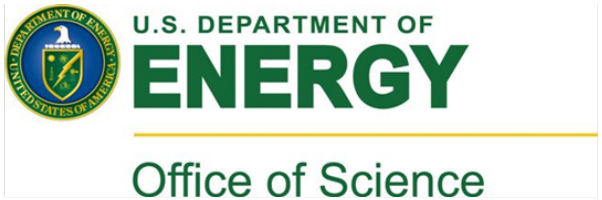
Historical view of the evolution of model performance (climate model errors) over time

Gerald Meehl

National Center for Atmospheric Research

Boulder Colorado



Biological and Environmental Research
Regional and Global Climate Modeling Program

Climate model evaluation today: multiple global earth system models run for hundreds of years, a move towards standardized diagnostics of many fields to compare to multiple observational data sets

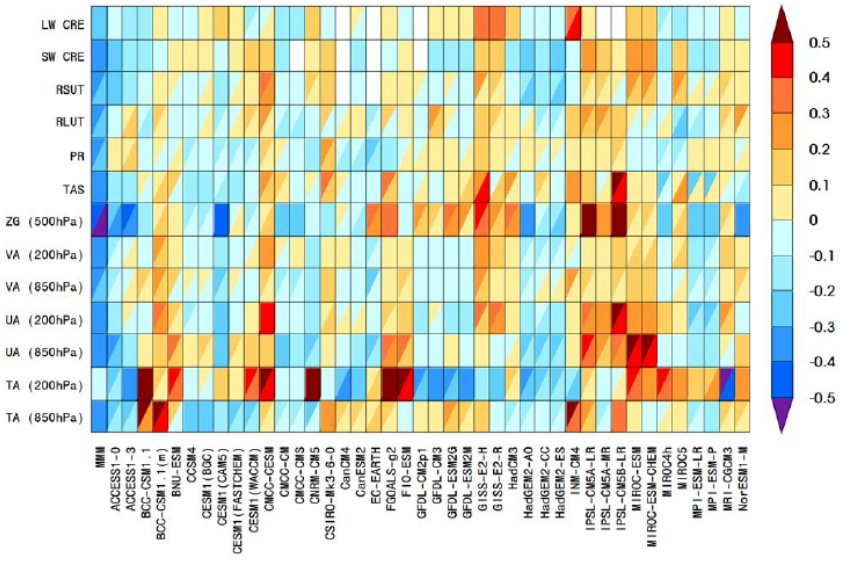


Figure 9.7: Relative error measures of CMIP5 model performance, based on the global seasonal-cycle climatology (1980–2005) computed from the historical experiments. Rows and columns represent individual variables and models, respectively. The error measure is a space–time root-mean-square error (RMSE), which, treating each variable separately, is portrayed as a relative error by normalizing the result by the median error of all model results (P. Gleckler, Taylor, & Doutriaux, 2008). For example, a value of 0.20 indicates that a model's RMSE is 20% larger than the median CMIP5 error for that variable, whereas a value of –0.20 means the error is 20% smaller than the median error. No colour (white) indicates that model results are currently unavailable. A diagonal split of a grid square shows the relative error with respect to both the default reference data set (upper left triangle) and the alternate (lower right triangle). The relative errors are calculated independently for the default and alternate data sets. All reference data used in the diagram are summarized in Table 9.3.

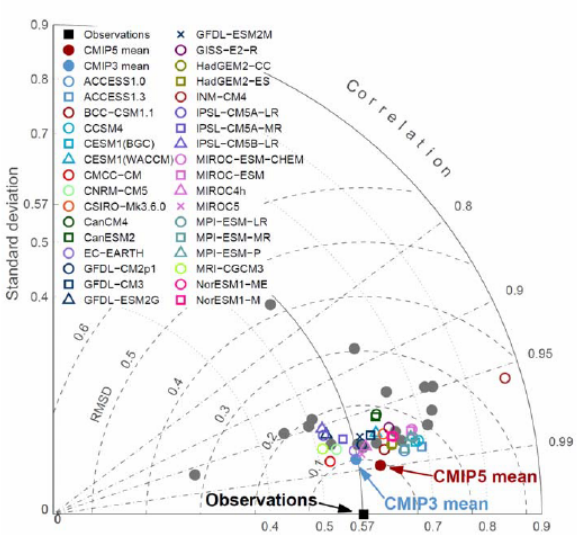


Figure 9.16: Taylor diagram for the dynamic sea-surface height climatology (1987–2000). The radial coordinate shows the standard deviation of the spatial pattern, normalised by the observed standard deviation. The azimuthal variable shows the correlation of the modelled spatial pattern with the observed spatial pattern. The root-mean square error with bias removed is indicated by the dashed grey circles about the observational point. Analysis is for the global ocean, 50°S–50°N. The reference dataset is AVISO, a merged satellite product (Ducet, Le Traon, & Reverdin, 2000), which is described in Chapter 3. One realisation per model is shown for each CMIP5 and CMIP3 model result. Grey filled circles are for individual CMIP3 models; other symbols as in legend.

Climate model evaluation today: multiple global earth system models run for hundreds of years, a move towards standardized diagnostics of many fields to compare to multiple observational data sets

But it wasn't always this way…

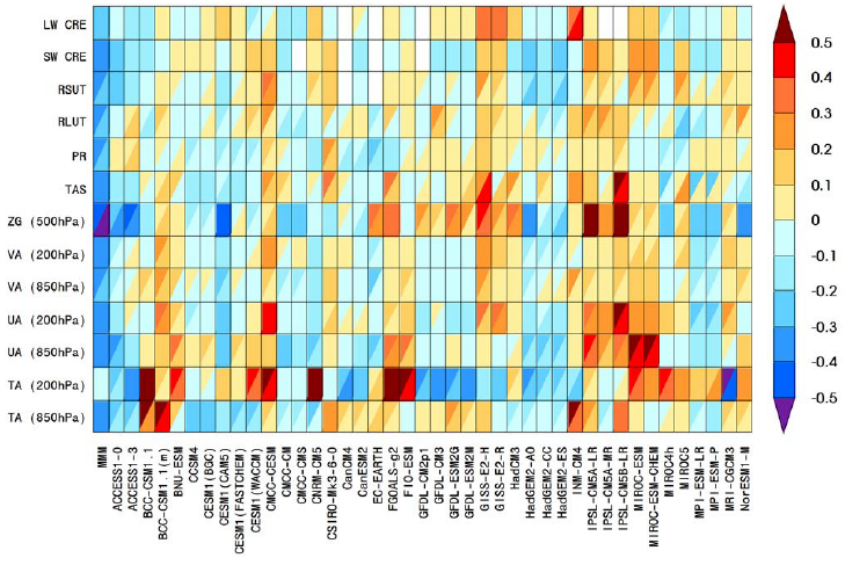


Figure 9.7: Relative error measures of CMIP5 model performance, based on the global seasonal-cycle climatology (1980–2005) computed from the historical experiments. Rows and columns represent individual variables and models, respectively. The error measure is a space–time root-mean-square error (RMSE), which, treating each variable separately, is portrayed as a relative error by normalizing the result by the median error of all model results (P. Gleckler, Taylor, & Doutriaux, 2008). For example, a value of 0.20 indicates that a model's RMSE is 20% larger than the median CMIP5 error for that variable, whereas a value of –0.20 means the error is 20% smaller than the median error. No colour (white) indicates that model results are currently unavailable. A diagonal split of a grid square shows the relative error with respect to both the default reference data set (upper left triangle) and the alternate (lower right triangle). The relative errors are calculated independently for the default and alternate data sets. All reference data used in the diagram are summarized in Table 9.3.

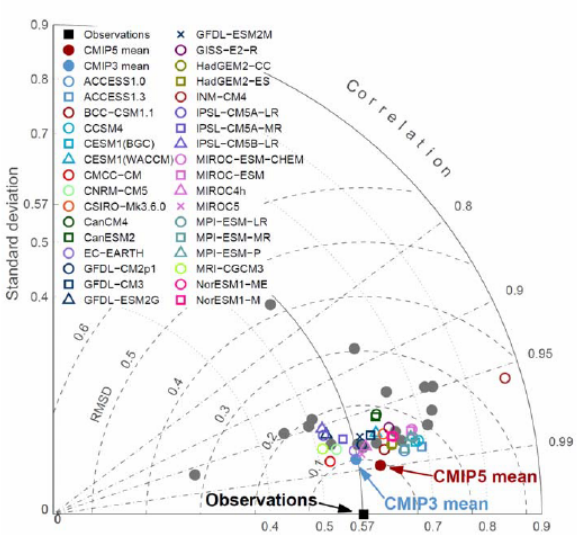
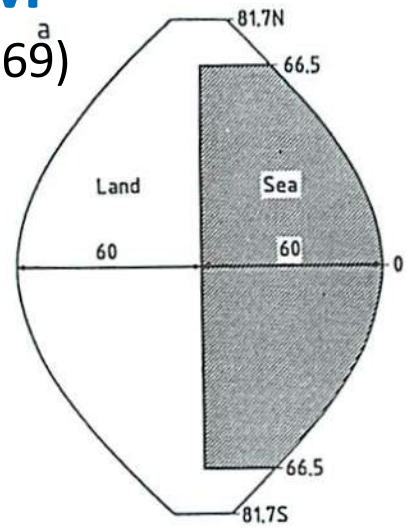
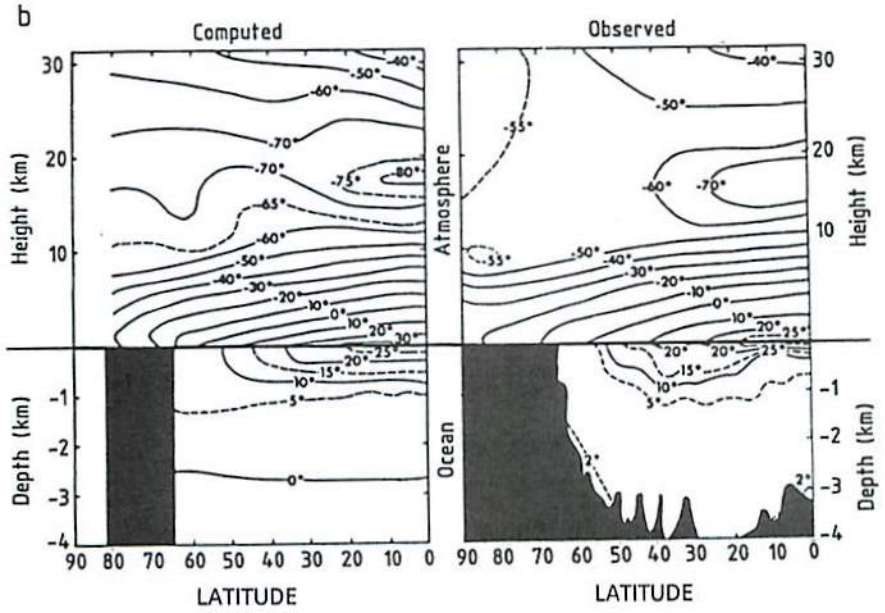


Figure 9.16: Taylor diagram for the dynamic sea-surface height climatology (1987–2000). The radial coordinate shows the standard deviation of the spatial pattern, normalised by the observed standard deviation. The azimuthal variable shows the correlation of the modelled spatial pattern with the observed spatial pattern. The root-mean square error with bias removed is indicated by the dashed grey circles about the observational point. Analysis is for the global ocean, 50°S–50°N. The reference dataset is AVISO, a merged satellite product (Ducet, Le Traon, & Reverdin, 2000), which is described in Chapter 3. One realisation per model is shown for each CMIP5 and CMIP3 model result. Grey filled circles are for individual CMIP3 models; other symbols as in legend.

In the beginning, there was the "sector coupled ocean-atmosphere GCM"

(Manabe and Bryan, 1969)





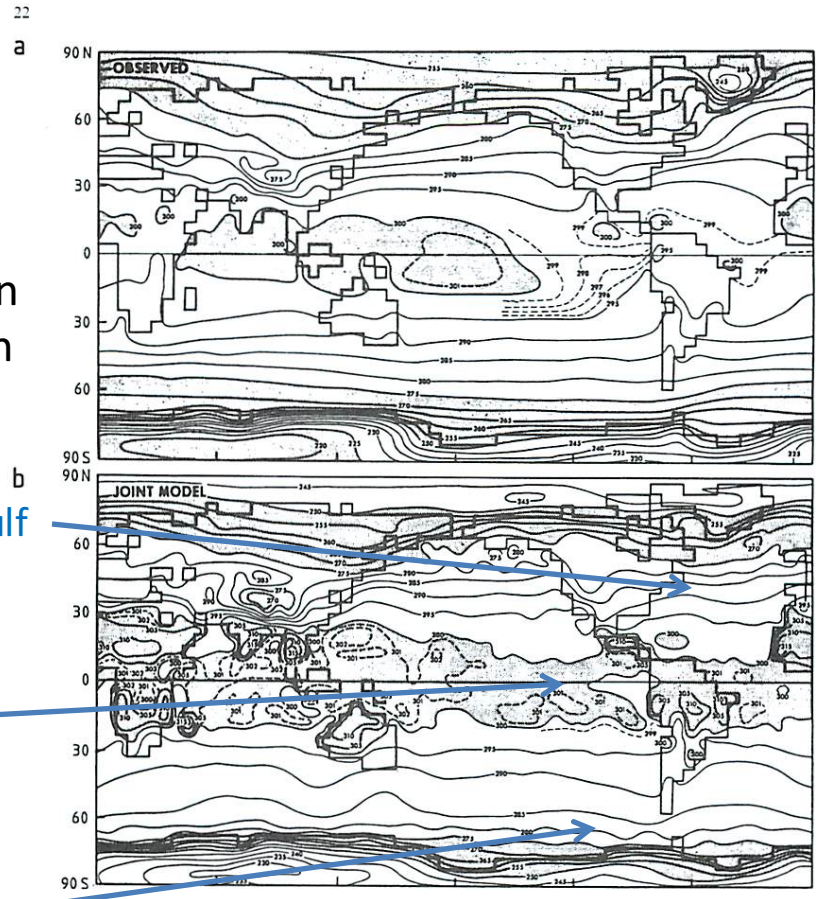
5° lat-lon grid in atmosphere and ocean, 9 level atmosphere, 5 ocean levels

"Asynchronous coupling" atmosphere run for a year, ocean for 100 years with annual mean solar forcing; models communicated every half year in atmosphere

Calculation took 1200 hours on a UNIVAC 1108, a huge computational effort at the time

As computer power increased, *global* coupled climate models emerged

(Manabe et al., 1975)



A few errors in the simulation were noted:

Poorly represented Gulf Stream?

Little east-west SST gradient across the Pacific?

Southern Ocean too warm?

5° lat-lon grid in atmosphere and ocean, 9 level atmosphere, 12 ocean levels

"Asynchronous coupling" atmosphere run for ~one year, ocean for 272 years with annual mean solar forcing; models communicated every half year in atmosphere

Compared to a "swamp" ocean, notable improvements from ocean dynamics

Limitations of computer power in the late-1970s necessitated "asynchronous coupling schemes"

GFDL

NCAR

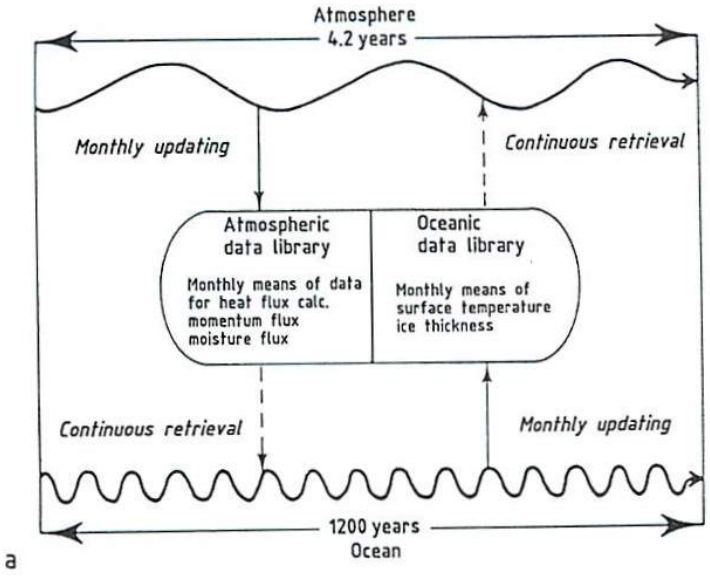


Fig. 3a, b. Asynchronous coupling schemes for seasonal cycle integrations for a GFDL coupled model (Manabe et al. 1979); b NCAR coupled model (Washington et al. 1980)

Late-1970s: Even in these early coarse-grid coupled models, now-familiar systematic errors were present

(2 member multi-model data set: computed minus observed SST)

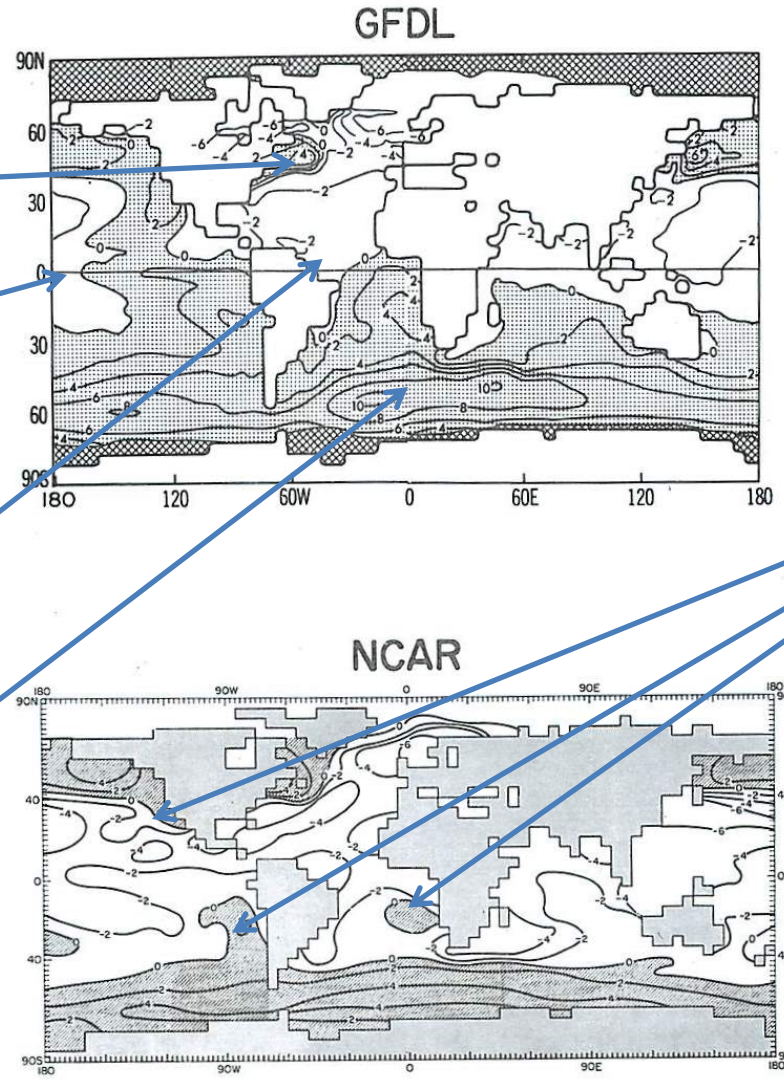
Common errors in both models:

Poorly represented Gulf Stream

Western equatorial Pacific too cold (eastern Pacific cold tongue extends too far west)

Western tropical Atlantic too cold

Southern Ocean too warm



Too warm in west coast stratus regimes

FIG. 8. Annual average mean sea surface temperature (SST) differences (computed minus observed) for GFDL (Manabe *et al.*, 1979, and Corrigendum) and NCAR coupled models. Positive areas (computed warmer than observed) are shaded.

(the first NCAR coupled model: Washington et al., 1980; 5 degree 8L atmosphere, 5 degree 4L ocean)

In the 1980s, spectral atmospheric models began to be used in coupled models, and interesting phenomena beyond climatological means began to be documented in comparison to observations, such as blocking

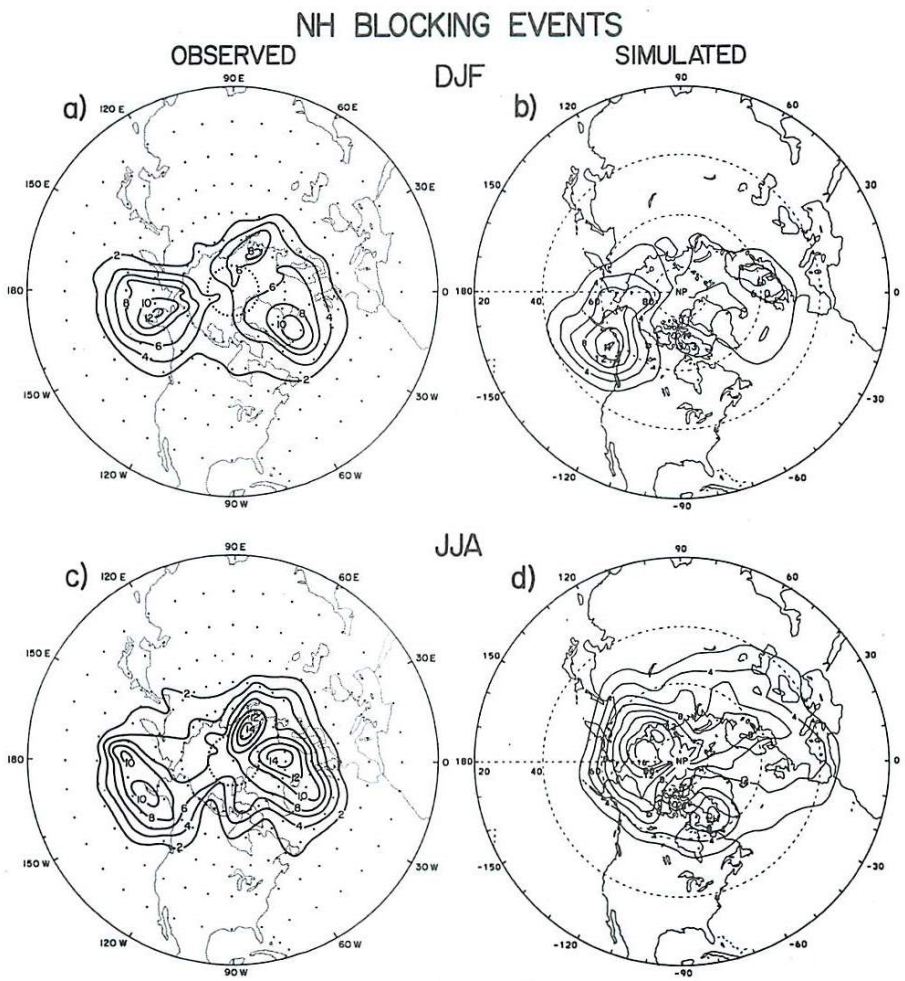


FIG. 7. Number of Northern Hemisphere 500 mb persistent height anomalies indicative of blocking events per 15 years in the observations and the NCAR CCM control simulation in winter (DJF): (a) observed, (b) simulated; and summer (JJA) (c) observed, (d) simulated. The criteria for persistent height anomalies for both observations and model are $T = 7$ days and $M = 200$ m in winter, 100 m in summer. Observations are from Shukla and Mo (1983).

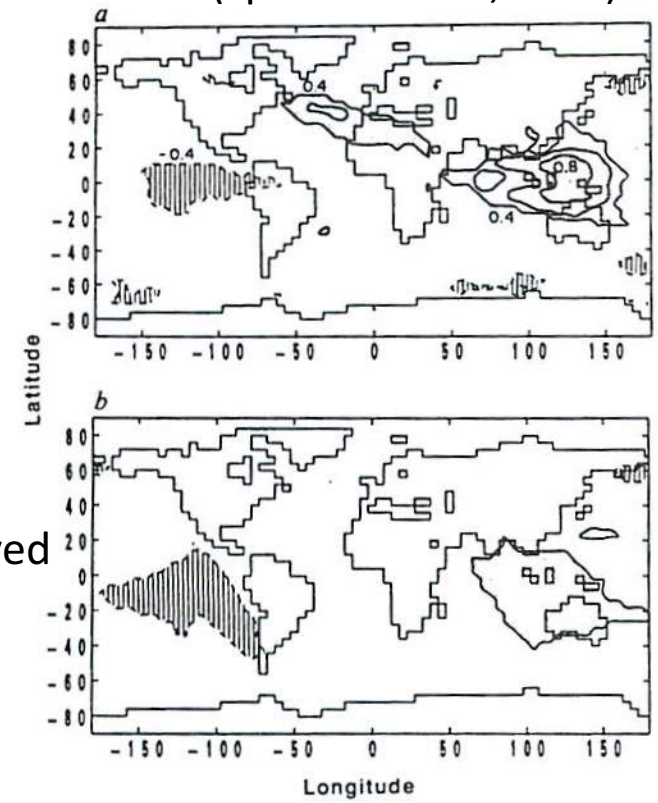
NCAR R15 (4.5° x 7.5°)
9 level coupled model
(Bates and Meehl, 1986)

Then something unexpected happened—coarse-grid coupled models in the late-1980s produced something that looked like the Southern Oscillation

OSU model (Sperber et al., 1987)

Model

Observed

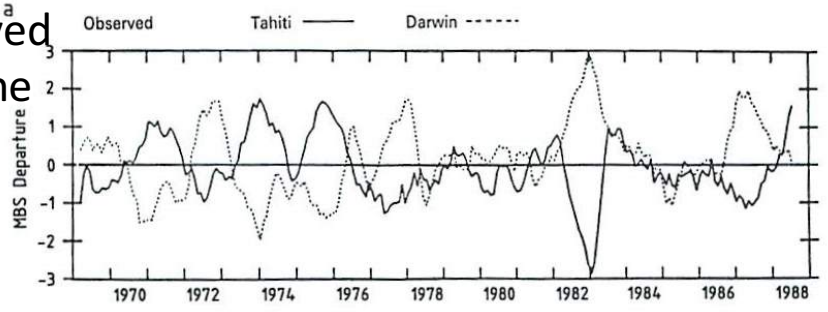


Correlation of Darwin SLP with all other grid points

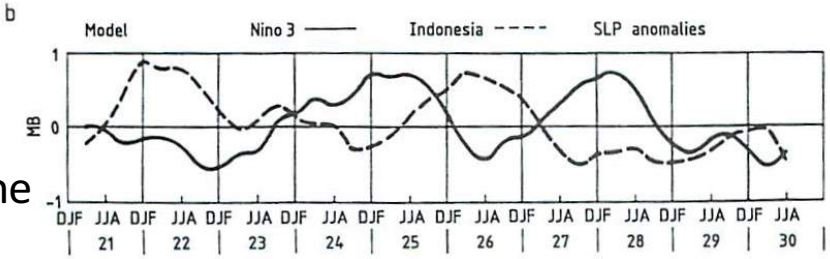
Fig. 12. a Correlation of sea-level pressure at a point near Darwin, Australia, with all other points in the coupled model; only areas greater than 95% significance are shown; b same as a, except for observed correlations (Sperber et al. 1987)

NCAR model (Meehl, 1990)

Observed SLP time series



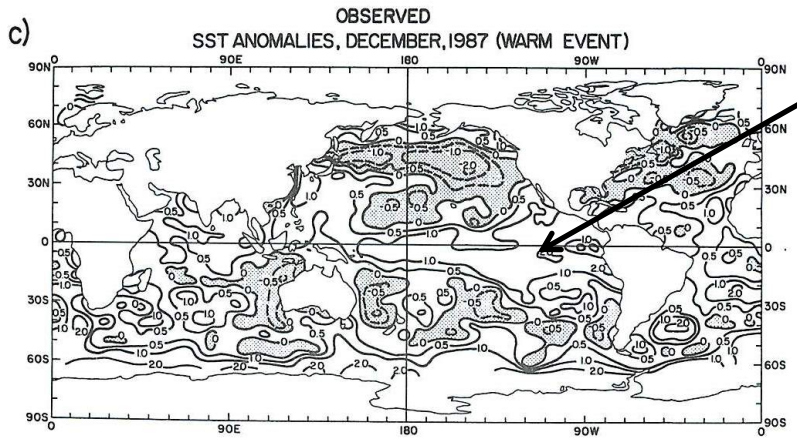
Model SLP time series

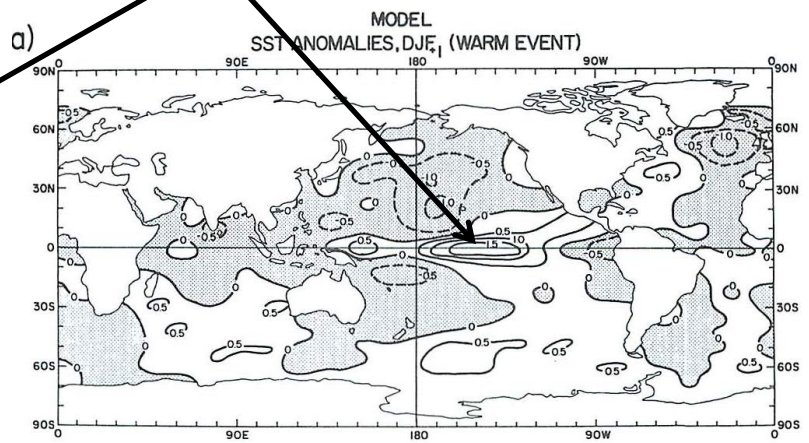


El Niño events were being spontaneously generated in a coarse-grid coupled model—that wasn't supposed to happen since the 5 degree ocean model could not produce internal ocean waves, thus no delayed action oscillator, the accepted El Niño mechanism at that time

NCAR model (Meehl, 1990)

Model El Niño events shifted west compared to observations





Model El Niño teleconnections similar to observations, but shifted west

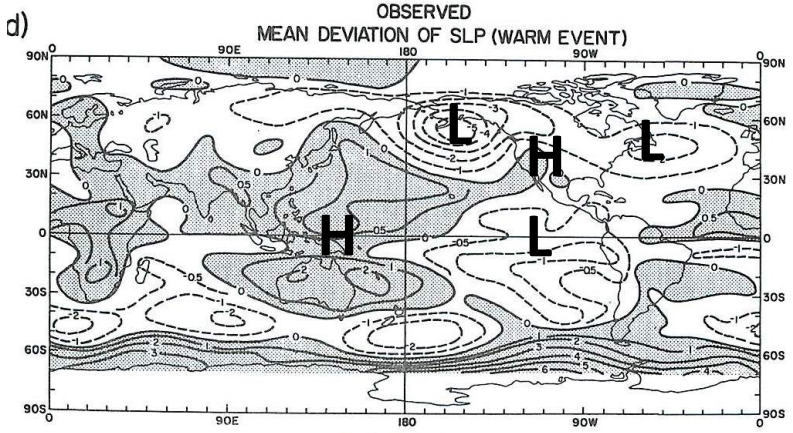


FIG. 10. (*Continued*)

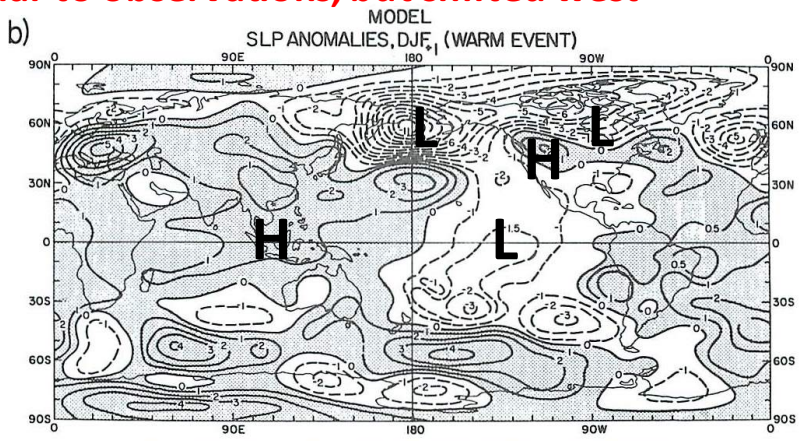
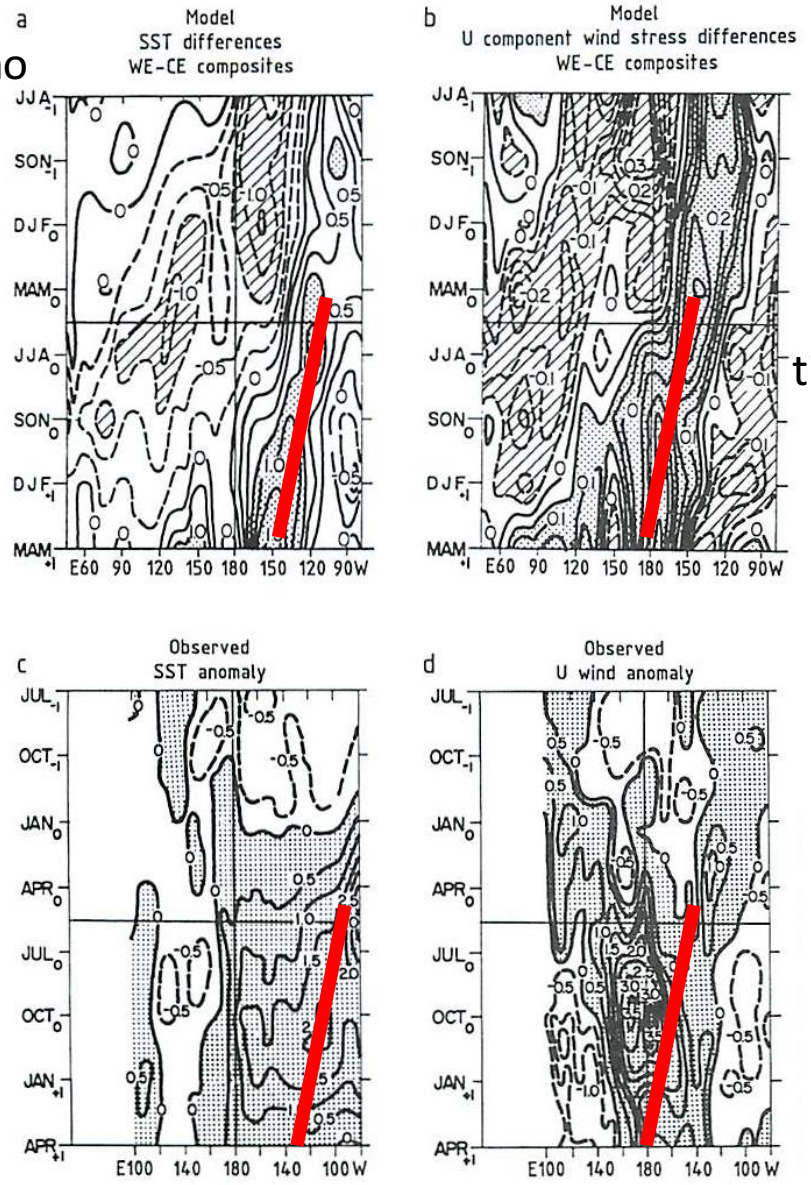


FIG. 10. As in Fig. 9 except for computed and observed SST and SLP anomalies representative of a warm event in the central equatorial Pacific.

And the model SST anomalies in the equatorial Pacific propagated from east to west, similar to a composite of observed events (1957, 1965, 1972), in association with westerly wind stress anomalies just to the west of the positive SST anomalies

Model El Niño



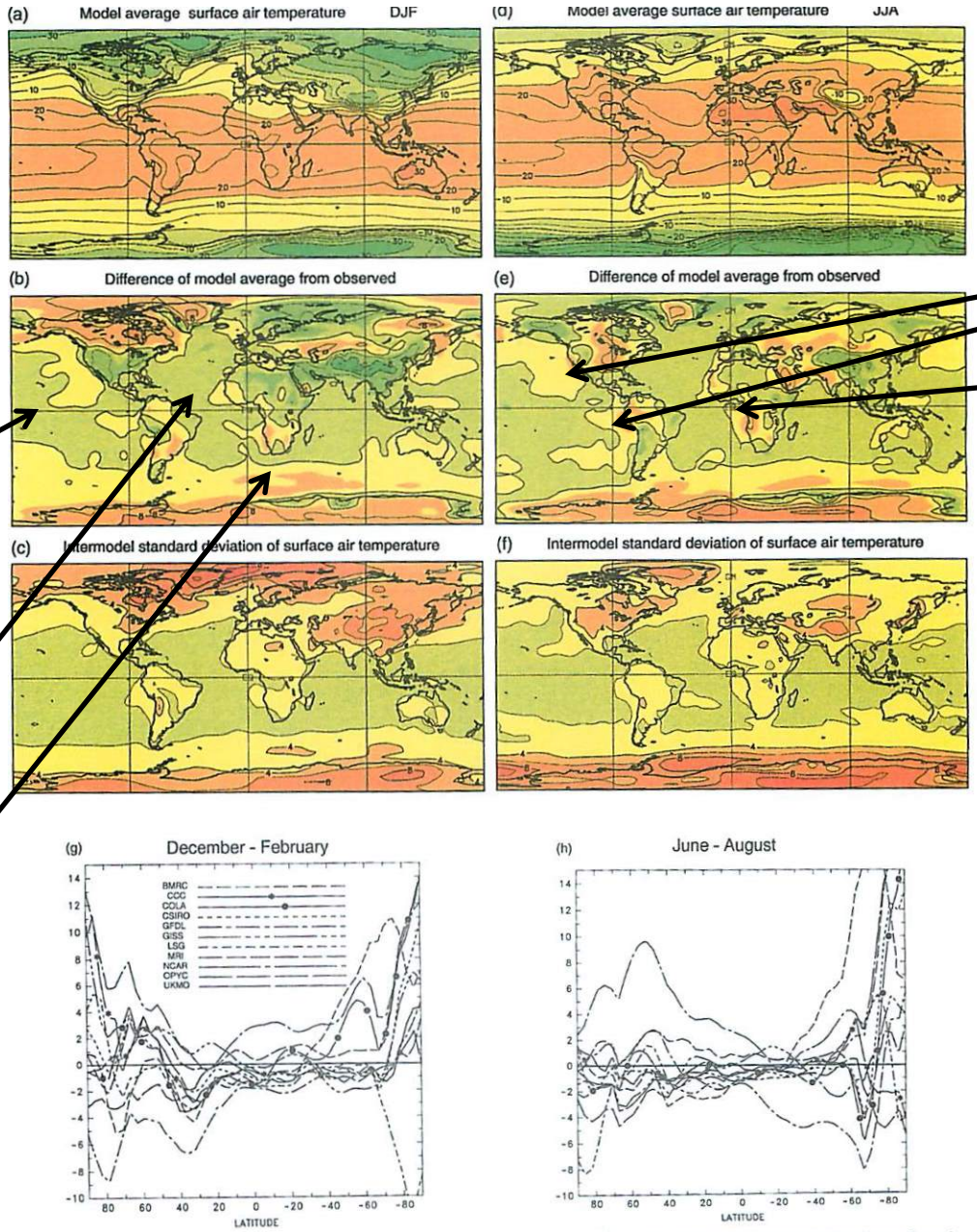
NCAR model (Meehl, 1990)

1995 IPCC Second Assessment Report (SAR), 9 models, pre-CMIP, R15 class (4.5° latitude x 7.5° longitude) Some flux corrected, some not—causes complications for model evaluation

Equatorial western Pacific too cold

Western tropical Atlantic too cold

Southern Ocean too warm



Too warm in west coast stratus regimes

Figure 5.1: (a) The average surface air temperature simulated by nine coupled models, (b) the difference of the model average from the observed data of Jenne (1975) and (c) the intermodel standard deviation for DJF. (d-f) As in (a-c) but for JJA. (g) The zonally averaged difference of eleven coupled models' surface air temperature from observations for DJF. (h) As in (g) but for JJA. Units °C.

A metric from the SAR 1995: interannual standard deviation for multi-century control runs
Greater over continents, less over tropical oceans, low amplitude El Niño

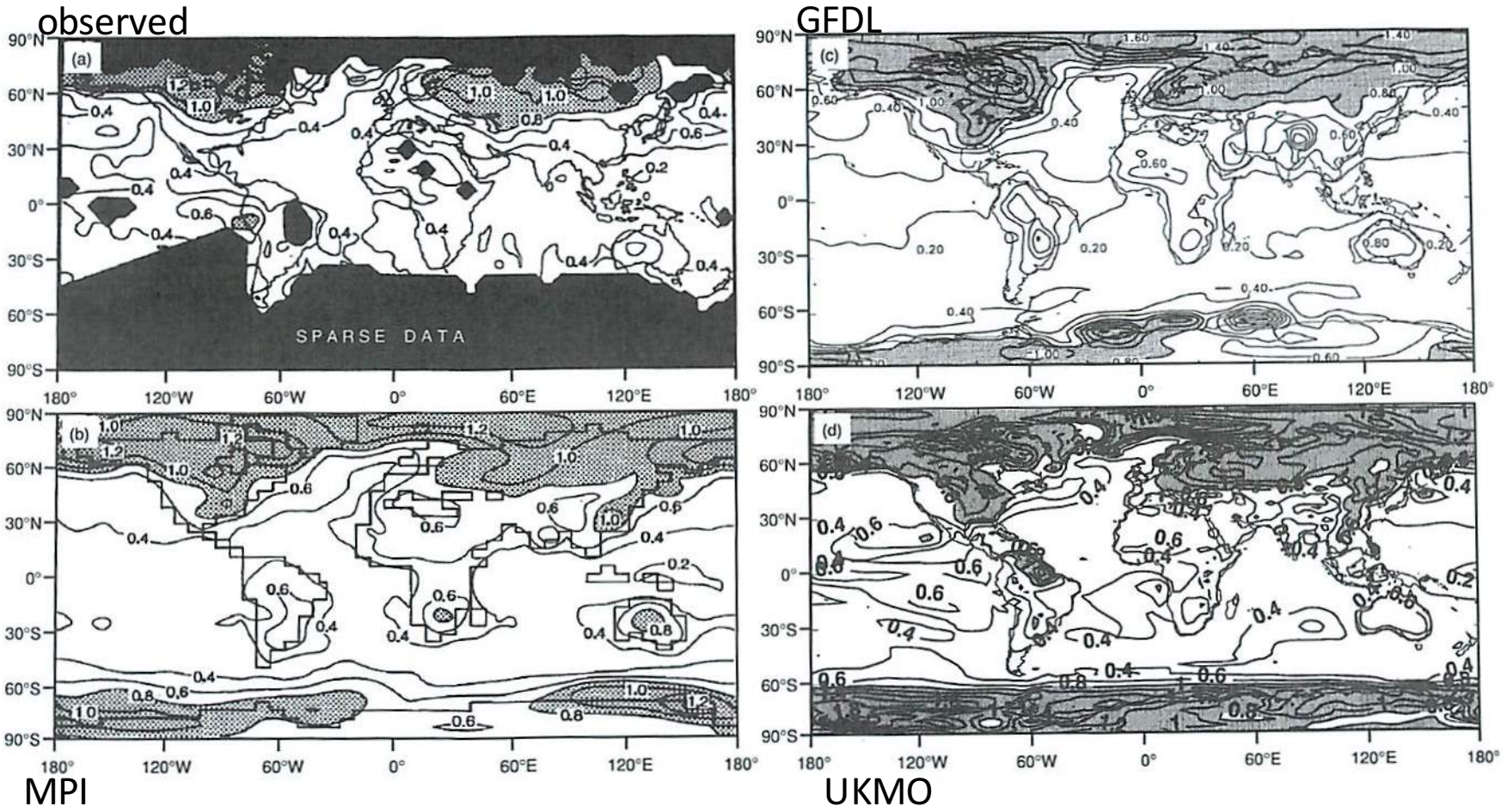


Figure 5.6: The standard deviation of the annual mean surface air temperature: (a) over 110 years of observations (Jones *et al.*, 1991), (b) over 1000 yr from the GFDL coupled model, (c) over 350 yr from the MPI/LSG coupled model, and (d) over 130 yr from the UKMO coupled model. Units °C.

IPCC Third Assessment Report (TAR) 2001; R15 (~5° resolution) to T42 (~3° resolution); ~20 models (CMIP1,2); flux corrected models had lower amplitude SST errors, but non-flux corrected models still had familiar systematic errors

Western tropical Pacific too cold (eastern Pacific cold tongue extends too far west)

Southern Ocean too warm

Flux-corrected models

non-flux-corrected models

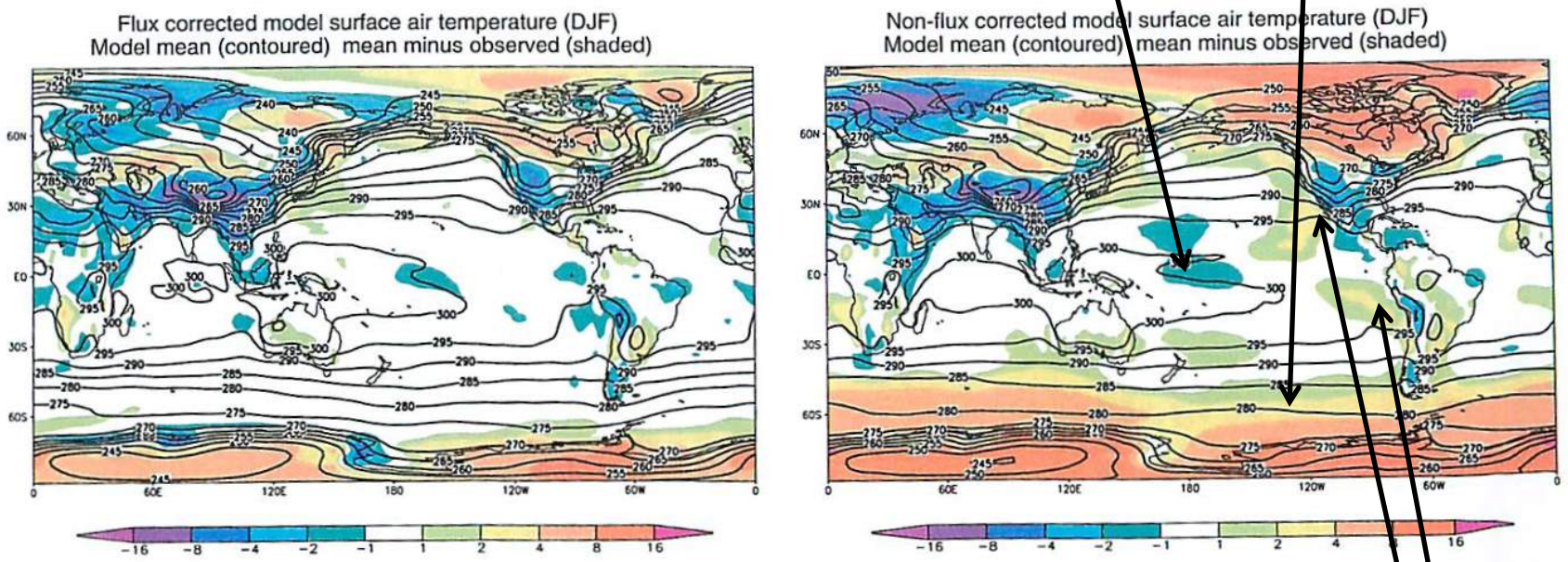


Figure 8.2: December-January-February climatological surface air temperature in K simulated by the CMIP1 model control runs. Averages over all models (upper left), over all flux adjusted models (lower left) and over all non-flux adjusted models (lower right) are shown together with

Too warm in west coast stratus regimes

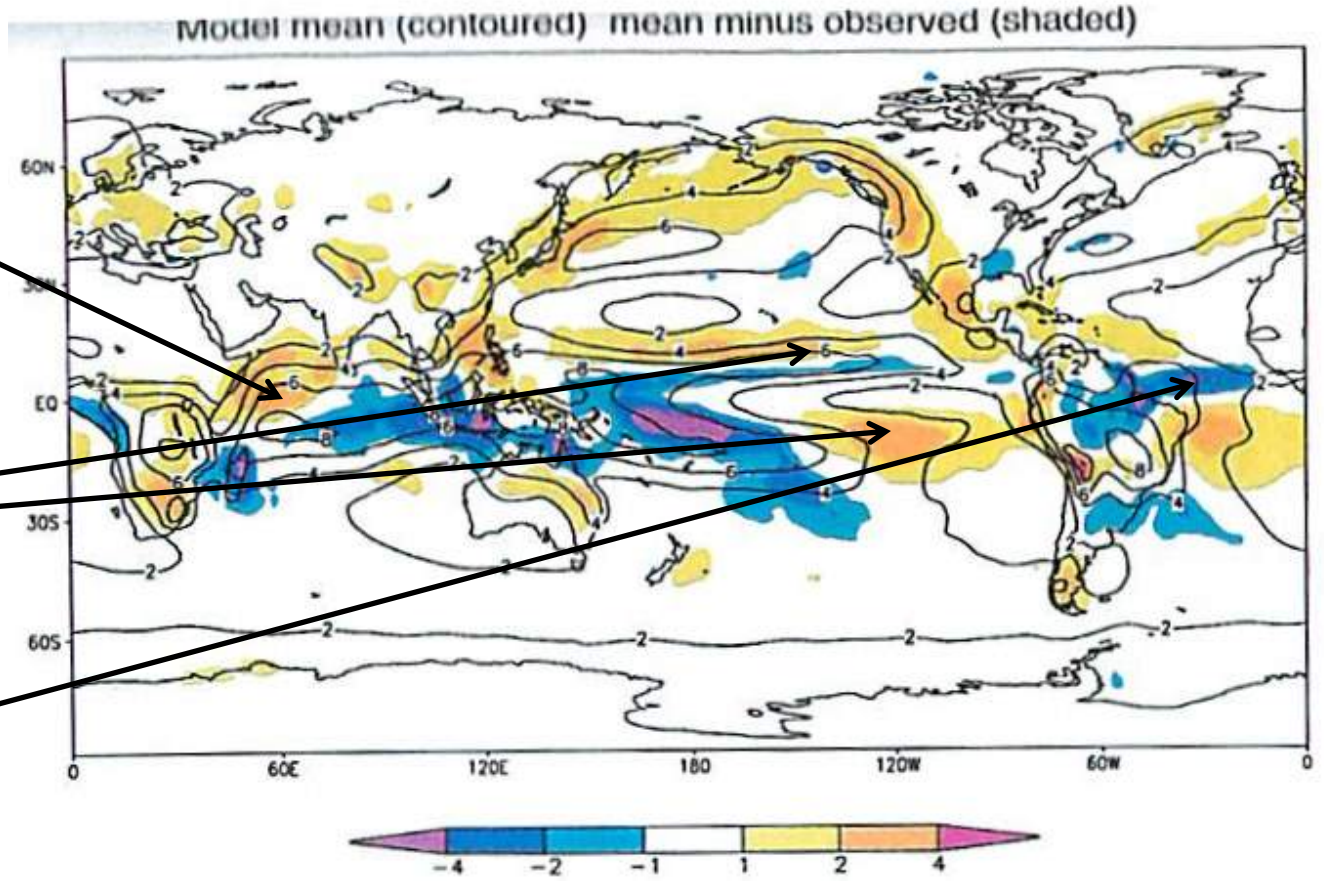
IPCC TAR, 2001; R15-T42 models (CMIP1,2)

Precipitation errors (colored shading)

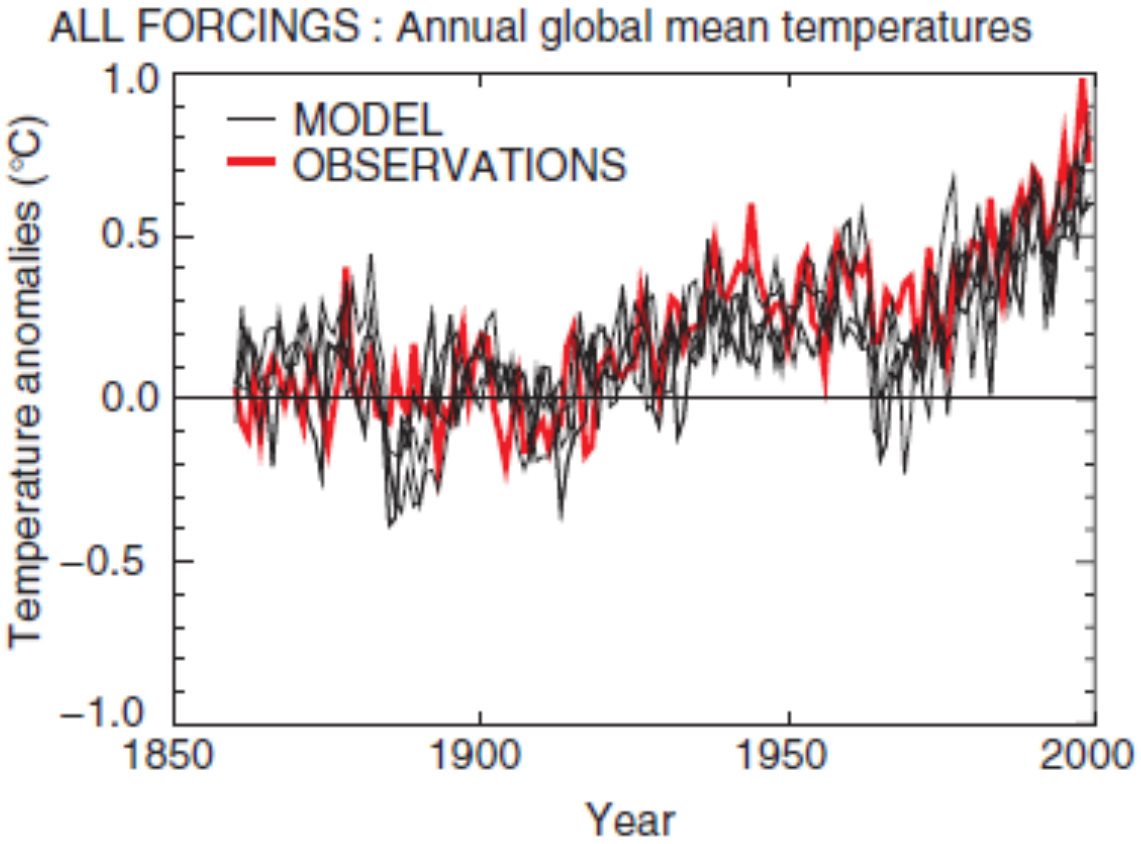
Too much rainfall too far west in tropical Indian Ocean

Double ITCZ

Too little rainfall in equatorial western Atlantic



The IPCC TAR (2001) showed a new way of evaluating a climate model: can it simulate the time series of 20th century temperature by including natural and anthropogenic forcings?



1998: The landmark AGCI session on extremes emerged from the IPCC Second Assessment Report (1995) where it was recognized that the impacts community wanted to study future changes in weather and climate extremes, but there was little research being done on this topic in the physical climate science community

It was decided to bring together climate modelers, observationalists, and impacts researchers in an AGCI session to chart a research agenda to study changes in weather and climate extremes for the first time

The 1998 session produced five articles in the March 2000 issue of the Bulletin of the American Meteorological Society…

An Introduction to Trends in Extreme Weather and Climate Events: Observations, Socioeconomic Impacts, Terrestrial Ecological Impacts, and Model Projections*

Gerald A. Meehl,[a] Thomas Karl,[b] David R. Easterling,[b] Stanley Changnon,[c] Roger Pielke Jr.,[a] David Changnon,[d] Jenni Evans,[e] Pavel Ya. Groisman,[b] Thomas R. Knutson,[f] Kenneth E. Kunkel,[c] Linda O. Mearns,[a] Camille Parmesan,[g] Roger Pulwarty,[h] Terry Root,[i] Richard T. Sylves,[j] Peter Whetton,[k] and Francis Zwiers[l]

ABSTRACT

Weather and climatic extremes can have serious and damaging effects on human society and infrastructure as well as on ecosystems and wildlife. Thus, they are usually the main focus of attention of the news media in reports on climate. There are some indications from observations concerning how climatic extremes may have changed in the past

Observed Variability and Trends in Extreme Climate Events: A Brief Review*

D. R. Easterling,+ J. L. Evans,# P. Ya. Groisman,+ T. R. Karl,+ K. E. Kunkel,@ and P. Ambenje&

ABSTRACT

Variations and trends in extreme climate events have only recently received much attention. Exponentially increasing economic losses, coupled with an increase in deaths due to these events, have focused attention on the possibility that these events are increasing in frequency. One of the major problems in examining the climate record for changes in extremes is a lack of high-quality, long-term data. In some areas of the world increases in extreme events are apparent,

Trends in Extreme Weather and Climate Events: Issues Related to Modeling Extremes in Projections of Future Climate Change*

Gerald A. Meehl,+ Francis Zwiers,# Jenni Evans,@ Thomas Knutson,& Linda Mearns,+ and Peter Whetton**

ABSTRACT

Projections of statistical aspects of weather and climate extremes can be derived from climate models representing possible future climate states. Some of the recent models have reproduced results previously reported in the Intergovernmental Panel on Climate Change (IPCC) Second Assessment Report, such as a greater frequency of extreme warm days and lower frequency of extreme cold days associated with a warmer mean climate, a decrease in diurnal temperature range associated with higher nighttime temperatures, increased precipitation intensity, midcontinent summer drying, decreasing daily variability of surface temperature in winter, and increasing variability of northern midlatitude summer surface temperatures. This reconfirmation of previous results gives an increased confidence in the credibility of the models.

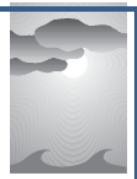
Human Factors Explain the Increased Losses from Weather and Climate Extremes*

Stanley A. Changnon,+ Roger A. Pielke Jr.,# David Changnon,@ Richard T. Sylves,& and Roger Pulwarty,**

ABSTRACT

Societal impacts from weather and climate extremes, and trends in those impacts, are a function of both climate and society. United States losses resulting from weather extremes have grown steadily with time. Insured property losses have trebled since 1960, but deaths from extremes have not grown except for those due to floods and heat waves. Data on losses are difficult to find and must be carefully adjusted before meaningful assessments can be made. Adjustments to historical loss data assembled since the late 1940s shows that most of the upward trends found in financial losses are due to societal shifts leading to ever-growing vulnerability to weather and climate extremes. Geographical locations of the large loss trends establish that population growth and demographic shifts are the major factors behind the increasing losses from weather–climate extremes. Most weather and climate extremes in the United States do not exhibit steady, multidecadal increases found in their loss values. Without major changes in societal responses to weather and climate extremes, it is reasonable to predict ever-increasing losses even without any detrimental climate changes. Recognition of these trends in societal vulnerability to weather–climate extremes suggests that the present focus on mitigating the greenhouse effect should be complemented by a greater emphasis on adaptation. Identifying and understanding this

Impacts of Extreme Weather and Climate on Terrestrial Biota*



Camille Parmesan,+,& Terry L. Root,# and Michael R. Willig@

ABSTRACT

Climate is a driver of biotic systems. It affects individual fitness, population dynamics, distribution and abundance of species, and ecosystem structure and function. Regional variation in climatic regimes creates selective pressures for the evolution of locally adapted physiologies, morphological adaptations (e.g., color patterns, surface textures, body shapes and sizes), and behavioral adaptations (e.g., foraging strategies and breeding systems). In the absence of humans, broad-scale, long-term consequences of climatic warming on wild organisms are generally predictable. Evidence from Pleistocene glaciations indicates that most species responded ecologically by shifting their ranges poleward and upward in elevation, rather than evolutionary through local adaptation (e.g., morphological changes). But these broad patterns tell us little about the relative importance of gradual climatic trends as compared to extreme weather events in shaping these processes. Here, evidence is brought forward that extreme weather events can be implicated as mechanistic drivers of broad ecological responses to climatic trends. They are, therefore, essential to include in predictive bio-

…culminating with a review article in Science (the 22 September 2000 issue)

These papers set the climate science and impacts community on a course of research that has become one of the central areas of study with respect to understanding past and future extremes that has been featured prominently in every subsequent IPCC assessment

SCIENCE'S COMPASS • REVIEW

REVIEW: ATMOSPHERIC SCIENCE

Climate Extremes: Observations, Modeling, and Impacts

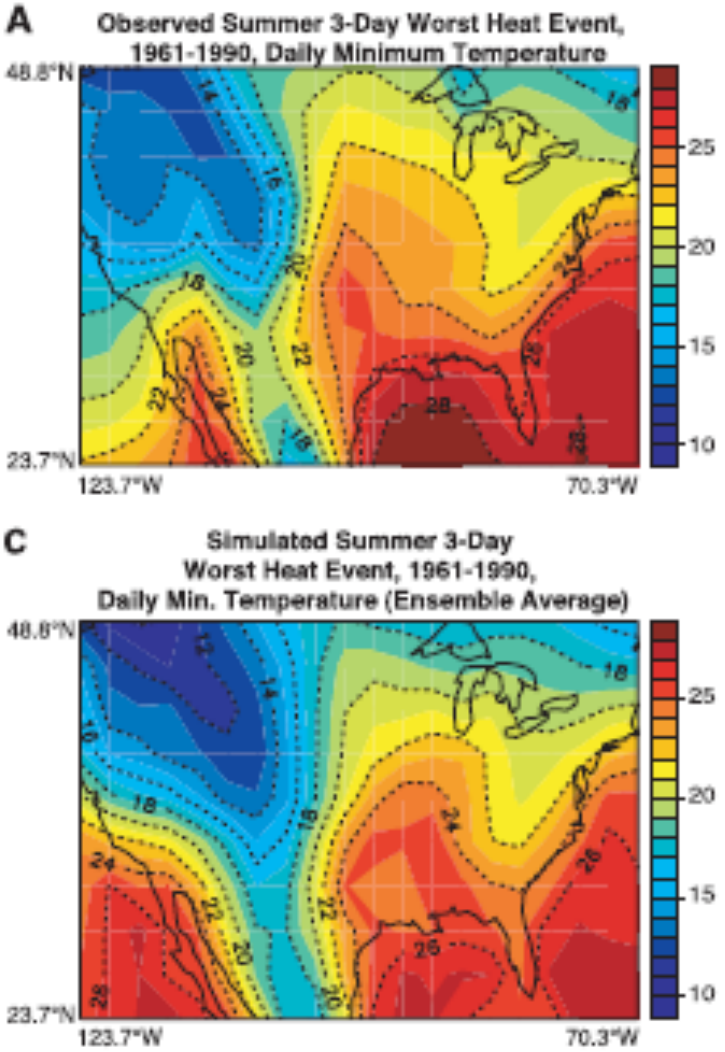
David R. Easterling,[1]* Gerald A. Meehl,[2] Camille Parmesan,[3] Stanley A. Changnon,[4] Thomas R. Karl,[1] Linda O. Mearns[2]

One of the major concerns with a potential change in climate is that an increase in extreme events will occur. Results of observational studies suggest that in many areas that have been analyzed, changes in total precipitation are amplified at the tails, and changes in some temperature extremes have been observed. Model output has been analyzed that shows changes in extreme events for future climates, such as increases in extreme high temperatures, decreases in extreme low temperatures, and increases in intense precipitation events. In addition, the societal infrastructure is becoming more sensitive to weather and climate extremes, which would be exacerbated by climate change. In wild plants and animals, climate-induced extinctions, distributional and phenological changes, and species' range shifts are being documented at an increasing rate. Several apparently gradual biological changes are linked to responses to extreme weather and climate events.

the 20th century (*2*), and that this increase is associated with a stronger warming in daily minimum temperatures than in maximums, leading to a reduction in the diurnal temperature range (*3*). Land surface precipitation has also increased over the same period in the mid- to high latitudes, but shows a decrease in the tropics and subtropics (*2*). Given these changes, it is expected that there would also be changes in what are now considered extreme events (*4*). Therefore, if there are indeed identifiable trends in certain extreme climatic events, such as extremes in temperature or precipitation, it would add to the

Motivated by that session, new papers started to be published with analyses of various types of extremes from climate models

For example, heat waves…
(Meehl and Tebaldi, 2004;
PCM, T42)



…and frost days:

(Meehl, Tebaldi and Nychka, 2004; PCM)

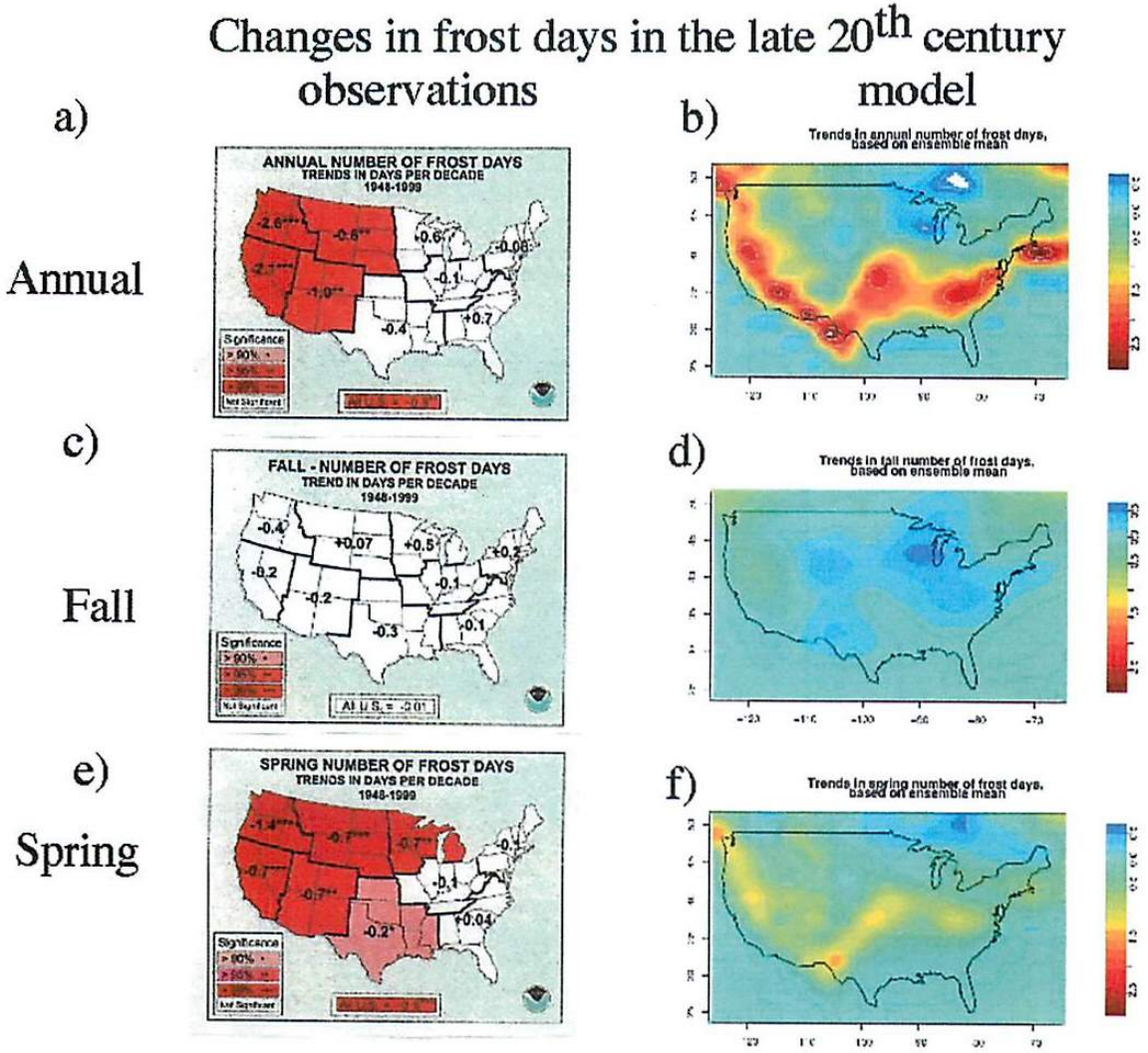


Fig. 1 Trends (days per decade) of changes in frost days for the USA for the period 1948–1999, **a** area-averaged annual observations from Easterling (2002); **b** four member ensemble annual mean from model simulation; **c** same as **a** except for fall (September–October–November) from observations; **d** same as **b** except for fall from model; **e** same as **a** except for spring (March–April–May) from observations; **f** same as **b** except for spring from model

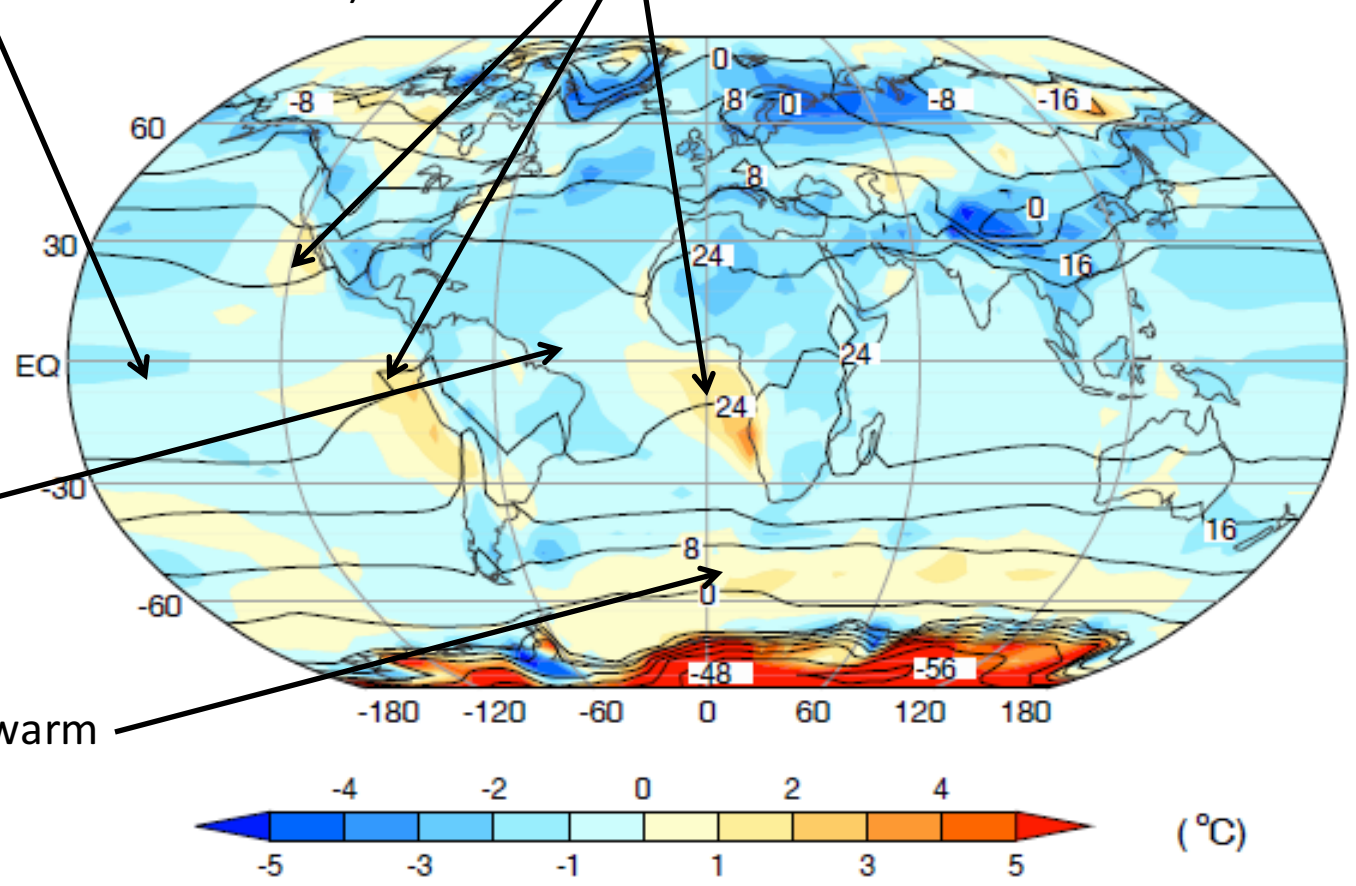
2007: IPCC Fourth Assessment Report

23 CMIP3 models; 1° to 2° resolution

Surface temperature errors similar to previous generations

Western tropical Pacific too cold (eastern Pacific cold tongue extends too far west)

Too warm in west coast stratus regimes

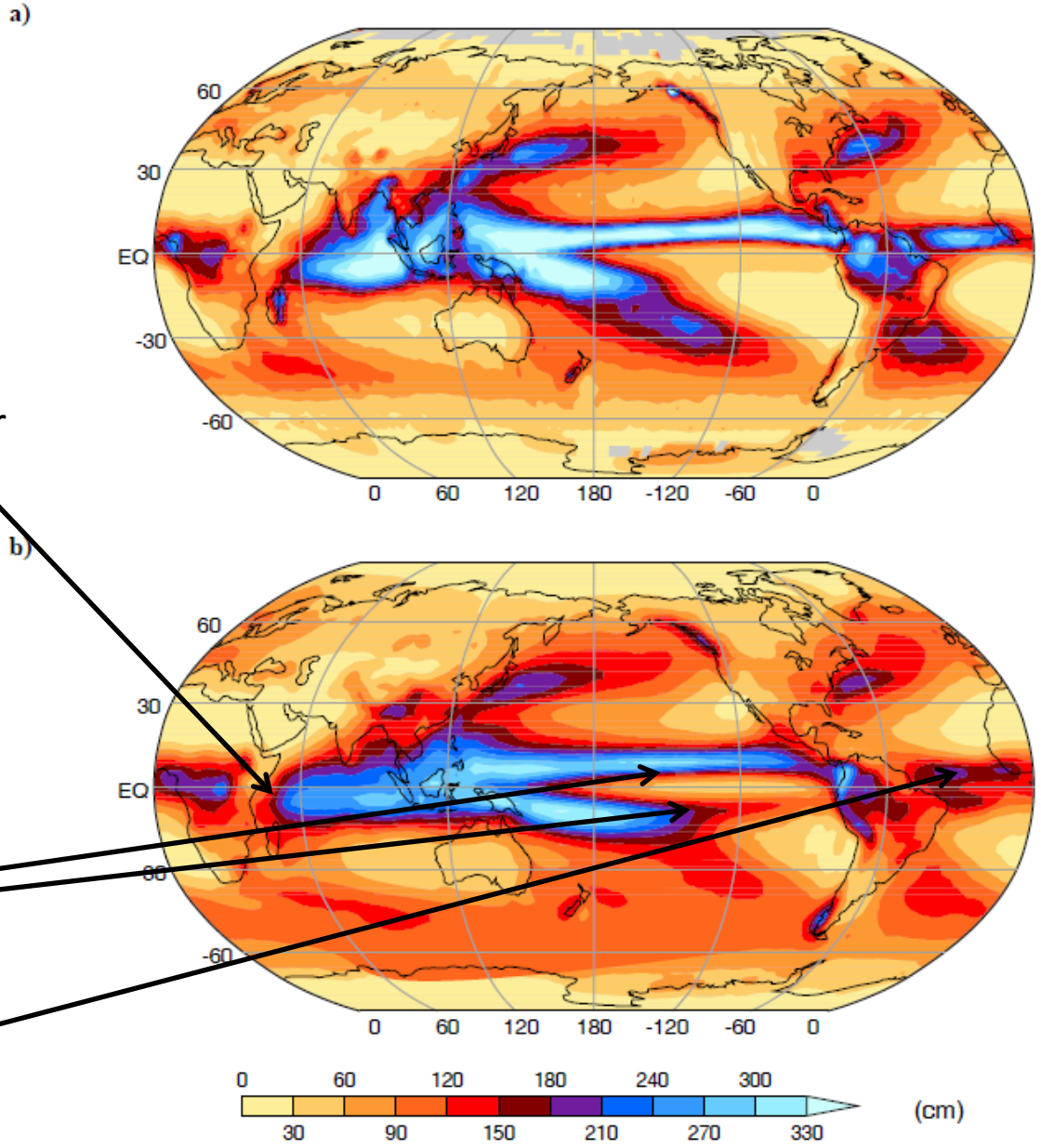


Western equatorial Atlantic too cold

Southern Ocean too warm

2007: IPCC Fourth Assessment Report

23 CMIP3 models; 1° to 2° resolution



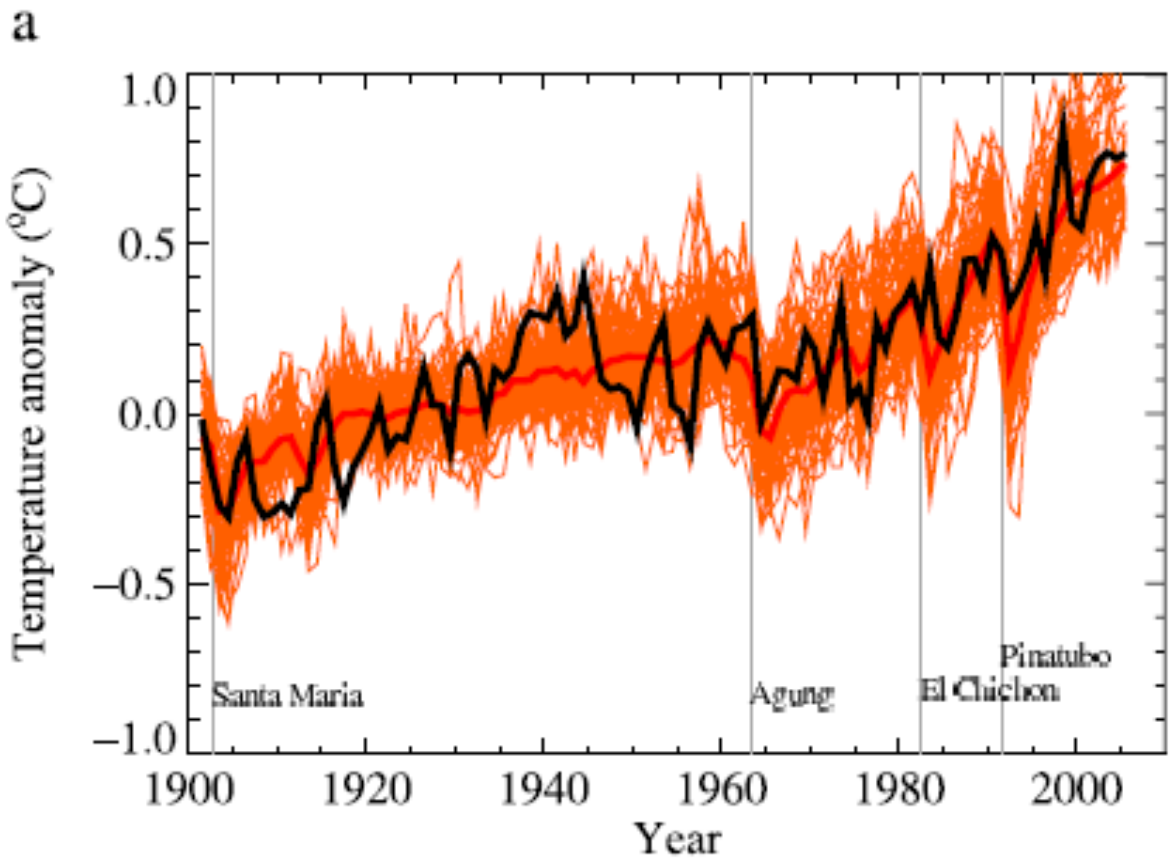
Too much rainfall too far west in tropical Indian Ocean

Double ITCZ

Too little rainfall in western equatorial Atlantic

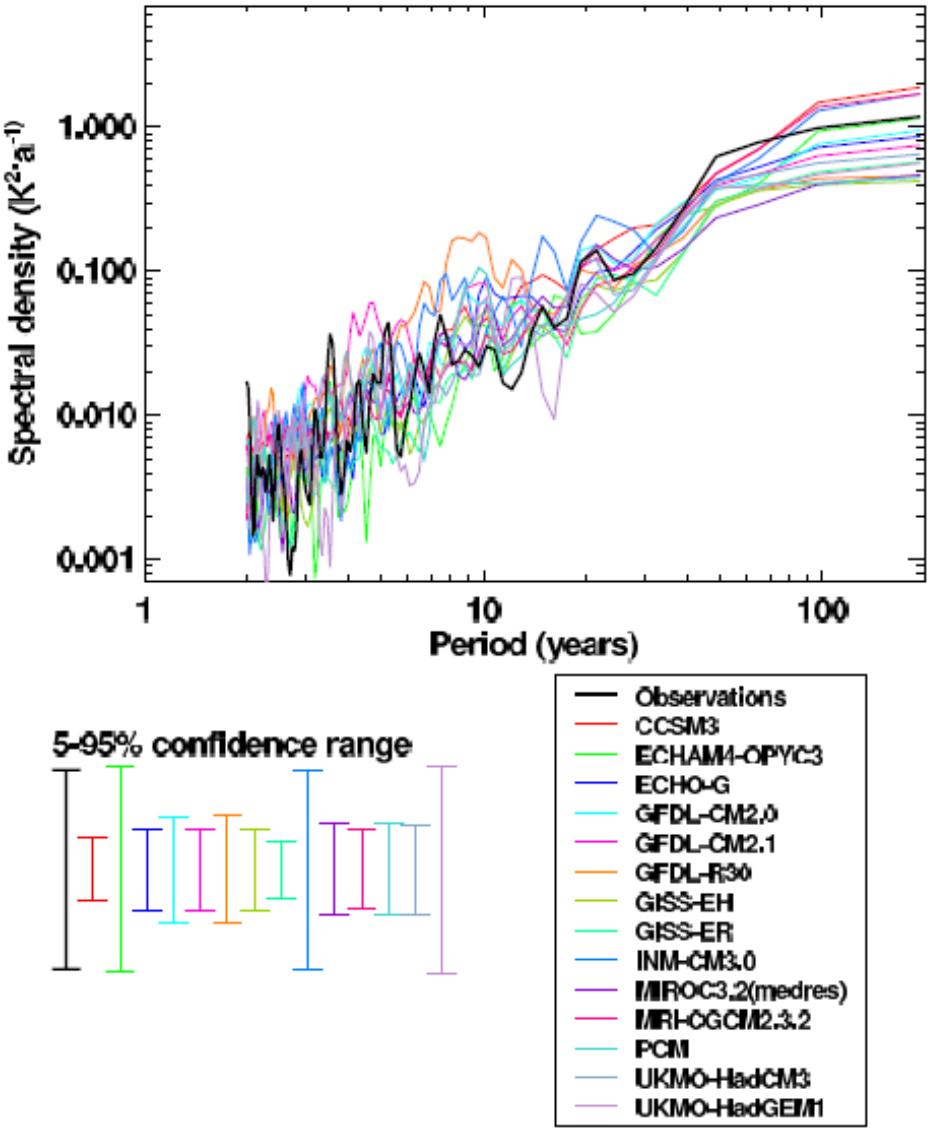
In the AR4 there was a multi-model ensemble of 20th century simulations to compare to observations

IPCC AR4 2007
23 CMIP5 models



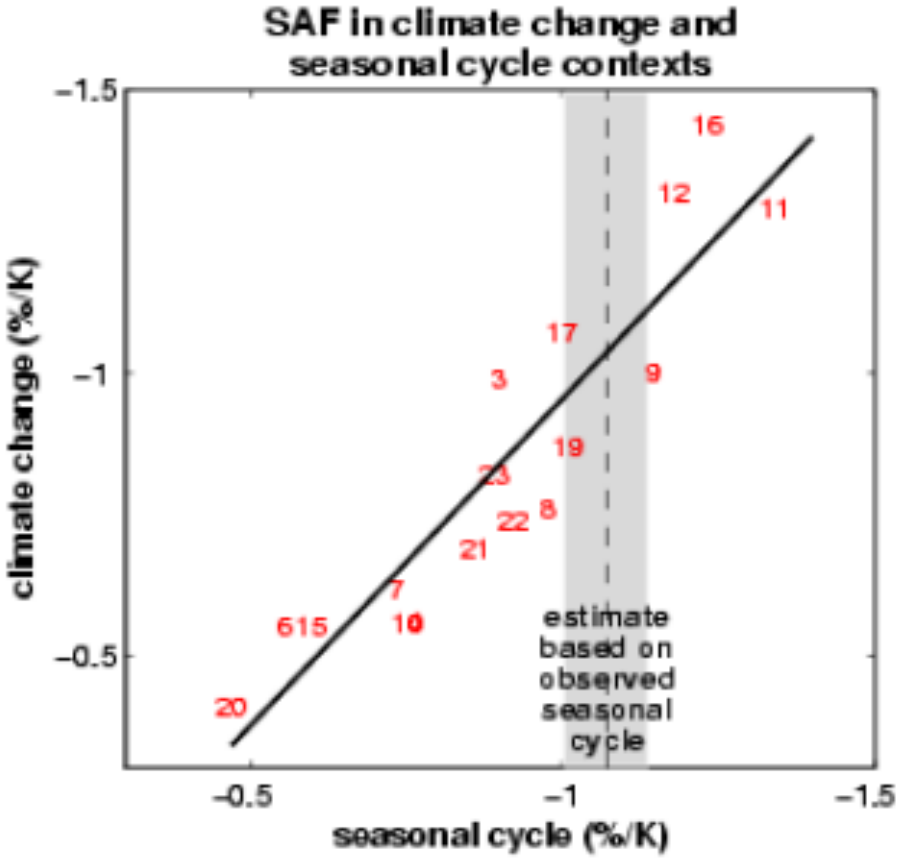
IPCC AR4 2007: 14 CMIP3 models with multi-century pre-industrial control runs to evaluate interannual to century timescale variability

"…the models have variance at global scales that is consistent with the observed variance at the 5% significance level on the decadal to inter-decadal time scales"



The concept of "emergent constraints" was assessed in the IPCC AR4 (from Hall and Qu, 2006), a new way to evaluate climate model simulations compared to observations with implications for feedbacks in the system

Seasonal cycle of snow albedo in observations is a good proxy for snow-albedo feedback (SAF)



IPCC AR5, 2013
1 degree class coupled models, CMIP5

Table 9.4: Summary of assessment of interannual to interdecadal variability in climate models. See also Figure 9.44.

	Short Name	Level of Confidence	Level of Evidence for Evaluation	Degree of Agreement	Model Quality	Difference with AR4 (including CMIP5 vs. CMIP3)	Section
Global SST variability	SST-var	*High*	*Robust*	*Medium*	Medium	Slight improvement in the tropics	9.5.3.1
North Atlantic Oscillation and Northern Annular Mode	NAO	*Medium*	*Medium*	*Medium*	High	No assessment	9.5.3.2
Southern Annular Mode	SAM	*Low*	*Limited*	*Medium*	Medium	No assessment	9.5.3.2
Atlantic Meridional Overturning Circulation Variability	AMOC-var	*Low*	*Limited*	*Medium*	Medium	No improvement	9.5.3.3
Atlantic Multi-decadal Variability	AMO	*Low*	*Limited*	*Medium*	Medium	No improvement	9.5.3.3
Atlantic Meridional Mode	AMM	*High*	*Medium*	*High*	Low	No assessment	9.5.3.3
Atlantic Niño	AN	*Low*	*Limited*	*Medium*	Low	Slight improvement	9.5.3.3
El Niño Southern Oscillation	ENSO	*High*	*Medium*	*High*	Medium	Slight improvement	9.5.3.4
Indian Ocean Basin mode	IOB	*Medium*	*Medium*	*Medium*	High	Slight improvement	9.5.3.4
Indian Ocean Dipole	IOD	*Medium*	*Medium*	*Medium*	Medium	No improvement	9.5.3.4
Pacific North American	PNA	*High*	*Medium*	*High*	Medium	Slight improvement	9.5.3.5
Tropical ENSO teleconnections	ENSOtele	*High*	*Robust*	*Medium*	Medium	No improvement	9.5.3.5
Pacific Decadal Oscillation	PDO	*Low*	*Limited*	*Medium*	Medium	No assessment	9.5.3.6
Interdecadal Pacific Oscillation	IPO	*Low*	*Limited*	*Medium*	High	No assessment	9.5.3.6
Quasi-Biennial Oscillation	QBO	*Medium*	*Medium*	*Medium*	High	No assessment	9.5.3.7

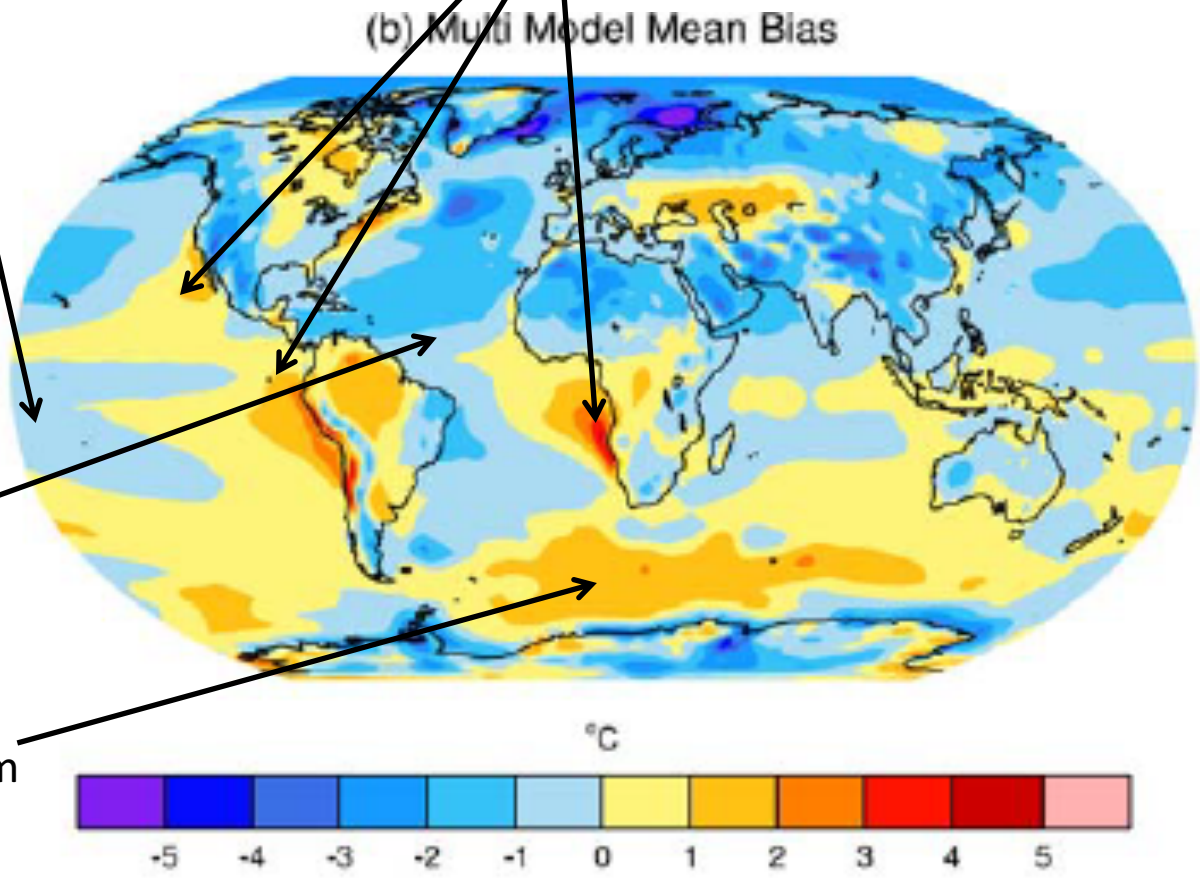
2013: IPCC Fifth Assessment Report

60 CMIP5 models; 1° resolution

Surface temperature errors still look familiar

Western tropical Pacific too cold (eastern Pacific cold tongue extends too far west)

Too warm in west coast stratus regimes



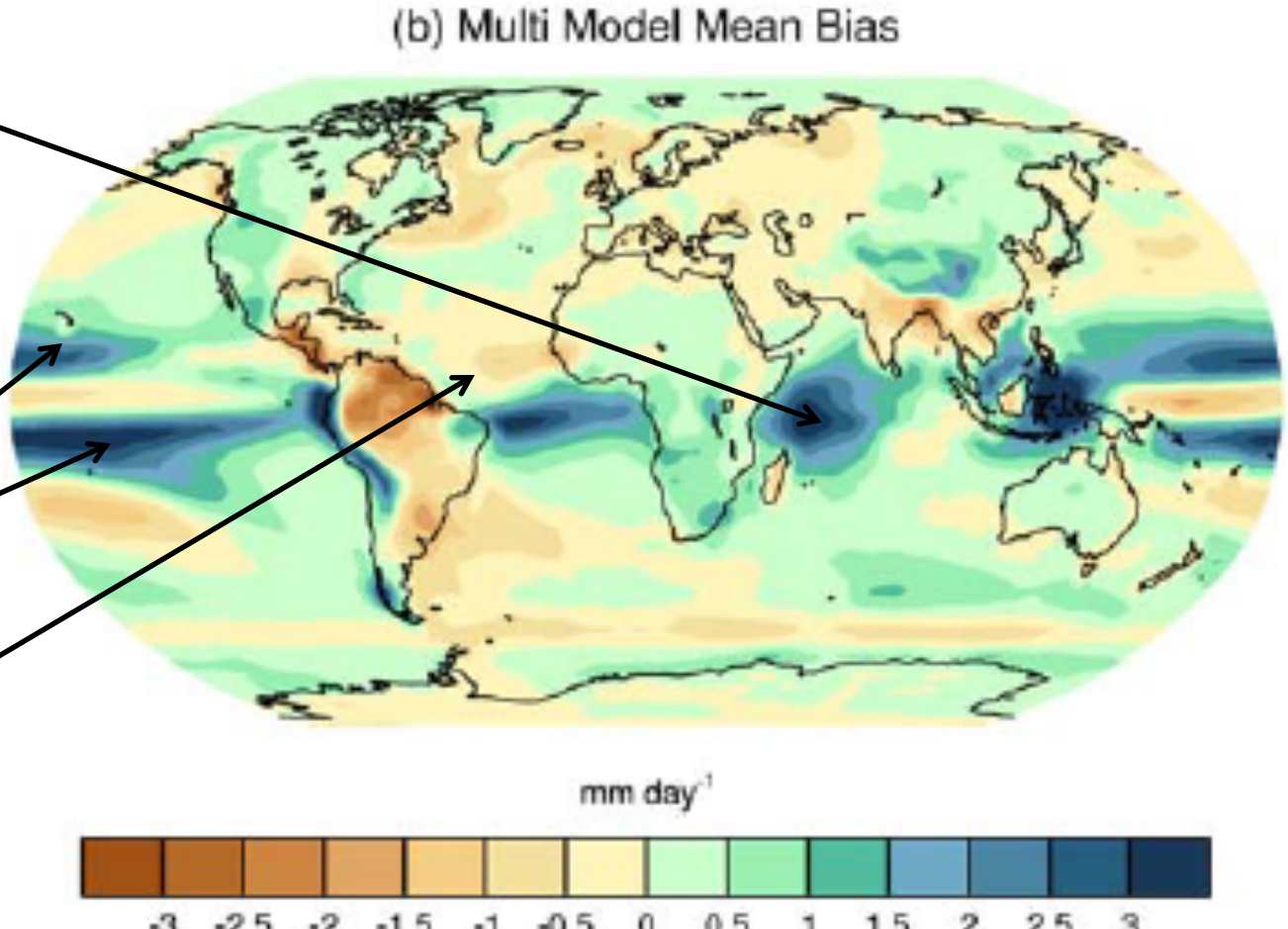
Western equatorial Atlantic too cold

Southern Ocean too warm

2013: IPCC Fifth Assessment Report

60 CMIP5 models; 1° resolution

Precipitation errors still look familiar



Too much rainfall too far west in tropical Indian Ocean

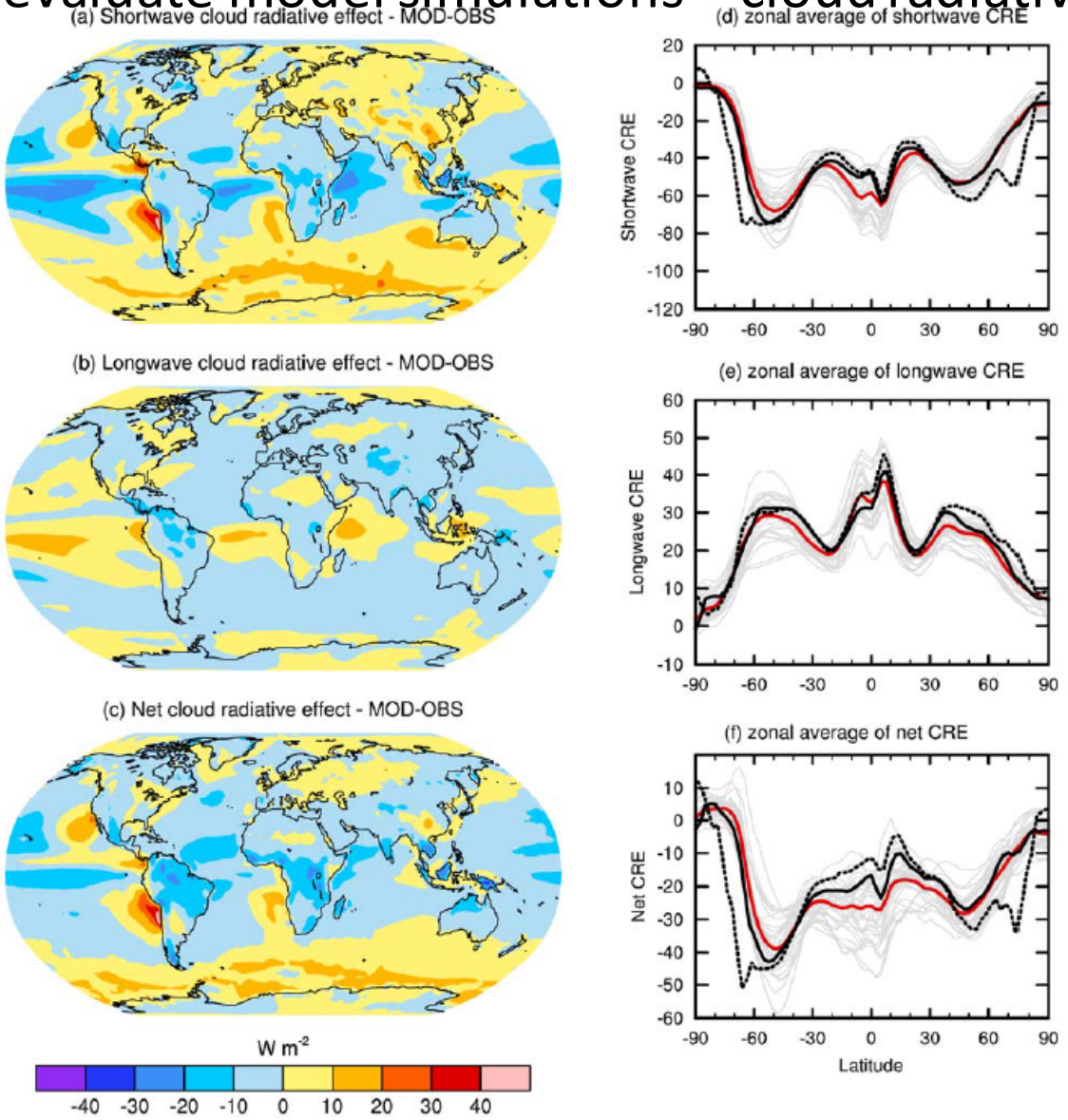
Double ITCZ

Too little rainfall in western equatorial Atlantic

2013: IPCC Fifth Assessment Report

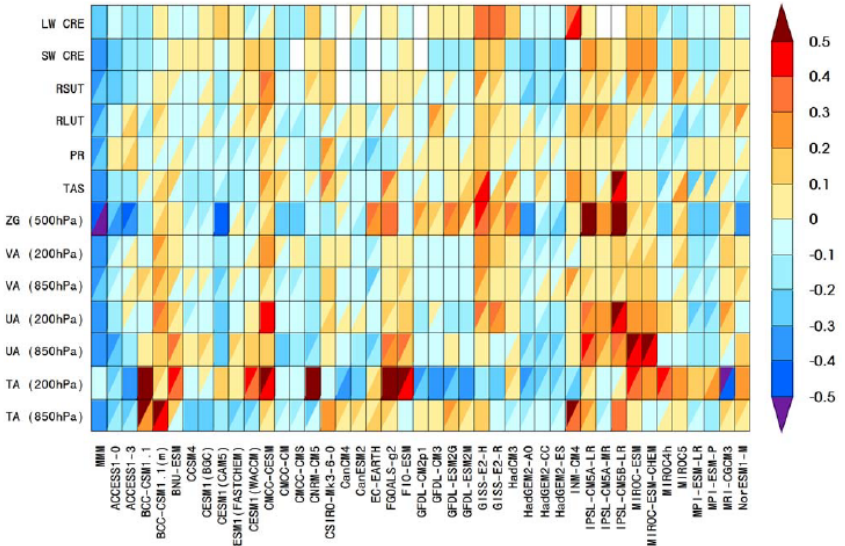
60 CMIP5 models; 1° resolution

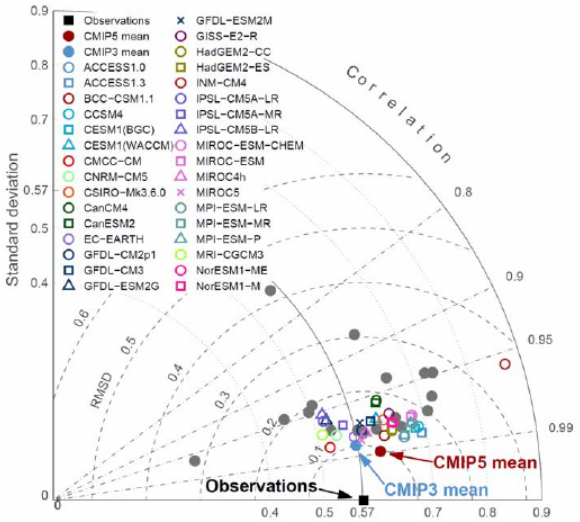
New ways to evaluate model simulations—cloud radiative forcing



Climate model evaluation today: multiple global earth system models run for hundreds of years, a move towards standardized diagnostics of many fields to compare to multiple observational data sets

The challenge: reduce model systematic errors (some have been in the models since the early days of climate modeling) by improving the models: **increase knowledge of observed processes, more and better observations and relevant model evaluation, represent those processes in models**





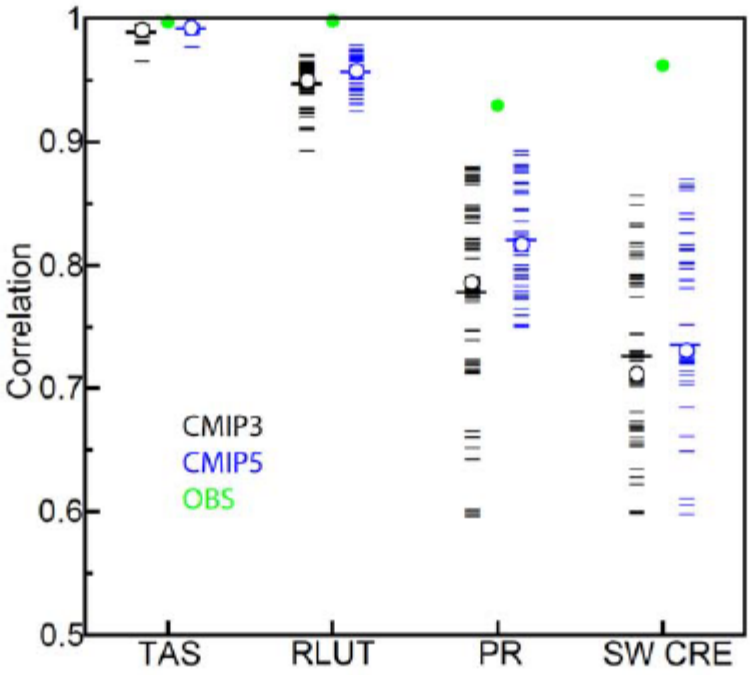


Figure 9.6: Centred pattern correlations between models and observations for the annual-mean climatology over the period 1980–1999. Results are shown for individual CMIP3 (black) and CMIP5 (blue) models as thin dashes, along with the corresponding ensemble average (thick dash) and median (open circle). The four variables shown are surface air temperature (TAS), top-of-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 realisation is used from each model from the CMIP3 20C3M and CMIP5 historical simulations.

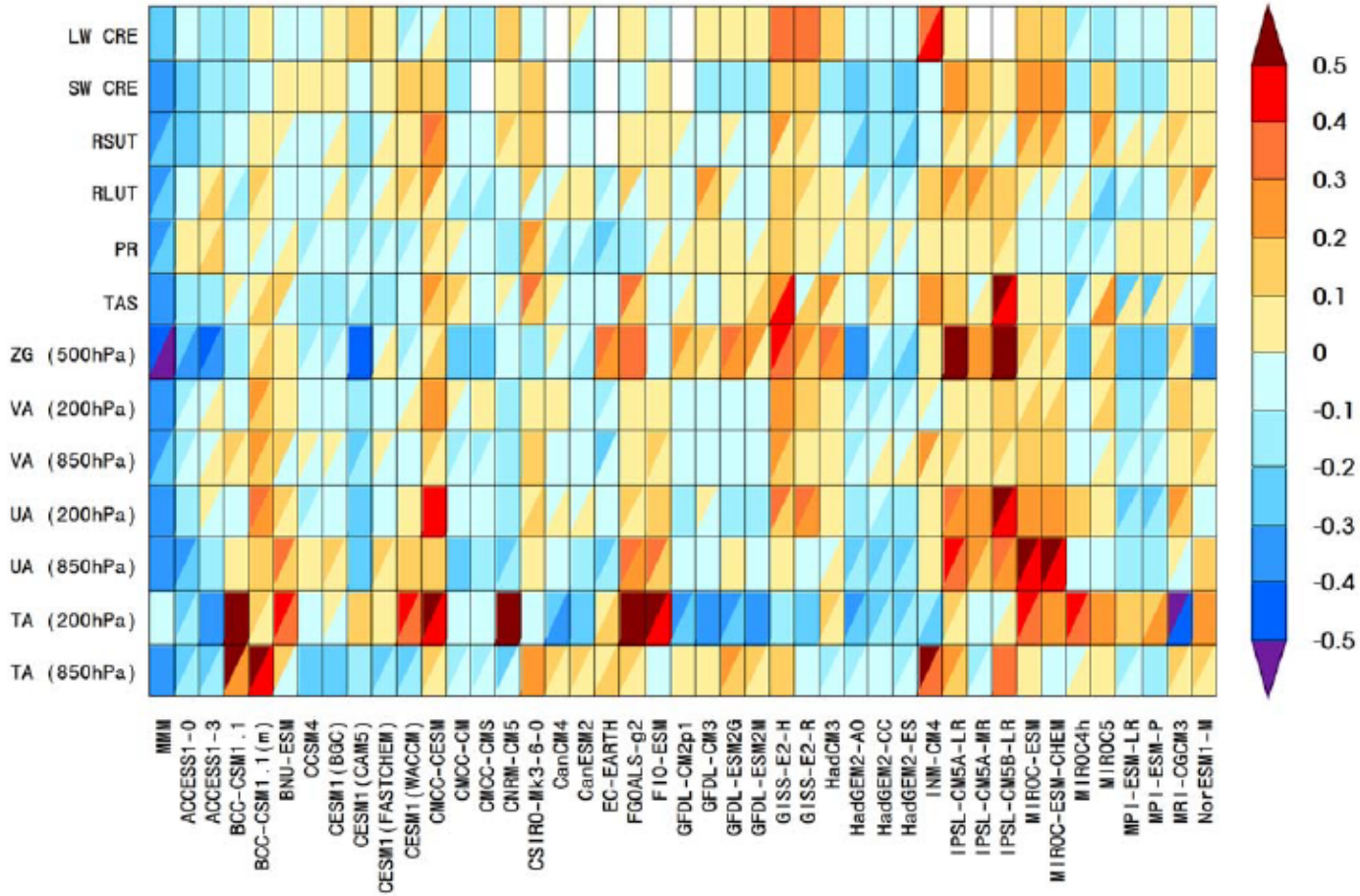


Figure 9.7: Relative error measures of CMIP5 model performance, based on the global seasonal-cycle climatology (1980–2005) computed from the historical experiments. Rows and columns represent individual variables and models, respectively. The error measure is a space–time root-mean-square error (RMSE), which, treating each variable separately, is portrayed as a relative error by normalizing the result by the median error of all model results (P. Gleckler, Taylor, & Doutriaux, 2008). For example, a value of 0.20 indicates that a model's RMSE is 20% larger than the median CMIP5 error for that variable, whereas a value of –0.20 means the error is 20% smaller than the median error. No colour (white) indicates that model results are currently unavailable. A diagonal split of a grid square shows the relative error with respect to both the default reference data set (upper left triangle) and the alternate (lower right triangle). The relative errors are calculated independently for the default and alternate data sets. All reference data used in the diagram are summarized in Table 9.3.

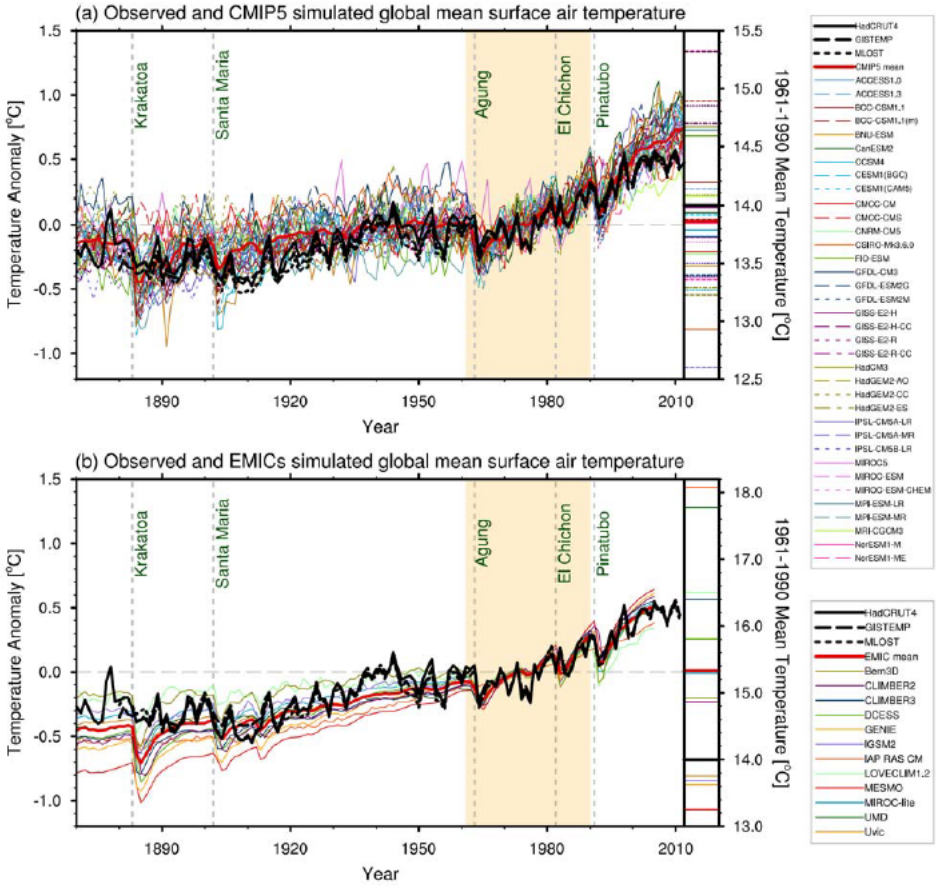


Figure 9.8: Observed and simulated time series of the anomalies in annual- and global-mean surface temperature. All anomalies are differences from the 1961–1990 time-mean of each individual time series. The reference period 1961–1990 is indicated by yellow shading; vertical dashed grey lines represent times of major volcanic eruptions. (a) Single simulations for CMIP5 models (thin lines); multi-model mean (thick red line); different observations (thick black lines). Observational data (see Chapter 2) are HadCRUT4 (Morice, Kennedy, Rayner, & Jones, 2012), GISTEMP (Hansen, Ruedy, Sato, & Lo, 2010), and MLOST (Vose et al., 2012) and are merged surface temperature (2 m height over land and surface temperature over the ocean). All model results have been sub-sampled using the HadCRUT4 observational data mask (see Chapter 10). Following the CMIP5 protocol (Taylor et al., 2012), all simulations use specified historical forcings up to and including 2005 and use RCP4.5 after 2005 (see Figure 10.1 and note different reference period used there; results will differ slightly when using alternative RCP scenarios for the post-2005 period). (a) Inset: the global-mean surface temperature for the reference period 1961–1990, for each individual model (colours), the CMIP5 multi-model mean (thick red), and the observations (thick black, P. D. Jones, New, Parker, Martin, and Rigor (1999)). Bottom: single simulations from available EMIC simulations (thin lines), from Eby et al. (2013). Observational data are the same as in (a). All EMIC simulations ended in 2005 and use the CMIP5 historical forcing scenario. (b) Inset: Same as in (a) but for the EMICs.

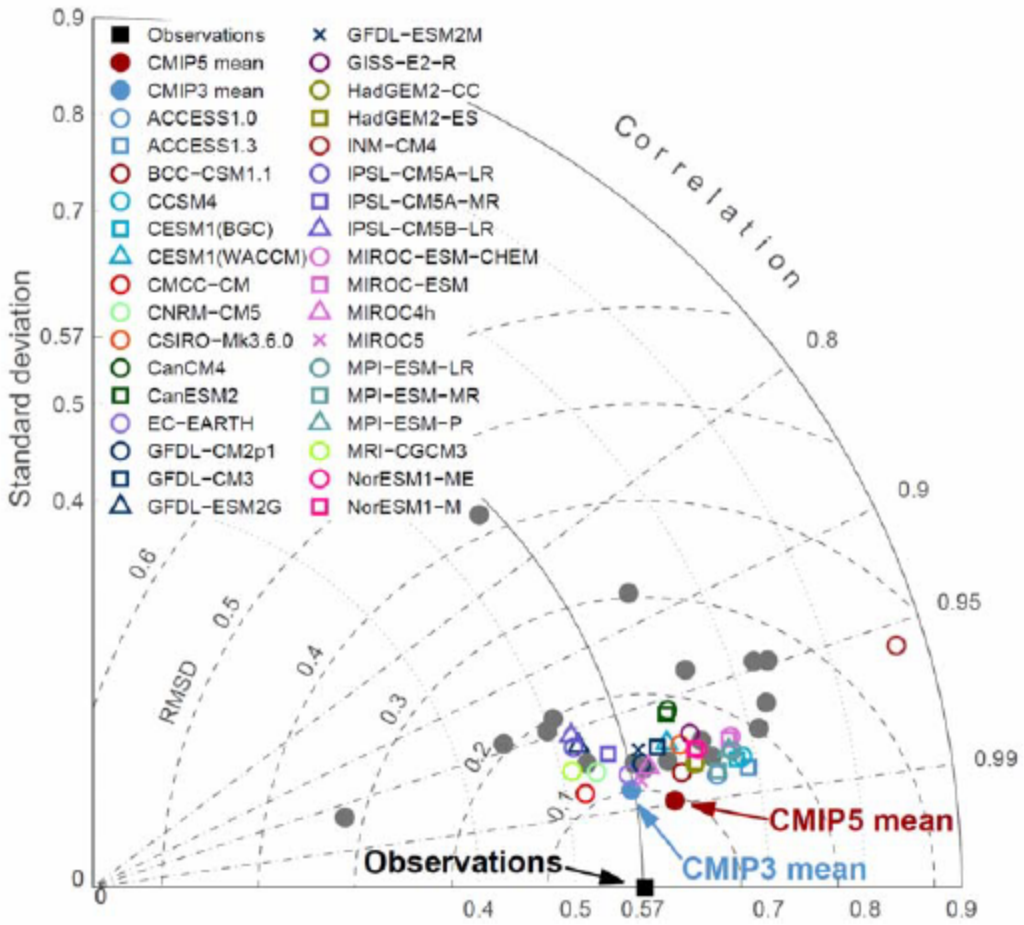
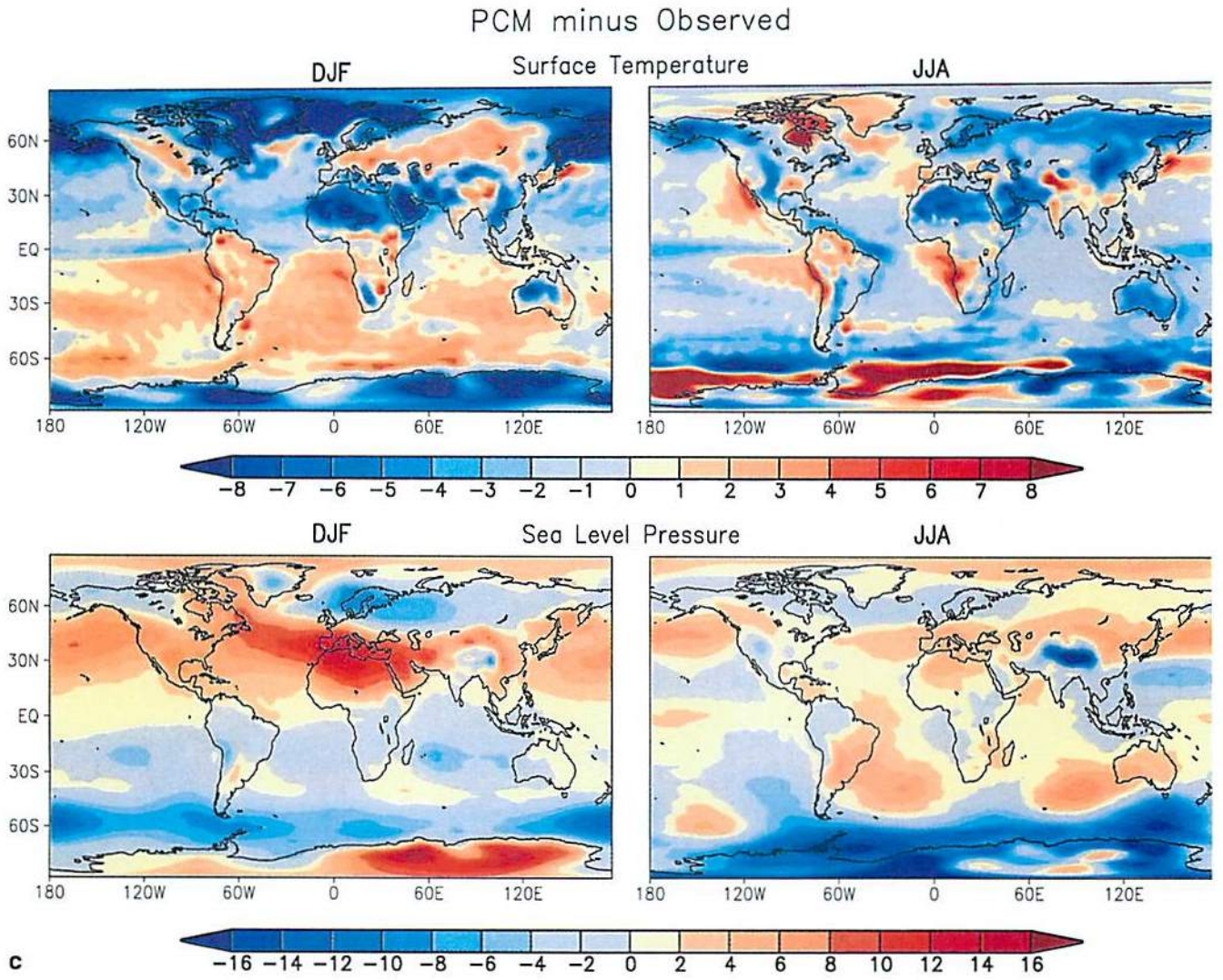


Figure 9.16: Taylor diagram for the dynamic sea-surface height climatology (1987–2000). The radial coordinate shows the standard deviation of the spatial pattern, normalised by the observed standard deviation. The azimuthal variable shows the correlation of the modelled spatial pattern with the observed spatial pattern. The root-mean square error with bias removed is indicated by the dashed grey circles about the observational point. Analysis is for the global ocean, 50°S–50°N. The reference dataset is AVISO, a merged satellite product (Ducet, Le Traon, & Reverdin, 2000), which is described in Chapter 3. One realisation per model is shown for each CMIP5 and CMIP3 model result. Grey filled circles are for individual CMIP3 models; other symbols as in legend.

The next generation of coupled model emerges in the late 1990s, and familiar systematic errors are still there (CMIP1/2 generation of models; IPCC TAR)
True model errors manifested in models without flux correction

PCM: T42 (~2.8 atmosphere, ~1 ocean)



1995 SAR, 9 models, pre-CMIP
R15 class (4.5° latitude x 7.5 ° longitude)
Some flux corrected, some not

AMIP models

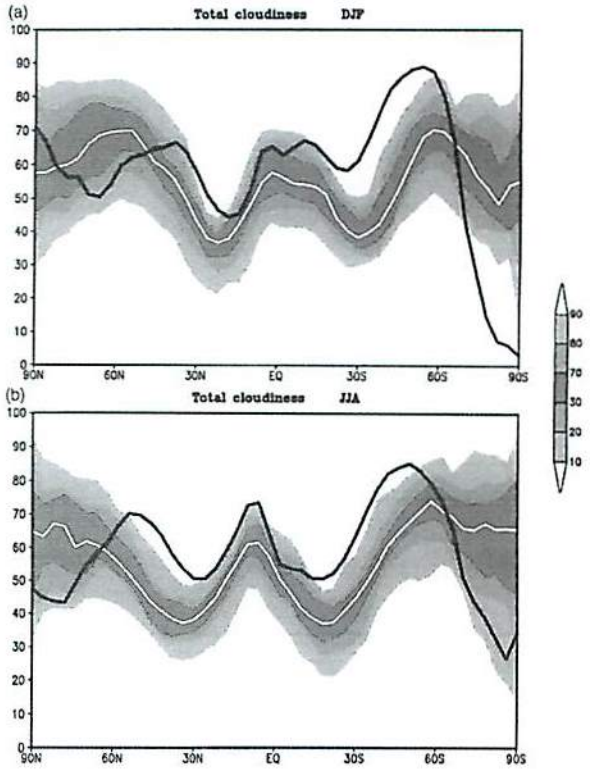


Figure 5.10: As in Figure 5.9 except for total cloudiness (%). The observational estimates are from ISCCP data for 1983 to 1990 (Rossow and Schiffer, 1991).

IPCC SAR, 1995

Snow and sea ice from models compared to observations

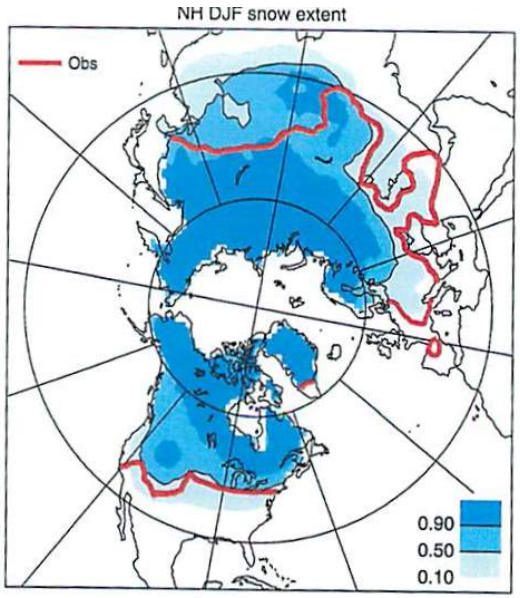


Figure 8.11: Illustration of the range of snow cover extent in CMIP1 model simulations listed in Table 8.3: Northern Hemisphere, DJF. The figure is constructed similarly to Figure 8.10 based on the prescribed 1 cm cutoff. The observed boundary is based on Foster and Davy (1988).

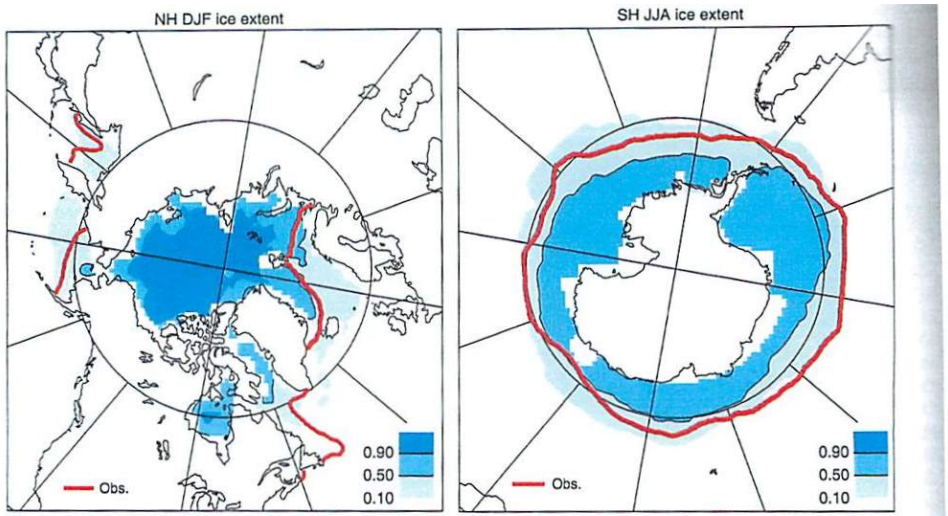


Figure 8.10: Illustration of the range of sea-ice extent in CMIP1 model simulations listed in Table 8.3: Northern Hemisphere, DJF (left) and Southern Hemisphere, JJA (right). For each model listed in Table 8.3, a 1/0 mask is produced to indicate presence or absence of ice. The fifteen masks were averaged for each hemisphere and season. The 0.5 contour therefore delineates the region for which at least half of the models produced sea ice. The 0.1 contour indicates the region outside of which only 10% of models produced ice, while the 0.9 contour indicates that region inside of which only 10% of models did not produce ice. The observed boundaries are based on GISST_2.2 (Rayner *et al.*, 1996) averaged over 1961 to 1990.

By the late-1980s, advances in computer power allowed performing full seasonal cycle simulations for tens of years, and spectral atmospheric models became popular (R15 coupled model, Washington and Meehl, 1989)

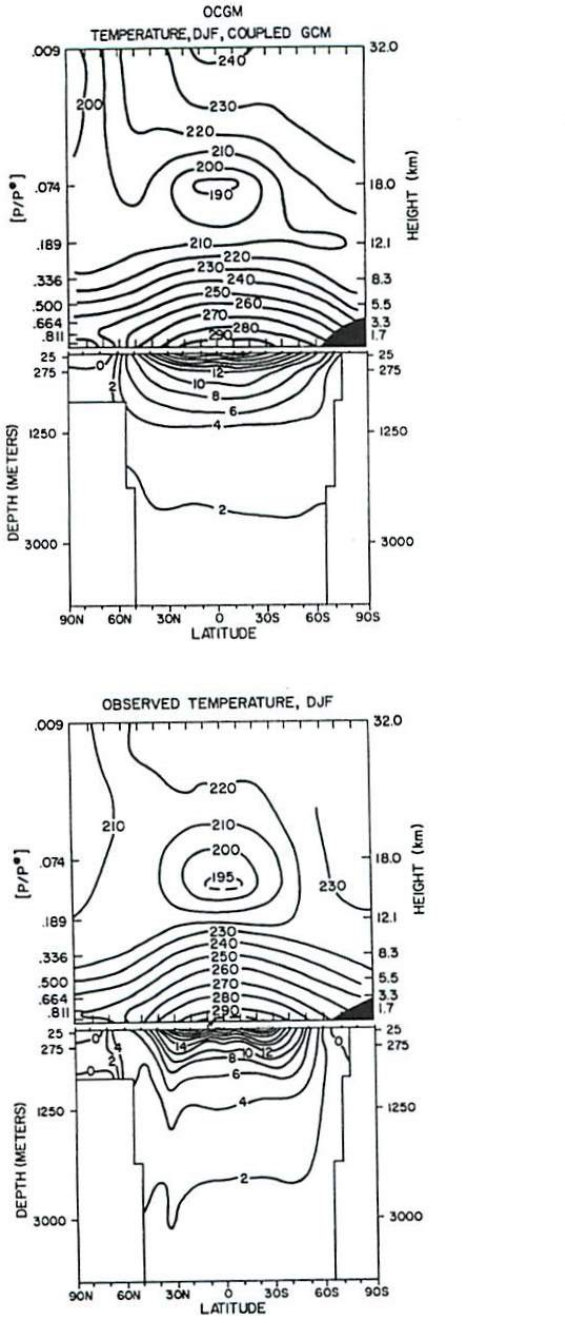


Fig. 5. Zonal mean temperatures for atmosphere and ocean for NCAR coupled model (*top*) and observations (*bottom*) (after Washington and Meehl 1989)