National Aeronautics and Space Administration
Goddard Institute for Space Studies

Goddard Space Flight Center
Sciences and Exploration Directorate
Earth Sciences Division

Subtle lessons from the art of model-observation confrontations

Gavin Schmidt, NASA GISS

CLIVAR Mar 2024

Philosophical Footnote #1

All models are wrong

(There are no scientific realists in climate model group foxholes…)

Philosophical Footnote #2

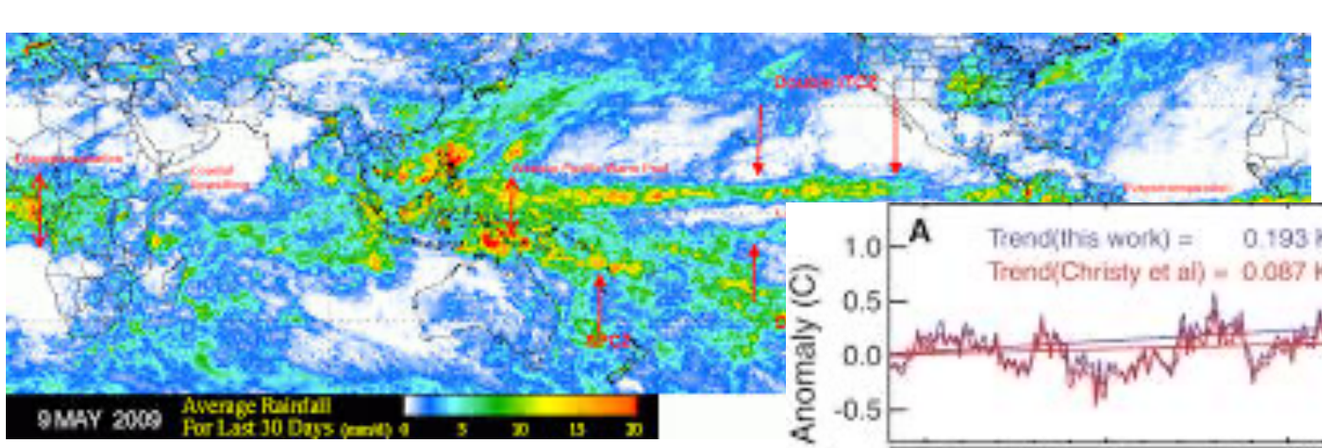
All observations and comparisons are based on models too

(It’s models all the way down!)

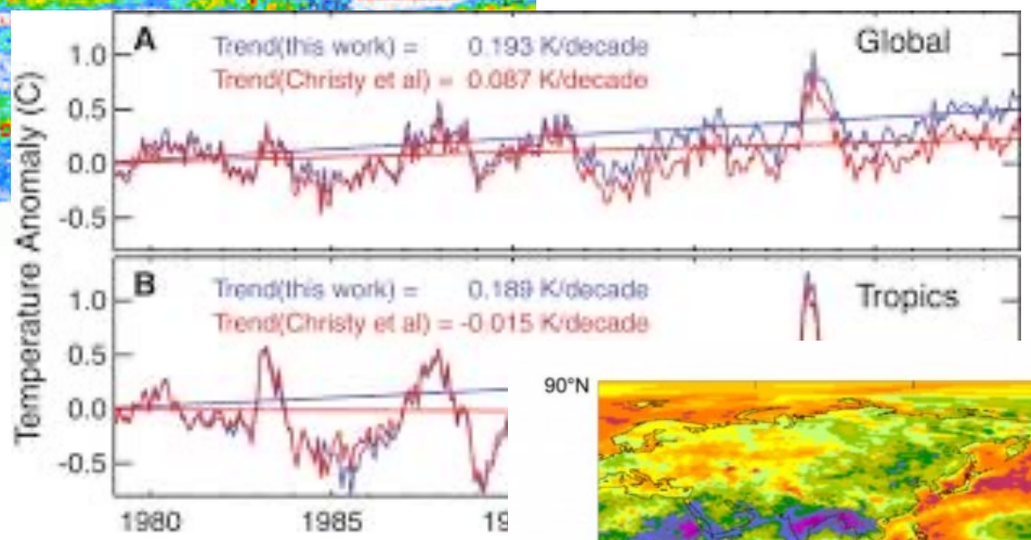
Goddard Institute for Space Studies

Climate science controversies

Most of the controversies related to modern climate change involve discrepancies between global scale climate models and observations (of various sorts).

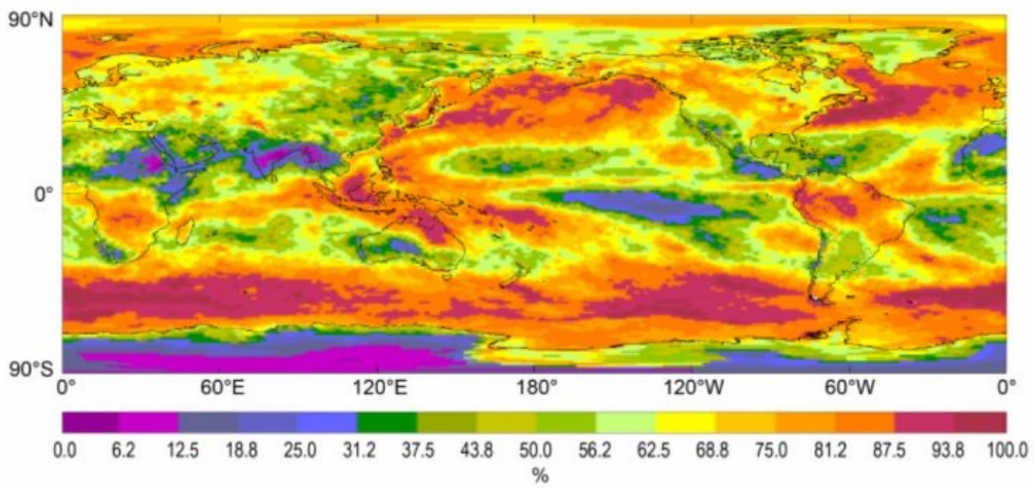


The double-ITCZ



The ‘hiatus’

MSU “cooling”



Mixed cloud phase

Goddard Institute for
Space Studies

There are many reasons for this…

Models are wrong

Observational data are imperfect

- noisy and biased
- Often w/o estimates of structural uncertainty

Comparisons are not appropriate

- Not like-with-like
- Unrecognized point-of-view biases
- Irreducible effects of chaos

Philosophical Footnote #3

This is called Duhem/Quine underdeteriminancy

Any model-data confrontation is a test of *multiple* hypotheses

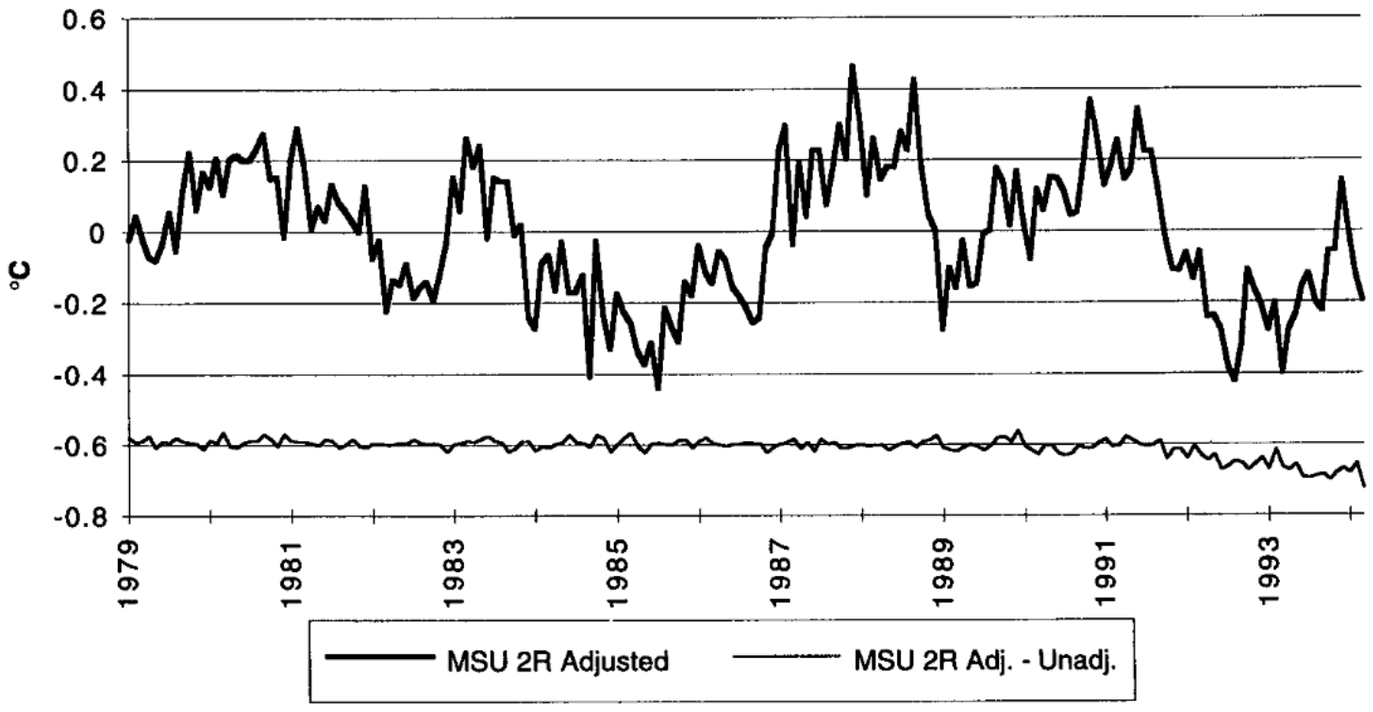
Goddard Institute for
Space Studies

Example 1: MSU 'cooling'

1992: A new method of tracking global atmospheric temperature is developed by Spencer and Christy

NASA press release (1997): "the satellite data are the best quality possible"

The problem? The MSU lower atmosphere data show cooling from 1979, while surface data and models show warming.



Christy et al 1995

Example 1: MSU ‘cooling’ (cont)

Goddard Institute for Space Studies

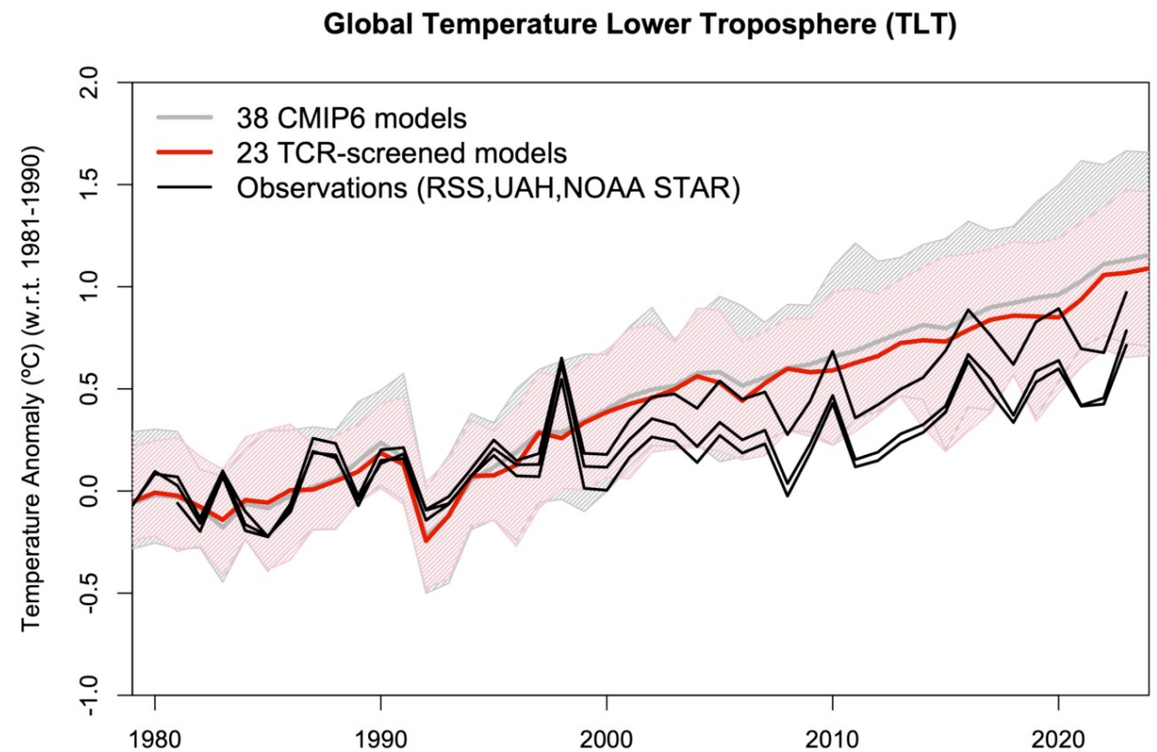
Originators of the satellite data (still) blame the models and the other data sets.

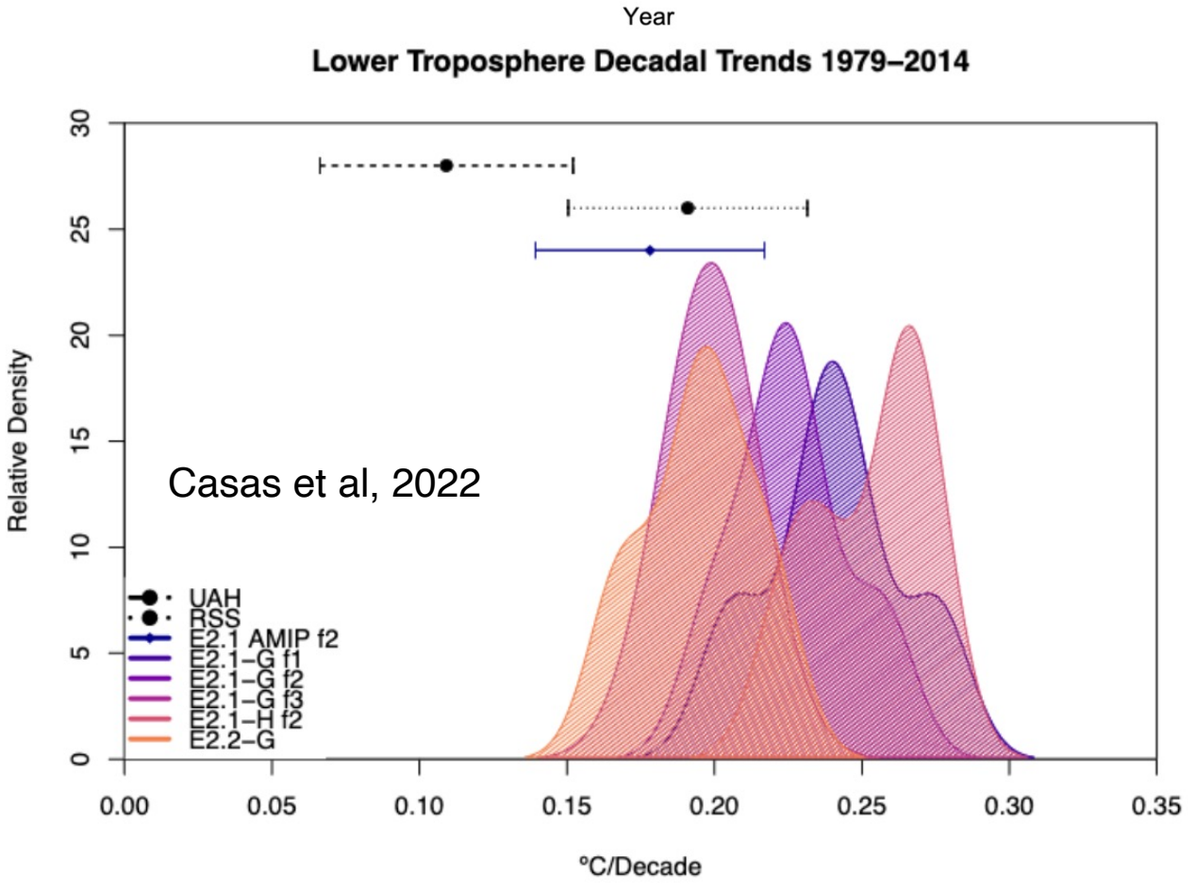
But no fixes/updates to either materially affect the problem

Independent replication of satellite products shows significant structural uncertainty (Mears et al, 2003)

New factors ‘orbital drift’, errors in diurnal correction, etc. lead to reversal of ‘observed’ trends in better concordance with models.

Deeper exploration of structural uncertainty in models and observations effectively eliminate discordance (but UAH is still an outlier)





Goddard Institute for
Space Studies

Example 2: The ‘hiatus’

2006: “Global Warming has stopped in 1998” claim based on short term trends in temperature data. Not taken seriously given expected internal variability. (A ‘*tolerable discrepancy*’)

Philosophical Footnote #4

‘Tolerable discrepanies’ arise inevitably when a relatively efficient theory is confronted with the complexity of the real world.

Only when the discrepancies are not expected and cannot be easily explained away do we get a discordance

Goddard Institute for
Space Studies

Example 2: The ‘hiatus’

2006: “Global Warming has stopped in 1998” claim based on short term trends in temperature data. Not taken seriously given expected internal variability. (A ‘*tolerable discrepancy*’)

2012: In two data sets (HadCRUT3 and NOAA GlobalTemp v3), short-term trends start to approach 2σ anomaly…

2013-2014: ~50 papers exploring why

- Misspecified forcings, biases in observations, insufficient model internal variability, comparison improvement etc.

2015-2016?

Example 2: The 'hiatus' (cont)

Goddard Institute for Space Studies

"All the News That's Fit to Print"

The New York Times

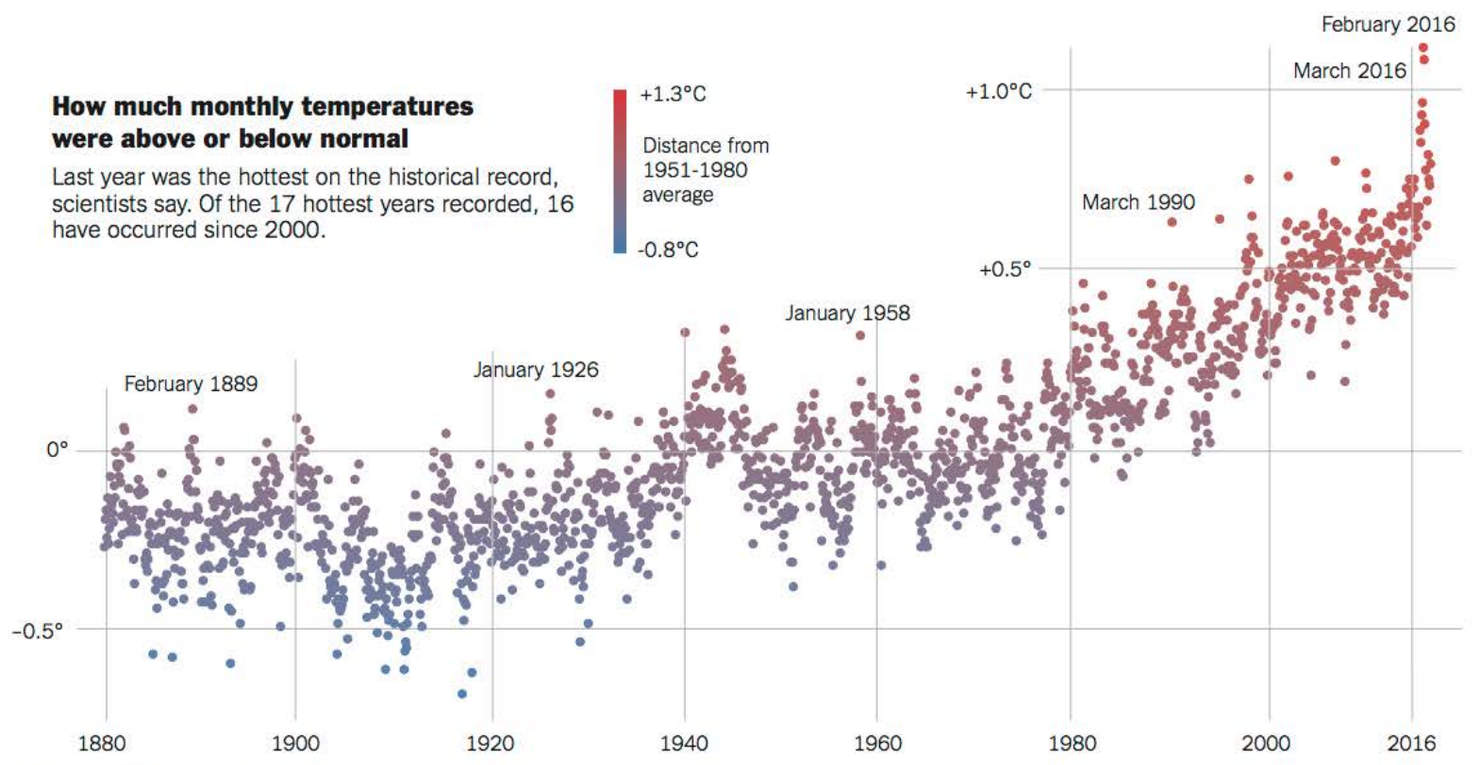
Late Edition

Today, partly sunny, milder in the afternoon, high 50. **Tonight,** partly cloudy, low 37. **Tomorrow,** some sunshine, then clouds, showers late, high 46. Weather map, Page A28.

VOL. CLXVI ... No. 57,482 NEW YORK, THURSDAY, JANUARY 19, 2017 $2.50

How much monthly temperatures were above or below normal

Last year was the hottest on the historical record, scientists say. Of the 17 hottest years recorded, 16 have occurred since 2000.



Source: NASA GISS Surface Temperature Analysis

JUGAL K. PATEL/THE NEW YORK TIMES

FOR THIRD YEAR, THE EARTH IN 2016 SET HEAT RECORD

Threat to Society and Nature Is Rising — Scale of Shift Startles Scientists

By JUSTIN GILLIS

Marking another milestone for a changing planet, scientists reported on Wednesday that the Earth reached its highest temperature on record in 2016, trouncing a record set only a year earlier, which beat one set in 2014. It is the first time in the modern era of global warming data that temperatures have blown past the previous record three years in a row.

The findings come two days before the inauguration of an American president who has called

gases.

"A single warm year is something of a curiosity," said Deke Arndt, chief of global climate monitoring for the National Oceanic and Atmospheric Administration. "It's really the trend, and the fact that we're punching at the ceiling every year now, that is the real indicator that we're undergoing big changes."

The heat extremes were especially pervasive in the Arctic, with temperatures in the fall running

Goddard Institute for
Space Studies

Example 2: The ‘hiatus’ (cont)

Consequences:

- updates to surface data sets removed the anomaly perceived in 2012.

    - notably, after HadCRUT4 and NOAA update (Karl et al, 2015) comparisons improved

        - SST/SAT blend is a better comparison than just SAT

- model forcings were revised.

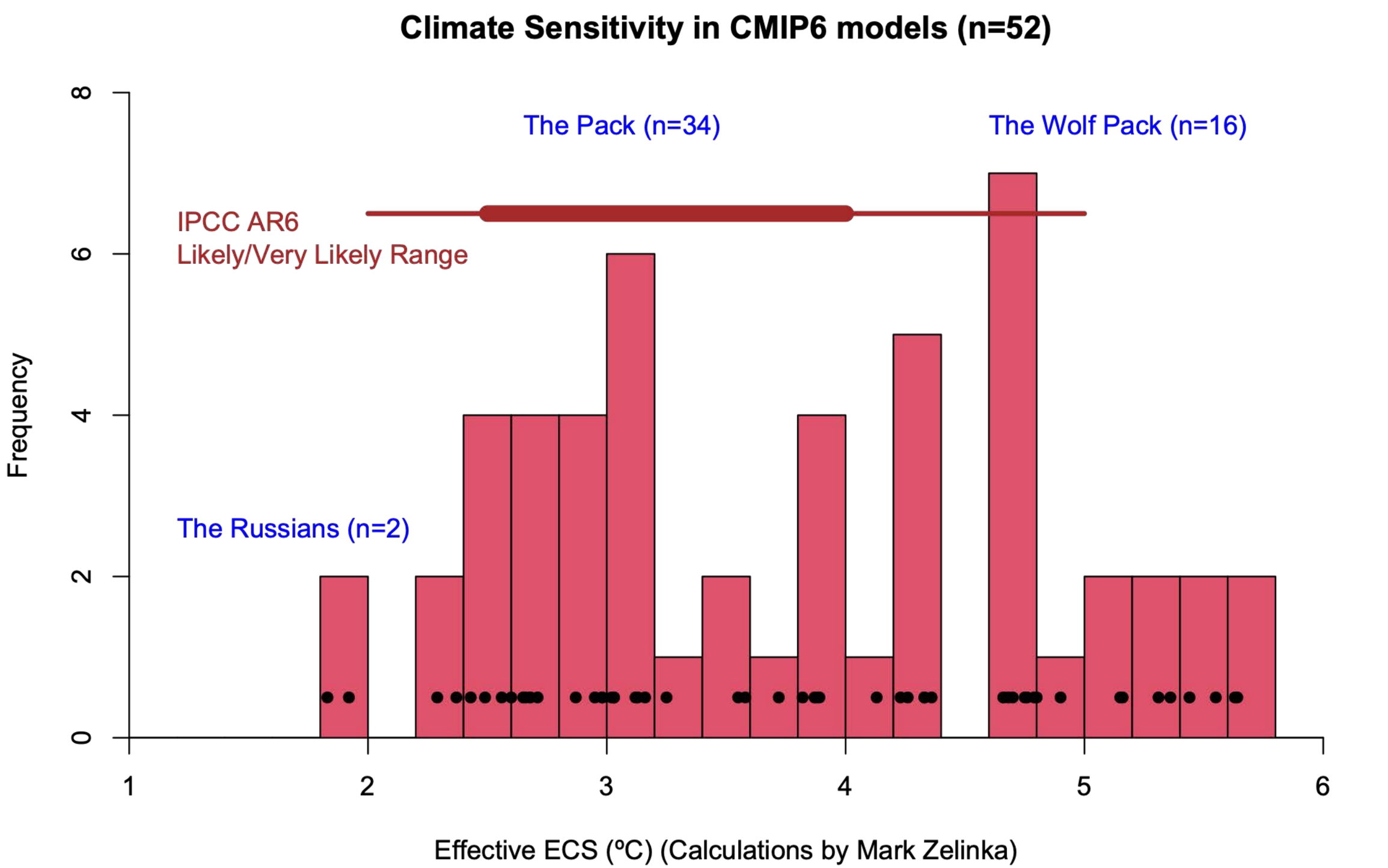
    - volcanoes, solar, maybe aerosols…

2022: “Global warming stopped in 2016”….

2025 (a prediction!): “Global warming stopped in 2024”…

Goddard Institute for Space Studies

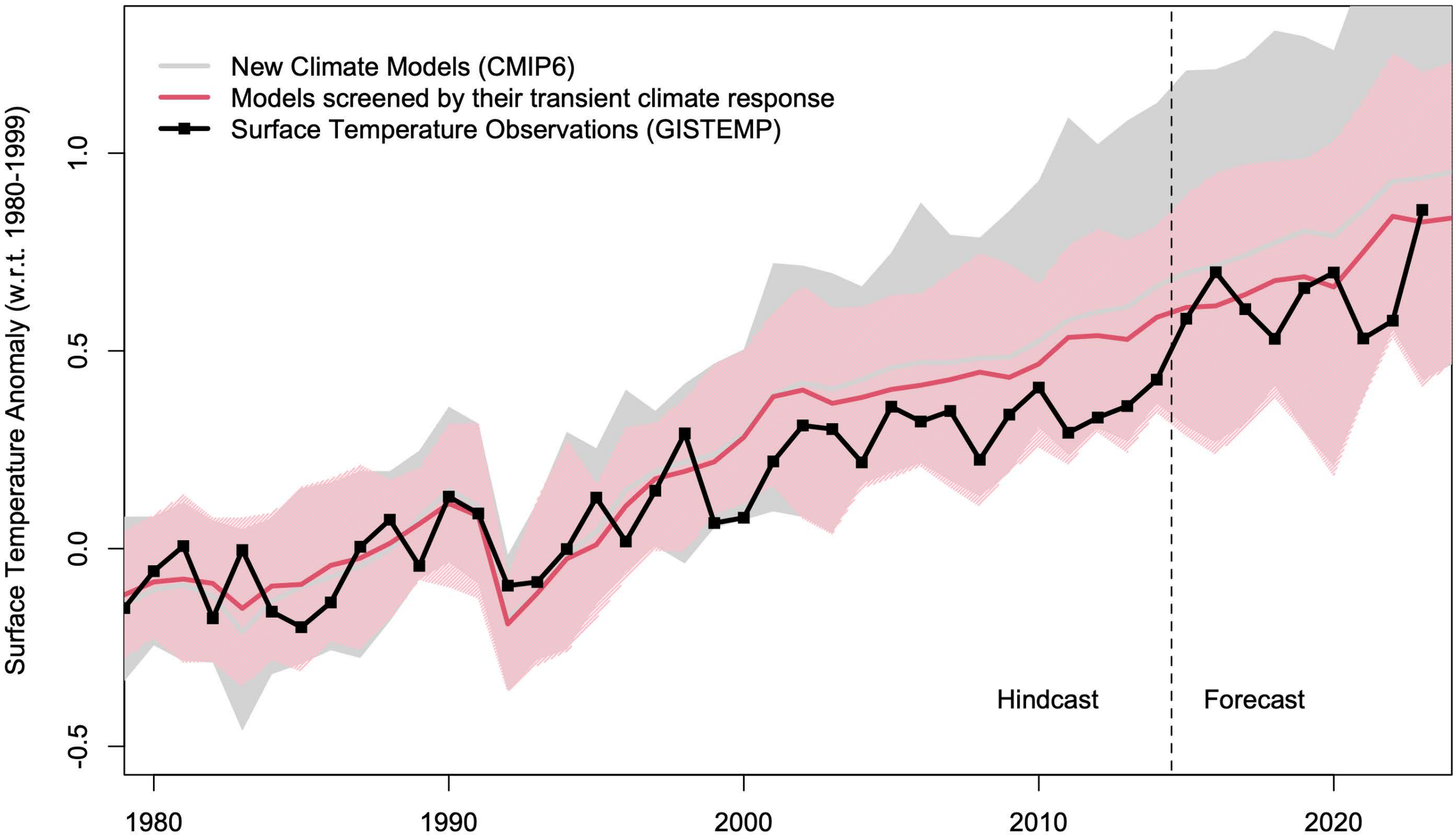
Example 3: CMIP6 produced some surprises…



This is reflected in the ensemble recent trends

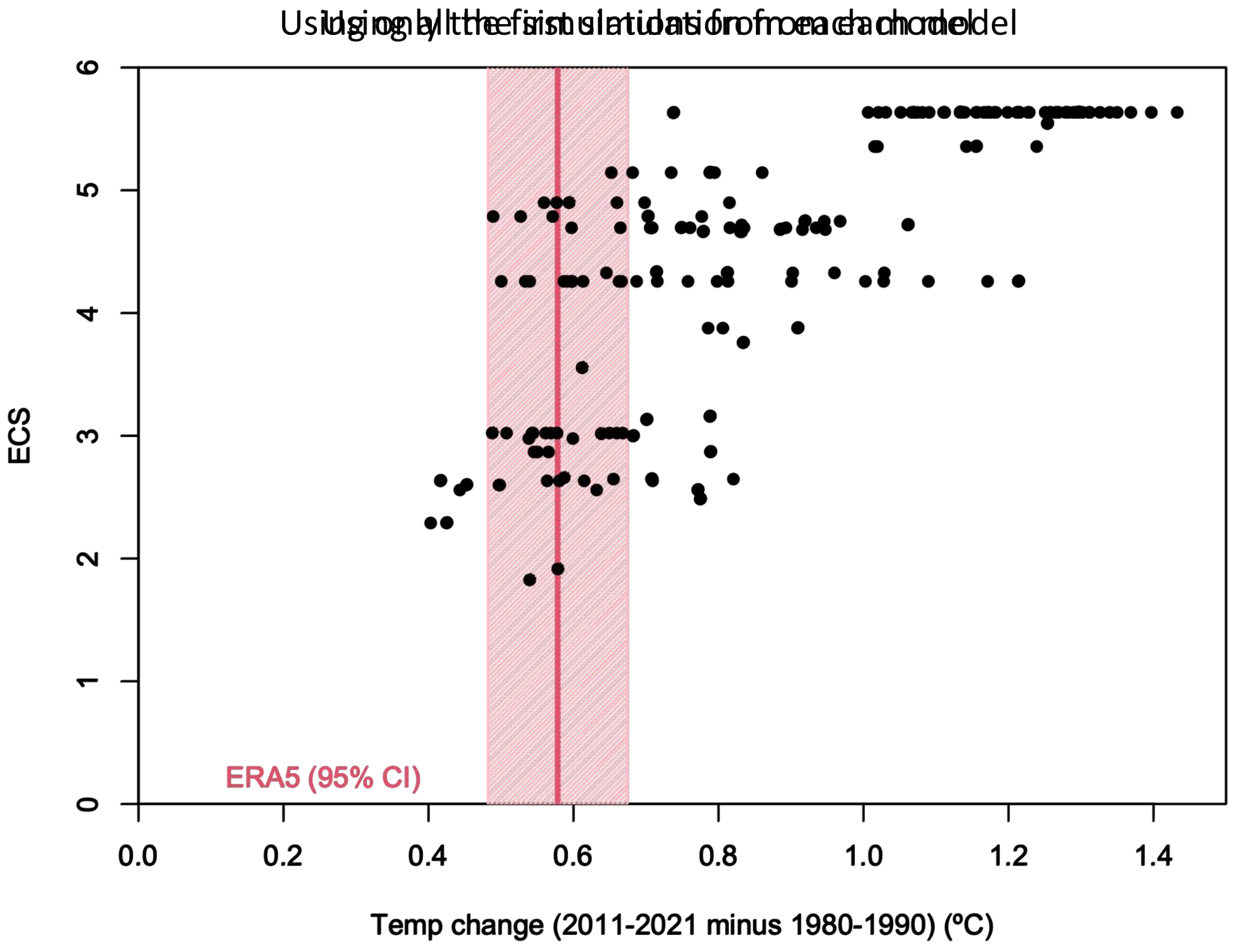
Goddard Institute for
Space Studies

Comparison of latest climate models to observations
(using first simulation from each model)



Goddard Institute for Space Studies

But that's not sufficient to assess the consistency of individual models



Goddard Institute for
Space Studies

What's an appropriate test?

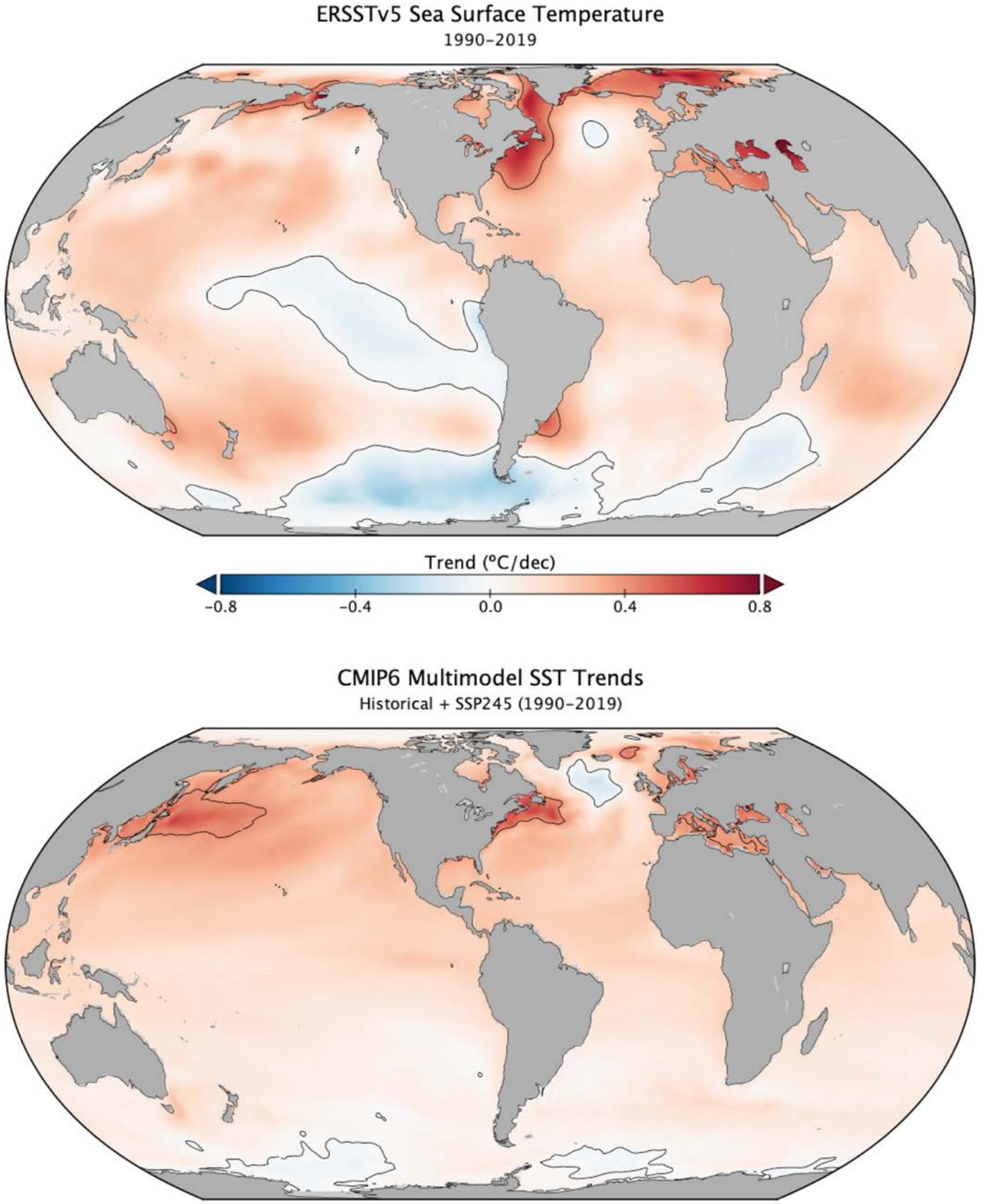
Do the observations plausibly come from the distribution given by a model (or the ensemble)? I.e. are the observations exchangeable with a model ensemble member?

$$d = |\overline{T_m} - \overline{T_o}| / \sqrt{s\{< T_m >\}^2 + s\{T_o\}^2}$$

where 'd' should look like a t-statistic, which takes into account both the ensemble spread and the observational uncertainty (Santer et al, 2008).

Goddard Institute for Space Studies

Example 4: Southern Ocean SST trends in recent decades



SST trends 1990-2019 differ from CMIP ensemble mean:

Eastern tropical Pacific cooling

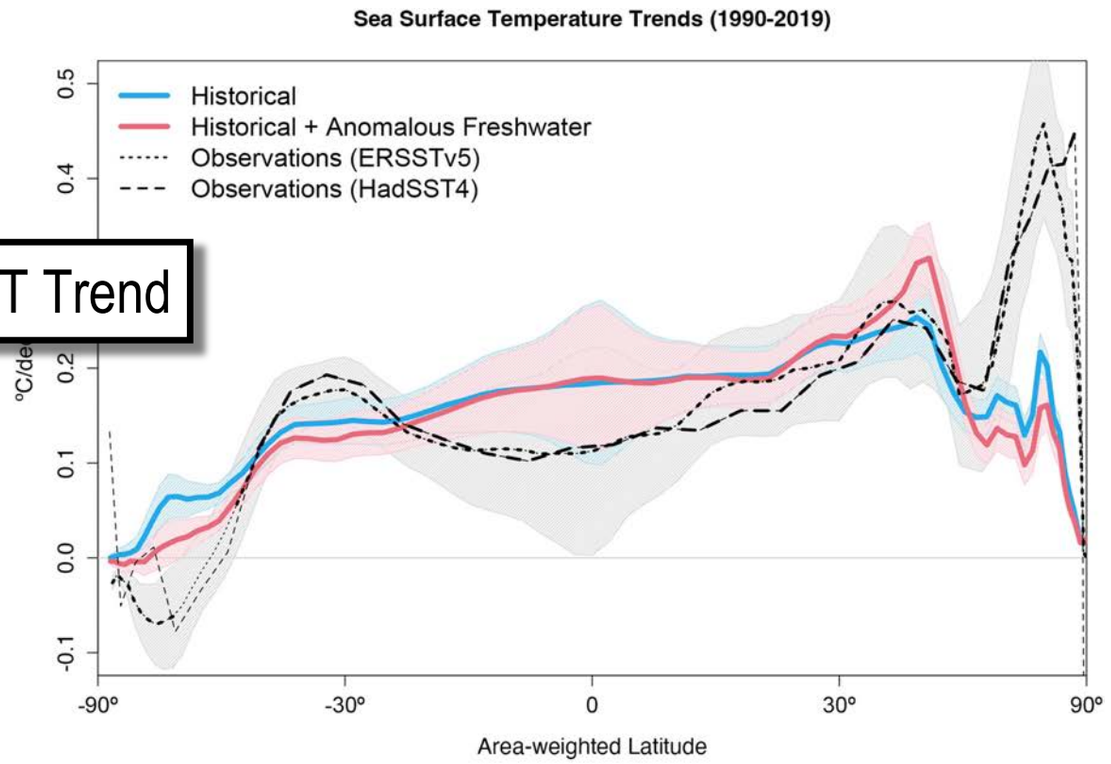
Southern Ocean cooling

Arctic Ocean/Barents Sea warming

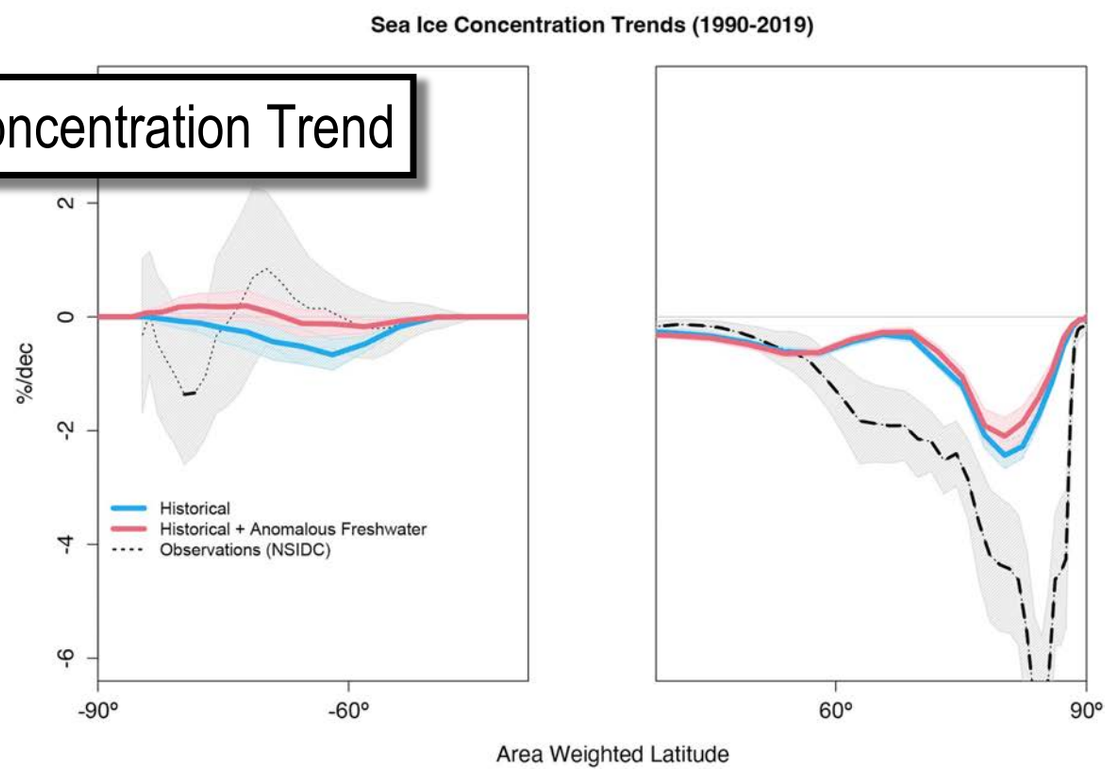
Goddard Institute for Space Studies

Anomalous freshwater from the ice sheets is missing in CMIP models

Zonal SST Trend



Sea Ice Concentration Trend



Add estimated anomalous freshwater from 1990 to 2019 in Southern Ocean and North Atlantic

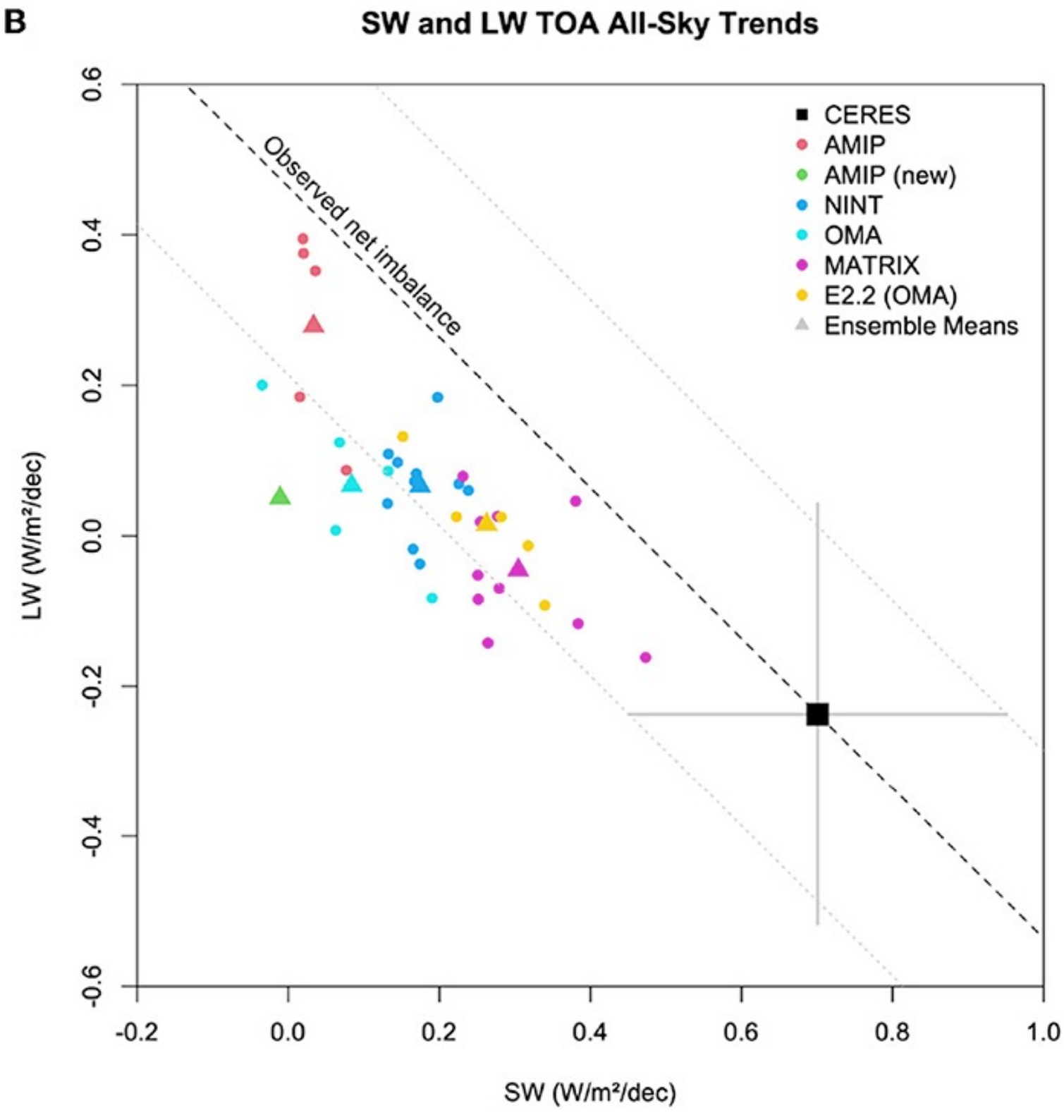
Increases Southern Ocean cooling

Reverses Antarctic sea ice extent trend

Schmidt et al (2023, GRL)

Example 5: Earth's Energy Imbalance

Goddard Institute for Space Studies



The balance of CERES TOA trends between the LW and SW seems anomalous w.r.t. GISS model varieties.

But, comparisons are not like-to-like - AMIP runs only until 2014, SST data is out of date, aerosol emissions only to 2019.

Observational uncertainty is poorly characterized.

CERESMIP project proposed to try and drill down - updates of SST, aerosol and other forcings, GHG- + aerosol-only runs etc. (Schmidt et al, 2023)

Goddard Institute for
Space Studies

To recap…

Models are wrong

Observational data are imperfect

- noisy and biased
- Often w/o estimates of structural uncertainty

Comparisons are not appropriate

- Not like-with-like
- Unrecognized point-of-view biases
- Irreducible effects of chaos

Philosophical Footnote #5

You can’t know ahead of time which hypothesis you are challenging with any discrepancy

Goddard Institute for Space Studies

Recent papers on this topic

Casas, et al. 2023: Understanding model-observation discrepancies in satellite retrievals of atmospheric temperature using GISS ModelE. doi:10.1029/2022JD037523.

Structural issues in model MSU trends

Jain, et al. 2023:, Importance of internal variability for climate model assessment. doi:10.1038/s41612-023-00389-0.

Importance of internal variability in comparisons

Schmidt, et al., 2023: Comment on "Advanced testing of low, medium, and high ECS CMIP6 GCM simulations versus ERA5-T2m" by N. Scafetta (2022). doi:10.1029/2022GL102530.

Importance of internal variability & obs. uncertainty in comparisons

Schmidt, et al., 2023: CERESMIP: A climate modeling protocol to investigate recent trends in the Earth's energy imbalance. doi:10.3389/fclim.2023.1202161.

Investigation of possible model/forcing issues in EEI trends

Schmidt et al., 2023: Anomalous meltwater from ice sheets and ice shelves is a historical forcing. *Geophys. Res. Lett.*, **50**, no. 24, e2023GL106530, doi:10.1029/2023GL106530.

Impact of missing model processes and noise on SST/Antarctic sea ice trends

Roach et al., 2023: Winds and meltwater together lead to Southern Ocean surface cooling and sea ice expansion. doi:10.1029/2023GL105948.