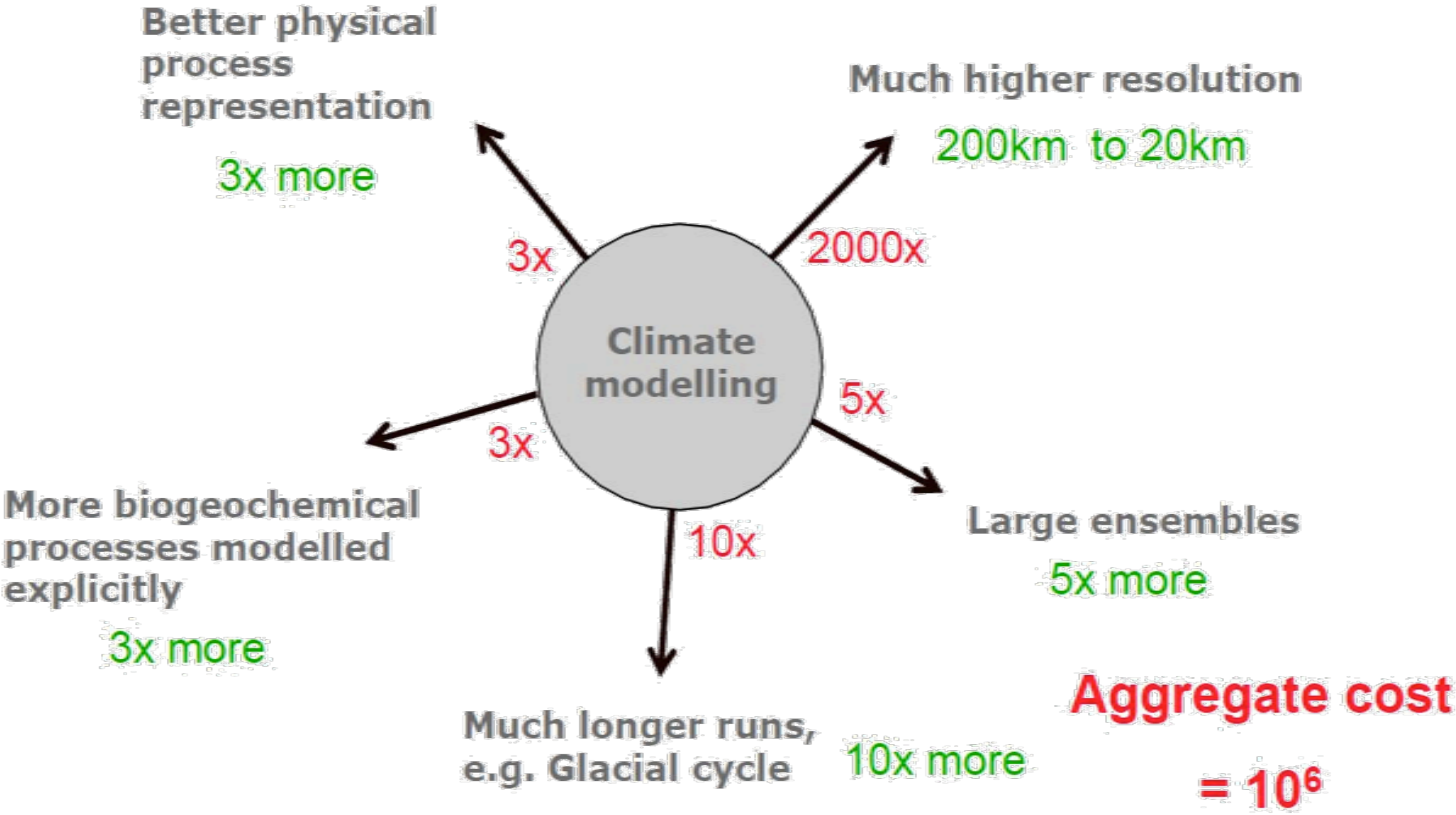
3. Examples of large-ensemble climate modelling



Enhanced HPC requirements by climate modelling

- Much larger ensemble size (repeat simulation with nearly identical initial and boundary conditions)
  - Uncertainty
  - Predictability
  - Data assimilation (model-data fusion)
- Higher resolution
  - Ocean: 10x higher resolution per horizontal dimension (to resolve processes with highest kinetic energy)
  - Atmosphere: 10x higher resolution per horizontal dimension (e.g., to represent hurricanes; to represent circulation regimes that are important for the response of clouds to climate change)
- Higher complexity (more model components)
  - Biogeochemical cycles & feedbacks (e.g., how much anthropogenic CO_2 will remain in the atmosphere?)
- Longer runs (ice age cycles)

What science trend will be most prominent? Analysis of options: what is desired?



Max-Planck-Institut für Meteorologie

(Brian Hoskins & Walter Zwieflhofer)

What science trend will be most prominent?
Analysis of options: practicalities (1)

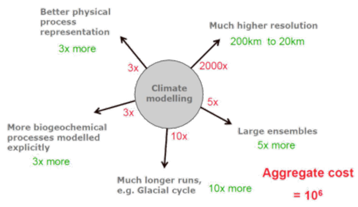
- Growth in all dimensions not possible to the extent desired
- Low-hanging fruits? In order of increasing effort (vis-à-vis current trends in HPC provision):
  - Ensembles: weak scalability & computationally trivially parallel, but post-processing requires HPC provision for data organization (entire output must be easily accessible; currently: must be held in one location)
  - Resolution: weak scalability, scientifically very interesting, but difficult and computationally most expensive
  - Complexity (e.g., chemical, biological)/additional physical processes: scientifically most interesting to many, unpredictable scalability
  - Longer runs: totally against technology trends, does not scale

What science trend will be most prominent?
Analysis of options: practicalities (2)

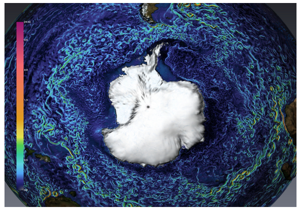
- Growth in all dimensions not possible to the extent desired:
- Data deluge (currently >10^2 petabytes per year across community)
  - Cost (storage does not get cheaper as fast as does computing)
  - Organization
- Required: massive reduction of data production per flop
  - Ever more sophisticated "online" diagnostics (during runtime)
- Necessary: more re-computing instead of storage, but not trivial
  - We simulate a chaotic system; simulation on a different HPC system might diverge; need to separate systematic from quasi-random differences between original and repeated simulation
  - Archived output is heavily used outside modelling centres; this trend is going to increase (e.g., climate impact research; climate services)

Outline

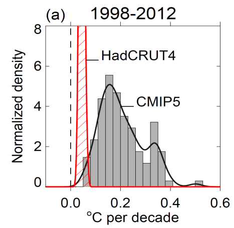
1. HPC requirements by climate science drivers



2. Example of high-resolution climate modelling: STORM

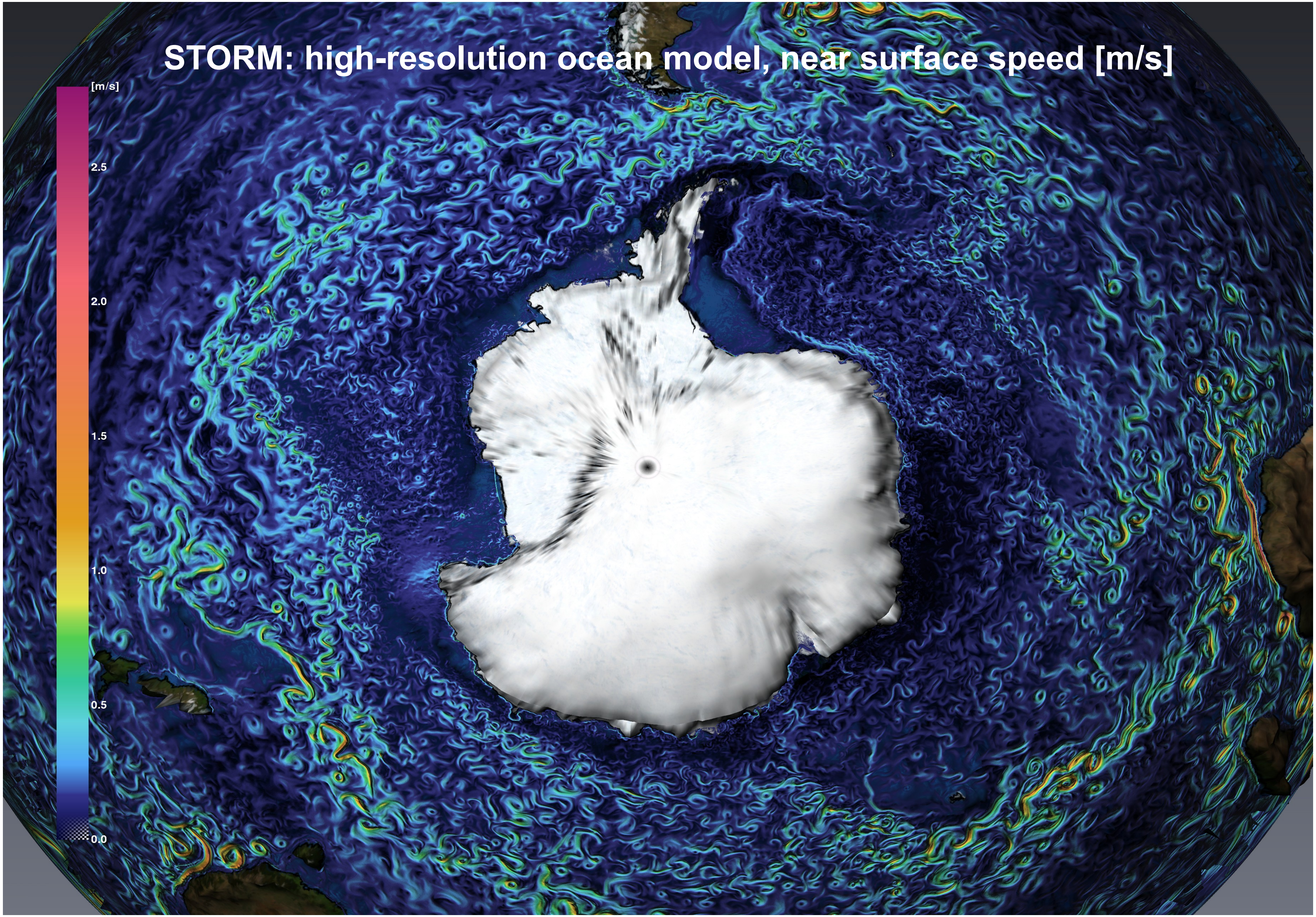


3. Examples of large-ensemble climate modelling



Fundamental reason for moving toward higher resolution

- A larger portion of the simulated system is moved from unresolved to resolved scales, from having to make plausible assumptions (parameterisations) to knowing the underlying equations


STORM: high-resolution ocean model, near surface speed [m/s]
[m/s]
2.5
2.0
1.5
1.0
0.5
0.0

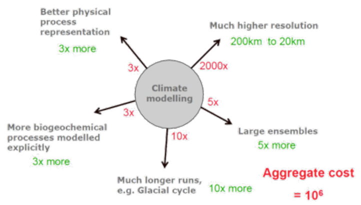
STORM: global high-resolution ocean & atmosphere simulation

Notable result:

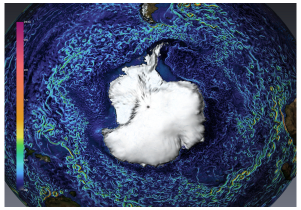
- Possible because high spatial resolution permits the representation of the most energetic processes (boundary currents; eddies)
  - Generation of eddies is expected to be a crucial process in converting potential to kinetic energy in the ocean (as it is in atmosphere)
- STORM: first quantitative analysis of the mechanical energy cycle in the global ocean (how is work performed on the ocean, and how is mechanical energy distributed and dissipated within the ocean)
  - Atmosphere (known): heat engine (internal energy → motion)
  - Ocean (new): wind mill (work → dissipation) und refrigerator (work → temperature differences); eddy generation crucial locally but less important for global circulation

Outline

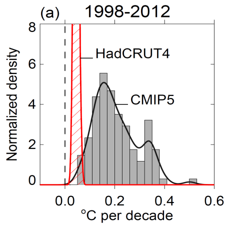
1. HPC requirements by climate science drivers



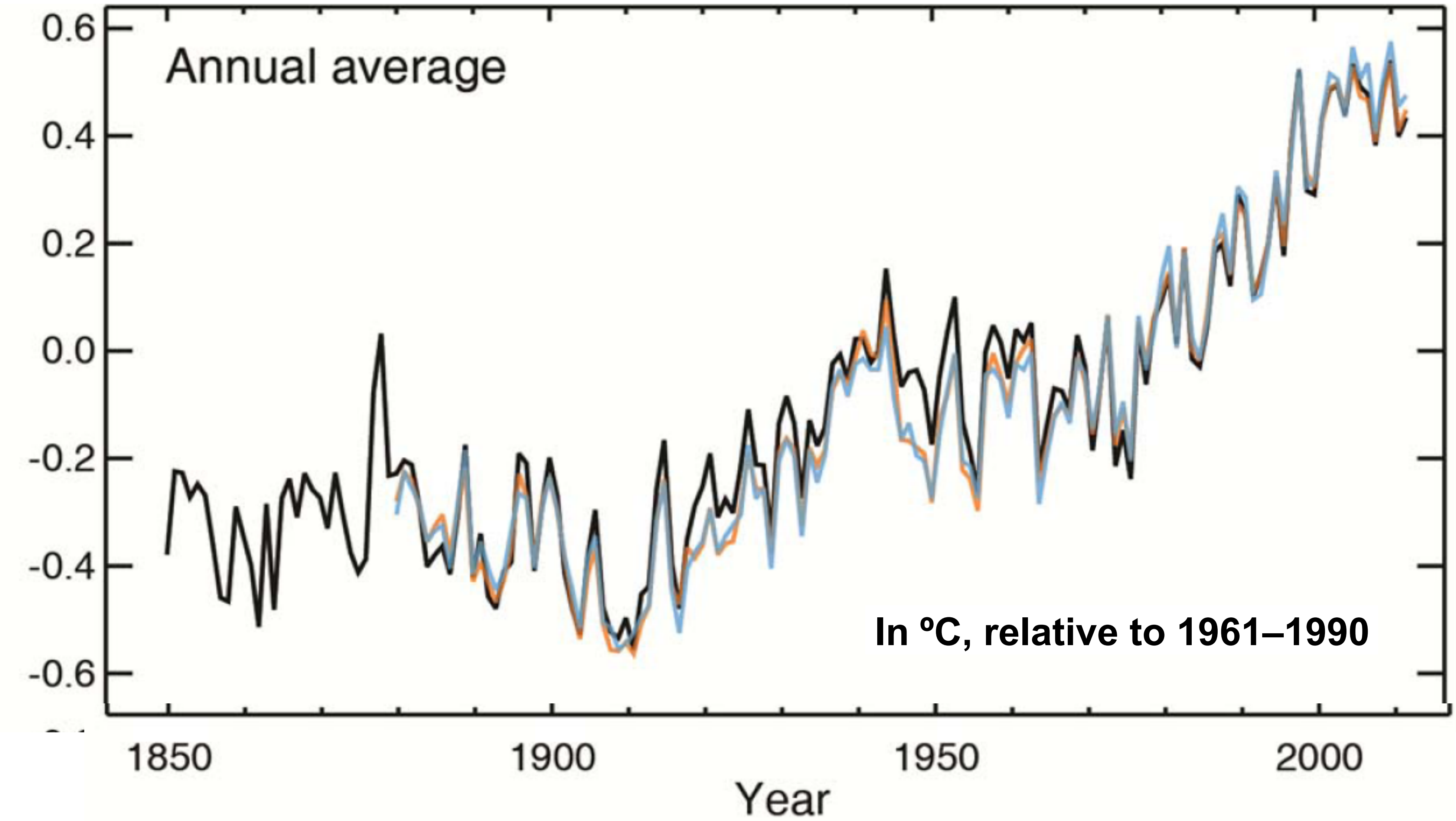
2. Example of high-resolution climate modelling: STORM



3. Examples of large-ensemble climate modelling

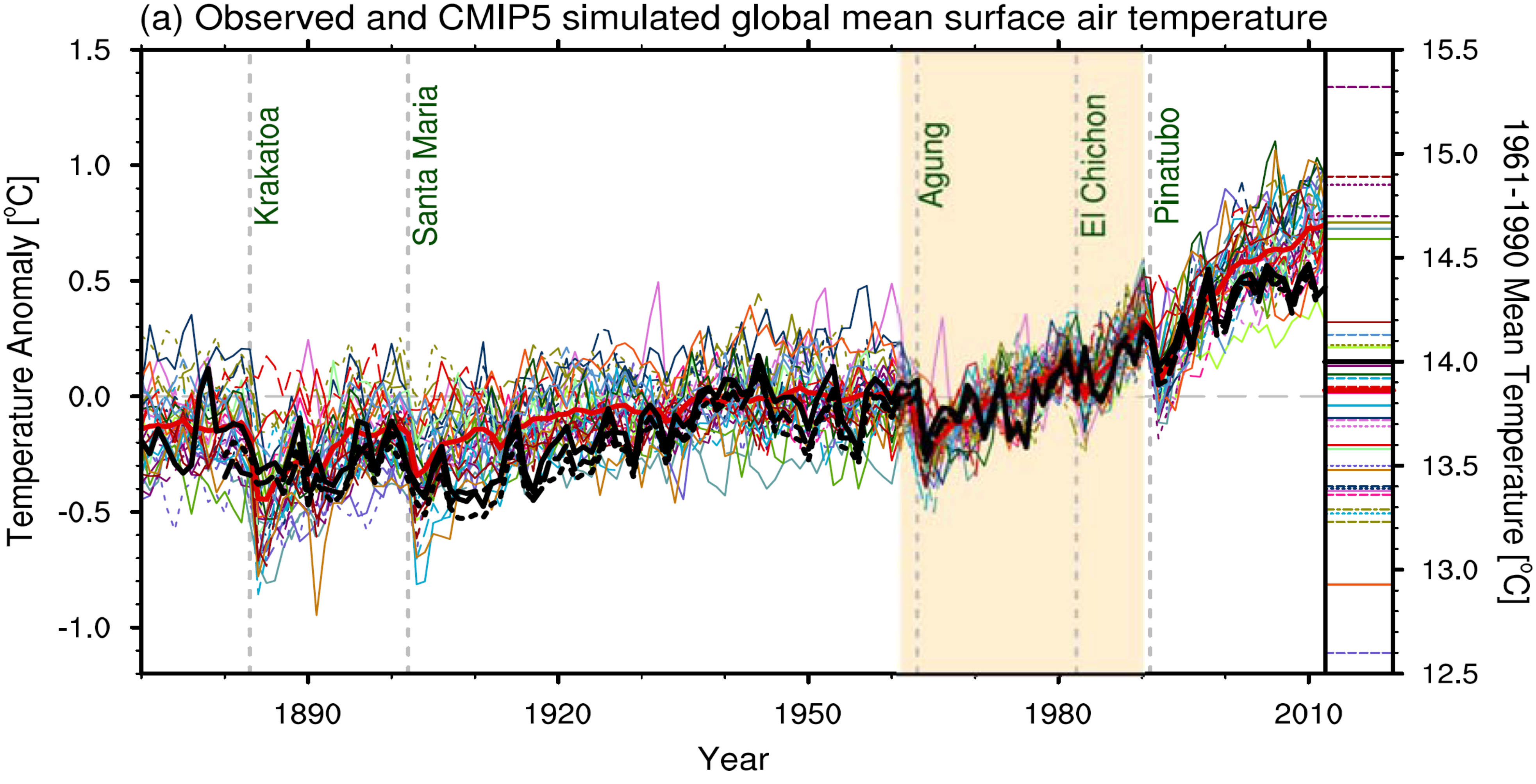


Observed global-mean surface temperature (GMST) shows long-term warming trend, overlaid by chaotic fluctuations



IPCC WG1, AR5, Figure SPM.1a

CMIP5 models (colours) reproduce observed long-term GMST over the 20th century (black), but not the reduction in trend over the past 15 years

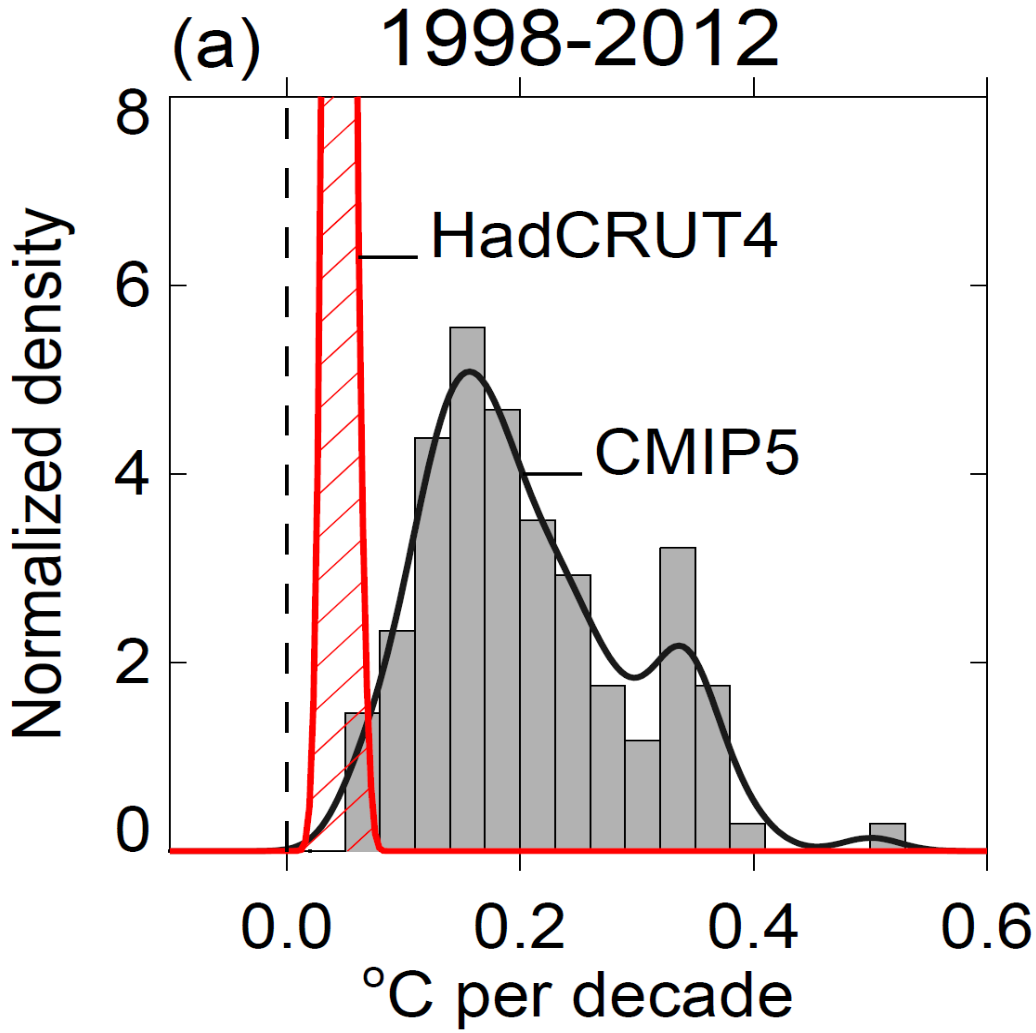


IPCC WG1, AR5, Figure 9.8

How to compare the single observed realisation of a chaotic system against model simulations? We need to know whether a difference is due purely to quasi-random effects

- Use large ensemble of simulations
- Hitherto unanswered question: if ensemble is constructed using different models, is difference among ensemble members due to differences in model physics or due to quasi-random effects?
- Working hypothesis in IPCC AR5: ensemble spread is measure of quasi-random effects

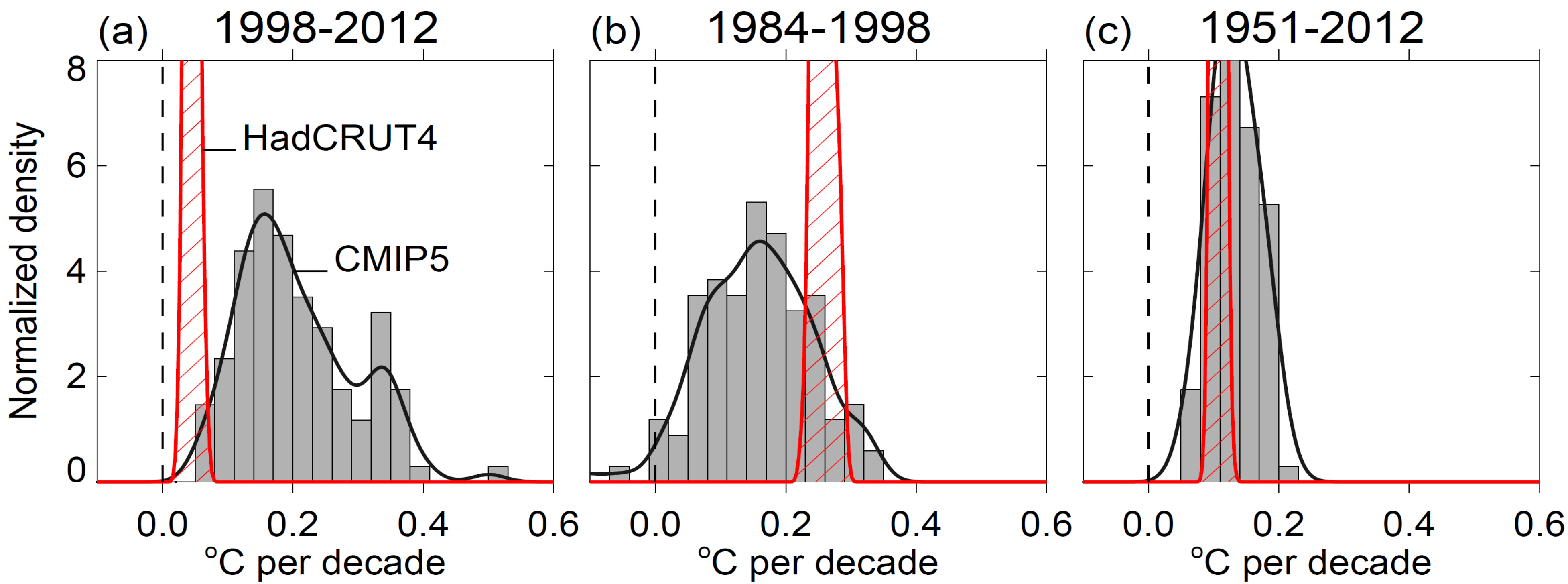
Over 1998–2012, GMST trend is higher than observed in 111 out of the 114 available CMIP5 historical simulations



Length of histogram bar displays how many simulations show a trend in the range defined by the width of the bar

Observations: HadCRUT4

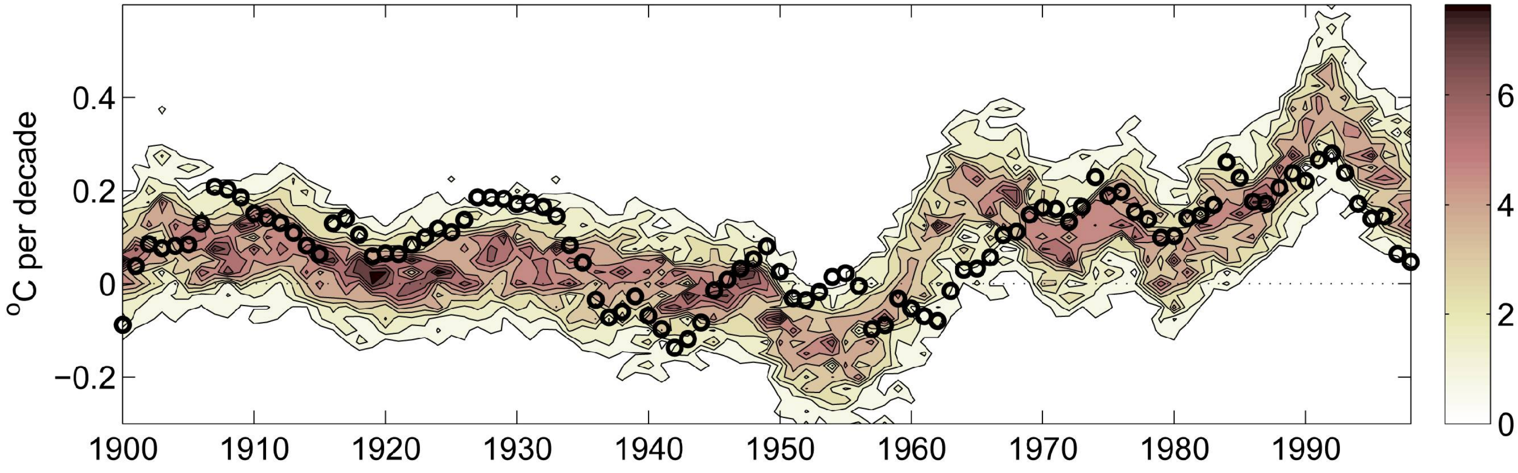
It matters when you look – observations show larger trend than models over 1984–1998. Long-term trend reproduced by models.



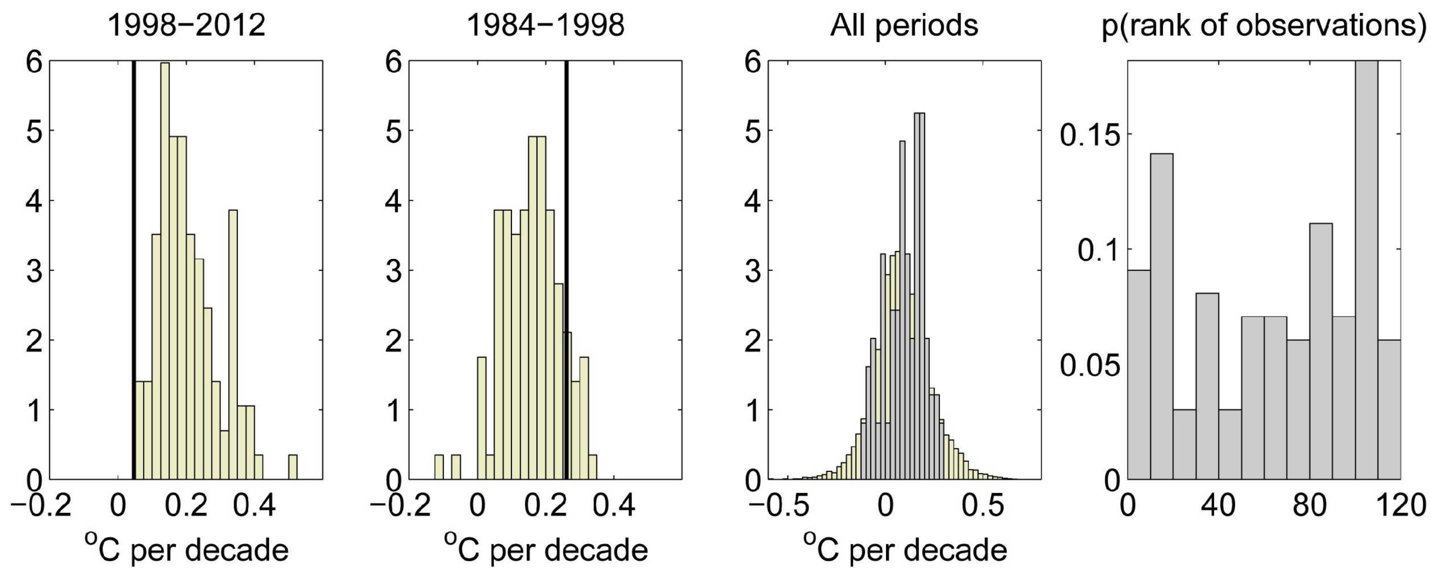
Length of histogram bar displays how many simulations show a trend in the range defined by the width of the bar

IPCC WG1, AR5, Box 9.2, Figure 1

15-year trends in global surface temperature: large ensemble necessary for quantitative comparison of models (colour) against observations (black/grey)



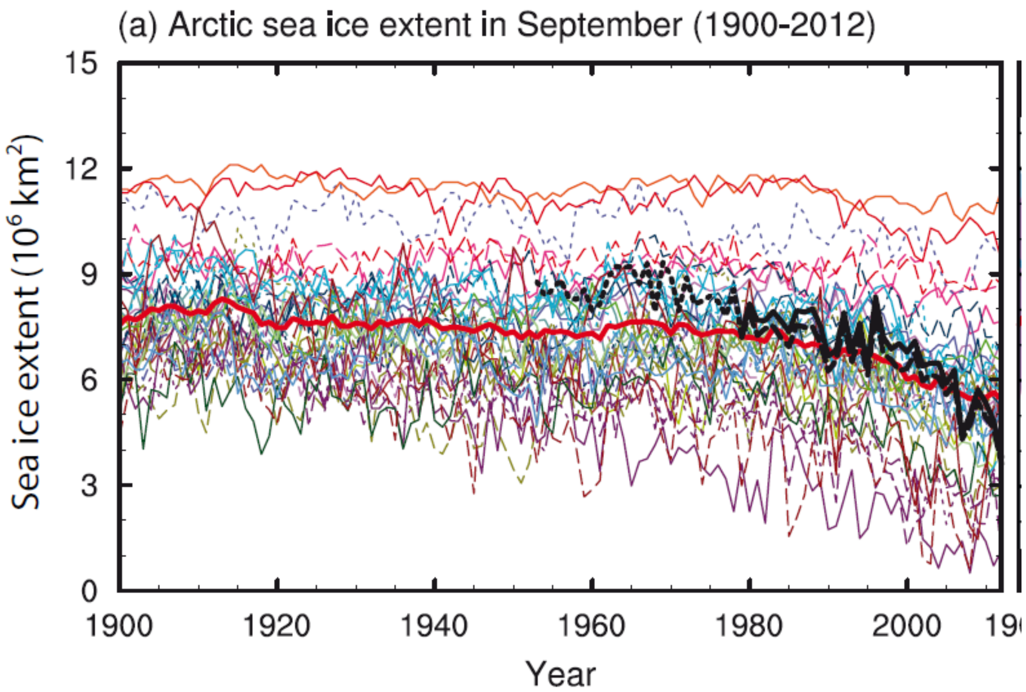
Number of models showing trend of this magnitude

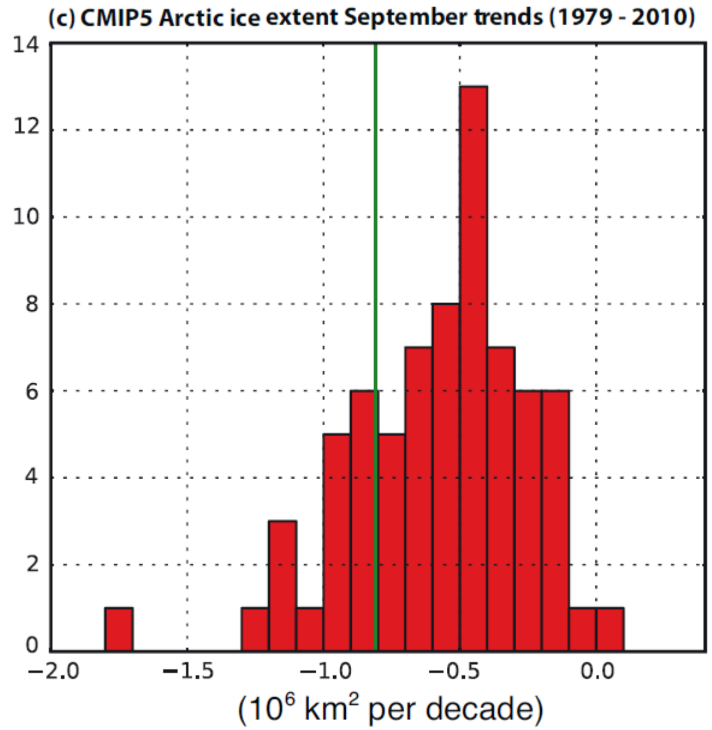


Observed trend could be “anywhere” if juxtaposed to model ensemble

Marotzke and Forster, in preparation

IPCC AR5: Trends in Arctic summer sea ice are consistent between model ensemble (colours) and observations (black)





Ensemble perspective essential to overcoming prejudice that models (red bars) underestimate the observed trend (green line)

Conclusions

- Heavy investment necessary to enable increased spatial resolution in climate models
  - Higher resolution does not automatically guarantee superior simulation
- Larger ensembles provide crucial capability to compare model simulations against observations
  - Simpler to implement on modern HPC systems than higher-resolution models
  - Data requirements are substantial

Thank you for your attention!