



LAWRENCE  
LIVERMORE  
NATIONAL  
LABORATORY

UCRL-TR-202550

# An Appraisal of Coupled Climate Model Simulations

*K. AchutaRao, C. Covey, C. Doutriaux, M. Fiorino,  
P. Gleckler, T. Phillips, K. Sperber, K. Taylor*

*Edited by D. Bader*

**August 16, 2004**

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

# Contents

<b>1</b>	<b>Introduction.....</b>	<b>1</b>
<b>2</b>	<b>Features of Appraisal Models .....</b>	<b>3</b>
<b>3</b>	<b>Climate Trends in Unforced and Forced Simulations.....</b>	<b>11</b>
<b>3.1</b>	<b>Control Run Secular Trends .....</b>	<b>11</b>
3.1.1	Annual and Global Mean Surface Air Temperature Power Spectra.....	11
3.1.2	Annual and Global Mean Surface Air Temperature Trends.....	12
3.1.3	Annual Average Global Maps of SST Trends .....	12
3.1.4	Annual and Global Mean Oceanic Temperature, Salinity and Sea-Ice Extent Trends.	20
<b>3.2</b>	<b>CMIP 1%/yr runs.....</b>	<b>28</b>
3.2.1	Summary .....	31
<b>4</b>	<b>Atmospheric Climatology and Annual Cycle .....</b>	<b>33</b>
<b>4.1</b>	<b>Overview .....</b>	<b>33</b>
<b>4.2</b>	<b>Zonal Mean Vertical Cross Sections.....</b>	<b>36</b>
4.2.1	Air Temperature .....	36
4.2.2	Zonal Wind.....	36
4.2.3	Specific Humidity .....	36
<b>4.3</b>	<b>Zonal Means, Statistical Plots and Maps.....</b>	<b>40</b>
4.3.1	Surface Air Temperature .....	40
4.3.2	Precipitation .....	41
4.3.3	Outgoing Longwave Radiation .....	42
4.3.4	Total Cloud Cover .....	43
4.3.5	Mean Sea-level Pressure .....	43
<b>4.4</b>	<b>The Amplitude and Phase of the Seasonal Cycle.....</b>	<b>54</b>
4.4.1	Introduction .....	54
4.4.2	Global Characteristics: T(2m), Precipitation and OLR .....	55
4.4.3	North America .....	59
4.4.4	South America, the Atlantic and Western Africa .....	63
4.4.5	Indo-Pacific .....	66
4.4.6	Europe/Middle East/ Northern Africa .....	69
<b>4.5</b>	<b>Tropical Biases in Rainfall, SST, and Windstress .....</b>	<b>72</b>
4.5.1	Introduction .....	72
4.5.2	The Models and Validation Data.....	72
4.5.3	The Biases .....	72
<b>4.6</b>	<b>The Köppen Climate Classification .....</b>	<b>78</b>
4.6.1	Introduction .....	78
4.6.2	The NRL Köppen Classification scheme .....	78
4.6.3	KC of the CMIP2+ Mean Model.....	79
4.6.4	The Transfer Matrix .....	80
4.6.5	BE—Big Error.....	81

4.6.6 Major Class Distribution Error .....	82
4.6.7 Summary of KC results .....	83
<b>5 The Mean State of the Ocean and Cryosphere .....</b>	<b>93</b>
5.1 Introduction .....	93
5.2 The Atlantic Ocean.....	97
5.3 The Pacific Ocean.....	102
5.4 The Indian Ocean .....	107
5.5 The Arctic Ocean.....	111
5.6 The Southern Ocean.....	115
5.7 Sea Ice.....	119
<b>6 Simulated Variability.....</b>	<b>121</b>
6.1 El Niño/Southern Oscillation.....	121
6.1.1 Introduction .....	121
6.1.2 The Models.....	121
6.1.3 Seasonal Cycle Phase Locking and Power Spectra .....	121
6.1.4 SOI and NIÑO3 Teleconnections.....	122
6.1.5 Summary .....	123
6.2 The North Atlantic Oscillation .....	127
6.2.1 Introduction .....	127
6.2.2 Annual Mean Climatology and Variance .....	127
6.2.3 NAO Teleconnections .....	128
6.2.4 Summary .....	130
6.3 Tropical Wavenumber–Frequency Spectra .....	139
6.4 The Madden–Julian Oscillation .....	146
6.4.1 Introduction .....	146
6.4.2 The Models and Validation Data.....	146
6.4.3 OLR and the MJO .....	147
6.4.4 MJO Convection: CMIP2+ vs. AMIPII .....	149
6.4.5 Summary .....	150
<b>7 Appraisal Summary.....</b>	<b>159</b>
<b>Acknowledgments .....</b>	<b>163</b>
<b>References.....</b>	<b>165</b>
<b>Appendix 1: Observationally Based Resources.....</b>	<b>177</b>
<b>Appendix 2: Interpretation of the Statistical Taylor Diagram.....</b>	<b>179</b>
<b>Appendix 3: Acronyms.....</b>	<b>183</b>

## List of Figures

Fig. 3.1. Power spectra of detrended globally and annually averaged surface air temperature simulated by CMIP2+ models and observed by Jones / IPCC. The curves are scaled so that the areas under them (if plotted linearly) equal the total variances about the mean of the detrended time series. The 95% confidence interval—based only on uncertainties due to finite sample size—is the same for all cases with the logarithmic scales used in this graphic. The models exhibit global variability that is typically less than observed, as expected for control run simulations.....	14
Fig. 3.2. Time series of annual mean surface air temperature from the last 200 years of CMIP2+ control runs integrations and observed by Jones / IPCC. Blue curves are Northern Hemisphere means, pink curves are Southern Hemisphere means, and black curves are global means. Black straight lines are linear least-square fits to global means. The models typically exhibit smaller trends than observed, as expected for control run simulations.....	15
Fig. 3.3. Sea surface temperature trends observed for the period 1898-1997 (HadSST database) and simulated by two of the three CMIP2+ models that are not flux-adjusted. Results of globally averaging these trends are given in Table 3.1 for the results shown in this figure and for all other CMIP2+ models. All the models exhibit trends that are considerably smaller than observed, as expected for control run simulations.....	19
Fig. 3.4. Average salinity over total ocean volume for the three CMIP2+ models that were not flux-adjusted.....	22
Fig. 3.5. Deviation of the globally and annually averaged ocean temperature from the first year (of data archived at PCMDI) as a function of depth (degrees C). The three models that are most suggestive of trends (BCM02, CCCma and PCM) are found to approach quasi-equilibrium later in the integrations (not shown). In each case, years 60–79 are found to be periods of relative stability, and are therefore used to derive climatologies for the mean climate Sections 4 (atmosphere) and 5 (oceans). .....	23
Fig. 3.6. Deviation of the Arctic Basin annual mean temperature from the first year (of data archived at PCMDI) as a function of depth (degrees C).....	24
Fig. 3.7. Deviation of the globally and annually averaged ocean salinity from the first year (of data archived at PCMDI) as a function of depth (ppt) .....	25
Fig. 3.8. Deviation of the Arctic Basin annual mean ocean salinity (ppt) from the first year (of data archived at PCMDI) as a function of depth (m) .....	26
Fig. 3.9. Northern and Southern Hemisphere 5-year running means of the percentage changes in the total sea ice surface area with respect to the mean of the first 5 years.....	27
Fig. 3.10. Differences between increasing atmospheric carbon dioxide scenarios and model control runs in the global and annual mean, for surface air temperature in K (top) and precipitation in mm/day (bottom), as a function of time in years.....	29
Fig. 3.11. Differences near the time of CO <sub>2</sub> doubling between increasing atmospheric carbon dioxide scenarios and model control runs, averaged over all models, for the December–January–February season (top row) and the June–July–August season (bottom row). The left-hand column shows surface air temperature differences in K; the right-hand column shows precipitation differences in mm/day.....	30

Fig. 4.1. Taylor diagrams (Appendix 2) showing differences in statistical measures of model agreement with observed climatology (Appendix 1) for several different fields. Comparisons are shown for the CMIP mean and median models (upper panel) and the median AMIP and CMIP models (lower panel). The statistics are based on sums computed over the four climatological mean seasons and over all grid cells, weighted by grid cell area. Note that for each field, the standard deviation and root-mean-square error statistics have been normalized by the respective observed standard deviation of the field. The following fields were considered: surface air temperature (TAS), precipitation (P), outgoing longwave radiation (OLR), reflected shortwave radiation (SW), longwave cloud radiative forcing ( $LW_{CRF}$ ), reflected shortwave clear-sky flux ( $SW_{clr}$ ), surface sensible heat flux over oceans (SH), surface latent heat flux over oceans (LH), eastward and northward components of surface wind stress over oceans (TAUU, TAUUV), zonal and meridional wind components at 200 hPa ( $U_{200}$ , $V_{200}$ ), geopotential height at 500 hPa ( $Z_{500}$ ), mean sea level pressure over oceans (PSL), specific humidity at 850 hPa ( $Q_{850}$ ), and total cloud fraction (CLT).....	35
Fig. 4.2. DJF and JJA air temperature. First row: Reference (dark line) and mean model (white line) zonal averages at 200 hPa, with $\pm 1\&2$ inter-model standard deviations (shaded); Second and third rows: Reference (ERA15) and mean model zonal cross sections; Fourth row: Mean model–Reference. ....	37
Fig. 4.3. DJF and JJA zonal wind (m/s). First row: Reference (dark line) and mean model (white line) zonal averages at 850 hPa, with $\pm 1\&2$ inter-model standard deviations (shaded); Second and third rows: reference and mean model zonal cross sections; Fourth row: Mean model–reference. ....	38
Fig. 4.4. DJF and JJA specific humidity (kg/kg). First row: Reference (dark line) and mean model (white line) zonal averages at 850hPa, with $\pm 1\&2$ inter-model standard deviations (shaded); Second and third rows: reference and mean model zonal cross sections; Fourth row: Mean model–reference. ....	39
Fig. 4.5. DJF and JJA surface air temperature (K). Upper panels: observed (dark line) and mean model (white line) zonal averages with $\pm 1\&2$ inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean. ....	44
Fig. 4.6. DJF and JJA surface air temperature (K). First row: Jones (observations); Second row: Multi-model ensemble mean; Third row: Multi-model ensemble observations; Bottom row: Multi model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots. ....	45
Fig. 4.7. DJF and JJA total precipitation (mm/day). Upper panels: observed (dark line) and mean model (white line) zonal averages with $\pm 1\&2$ inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean. ....	46
Fig. 4.8. DJF and JJA total precipitation (mm/day). First row: CMAP (Observations); Second row: Multi-model ensemble mean; Third row: Multi-model ensemble observations; Bottom row: Multi-model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots. ....	47
Fig. 4.9. DJF and JJA outgoing longwave radiation ( $Wm^{-2}$ ). Upper panels: observed (dark line) and mean model (white line) zonal averages with $\pm 1\&2$ inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean. ....	48

Fig. 4.10. DJF and JJA outgoing longwave radiation ( $\text{Wm}^{-2}$ ). First row: ERBE (Observations); Second row: Multi-model ensemble mean; Third row: Observations minus Multi-model ensemble; Bottom row: Multi model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots.....	49
Fig. 4.11. DJF and JJA total cloud cover (percent). Upper panels: observed (dark line) and mean model (white line) zonal averages with $\pm 1\&2$ inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean.....	50
Fig. 4.12. DJF and JJA total cloud cover (percent). First row: ISCCP (Observations); Second row: Multi-model ensemble mean; Third row: Observations minus Multi-model ensemble; Bottom row: Multi model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots.....	51
Fig. 4.13. DJF and JJA sea level pressure (hPa). Upper panels: observed (dark line) and mean model (white line) zonal averages with $\pm 1\&2$ inter-model standard deviations (shaded); Middle panels: ocean pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean.....	52
Fig. 4.14. DJF and JJA mean sea level pressure (hPa). First row: ERA15 (Reanalysis); Second row: Multi-model ensemble mean; Third row: reanalysis minus Multi-model ensemble; Bottom row: Multi model ensemble inter-model standard deviation. Note the nonlinear scales are used for all plots.....	53
Fig. 4.15. Surface air temperature (K). Top row: Observational (Jones) annual amplitude and phase; Second row: Multi-model ensemble annual amplitude and phase; Third row: Observational ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle; Bottom row: multi-model ensemble ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle. Note the nonlinear scales.....	56
Fig. 4.16. Total precipitation (mm/day). Top row: Observational (CMAP) annual amplitude and phase; Second row: Multi-model ensemble annual amplitude and phase; Third row: Observational ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle; Bottom row: multi-model ensemble ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle.....	57
Fig. 4.17. Outgoing longwave radiation ( $\text{Wm}^{-2}$ ). Top row: Observational (ERBE) annual cycle amplitude and phase; Second row: Multi-model ensemble annual amplitude and phase; Third row: Observational ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle; Bottom row: multi-model ensemble ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle.....	58
Fig. 4.18. North American annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{Wm}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 o'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.....	61
Fig. 4.19. Southwest US regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.....	62

Fig. 4.20. South America, the Atlantic and Western Africa annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{Wm}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 O'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.....	64
Fig. 4.21. Nordeste regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.....	65
Fig. 4.22. Indo-Pacific annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{Wm}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 o'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.....	67
Fig. 4.23. India regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.....	68
Fig. 4.24. Europe/Middle East/North Africa annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{Wm}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 o'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.....	70
Fig. 4.25. Sahel regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.....	71
Fig. 4.26. (a) Annual mean CMAP rainfall and NCEP/NCAR surface windstress for 1979–95. Differences (non-flux-adjusted coupled model–observations) of simulated rainfall and windstress between (b) CCSM2.0, (c) HadCM3, and (d) PCM, and the observations in (a). .....	75
Fig. 4.27. Differences (flux-adjusted coupled model–observations) of simulated rainfall and wind stress between (a) CCCma_CGCM2, (b) CSIRO_Mk2, (c) ECHAM4_OPYC, (d) ECHO-G, (e) GFDL_R30_c, (f) HadCM2, and (g) MRI_CGCM2.3, and the observations in 4.26a. ....	76
Fig. 4.28. Differences (AMIP II model–observations) of simulated rainfall and windstress between (a) CAM2.0, (b) HadAM3, (c) ECHAM4, and (d) GFDL_R30, and the observations in Fig. 4.26a. ....	77
Fig. 4.29. Full (27) KC classes for the mean model (top panel, “mean-c03a”) for 20-y annual cycle of model years 60–79 and the observational standard (bottom panel “Observed”). .....	86
Fig. 4.30. The 8 BE (big error) types (hot, hot/wet, wet, cold/wet, cold, cold/dry, dry, hot/dry) and seasonality (annual, summer, winter) for the mean model.....	87

Fig. 4.31. Comparison of Control (constant CO <sub>2</sub> , A) and perturbation (increasing CO <sub>2</sub> , B) runs for 11 models and the mean (mean-c03a) and median model (median-c03a), based on the 20-year annual cycle for years 60-79 <i>over land only</i> . The length of the bar proportional to the percent total area <i>on</i> is (green), <i>above</i> (warmer/dryer/) is red) and <i>below</i> (cooler/wetter, blue) the diagonal of the transfer matrix. Also displayed are the ERA-40 reanalysis (“era40”) and two observational KCs “CAI_L-W” (CAI v Legates and Willmott [1990a?] surface air temperature) and “CAI_CMAP” (CMAP v CAI precipitation). Total of the blue and red area (off-diagonal error) is printed above the model name and error > 15% is bold.....	88
Fig. 4.32. As in Fig. 4.31 except for the perturbation (test input climatograph) versus the control model (reference climatograph).....	89
Fig. 4.33. As in Fig. 4.31 except for % area with big errors: 1) hot/red, 2) dry/yellow, 3) wet/green; and 4) cold/blue. ....	90
Fig. 4.34. As in Fig. 4.32 except for big errors.....	91
Fig. 4.35. As in Fig. 4.31 except for total distribution error by the 5 major KC classes (A = purple, B = rust, C = green, D = brown, E = blue). Size of the bar indicates the magnitude of the difference and the position of the bar above (more area in the Model compared to the Obs) and below (less area) the 0 % line indicates direction, i.e., bars above 0 mean “too much” and below “too little” area of the color-coded class. The total error (sum of the magnitudes of the 5 errors) is listed above model name on x axis with bold number indicating total distribution error >15%. Note that the length of the bar above and below 0 is the same (total difference is 0.0).....	92
Fig. 5.1. Sea surface temperature time mean difference between the CMIP2+ models and Levitus (degrees C). Note the large values at high latitude are a result of comparing the model skin temperature on top of sea-ice vs. the observed estimate below the ice.....	95
Fig. 5.2. Sea surface salinity time mean difference between the CMIP2+ models and Levitus (ppt).....	96
Fig. 5.3. Potential temperature in the Atlantic Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.....	98
Fig. 5.4. Salinity in the Atlantic Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.....	99
Fig. 5.5. Salinity section at 35°N in the Atlantic Ocean. Note the different vertical scales for 0–1000 m and 1000–5000 m.....	100
Fig. 5.6. Meridional current at 35°N in the Atlantic Ocean. Note the non-uniform contour increments.....	101
Fig. 5.7. Potential temperature in the Pacific Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.....	103
Fig. 5.8. Salinity in the Pacific Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m. ....	104
Fig. 5.9. Equatorial Pacific (averaged 2°S–2°N) upper ocean temperature.....	105
Fig. 5.10. Cross section of zonal currents at 145°W in the Pacific Ocean (m/s). ....	106
Fig. 5.11. Potential temperature in the Indian Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.....	108
Fig. 5.12. Salinity in the Indian Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m. ....	109
Fig. 5.13. Northward current in the Indian Ocean at 35°S. Note the non-uniform contour increments. ....	110
Fig. 5.14. Potential temperature in the Arctic Ocean: zonal-average versus depth .....	112

Fig. 5.15. Salinity in the Arctic Ocean: zonal-average versus depth.....	113
Fig. 5.16. Northward current through the Fram Strait and the Barents Sea at 80°N.....	114
Fig. 5.17. Potential temperature in the Southern Ocean: zonal-average versus depth.....	116
Fig. 5.18. Salinity in the Southern Ocean: zonal-average versus depth.....	117
Fig. 5.19. Southern Ocean zonal current at 60°W.....	118
Fig. 5.20. Sea-ice concentration (percent) for the CMIP 20-year climatologies and observations.....	119
Fig. 5.21. Annual cycle in total sea-ice area ( $10^6 \text{ km}^2$ ) for the 20-year climatologies of the CMIP simulations and observationally based estimates.....	120
Fig. 6.1. Monthly standard deviations of NIÑO3 surface air temperature anomalies.....	124
Fig. 6.2. The maximum entropy power calculated for the surface air temperature anomalies (sea surface temperature anomalies from HadISST 1.1) from monthly mean climatology for the NIÑO3 region (5°S–5°N, 150°W–90°W). The vertical lines represent 2- and 7-year periods.....	124
Fig. 6.3. The lag correlation between surface air temperature anomalies in the NIÑO3 region and the SOI as defined above plotted for all models and reanalyses and between the HadISST 1.1 sea surface temperature and CRU SOI datasets.....	125
Fig. 6.4. The evolution of the surface air temperature anomaly in the NIÑO3 region for a composite warm event in models, reanalyses and the HadISST 1.1 dataset. The shaded area represents the one standard deviation envelope of the observed NIÑO3 sea surface temperature anomaly for warm events in the HadISST 1.1 dataset.....	125
Fig. 6.5. December–February warm event composites of (a) Surface air temperature, (b) Sea-level pressure, and (c) Precipitation from MRI_CGCM2.3.....	126
Fig. 6.6. Annual mean sea-level pressure climatology (hPa): (a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) model–NCEP/NCAR reanalysis. Variance of monthly mean sea-level pressure anomalies (hPa <sup>2</sup> ), (d) NCEP/NCAR reanalysis (1979–95), (e) 20 years of CCSM2.0, (f) model–NCEP/NCAR reanalysis.....	132
Fig. 6.7. Annual mean surface air temperature climatology (K): (a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) model–NCEP/NCAR reanalysis. Variance of monthly mean surface air temperature anomalies (K <sup>2</sup> ), (d) NCEP/NCAR reanalysis (1979–95), 20 years of CCSM2.0, (f) model–NCEP/NCAR reanalysis.....	133
Fig. 6.8. Linear regression of the NAO principal component with monthly anomalies of sea-level pressure (hPa) and 850-hPa wind (ms <sup>-1</sup> ): (a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) CCSM2.0–NCEP/NCAR reanalysis. The regressions have been scaled by a one-standard-deviation perturbation of the respective principal components, and plotted where they are at least 5% significant.....	134
Fig. 6.9. Linear regression of the NAO principal component with monthly anomalies of surface air temperature (K): (a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) CCSM2.0–NCEP/NCAR reanalysis. The regressions has been scaled by a one-standard-deviation perturbation of the respective principal components, and plotted where they are at least 5% significant.....	135
Fig. 6.10. Zonal mean (90°W–40°E) of the regression of the observed and simulated NAO principal components with monthly anomalies of sea-level pressure (hPa). The regressions have been scaled by a one-standard-deviation perturbation of the respective principal components. The zonal averaging has been done prior to masking for statistical significance.....	136

Fig. 6.11. Meridional mean ( $50^{\circ}\text{N}$ – $85^{\circ}\text{N}$ ) of the regression of the observed and simulated NAO principal components with monthly anomalies of surface air temperature (K). The regressions have been scaled by a one-standard-deviation perturbation of the respective principal components. The zonal averaging has been done prior to masking for statistical significance.....	137
Fig. 6.12. Scatter-plot of the standard deviation of the NAO principal components with respect to their lag-1 autocorrelation. The result(s) from NCEP/NCAR reanalysis is given in blue; those from the coupled CMIP2+ and AMIP II models are given in red and black, respectively.....	138
Fig. 6.13. Frequency–wavenumber plots of OLR between $15^{\circ}\text{N}$ – $15^{\circ}\text{S}$ decomposed into eastward- and westward-propagating wavenumbers for the antisymmetric and symmetric spectra: (a–b) AVHRR OLR, (c–d) NCEP/NCAR Reanalysis (1979–97) OLR, (e–f) CAM2.0 OLR, (g–h) CSIRO_Mk2 OLR, (i–j) ECHAM4_OPYC3 OLR, (k–l) ECHO-G OLR, (m–n) GFDL_R30_c OLR, (o–p) HadCM3 OLR, and (q–r) PCM OLR. ....	141
Fig. 6.14. November–March outgoing longwave radiation (OLR): climatology of (a) AVHRR and (b) ECHO-G; daily variance of (c) AVHRR, and (d) ECHO-G; 20–100-day bandpass-filtered variance (e) AVHRR, (f) ECHO-G, and percent of daily variance explained by the 20–100-day bandpass-filtered OLR (g) AVHRR and (h) ECHO-G.....	151
Fig. 6.15. Lead-lag correlation between PC-1 and PC-2 for each winter. Positive correlations at positive time lags indicate that convection over the Indian Ocean leads that over the western Pacific Ocean. The solid black curve is the average over all years of data. For the model, the dashed black curve is the average for years that lie in the observed phase-space (upper-right quadrant) of Fig. 6.14c. (a) AVHRR, (b) ECHO-G, (c) Phase-space of the maximum positive correlation and its associated time lag for each year of data.....	152
Fig. 6.16. Lag-0 linear regressions of PC-1 with 20–100-day filtered OLR ( $\text{W m}^{-2}$ ): (a) AVHRR OLR, (b) ECHO-G. Panels (c) and (d) are as (a) and (b), but for regressions using PC-2. Data are plotted for a one-standard-deviation perturbation of the respective principal components where the fit is 5% significant or better, assuming each pentad is independent. ....	153
Fig. 6.17. Longitude–time lag plots of the linear regression of PC-1 with $5^{\circ}\text{N}$ – $5^{\circ}\text{S}$ averaged 20–100-day bandpass-filtered (a) AVHRR OLR ( $\text{W m}^{-2}$ ), (c) SST and ground temperature (K), and (e) 10-m zonal wind ( $\text{ms}^{-1}$ ). Contours of the OLR regression are plotted on each panel in increments of $2.5 \text{ W m}^{-2}$ . Panels (b), (d), and (f) are as (a), (c), and (e) but for ECHO-G. Time lags run from -25 to 25 days. The vertical dashed line gives the longitude of strongest convection in EOF-1 (Fig. 6.15a), and the horizontal dashed line corresponds to zero time lag. Data are plotted for a one-standard-deviation perturbation of the respective principal components where the fit is 5% significant or better assuming each pentad is independent. On each plot isolines of the OLR regression are plotted (negative values correspond to enhanced convection). ....	154
Fig. 6.18. As Fig. 6.16 but for (a–b) latent heat flux ( $\text{W m}^{-2}$ ), (c–d) 1000-hPa divergence ( $\text{s}^{-1}$ ), and (e–f) 1000-hPa specific humidity ( $\text{kg kg}^{-1}$ ). ....	155
Fig. 6.19. Longitude–height cross sections of zero-time-lag linear regressions of PC-1 with $5^{\circ}\text{N}$ – $5^{\circ}\text{S}$ averaged 20–100-day bandpass-filtered (a) divergence ( $\text{s}^{-1}$ ), (c) specific humidity ( $\text{kg kg}^{-1}$ ), and (e) zonal wind/vertical velocity vectors [note: the vertical velocity ( $\text{Pa s}^{-1}$ ) has been multiplied by -100 to give scaling compatible with the u-wind ( $\text{ms}^{-1}$ )] and contours of the u-wind in increments of $0.5 \text{ ms}^{-1}$ . Panels (b), (d), and (f) are as (a), (c), and (e), but for ECHO-G. Note: in (f) the vertical velocity was unavailable, so the vectors are omitted. The vertical dashed line at $125^{\circ}\text{E}$ is the longitude of strongest convection in Fig. 6.15a. Wind vectors are plotted at every other grid point for clarity. Data are plotted for a one-standard-deviation perturbation of PC-1 where the fit is 5% significant or better, assuming each pentad is independent. ....	156

Fig. 6.20. Time-lag versus height plots of linear regressions of PC-1 at 125°E (5°N–5°S averaged) of 20–100-day bandpass-filtered (a) divergence ( $s^{-1}$ ), and (c) specific humidity ( $kg\ kg^{-1}$ ). Panels (b) and (d) are as (a) and (c), but for ECHO-G. Data are plotted for a one-standard-deviation perturbation of PC-1 where the fit is 5% significant or better assuming each pentad is independent.....	157
Fig. 6.21. As Fig. 6.16, but for longitude–time lag plots of the linear regression of PC-1 with 5°N–5°S averaged 20–100-day bandpass-filtered OLR ( $Wm^{-2}$ ) from the additional CMIP2+ models (a) CCSM2.0, (b) CSIRO_Mk2, (c) ECHAM4_OPYC, (d) GFDL_R30_c, (e) HadCM3, and (f) PCM.....	158
Fig. A.2.1. Sample Taylor diagram displaying a statistical comparison between eight model estimates of the June/July/August precipitation fields and observations.....	179

## List of Tables

Table 2.1. Salient features of coupled ocean-atmosphere general circulation models (OAGCMs) entered in the 2004 PCMDI appraisal are listed alphabetically by model acronym along with the approximate year of the respective simulations (“vintage”). Also listed are the respective sponsoring institutions, the horizontal and vertical resolution of the model atmosphere and ocean as well as the pressure of the atmospheric top and the oceanic vertical coordinate and upper boundary condition. The representation of sea ice structure and dynamics, as well as the representation of soil moisture, vegetation, and runoff (e.g. discharge of fresh water to the ocean via a river routing model) also are indicated, with citations of references that describe further details of components of the coupled OAGCMs. See also the explanatory notes which follow the table on the next page. ....	6
Table 2.2. Features of the respective control and perturbed (1% per year increasing CO <sub>2</sub> ) experiments of the coupled ocean-atmosphere general circulation models (OAGCMs) entered in the 2004 PCMDI appraisal are listed alphabetically by model name. Brief descriptions of the spin-up procedures/durations and flux adjustments (if any) employed, along with pertinent references are included. Also listed are the archived control run and perturbed run years, <i>both expressed relative to the end of the coupled spin-up period</i> . See also the explanatory notes which follow the table on the next page. ....	9
Table 3.1. SST trends for the CMIP2+ models (K/century). ....	13
Table 4.1. Annual mean rainfall rate (mm day <sup>-1</sup> ), root-mean-square difference (RMSD), pattern correlation (R) and linear slope for 130°E–90°W, 20°N–20°S relative to CMAP (1979–95; ocean only). The models are grouped into those that did not use flux adjustment, those that used flux adjustment (shaded), and the AMIP II models (last 4 entries). ....	73
Table 4.2. Annual mean sea-surface temperature (C), root-mean-square difference (RMSD), pattern correlation (R) and linear slope for 130°E–90°W, 20°N–20°S) relative to Reynolds SST (1979–95; ocean only). Models that used flux adjustment are shaded. ....	73
Table 4.3. Description of the Köppen Climate Classes—major (5), main (11) and full (all 27). ....	84
Table 4.4. Transfer matrix between the mean model (Model) and observations (Obs) in % surface land area for the major KC classes. Values on the diagonal (shaded) are the % area the Model and Obs classes agree; values above the diagonal where the model class is lower (toward A); and values below the diagonal elements where the model class is higher (toward E). Above-diagonal area can be interpreted as a shift toward a warmer, moister climate (E→A), and below diagonal area as a shift toward a cooler, dryer climate (A→E). The total area in each class for the Model and Obs is shown in the last two columns. For example, the observed total area for class B (dry) is 26.6%, but for the mean model it is 21.3%, or 5.3% less than observed. ....	85
Table 6.1. Reanalyzed, CMIP2+, and AMIP II (the last 3 rows) NAO characteristics. The columns give the observation/model designation, the standard deviation of PC-1 (hPa), the one-month lag correlation of PC-1, RL1, and the root-mean-squared error, RMSE (hPa), and pattern correlation, Rpat, of the linear regression of PC-1 with sea-level pressure (scaled by a one-standard-deviation perturbation of PC-1). Shaded models used the same atmospheric component in their CMIP2+ and AMIP II simulations. ....	131
Table 6.2. Observed, reanalyzed, CMIP2+, and AMIP II (the last 4 rows) model characteristics of the Madden-Julian Oscillation for years when eastward propagation of convection dominates (see Fig. 6.14). Given are the standard deviations of PC-1 and PC-2, the maximum positive correlation of PC-1 relative to PC-2, R, the time lag (days) at which it occurred, and the fraction of years for which the PCs had a lead-lag relationship consistent with the observations. Shaded models used the same atmospheric component in their CMIP2+ and AMIP II simulations. ....	150



# 1 Introduction

*D. Bader*

In 2002, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) proposed the concept for a state-of-the-science appraisal of climate models to be performed approximately every two years. Motivation for this idea arose from the perceived needs of the international modeling groups and the broader climate research community to document progress more frequently than provided by the Intergovernmental Panel on Climate Change (IPCC) Assessment Reports. A committee of external reviewers, which included senior researchers from four leading international modeling centers, supported the concept by stating in its review:

“The panel enthusiastically endorses the suggestion that PCMDI develop an independent appraisal of coupled model performance every 2-3 years. This would provide a useful ‘mid-course’ evaluation of modeling progress in the context of larger IPCC and national assessment activities, and should include both coupled and single-component model evaluations.”

Subsequently, the PCMDI presented the appraisal concept to the World Climate Research Programme’s Working Group on Coupled Modeling (WGCM) in October 2002 and received encouragement from the WGCM to pursue the idea. In many respects, this appraisal continues the seminal work begun by WGCM in its first report to the WCRP (Gates et al., 1993). This first appraisal would have been impossible without the collaboration and participation of the international modeling groups. They have trusted the PCMDI with the output from their model simulations, which form the foundation for the appraisal. As we have done in the past with the Atmospheric Model Intercomparison Project (AMIP) and the Coupled Model Intercomparison Project (CMIP), the PCMDI’s charter is to provide insightful diagnostics of community simulations taken collectively, and not to make judgments on individual models. We must state clearly that this appraisal is a “snapshot” of the climate models as represented by the simulations in our database, many of which are a few years old. Model development is an ongoing activity, and nearly all of the models included in our database have been improved since these simulations were completed. Further, the pace of model development differs among modeling institutions. Consequently, while one group’s model may be viewed as “lagging” its peers in terms of completeness or sophistication, that same group may be viewed as leading the field in the near future following a period of intense development.

This first appraisal effort endeavors to accomplish several objectives. The modeling community can use the appraisal as a baseline upon which future model improvements can be documented and evaluated. Additionally, it is a reference document for the climate researchers and others who rely on models and model simulations to provide new knowledge about the state of the climate system and projected climate change. Broadly defined, this group includes those who contribute to and read the IPCC Working Group 1 Scientific Assessment of Climate Change Reports. This and later appraisals will provide them with an assessment of the strengths and weaknesses of current climate models, as well as documentation of model improvement over time. For this reason, we have focused on the aspects of climate model simulations that we feel are the most relevant in the study of century-scale climate change, including the simulation of the current climate and climate variability.

We have restricted this appraisal to analyses that can be completed using the simulations contributed as part of the Coupled Model Intercomparison Project's "2+" experiment (CMIP2+, Covey et al., 2003). It must be noted that the CMIP2+ experiment was not designed to meet the needs of this appraisal, but rather provided the opportunity for the PCMDI to undertake and complete a comprehensive study of coupled climate models collectively. Consequently, there are aspects of model behavior that were beyond the scope of the CMIP2+ experiment, particularly cloud forcing and feedback, which are not adequately addressed in this first appraisal.

Chapter 2 provides background on the models used to produce the simulations in the database. The heart of the appraisal starts in Chapter 3 with an analysis of the global trends in surface temperatures for the unforced control simulations, which provides an estimate of the coupled models' long-term dynamical behavior, and any tendencies to "drift" to different steady-states. Following the examination of the control simulations, we present a similar surface temperature analysis of simulations from an idealized climate change scenario with slowly increasing atmospheric CO<sub>2</sub> concentration. An extensive analysis of the models' steady-state climatologies from the control runs is found in Chapter 4. In this chapter we look at the three-dimensional atmospheric structure as well as some aspects of region-specific surface climate. Chapter 5 contains results from initial studies of ocean circulation. We conclude the appraisal in Chapter 6 with an extensive analysis of the models' ability to simulate three important observed modes of climate variability. A brief summary of the diagnostics is presented in Chapter 7.

The PCMDI will start work on the next appraisal almost immediately after completion of this version, in part to include simulations from models that have recently completed a development cycle. We intend to expand the analysis to give a more comprehensive evaluation of the ocean, land-surface hydrology, cloud-climate feedbacks and possibly additional features that may be suggested by the community in response to this first appraisal. Future appraisals are planned at two-year intervals.

## 2 Features of Appraisal Models

*T. Phillips*

The climate simulations analyzed in this PCMDI appraisal of coupled ocean-atmosphere models are drawn from submissions to the second phase of the Coupled Model Intercomparison Project, designated as CMIP2. The CMIP2 intercomparison comprises paired runs from each participating coupled model: 1) a simulation of present-day climate and 2) a simulation of a perturbed climate resulting from increasing the concentration of atmospheric carbon dioxide (CO<sub>2</sub>) at a rate of 1 percent per year for a minimum period of 80 years. (See further details of the CMIP2 numerical experiments and minimal data requirements at <http://www-pcmdi.llnl.gov/cmip/announ.htm>.) Moreover, a subset of the model submissions (denoted as CMIP2+) include time series of substantially more climate variables than the minimal requirements for CMIP2. These expanded data holdings afford the opportunity to more comprehensively intercompare coupled model simulations of present-day and CO<sub>2</sub>-perturbed climates than has previously been possible. Thus, the CMIP2+ simulations are the focus of this PCMDI appraisal.

The acronyms of the eleven coupled atmosphere-ocean general circulation models (AOGCMs) which produced these CMIP2+ simulations are listed in the first column of Table 2.1. A comparison with previous lists of this type (e.g., Table 1 in Covey et al., 2003) reveals that two of these models (BCM and ECHO-G) are new to CMIP2, while four (CCCma\_CGCM2, CCSM2.0, MRI\_CGCM2.3, and PCM) are current versions of models that are already participating in CMIP2. (Thus far, only time series of ocean variables from the BCM model have been provided to PCMDI; hence this simulation is not considered in the sections of this appraisal which pertain to different aspects of the atmospheric climates of the CMIP2+ models.)

The CMIP2+ models typify the current generation of coupled AOGCMs, and they display certain broad similarities. For example, the dynamics of the atmosphere and ocean components of the respective coupled models are all based on the global primitive equations—coupled partial differential equations that predict fluid flow on a rotating sphere. In every CMIP2+ model, also, the temperature and thickness of sea ice as well as the temperature and hydrology of the land are predicted variables. All the CMIP2+ models thus account, in some fashion, for the coupling of major components of the global climate system, and so are suited for examining potential secular global climate change. In this respect, the CMIP2+ models all contrast with simpler coupled ocean-atmosphere models that are more geographically and/or physically constrained, and thus are designed for more limited objectives (e.g. for prediction of ENSO phenomena, the climate of a particular season, etc.).

Nevertheless, there are numerous differences among these CMIP2+ entries, some of which are enumerated in Table 2.1, along with citations of references that detail these (and many other) feature differences. For example, there is considerable variation in the “vintage” of the submitted runs (column 1). The majority of these simulations were generated by models that are of a mid-to-late 1990s vintage (CSIRO\_Mk2, ECHAM4\_OPYC3, ECHO-G, GFDL\_R30\_c, HadCM2, HadCM3, and PCM), while the remainder (BCM, CCCma\_CGCM2, CCSM2.0, and MRI\_CGCM2.3) were produced more recently. In addition, there is considerable national diversity evident in the model

sponsors (column 2 of Table 2.1), with groups in Australia, Canada, Germany, Japan, Norway, the UK and the USA represented.

Except for HadCM2 and HadCM3, the atmospheric components of the coupled models (column 3 of Table 2.1) represent state variables as spectral quantities (i.e. as coefficients of a series of spherical harmonic functions at progressively finer spatial scales). Ocean variables, however, are universally formulated in finite-difference form (i.e. as quantities on a latitude-longitude grid), but with some variation in the vertical coordinate: most ocean models assign depth for this dimension, but two models (BCM, ECHAM4\_OPYC3) instead use density (column 4 of Table 2.1). Most ocean models also have a rigid-lid upper boundary condition, but in some (BCM, CCSM2.0, ECHAM4\_OPYC3, ECHO-G, and PCM) the ocean top is treated as a free surface.

There is also a fairly large range in atmospheric horizontal resolution among the appraisal models: from spectral R21 ( $\sim 3.2^\circ \times 5.6^\circ$ ) for the CSIRO model to spectral T63 ( $\sim 1.9^\circ \times 1.9^\circ$ ) for BCM. There is considerable variation in vertical resolution as well (from L9 for CSIRO\_Mk2 to L31 for BCM). The atmospheric model top also ranges between 21 hPa in CSIRO-Mk2 to 0.4 hPa in MRI\_CGCM2.3, where the presence of 16 vertical levels (out of a total of 30) above 200 hPa also permits rudimentary simulation of stratospheric circulations. There is a somewhat lower range in ocean vertical resolution across the majority of models. (Most have  $\sim 20$  ocean levels, but ECHAM4\_OPYC3 has only 11, while CCSM2.0 has 40.) Somewhat greater variations are to be found in ocean horizontal resolution (e.g.,  $3.2^\circ \times 5.6^\circ$  for CSIRO\_Mk2 and  $0.5\text{--}0.67^\circ \times 0.67^\circ$  in PCM), but with many models specifying tighter latitudinal resolution ( $\sim 0.3\text{--}0.5^\circ$ ) near the Equator in order to more accurately simulate equatorial wave dynamics.

Sea ice dynamics also is simulated in all the models, with the majority doing so via an explicit rheology (column 5 of Table 2.1), while the rest (GFDL\_R30\_c, HadCM2, HadCM3, MRI\_CGCM2.3) treat the motion of the ice as being governed by a free-drift approximation (i.e. setting the internal pressure of the ice to zero). Except for GFDL\_R30\_c, all of the models account for ice leads as well (fractional ice coverage of a grid box), either as a consequence of the chosen rheology or as a parameterized feature.

There is a broad spectrum of complexity in representing land hydrology (column 6 of Table 2.1). Two models (CCCma\_CGCM2 and GFDL\_R30\_c) adopt a classic “bucket” formulation of soil moisture that accounts only implicitly for the effects of vegetation (albeit with geographical variation in the bucket’s moisture capacity, in the case of CCCma\_CGCM2). Two other models (ECHAM4\_OPYC3 and ECHO-G) also employ a bucket representation of soil moisture, but include a vegetation canopy that intercepts and re-evaporates a fraction of the precipitation, while imposing a resistance on the evapotranspiration of soil moisture by plants. In addition to including a canopy, the majority of the models (BCM, CCSM2.0, CSIRO\_Mk2, HadCM2, HadCM3, MRI\_CGCM2.3, and PCM) also employ a more complex representation of soil moisture than a simple bucket (e.g. diffusion of water through multiple soil layers). Except for three models (CSIRO\_Mk2 and PCM), the routing of freshwater runoff via rivers to the ocean also is represented in a rudimentary way.

Selected aspects, also, of the control and perturbed simulations of each CMIP2+ model are included in Table 2.2. In column 2, the procedures for spinning up the coupled models so as to obtain initial conditions for their respective control simulations are

briefly described, with cited references that elaborate these details. While the details of the spin-up procedures are seen to be quite varied, most of the modeling groups first spin up the atmosphere and ocean components in stand-alone mode (although often using information derived from the atmospheric spin-up as boundary forcing for the ocean) before initiating the spin-up of the coupled system.

The spin-up procedures of the CCSM2.0 and HadCM3 models are the chief exceptions to such a sequential procedure. In these models, spin-up of the coupled system proceeds directly after initialization of the atmosphere and ocean components from relevant climatologies. In particular, the estimates of climatological ocean temperature and salinity provided by Levitus and co-workers (e.g. Levitus, 1982; Levitus and Boyer, 1994; Levitus et al., 1995) figure prominently in the spin-up procedures of virtually all the models. Because the typical coupled spin-up period is less than ~150 years, the memory of these ocean climatologies probably is substantially retained in the corresponding control runs (especially in their early phases). Possible exceptions include the control simulations of the CCSM2.0, GFDL\_R30\_c, HadCM2, and HadCM3 models, which follow after coupled spin-ups of 400–900 years duration, and therefore may display greater drift from the prescribed ocean climatologies.

Most of the CMIP2+ models also employ some types of ocean-atmosphere flux adjustments (usually computed as part of the spin-up procedures) in order to limit nonphysical coupled climate drift. Three models (CCSM2.0, HadCM3, and PCM) do not apply any flux adjustment, however; whether other models adjust surface fluxes of heat, fresh water (“water”), and momentum (“winds”) is noted in column 3 of Table 2.2 along with pertinent references.

Many of the control runs supplied to the PCMDI data archive begin immediately after (i.e. in year “1”—see column 4 of Table 2.2) the coupled spin-up period. However, the start of some archived control runs is ~50–300 years after the completion of the coupled spin-up (e.g., CCCma(CGCM2, CSIRO-Mk2, ECHAM4\_OPYC3, ECHO-G, and HadCM3). The length of the archived control runs also ranges widely across the models: from 80 years (HadCM2) to 650 years (CCSM2.0). The diversity of the control runs, as well as the varying lengths of the corresponding spin-up periods, have potential impacts on the trends in ocean variables that are analyzed in Section 3 of this report. *Note, however, that for consistency in analyzing climatological features of the model control runs (e.g., as in Sections 4–6 of this document), the relevant statistics were computed only over years 60–79 from the beginning of each archived control run, irrespective of its total length.*

Finally, the archived perturbed runs are also seen to vary considerably in length, from ~80 to 150 years (column 5 of Table 2.2). They display obvious variation in another respects as well: while most of the supplied perturbed simulations have the same starting point as their associated control run, several (e.g. BCM, GFDL\_R30\_c, and PCM) instead “branch off” from the control at a later point.

The reader is directed to the reference citations in Tables 2.1 and 2.2 for more details on the many other inter-model/experimental differences that cannot be captured in a few tables. Further information on selected features of some of the CMIP2+ models also may be found online at <http://www-pcmdi.llnl.gov/projects/cmip/>.

Table 2.1. Salient features of coupled ocean-atmosphere general circulation models (OAGCMs) entered in the 2004 PCMDI appraisal are listed alphabetically by model acronym along with the approximate year of the respective simulations (“vintage”). Also listed are the respective sponsoring institutions, the horizontal and vertical resolution of the model atmosphere and ocean as well as the pressure of the atmospheric top and the oceanic vertical coordinate and upper boundary condition. The representation of sea ice structure and dynamics, as well as the representation of soil moisture, vegetation, and runoff (e.g. discharge of fresh water to the ocean via a river routing model) also are indicated, with citations of references that describe further details of components of the coupled OAGCMs. See also the explanatory notes which follow the table on the next page.

<b>Model, Vintage</b>	<b>Sponsor, Country</b>	<b>Atmosphere: Resolution References</b>	<b>Ocean: Resolution Z Coord., TopBC References</b>	<b>Sea Ice: Dynamics, Structure References</b>	<b>Land: Soil, Plants, Rivers References</b>
BCM, 2002	University of Bergen (UB), Norway	top = 10 hPa T63 (1.9°×1.9°) L31 Deque et al., 1994	0.8-2.4°×2.4° L24 density, free sfc. Bleck et al., 1992	rheology, leads Drange&Simonsen, 1996; Hibler, 1979	layers, canopy, routing Douville et al., 1995; Mahfouf et al., 1995
CCCma_CGCM2, 2001	Canadian Centre for Climate Modelling & Analysis (CCCma), Canada	top = 5 hPa T32 (3.7°×3.7°) L10 McFarlane et al., 1992	1.9°×1.9° L29 depth, rigid lid, Pacanowski et al., 1993	rheology, leads Flato&Hibler, 1990	bucket, routing McFarlane et al., 1992
CCSM2.0, 2002	National Center for Atmospheric Research (NCAR), USA	top = 2.9 hPa T42 (2.8°×2.8°) L26 Collins et al., 2003	0.3-1.0°×1.0° L40 depth, free sfc. Smith&Gent, 2002	rheology, leads Briegleb et al., 2002	layers, canopy, routing Bonan et al., 2002; Branstetter&Erickson, 2003
CSIRO_Mk2, 1997	Commonwealth Scientific & Industrial Research Organization (CSIRO), Australia	top = 21 hPa R21 (3.2°×5.6°) L9 McGregor et al., 1993	3.2°×5.6° L21 depth, rigid lid Hirst et al., 2000	rheology, leads O'Farrell, 1998	layers, canopy Kowalczyk et al., 1991, 94
ECHAM4_OPYC3, 1996	Max Planck Institut fur Meteorologie (MPI), Germany	top = 10 hPa T42 (2.8°×2.8°) L19 Roeckner et al., 1996b	0.5-2.8°×2.8° L11 density, free sfc. Oberhuber, 1993	rheology, leads Oberhuber, 1993	bucket, canopy, routing, Roeckner et al., 1996b
ECHO-G, 1999	Model & Data Group (M&D), Germany	top = 10 hPa T30 (3.9°×3.9°) L19 Roeckner et al., 1996b	0.5-2.8°×2.8° L20 depth, free sfc. Wolff et al., 1997	rheology, leads Wolff et al., 1997	bucket, canopy, routing Roeckner et al., 1996b

Table 2.1. (continued)

Model, Vintage	Sponsor, Country	Atmosphere: Resolution References	Ocean: Resolution Z Coord., Top BC References	Sea Ice: Dynamics, Structure References	Land: Soil, Plants, Rivers References
GFDL_R30_c, 1996	Geophysical Fluid Dynamics Laboratory (GFDL), USA	top = 15 hPa R30 ( $2.3^\circ \times 3.8^\circ$ ) L14 Delworth et al., 2002	$1.9^\circ \times 2.3^\circ$ L18 depth, rigid lid Pacanowski et al., 1993	free drift, no leads Delworth et al., 2002	bucket, routing Milly, 1992
HadCM2, 1995	Meteorological Office (MO), UK	top = 5 hPa $2.5^\circ \times 3.8^\circ$ L19 Cullen, 1993; Hewitt&Mitchell, 1996	$2.5^\circ \times 3.8^\circ$ L20 depth, rigid lid Bryan, 1969; Cox, 1984	free drift, leads Cattle&Crossley, 1995	layers, canopy, routing Warrilow et al., 1986; Gregory&Smith, 1990
HadCM3, 1997		top = 5 hPa $2.5^\circ \times 3.8^\circ$ L19 Pope et al., 2000	$1.5^\circ \times 1.5^\circ$ L20 depth, rigid lid Gordon et al., 2000	free drift, leads Cattle&Crossley, 1995	layers, canopy, routing Cox et al., 1999
MRI_CGCM2.3, 2002	Meteorological Research Institute (MRI), Japan	top = 0.4 hPa T42 ( $2.8^\circ \times 2.8^\circ$ ) L30 Yukimoto et al., 2001	$0.5\text{--}2.0^\circ \times 2.5^\circ$ L23 depth, rigid lid Yukimoto et al., 2001	free drift, leads Mellor&Kantha, 1989	layers, canopy, routing Sellers et al., 1986; Sato et al., 1989
PCM, 1999	Department of Energy(DOE), USA	top = 2.9 hPa T42 ( $2.8^\circ \times 2.8^\circ$ ) L18 Kiehl et al., 1998	$0.5\text{--}7^\circ \times 0.7^\circ$ L32 depth, free sfc. Maltrud et al., 1998	rheology, leads Zhang et al., 1999	layers, canopy Bonan, 1998

**Notes**

**Atmosphere: Top pressure, Resolution, References**—The pressure at the top of the model atmosphere is listed in hecto-Pascals (hPa). In models with finite-difference representations of atmospheric variables, the horizontal resolution is expressed as the size of a model grid box in degrees latitude by degrees longitude. In models employing spectral representations, the horizontal resolution is expressed as the truncation (following triangular T, or rhomboidal R schemas) wave number followed by a roughly equivalent latitude by longitude grid spacing. Vertical resolution is expressed in terms of the number of vertical levels L. Cited references describe other details of the atmospheric model.

**Ocean: Resolution, Z Coord. (Vertical Coordinate), Top BC (Boundary Condition) , References**—Horizontal resolution is expressed as the size of a grid box in degrees latitude x degrees longitude, where a range of latitude increments is listed for those models with finer latitudinal grid spacing near the Equator. Vertical resolution is indicated by the number of vertical levels L, which may follow either a depth or density coordinate, as also listed. Whether the upper boundary of the ocean model is a rigid lid or free surface is also indicated. Cited references describe other details of the ocean model.

**Sea Ice: Dynamics, Structure, References**—It is indicated whether sea ice dynamics are simulated by an explicit rheology (“rheology”) or instead are represented by a free drift (“free drift”) approximation (i.e. internal pressure of ice assumed to be zero). Whether an open-water fraction of the grid box (“leads”) is accounted for is also noted. Cited references describe other details of the sea ice model.

**Land: Soil, Plants, Rivers, References**—Whether soil moisture is modeled by a single-layer “bucket” or by a multi-layered scheme (“layers”) and whether a vegetation canopy is explicitly represented are indicated. The presence of a routing scheme (“routing”) to simulate river discharge to the ocean model is also noted. Cited references provide further details on these and other features of the land model.

Table 2.2. Features of the respective control and perturbed (1% per year increasing CO<sub>2</sub>) experiments of the coupled ocean-atmosphere general circulation models (OAGCMs) entered in the 2004 PCMDI appraisal are listed alphabetically by model name. Brief descriptions of the spin-up procedures/durations and flux adjustments (if any) employed, along with pertinent references are included. Also listed are the archived control run and perturbed run years, *both expressed relative to the end of the coupled spin-up period*. See also the explanatory notes which follow the table on the next page.

Model	Control Run Initialization: Spin-up Procedures/Durations (References)	Flux Adjustments: Heat, Water, Winds References	Years Archived:	
			Control	Perturbed
BCM	Ocean spin-up with atmospheric forcings for 125 yrs with restoration toward climatologies; atmosphere spin-up using last 10 days of ocean spin-up; coupled spin-up for 25 yrs with SSS/SST restored toward climatologies (Furevik et al., 2003).	heat, water Furevik et al., 2003	1-300	101-180
CCCma_CGCM2	Atmosphere spin-up of 30 yrs with coupling to slab ocean; ocean spin-up for ~ 4000 yrs with these atmospheric forcings and with SST/SSS restored toward observed climatologies; coupled spin-up for 50 yrs with restoring terms (Kim et al., 2002).	heat, water Flato et al., 2000	51-130	51-130
CCSM2.0	Ocean and atmosphere initialized from realistic January climatologies, followed by coupled spin-up for 350 years with constant 1990 forcings (Kiehl&Gent, 2004)	no adjustments Kiehl&Gent, 2004	1-650	?
CSIRO_Mk2	Ocean spin-up from global-average climatology for 2170 yrs with asynchronous, and for 710 yrs with synchronous forcings; atmosphere (with sea ice) spin-up for multiple decades; coupled spin-up for 105 yrs with flux adjustments (Gordon & O'Farrell, 1997).	heat, water, winds Gordon & O'Farrell, 1997	351-450	351-450
ECHAM4_OPYC3	Atmosphere spin-up for 20 yrs with monthly climatological SSTs; ocean spin-up for 500 yrs with these atmospheric forcings and SST/ SSS restored toward climatologies, then for 500 yrs with forcings from daily anomalies of heat and freshwater sfc fluxes; coupled spin-up for 100 yrs with SST/SSS restored toward climatologies (Roeckner et al., 1996a).	heat, water Roeckner et al., 1996a	150-299	?
ECHO-G	Atmosphere spin-up for 18 yrs with climatological monthly SSTs; ocean spin-up for 2034 yrs with forcings from this atmospheric run and with SST/SSS restored toward climatologies; coupled spin-up for 155 years with SST/SSS restored toward climatologies (Min et al., 2004).	heat, water (zero global means) Min et al., 2004	310-409	310-387

Table 2.2. (continued)

Model	Control Run Initialization: Spin-up Procedures/Durations (References)	Flux Adjustments: Heat, Water, Winds References	Years Archived:	
			Control	Perturbed
GFDL_R30_c	Atmosphere spin-up for 80 yrs with climatological monthly SSTs; ocean spin-up from rest for ~3000 yrs with acceleration of deep ocean and SST/SSS restored toward climatologies; coupled integrated for 900 yrs with flux adjustments (Delworth et al., 2002).	heat, water Delworth et al., 2002	1-300	51-130
HadCM2	Ocean initialized from climatological T/S, with spin-up of 510 yrs mostly in coupled mode and alternating SST-SSS restoration toward climatologies with flux adjustments (Johns et al., 1997).	heat, water Johns et al., 1997	1-80	1-80
HadCM3	Atmosphere and ocean initialized from observed climatologies, followed by coupled spin-up of 400 yrs (Gordon et al., 2000).	no adjustments Gordon et al., 2000	101-180	101-180
MRI_CGCM2.3	Ocean spin-up for 335 yrs, in partially coupled mode with forcing by atmospheric wind stresses, and with SST/SSS restored toward climatologies; fully coupled spin-up for 95 more yrs with flux adjustments (Yukimoto et al., 2001; Noda et al., 2001).	heat, water Yukimoto et al., 2001; Yukimoto&Noda, 2003	1-150	1-150
PCM	Atmosphere spin-up for 10 yrs with climatological monthly SSTs; ocean run with restoration toward observed T/S, then for 86 yrs with repeated forcings from yrs 5-10 of atmospheric run and accelerated deep-ocean T/S; coupled spin-up for 50 yrs (Washington et al., 2000).	no adjustments Washington et al., 2000	1-300	151-232

**Notes**

**Control Run Initialization: Spin-up Procedures/Durations (References)**—Spin-up procedures for atmosphere/ocean/coupled system and their respective durations in simulation years, with relevant references for further details.

**Flux Adjustments: Heat, Water, Winds, References**—Inclusion of adjustments in surface fluxes of heat (“heat”), fresh water (“water”), and/or momentum (“winds”) for the coupled ocean-atmosphere system are noted. Cited references describe details of the flux adjustments applied as well as other information on the coupled model system.

**Years Archived: Control and Perturbed**—Archived control run and perturbed run years, *expressed relative to the end of the coupled spin-up period*.

### 3 Climate Trends in Unforced and Forced Simulations

C. Covey, P. Gleckler

#### 3.1 Control Run Secular Trends

##### 3.1.1 Annual and Global Mean Surface Air Temperature Power Spectra

This section introduces the behavior of the CMIP2+ models by examining the trends they simulate in air temperature and other quantities near the Earth's surface. Historically these trends were a problematic feature of coupled ocean-atmosphere GCMs. When coupled GCMs came into widespread use in the 1980s, one concern was that unless nonphysical "flux adjustment" terms were inserted in their equations, simulated climates would drift to completely unrealistic states. The mid-1990s IPCC assessment report noted this problem (Gates et al., 1996), but the subsequent IPCC Third Assessment Report (TAR) stated, "Some non-flux adjusted models are now able to maintain stable climatologies of comparable quality to flux adjusted models" (McAvaney et al., 2001). Here we document the situation for the most recent generation of models.

First, however, we briefly consider the simulated climate variability that remains after secular trends are removed. (A thorough examination of climate variability is deferred to Section 6 below.) Figure 3.1 shows power spectra of global and annual mean surface air temperature. Each of the CMIP2+ models are represented except for the BCM, for which atmosphere output is not yet available. Observations compiled for the IPCC TAR are also included. The spectral density (vertical scale) gives the amount of temperature variance at each cyclic period (horizontal scale). We calculated the spectral density by methods described in Jenkins and Watts (1968), using auto-covariance with lags up to 1/4 the length of each time series and a Tukey window 1/10 as long as each time series.

The resulting CMIP2+ spectra displayed in the figure are quite similar to results from the earlier CMIP2 generation of models. The similarity is not surprising. As we discuss in the previous section, the CMIP2 and CMIP2+ models are an overlapping set. Power spectra for global and annual mean surface air temperature from the CMIP2 models are given in Fig. 17 of Covey et al. (2003), hereafter referred to simply as the CMIP2 overview. In both the CMIP2 overview and the present results, the models simulate variability that is generally less than or equal to that of the observations. The main exceptions to this rule (for both CMIP2 and CMIP2+) come from the HadCM2 and especially the GFDL\_R30\_c models at periods of ~10 years.

One would expect simulated variability that is less than observed—at least on longer time scales—for model control runs. By definition control runs lack the time-evolving climate forcing (both natural and anthropogenic) that is implicit in the observed record. Thus control run simulations should give a lower bound to the actual climate system's variability. Some of the discrepancy, however, may be due to problems with the models themselves if they underestimate ENSO and other phenomena at shorter (<10-year) periods. We discuss this issue in Section 6.

### **3.1.2 Annual and Global Mean Surface Air Temperature Trends**

Turning to the long-term trends of surface air temperature, Fig. 3.2 shows time series of annual and area-averaged means from each of the CMIP2+ models (except the BCM) and the IPCC TAR observations. Area averages are given separately for the Northern Hemisphere, the Southern Hemisphere and the entire globe. Note that during 1961-1991, the observed absolute temperature averaged 287 K. The global mean temperatures of the CMIP2+ models lie between 285 and 287 K. This result compares with 284-290 K reported in the CMIP2 overview.

When one examines secular trends in the CMIP2+ database, it is important to keep in mind that output from different models may come from different time segments within long control runs. For example, the CCCma CGCM2, ECHO-G and HadCM3 output in the CMIP2+ database is from near the beginning of these models' control runs, whereas control run output from the CCSM2.0, GFDL R30c and PCM extends for 300 simulated years. Institutions chose the time segment for their CMIP2+ control run output to match the time period for their 1% per year increasing atmospheric carbon dioxide scenario. This choice was necessary in order to permit comparison of the control run and increasing CO<sub>2</sub> output, but secular trends are typically smaller during later times in model control runs. Accordingly, the results discussed below represent an upper bound on control run secular trends, and it would be inappropriate to draw conclusions from these results about which models are inherently more stable.

The observed global mean linear trend is 0.4 K/century for the full 138 years of available data (and 0.7 K/century for the 20<sup>th</sup> century). With one exception, the CMIP2+ model control runs exhibit global mean linear trends that are a factor of 5 or more smaller in magnitude. As noted above, changes that are smaller than observed are expected for control runs, because these simulations don't include changes in climate forcing. The mean magnitude  $\pm$  standard deviation of the global mean trends for the CMIP2+ models is  $0.08 \pm 0.12$  K/century, similar to the CMIP2 results. The three CMIP2+ models that are not flux-adjusted—CCSM2.0, HadCM3 and PCM—have global mean trends of -0.2, 0.6 and 0.2 K/century, respectively. These trends are comparable to those of the CMIP2 non-flux-adjusted models (mean magnitude 0.3 K / century) and are much smaller than those of the CMIP1 non-flux-adjusted models (mean magnitude 1.0 K/century).

### **3.1.3 Annual Average Global Maps of SST Trends**

Evidently the newer generation of non-flux-adjusted models is able to maintain surface temperature "climate drift" within acceptable bounds for century-scale simulations, at least in a global or hemispheric mean sense. Does this statement also apply when we look at the individual grid points of models? Table 3.1 summarizes the sea surface temperature trends for the CMIP2+ models in two ways. The mean over the grid points is a global area average. It gives results for SST that are similar to those for global mean surface air temperature discussed above. The root-mean-square (RMS) over grid points gives the typical magnitude of surface temperature trend at an individual grid point of each model (or observations). As with global and hemispheric means, the RMS trends are smaller than observed, as expected for control run simulations. This statement is true for both flux-adjusted and non-flux-adjusted models. The non-flux-adjusted CCSM2.0 and PCM are particularly noteworthy in this respect. SST trends of the third CMIP2+

non-flux-adjusted model, HadCM3, are also smaller than observed but are greater than those of the CCSM2.0 and PCM. As noted above, HadCM3 output in the CMIP2+ database comes from the beginning rather than the end of its control run, exaggerating its apparent climate drift.

A more detailed comparison of trends from the CCSM2.0 model, the PCM and observation is provided by Fig. 3.3. This shows that in the CCSM2.0 simulation, there are about as many areas of warming as of cooling, and the magnitude of the SST trend is everywhere smaller than 1 K/century. Trends for the PCM are somewhat greater (up to ~1.6 K/century), with cooling near the North Pole and warming near the South Pole, but again there are about equal areas of warming and cooling. (Note that in the models, “sea surface temperature” near the poles is really ice surface temperature.) For observations during 1898–1998, in contrast, the linear trend at nearly all grid points is warming and the magnitude of the trend exceeds 2 K/century at some locations. Presumably this trend is mainly due to anthropogenic global warming, believed to dominate the climate record since the mid-20<sup>th</sup> century and not included in model control runs. Trend maps for the other CMIP2+ models (not shown) confirm this qualitative difference between control runs and observations.

Table 3.1. SST trends for the CMIP2+ models (K/century).

Model or observed dataset	Area-averaged mean over grid points	Area-averaged RMS over grid points
Observed 1956-2001 <sup>a</sup>	0.598	1.07
BCM_version1	0.157	0.14
CCCma(CGCM2)	0.397	0.72
CCSM2.0 <sup>b</sup>	-0.031	0.15
CSIRO_Mk2	-0.008	0.25
ECHAM4_OPYC3	0.081	0.87
ECHO-G	0.114	0.61
GFDL_R30_c	-0.029	0.36
HadCM2	-0.014	0.46
HadCM3 <sup>b</sup>	0.017	0.80
MRI_CGCM2.3	0.017	0.11
PCM <sup>b</sup>	0.036	0.28

<sup>a</sup> As compiled for AMIP boundary conditions.

<sup>b</sup> Not flux-adjusted.

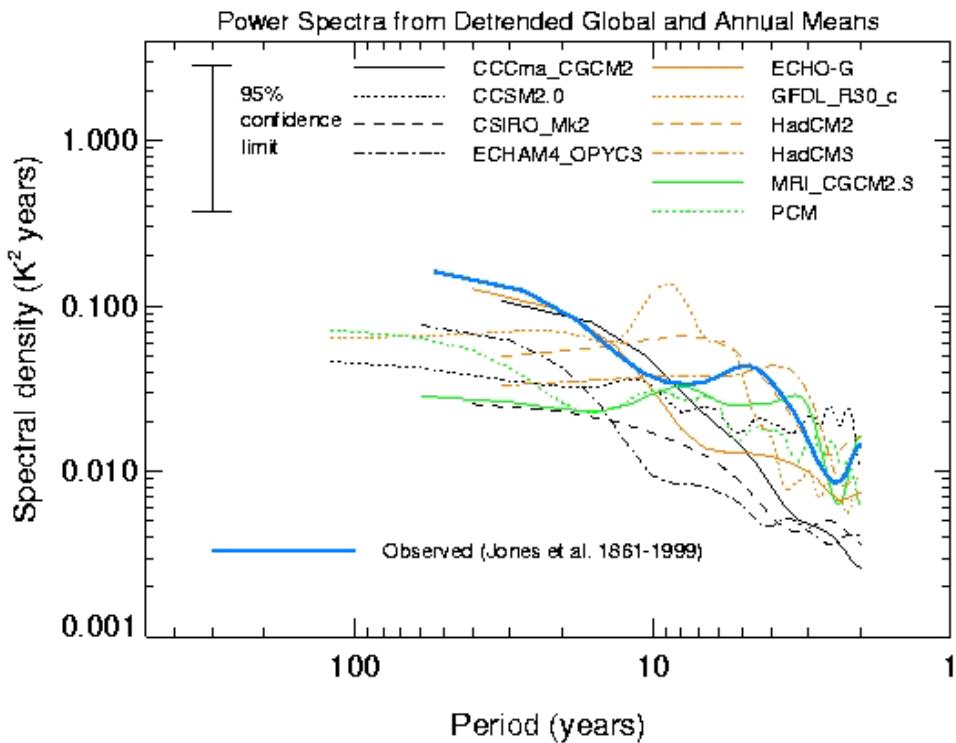


Fig. 3.1. Power spectra of detrended globally and annually averaged surface air temperature simulated by CMIP2+ models and observed by Jones / IPCC. The curves are scaled so that the areas under them (if plotted linearly) equal the total variances about the mean of the detrended time series. The 95% confidence interval—based only on uncertainties due to finite sample size—is the same for all cases with the logarithmic scales used in this graphic. The models exhibit global variability that is typically less than observed, as expected for control run simulations.

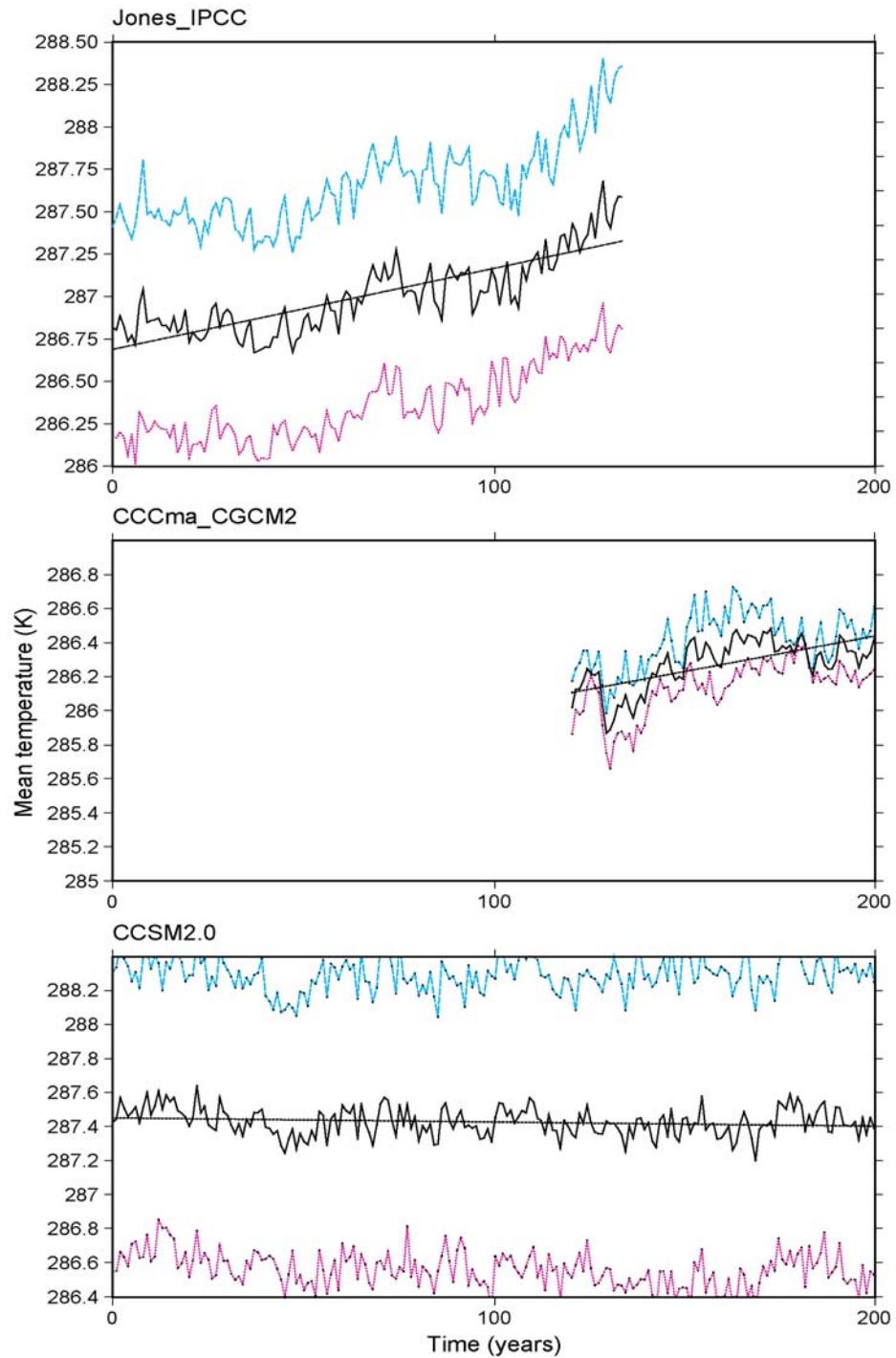


Fig. 3.2. Time series of annual mean surface air temperature from the last 200 years of CMIP2+ control runs integrations and observed by Jones / IPCC. Blue curves are Northern Hemisphere means, pink curves are Southern Hemisphere means, and black curves are global means. Black straight lines are linear least-square fits to global means. The models typically exhibit smaller trends than observed, as expected for control run simulations.

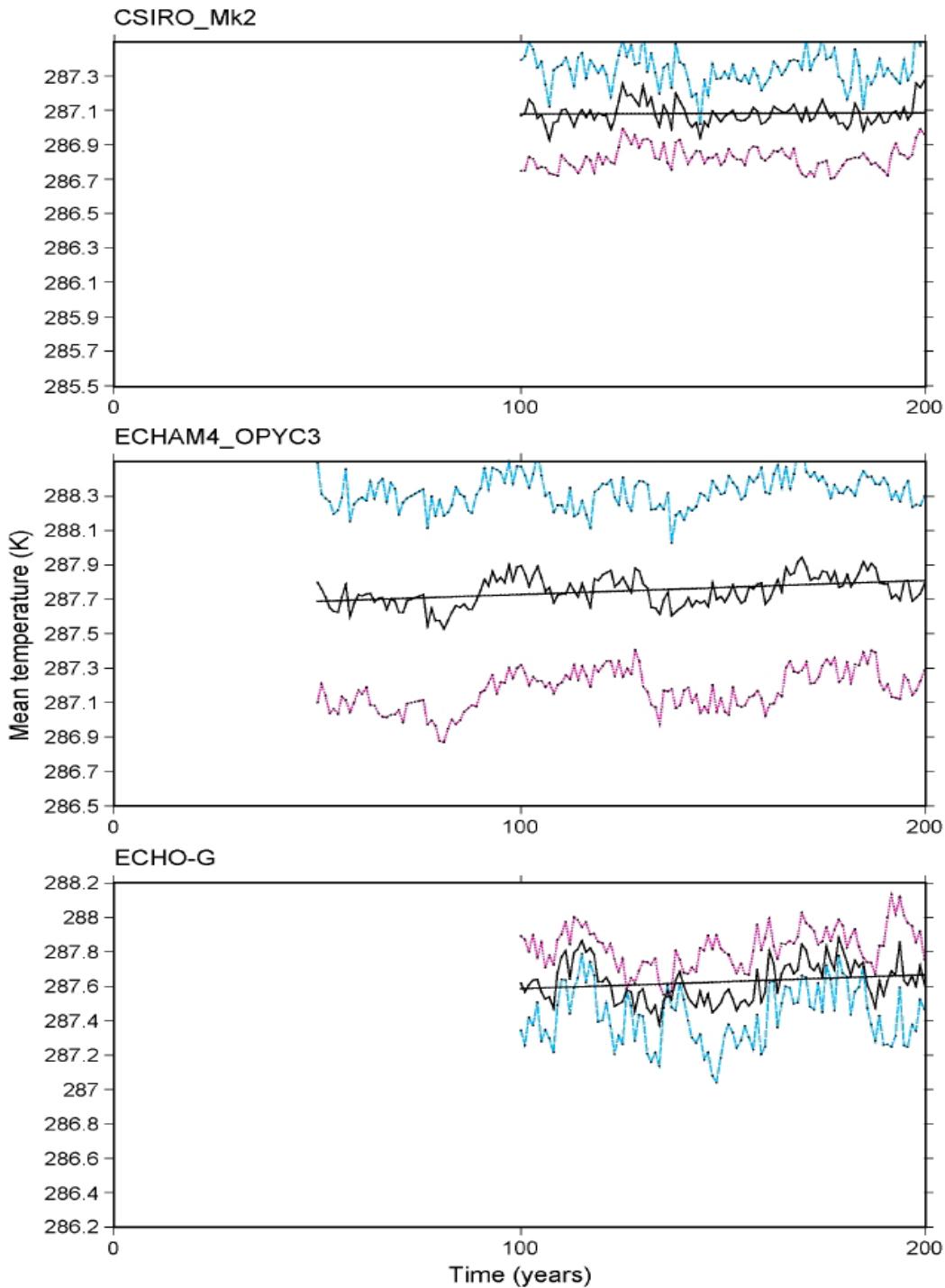


Fig. 3.2 (cont.). Time series of annual mean surface air temperature from the last 200 years of CMIP2+ control runs integrations and observed by Jones / IPCC. Blue curves are Northern Hemisphere means, pink curves are Southern Hemisphere means, and black curves are global means. Black straight lines are linear least square fits to global means. The models typically exhibit smaller trends than observed, as expected for control run simulations.

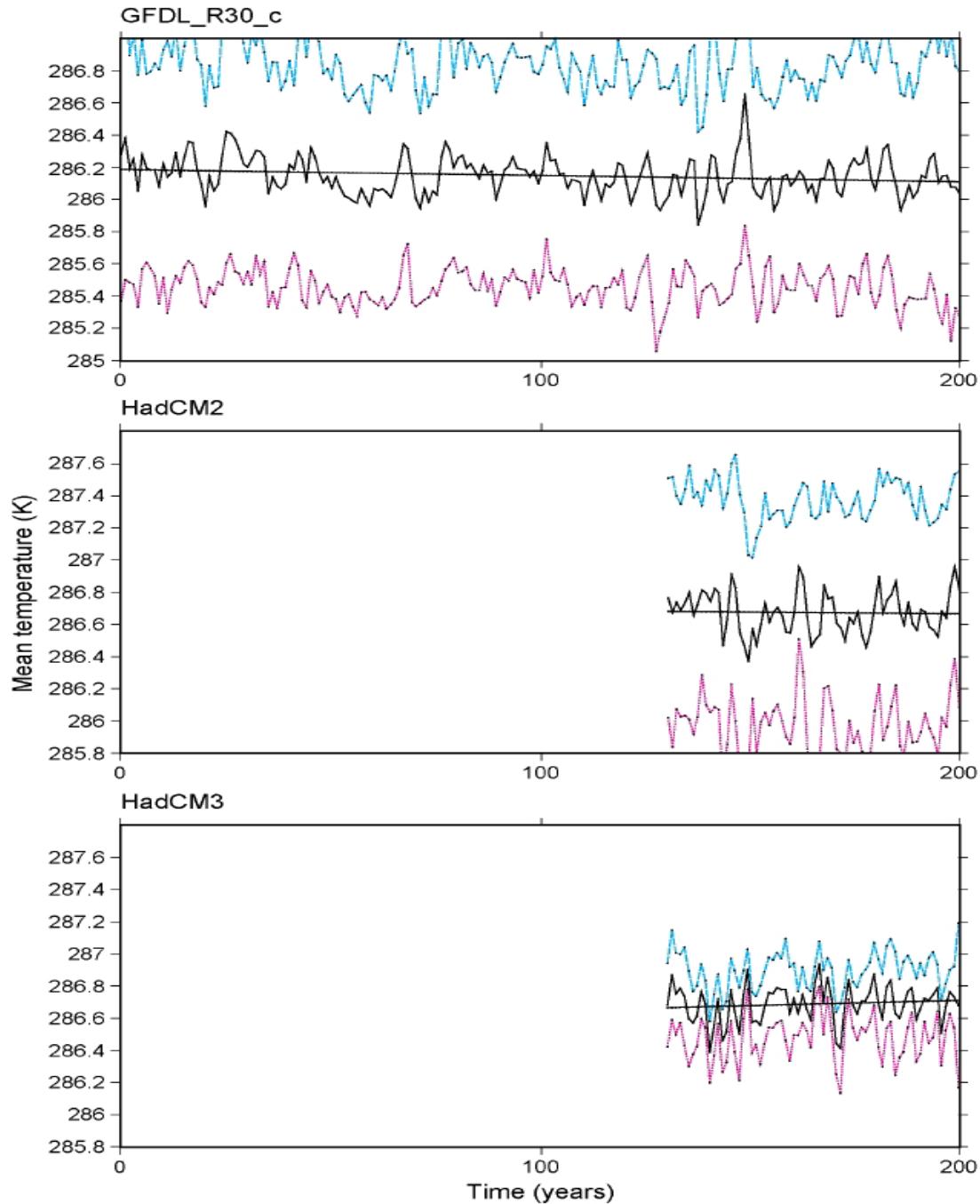


Fig. 3.2 (cont.). Time series of annual mean surface air temperature from the last 200 years of CMIP2+ control runs integrations and observed by Jones / IPCC. Blue curves are Northern Hemisphere means, pink curves are Southern Hemisphere means, and black curves are global means. Black straight lines are linear least-square fits to global means. The models typically exhibit smaller trends than observed, as expected for control run simulations.

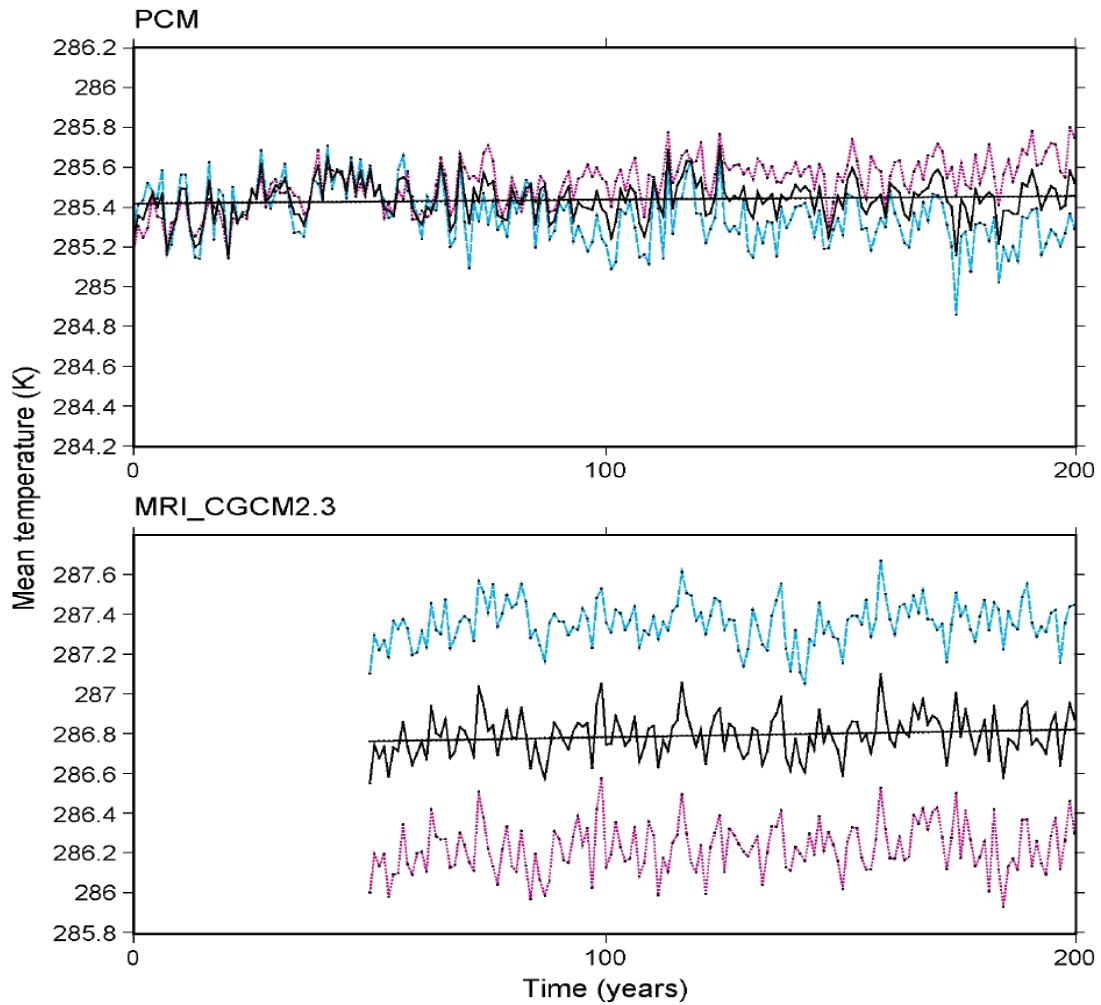


Fig. 3.2 (cont.). Time series of annual mean surface air temperature from the last 200 years of CMIP2+ control runs integrations and observed by Jones / IPCC. Blue curves are Northern Hemisphere means, pink curves are Southern Hemisphere means, and black curves are global means. Black straight lines are linear least-square fits to global means. The models typically exhibit smaller trends than observed, as expected for control run simulations.

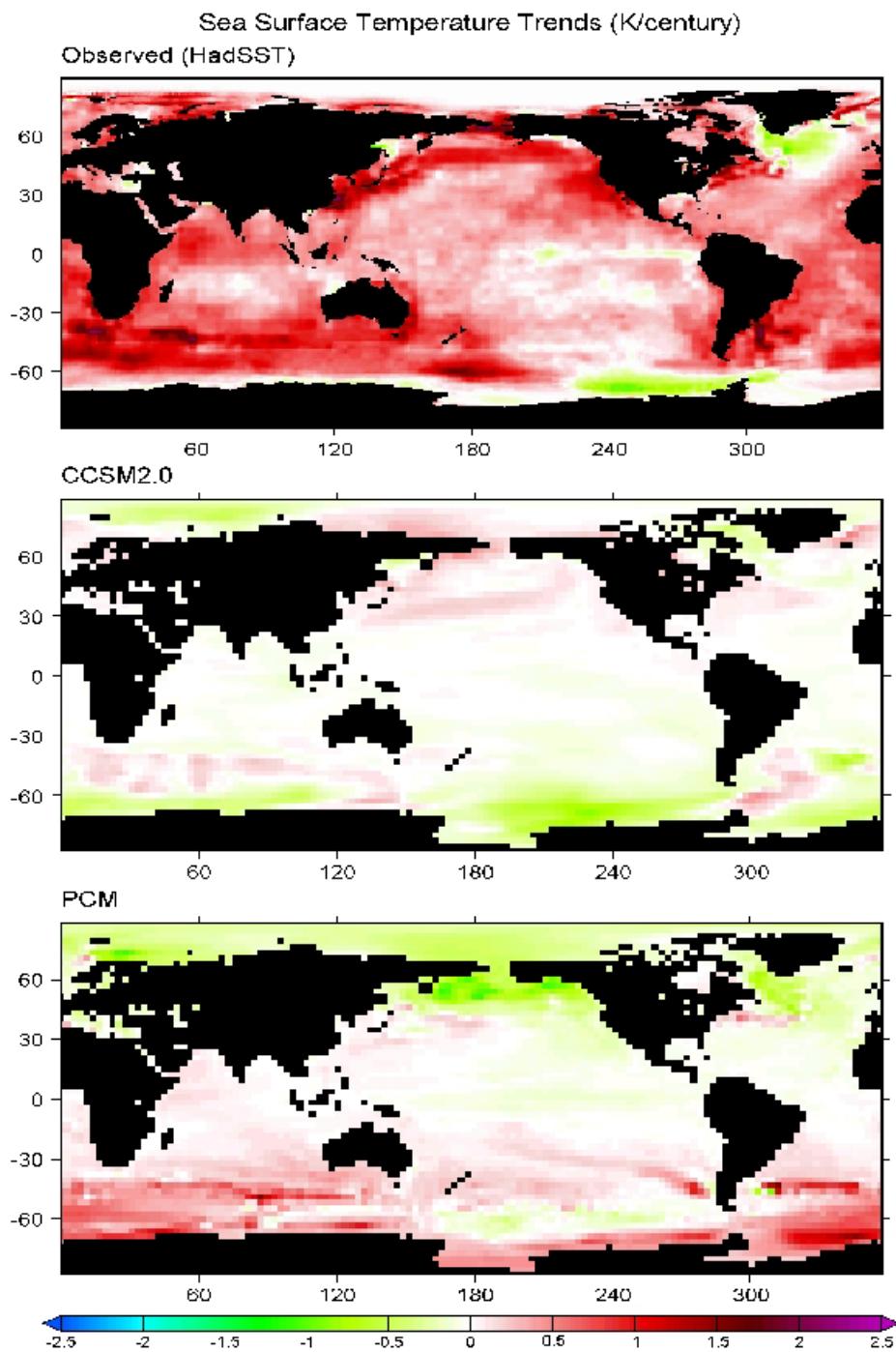


Fig. 3.3. Sea surface temperature trends observed for the period 1898-1997 (HadSST database) and simulated by two of the three CMIP2+ models that are not flux-adjusted. Results of globally averaging these trends are given in Table 3.1 for the results shown in this figure and for all other CMIP2+ models. All the models exhibit trends that are considerably smaller than observed, as expected for control run simulations.

### **3.1.4 Annual and Global Mean Oceanic Temperature, Salinity and Sea-Ice Extent Trends**

Figure 3.4 shows total ocean volume averaged salinity over time for the CCSM2.0 as well as for the other two CMIP2+ models that were not flux adjusted, HadCM3 and the PCM. In all three models, the change in averaged salinity over  $\sim 100$  years or more is exceedingly small. The HadCM3 exhibits a change of only 0.0005 ppt over 80 years despite the fact that its CMIP2+ output was taken from the beginning of its control run. Both the CCSM2.0 and PCM exhibit considerably smaller rates of salinity drift after the first few decades of their control runs. All three rates are much less than the steady 0.04 ppt/century increase reported by Bryan (1998) for the CSM1. Similarly encouraging results (not shown) are evident for total ocean heat content. Long-term trends of this quantity are in the range  $1\text{--}5 \times 10^{23}$  J / century for the three CMIP2+ non-flux-adjusted models. This rate of change may be somewhat greater than observed ( $\sim 10^{23}$  J over 40 years, according to Barnett et al., 2001) but amounts to no more than 0.03% of total ocean heat content per century.

For the rest of this section we continue our examination of trends in the CMIP2+ control runs by looking at ocean temperature and salinity as a function of depth. We restrict ourselves to the first 80 years from each simulation in the PCMDI data base for several reasons. First, 80 years was the minimum amount of control run data requested for CMIP2+ and therefore all that was provided by some groups. Another reason is that the evaluation of the simulated mean climate in Sections 4 (atmosphere) and 5 (ocean) are based on 20-year climatologies computed from years 60–79 (relative to the beginning of the portion of the integrations provided to PCMDI).

Figure 3.5 shows the anomaly of the globally and annually averaged ocean temperature as a function of depth, with respect to the first year in the time series. In most cases the departures from the initial year are very small (less than 0.05 C), suggesting that the simulations have reached a state of quasi-equilibrium. Although the BCM02, CCCma\_CGCM2, and PCM reveal some drift at intermediate depths, our examination of longer portions of these runs (not shown) suggests that they too have largely achieved quasi-equilibrium by years 60–79 (the climatology period of this appraisal).

In Section 5 we will demonstrate that basin-by-basin, the mean climate temperature and salinity of the models are in qualitative agreement with one another and the Levitus data set. The Arctic Ocean is, however, an important exception. Because of this, and the fact that the Arctic Basin is uniquely isolated, results analogous to Fig. 3.5 are shown in Fig. 3.6, but this time averaged over the Arctic Ocean. While some of the variability seen in these plots can be characterized as secular trends, decadal scale variations are also evident, e.g., the intermediate waters of GFDL\_R30, HadCM3 and ECHO\_G.

Figures 3.7 and 3.8 are salinity anomaly analogues to the temperature Figs. 3.5 and 3.6. In the global case (Fig. 3.7), the departure from the first year is quite small for most models (less than 0.005 ppt), especially below 1000 m. In relative (percentage) terms, salinity anomalies are small compared to that of temperature. Changes in Arctic Ocean salinity are also relatively small, but intriguing nonetheless. We can make no claims whatsoever concerning the realism of these basin-averaged salinity variations. These plots do, however, suggest that our choice of a climatology period is not very sensitive to salinity variations.

In Fig. 3.9 we take a cursory look at the stability of sea-ice extent in several of the CMIP2+ models. The Northern and Southern Hemisphere plots depict 5-year running means of percentage changes in the total sea ice surface area with respect to the 80-year time mean (the same 80 years shown in the temperature and salinity plots above). None of the models exhibit changes of more than 15% in total surface area during the 80-year period, and at the end of this period all of them are within 5% of the 80-year time mean. Only several of the models are shown because of problems in the data (or lack thereof) in some of the models.

As noted throughout this appraisal we must reiterate that, in general, control run output from the models used here comes from different control runs time segments (see Table 2.2). Conclusions concerning the relative stability of these models should not be drawn on the results of this section. The analysis does, however, put into context the results of Sections 4 and 5, with the important point being that all models are relatively stable during the period chosen to examine the mean climate.

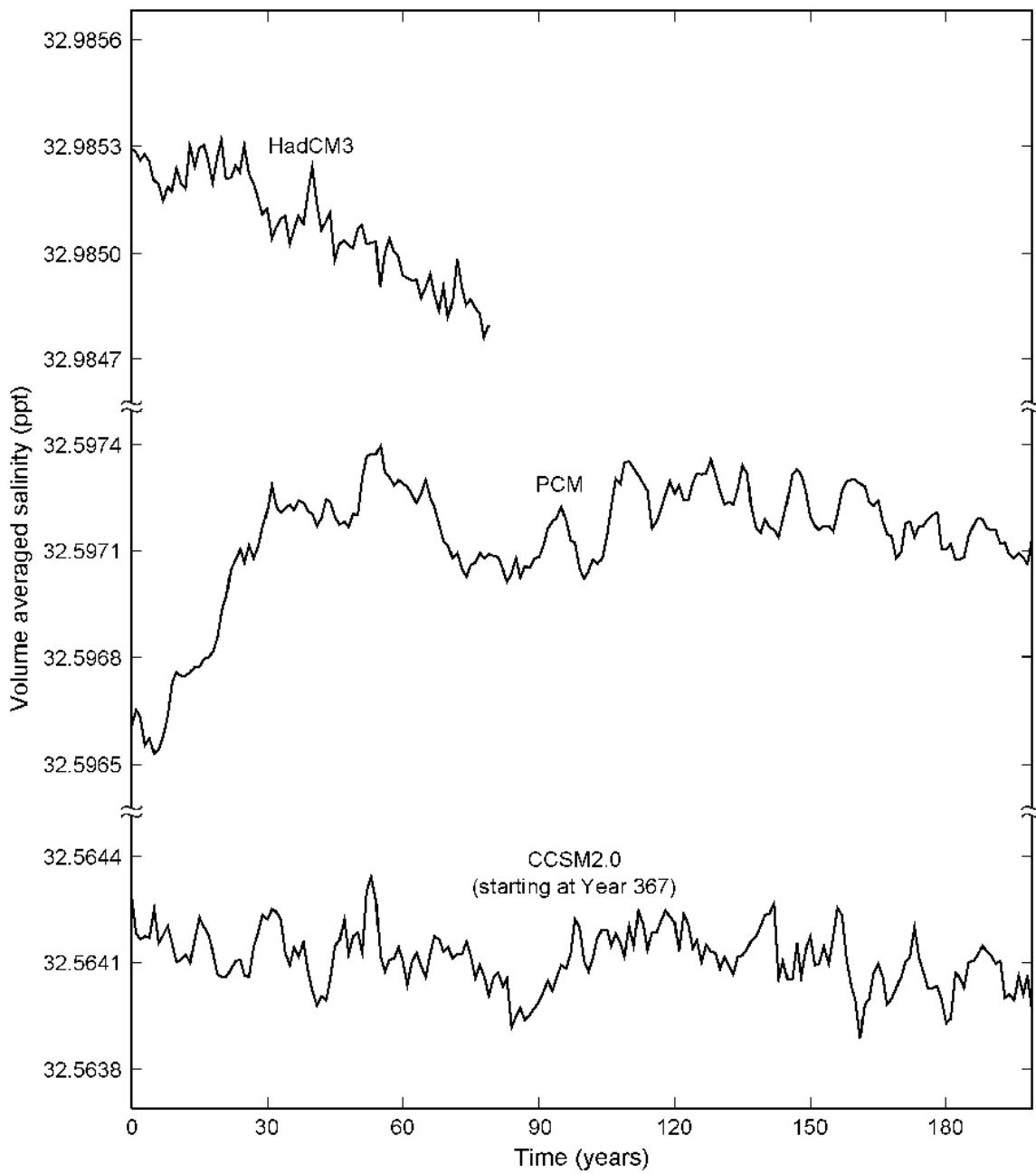


Fig. 3.4. Average salinity over total ocean volume for the three CMIP2+ models that were not flux-adjusted.

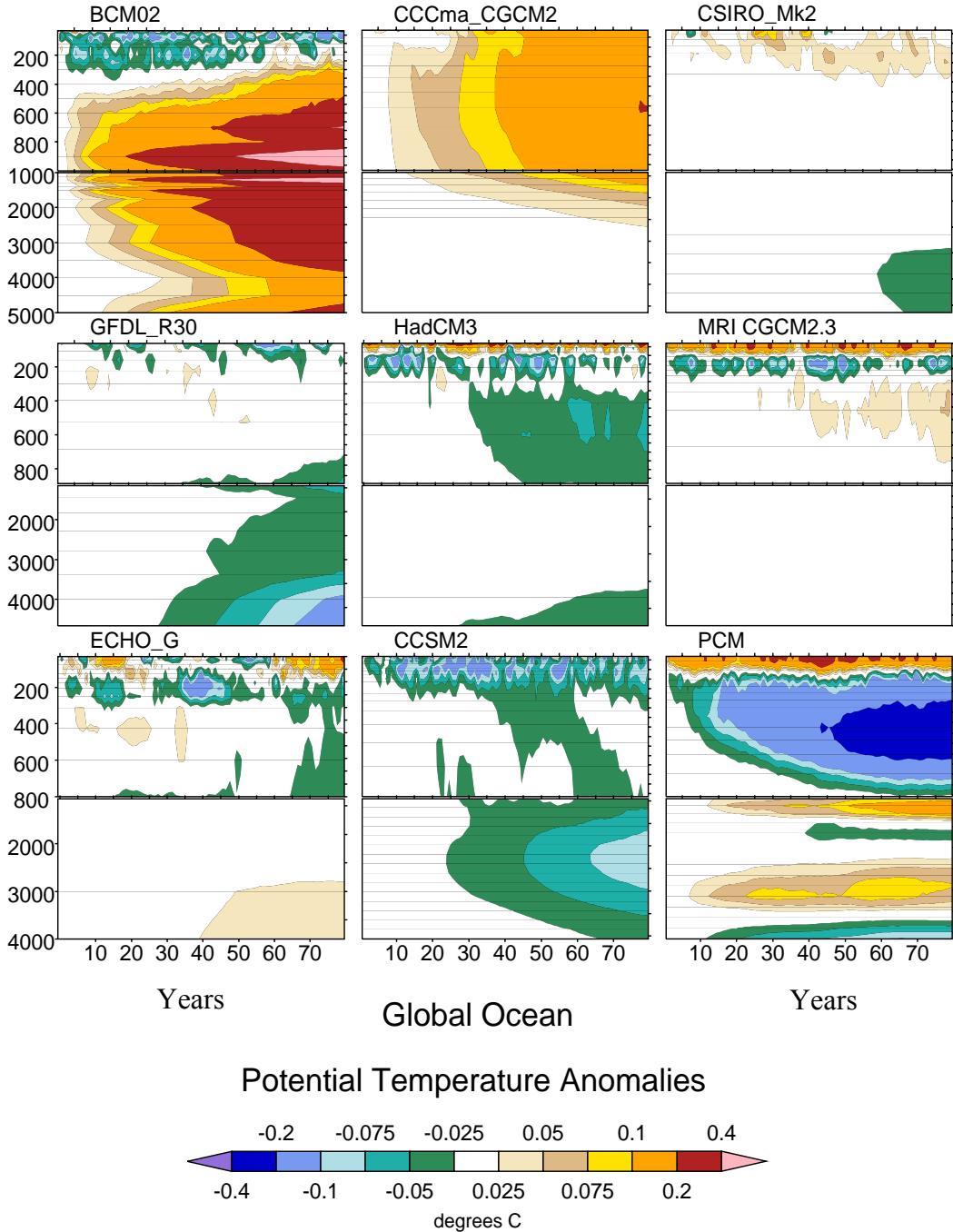


Fig. 3.5. Deviation of the globally and annually averaged ocean temperature from the first year (of data archived at PCMDI) as a function of depth (degrees C). The three models that are most suggestive of trends (BCM02, CCCma and PCM) are found to approach quasi-equilibrium later in the integrations (not shown). In each case, years 60–79 are found to be periods of relative stability, and are therefore used to derive climatologies for the mean climate Sections 4 (atmosphere) and 5 (oceans).

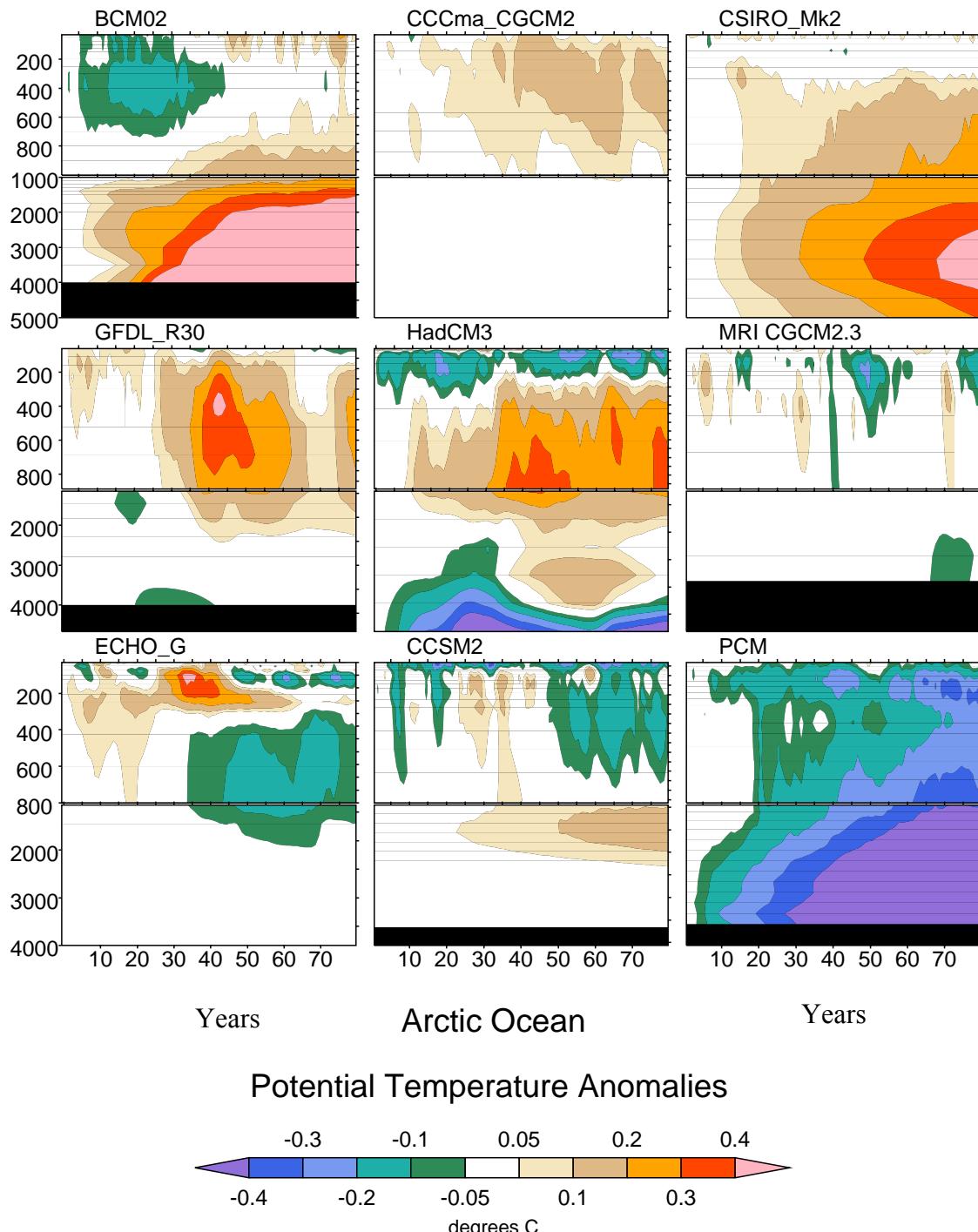


Fig. 3.6. Deviation of the Arctic Basin annual mean temperature from the first year (of data archived at PCMDI) as a function of depth (degrees C).

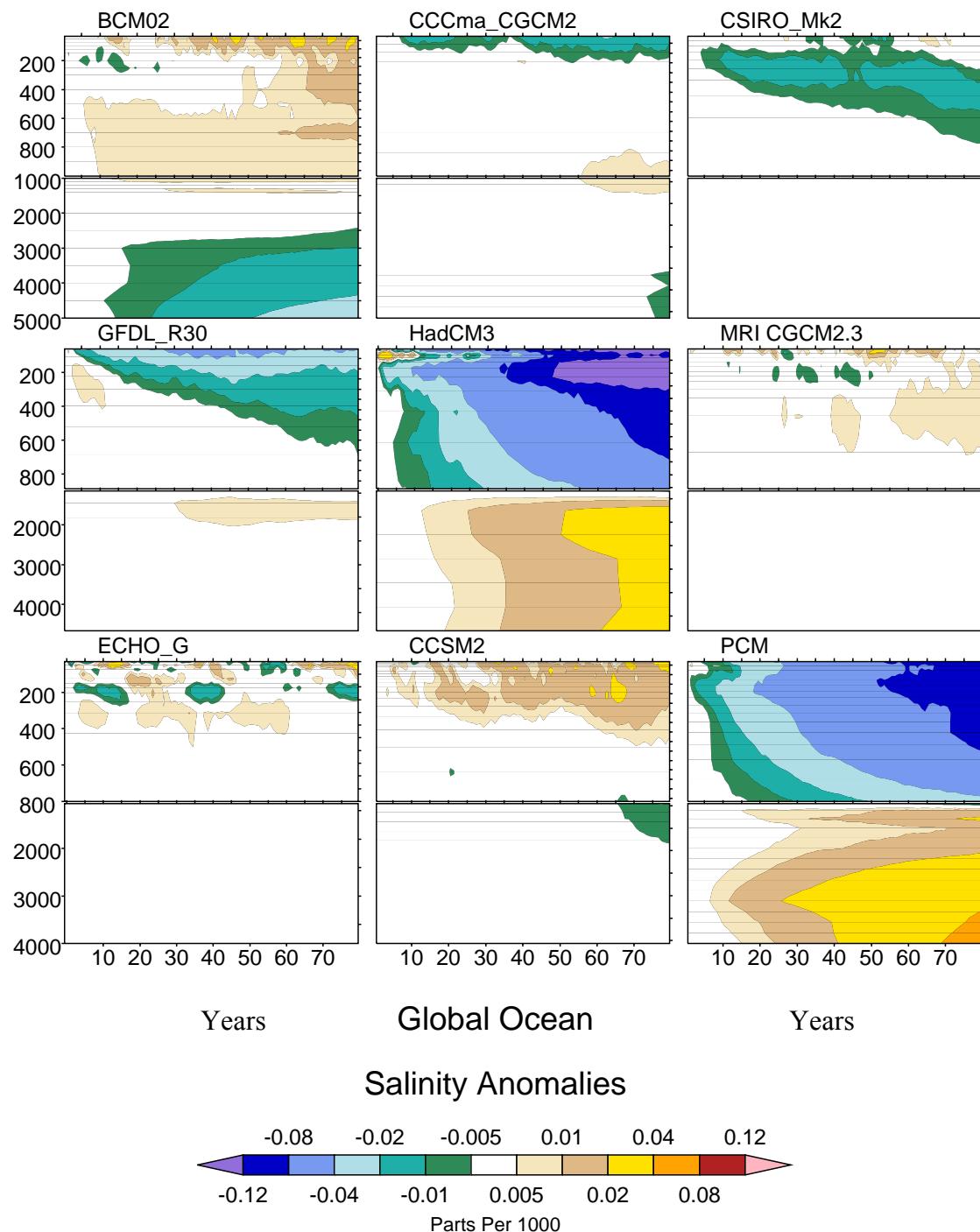


Fig. 3.7. Deviation of the globally and annually averaged ocean salinity from the first year (of data archived at PCMDI) as a function of depth (ppt).

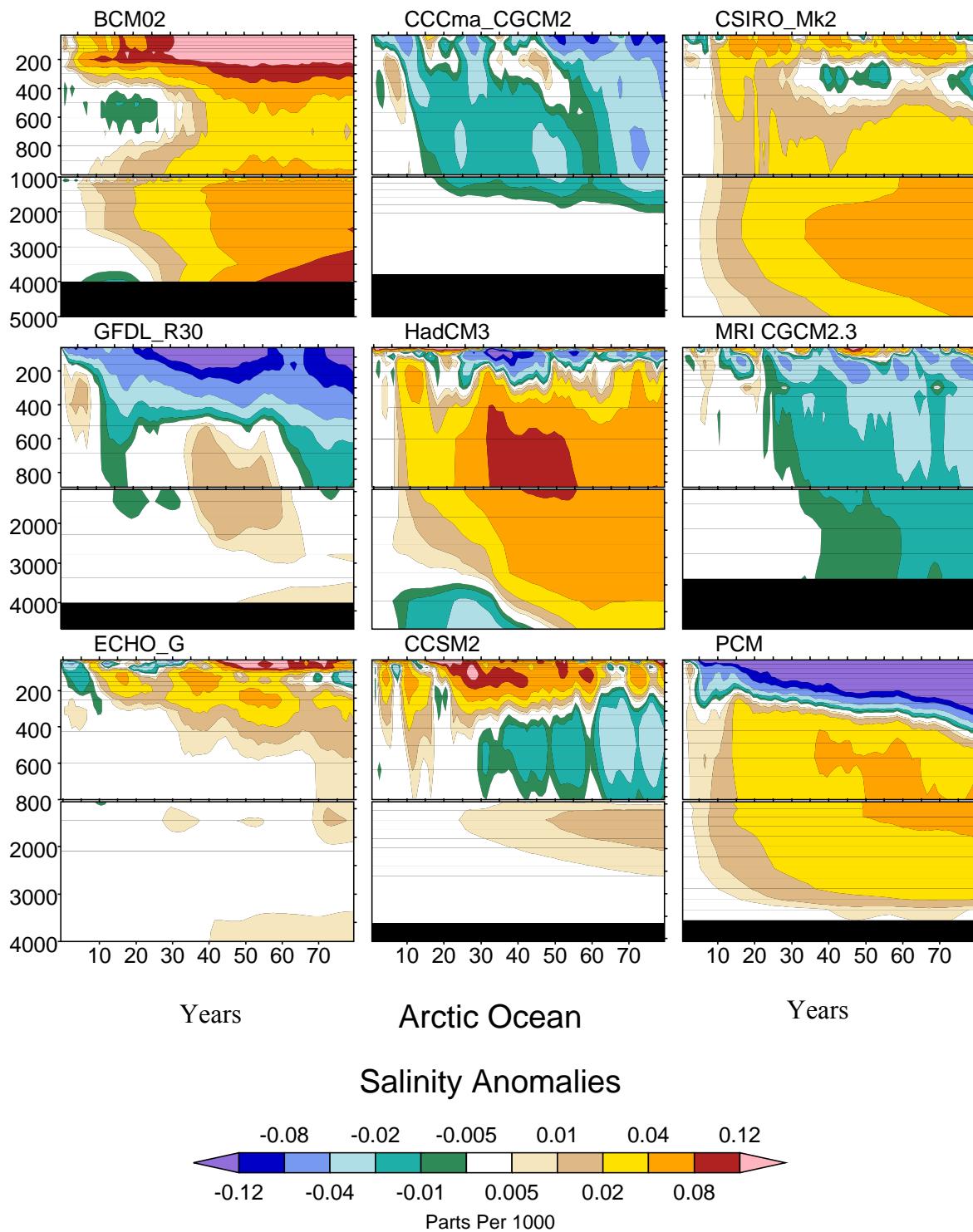


Fig. 3.8. Deviation of the Arctic Basin annual mean ocean salinity (ppt) from the first year (of data archived at PCMDI) as a function of depth (m).

## Deviations in the Annual Mean Total Sea Ice Coverage

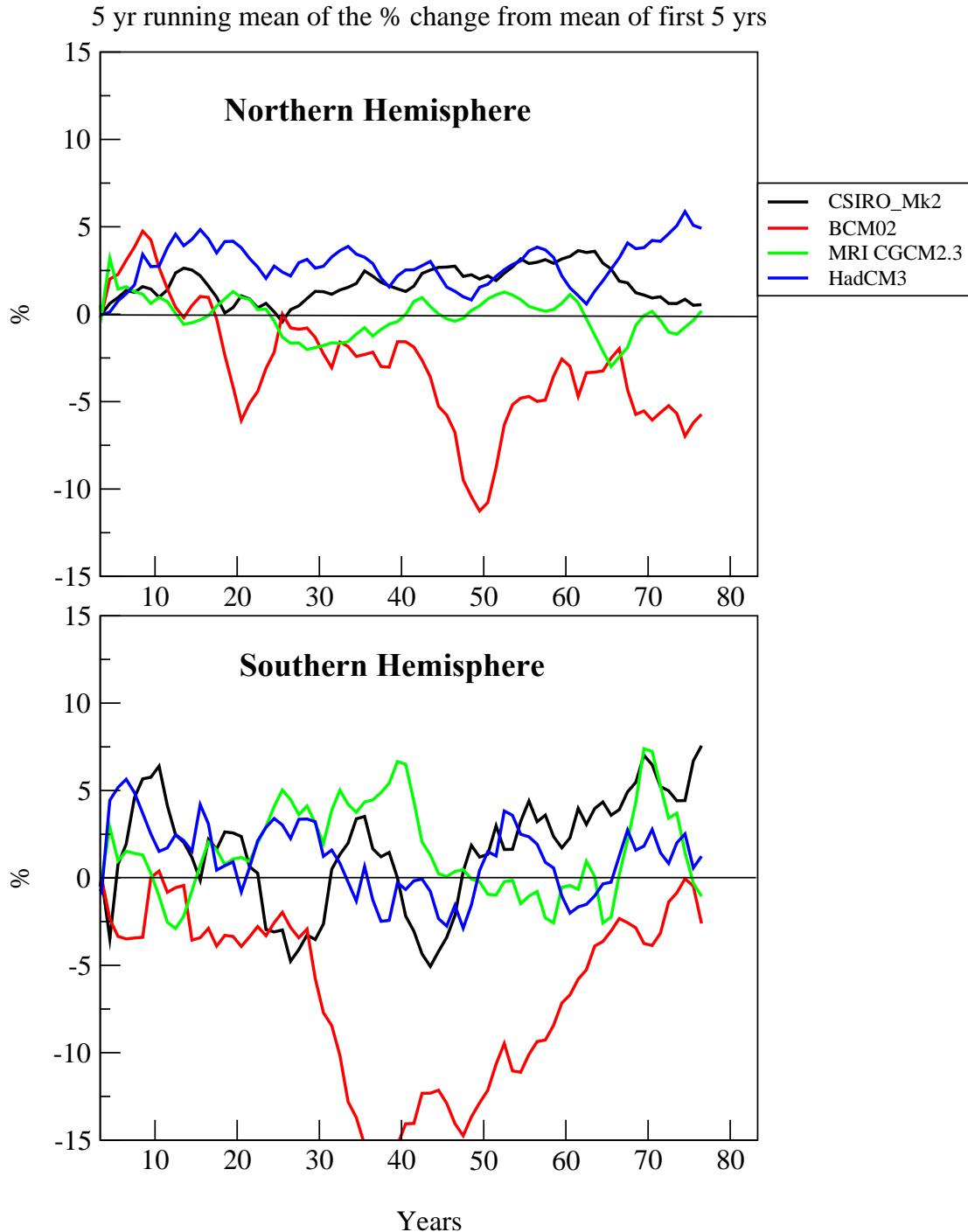


Fig. 3.9. Northern and Southern Hemisphere 5-year running means of the percentage changes in the total sea ice surface area with respect to the mean of the first 5 years.

### 3.2 CMIP 1%/yr runs

We now shift our attention from control run simulations to an idealized scenario of global warming. In this scenario, the atmospheric concentration of carbon dioxide increases at a rate of 1% per year, doubling after about 70 years. Starting concentrations of atmospheric CO<sub>2</sub> vary among the models but are typically set at late 20<sup>th</sup> century values. All other climate forcing factors (other greenhouse gases, aerosols, etc.) remain constant in these simulations. This scenario is obviously a highly idealized one and should not be considered a realistic projection of the future, except perhaps in an order-of-magnitude sense. Nevertheless it is useful to examine the response of the models to this common forcing.

Figure 3.10 shows the global and annual average difference between 1%/yr increasing CO<sub>2</sub> runs and control runs for surface air temperature (top) and precipitation (bottom). As in the CMIP2 overview, a gradual increase in both quantities is apparent together with a spread over the models in the rates of increase, particularly in the case of precipitation. For surface air temperature, the spread among the CMIP2+ models shown here is somewhat greater than the spread found in the CMIP2 overview. The main difference from the CMIP2 results is the addition of one or two new models that obtain global mean warming at the high end: ~3 K after 80 years of CO<sub>2</sub> increase. Still, the general results shown here are consistent with both the CMIP2 overview and many other studies over the years (e.g., Fig. 9.3 in Cubasch et al. 2001). Global mean warming at the time of CO<sub>2</sub> doubling is in the range 1–3 K, and the associated global mean precipitation increase can be anywhere from near zero to ~0.15 mm/day (about a 5% increase from the present day value) depending on the model. These results emphasize the continuing significant uncertainty in model projections of future climate that arise from model formulation.

To exhibit the seasonal and latitude-longitude structure of these temperature and precipitation changes, Fig. 3.11 shows model means around the time of CO<sub>2</sub> doubling. The corresponding annual mean results for the CMIP2 overview are presented in that document's Figs. 2 and 4. In the model mean, poleward amplification of global warming is evident in Northern Hemisphere winter at northern high latitudes (upper left of Fig. 3.11) but is not nearly as prominent in Southern Hemisphere winter (lower left of Fig. 3.11). The CMIP2 overview found that high-latitude amplification of the warming was common to most of the models in the Northern Hemisphere. In Southern Hemisphere high latitudes, however, the CMIP2 results indicated that the warming was weaker and less consistent among models. The same conclusion may follow from the CMIP2+ results shown here. Indeed, the model mean warming amounts in Fig. 3.6 are actually negative—indicating cooling—at one or two locations off the coast of Antarctica. Model mean changes in precipitation (right side of Fig. 3.11) include increases at higher latitudes and complicated patterns of increase and decrease at lower latitudes. It is worth remembering from the CMIP2 overview that only the high-latitude precipitation changes were consistent across models.

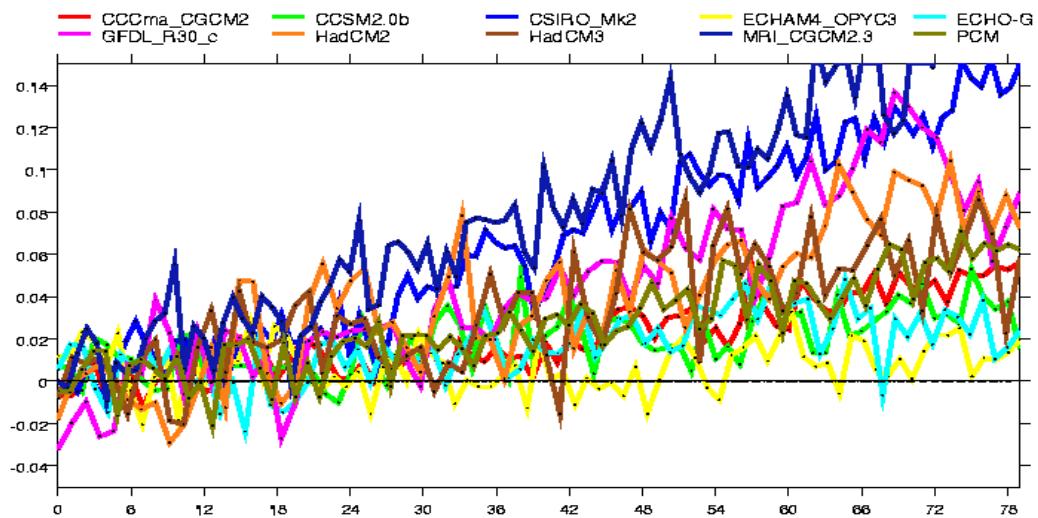
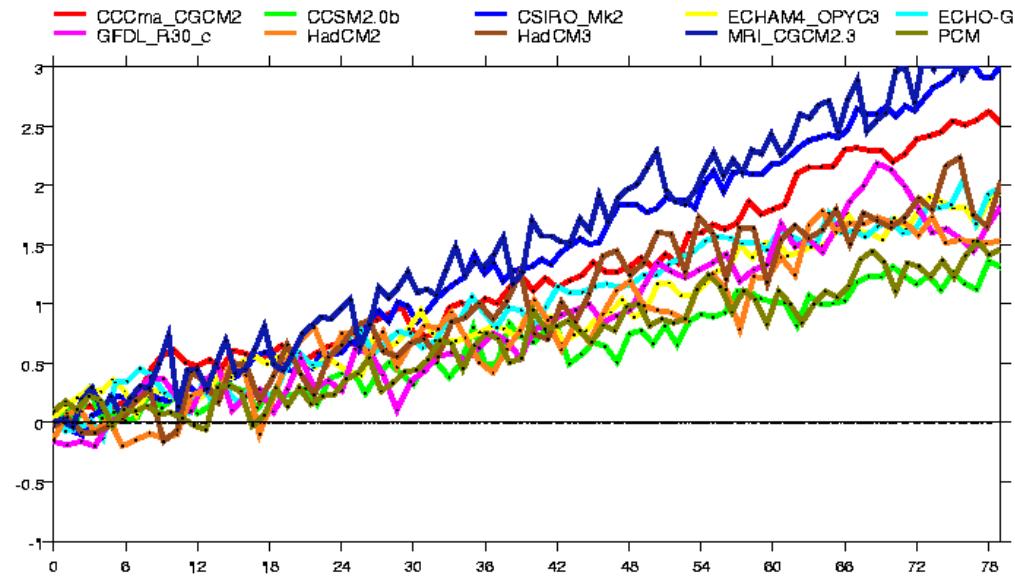


Fig. 3.10. Differences between increasing atmospheric carbon dioxide scenarios and model control runs in the global and annual mean, for surface air temperature in K (top) and precipitation in mm/day (bottom), as a function of time in years.

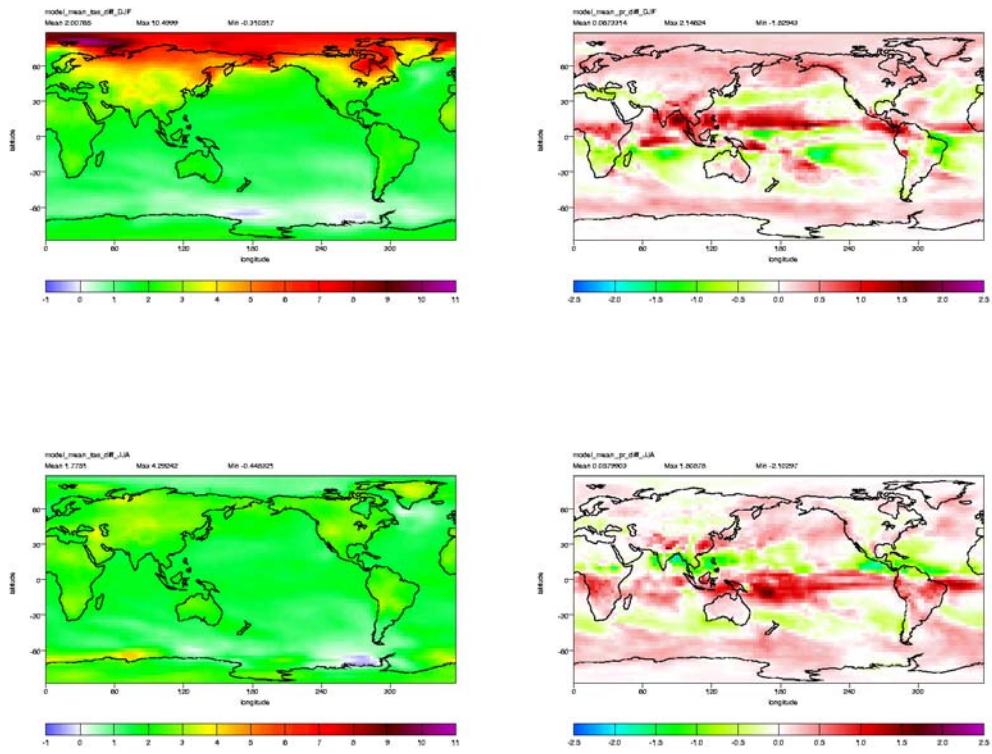


Fig. 3.11. Differences near the time of  $\text{CO}_2$  doubling between increasing atmospheric carbon dioxide scenarios and model control runs, averaged over all models, for the December–January–February season (top row) and the June–July–August season (bottom row). The left-hand column shows surface air temperature differences in K; the right-hand column shows precipitation differences in mm/day.

### **3.2.1 Summary**

The typical first application of a new GCM is a long control run simulation in which climate forcing factors (from solar brightness, atmospheric carbon dioxide, etc.) are held constant. Traditionally, one compares long-term means from control run output with observations in the hope of obtaining close agreement—as we do in the bulk of this report. In the real world, however, climate forcing changes over time. Such changes include anthropogenic increases in both aerosols and carbon dioxide and other greenhouse gases, as well as natural variations in the Sun’s output, volcanic eruptions, and many other quantities. Thus a well-behaved model control run should exhibit *smaller* than observed trends of surface temperature and other climatic variables. Quite the opposite situation prevailed in the early days of climate modeling. Coupled ocean-atmosphere GCM simulations drifted relatively quickly and steadily unless constrained by nonphysical flux adjustments (and in some cases did so even with flux adjustments). In recent years the situation has improved dramatically. This improvement was documented in the most recent IPCC assessment report and is confirmed by the results given above. Although most of the CMIP2+ models employ flux adjustments, both the flux adjusted and the non-flux adjusted models exhibit acceptably small “climate drift” for century-scale simulations.

As noted in the CMIP2 overview, small rates of climate drift at the surface do not rule out the existence of problematic long-term drift in the deep ocean, which could threaten the viability of model simulations carried out for much longer than a century. The CMIP2+ ocean trend results, however, give at least preliminary cause for optimism in this regard. With the exception of the Arctic basin, deep ocean temperature and salinity trends are quite small over 80 years or more. The fact that several of the CMIP2+ models have been integrated for 1000 simulated years without egregious problems (e.g., Kiehl and Gent 2004) further attests the stability of modern coupled climate models, even in the absence of flux adjustments.

Finally, a brief examination of forced climate trends from a common scenario of 1% per year increasing atmospheric carbon dioxide shows (not surprisingly) that the CMIP2+ models respond in similar fashion to the earlier CMIP2 models. Differences among models in the amount of global warming are substantial, spanning about a factor of 2 by the time of CO<sub>2</sub> doubling. Differences in the precipitation response are even more striking and include, for a few models, small global mean change.



## 4 Atmospheric Climatology and Annual Cycle

*P. Gleckler, K. Sperber, M. Fiorino, K. Taylor*

### 4.1 Overview

The coupled models evaluated in this report are compared with available observationally based estimates. Our objective is to examine these models collectively, and to highlight their strengths and weaknesses as a whole. The control simulations used here were all at least 80 years in length. Except where noted the analyses herein are based on 20-year climatologies computed from years 60–79 (relative to the beginning of the portion of the integrations provided to PCMDI). It should be noted that the simulations provided to PCMDI were spun-up for various lengths of time, and thus in general the climatologies are not calculated from the same point in each of the integrations. We believe this inconsistency is acceptable because as shown in Section 3, the “drift” in these control simulations is small. Furthermore, tests have demonstrated that our results (for the mean climate) do not differ quantitatively if we choose another averaging period (e.g., years 10–29).

We will limit our evaluation to fields that most modeling groups were able to provide, and for which there are useful observations available for reference. As discussed in Appendix 1 (Observations), it is important to note that the accuracy of the observational estimates depends very much on the quantity being measured (or derived). For some fields (e.g., total precipitation and ocean surface fluxes), the accuracy of the measurement also varies greatly by location.

Our analysis of the simulated mean climate will focus on global and large scales, although several examples at the regional scale are included. Traditional seasonal maps and zonal averages are shown for selected fields. Harmonic analysis (e.g., Hsu and Wallace, 1976) is employed to provide additional insight on the progression of the seasonal cycle. Error statistics are used to gauge overall differences between simulated and observed fields, and are summarized in Taylor diagrams (see Appendix 2). Regional-scale area-averaged time series are included to illustrate key seasonal processes (e.g., the Southwest U.S. monsoon and Nordeste rainfall).

In Sections 4.2-4.3 we will highlight the results of a *multi-model ensemble mean model*. Our objective in this section is to evaluate the collective performance of coupled climate models, not specifically any individual model. PCMDI is providing analogous results directly to modelers that do highlight their model (in the context of others), and is planning to establish an electronic report series dedicated to the evaluation of individual simulations.

Although the use of a mean model can sometimes be misleading because outliers may unduly influence the result, this does not appear to be the case here. As demonstrated by a Taylor diagram<sup>1</sup> shown in Fig. 4.1 (upper panel), the multi-model mean and multi-model median global climate patterns for 16 different fields show similar agreement with observations (in terms of correlation, root-mean-square error, and the spatial-temporal standard deviation). Another characteristic of the multi-model mean that will become

---

<sup>1</sup> See Appendix 2 and Taylor (2001) for further explanation.

evident in subsequent Taylor diagrams is that it generally compares better with observations than does any individual simulation, at least in terms of its statistical agreement with the observed mean climate.

In this appraisal we analyze coupled model simulations, and it is not generally possible to determine whether the ocean model or the atmospheric model is primarily responsible for model error. In this section, we provide some evidence that coupled model simulations of present climate compare with observations about as well as stand-alone atmospheric models (run with prescribed sea ice and sea surface temperature, according to the AMIP protocol). Figure 4.1 (lower panel) summarizes these results by considering 16 different climatological fields and comparing model results with the observed climatological cycle. The overall impression conveyed by the figure is that the median CMIP2+ model is neither systematically better nor worse than the median AMIP model. We cannot, however, attribute all the differences to the coupling of an ocean to the atmosphere, since the CMIP and AMIP ensembles comprise different sets of recent atmospheric models.<sup>2</sup>

It is important to note that our CMIP2+ multi-model ensemble includes all models used in this report. It has already been pointed out that this collection is a mix of flux-adjusted and unadjusted simulations. To complicate matters further, five of the models are adjusted monthly, while two others have only annual-mean adjustments applied. The remaining three simulations have no flux adjustment at all (see Section 2).

The results of this chapter are presented as follows: Global-scale characteristics are shown for selected cross-sections of three-dimensional atmospheric variables (Section 4.2), followed by zonal means, Taylor diagrams and maps of two-dimensional variables (Section 4.3). In Section 4.4 the annual cycle of four selected large-scale domains is then examined more closely with harmonic analysis and selected regional time series. An examination of prominent tropical biases frequently seen in coupled model simulations is described in Section 4.5. Finally, the Köppen classification scheme is used in Section 4.6 to complement the analysis of surface air temperature and precipitation.

---

<sup>2</sup> The median model results are computed from all available output for CMIP2+, and model results submitted between 1998 and 2002 for AMIP. The atmospheric models in CMIP2+ are generally not the latest versions available from the contributing modeling centers, and the ensemble of modeling centers contributing results to CMIP2+ are a subset of those contributing recent AMIP simulations.

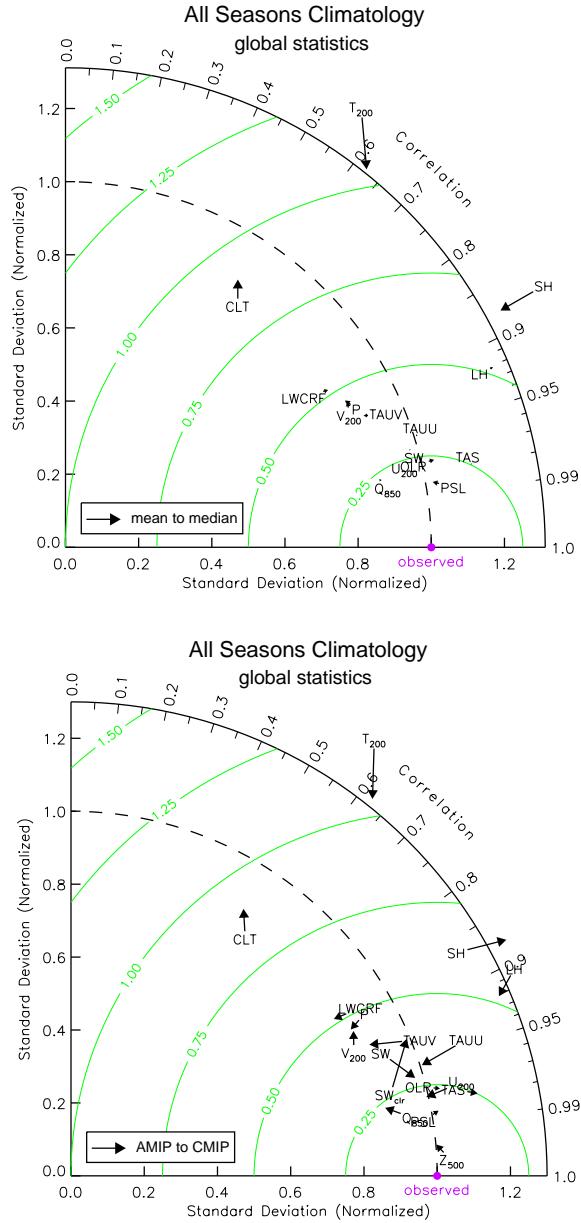


Fig. 4.1. Taylor diagrams (Appendix 2) showing differences in statistical measures of model agreement with observed climatology (Appendix 1) for several different fields. Comparisons are shown for the CMIP mean and median models (upper panel) and the median AMIP and CMIP models (lower panel). The statistics are based on sums computed over the four climatological mean seasons and over all grid cells, weighted by grid cell area. Note that for each field, the standard deviation and root-mean-square error statistics have been normalized by the respective observed standard deviation of the field. The following fields were considered: surface air temperature (TAS), precipitation (P), outgoing longwave radiation (OLR), reflected shortwave radiation (SW), longwave cloud radiative forcing ( $LW_{CRF}$ ), reflected shortwave clear-sky flux ( $SW_{clr}$ ), surface sensible heat flux over oceans (SH), surface latent heat flux over oceans (LH), eastward and northward components of surface wind stress over oceans (TAUU, TAUV), zonal and meridional wind components at 200 hPa ( $U_{200}$ ,  $V_{200}$ ), geopotential height at 500 hPa ( $Z_{500}$ ), mean sea level pressure over oceans (PSL), specific humidity at 850 hPa ( $Q_{850}$ ), and total cloud fraction (CLT).

## 4.2 Zonal Mean Vertical Cross Sections

### 4.2.1 Air Temperature

The simulated air temperature is compared to the ERA15 reanalysis in Fig. 4.2, with the top panel (zonal mean at 200 hPa) also showing the NCEP-NCAR reanalysis. At 200 hPa the models differ considerably among themselves (grey shading), with the well known upper troposphere summertime cold bias in the models being clearly visible in both the 200-hPa zonal means and the difference plots of the zonal mean cross sections. Apart from that, however, throughout the troposphere the multi-model ensemble mean model (hereafter referred to as the mean model) agrees qualitatively with the reanalysis, with differences of less than 2 degrees C outside the cold bias regions.

### 4.2.2 Zonal Wind

The zonal wind is shown in Fig. 4.3, the layout of which is identical to Fig. 4.2 including the reference data sets used (ERA15 and NCEP-NCAR reanalyses). The mean model agrees very well with both reanalyses in the Northern Hemisphere (NH) extra-tropics during the DJF season (i.e., the December, January, February average), with relatively little spread (grey shading) among the models. During this season there is less consistency among the models with the reference data sets in the Southern Hemisphere (SH), where the core of the summertime jet maxima in the models is on the order of 10 m/s stronger in the mean model. The mean model continues to agree well with the reference data in the NH during the austral winter, but again with larger discrepancies in the SH. Apart from the differences seen in the jet strengths, and consistent with the good agreement seen in Fig. 4.2, the mean model captures virtually all the prominent features seen in the observed cross sections. The systematic differences in the extra-tropical stratospheric jet are consistent with the cold bias (Fig. 4.2) via the thermal wind equation.

### 4.2.3 Specific Humidity

The simulated specific humidity is shown in Fig. 4.4 with the top panel zonal means calculated at 850 hPa. The accuracy of specific humidity estimates decreases substantially with height (much more than is the case for temperature or winds), and thus our focus here is more on the lower troposphere. Here we see less agreement between the two reanalyses, particularly in the lower latitudes where the humidity is greatest. At 850 hPa the mean model agrees better with the NCEP-NCAR reanalysis than it does with ERA15 which is moister in the tropics, but note that the differences between the reanalyses are comparable to the inter-model spread. The cross-section differences suggest that the models are too dry in the tropical and subtropical lower troposphere, except perhaps at the lowest levels where the situation is reversed. Elsewhere, comparison with ERA15 suggests that the mean model is too moist. Note that comparison with NCEP-NCAR reanalysis would yield a smaller bias here.

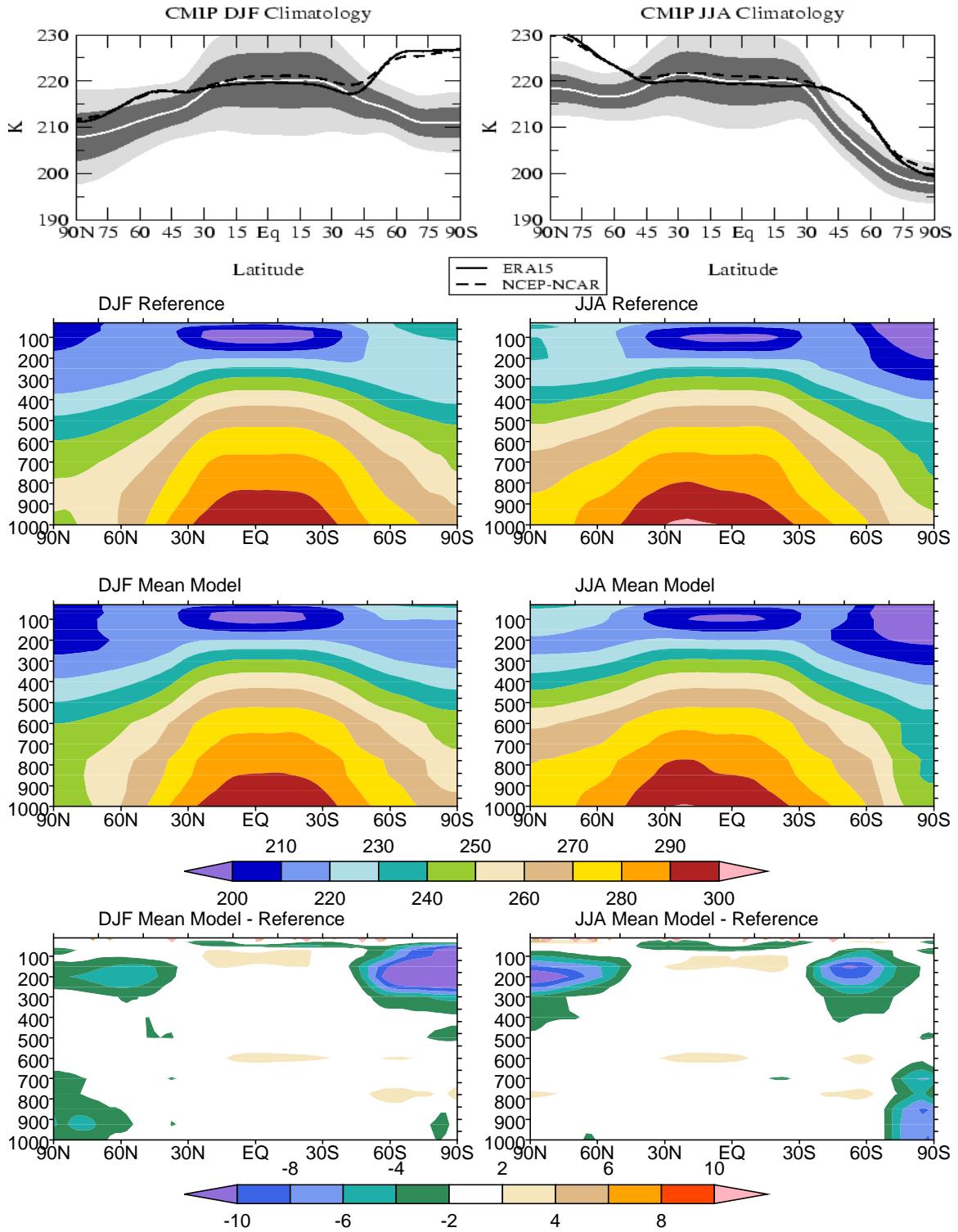


Fig. 4.2. DJF and JJA air temperature. First row: Reference (dark line) and mean model (white line) zonal averages at 200 hPa, with  $\pm 1\& 2$  inter-model standard deviations (shaded); Second and third rows: Reference (ERA15) and mean model zonal cross sections; Fourth row: Mean model–Reference.

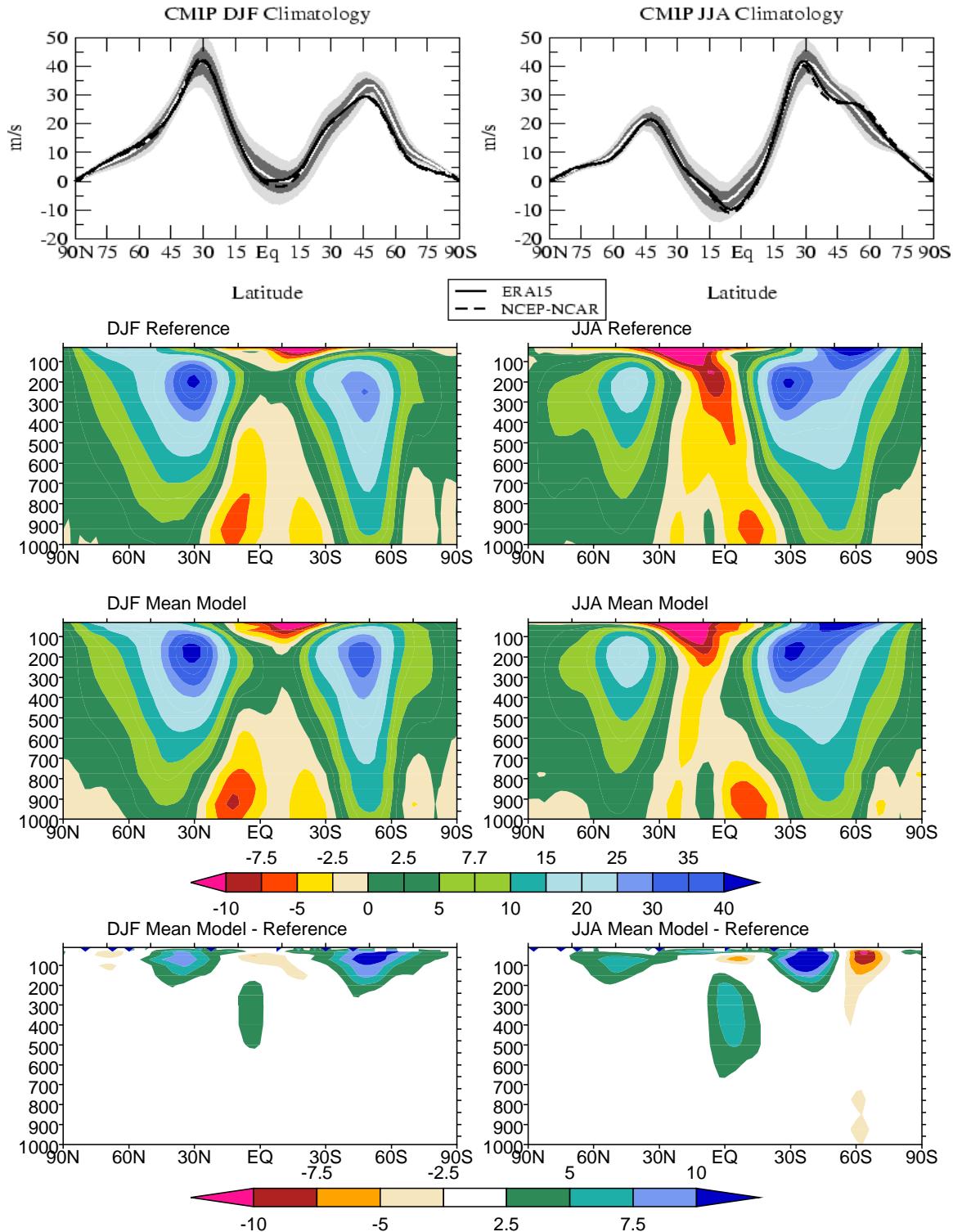


Fig. 4.3. DJF and JJA zonal wind (m/s). First row: Reference (dark line) and mean model (white line) zonal averages at 850 hPa, with  $\pm 1\&2$  inter-model standard deviations (shaded); Second and third rows: reference and mean model zonal cross sections; Fourth row: Mean model–reference.

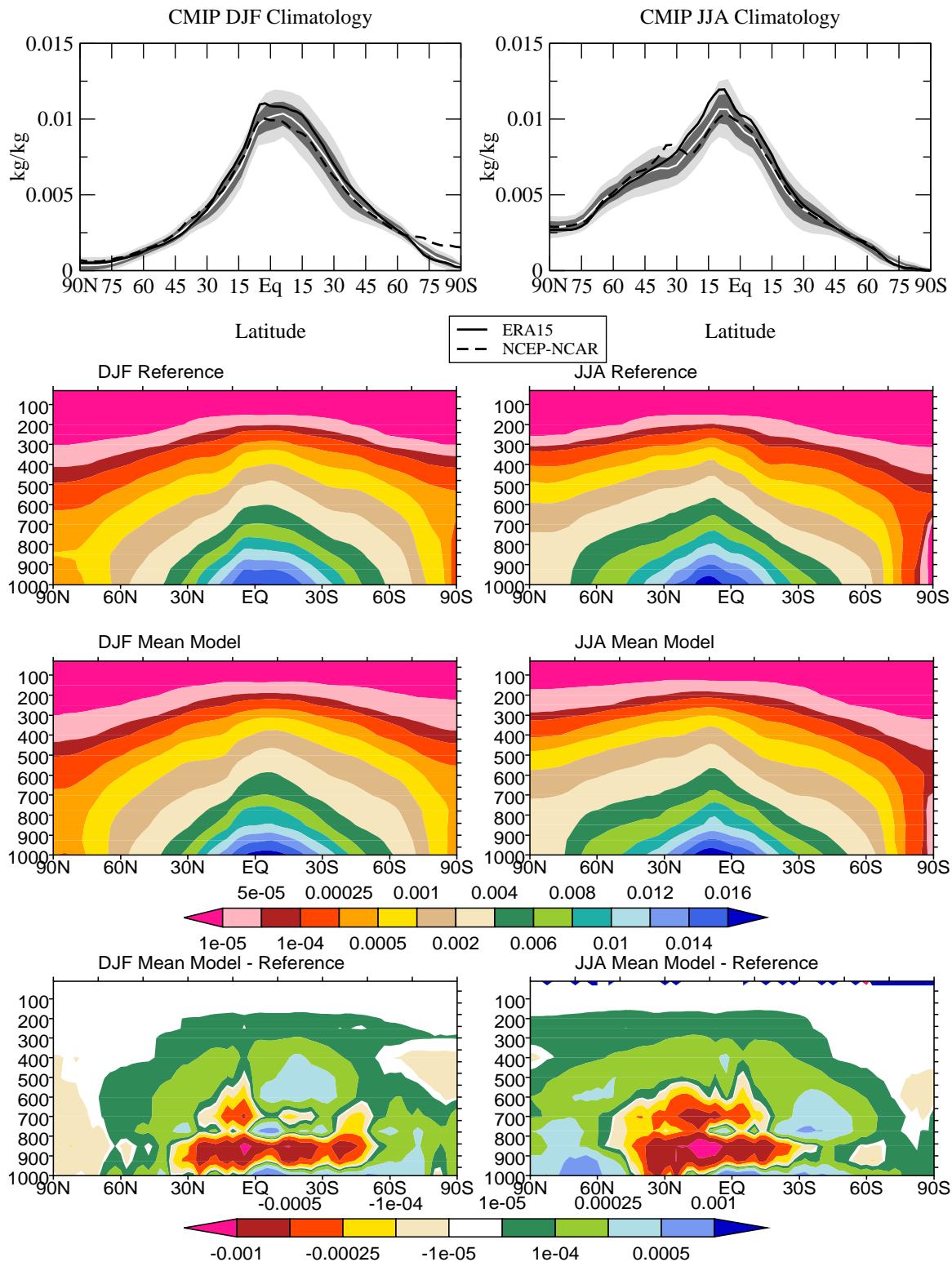


Fig. 4.4. DJF and JJA specific humidity (kg/kg). First row: Reference (dark line) and mean model (white line) zonal averages at 850hPa, with  $\pm 1\&2$  inter-model standard deviations (shaded); Second and third rows: reference and mean model zonal cross sections; Fourth row: Mean model–reference.

## 4.3 Zonal Means, Statistical Plots and Maps

In Figs. 4.5-14 we examine zonal mean and global scale features of surface air temperature, precipitation, outgoing longwave radiation, total cloud cover and mean sea level pressure. Also included are statistical summaries in terms of Taylor diagrams.

### 4.3.1 Surface Air Temperature

The simulated zonal mean surface (2-m) air temperature, shown on the top panel of Fig. 4.5, is compared with Jones' observational data. Note the discontinuous feature of the Jones data at higher latitudes is a result of missing data. The zonal mean surface air temperature of the mean model agrees quite well with the observational estimate, especially equatorward of 60 degrees. However, the observations are a few degrees warmer at all latitudes. We must keep in mind, however, that most of the models used in the mean model are flux-adjusted, and thus over the ocean are constrained to agree well with observations.

In the second row of Fig. 4.5 Taylor diagrams based on the total spatial statistics are shown for the DJF and JJA seasons. These figures distinguish between models that have been flux-adjusted and those that have not. For some of the coupled models, AMIP simulations are available for their atmospheric components, and in these cases arrows have been drawn from the statistics representing the AMIP simulation to the statistics representing the CMIP simulation. It can be seen that in most, but not all, cases the agreement with observations degrades slightly upon coupling (arrows point more or less away from the observed target marked "Jones" along the abscissa). Although an effort has been made to include only AMIP to CMIP results where the atmospheric model is unchanged, definitive documentation is lacking. We suspect that in some cases minor modifications were made to the atmospheric model, so not all of the difference can be attributed to coupling alone. This clearly illustrates why it will be valuable in the future to require that for every new CMIP simulation performed, an AMIP simulation also be carried out without modification of the atmospheric component of the model. Then it will be possible to truly isolate the effects of coupling.

CMIP models without a "sister" AMIP simulation are shown as red or blue dots on the Taylor diagram, and yellow dots on the figure are the statistics from recent AMIP simulations. This shows that CMIP simulations of the DJF and JJA global temperature patterns are generally within (and in some cases better than) most simulations by stand-alone atmospheric models.

Finally, note that on the lower panel of Fig. 4.5, which compares patterns after the zonal mean has been removed, the observed standard deviation is only 1/5<sup>th</sup> as large as the panel above it. This indicates that most of the variance in the spatial patterns is due to the zonal mean pattern (top panel); the longitudinal variations, including land sea contrast, are more evident in the lower panel. The models generally are able to simulate the zonal mean temperature quite well (correlations generally greater than 0.97, not shown), but vary in their ability to simulate longitudinal variations, with correlations ranging from less than 0.7 to more than 0.8.

In the full DJF and JJA comparisons (middle row of Fig. 4.5) we see that the correlation between the observational product (Jones) and simulations is quite high, in most cases at

least 95%. The coupled models without flux adjustment appear to agree slightly better with the observations than the collection of flux-adjusted models. In most cases there is more spatial variance in the models than in the Jones reference. There is more spread between the models when the zonal mean has been removed (last row of Fig. 4.5), but note the green contours indicate that the pattern RMS accounts for only about half of the total RMS in this case, implying that the zonal mean component dominates the overall error. (Recall that the RMS component errors add quadratically to yield the squared total error.)

Global DJF and JJA maps of surface air temperature are shown in Fig. 4.6. Note the white areas at high latitudes in the reference (Jones) dataset represent missing data. In a qualitative sense, the mean model compares extremely well with the observations, as evidenced by comparing the patterns in the first two rows of Fig. 4.6. Difference maps (third row) do however illuminate important differences. Consistent with the zonal means shown in the previous figure, over the oceans the reference dataset is at least 0.5°C warmer and in some areas (e.g., JJA North Atlantic) is between 1.5–2°C warmer. More pronounced differences are seen over land, particularly in mountain areas and high plateaus where the mean model is more than 10°C cooler than the reference data. Conversely, the mean model is warmer than the observations in relatively few areas. The cause of this systematic discrepancy is not obvious, but as mentioned previously it is important to point out that models do not directly solve for air temperature at a height of 2 m, and the methods by which interpolation is performed between the lowest model layer and the surface temperature is not the same for all models. Having interpolated the models to a common (T42) grid, we have calculated the point-by-point inter-model standard deviation to estimate where models are in agreement with one another, and where they are not. This is shown in the bottom row of Fig. 4.6. Over the oceans the majority of the models are within 1°C of one another, but the situation is quite different over land where the spread across models is much greater. Note that apart from the higher latitudes, the inter-model spread is not noticeably larger in areas of high elevation.

### 4.3.2 Precipitation

We turn our attention to the simulated total precipitation in Figs. 4.7–8. As evidenced in the zonal means, there are some important differences between the two observational estimates (Xie-Arkin and CMAP), although they are not independent of one another. Not unexpectedly, the inter-model spread is also much larger than the case of the surface air temperature. With substantial observational uncertainties aside, most of the models (and therefore the mean model) overestimate the mid-latitude precipitation except in the NH summer. In the tropics, the mean model compares well with the data in DJF, but substantially underpredicts the precipitation in JJA. Turning our attention to the statistical plots, we see in the global field DJF and JJA (middle row) figures that most of the models cluster around the dashed black line, suggesting that their spatial variance compares well with the CMAP data set. The pattern correlation of most of the models ranges between 0.7 and 0.9. The results are similar for the case when the zonal mean component has been removed (bottom row), but in this case the reduction in the pattern RMS is not as prominent as was the case for surface air temperature.

As illustrated in all four Taylor diagrams in Fig. 4.7—and a frequently reoccurring theme—the multi model ensemble mean compares better with the observational estimate than any of the individual models (in that the green concentric contours centered about

the CMAP observational point represent the pattern RMS error). However, the mean model does have less spatial variability than the CMAP data set (since it lies well inside the dashed quarter circle), as might be expected from the averaging process that tends to smooth out smaller-scale features. It is interesting to note that there is almost no difference between the AMIP and CMIP multi-model means (AMIP and CMIP respectively being the tail and head of the black arrows) in spite of the apparently large differences shown for individual models, and even though the collection of models used in the two ensembles is quite different. Also note that unlike the surface temperature, the precipitation pattern is not dominated by its zonal mean variance (as deduced from the standard deviation scales in the Taylor diagrams).

The DJF and JJA spatial distributions are shown in Fig. 4.8. It should be re-emphasized that there is significant uncertainty in the precipitation estimates; however, the large-scale patterns are probably more accurately represented than the absolute values. There are encouraging similarities between the patterns of observational estimate and the mean model. However, in the subtropical eastern oceans where there is very little precipitation, the mean model tends to produce too much rainfall, whereas in the areas of tropical convergence maxima areas the mean model underestimates the (convective) precipitation, particularly in JJA as we saw in the top row zonal average of Fig. 4.7. Inter-model variations are, not surprisingly, largest in the tropical convergence areas, and to a lesser extent in the mid-latitude storm tracks (i.e., in regions of generally high precipitation).

#### **4.3.3 Outgoing Longwave Radiation**

The mean model zonal average outgoing longwave radiation (OLR) is shown on the top panel of Fig. 4.9 with the ERBE observational data. The mean model agrees quite well with the observations, but note the significant inter-model spread, which at some latitudes is as much as  $25 \text{ Wm}^{-2}$  at the  $\pm 2$  standard deviation level. In the statistical diagrams of Fig. 4.9 we see again that the mean model stands out, exhibiting better agreement with the observations than any individual model. The arrows typically indicate that the agreement with observations tends to decrease when the atmospheric models are coupled to the oceans. Some CMIP models, however, simulate the OLR as well or better than the majority of AMIP models. As was the case for surface air temperature and precipitation, there is very little difference between the AMIP and CMIP median models (despite the fact that they comprise a very different collection of models).

From a global perspective, it can be argued that OLR is more accurately measured than surface air temperature or precipitation. This in turn suggests that the results presented here are likely more robust than those in Figs. 4.5–4.8, and that we can more precisely identify apparent systematic model biases. Although the distributions of the mean model OLR (Fig. 4.10) agree qualitatively with the ERBE data, the difference maps between the mean model and observations do highlight important biases, most of which occur in the tropics or subtropics where differences of either sign are as much as  $25 \text{ Wm}^{-2}$ . Problems in simulating clouds are undoubtedly responsible for these biases. It is interesting to note that there is little correlation between these low-latitude bias maxima in the mean model and the inter-model standard deviation depicted in the last row of Fig. 4.10.

The zonal mean precipitation from the CMIP2 overview suggested that many models suffered from a nonphysical “double inter-tropical convergence zone.” This bias is not

evident on a global scale with these (CMIP2+) somewhat newer results, however the problem is clearly revealed in the tropical biases section (4.5).

#### **4.3.4 Total Cloud Cover**

The zonal average total cloud cover is shown in Fig. 4.11. This quantity is not sufficient for evaluating the simulated cloud-radiative properties, but it is the traditional first comparison of simulated clouds to observations, and it can augment the information we obtain from clear and cloudy sky radiative fluxes. Because total cloud cover as derived from models is generally not consistent with the observed definition, new diagnostic tools have been developed to enable better model-to-observation comparisons, e.g., Klein and Jakob (1999), Webb et al. (2001), and Zhang et al. (2004). In our evaluation of the CMIP2+ data, we are limited (by the model output available) to presenting traditional total cloud cover comparisons.

In the upper panel in Fig. 4.11 the mean model agrees reasonably with the observational estimates (ISCCP) in the lower latitudes, but underestimates cloud fraction in the extratropics, especially in the Southern Hemisphere. In the high latitudes, there is a very large spread among the models, and ISCPP suggests the models have too many clouds in the polar regions. The error statistics in the Taylor diagrams of Fig. 4.11 show much lower pattern correlations than we have seen for the other fields, and in the boreal winter most models seem to underestimate the spatial variability. The total cloud cover patterns are shown in Fig. 4.12. At first glance the mean model compares with ISCCP reasonably well in a qualitative sense, but upon closer examination there are serious discrepancies. There are the well-known biases, e.g., in eastern oceans where models fail to form marine stratus, near the tropical Pacific cold tongue where the models over-predict clouds, and over Eurasia in boreal winter where the models are too cloudy. In a gross sense, the mean model does not compare well with this observed estimate, but it should be emphasized further that more innovative approaches to diagnose simulated clouds and their effects are under development. From the bottom panel of Fig. 4.12 it is clear that there are few places where the models agree with one another.

#### **4.3.5 Mean Sea-level Pressure**

The mean sea level pressure of the mean model is examined in Fig. 4.13. Equator-ward of 60 degrees latitude the models agree well with the two reanalyses, but at higher latitudes there are significant discrepancies, which is consistent with the earlier results from CMIP2. Note ERA15 and the NCEP reanalysis agree with one another in the Arctic, but not in the high latitudes of the SH, arising in part from difficulties in extrapolating surface pressure to sea level. The Taylor diagrams suggest both the coupled and uncoupled simulations correlate very well with ERA15. The mean CMIP and AMIP models stand out, as they have for other quantities. Spatial patterns are shown in Fig. 4.14. The mean model looks very much like the ERA15, but there are some notable biases in North Africa, the circumpolar current, etc. The inter-model standard deviation is small, with maxima in high-altitude areas where, again, differences in extrapolation methods may be partially responsible.

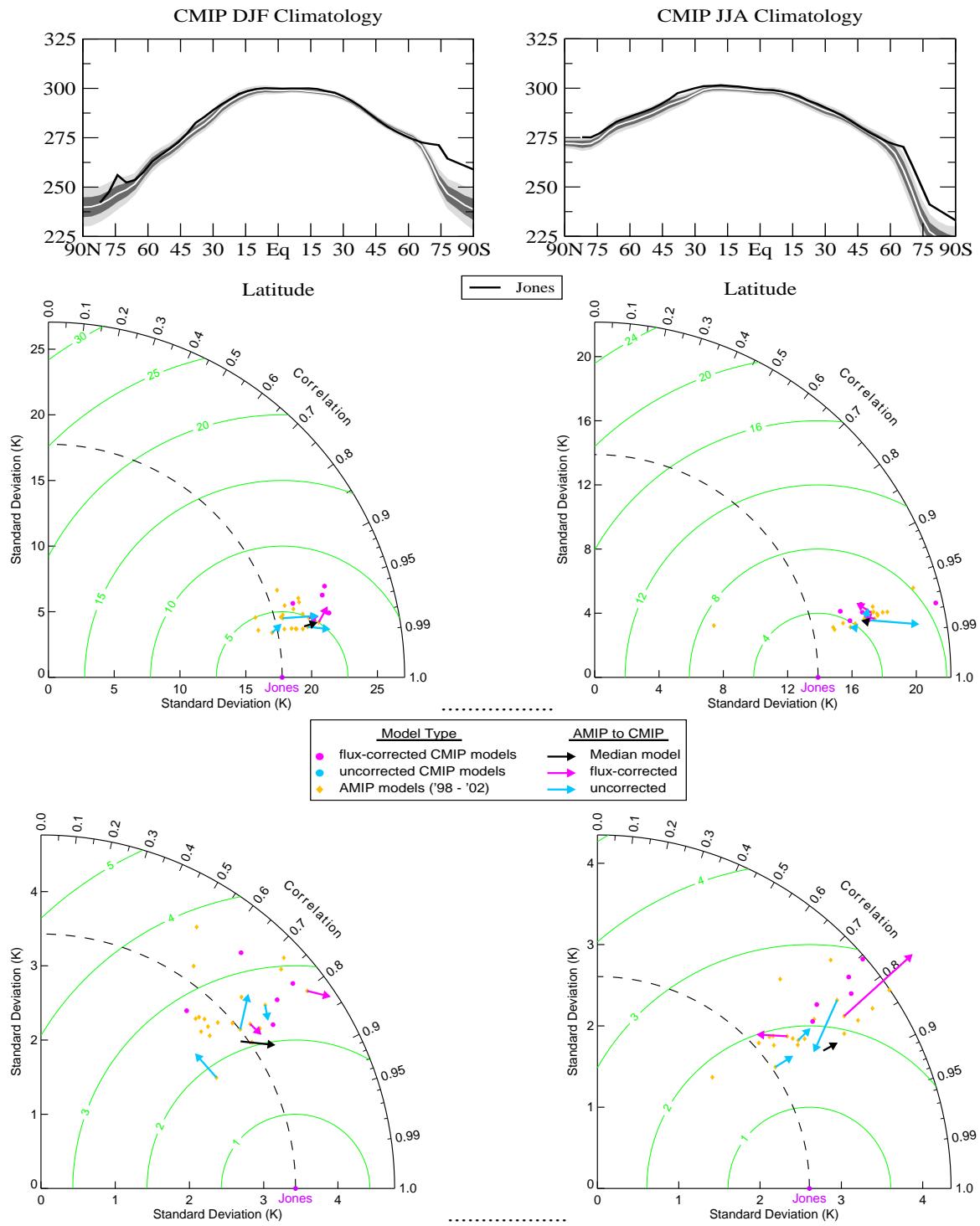


Fig. 4.5. DJF and JJA surface air temperature (K). Upper panels: observed (dark line) and mean model (white line) zonal averages with  $\pm 1\&2$  inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean.

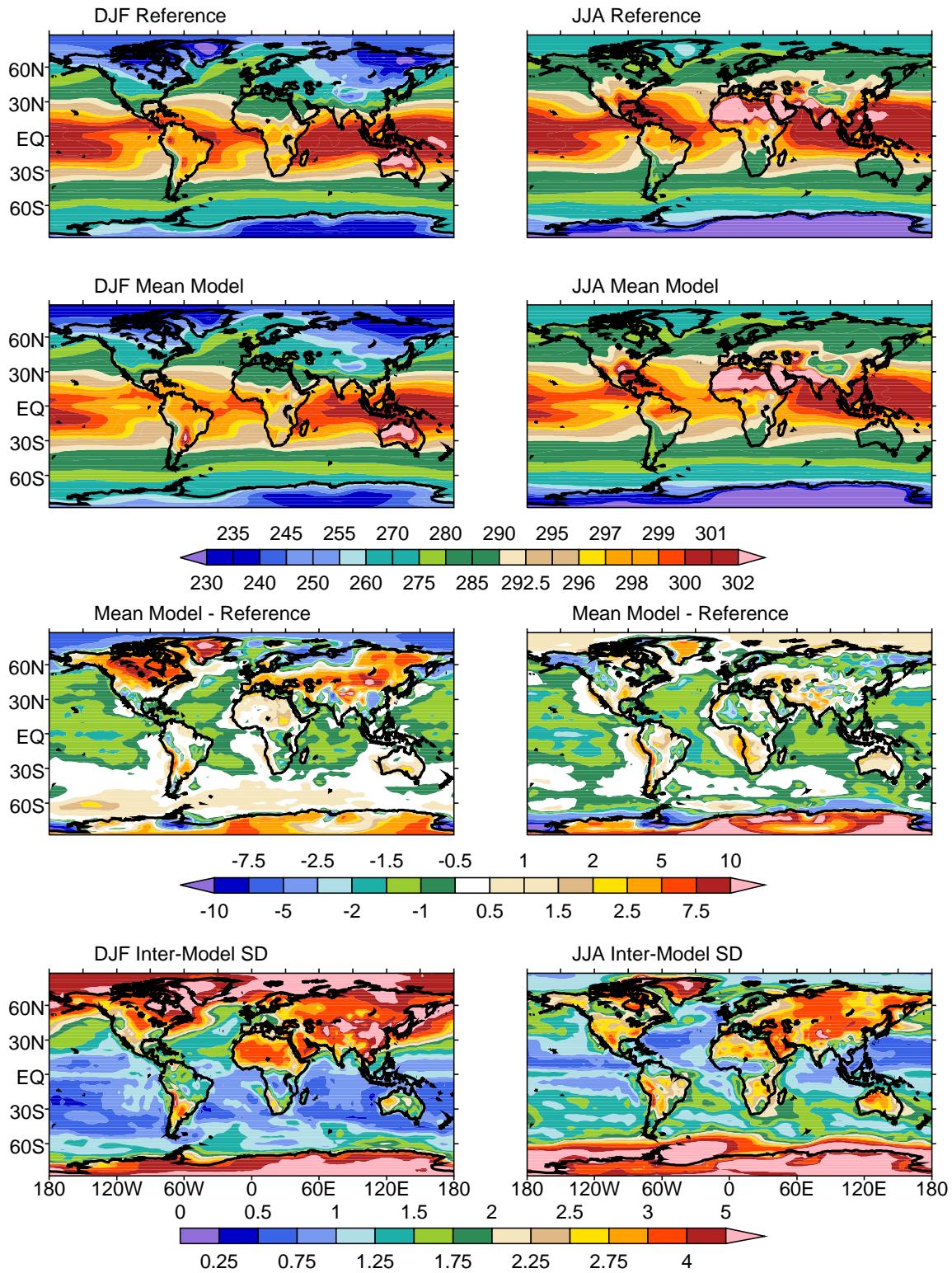


Fig. 4.6. DJF and JJA surface air temperature (K). First row: Jones (observations); Second row: Multi-model ensemble mean; Third row: Multi-model ensemble observations; Bottom row: Multi model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots.

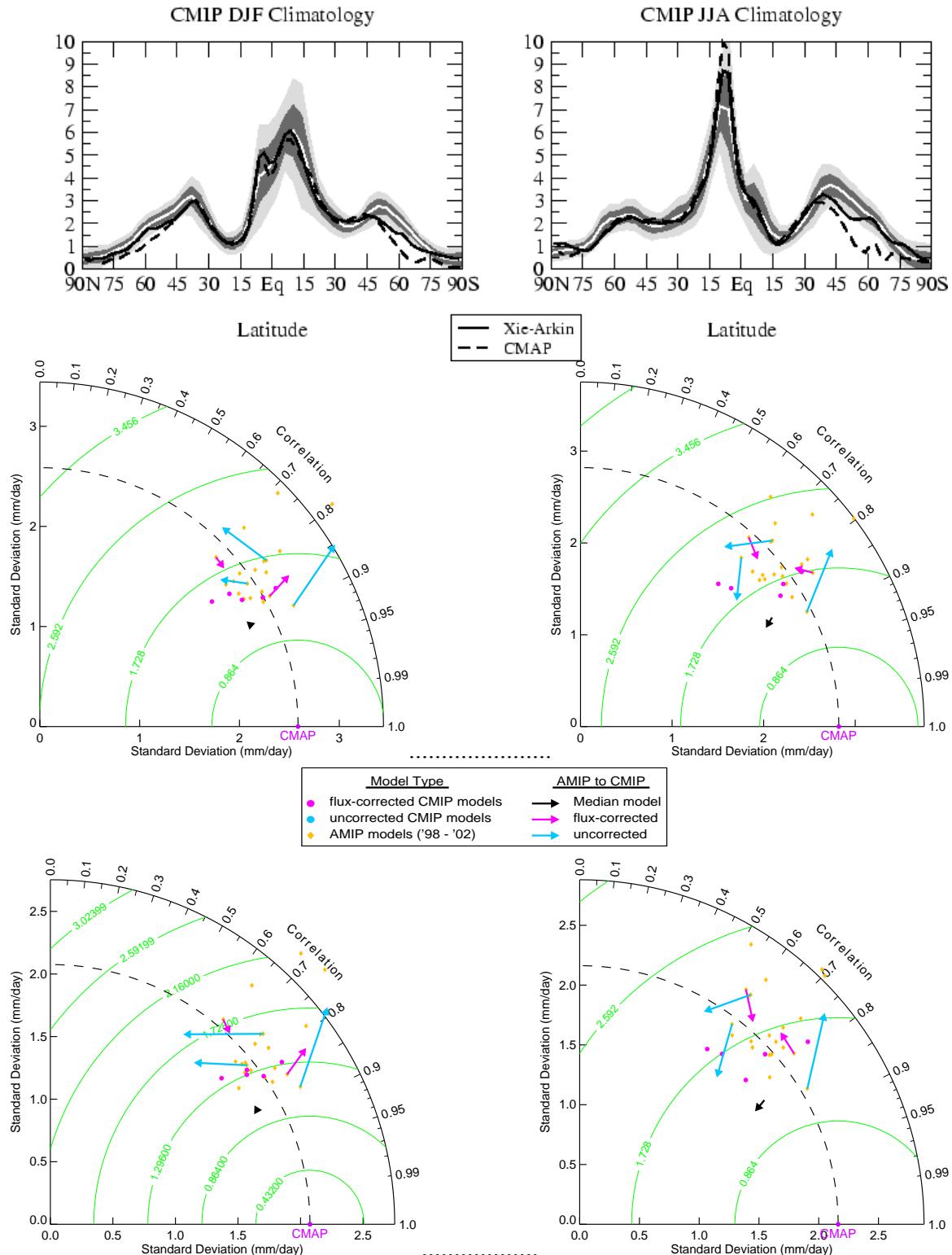


Fig. 4.7. DJF and JJA total precipitation (mm/day). Upper panels: observed (dark line) and mean model (white line) zonal averages with  $\pm 1\&2$  inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean.

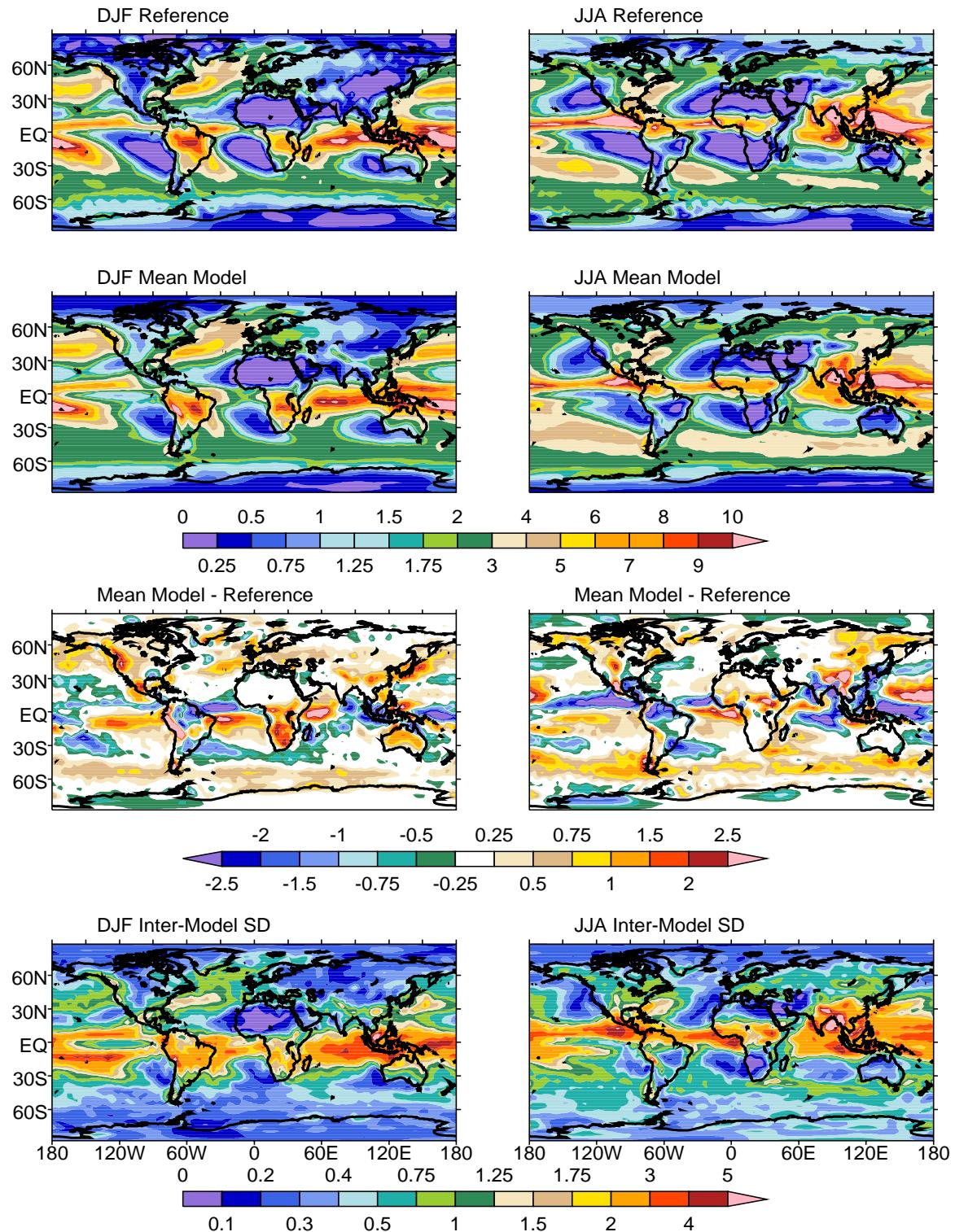


Fig. 4.8. DJF and JJA total precipitation (mm/day). First row: CMAP (Observations); Second row: Multi-model ensemble mean; Third row: Multi-model ensemble observations; Bottom row: Multi-model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots.

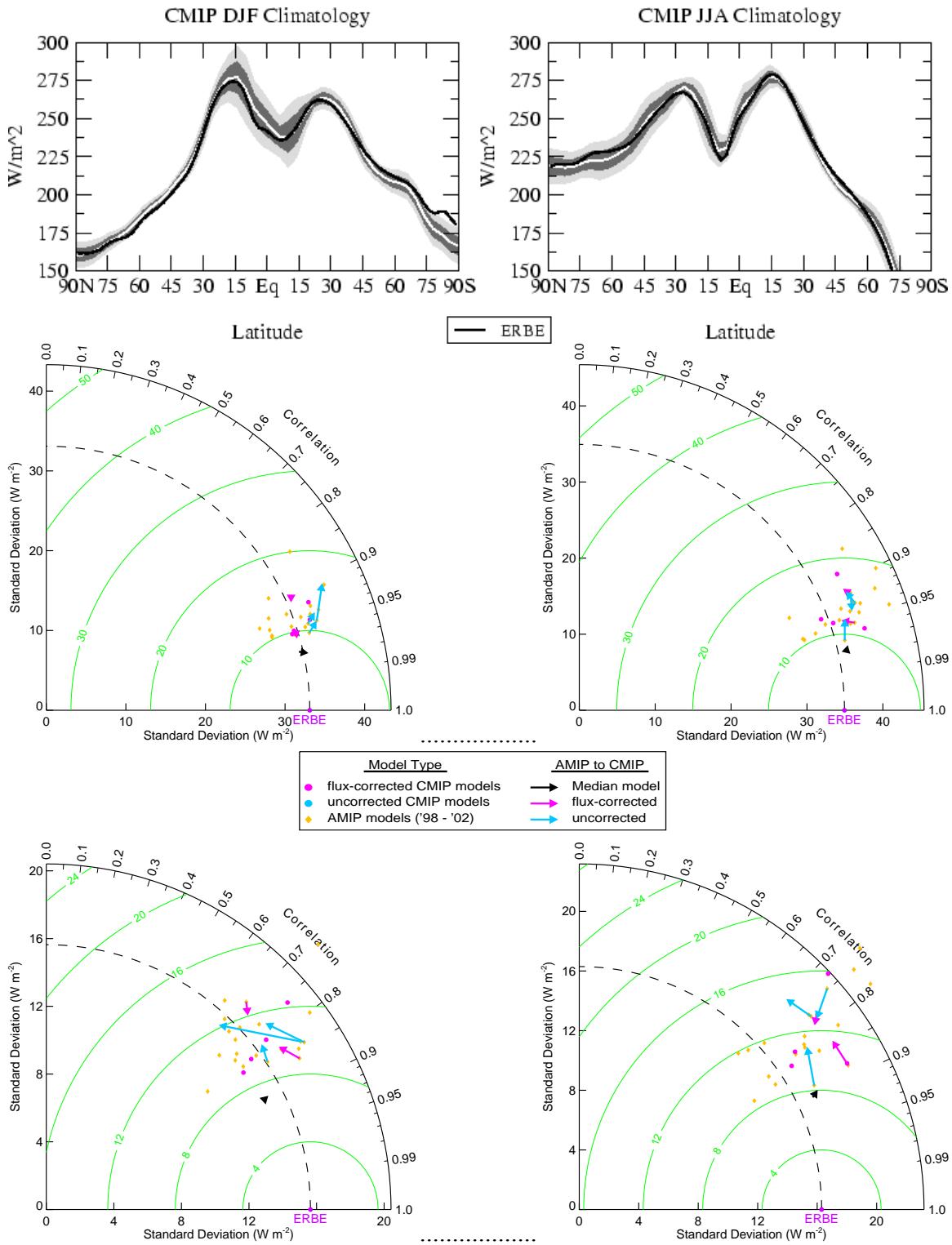


Fig. 4.9. DJF and JJA outgoing longwave radiation ( $\text{W m}^{-2}$ ). Upper panels: observed (dark line) and mean model (white line) zonal averages with  $\pm 1\&2$  inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean.

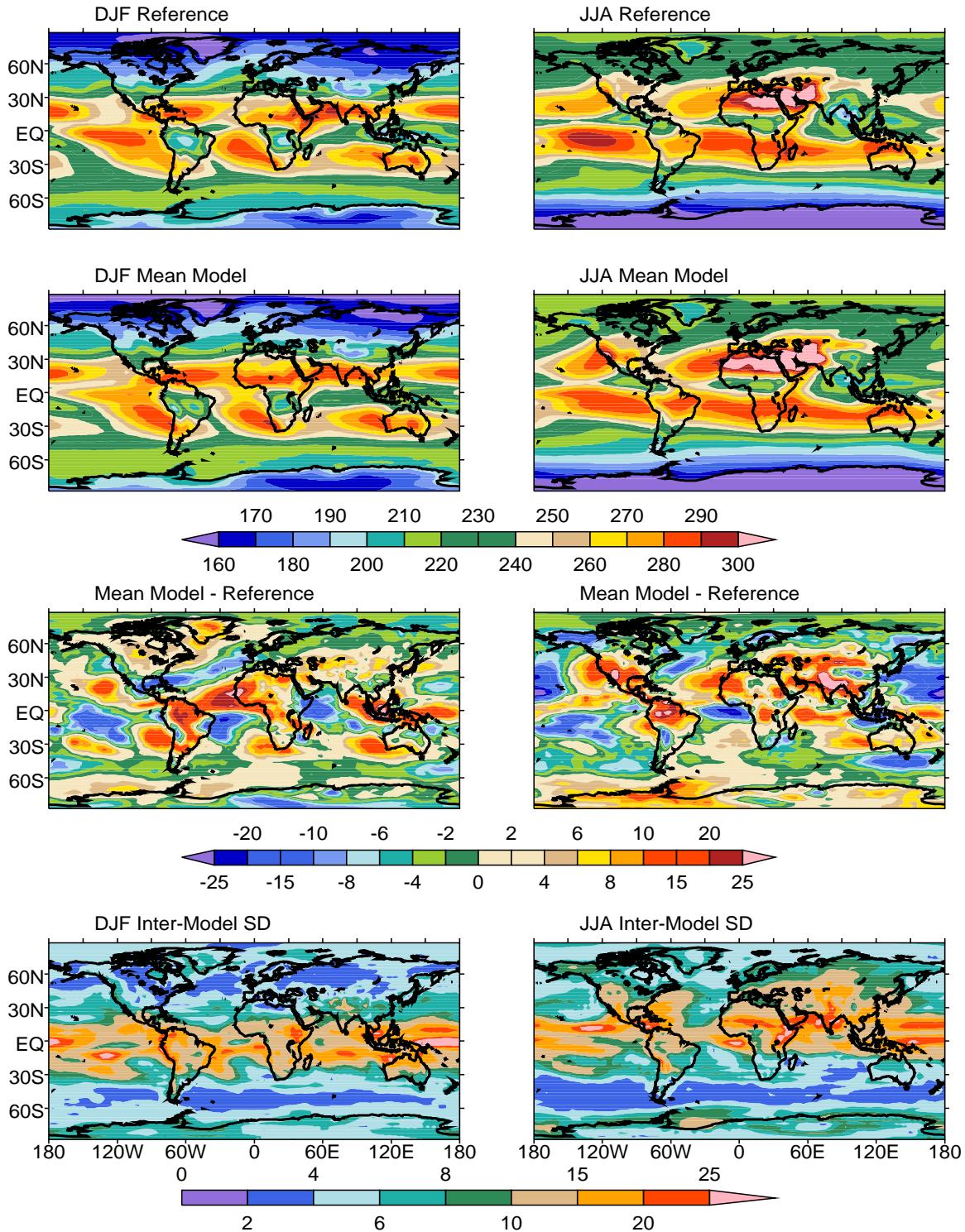


Fig. 4.10. DJF and JJA outgoing longwave radiation ( $\text{Wm}^{-2}$ ). First row: ERBE (Observations); Second row: Multi-model ensemble mean; Third row: Observations minus Multi-model ensemble; Bottom row: Multi model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots.

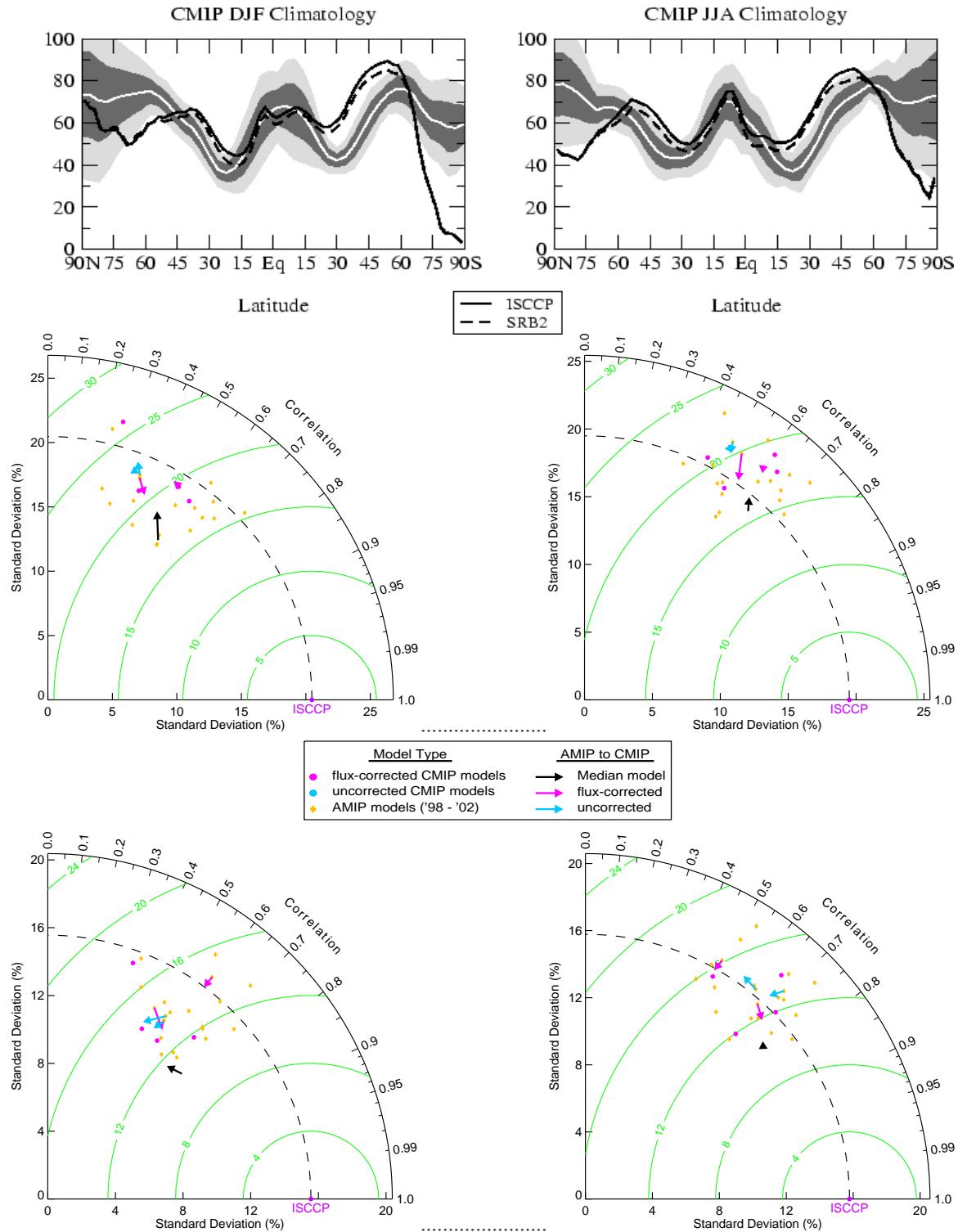


Fig. 4.11. DJF and JJA total cloud cover (percent). Upper panels: observed (dark line) and mean model (white line) zonal averages with  $\pm 1\&2$  inter-model standard deviations (shaded); Middle panels: spatial pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean.

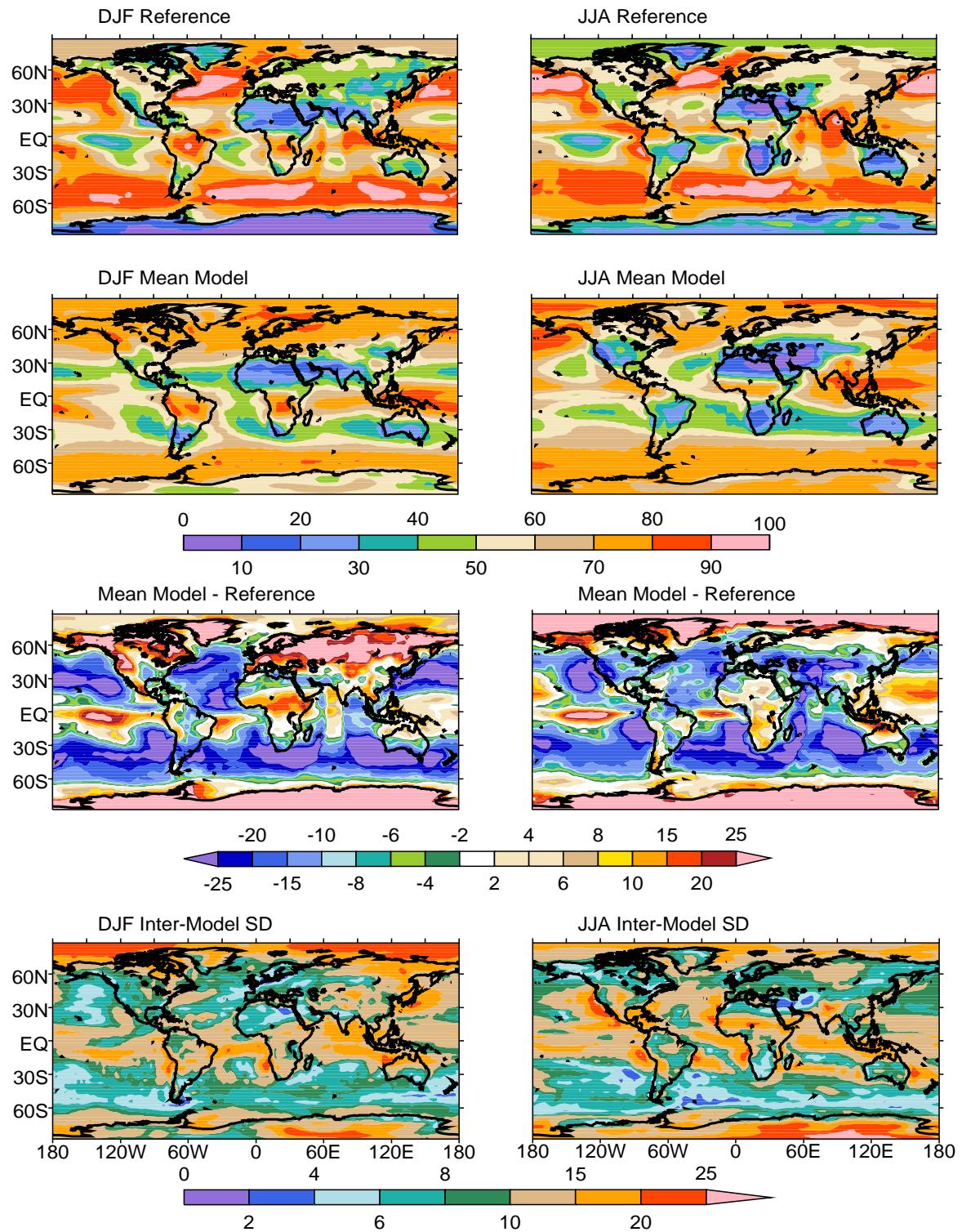


Fig. 4.12. DJF and JJA total cloud cover (percent). First row: ISCCP (Observations); Second row: Multi-model ensemble mean; Third row: Observations minus Multi-model ensemble; Bottom row: Multi model ensemble inter-model standard deviation. Note that nonlinear scales are used for all plots.

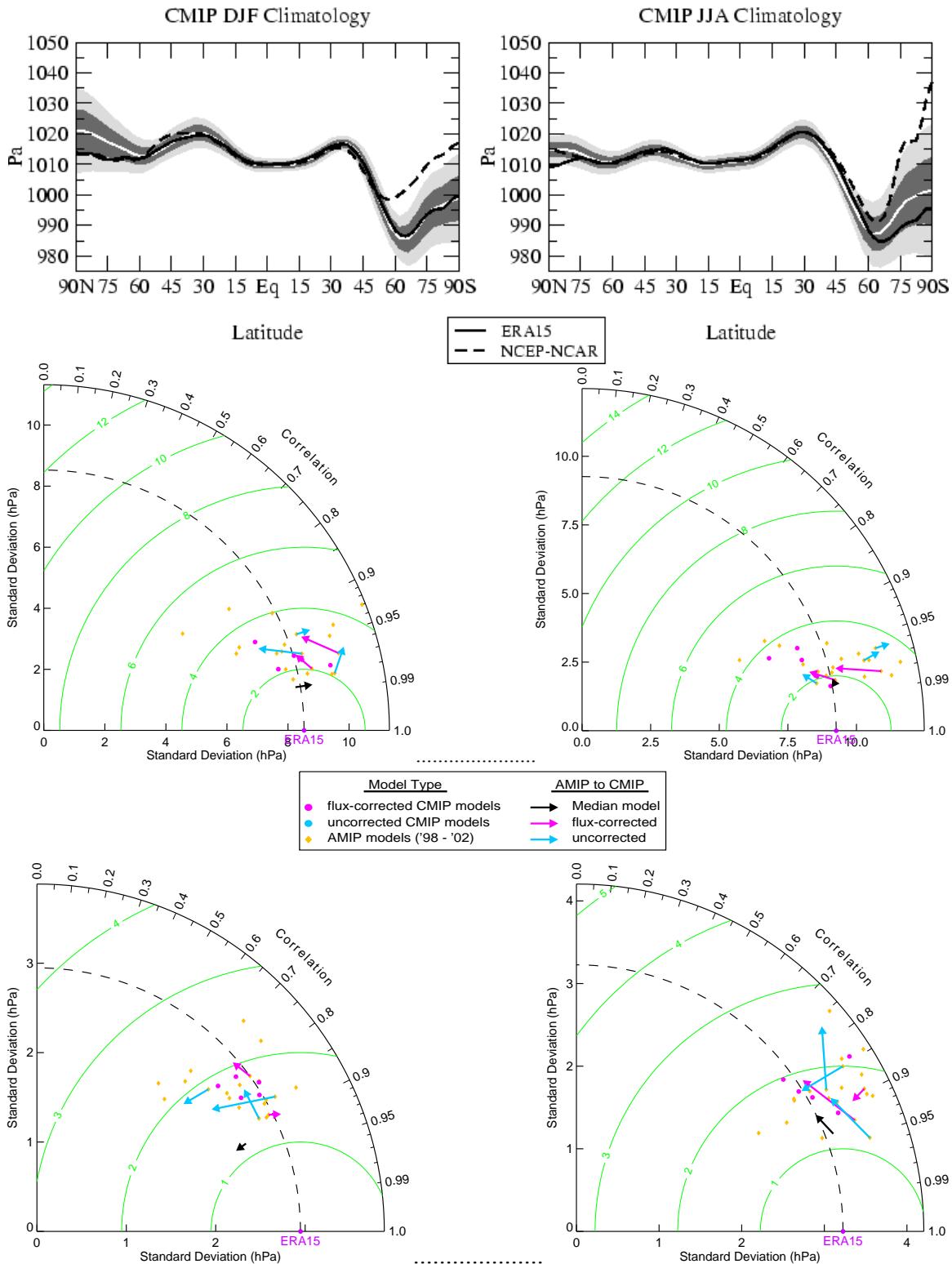


Fig. 4.13. DJF and JJA sea level pressure (hPa). Upper panels: observed (dark line) and mean model (white line) zonal averages with  $\pm 1\&2$  inter-model standard deviations (shaded); Middle panels: ocean pattern statistics based on all latitude and longitude grid-cell values (area weighted); Lower panels: spatial statistics for departures from the zonal mean.

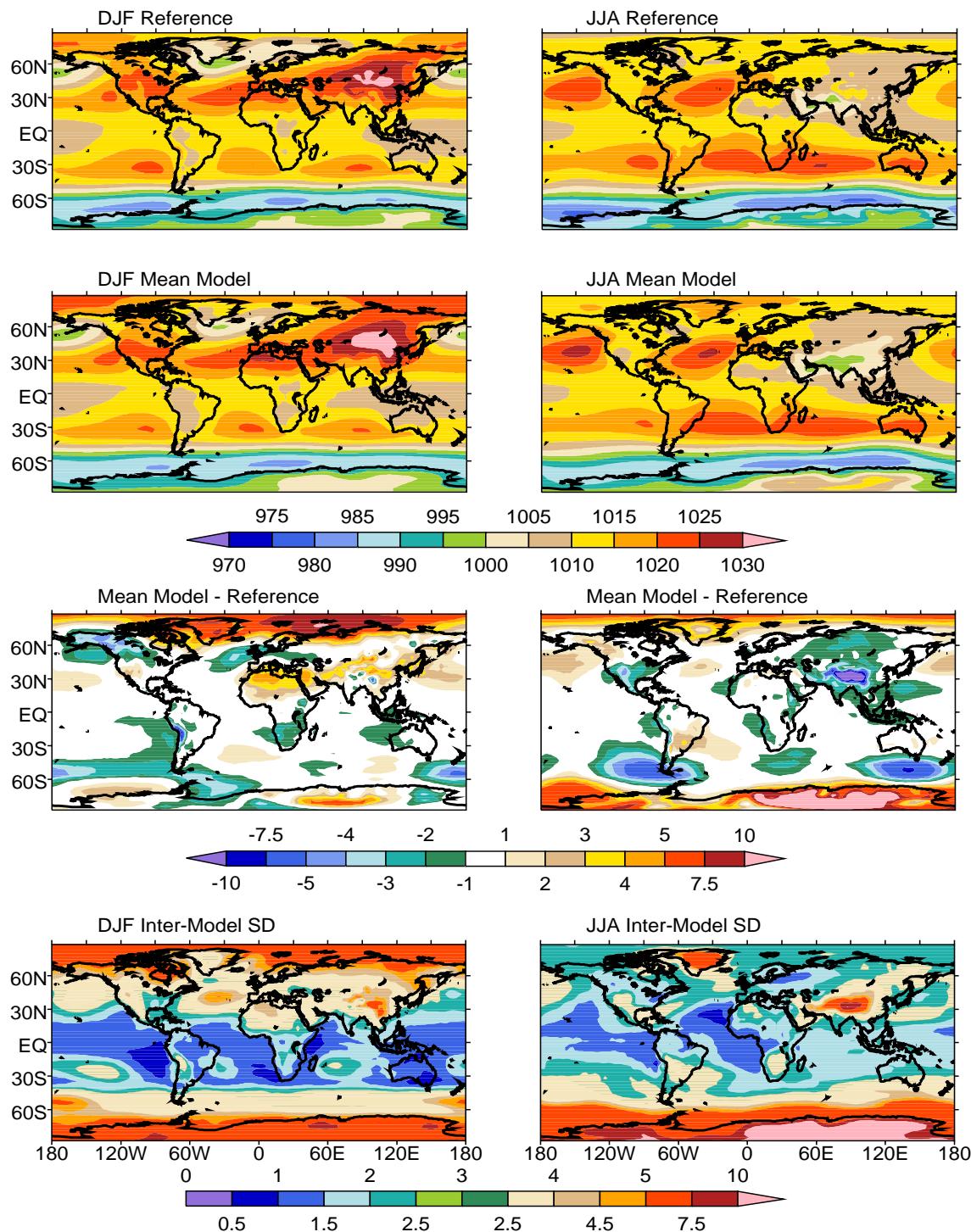


Fig. 4.14. DJF and JJA mean sea level pressure (hPa). First row: ERA15 (Reanalysis); Second row: Multi-model ensemble mean; Third row: reanalysis minus Multi-model ensemble; Bottom row: Multi model ensemble inter-model standard deviation. Note the nonlinear scales are used for all plots.

## 4.4 The Amplitude and Phase of the Seasonal Cycle

### 4.4.1 Introduction

In this section we will examine the mean annual cycle for surface air temperature, precipitation, and outgoing longwave radiation. We begin by taking a cursory look at the global-scale, including an illustration of the relative importance of the semi-annual cycle. We will then look more closely at the annual cycle in four large-scale domains encompassing: 1) North America, 2) South America, Atlantic Ocean, and Western Africa, 3) the Indo-Pacific, and 4) Europe, the Middle East and Northern Africa. In each of these selected regions we will examine characteristics of the annual cycle (via harmonic analysis described below). Within each region we will also highlight one key regional-scale sub-domain with area-averaged time series of our selected variables.

Harmonic analysis has proven to be a useful tool for evaluating the amplitude and phase of the annual cycle (e.g., Hsu and Wallace, 1976; Kirkyla and Hameed, 1989; Gates et al., 1999). A monthly mean annual time series at any (x-longitude, y-latitude) location can be represented as:

$$X(x, y, t) = \bar{X}(x, y) + \sum_{i=1}^{N/2} A_i(x, y) \cos\left(\frac{360it}{P} + \phi_i(x, y)\right) \quad (\text{Eq. 4.4.1})$$

where  $\bar{X}$  is the arithmetic (in this case annual) mean,  $A_i$  are the amplitudes of the harmonics,  $\phi_i$  are the phase angles of the corresponding harmonics,  $N$  is the total number of observations (in our case  $N = 12$  months),  $P$  is the period of the observation (one year) and  $t$  is the time. The annual cycle requires equal spacing between increments, which can be approximated as twelve intervals of 30.44 days. All of the observational and model data used here was only available as monthly calendar means, but we agree with previous investigators that this mismatch does not greatly affect our results. The resulting series is reduced to six harmonics, the sum of which completely describe the annual cycle. The first harmonic has one maximum and one minimum, and thus describes the tendency annual cycle. The second harmonic represents the semi-annual tendency and has two maxima and minima. The third describes features repeating three times each year, etc.

Our primary interest here will be the annual cycle, but the relative influence of each harmonic can be easily evaluated by comparing ratios, e.g.,  $A_2/A_1$ . A large first harmonic reveals a strong annual cycle component, and likewise a comparatively large second harmonic suggests an import influence of the semi-annual cycle. The phase angle is indicative of the time of the year that a maximum (or minimum) of a harmonic occurs.

The six harmonics that represent the annual curve are orthogonal to one another, and thus tend to describe different regime features. Generally speaking, longer period harmonics (e.g.,  $A_1$  and  $A_2$ ) represent large-scale features of the atmospheric circulation, whereas the shorter-period harmonics capture more local phenomenon.

Although our focus here is on the large scale, we will be particularly interested in exploiting this analysis for capturing and highlighting important boundaries and areas of transition that may not be detected by more traditional analysis of the mean climate.

#### **4.4.2 Global Characteristics: $T(2m)$ , Precipitation and OLR**

In Fig 4.15 we examine the amplitude and phase of the seasonal cycle of surface air temperature. In this and similar figures that follow there are no measures of inter-model spread. From the first two rows of Fig. 4.15, we see that the first harmonic amplitude and phase (i.e., the time of maximum temperature) of the mean model qualitatively agrees very well with the Jones data set both over oceans and land. Closer examination does reveal important discrepancies in both amplitude (e.g., North America and the greater Sahara) and phase (e.g., premature temperature maxima in the Amazon Basin). We will take a closer look at such differences when we turn our attention to our selected large-scale domains. Correct simulation of the semi-annual cycle can be expected to be more difficult. Nevertheless, apart from some notable differences (such as the underestimate of  $A_2/A_1$  over Eurasia seen in the bottom half of Fig 4.15), the mean model captures quite well many gross scale features seen in the observed semi-annual phase and amplitude.

The amplitude and phase of the annual and semiannual precipitation is shown in Fig. 4.16. Note the similarity between the annual cycle amplitude patterns here and the DJF and JJA seasonal means shown in Fig 4.8. The gross features of the mean model compare well with the observational annual amplitude, including the locations of sharp transition regions. The amplitude of the mean model appears too strong in the mid-latitudes of both the NH and SH, particularly in the  $30^{\circ}$ - $40^{\circ}$  latitude band, e.g., in the Southern Ocean southwest of Australia. In contrast, the mean model underestimates the amplitude in the tropical convergence zones, which appear to be too confined, even after interpolating the models to a common grid (T42) and averaging them together. This suggests, at least in the case of the tropical Pacific and Atlantic Oceans, that the simulated ITCZ may not migrate with the seasons as much as is expected (refer to the tropical biases section 4.5). We find that despite some important differences, it is remarkable how well the gross features of the annual cycle phase of the mean model agree with the observations. The semi-annual features are much more complex. In most areas where the semi-annual amplitude is comparable (or exceeds) the annual amplitude, it appears this is not a result of annual amplitude minimum. When we look more regionally, however, we will see many examples where this is not the case, and where in fact, the  $A_2/A_1$  ratio is large primarily because  $A_1$  is small.

Global amplitude and phase plots of OLR are shown in Fig. 4.17. The annual cycle amplitude of the mean model captures most of the main features of the observations, including various maxima seen over the continents, and the locations of sharp transition zones. One important exception is the North Atlantic Ocean, where the patterns and magnitude are quite different between ERBE and the mean model. As expected, the annual cycle phase closely matches that of the surface air temperature (see Fig. 4.5) and once more there is good agreement between the mean model and ERBE, even in regions of sharp transition. Note that most of the regions of maximum semi-annual amplitude are actually artifacts of underestimating the annual cycle amplitude. There are, however, regions where the semiannual cycle may be important, e.g., over North Africa, and the Middle East, where  $A_2/A_1$  ratios greater than 20% suggest the semiannual cycle is important to some degree.

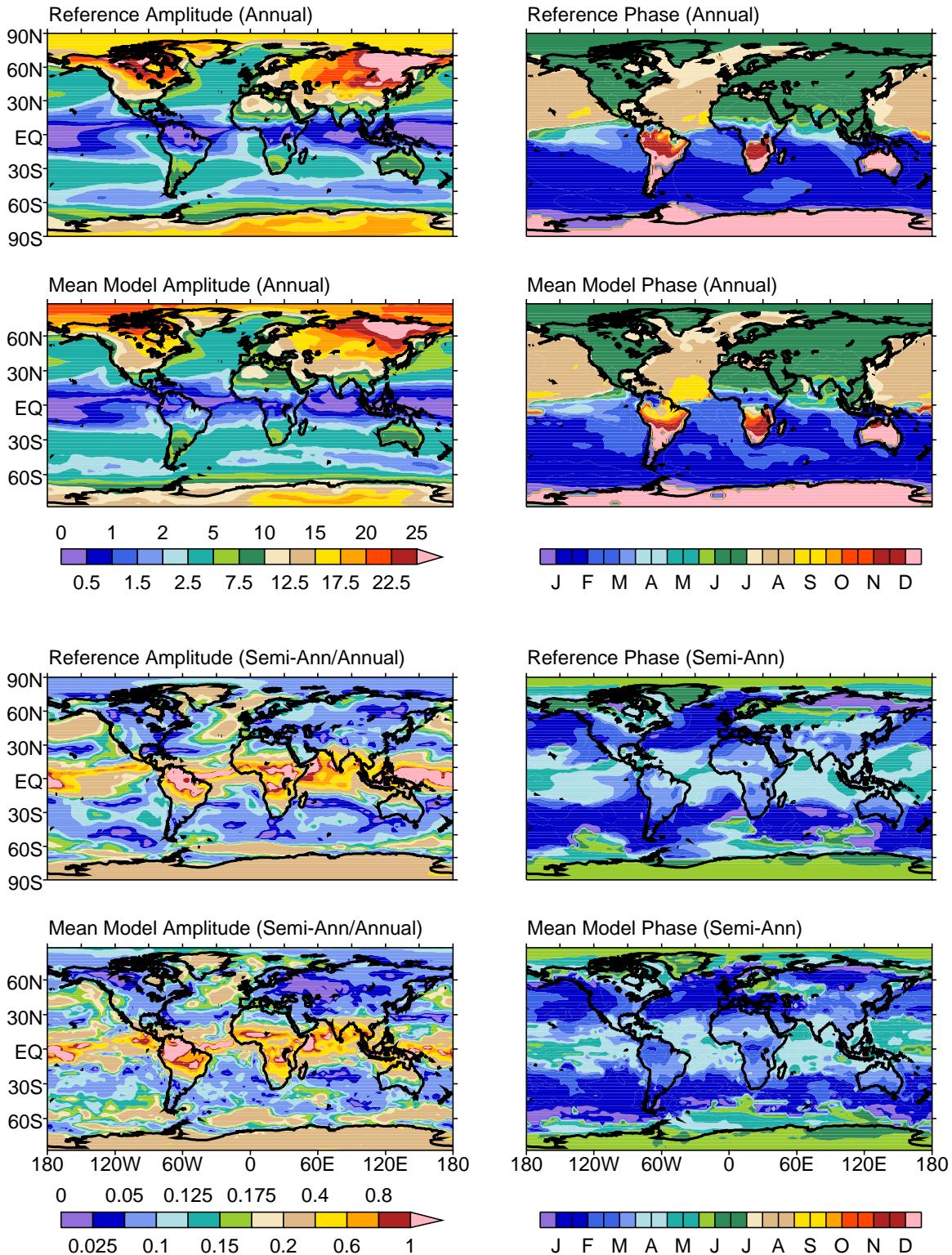


Fig. 4.15. Surface air temperature (K). Top row: Observational (Jones) annual amplitude and phase; Second row: Multi-model ensemble annual amplitude and phase; Third row: Observational ratio of semi-annual to annual amplitude ( $A_2/A_1$ ) , and phase of semi-annual cycle; Bottom row: multi-model ensemble ratio of semi-annual to annual amplitude ( $A_2/A_1$ ) , and phase of semi-annual cycle. Note the nonlinear scales.

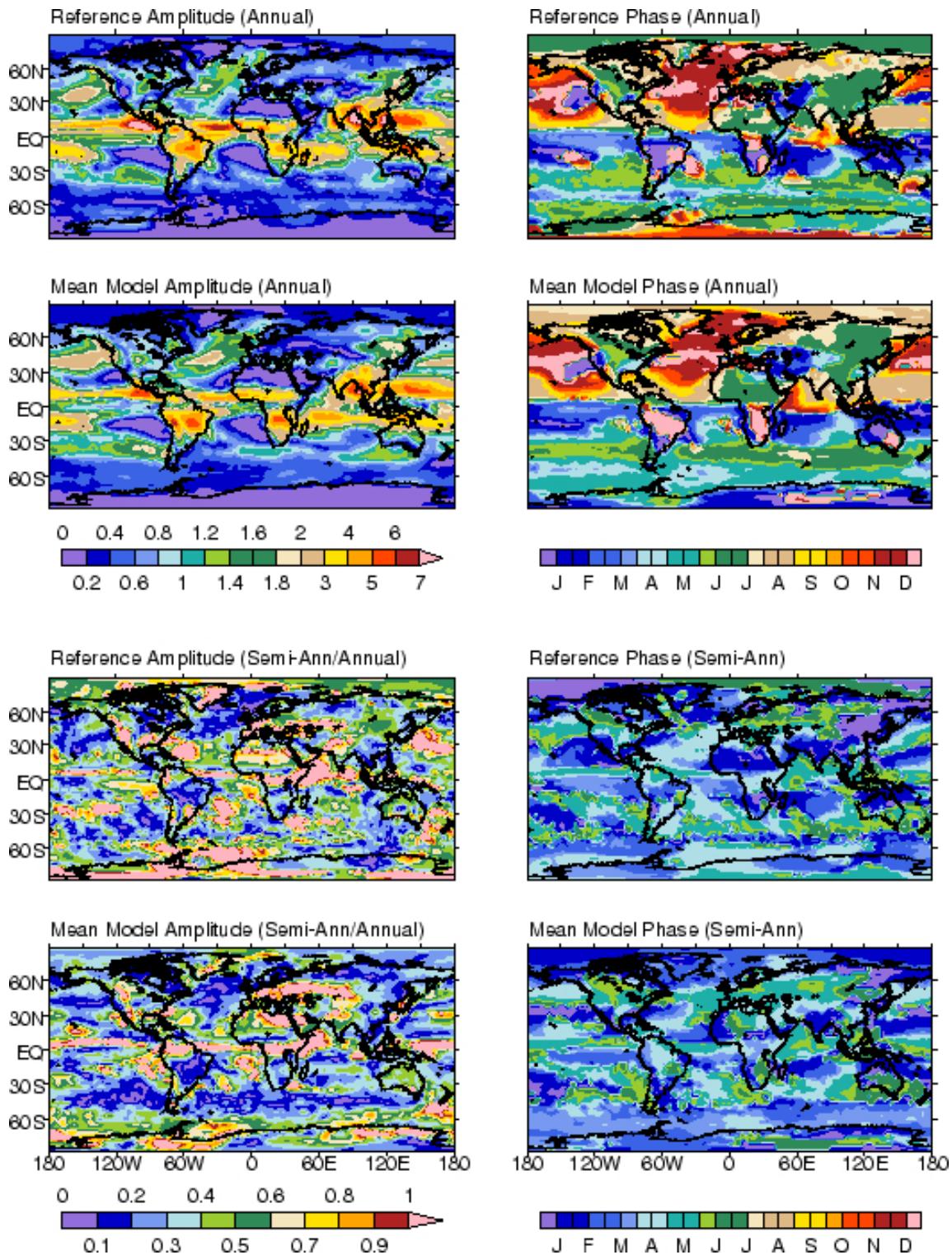


Fig. 4.16. Total precipitation (mm/day). Top row: Observational (CMAP) annual amplitude and phase; Second row: Multi-model ensemble annual amplitude and phase; Third row: Observational ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle; Bottom row: multi-model ensemble ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle.

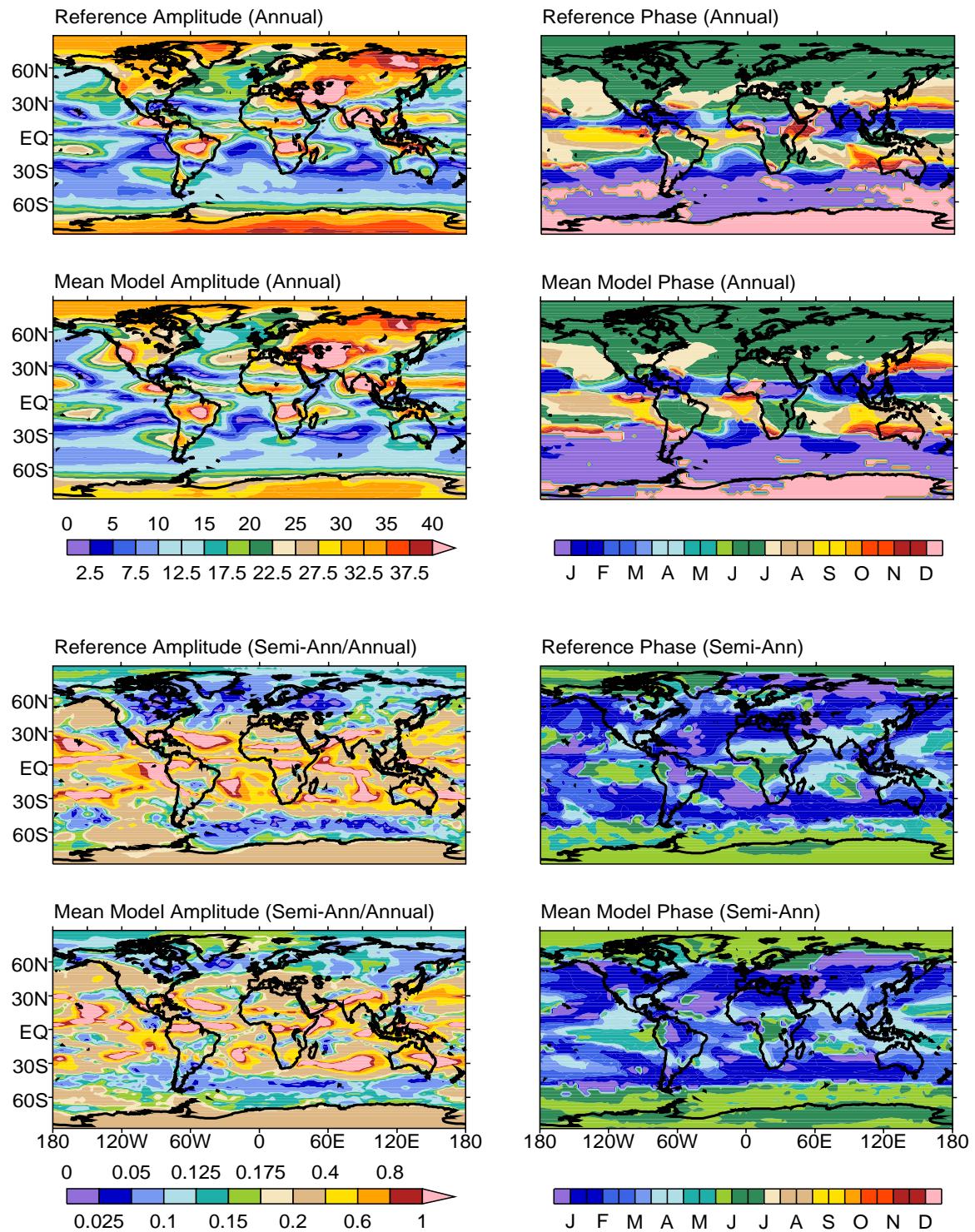


Fig. 4.17. Outgoing longwave radiation ( $\text{Wm}^{-2}$ ). Top row: Observational (ERBE) annual cycle amplitude and phase; Second row: Multi-model ensemble annual amplitude and phase; Third row: Observational ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle; Bottom row: multi-model ensemble ratio of semi-annual to annual amplitude ( $A_2/A_1$ ), and phase of semi-annual cycle.

#### **4.4.3 North America**

##### **4.4.3.1 Surface (2 m) Air Temperature, Precipitation and OLR**

Figure 4.18 illustrates how the annual cycle amplitude of surface air temperature increases from the tropics to the extra-tropics in the North American Region, with the largest amplitudes over the continental interior, away from the moderating effect of the oceans. Qualitatively the amplitude of the mean model and reference data agree quite well; however, the TAS difference plot in Fig. 4.18 clearly illustrates that the mean model underestimates the amplitudes in the northwestern areas of the continent and in the eastern tropical Pacific. Along the western coastline and in the southern U.S., the TAS annual amplitude of the mean model appears to be overestimated. Over much of the ocean areas the mean model agrees with the observations to within 0.5°C, not surprising given that many of the models employ flux adjustments. One exception is off the west coast in the areas where there is an observed active annual cycle of marine stratus. Here the mean model overestimates the TAS annual cycle, perhaps as a result of the models' inability to capture the boreal summer marine stratus and thereby leading to excessive surface temperatures.

Examining the vectors in the TAS plot (Fig. 4.18), it is clear that the mean model realistically captures the boreal summer maximum over the continent as well as over the oceans, where north of 15°N, the phasing lags the continent due to the oceans' thermal inertia. However, in the subtropics west of Central America there are notable differences (both lag and lead) between the mean model and reference data, but note here the TAS annual amplitude is less than 2°C.

The observed tropical rainfall annual cycle amplitude and phase are qualitatively captured by the mean model, except very near the equator. Further north, the amplitude minimum west of Baja California is underestimated, while the maximum over the Pacific near 45°N is overestimated. The west coast amplitude maximum penetrates too far inland, possibly owing to the models' inability to resolve the Rocky Mountains, although the physical interactions that influence moisture transport are of paramount importance. Farther east and similar to observations, the mean model produces a maximum over the north-central U.S., and a minimum over the Gulf coast. However, the mean model fails to capture their southwest-to-northeast tilts.

As illustrated in the precipitation difference map of Fig. 4.18, the mean model accurately represents the phase in many areas. The eastern tropical Pacific is a notable exception, but there are other smaller-scale discrepancies. In a gross sense, the phase appears better captured in the western areas (both ocean and land) of this domain.

As with TAS, the mean model annual cycle amplitude and phase of outgoing longwave radiation captures all of the predominant features observed. However, several biases are illustrated in the OLR difference map. Most glaring is the excessive OLR amplitude extending off the coast of California to southern U.S. This amplitude bias closely corresponds to what we see in TAS and to some extent in precipitation. Just to the south there is a weak amplitude minimum in the mean model over Central America. These biases may be related via circulation patterns in the mean model being slightly shifted southward in a region where there are rather sharp North-South gradients. And this in turn could be related to the horizontal (North-South) resolution of the models being too

coarse. To the north, and throughout Canada, the OLR amplitude is systematically underestimated, and consistent with the weak amplitude of TAS.

In both the mean model and the observations, over the continent the OLR phase closely matches that of TAS. Over the oceans it is more consistent (but out of phase by six months) with the precipitation amplitude.

#### 4.4.3.2 Southwest U.S. Seasonal Cycle

The North American Monsoon is an important regional influence that extends from the eastern Pacific, north to Central America, Mexico and the southwest U.S. As seen in Fig. 4.18, the mean model has difficulty in representing the transition of the annual cycle from a west coast maximum in rainfall to a minimum over the southwest U.S. It is instructive to examine the seasonal cycle over the southwest U.S. where the North America monsoon is observed. As seen in Fig. 4.19, the onset of rainfall over this region is characterized by a sharp onset from June to July in observations. The individual models exhibit great difficulty in capturing this onset, which begins one month early in two models and one or two months later than observed in two other models. Also, all of the models are characterized by a strong semi-annual component. The seasonal cycle of simulated cloud fraction is out of phase with the observations. Consistent with the earlier results, the seasonal cycle of OLR is dominated by the surface air temperature, although a systematic bias of 15–20 Wm<sup>-2</sup> of excessive OLR is evident during the boreal summer and autumn. This occurs despite the more uniform spread of simulated surface air temperature about the observations at this time. During boreal winter the models tend to underestimate the surface air temperature, with the spread being ~7 K throughout most of the year.

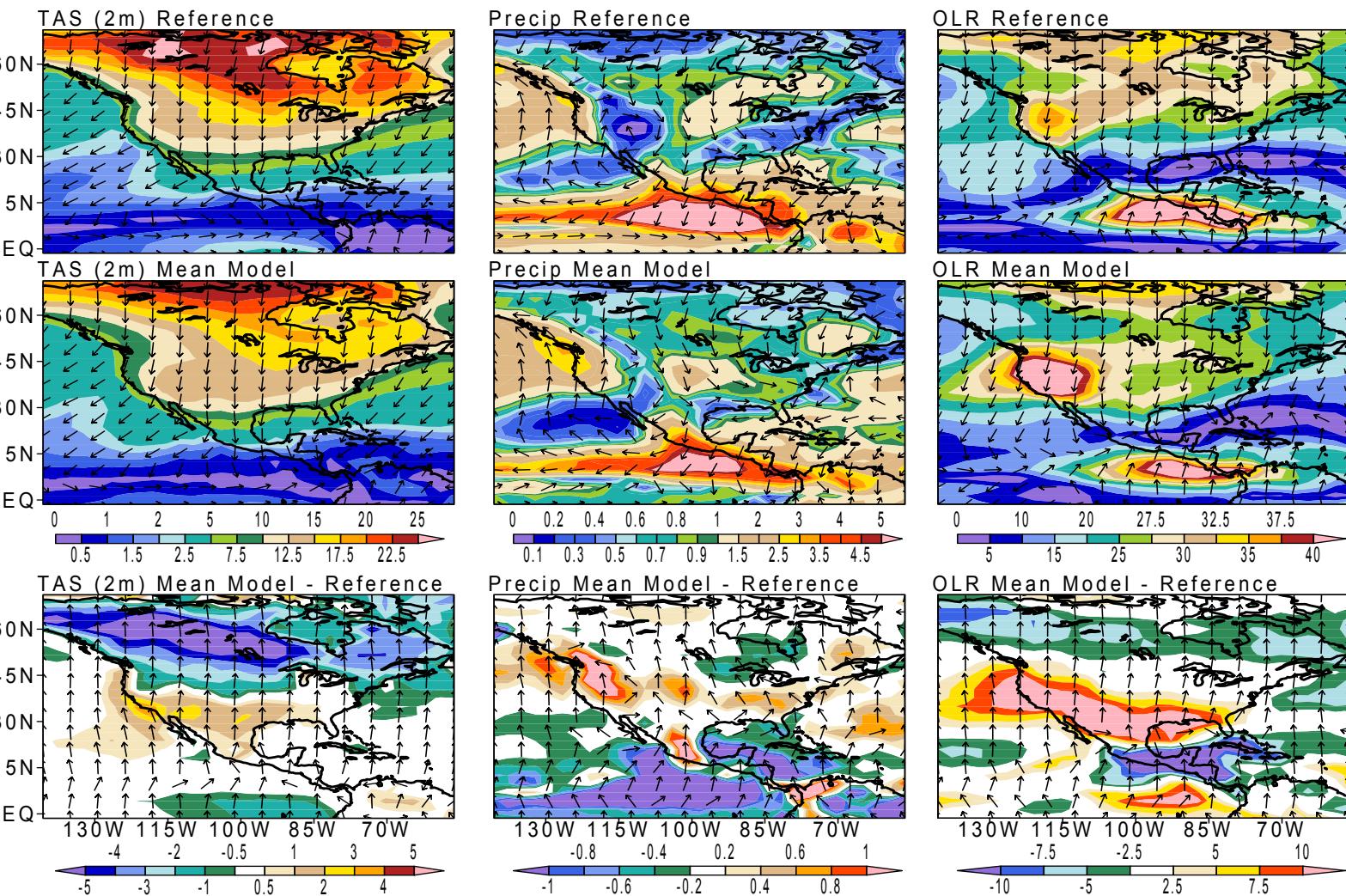


Fig. 4.18. North American annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{Wm}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 o'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.

## Southwestern U.S.

Averaging Domain: 30N, 37.5N, 105W, 115W

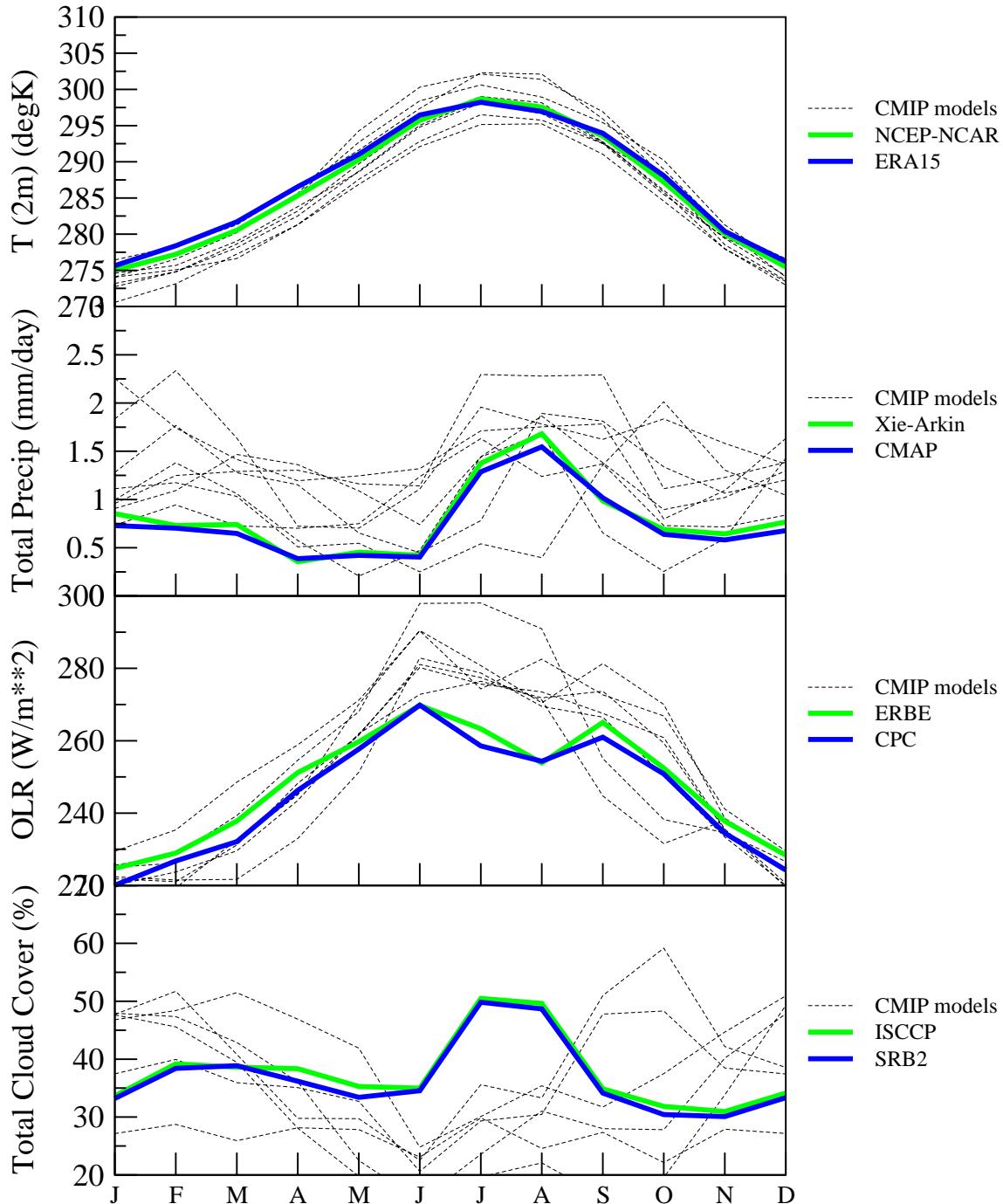


Fig. 4.19. Southwest US regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.

#### **4.4.4 South America, the Atlantic and Western Africa**

##### **4.4.4.1 Surface (2 m) Air Temperature, Precipitation and OLR**

Overall, the amplitude and phase of the annual cycle of surface air temperature is accurately simulated in this domain (Fig. 4.20). Southward of 15°S, the temperature over land closely follows the seasonal cycle of insolation, while that over the ocean lags. In the tropical convergence zone the phase transitions to the autumn in both hemispheres. The mean model's annual cycle amplitude agrees well with the observations, but it is overestimated in central Brazil, and underestimated near the tip of the South American continent and over the Gulf of Guinea. It is also overestimated in the Atlantic Ocean between 50–60°S.

As in the observed precipitation, the mean model has a large annual cycle amplitude over the continents, and over the Northern Hemisphere tropical regions where the ITCZ traverses during the course of the year with a maximum around 10°N. The mean model overestimates the amplitude over Peru, but underestimates it near the Amazon outlet in Brazil, the Gulf of Guinea, central Argentina, and in the ITCZ. The ability of the mean model to realistically capture the observed phase is mixed. Over much of South America the phase is reasonable, as well as in the NH and SH subtropics over land and ocean. There are, however, some important discrepancies over much of the South Atlantic, e.g., a two-month mean model lead between 45°S and 55°S.

The annual cycle amplitude and phase of the tropical OLR are consistent with that of the rainfall. Southward of 30°S they are more consistent with that of the surface air temperature. Over much of the domain the phase closely matches observations, with exceptions being west of the Andes, off the Guinea coast, and in some open ocean areas of the South Atlantic. Along the equator the annual cycle amplitude is underestimated.

##### **4.4.4.2 Nordeste, Brazil Seasonal Cycle**

The Nordeste region (Fig. 4.21) of northeast Brazil is well known for its austral autumn rainfall maximum. With two exceptions, the pronounced seasonality of the rainfall is captured by the CMIP2+ models, though the spread among the models is more than a factor of 4 during the rainy season. Concomitant with the pronounced annual cycle of rainfall is that of OLR and total cloud cover. As for the rainfall, the model spread in estimating the OLR and cloud cover is large. Most models do, however, capture the weak seasonal cycle of surface air temperature, including the austral spring temperature maximum.

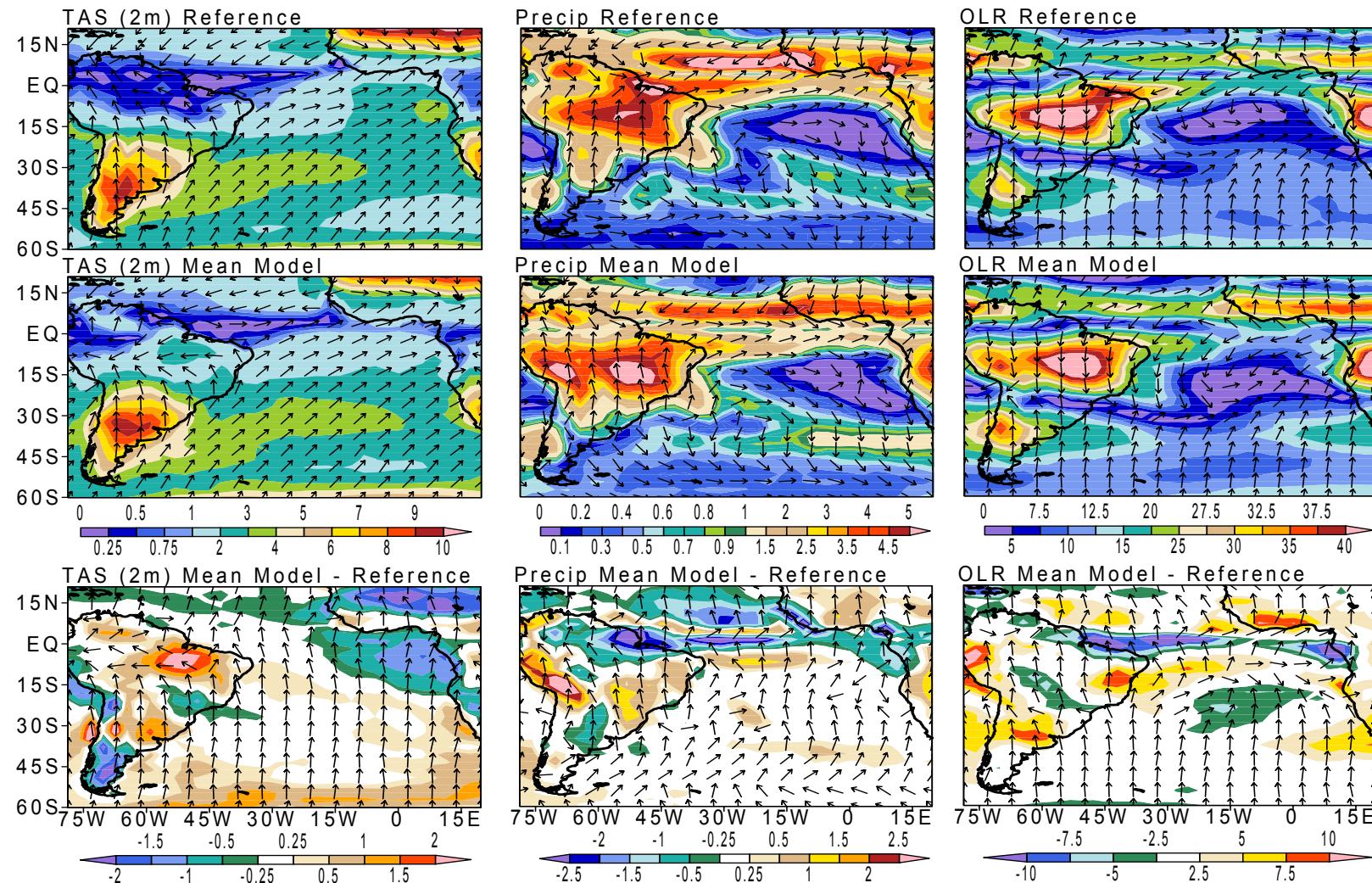


Fig. 4.20. South America, the Atlantic and Western Africa annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{W m}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 O'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.

## Nordeste Region

Averaging Domain: 10S, 4N, 42.5E, 35E (land only)

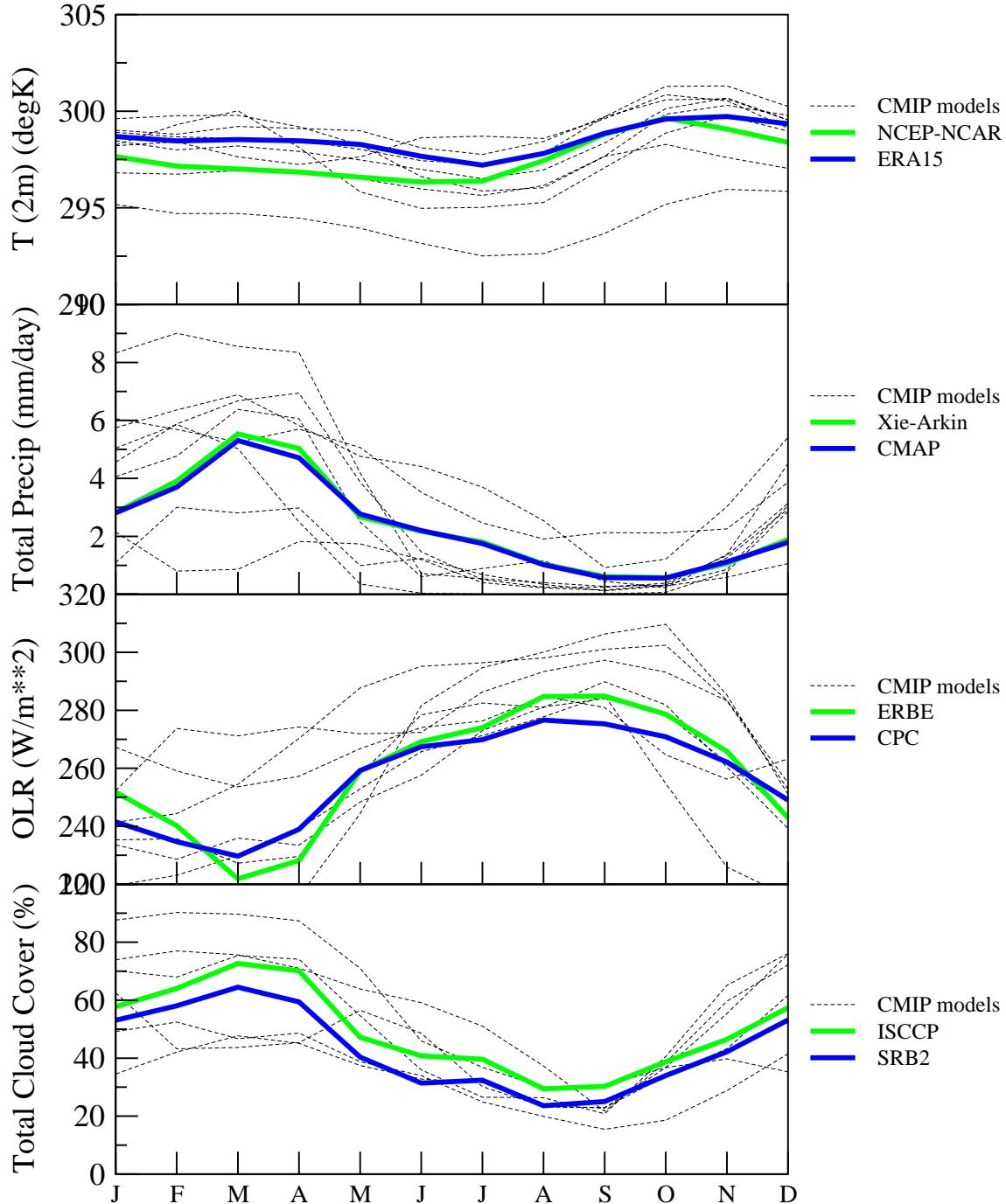


Fig. 4.21. Nordeste regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.

#### **4.4.5 Indo-Pacific**

##### **4.4.5.1 Surface (2 m) Air Temperature, Precipitation and OLR**

The dominant feature of the surface air temperature annual cycle harmonic in the Indo-Pacific region is the tropical minimum and the increasing amplitude poleward (Fig. 4.22). The mean model captures the amplitude gradients quite well, though the maximum over the interior desert of Australia is underestimated, while that over Asia is overestimated. The opposition in phase between the Northern and Southern Hemispheres is well represented, including the transition to an April maximum near the equator. Importantly, the mean model represents the late boreal spring maximum in temperature over India and Southeast Asia that precedes the onset of the monsoon. The largest phase errors occur over the equatorial west Pacific Ocean.

The amplitude of the annual harmonic of precipitation represents the monsoon influence that dominates the Indo-Pacific region (Fig. 4.22). Near India the mean model does not capture the rainfall maxima adjacent to the western Ghats, perhaps due to the inability of the coarse horizontal resolution to resolve the observed orography, and over the Bay of Bengal, and the west Pacific. Additionally, the rainfall maximum over the Arabian Sea occurs later than observed. From the Bay of Bengal to the west Pacific, the mean model captures the phase transition from early to late boreal summer. Additionally, the monsoon transition southeastward from the Bay of Bengal to Australia is resolved, indicating the ability to simulate the evolution of the large-scale monsoon system. The shortcomings that do exist are relatively minor, including the incorrect phase of the rainfall maxima near the equator north of New Guinea, and the overestimate of the rainfall amplitude over the southern Indian Ocean.

Not surprisingly, the annual cycle amplitude of OLR (Fig. 4.22) is consistent with that for precipitation, as is the phasing (note: the phase shows the month of maximum OLR, therefore the month of minimum OLR occurs 6 months later and is consistent with convection), with the model representing well the observed amplitude and phase.

##### **4.4.5.2 India Seasonal Cycle**

Time series of key variables over the Indian subcontinent show the performance of the individual models (Fig. 4.23). The bulk of the models represent the pre-monsoon heating that occurs in May. After the monsoon onset the temperature drops due to the increased soil moisture and increased cloud cover. The largest discrepancy between the models and observations is with respect to monsoon rainfall. Numerous models underestimate the boreal summer monsoon rainfall, and the onset and maximum rainfall tend to occur too late, consistent with the annual cycle phasing discussed previously. Despite the inhomogeneity of the Indian monsoon rainfall, the OLR and cloud cover are better represented. Consistent with the underestimate of rainfall over India, the OLR for most models is not as low as observed. However, the cloud cover is more equitably distributed about the observations, with the cloud fraction spanning the range of 60–90% during the monsoon season.

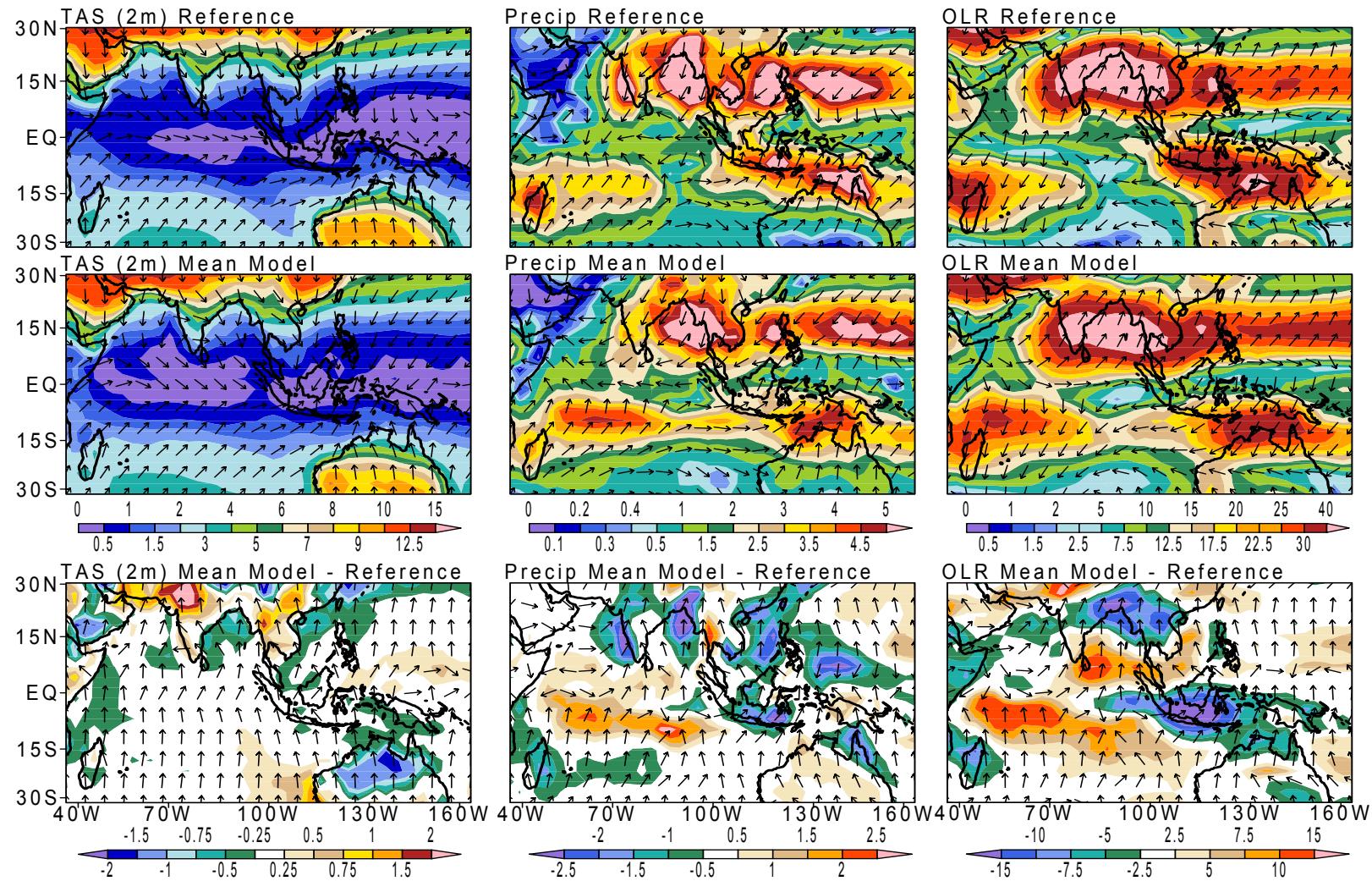


Fig. 4.22. Indo-Pacific annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{W m}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 o'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.

## India Region

Averaging Domain: EQ, 22.5N, 70E, 87E (land only)

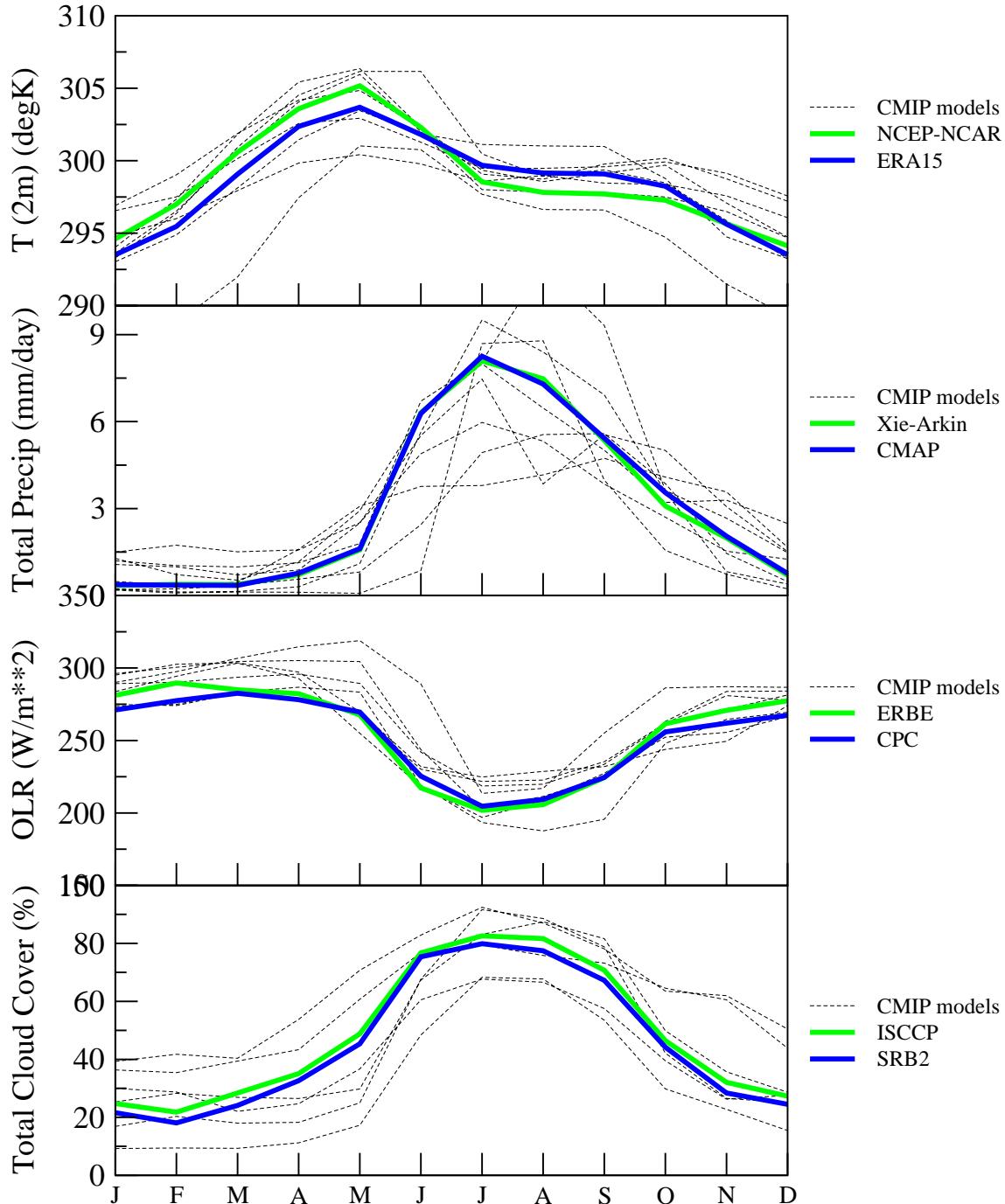


Fig. 4.23. India regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.

## **4.4.6 Europe/Middle East/ Northern Africa**

### **4.4.6.1 Surface (2 m) Air Temperature, Precipitation and OLR**

As expected, the largest amplitude of the annual cycle of surface air temperature occurs over the continents (Fig. 4.24). The mean model represents this well, though the amplitude over Europe, Egypt, and interior Russia is in places slightly underestimated. The model excellently captures the boreal summer phase over land north of 15°N and the phase lag over the ocean due to its larger heat capacity. South of 15°N the late winter/boreal spring phasing is well captured, though the amplitude is small.

The simulated annual cycle amplitude of precipitation is well simulated. However, that in the tropical convergence zone is underestimated, and that over the northern Atlantic is overestimated. Additionally, the simulated phase transition from boreal summer over Africa to boreal autumn over the Atlantic near 15°N to a winter maximum further north agrees with observations. Over Eurasia the model amplitude is weaker than observed, and the peak precipitation occurs too early.

The annual cycle amplitude and phasing of the OLR in the tropical convergence zone is consistent with the timing of the rainfall maximum. Alternatively, over the remainder of the continental land masses, where the annual cycle of rainfall is weak, it is the phasing of the temperature maximum that dominates the OLR. The mean model agrees well with the observations in terms of amplitude (though it is overestimated to the west of Spain and Portugal) and phase.

### **4.4.6.2 Sahel Seasonal Cycle**

The performance of the individual models over the Sahel is shown in Fig. 4.25. For surface air temperature the largest discrepancy among the models is during boreal winter, while for precipitation the largest spread occurs during the boreal summer rainy season. The OLR and cloud cover are consistent with the seasonal cycle of rainfall, but the inter-model spread is nearly uniform throughout the year.

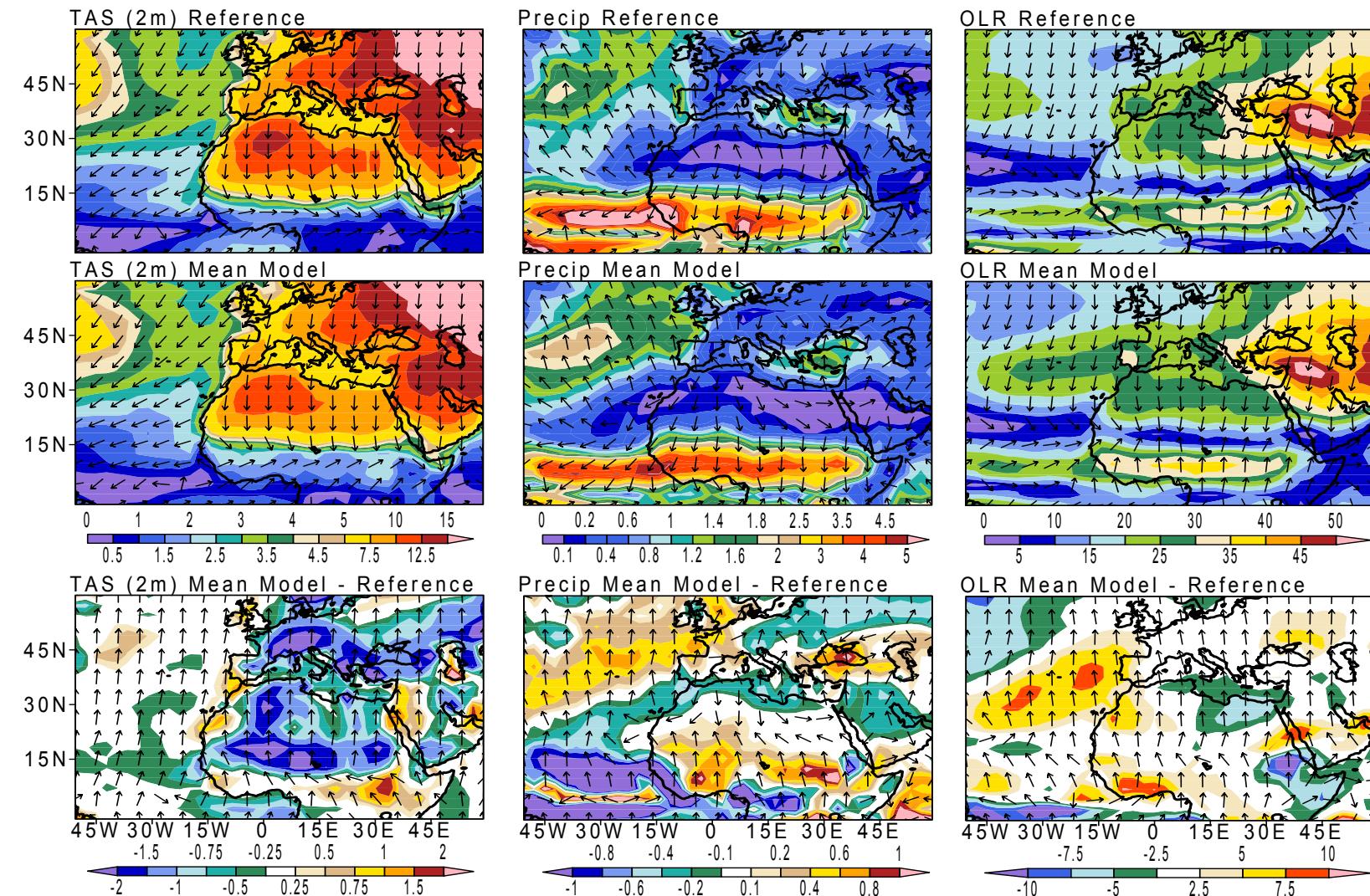


Fig. 4.24. Europe/Middle East/North Africa annual cycle amplitude (color) and phase (vectors indicate month of maximum): surface air temperature ( $^{\circ}\text{C}$ , left column), precipitation (mm/day, middle column), outgoing longwave radiation ( $\text{Wm}^{-2}$ , right column), reference data (top row), mean model (middle row) and mean model–reference (bottom row). In the top and middle rows the vector orientation is January at noon and April at 3 o'clock. In the bottom row the vector orientation is zero difference at noon and mean model leading the reference by 3 months at 3 o'clock. Vectors are only shown for every other grid cell. Note the nonlinear scales.

## Sahel

Averaging Domain: 12N, 18N, 15W, 38E (land only)

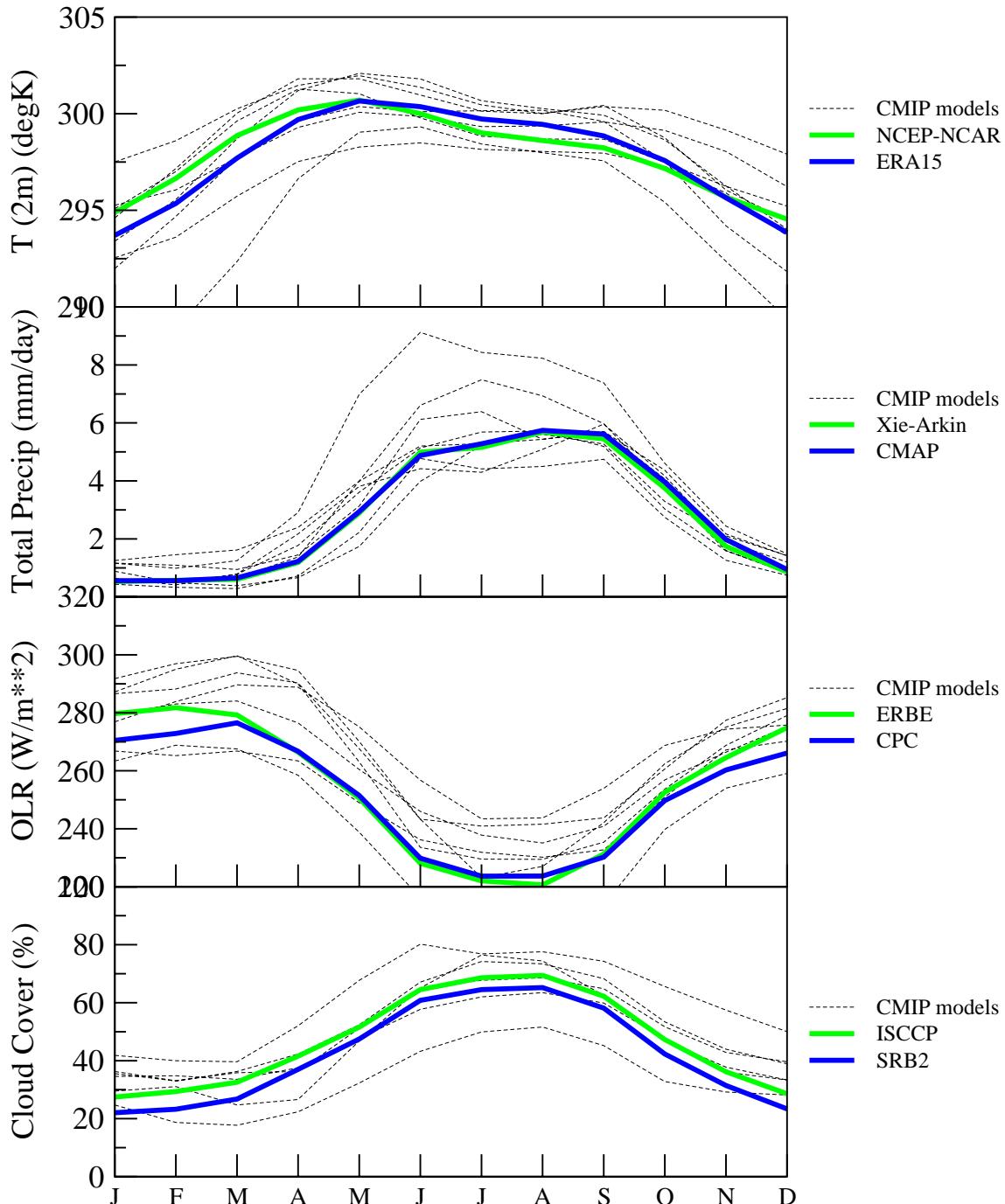


Fig. 4.25. Sahel regional-average time series for surface air temperature, total precipitation, outgoing longwave radiation and total cloud cover.

## 4.5 Tropical Biases in Rainfall, SST, and Windstress

### 4.5.1 Introduction

The goal of this section is to characterize the biases in the time mean and annual cycle of tropical rainfall with respect to sea-surface temperature (SST) and windstress. The analysis is geared toward examining the “split” intertropical convergence zone (ITCZ) in the Pacific Ocean in both coupled ocean-atmosphere general circulation models (GCMs) and atmospheric GCMs integrated with observed SST (e.g., AMIP II integrations). As demonstrated earlier in this section, comparison of AMIP with CMIP provides insight into the origins of coupled model behavior; this is especially true for tropical biases. Systematic biases in the Atlantic and in the subtropics are also briefly discussed.

### 4.5.2 The Models and Validation Data

Ten models from the CMIP2+ database are analyzed based on the last 10 years of contributed data. The flux adjusted and non-flux-adjusted models are grouped separately. Four AMIP II models that utilize the same atmospheric GCM as the CMIP2+ models are analyzed to assess the role that air-sea interaction plays in exacerbating the errors present in the AMIP II integrations. The AMIP II integrations and observed data are for the period 1979–95. All data have been analyzed on a  $2.5^\circ \times 2.5^\circ$  grid ( $0^\circ$ – $357.5^\circ$ ,  $90^\circ\text{S}$ – $90^\circ\text{N}$ ).

As in the previous sections, we use multiple observed datasets to assess model performance with respect to observational uncertainty. The rainfall is from the Climate Prediction Center Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1996, 1997), and the Global Precipitation Climatology Project (GPCP; Huffman et al., 1995, 1997). We use SSTs from Reynolds and Smith (1994) and the Hadley Centre (HadISST; Rayner et al., 2000, 2003). Finally, we employ the windstress from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al., 1996).

### 4.5.3 The Biases

Table 4.1 contains statistics of annual mean precipitation rate over the tropical Pacific ( $130^\circ\text{E}$ – $90^\circ\text{W}$ ,  $20^\circ\text{N}$ – $20^\circ\text{S}$ ). The GPCP rainfall is 0.83 mm/day drier than the CMAP rainfall. This difference is due predominantly to the bias component since the pattern correlation between the two datasets is 0.99. The bias is mainly due to lower rainfall estimates over the ocean, where GPCP rainfall is approximately 15% below atoll gauge data (G. Huffman 2000, personal communication). With the exception of HadAM3, the root-mean-squared difference (RMSD) of the simulated rainfall is greater than the observational uncertainty. HadAM3 also has the largest pattern correlation with the CMAP data. Compared with models that used flux adjustment, the non-flux-adjusted models typically have larger RMSD and lower pattern correlations for rainfall (Table 4.1) and SST (Table 4.2). Compared with their non-flux adjustment counterparts the AMIP II models (CAM2.0 and HadAM3) better represent rainfall. Conversely, the rainfall from the AMIP II models ECHAM4 and GFDL\_R30 was less well represented compared with their flux-adjusted CMIP2+ integrations.

Table 4.1. Annual mean rainfall rate ( $\text{mm day}^{-1}$ ), root-mean-square difference (RMSD), pattern correlation (R) and linear slope for  $130^{\circ}\text{E}$ – $90^{\circ}\text{W}$ ,  $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$  relative to CMAP (1979–95; ocean only). The models are grouped into those that did not use flux adjustment, those that used flux adjustment (shaded), and the AMIP II models (last 4 entries).

Data/Model	Mean	RMSD	R	Slope
CMAP	4.55	0.00	1.00	1.00
GPCP	3.72	1.17	0.99	0.75
CCSM2.0	3.97	2.01	0.79	0.55
HadCM3	4.50	2.32	0.80	1.01
PCM	4.35	3.21	0.35	0.27
CCCma_CGCM2	4.25	1.95	0.78	0.60
CSIRO_Mk2	4.43	1.54	0.88	0.66
ECHAM4_OPYC3	4.59	1.45	0.91	1.04
ECHO-G	4.44	1.42	0.89	0.84
GFDL_R30_c	3.81	1.88	0.83	0.66
HadCM2	5.18	1.84	0.87	0.96
MRI_CGCM2.3	4.37	1.46	0.89	0.86
CAM2.0	4.68	1.42	0.90	0.71
HadAM3	4.86	1.05	0.95	0.95
ECHAM4	4.86	1.67	0.86	0.86
GFDL_R30	3.67	2.13	0.78	0.61

Table 4.2. Annual mean sea-surface temperature (C), root-mean-square difference (RMSD), pattern correlation (R) and linear slope for  $130^{\circ}\text{E}$ – $90^{\circ}\text{W}$ ,  $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$  relative to Reynolds SST (1979–95; ocean only). Models that used flux adjustment are shaded.

Model	Mean	RMSD	R	Slope
Reynolds	27.26	0.00	1.00	1.00
HadISST	27.29	0.30	0.98	0.98
CCSM2.0	26.85	0.87	0.88	0.82
HadCM3	27.06	1.12	0.86	1.13
PCM	26.46	1.41	0.71	0.61
CCCma_CGCM2	26.45	1.00	0.93	0.85
CSIRO_Mk2	27.05	0.50	0.96	0.95
ECHAM4_OPYC3	27.09	0.36	0.98	1.01
ECHO-G	27.14	0.47	0.96	0.87
GFDL_R30_c	27.33	0.35	0.98	0.93
HadCM2	26.84	0.62	0.97	1.06
MRI_CGCM2.3	26.94	0.52	0.97	0.95

The climatological mean spatial patterns of the rainfall and windstress from CMAP and NCEP/NCAR reanalysis are given in Fig. 4.26a. Differences of the non-flux adjusted models minus observations are given in Fig. 4.26b-d. They all exhibit a “split” ITCZ with the equatorial western Pacific rainfall being underestimated in conjunction with a westward extension of the cold tongue (not shown) and easterly windstress anomalies. Excessive rainfall is found near  $10^{\circ}\text{S}$ , though the maximum can occur either in the western or eastern Pacific, and it tends to be associated with an overestimate of SST (not shown). Five of seven of the flux-adjusted coupled models (Figs. 4.27c-g) are also dominated by easterly windstress anomalies over the equatorial Pacific, though the sign of the rainfall anomaly depends on whether the anomalous flow there is convergent or divergent. In retrospect, the “split” ITCZ in the non-flux-adjusted models is evident in their AMIP counterparts (Figs. 4.28a-b), though air–sea interaction exacerbates and modifies the error structure. The error in the flux-adjusted models is similar to that in their AMIP integrations (Figs. 4.28c-d).

Examination of time–latitude plots of rainfall at  $170^{\circ}\text{W}$  indicates the “split” ITCZ to be present throughout the course of the annual cycle in the non-flux-adjusted models (not shown). At  $120^{\circ}\text{W}$ , a subset of the non-flux-adjusted models incorrectly have a shift of the ITCZ from the Southern Hemisphere to the Northern Hemisphere.

Additionally, errors that affect all models include the northerly windstress anomalies adjacent to the west coast of South America, and too much (little) rainfall in the southern (northern) subtropical Atlantic Ocean, in conjunction with the anticyclonic flow in the southern Atlantic being too weak.

## CMIP2+ non-flux adjusted

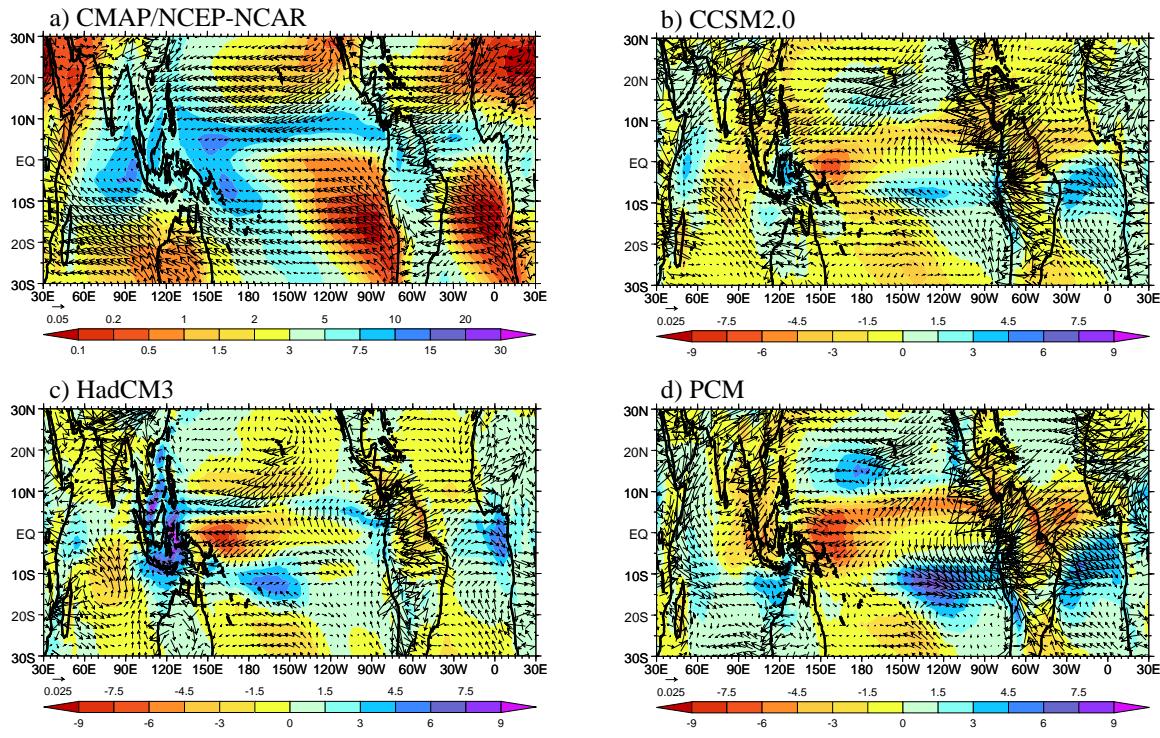


Fig. 4.26. (a) Annual mean CMAP rainfall and NCEP/NCAR surface windstress for 1979–95. Differences (non-flux-adjusted) coupled model–observations) of simulated rainfall and windstress between (b) CCSM2.0, (c) HadCM3, and (d) PCM, and the observations in (a).

## CMIP2+ flux adjusted

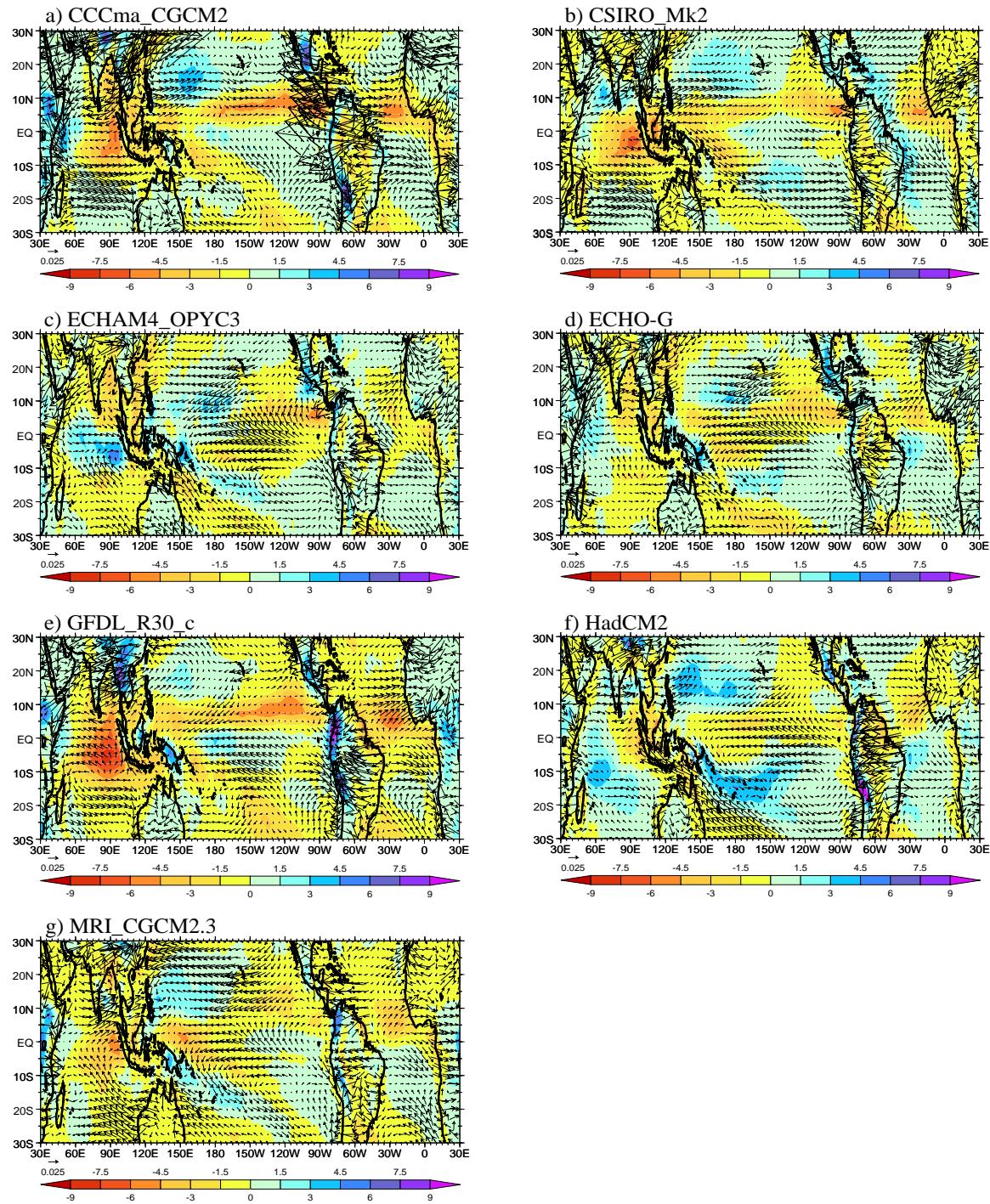


Fig. 4.27. Differences (flux-adjusted coupled model–observations) of simulated rainfall and wind stress between (a) CCCma\_CGCM2, (b) CSIRO\_Mk2, (c) ECHAM4\_OPYC, (d) ECHO-G, (e) GFDL\_R30\_c, (f) HadCM2, and (g) MRI\_CGCM2.3, and the observations in 4.26a.

## AMIP II

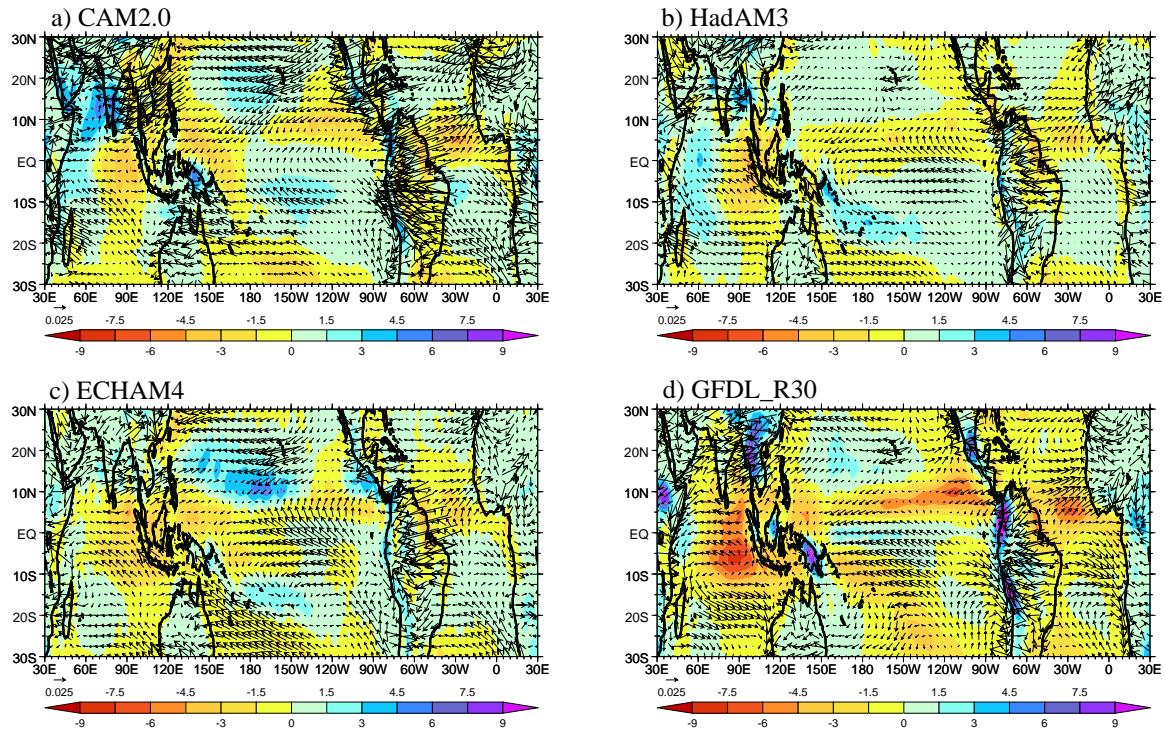


Fig. 4.28. Differences (AMIP II model–observations) of simulated rainfall and windstress between (a) CAM2.0, (b) HadAM3, (c) ECHAM4, and (d) GFDL\_R30, and the observations in Fig. 4.26a.

## 4.6 The Köppen Climate Classification

### 4.6.1 Introduction

Strictly speaking, GCMs do not directly simulate “climate.” Rather, the GCM solution, consisting of high-frequency (e.g., every 30 min) output of atmosphere/ocean state variables (e.g., winds and currents) and model-diagnosed physical forcing (e.g., precipitation), is filtered to extract lower-frequency “climate” variability and then only a subset and/or combination of these filtered data are used to define the climate state. Thus, climate modeling is application-dependent involving data reduction and selection, e.g., if one is interested in the global-scale, decadal trend of air temperature, then the full (and very large) GCM solution is averaged in space and time to produce a much smaller data set that defines the climate variable of interest.

In previous sections we examined climate variables that are closely related to the model solution and are thus more technical and model-centric. While such technical evaluations are essential for improving the physical realism of the GCMs, diagnostics that distill and translate the technical solutions into non-modeler terms are also important for impact assessments of climate change on society. Further, such diagnostics provide model developers and users with new points of view on model behavior.

In this section we consider an extreme form of data reduction, the Köppen (1923) climate classification (KC) system in which the GCM output is reduced to a single number. The KC *type* or *class* represents climate in a more geographical sense and attempts to relate the distribution of terrestrial plant species to the meteorological variables of surface air temperature and precipitation. The KC class also corresponds to geographical notions of climate such as “desert” or “tropical rainforest.”

The KC scheme has been used in climate model evaluations in the past (e.g., Lohmann et al., 1992, and Hogan, 1995) and while interest has been limited in the scientific community, we believe that a comprehensive appraisal of climate modeling should include KC-type measures of model performance.

We present results from the CMIP2+ project using an extended version of the KC scheme that not only shows the ability of the models to simulate present-day climate, but also gives practical estimates of climate change through a comparison of fixed- and varying- $\text{CO}_2$  model integrations.

### 4.6.2 The NRL Köppen Classification scheme

The KC algorithm takes for input a “climatograph” and outputs the KC class. The climatograph is a long-term, or climatological mean, annual cycle (as discussed in section 4.4) of surface air temperature and precipitation. In concrete terms, for each latitude-longitude model grid point we input 24 numbers (two variables and 12 months) and a single KC class is output. This output is in the form of a three-letter code (and numerical index). The first letter is the major class defining a very basic climate type. There are five major classes: **A**—Tropical; **B**—Dry; **C**—Temperate; **D**—Snowy; and **E**—Polar. Subsequent letters refine the climate classification within the major climate zone. For example, the part of California where PCMDI/LLNL is located experiences the **Cs** or “Mediterranean” climate with dry summers and rainy winters and, more specifically, the

**Csa** climate because the summer is long and hot. Table 4.3 gives a description of the 27 classes calculated by the KC scheme.

The specific KC algorithm used here comes from Hogan (1995) of the Naval Research Laboratory (NRL), Monterey, California with extensions that help summarize the results across the models. We will refer to this implementation as the NRL KC scheme. The NRL KC application uses the formulae of Geiger and Pohl (1954) to find the KC for two input climatographs—a test and a reference.

In addition to the output of two KC maps (a grid with an index for each class) for the test and reference, a “difference” class between the two is calculated. This difference class is found in two ways: 1) by perturbing the surface air temperature and/or precipitation of the test climatograph until its KC matches the reference KC; and 2) by selective replacement of the test climatograph with the reference. The perturbation/replacements are applied during the summer, winter and throughout the year to define the “seasonality” of the difference. If the modified KC matches the reference, and the modified climatograph is outside the reference by a specified bound, then the difference is defined as “big” and the nature of the climatograph difference defines the difference classification. For example, if the simulated summer air temperature is cooled by a bound of 1°C to produce a KC that matches the reference, and the perturbed climatograph temperature is outside the reference, then the difference class is “big” and “hot.” We used the same bounds as Hogan (1995) of 1°C for temperature and 5 mm/mo for precipitation.

The difference classification code is similar to the KC code and has 8 “big error” (BE) types: 1) hot (**H**); 2) hot/wet (**HW**); 3) wet (**W**); 4) cold/wet (**CW**); 5) cold (**C**); 6) cold/dry (**CD**); 7) dry (**D**); and 8) hot/dry (**HD**). The difference code also contains a seasonality indicator for: 1) summer (**s**); 2) winter (**w**); and 3) annual (**a**) to show when the modified climatograph matched the reference. For example, a BE of **HWa** would be “Hot” and “Wet” over the entire annual cycle (i.e., no summer/winter seasonality to the difference).

If the reference climatograph is based on observations, then the KC difference can be considered an “error.” However, climate-change integrations can be compared against their control integrations to assess the impact of the climate-change forcing. For example, in the case of increased CO<sub>2</sub>, the climate-change integration might show that the climate of southern Canada is changing from “snowy” (class **D**) to a milder and agriculturally more favorable temperate/cold (class **Cw**) climate. Thus, the NRL KC scheme has application to model both diagnosis and climate impact studies.

#### **4.6.3 KC of the CMIP2+ Mean Model**

The model climatographs come from the annual cycle taken over a 20-year period centered on year 70 (i.e., the 20-year period from years 60–79) of the GCM integration and this period is the same as used in the preceding annual cycle section. Although 20 years is somewhat shorter than the WMO standard of 30, we found little variation in the KC classes using 30- versus 20-year annual cycles.

Consider first the complete KC (all 27 classes as described in Table 4.3) for the “mean model.” The mean model is the ensemble average of 11 CMIP2+ model climatographs

(not a real GCM solution) and has been used in this appraisal to represent the class of coupled GCMs circa 2002.

Figure 4.29 shows the KC for the mean model and an observed climatograph that results from a merge of three observational data sets: 1) CPC (Climate Prediction Center) Merged Analysis of Precipitation (CMAP, Xie and Arkin, 1997) for precipitation over the ocean; the ECMWF ERA-40 (Simmons and Gibson, 2000) reanalysis for surface air temperature over the ocean; and Climate-Aided Interpolation (CAI, Willmott and Matsuura, 2000) for both precipitation and surface air temperature over land. The merged data provide global coverage, though emphasis will be on the behavior over land as the KC scheme was designed to relate surface meteorology to terrestrial plant communities.

The most striking aspect of the KC plot is the overall high degree of similarity between the mean model and the observations, signifying a largely successful simulation of the annual cycle of surface air temperature and precipitation. However, on closer examination there are some notable differences in the mean model: 1) excessive dryness in the Southern Hemisphere subtropics ( $\sim 30^{\circ}\text{S}$ ) as the observations are more steppe climate (**Bs**) over Australia where observations show a desert (**Bw**) climate; 2) greater wetness over Mexico, the southwest U.S. and Southern Africa; and 3) large differences over areas of high elevation.

To highlight these differences, the difference class from the NRL KC scheme is displayed in Fig. 4.30 only for **BE**. Consistent with the mean model KC in Fig. 4.29, the mean model **BE** over the Southern Hemisphere subtropical ridges is wet (Fig. 4.29, top panel, green blocks), with the error occurring predominantly during austral winter and over the annual mean in the seasonality plot (Fig. 4.30, bottom panel). We also note that the northern tropical Atlantic is too dry (yellow blocks), and that this error tends to occur in the boreal winter.

#### 4.6.4 The Transfer Matrix

The difference classification is a unique output from the NRL KC application, but another way of characterizing KC differences is the “transfer matrix” (Lohmann et al., 1992). This matrix is a contingency table with the percent surface area where the model and observed KC agree (along the diagonal) and disagree (off-diagonal) in the five major KC classes; for example, the percent area where the model is **B** (steppe), but the observation is **A** (tropical). The off-diagonal components represent climate shifts or disagreements in KC class.

The transfer matrix for the mean model and observations over land is given in Table 4.4, along with the percent area in each major class (last two columns). The largest discrepancy is for class **B**, associated with desert and steppe conditions, where the mean model is cooler and drier than observed. The area of disagreement is 5.3% (26.6%–21.3% of the total land area in the last two columns). Values above the diagonal indicate regions where the model climatograph is in a class closer to **A** (more tropical), but where the observations are in a class closer to **E** (polar). Similarly, values below the diagonal indicate areas where the model is cold (toward **E**), but the observed climate is warmer (toward **A**).

In Fig. 4.31 we display the total percent area above(warm) / on(correct) / below(cold) the diagonal for each of the CMIP2+ fixed CO<sub>2</sub> integrations (the Control), and each of the CMIP2+ climate change runs (increasing CO<sub>2</sub> 1%/year called Perturb) relative to the merged observational data set. This summing of area reduces the full transfer matrix (27×27) to 3 numbers and is a global measure of the KC differences visually seen in Fig. 4.29. We also show the mean and median CMIP2+ models, the ERA-40 reanalysis (1980–1999), and two observed data sets for surface air temperature and precipitation. The values for ERA-40 and the two observations provide a measure of observational uncertainty against which the significance of model errors can be gauged.

For the Control versus observations we find that 80% of the model land area is on the diagonal, whereas the various observations show a commonality of 85–90%. That is, the models correctly simulate the climate class over 80% of the land area which is quite impressive given the differences in the treatment of land surface hydrology among the models. However, the models also tend to be shifted toward cooler/wetter climates (blue bars or below-the-diagonal), consistent with the mean model results in Table 4.4, in which the area below the diagonal is 9.9%, while area above is 4.9%.

The bottom panel of Fig. 4.31 shows the same three-component KC difference as in the top panel, but now for the idealized climate-change simulations (Perturb) compared with present-climate observations. The input climatographs in the Perturb run come from the same model years as used for the control (years 60–79). The off-diagonal area is now more evenly distributed between above and below area, or a shift, relative to the control in the top panel, towards warmer climates.

To more clearly show the warming shift, we set the test to the Perturb run and the reference to the Control when running the NRL KC application so that the resulting transfer matrix measures differences relative to the control integration, i.e., model climate change forced by increasing CO<sub>2</sub>. In Fig. 4.32 we find the area above the diagonal is almost uniformly 10% with virtually no cooling area (blue bar). The response to the CO<sub>2</sub> forcing, in a climate shift sense, is very consistent among the models.

#### **4.6.5 BE—Big Error**

While the off-diagonal area in the transfer matrix gives a general sense of the KC difference between the test and reference, no information on physical climate change is available, such as “wetter” (more precipitation) or “warmer” (higher surface air temperature). For physical characterization of climate differences we use the special output of the NRL KC scheme—the land surface area with BE class. In the transfer matrix plots we displayed three values, but for the BE summary plot in Fig. 4.33 we show four; namely, the percent area with big “Hot” (red); “Dry” (yellow); “Wet” (green) and “Cold” (blue). Note that the total area of BE is not the same as the sum of the off-diagonal areas of the major classes, as the BE classification is based on the all 27 KC classes. For example, a grid box with a major class of “C” (temperate) in both the test and the reference may still have a BE class due to excessive summer precipitation.

As in Fig. 4.31, we show the percent land area with BE for the Control (A or top panel) and Perturb (B or bottom panel). In addition to the four BE areas, the total area with BE is printed above the model label, and in the case of the mean model (“Mean\_CMIP”), 19% of the total land surface area has BE and most of this error is of the too-wet (9%

green) variety. The nature of the KC differences between the models is now more apparent.

Some model Control runs in Fig. 4.33 have strong cold errors (long blue bars) with little Hot error while others have a higher fraction BE area in wetness errors (green and yellow bars). However, as was found in the transfer matrix plots in Fig. 4.31, the relative change between the Control (top panel) and the climate-change Perturb run (bottom) is clearly toward more area with Hot (red bars) BE. Some increase in area with dry BE (more dry, yellow, in perturb versus control) can be seen, but the strongest signal is in surface air temperature. Note also that the mean model nicely captures the general response of the individual models, and more impressive is the closeness of the mean model to the ECMWF reanalysis (“era40”).

Reanalysis can be viewed as a CMIP integration forced by both ocean (sea surface temperature and sea ice concentration) and atmosphere observations. The finding that the mean model land area with BE is only 4% greater than in reanalysis (19% versus 15%) suggests that a large portion of the basic climate of the mean GCM is captured by the model physics or that the annual cycle of precipitation and surface air temperature can be simulated by coupled GCMs. This mean model-reanalysis difference is even more remarkable given that the observations (“CAI\_LW” and “CAI\_CMAP”) show total BE area on the order of 5–10%, or that mean model-reanalysis difference is on the order of observational uncertainty.

The problem, however, is that the mean model does not exist (i.e., is not a computer code that can be run) and that there is a wide range in the distribution of BE among the individual models. The success of the mean must come from a compensation of errors; the modeling challenge is to find and correct these compensating errors in the individual models.

As we found in the transfer matrix plots (Fig. 4.32) for the perturb run relative to the control, we see a clear shift toward warmer climates in Fig. 4.34. The land area with Hot BE (not really an error since we are comparing the climate-change Perturb run with the Control and not with observations) ranges from 5–15% (the red bar) with very little precipitation difference class and no model with areas of Cold (blue) differences. Thus, the model climate change is predominantly a warming with some drying. However, the model response to the increasing CO<sub>2</sub> varies by a factor of 3, from 5 to 15%.

#### 4.6.6 Major Class Distribution Error

We next characterize the KC differences using the difference in total land area between the model and the observed KC for the five major classes, i.e., to difference the last two columns of the transfer matrix table giving the total percent area model and observed (see Table 4.4). This difference indicates both a mismatch in major KC class and errors in distribution among the classes. For example, if the difference in area of **A** (tropical) is -5% and for **B** (steppe) +5%, the model would have too little **A** and too much **B** or an error in spatial distribution of the climate zones.

This distribution error is plotted in Fig. 4.35 with color-coded bars for the 5 major classes. Bar length indicates the magnitude of the difference and the type or direction of the shift indicated by plotting below the zero line (off-diagonal class shift toward polar

conditions) and above (toward tropical) the zero line. For the mean model, we see virtually no difference in the total area of **A** (tropical, purple) and **E** (polar, blue), but about a -6% difference in the dry class (**B**) and an overestimate of the temperate (**C**) and snowy (**D**) areas relative to observations. As in the BE comparison we see a wide range of differences in both magnitude and direction for the individual models, though few models have too much tropical (**A**) area.

A curious result is that the differences (length of the bars) tend to decrease in the climate-change simulations, suggesting that the increasing CO<sub>2</sub> is driving the models to a climate closer to the observed current climate. This is consistent with the general too cool/wet bias seen in the on/above/below diagonal elements of the transfer matrix for the Control runs in the top panel of Fig. 4.31. Also, the models tend to lose polar area (**E** or blue < 0) in the warmed climate. However, the difference area actually increases for the mean model, which is interpreted as the mean of the Perturb runs moving away from the current climate. This response suggests that the mean model may be a more robust indicator of forced climate change than may be found from individual models.

#### 4.6.7 Summary of KC results

We have applied an extended version of the Köppen climate (KC) classification system to 11 CMIP2+ Control (fixed CO<sub>2</sub>) and Perturb (1%/y CO<sub>2</sub> increase) integrations, and the 11-member ensemble mean, to characterize model “climate” in geographical terms relevant to both the GCM modeling community and non-modelers. As a group, the CMIP2+ models successfully simulated the five major KC classes over ~80% of the land surface area compared with current-climate observations.

The KC scheme used here also calculates a difference class between two KC outputs that diagnoses the nature of difference (e.g., too Hot; too Dry) and its seasonality (e.g., summer time). The fraction of land area with large difference class errors is ~30% with the mean model showing significantly less (19%) area. The higher accuracy of the mean model suggests that there is a compensation of errors among the 11 models and that the mean model would be a good benchmark for measuring GCM improvement.

The KC scheme was also used to measure climate change by comparing Control and Perturb experiments. Both the difference class and shifts in area covered by the major KC classes showed a consistent response of the models to increasing CO<sub>2</sub>—warming with little change in precipitation. However, the net area of substantial warming (area with a Hot difference class) varied among the models by a factor of 3.

Table 4.3. Description of the Köppen Climate Classes—major (5), main (11) and full (all 27).

Major	Main	Full	Description
<b>A</b>		<i>Tropical</i>	
	<b>Af</b>	<b>Af</b>	rainforest
	<b>Am</b>	<b>Am</b>	rainforest/monsoon
	<b>Aw</b>	<b>Aw</b>	wet dry tropical savanna
<b>B</b>		<i>Dry</i>	
	<b>Bs</b>		<b>steppe (semi-arid, grass covered plain)</b>
		<b>Bsh</b>	dry, hot steppe
		<b>Bsk</b>	dry, midlatitude steppe
	<b>Bw</b>		<b>desert</b>
		<b>Bwh</b>	hot, desert
		<b>Bwk</b>	cool, desert
<b>C</b>		<i>Temperate—moist with mild winters</i>	
	<b>Cw</b>		<b>dry winter, wet summer</b>
		<b>Cwa</b>	long and hot summers
		<b>Cwb</b>	long and cool summers
		<b>Cwc</b>	short and cool summers
	<b>Cs</b>		<b>Mediterranean—hot summer, wet winter</b>
		<b>Csa</b>	long and hot summers
		<b>Csb</b>	long and cool summers
		<b>Csc</b>	short and cool summers
	<b>Cf</b>		<b>wet all seasons</b>
		<b>Cfa</b>	long and hot summers
		<b>Cfb</b>	long and cool summers
		<b>Cfc</b>	short and cool summers
<b>D</b>		<i>Snowy—moist with cold winters</i>	
	<b>Dw</b>		<b>dry winter</b>
		<b>Dwa</b>	long and hot summers
		<b>Dwb</b>	long and cool summers
		<b>Dwc</b>	short and cool summers
		<b>Dwd</b>	short summer, severe winter
	<b>Df</b>		<b>wet all seasons</b>
		<b>Dfa</b>	long and hot summers
		<b>Dfb</b>	long and cool summers
		<b>Dfc</b>	short and cool summers
		<b>Dfd</b>	short summer, severe winter
<b>E</b>		<i>Polar</i>	
		<b>Et</b>	polar
		<b>Ef</b>	polar

Table 4.4. Transfer matrix between the mean model (Model) and observations (Obs) in % surface *land* area for the major KC classes. Values on the diagonal (shaded) are the % area the Model and Obs classes agree; values above the diagonal where the model class is lower (toward **A**); and values below the diagonal elements where the model class is higher (toward **E**). Above-diagonal area can be interpreted as a shift toward a warmer, moister climate (**E**→**A**), and below diagonal area as a shift toward a cooler, dryer climate (**A**→**E**). The total area in each class for the Model and Obs is shown in the last two columns. For example, the observed total area for class **B** (dry) is 26.6%, but for the mean model it is 21.3%, or 5.3% less than observed.

Class	Obs					Total Area	
	A	B	C	D	E	Model	Obs
Model	A	18.5	1.0	0.8		20.2	20.4
	B	0.3	20.3	0.6	0.1	21.3	26.6
	C	1.6	2.8	13.0	0.6	18.1	15.3
	D		2.4	0.9	19.6	1.6	24.5
	E		0.2		1.7	13.9	15.8
							15.7

Koeppen Class (full): mean-c03a :: CONTROL yr=60–79  
 Obs pr: CMAP (CPC) ; tas: ERA40(ocn)+CAI(land)  
 XfrMtx: type(#)[Below:On:Above] :: MAJOR(5)[ 5:92; 3] MAIN(11)[10:79:11] Below->Cold;Above->Warm

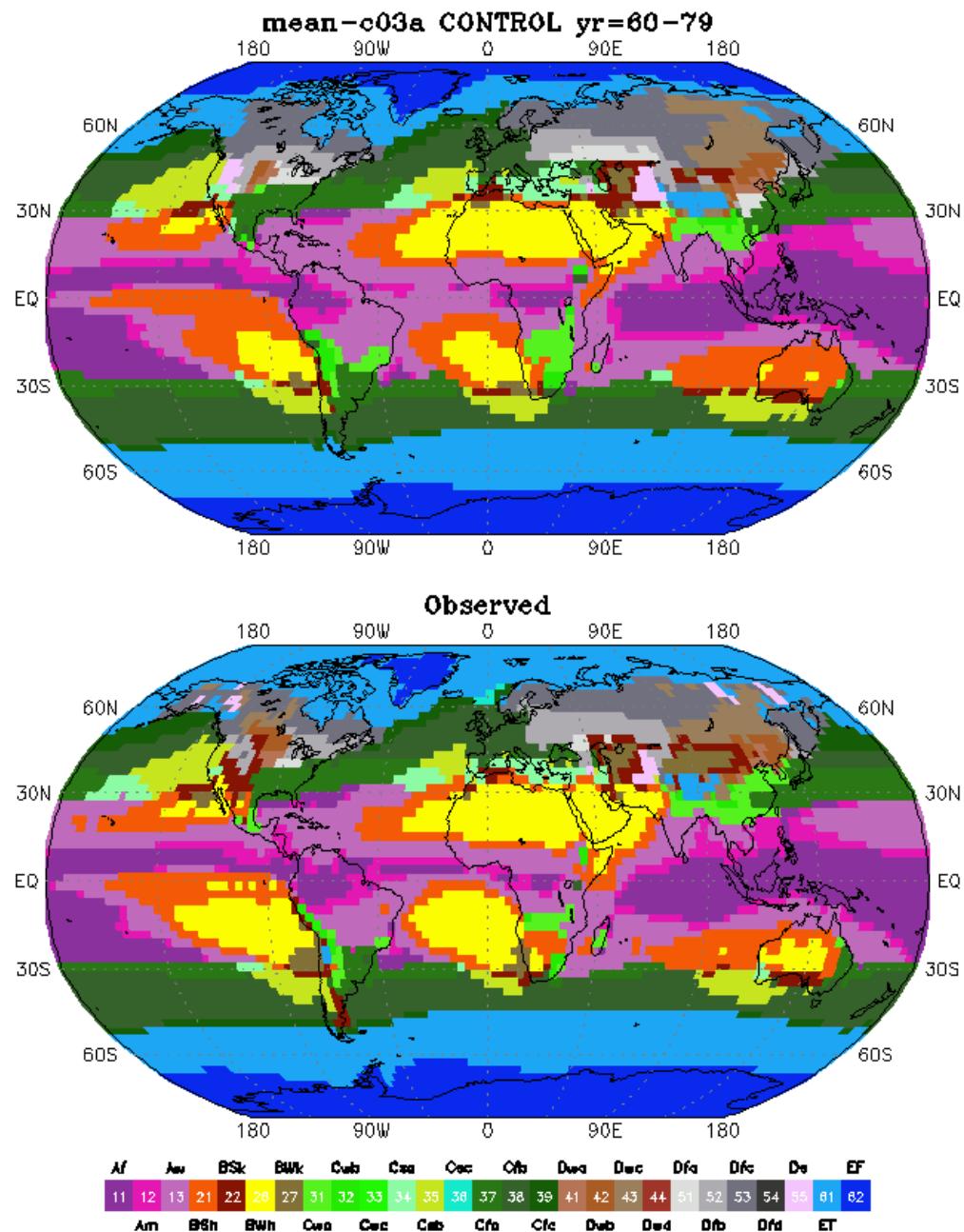


Fig. 4.29. Full (27) KC classes for the mean model (top panel, “mean-c03a”) for 20-y annual cycle of model years 60–79 and the observational standard (bottom panel “Observed”).

**Climatograph Error: mean-c03a :: CONTROL yr=60-79**

Obs pr: CMAP (CPC) ; tas: ERA40(ocn)+CAI(land) pr<sub>r</sub>=5.0 tas<sub>r</sub>=1.0

BigError: type(%area) :: Wet( 6) Dry( 2) Hot( 1) Cold( 2) :: Ann( 5) Sum( 8) Win(13)

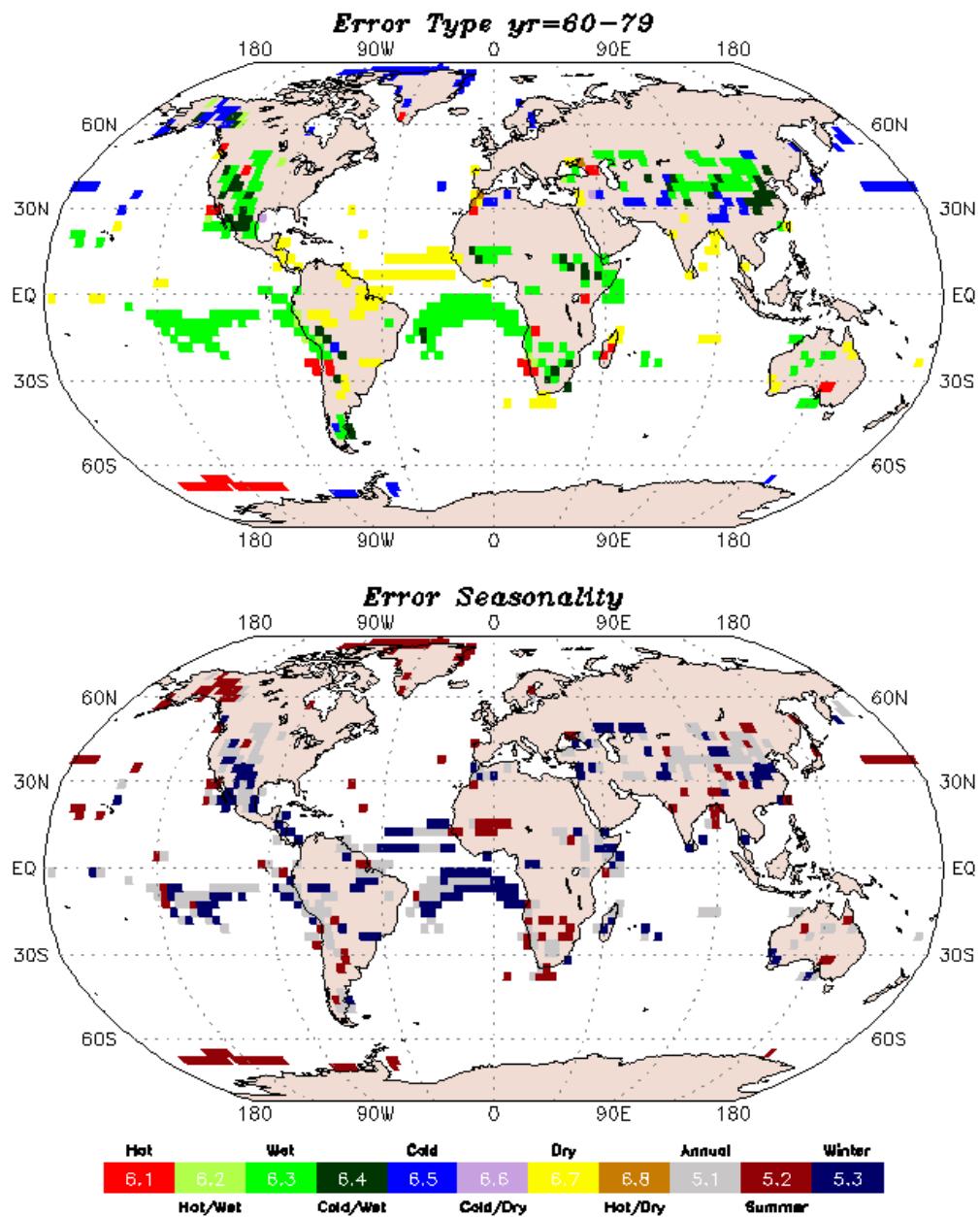


Fig. 4.30. The 8 BE (big error) types (hot, hot/wet, wet, cold/wet, cold, cold/dry, dry, hot/dry) and seasonality (annual, summer, winter) for the mean model.

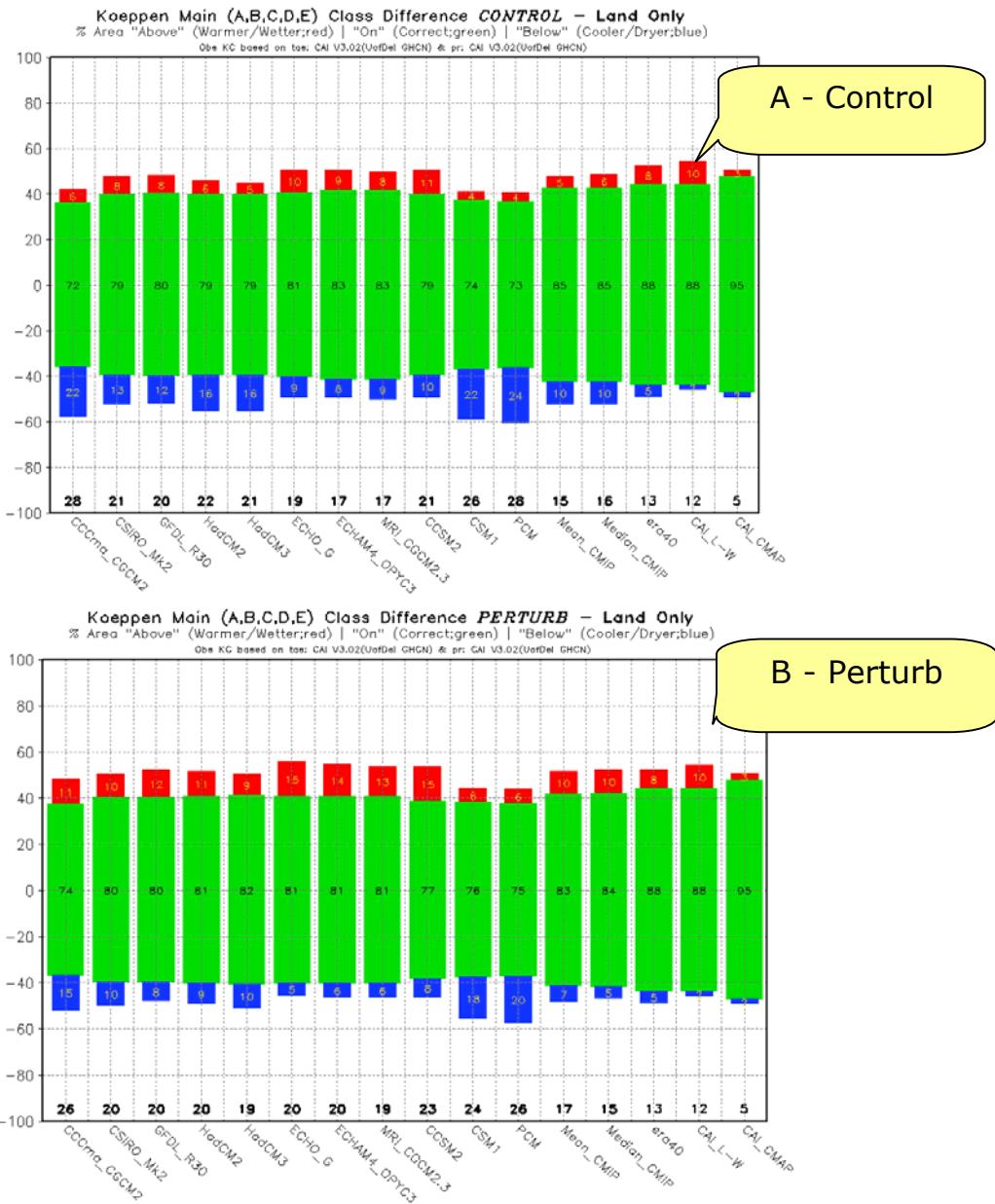


Fig. 4.31. Comparison of Control (constant CO<sub>2</sub>, A) and perturbation (increasing CO<sub>2</sub>, B) runs for 11 models and the mean (mean-c03a) and median model (median-c03a), based on the 20-year annual cycle for years 60-79 over *land only*. The length of the bar proportional to the percent total area *on* is (green), *above* (warmer/drier/) is red) and *below* (cooler/wetter, blue) the diagonal of the transfer matrix. Also displayed are the ERA-40 reanalysis ("era40") and two observational KCs "CAI\_L-W" (CAI v Legates and Willmott [1990a] surface air temperature) and "CAI\_CMAP" (CMAP v CAI precipitation). Total of the blue and red area (off-diagonal error) is printed above the model name and error > 15% is bold.

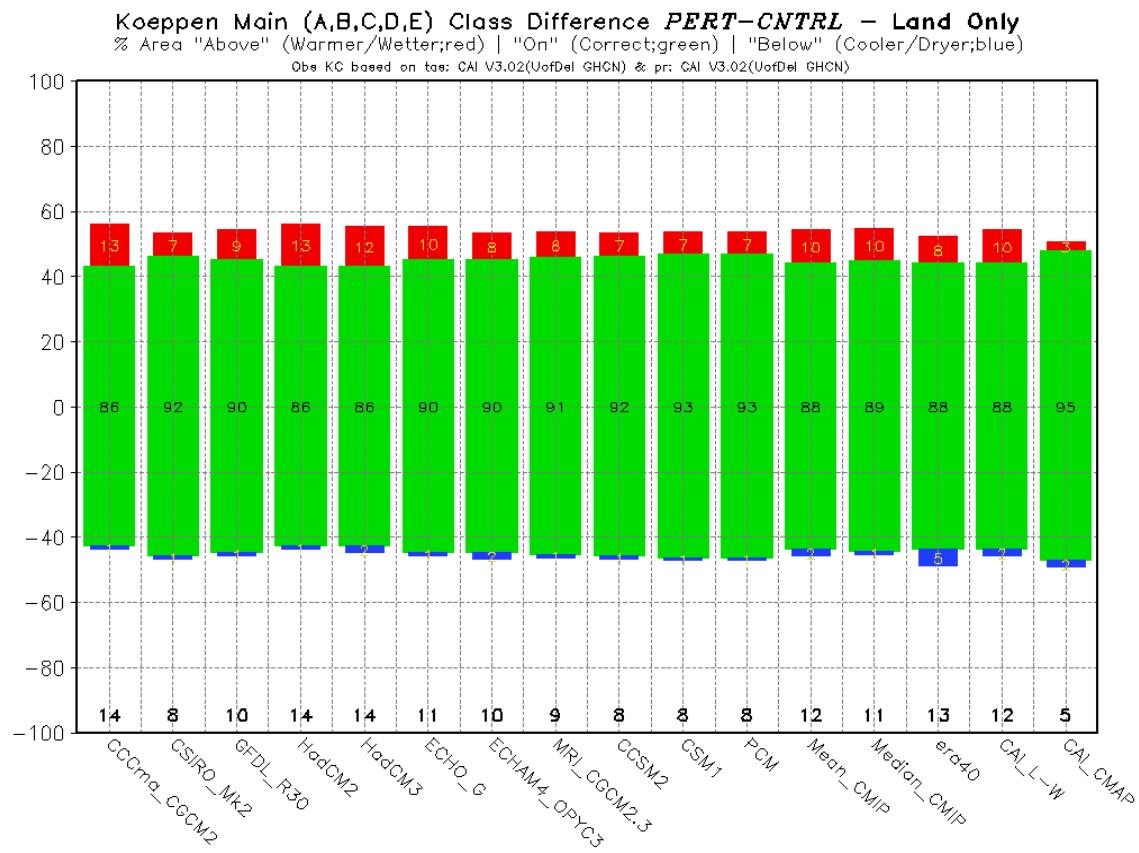


Fig. 4.32. As in Fig. 4.31 except for the perturbation (test input climatograph) versus the control model (reference climatograph).

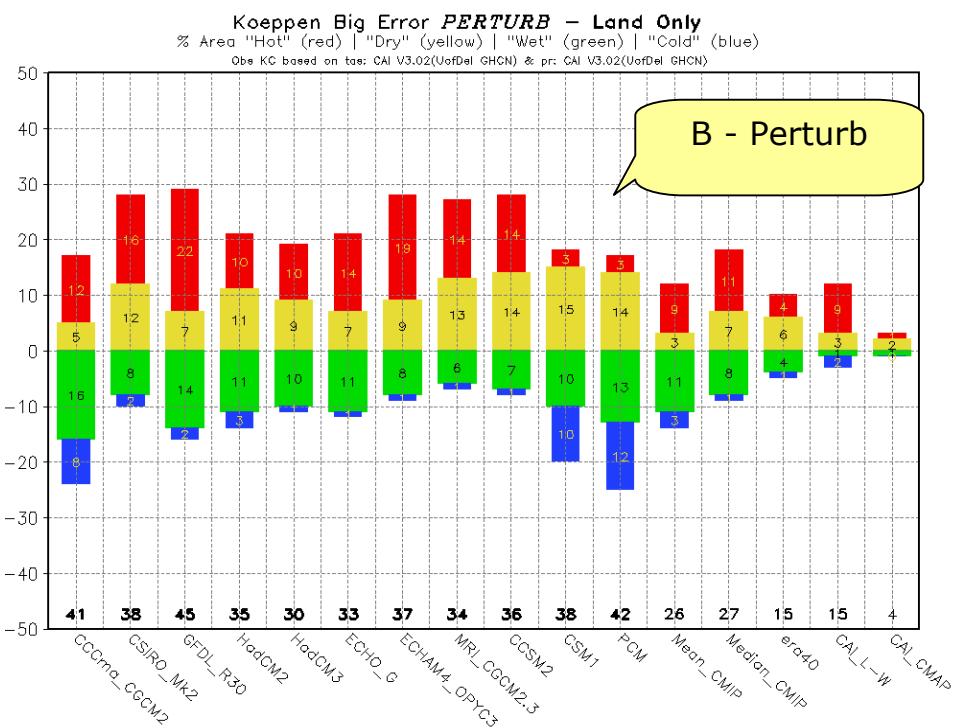
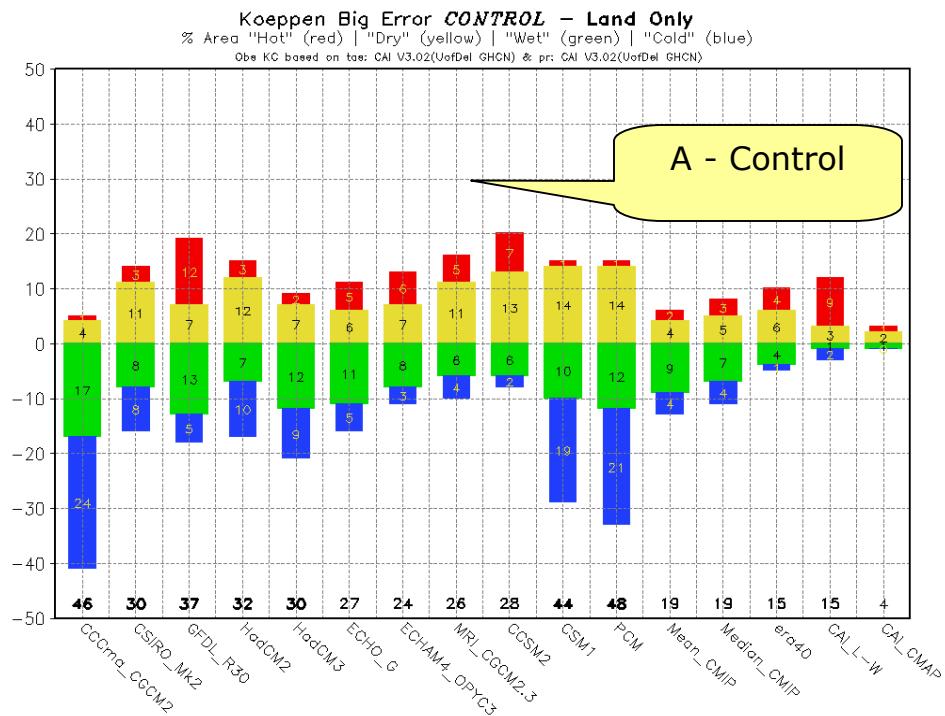


Fig. 4.33. As in Fig. 4.31 except for % area with big errors: 1) hot/red, 2) dry/yellow; 3) wet/green; and 4) cold/blue.

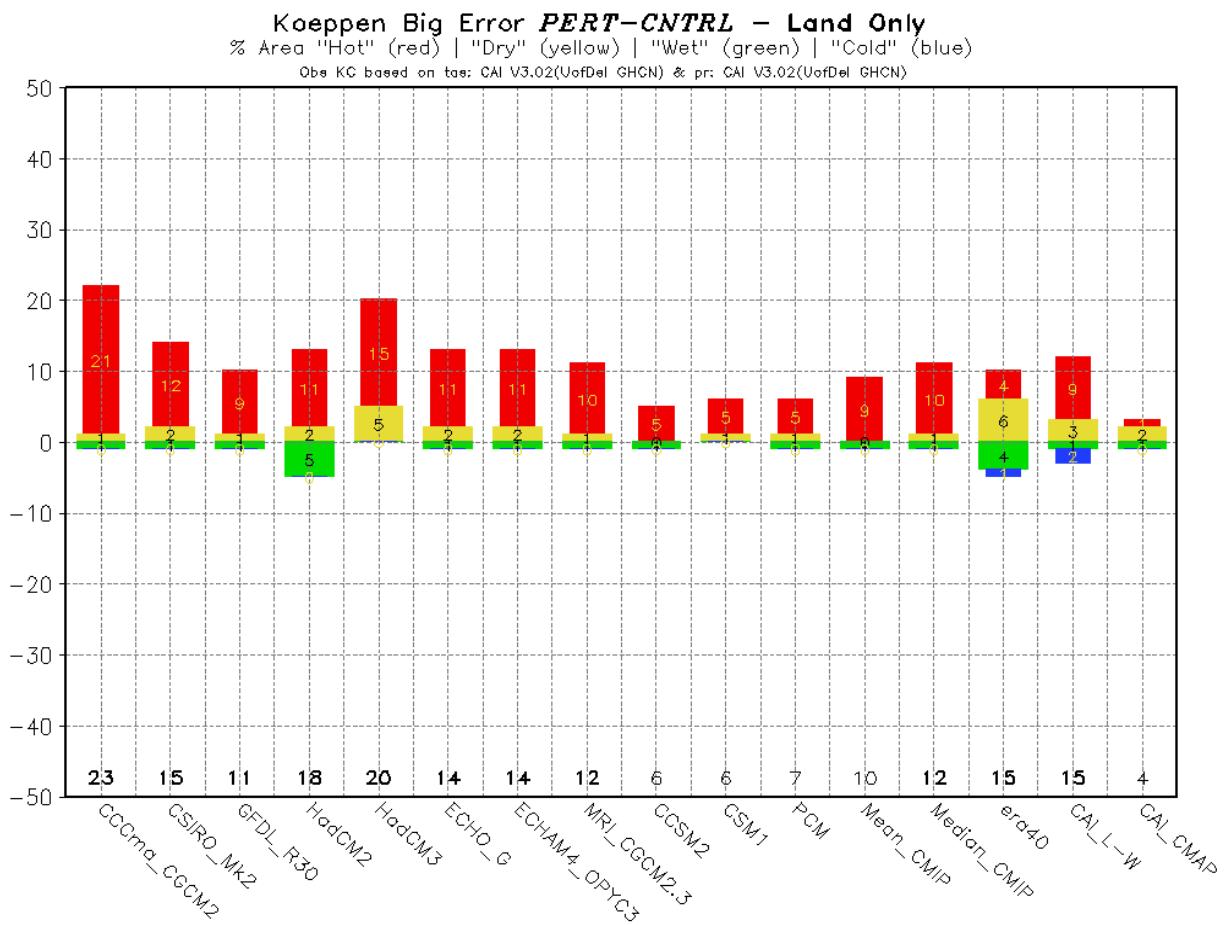


Fig. 4.34. As in Fig. 4.32 except for big errors.

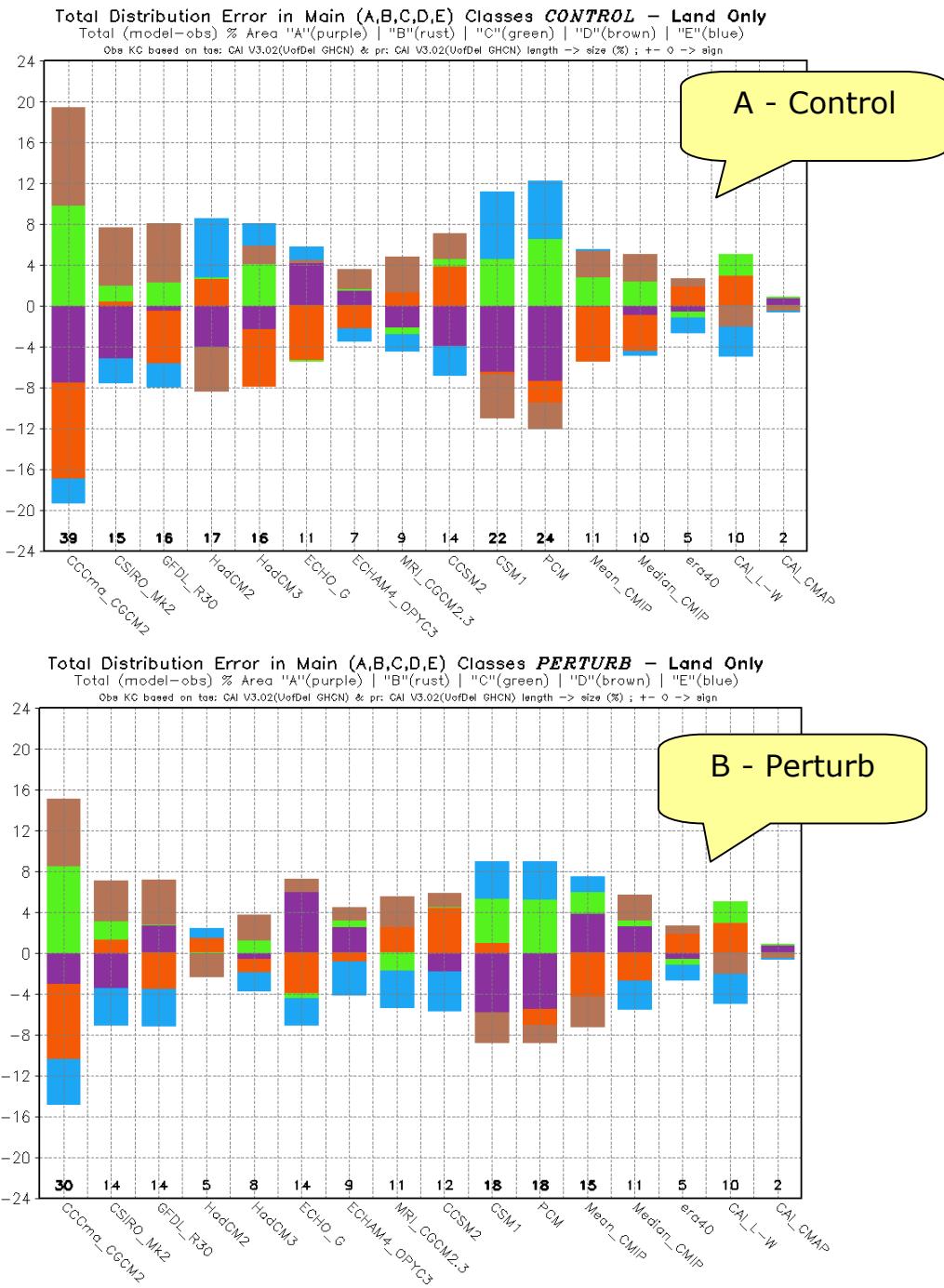


Fig. 4.35. As in Fig. 4.31 except for total distribution error by the 5 major KC classes (A = purple, B = rust, C = green, D = brown, E = blue). Size of the bar indicates the magnitude of the difference and the position of the bar above (more area in the Model compared to the Obs) and below (less area) the 0 % line indicates direction, i.e., bars above 0 mean “too much” and below “too little” area of the color-coded class. The total error (sum of the magnitudes of the 5 errors) is listed above model name on x axis with bold number indicating total distribution error >15%. Note that the length of the bar above and below 0 is the same (total difference is 0.0).

## 5 The Mean State of the Ocean and Cryosphere

*P. Gleckler, K. AchutaRao, K. Sperber*

### 5.1 Introduction

With an increasing focus on evaluating results from century-scale coupled ocean-atmosphere integrations, diagnosis of the simulated ocean is a new but important priority for PCMDI. In this section we take a first step, by examining the mean state of the simulated ocean and sea-ice from the models described in Section 2. Although the analysis below is limited to basic features of the major ocean basins, to date the ocean component from a large number of coupled models has seldom been intercompared in such detail.

For consistency with the atmospheric results, we restrict our analysis to climatologies of the same 20-year period used in the evaluation of the atmosphere component of the same models (see Section 4.1). We utilize the data sets of temperature and salinity (Levitus et al., 1998), hereafter referred to as WOA98. The WOA98 data are available on a  $1^\circ \times 1^\circ$  latitude-longitude grid at 33 depth levels, and include a mask that defines different ocean regions and seas. For the zonal average versus depth plots shown in this section we have interpolated the model climatologies to the WOA98 latitude-longitude-depth grid, but for all others we make use of the model data on the original grid.

In Section 4 our look at the mean climate of the atmosphere was restricted to the multi-model ensemble mean. For several reasons, here we have reverted to looking at results from individual models. Intercomparisons of the atmospheric mean state are now fairly routine, many existing biases are well known, and the mean model helps to consolidate results. For the simulated ocean, however, such comparisons are relatively new, and it is instructive to become familiar with where individual models generally agree and disagree. Additionally, in the atmosphere we have demonstrated that the mean state of the multi-model ensemble compares better with observations than any individual model, something that has yet to be established in ocean models, and which will likely prove more challenging because of the paucity of observations.

A major caveat about the present results is that the ocean component models have been run for varying time periods from multiple decades to many centuries. Efforts are under way (see Section 2) to document details of the spin-up procedures and length of these simulations. Presumably some of the models were initialized with Levitus data, and since the time scale of the deep ocean is many centuries, short integrations will differ little from Levitus at depth. Only long integrations will show the good or bad effects of modeled circulation and mixing of the deep properties. The intercomparison of the coupled models is, however, less dependent on time integration in the uppermost ocean. In future PCMDI appraisals, identical time periods of ocean integration will enable such models to be more carefully compared.

In the following sections we take a cursory look at the simulated Atlantic, Pacific, Indian, Arctic and Southern Oceans (as defined by WOA98). In each case, we examine the time-mean potential temperature<sup>3</sup> and salinity as a function of latitude (zonal mean) and

---

<sup>3</sup> For brevity, we will often refer to potential temperature as temperature.

depth, leaving an examination of density for a more in-depth model evaluation. Several additional sections of particular interest will also be shown, including time-averaged ocean currents. For the case of the tropical Pacific, we will compare the zonal currents of the models with data from the Simple Ocean Data Assimilation (SODA). SODA is an application of data assimilation using a forecast model based on GFDL ocean physics driven by observed historical meteorology—winds, heating, and rainfall-evaporation—assimilating historical observations of temperature, salinity, sea level, SST, and surface current (Carton et al., 2000a, 2000b).

At the ocean surface, all circulations are ultimately driven by wind stress and fluxes of heat and freshwater. The tropical wind stress from many CMIP2+ and AMIP simulations was highlighted in Section 4.5. Before looking at the simulated ocean at depth, we first examine the 20-year-time mean sea surface temperature and salinity. In Fig. 5.1 the difference between the Levitus SST climatology and the simulated surface temperature is shown for most of the CMIP2+ models. The three models without flux correction (shown in the right hand column) clearly have larger biases than the flux-corrected models. In some areas of the Southern Ocean a few of the flux adjusted models appear to deviate from the observations by 2°C or more. Note the large differences in the sea-ice regions are due to the fact that we are comparing each model’s surface skin temperature on the ice surface with the observed estimate of the surface temperature (below the ice). The same calculations were done comparing the models with the AMIP SSTs (not shown), yielding almost identical results. In Fig. 5.2, surface salinity difference maps (models—Levitus) are shown for our chosen 20-year period. Here the non-flux-adjusted models are shown in the top row, and there is one model for which the data is not available. Note the large differences between nearly all the models and Levitus in the Arctic. In the tropics and mid latitudes, the models without flux adjustments stand out more clearly than in Fig. 5.1.

Unfortunately, we do not have all components of the surface energy balance for most of these simulations, and therefore cannot present the implied heat transport in each ocean basin. Other derived ocean fields such as meridional overturning streamfunctions, heat and salt transports, and global maps of mean kinetic energy at key depths are planned for future appraisals.

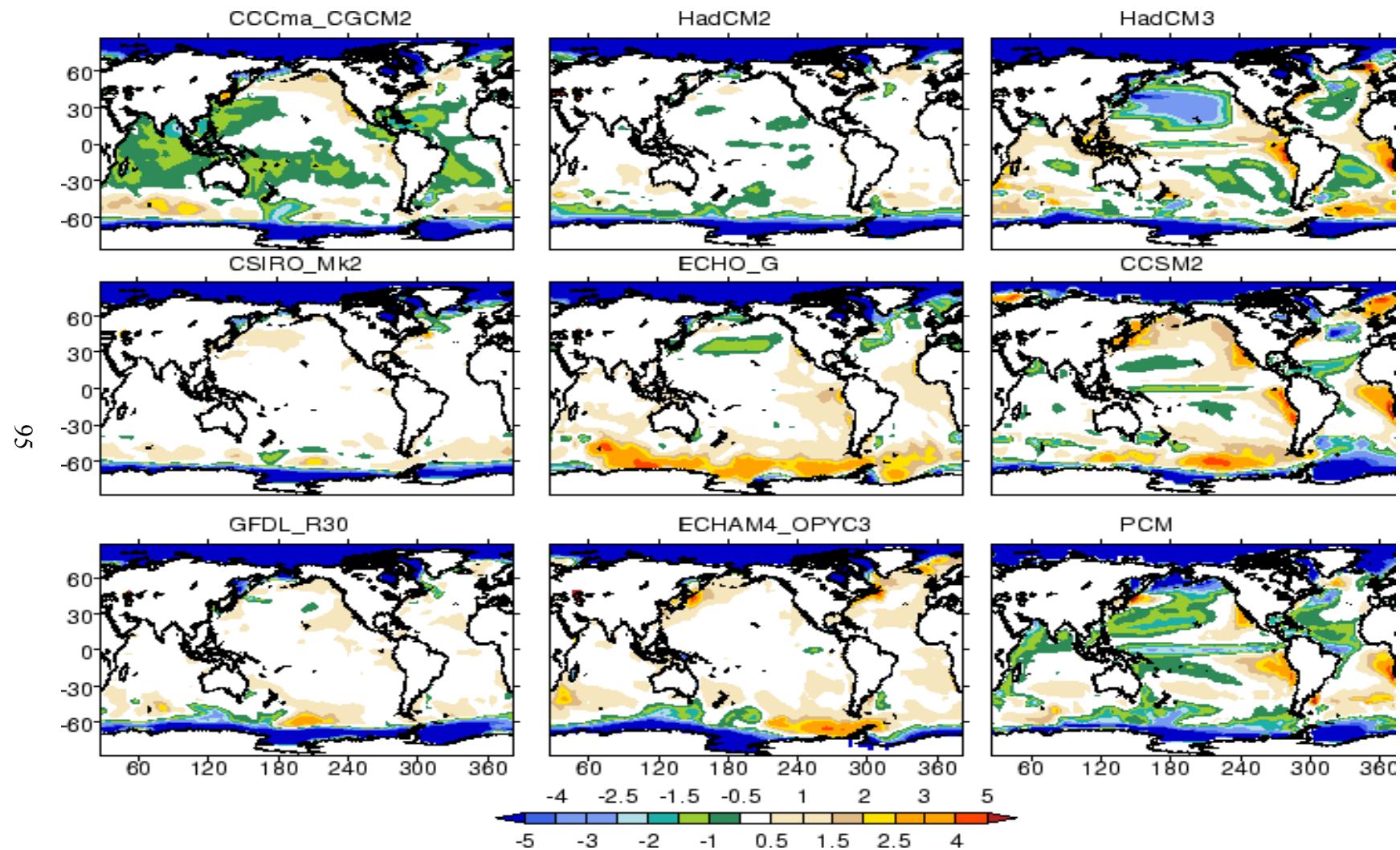


Fig. 5.1. Sea surface temperature time mean difference between the CMIP2+ models and Levitus (degrees C). Note the large values at high latitude are a result of comparing the model skin temperature on top of sea-ice vs. the observed estimate below the ice.

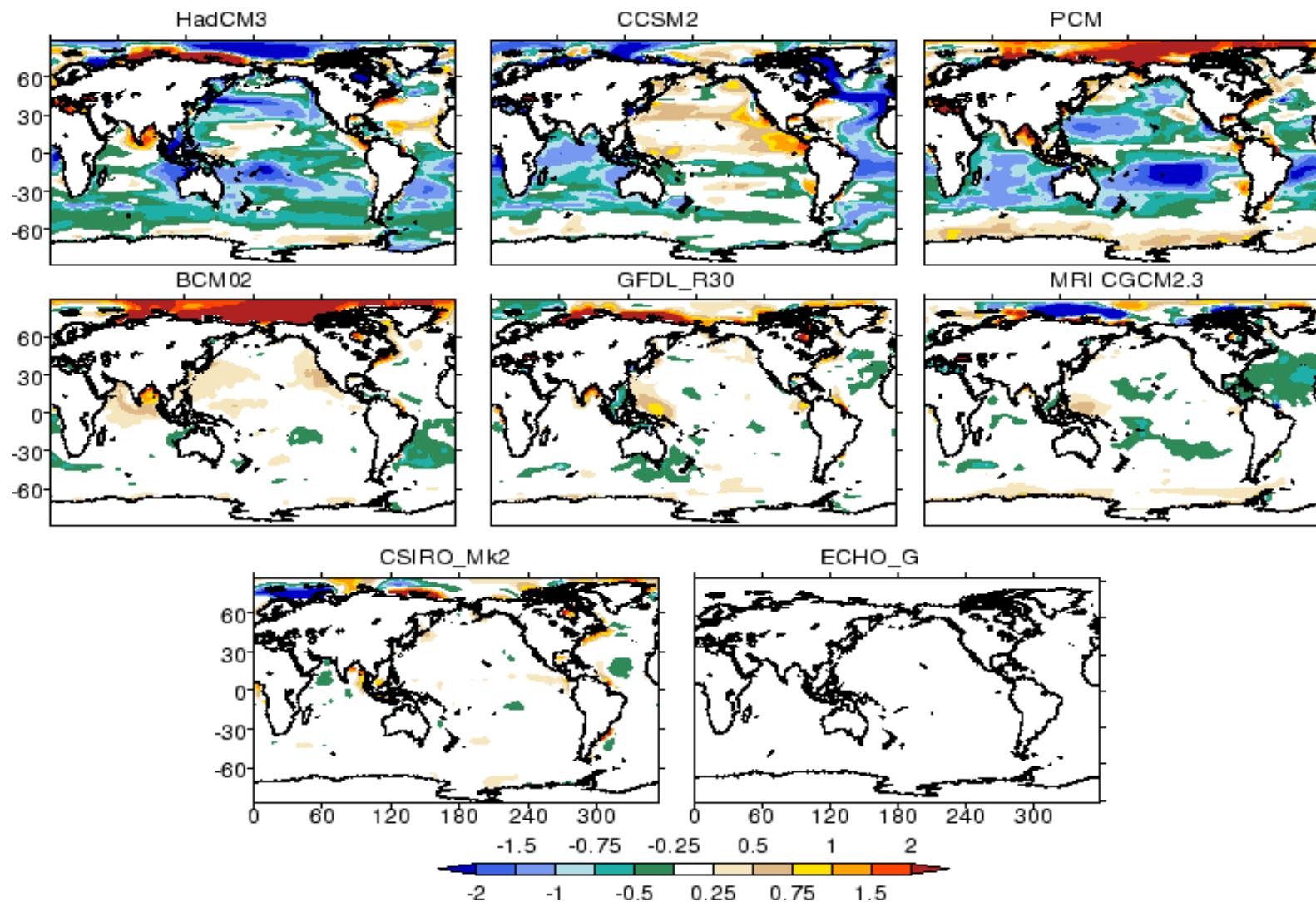


Fig. 5.2. Sea surface salinity time mean difference between the CMIP2+ models and Levitus (ppt).

## 5.2 The Atlantic Ocean

Figure 5.3 depicts the potential temperature of the zonally averaged latitude–depth cross section for the WOA98 climatology and the nine models for which potential temperature is available. The separate scaling of the vertical profile into the upper (0–1000 m) and deep (1000–5000 m) ocean will be used in most of the figures of this section. The permanent (non-seasonal) thermocline is deeper in the subtropical latitudes, with some shoaling in the tropics (see discussion of the Pacific Ocean below). The north–south asymmetry in the thermocline stems from the warm water transport in the Gulf Stream. At higher latitudes, the vertical temperature gradient is much weaker.

Because the mixing of different water masses is generally slow, some characteristics of the thermohaline circulation can be inferred from the temperature and salinity sections. For example, in the higher latitudes of the North Atlantic the weak vertical temperature gradient results from the density-driven formation of North Atlantic Deep Water (NADW). With strong surface cooling during winter and the brine rejection that accompanies sea-ice formation, conditions in some areas of the North Atlantic are ideal for fueling deep convection. In a simplistic sense, the formation of NADW is illustrated by the 2.5–5°C blue shading. The darker blue interval (values below 2.5°C) represents the denser Antarctic bottom water (AABW; see section on the Southern Ocean below).

In Fig. 5.4 the salinity of the zonally averaged latitude–depth cross section is shown for the WOA98 climatology and the models for which salinity is archived at PCMDI. The most striking feature is the shallow water salinity maximum near 30°N. This is a result of the Mediterranean Sea outflow of relatively warm and salty water, the influence of which is also shown in the longitude–depth cross section at 35°N in Fig. 5.5. Note also the salinity minimum in Fig. 5.4 originating at the surface near 45–50°S. Despite the low salinity, this Antarctic Intermediate Water (AAIW) is relatively cold and dense, sinking to intermediate and deep levels and then spreading toward the equator (see Fig. 5.3). All models have problems simulating the AAIW in the Atlantic.

The meridional current at 35°N in the Atlantic Ocean is shown in Fig. 5.6, illustrating the western and eastern portions of the North Atlantic anti-cyclonic subtropical gyre. The northward Gulf Stream is clearly defined and swift, whereas the *return flow in the east* is comparatively broad and shallow. There are some important differences between the models, but in general they do capture similar characteristics. Future comparisons will be made more useful by exploiting transects taken as part of the World Ocean Circulation Experiment (WOCE).

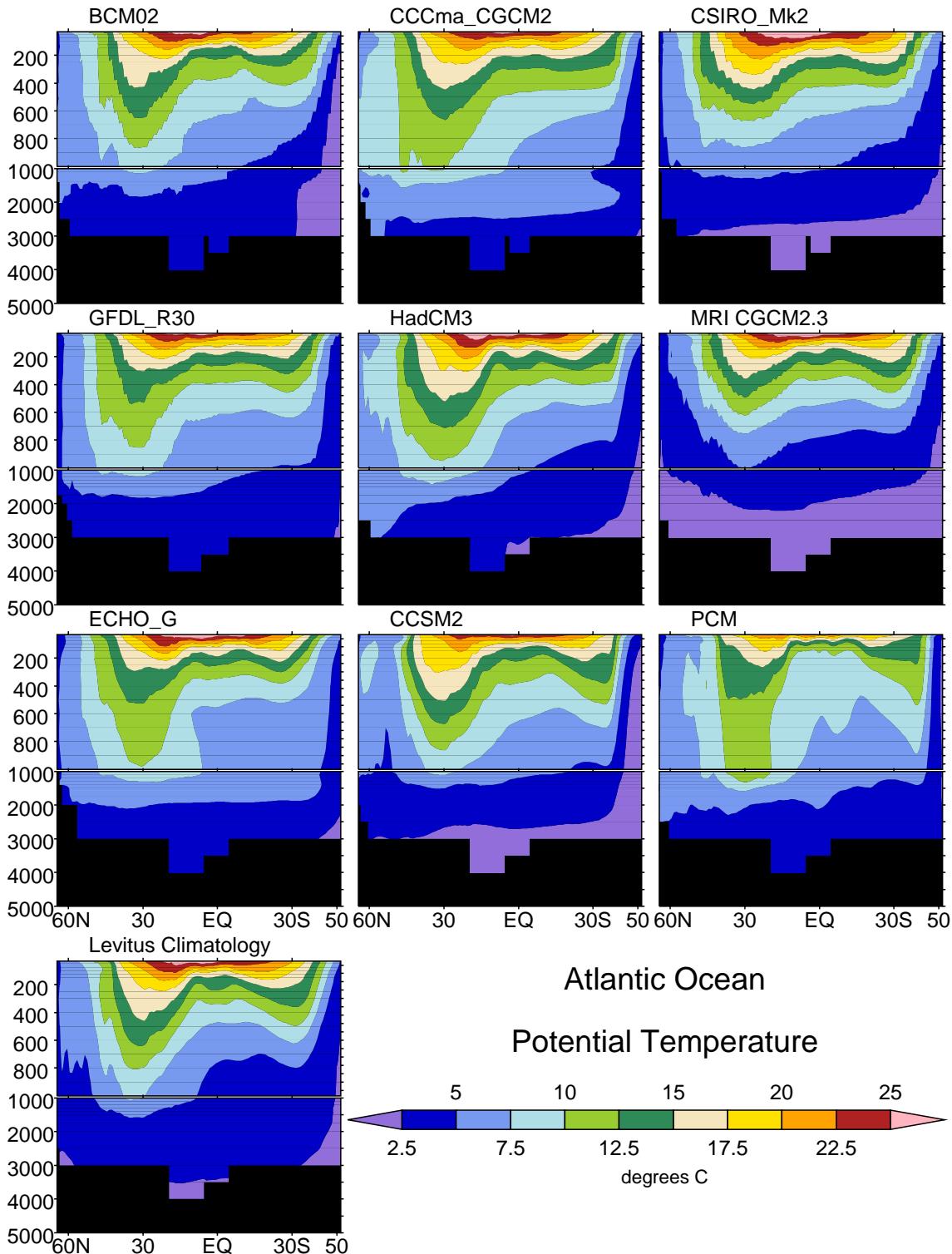


Fig. 5.3. Potential temperature in the Atlantic Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.

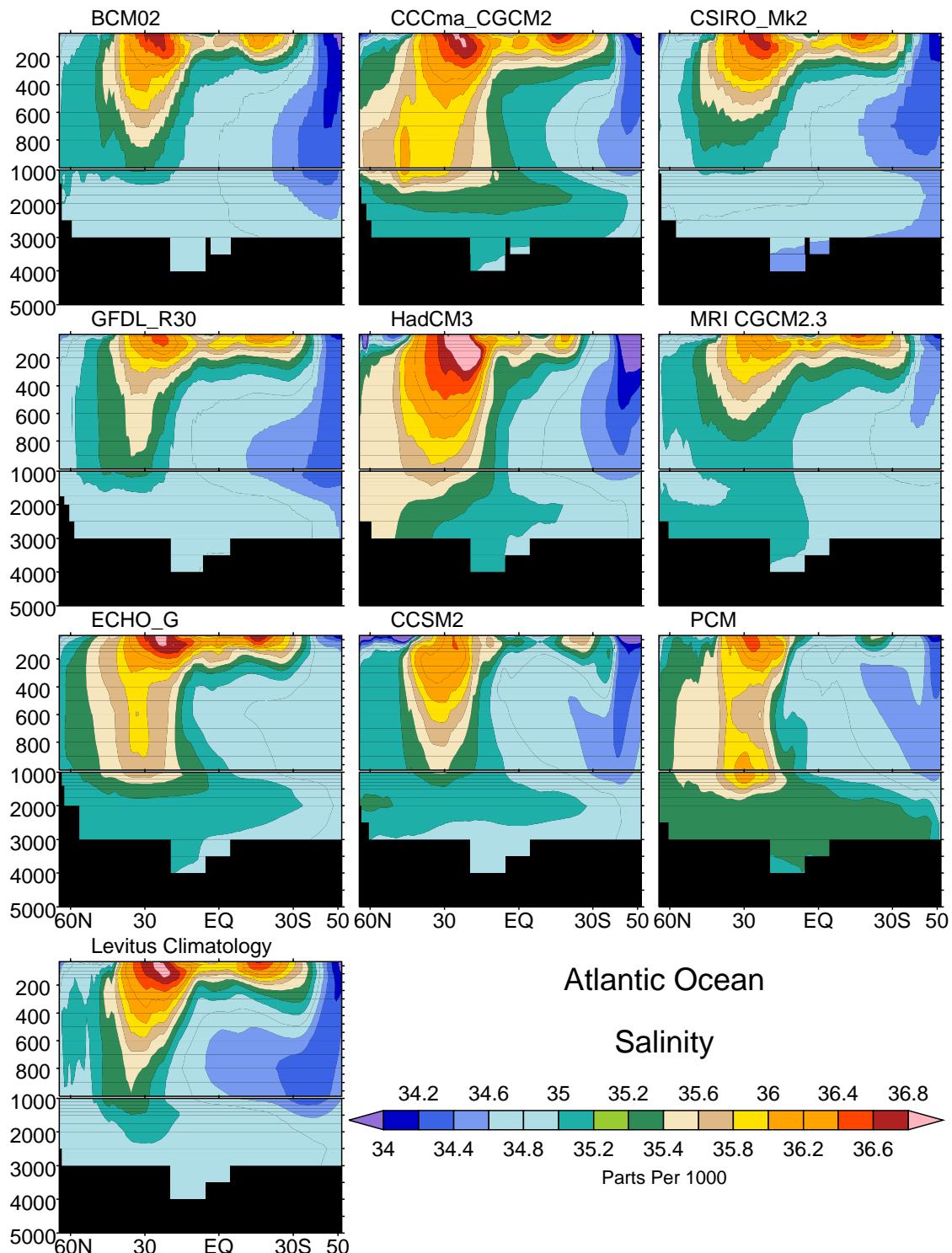


Fig. 5.4. Salinity in the Atlantic Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.

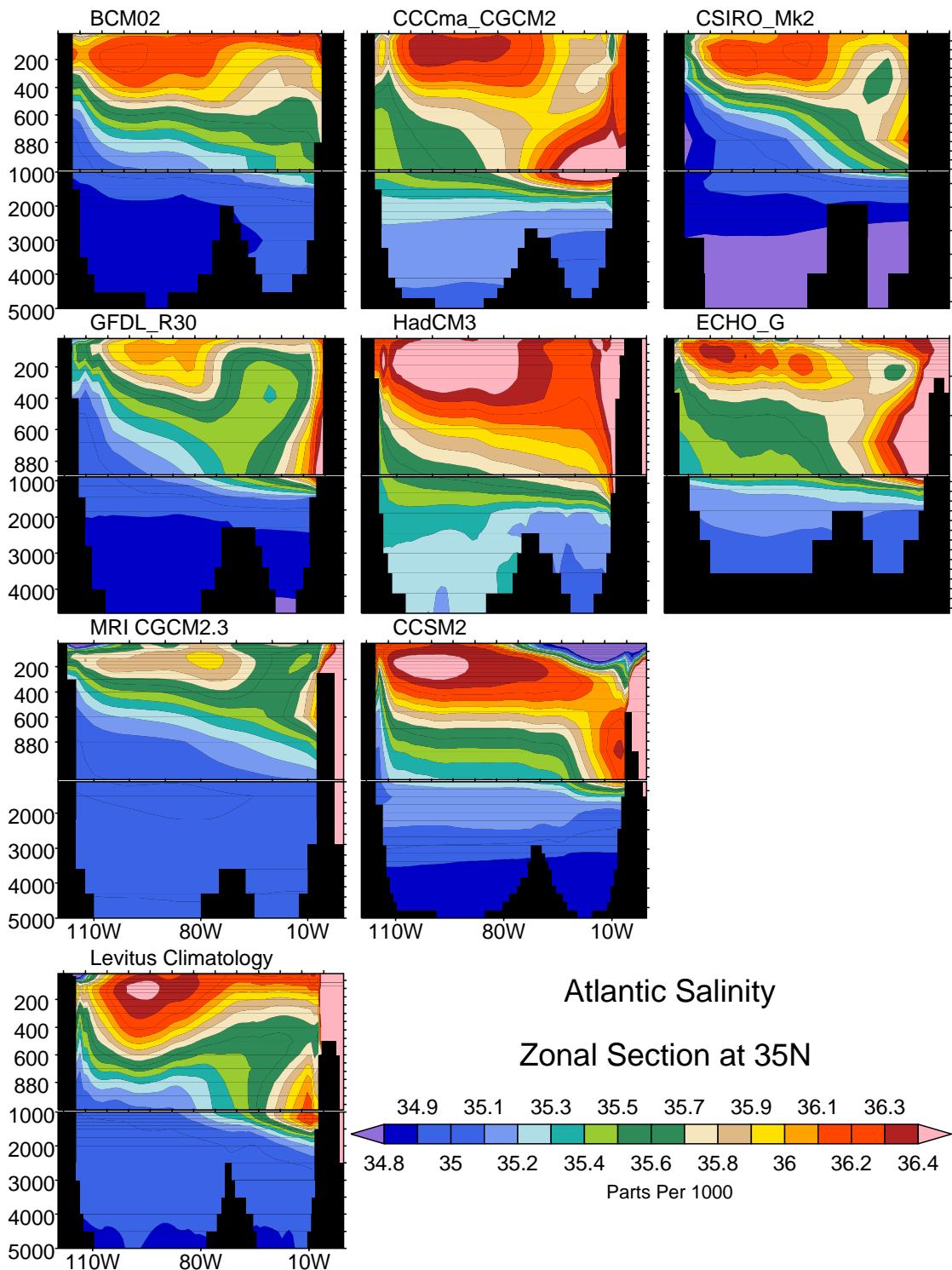
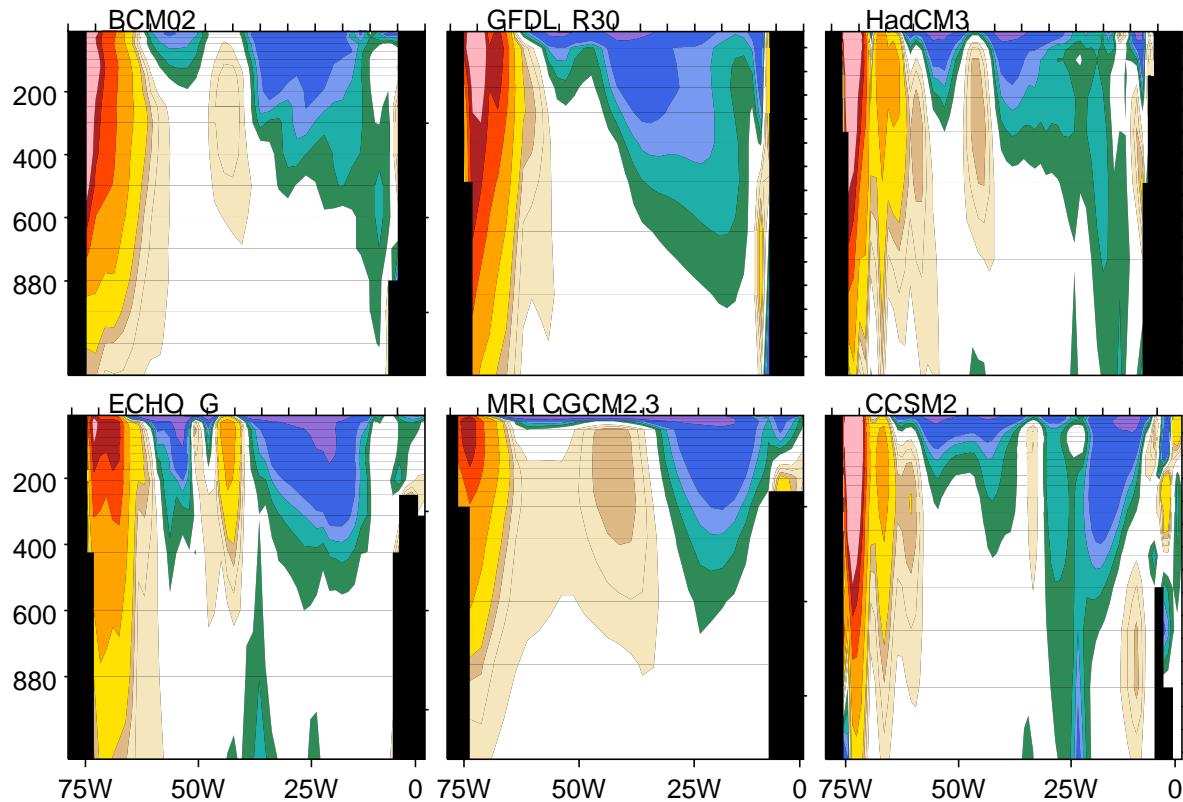


Fig. 5.5. Salinity section at 35°N in the Atlantic Ocean. Note the different vertical scales for 0–1000 m and 1000–5000 m.



Atlantic Northward Current

Section at 35N

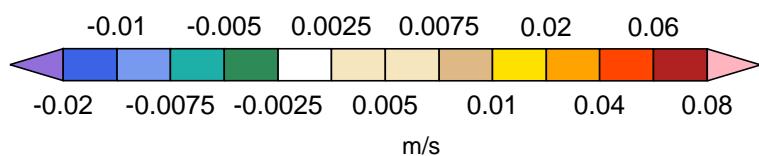


Fig. 5.6. Meridional current at 35°N in the Atlantic Ocean. Note the non-uniform contour increments.

### 5.3 The Pacific Ocean

Thermohaline forcing in the Pacific is much weaker than in the Atlantic. The fact that there is no deep water (and only weak intermediate) sources in the North Pacific is evidenced in Fig. 5.7, where contours of the dense deep waters do not reach the surface in the higher latitudes. However, deep water masses from the Southern and the Atlantic Oceans do extend into the Pacific, where they are thought to be among the oldest. The predominant circulations in the Pacific Ocean circulation are wind-driven gyres. In Fig. 5.8 we see that the Pacific is much less saline than the Atlantic, which is in large part due to differences in the surface freshwater (precipitation–evaporation) flux. Over large areas of the Pacific, precipitation is in excess of surface evaporation, while the contrary is true in the Atlantic. The surface salinity minimum near the equator (or more precisely, the ITCZ) results from excess precipitation. Contrary to the Atlantic Ocean, the Southern Pacific is more saline than the North Pacific, with a distinct surface maximum near 30°S. Pacific Ocean salinity and temperature are relatively homogeneous below about 2000 m.

Temperature along the equator (averaged 2°S–2°N) in the Pacific is shown in Fig. 5.9. On the equator, the effect of rotation on the circulation vanishes and the relatively steady trade winds transport surface waters westward, resulting in the sea surface in the western Pacific being about 50 cm higher than in the east. These surface winds also stimulate growth of the mixed layer where temperature and salinity are defined to be constant with depth. The relatively warm surface waters in the mixed layer are separated from the colder deep ocean water by the thermocline, which in the western Pacific is 100 meters or more in depth below the surface. In Fig. 5.9 the 20°C isotherm is used as a proxy of the thermocline. In the eastern Pacific, the upwelling off the coast of South America replaces the wind-blown warm waters with cold deep waters, resulting in a much shallower thermocline (~10–50 m deep). Moving west along the equator, the easterly winds also promote upwelling of cold water, which decreases to the west as the winds weaken. The distinct east–west gradient seen in the WOA98 data (Fig. 5.9) is captured to varying degrees by the models.

Ocean currents in the tropical Pacific are also principally driven by the trade winds, which cause poleward Ekman transport near the surface on either side of the equator. As a result, the pycnocline shoals near the equator which in turn drives geostrophic flow near the surface, much in the way currents are driven along eastern boundaries. In Fig. 5.10, these westward near-surface currents are evident a few degrees poleward on either side of the equator in the Simple Ocean Data Assimilation (SODA) ocean reanalysis, with the current quite a bit stronger on the southern side. The strong North Equatorial Countercurrent at 5–10°N is driven by cyclonic wind stress (curl) associated with the ITCZ. Atmospheric components of coupled models that incorrectly simulate a double ITCZ tend to produce a (false) equatorial countercurrent in the South Pacific. The Equatorial Undercurrent, strongest in the SODA data at 150–200 m, is driven by the tropical east–west surface height (and hence pressure) difference. The complexities of the tropical oceans require adequate resolution in models. In general, the models with higher resolution more closely compare with the zonal currents in the SODA data.

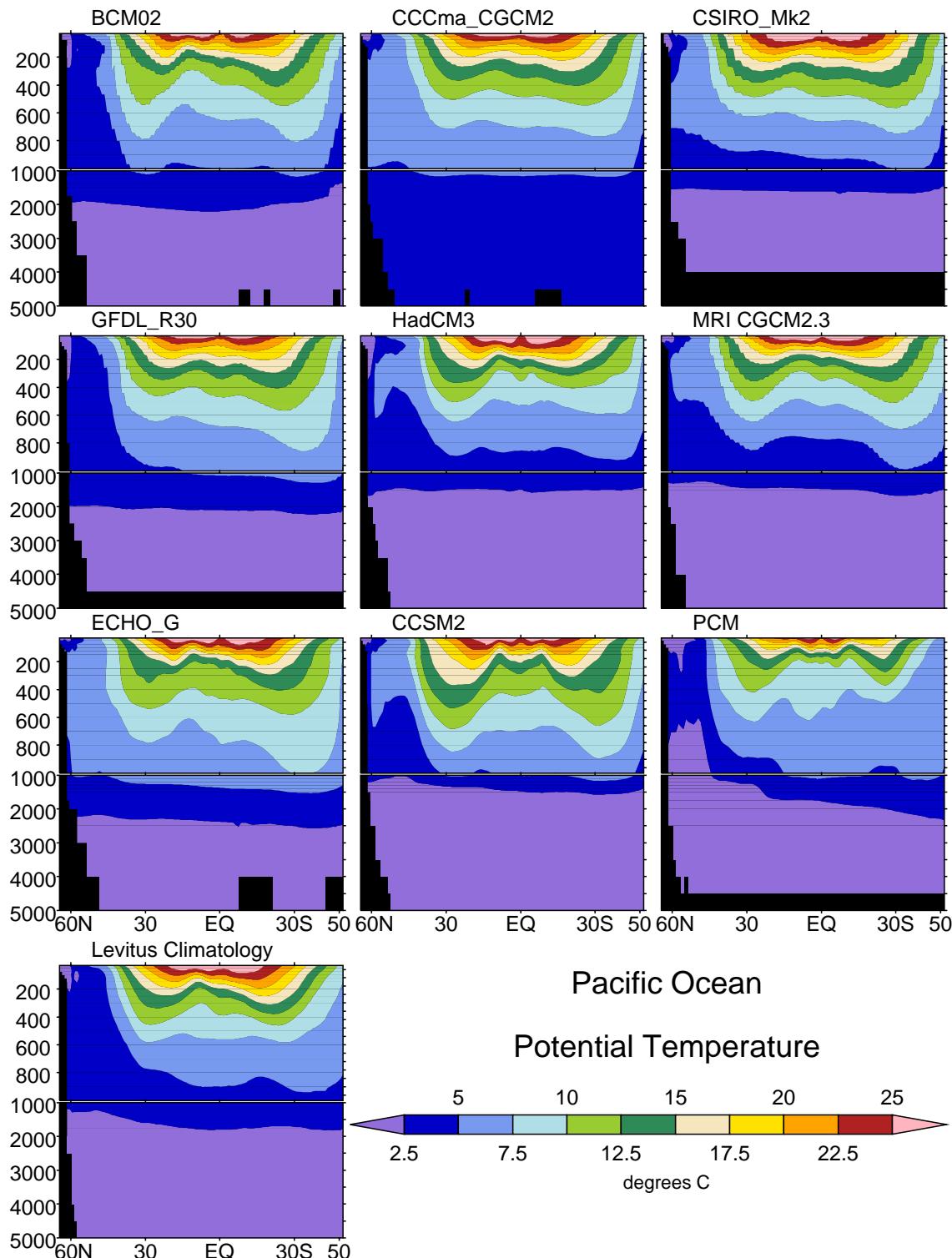


Fig. 5.7. Potential temperature in the Pacific Ocean: zonal-average versus depth.  
Note the different vertical scales for 0–1000 m and 1000–5000 m.

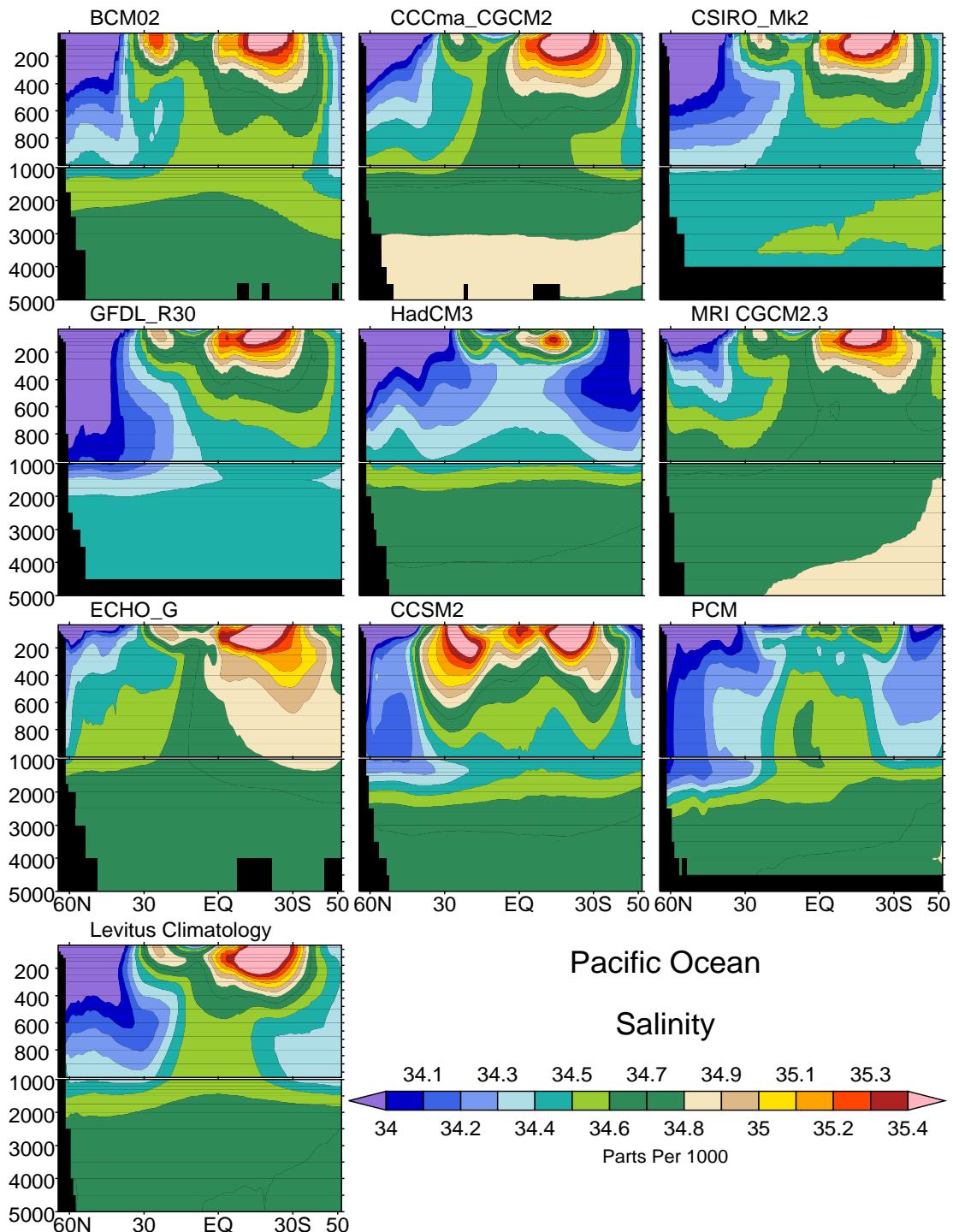


Fig. 5.8. Salinity in the Pacific Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.

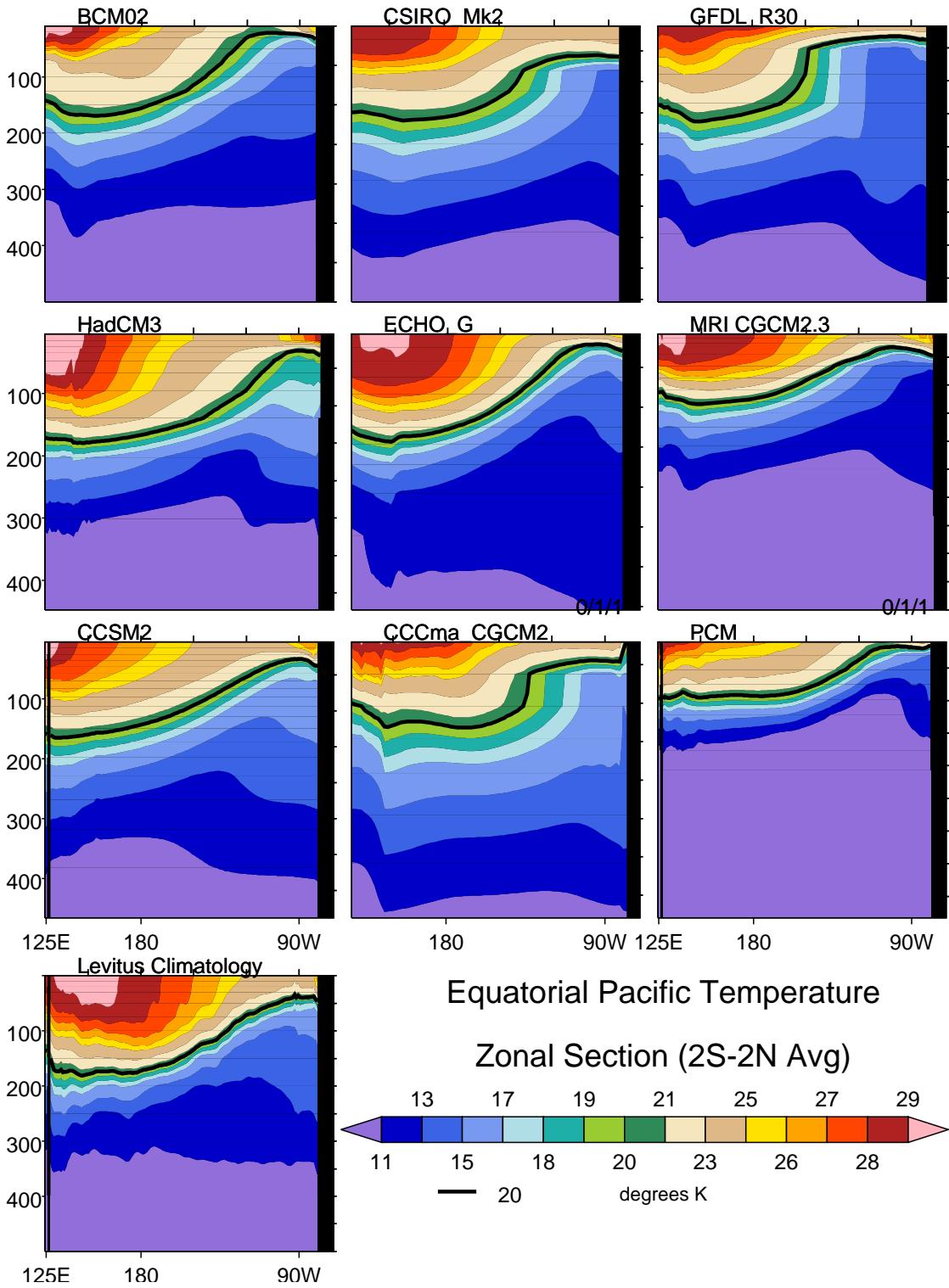


Fig. 5.9. Equatorial Pacific (averaged 2°S–2°N) upper ocean temperature.

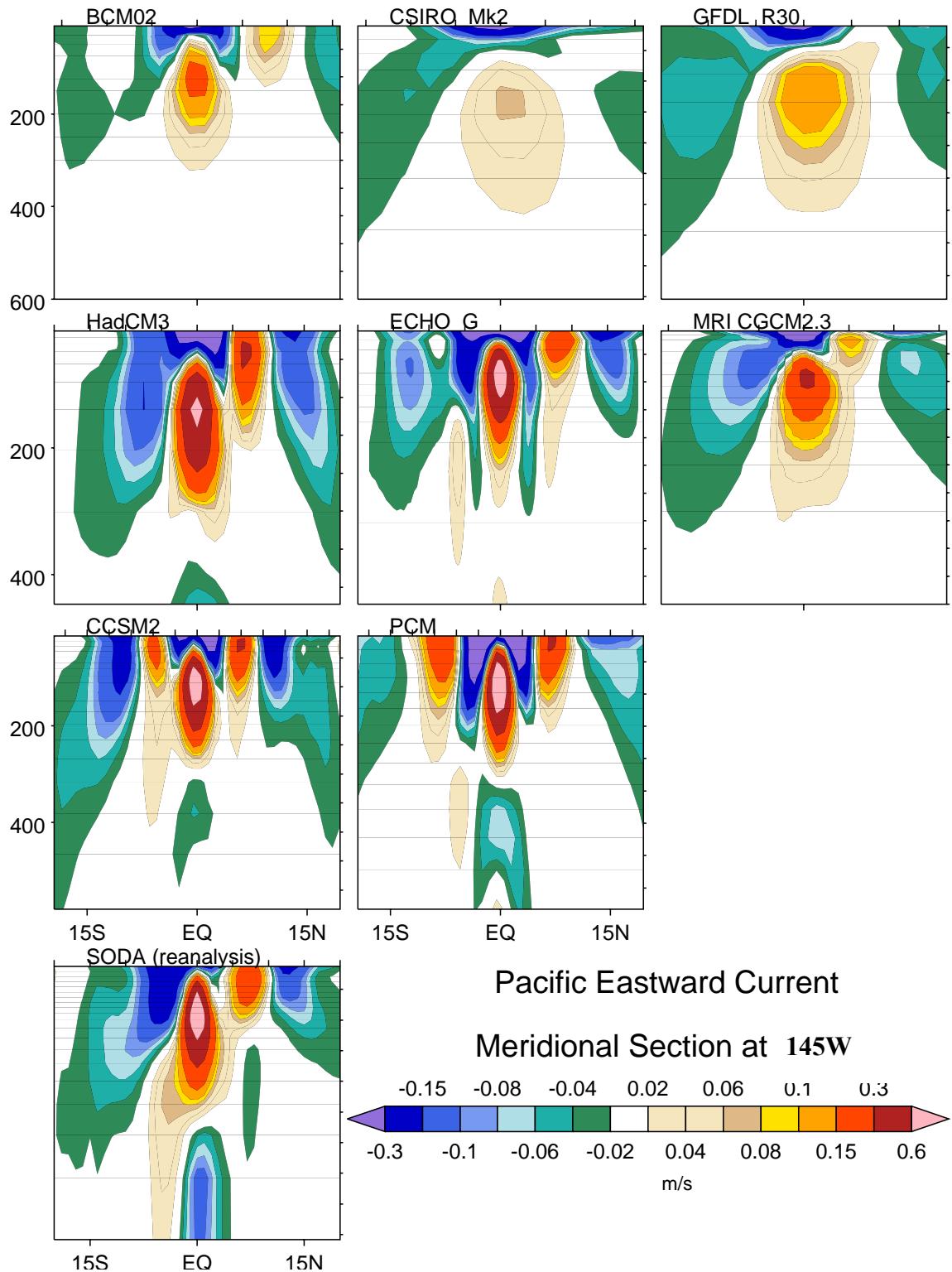


Fig. 5.10. Cross section of zonal currents at 145°W in the Pacific Ocean (m/s).

## 5.4 The Indian Ocean

The surface waters of the northern Indian Ocean are among the world oceans' warmest. They are consequently quite stable as far south as 30°S, as evidenced by the stratification in the upper 500 m, as seen in Fig. 5.11. Thus, deep water formation in the Indian Ocean is negligible. The cold deep water in the Indian Ocean appears to have its origin from the Southern Ocean near 40-50°S.

The salinity of the Indian Ocean is shown in Fig. 5.12, with a salinity maximum near 20°N originating in the Arabian Sea. It extends to a depth of nearly 1 km, in part owing to deep-reaching saline water originating in the Red Sea. West of Australia and across the Indian Ocean, a second salinity maximum results from there being more surface evaporation than precipitation. While all of the models capture the major features of the Indian Ocean salinity, there is more discrepancy than in the Atlantic or Pacific Oceans.

Figure 5.13 shows the meridional current at 32°S in the Indian Ocean, chosen in part because of transects taken during the World Ocean Circulation Experiment (WOCE, not shown here), and plans for repeat cruises. The Agulhas western boundary current is not as distinct as that of the Gulf Stream (see section above on the Atlantic Ocean), which in part may be due to the complex bathymetry of the Indian Ocean and the presence of Madagascar a few degrees off the east coast of Africa. In any case, as in the subtropical gyre of the North Atlantic, the models generally capture the basic features of this wind-driven circulation: intense western boundary currents contrast with the relatively broad and shallow southward return flow in the eastern portion of the gyre.

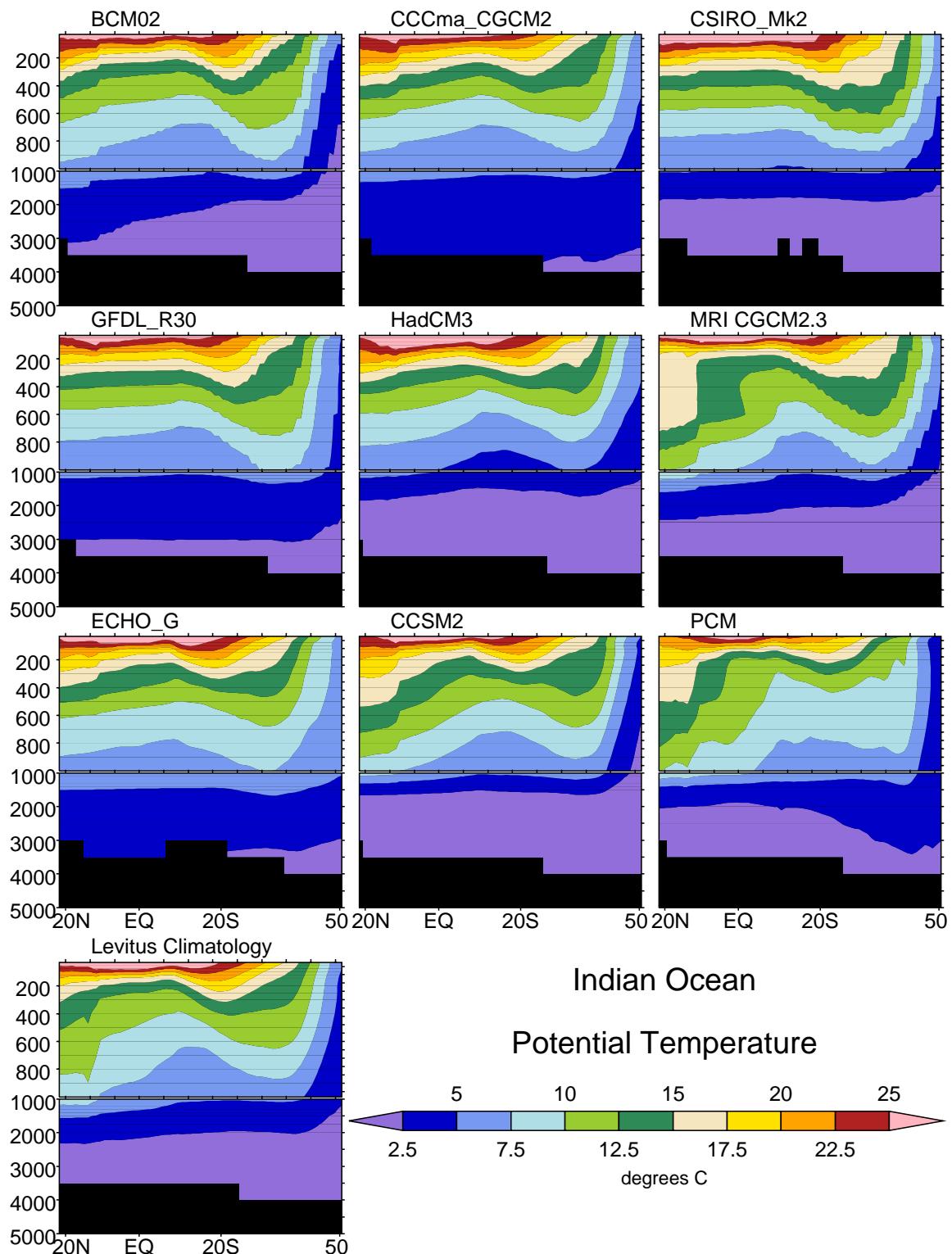


Fig. 5.11. Potential temperature in the Indian Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.

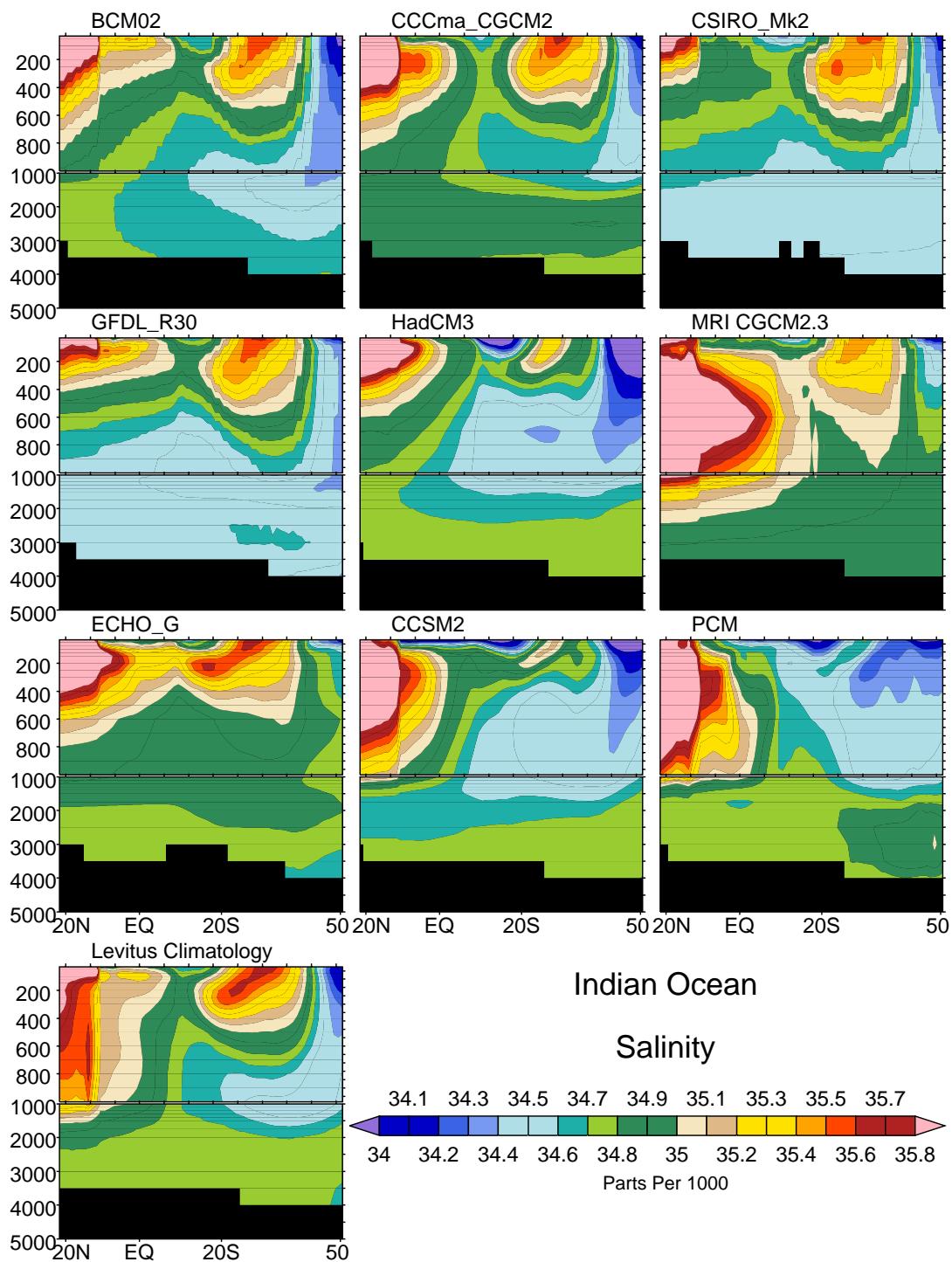
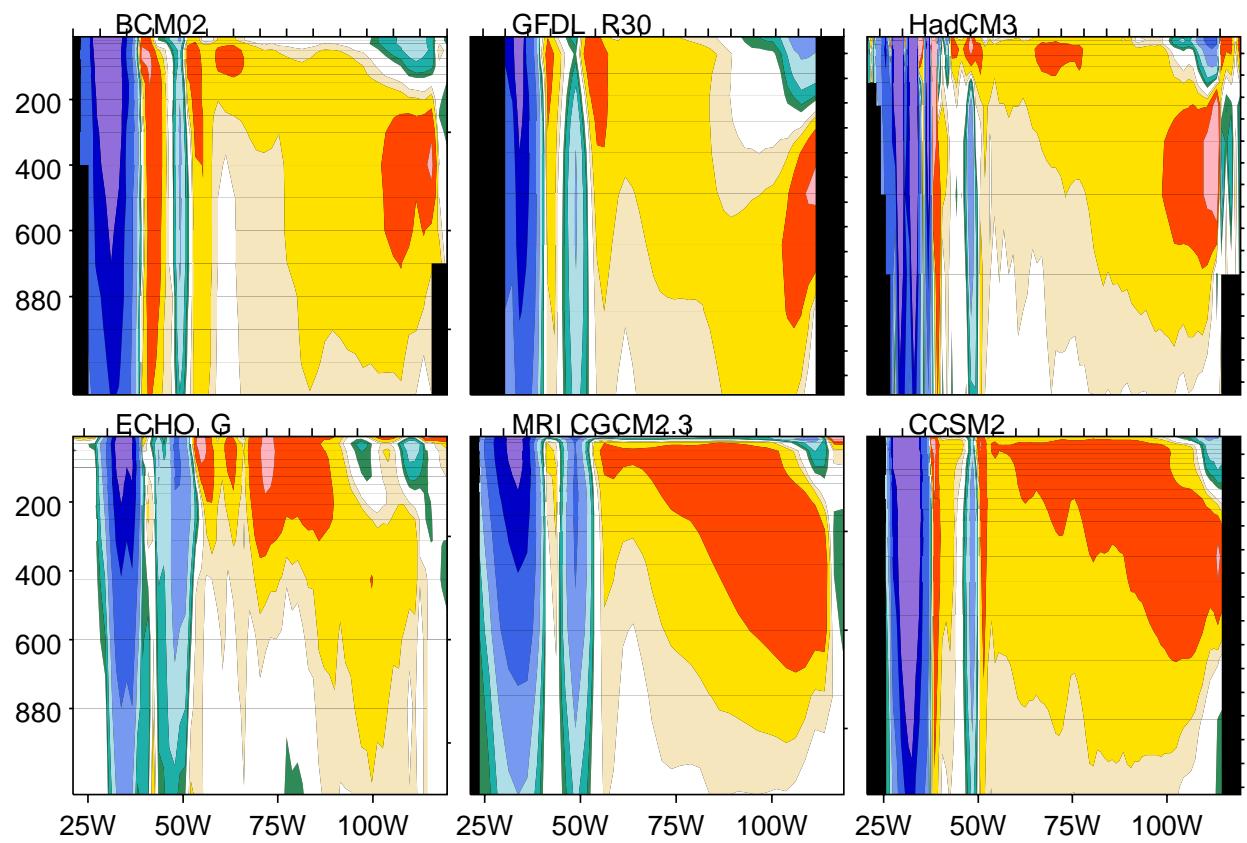


Fig. 5.12. Salinity in the Indian Ocean: zonal-average versus depth. Note the different vertical scales for 0–1000 m and 1000–5000 m.



### Indian Ocean Northward Current

#### Zonal Section at 32S

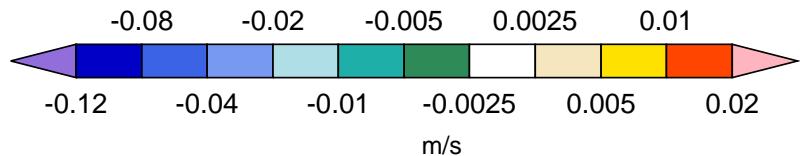


Fig. 5.13. Northward current in the Indian Ocean at 35°S. Note the non-uniform contour increments.

## 5.5 The Arctic Ocean

The potential temperature in the Atlantic sector of the Arctic Ocean is shown in Fig. 5.14. There is general agreement with relatively cold surface waters north of 70°N, and the indication of heat transport from the south in the subsurface Atlantic Layer. Apart from this however, the differences between the models and WOA98 are much more evident than seen in the other oceans we have examined. (The differences relative to other polar climatologies may be less, but these have not been determined.) Note, however, the temperature scale only ranges from -2° to 6°C and so the models differ by no more than a few degrees. While the heat balance is very sensitive to the insulating effects of sea-ice, these plots suggest that the temperature (and salinity) characteristics of the Arctic Ocean may be very sensitive to inter-basin exchanges (see below).

The Arctic salinity is similar to that of temperature (Fig. 5.15): Relatively fresh surface waters at the highest latitudes, and significant differences between the models and WOA98 at depth. As with temperature, note that some of the most visible differences in Arctic salinity are in fact quite small (less than 1 part per thousand). The WOA98 temperature data is suggestive of a positive near-surface northward flux of salinity into the Arctic Basin, which in many models appears exaggerated.

Via the Nordic Seas, the Fram Strait (between Greenland and Spitsbergen Island) is the only deep passage between the Arctic Ocean and the North Atlantic Ocean. In Fig. 5.16, the meridional current is shown in the Fram Strait and Barrents Sea (20°E–20°W) at 80°N. The southward-flowing East Greenland Current, generally in the western section of the Fram Strait, is the most significant outflow from the Arctic Ocean. Immediately to the east the northward West Spitsbergen Current flows around the Yermak Plateau at about 20°E. These Arctic–Atlantic exchanges, and those in the Barents Sea (20°E–55°E), differ significantly between the models.

Further tests are necessary to determine why the Arctic model vs. Levitus differences appear so much larger than the other major ocean basins. Possibilities to consider: inter-basin exchanges, the insulating effects of sea-ice, faster model spin-up relative to other basins, and the use of rotated poles in some models.

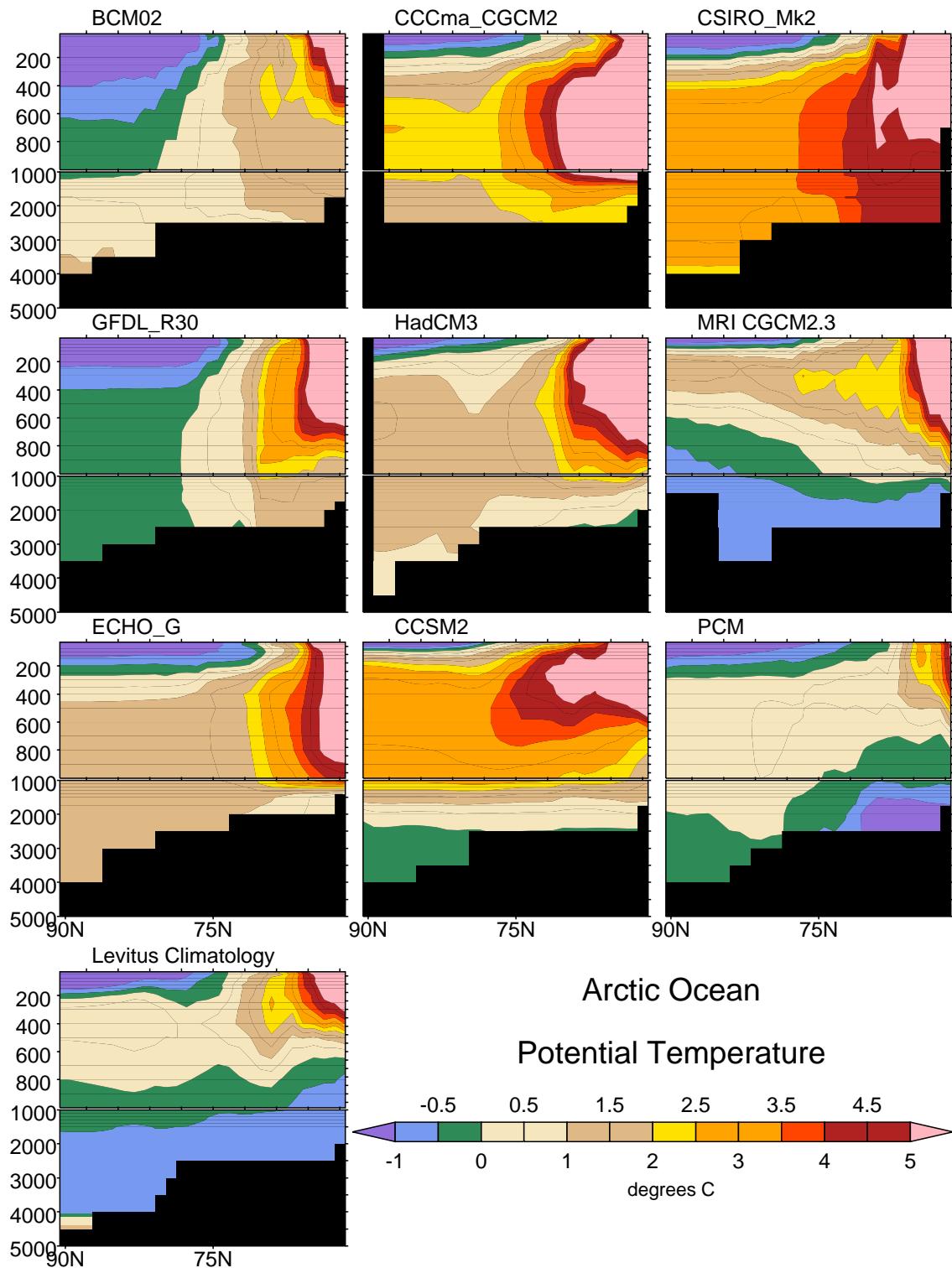


Fig. 5.14. Potential temperature in the Arctic Ocean: zonal-average versus depth

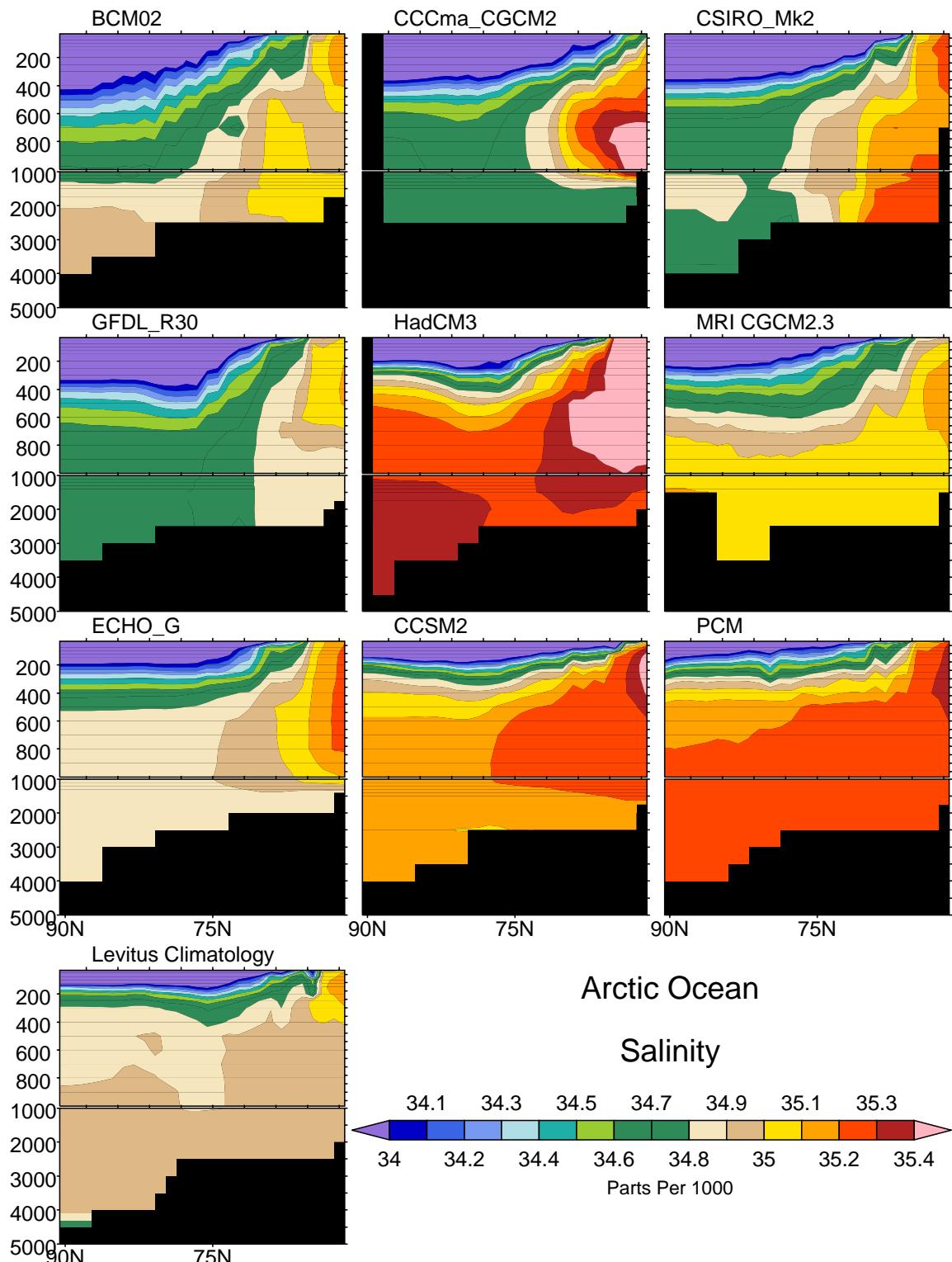
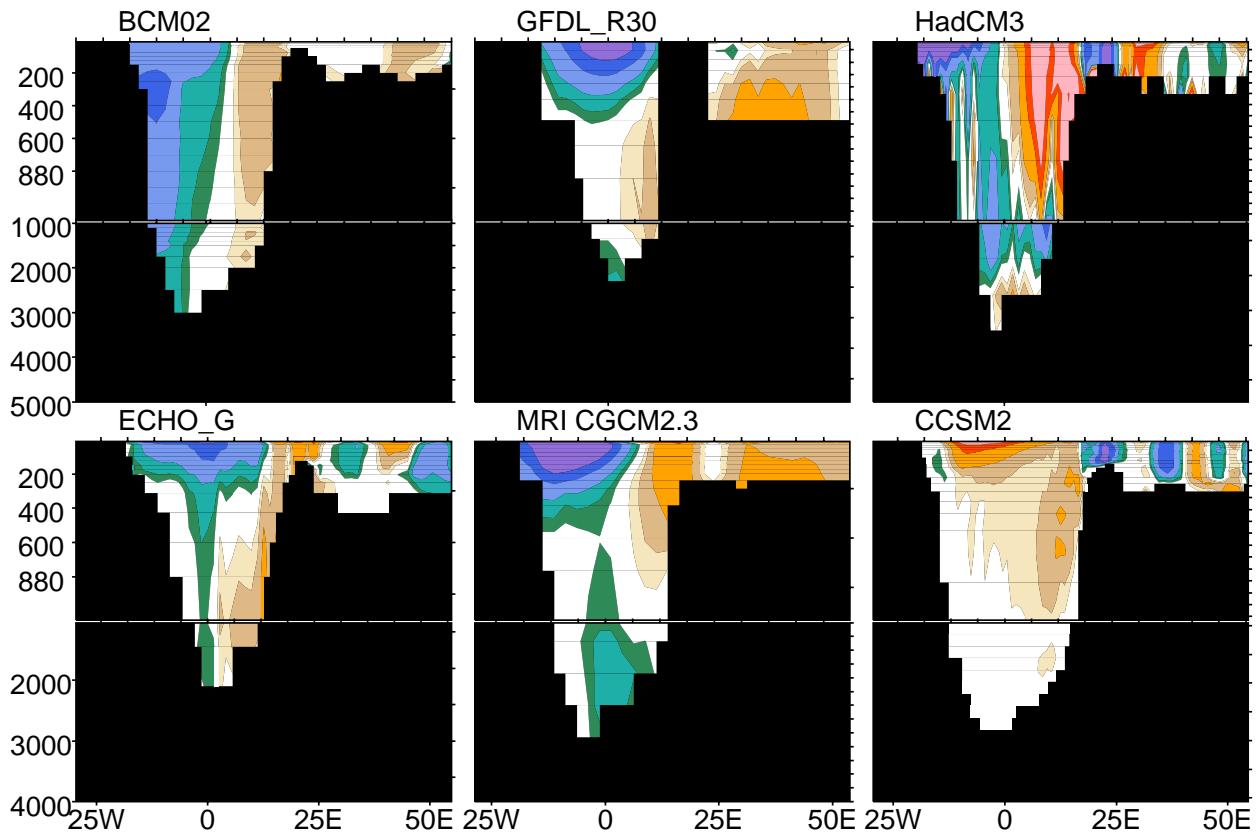


Fig. 5.15. Salinity in the Arctic Ocean: zonal-average versus depth.



Northward Current

Section at 80N

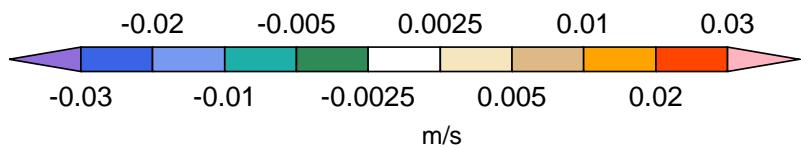


Fig. 5.16. Northward current through the Fram Strait and the Barents Sea at 80°N.

## 5.6 The Southern Ocean

The zonal averaged temperature versus depth for the Southern Ocean is shown in Fig. 5.17. A unique feature of the Southern Ocean is the wind-driven Antarctic Circumpolar Current (ACC), which facilitates the eastward redistribution of heat (and other properties) between the Atlantic, Pacific and Indian Oceans. The lack of temperature stratification between 50°S and 60°S in both WOA98 and the models is suggestive of Antarctic Intermediate Water (AAIW) formation. South of 70°S, most of the ocean area is limited to the Weddell and Ross Seas, where most Antarctic Bottom Water (AABW) forms locally. This is implied by the -1.5°C to -2°C contour in WOA98 data, and to varying degrees is captured by the models.

The surface waters of the Southern Ocean are relatively fresh compared with most ocean areas, the North Pacific being an important exception. This can be seen by comparing both the WOA98 data and the models in Fig. 5.18 with the salinity figures of the Atlantic, Pacific and Indian Oceans (Figs. 5.4, 5.8, 5.12). Note the brine rejection due to sea-ice formation is not evident in the shallows of the WOA98 data, perhaps because this localized phenomenon is smoothed with zonal averaging. In the deepest waters, however, the relatively saline and cold water is suggestive of AABW. As evidenced in Fig. 5.18, no model has a latitude–depth pattern of Southern Ocean salinity that agrees with the Levitus data.

The zonal current in the Southern Ocean is shown in Fig. 5.19 at 60°W (about 5 degrees west of the narrows of Drakes Passage). The Antarctic Circumpolar Current (ACC) is clearly evident in all the models, all with a maximum surface current between 50°S–60°S. The maximum intensity of the ACC does, however, vary between the models.

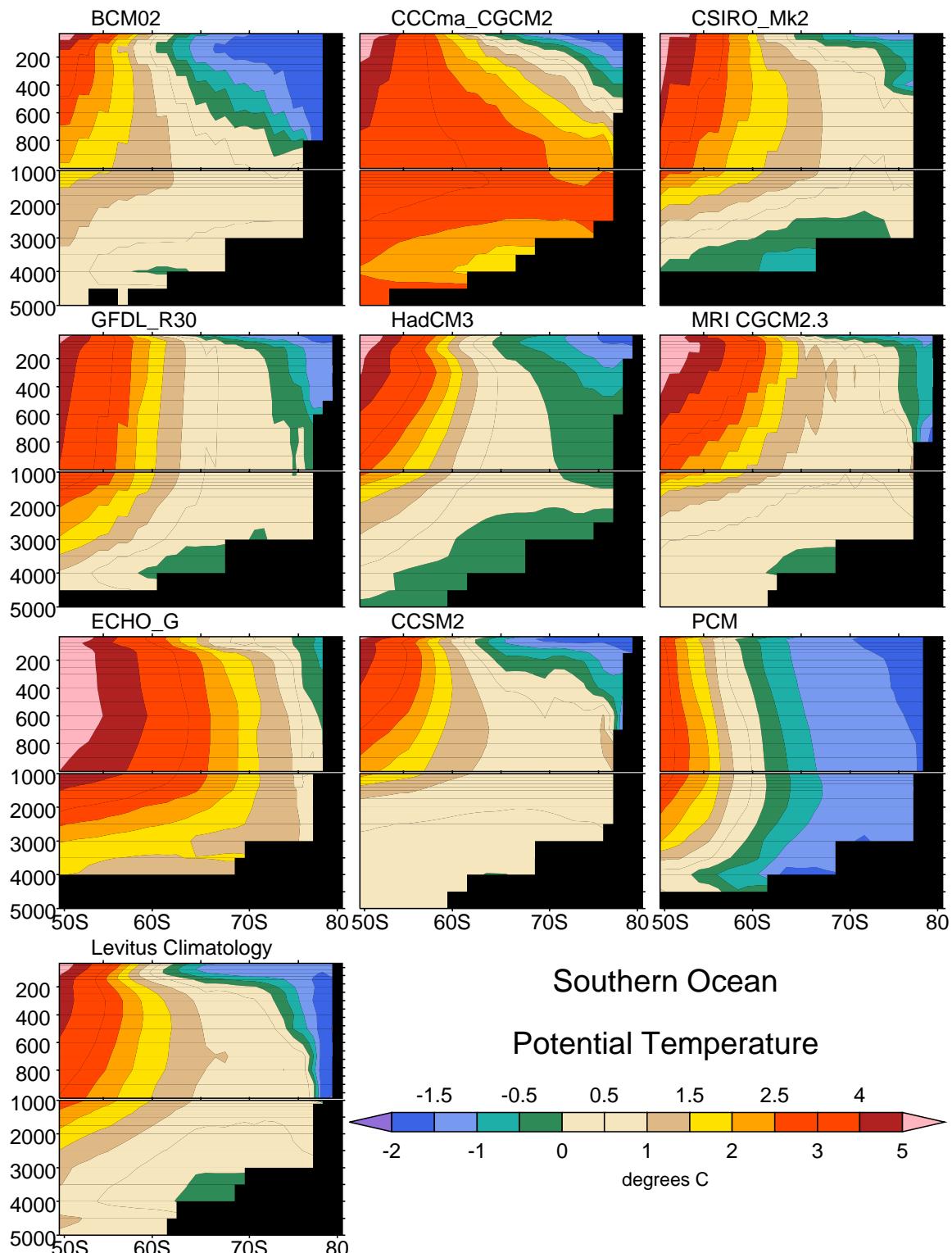


Fig. 5.17. Potential temperature in the Southern Ocean: zonal-average versus depth.

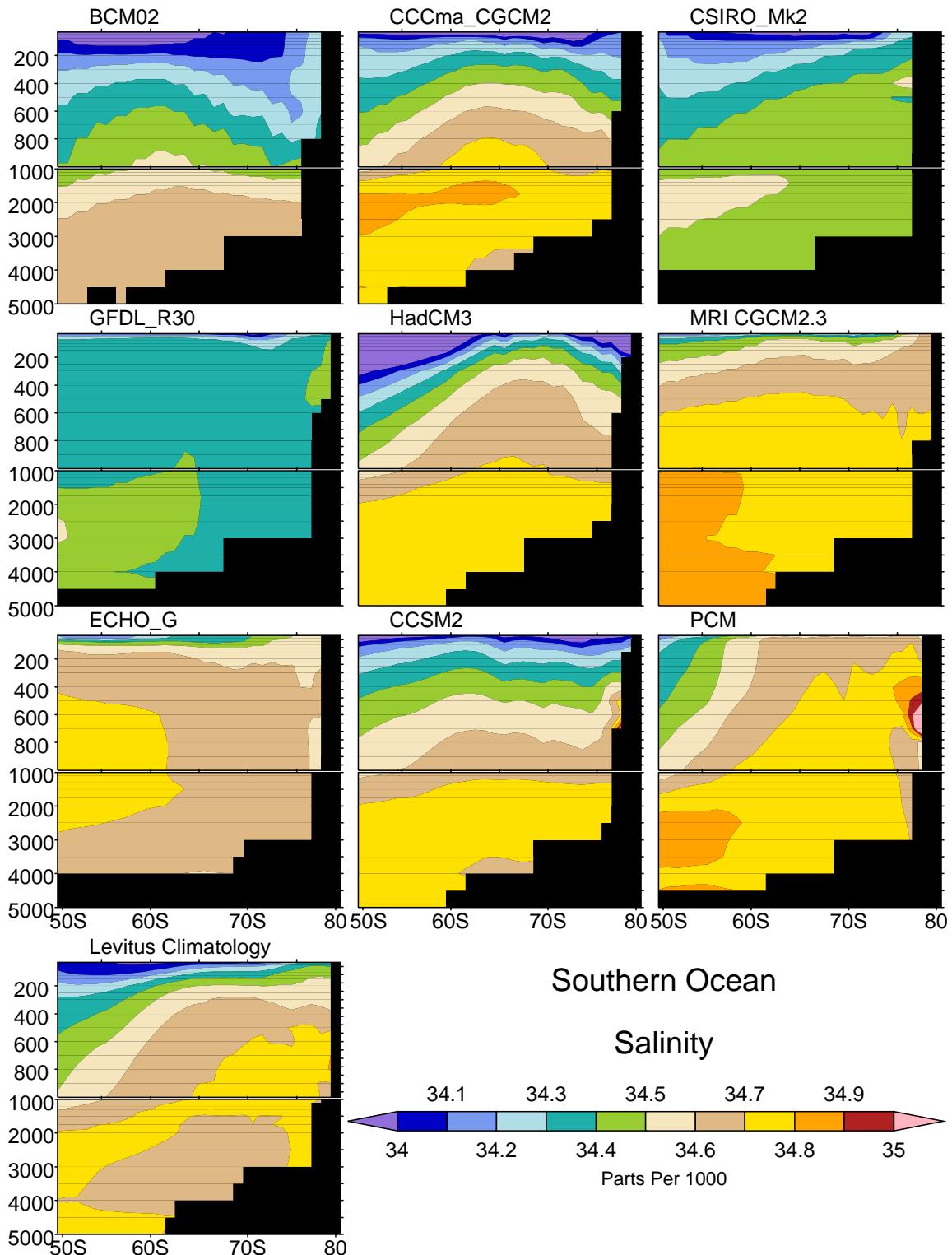
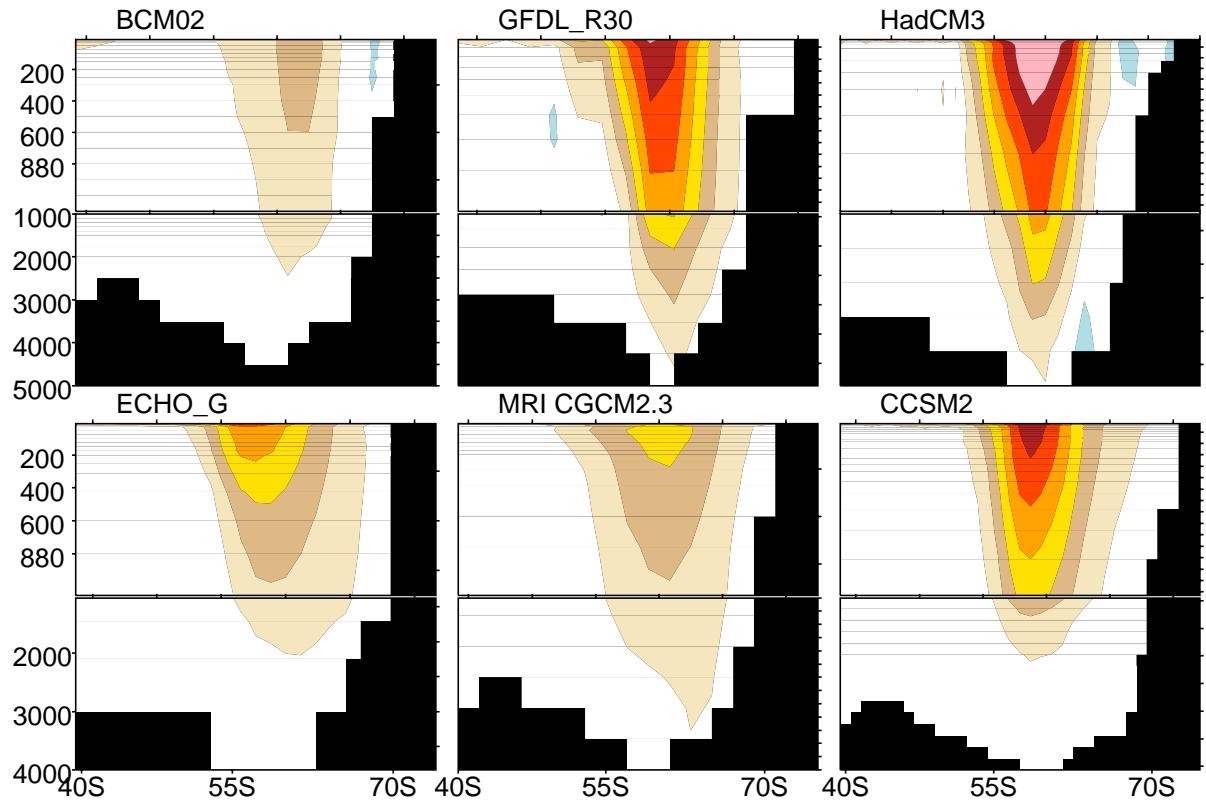


Fig. 5.18. Salinity in the Southern Ocean: zonal-average versus depth



Southern Ocean Zonal Current

Section at 60W

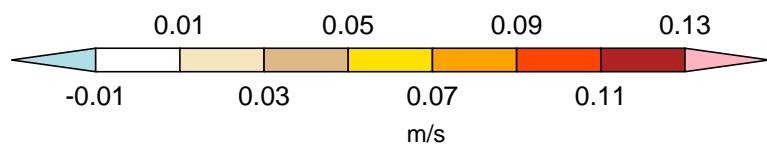


Fig. 5.19. Southern Ocean zonal current at 60°W.

## 5.7 Sea Ice

The sea-ice concentration for three models and an observationally based data set (1979–1999 climatology used in AMIP; R. Grumbine, NCEP) are shown in Fig. 5.20 with February and August representing the extreme seasons in both hemispheres. Note that the sea-ice concentration is available for only a few models in the CMIP2+ database.

Much of the Arctic Ocean is at least 1000 m deep, and generally covered with sea-ice throughout the year. In the continental shelf regions of the Arctic (about one third of the basin), the waters are much shallower and in Boreal summer some are ice-free. A large part of the seasonal variation in total Arctic ice-cover is in the Bering Sea (along continents and the Aleutian Islands), and in the Sea of Okhotsk. The distribution of sea-ice near Antarctica is more symmetric and compact than that of the Arctic, and has a much larger seasonal variation in total area covered. The annual cycle of total sea-ice area is shown in Fig. 5.21. The amplitude of the seasonal cycle is much larger in the Antarctic, with only about 3–4 million square km<sup>2</sup> left at the end of summer, compared with 7 million km<sup>2</sup> in the Arctic.

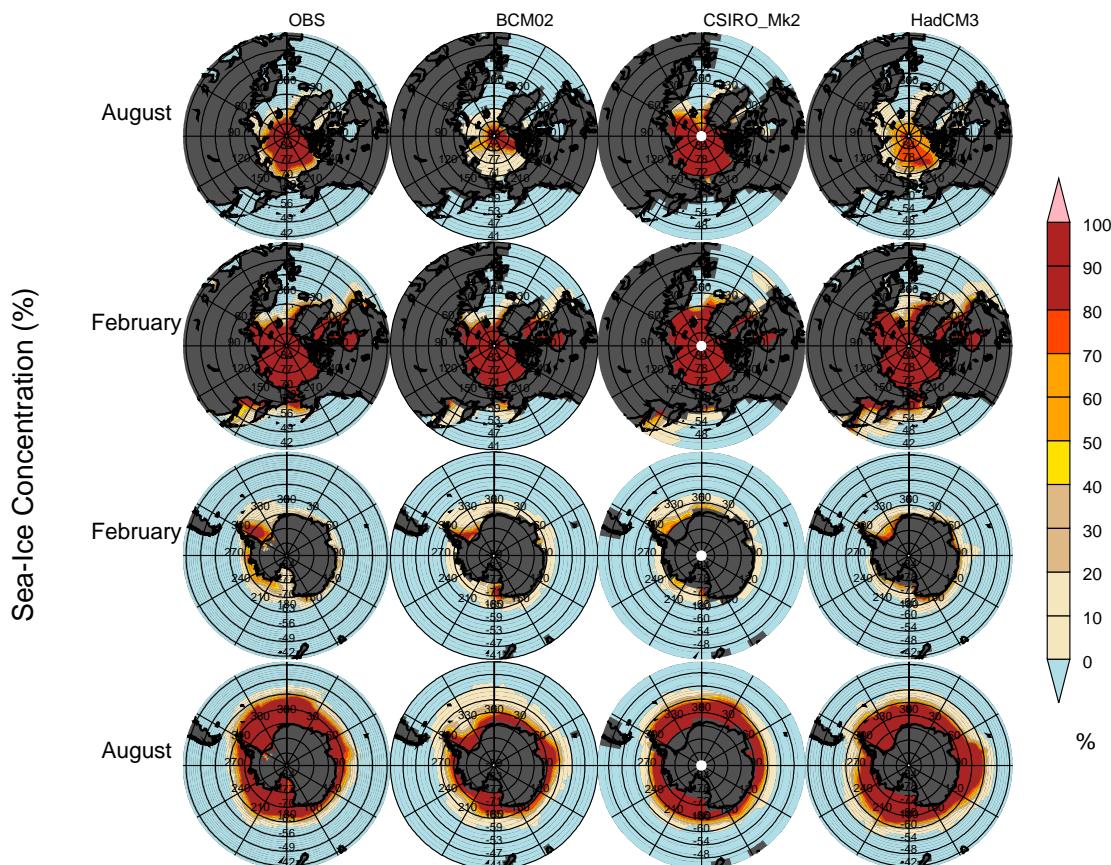


Fig. 5.20. Sea-ice concentration (percent) for the CMIP 20-year climatologies and observations.

## Annual Cycle of Sea-Ice Cover

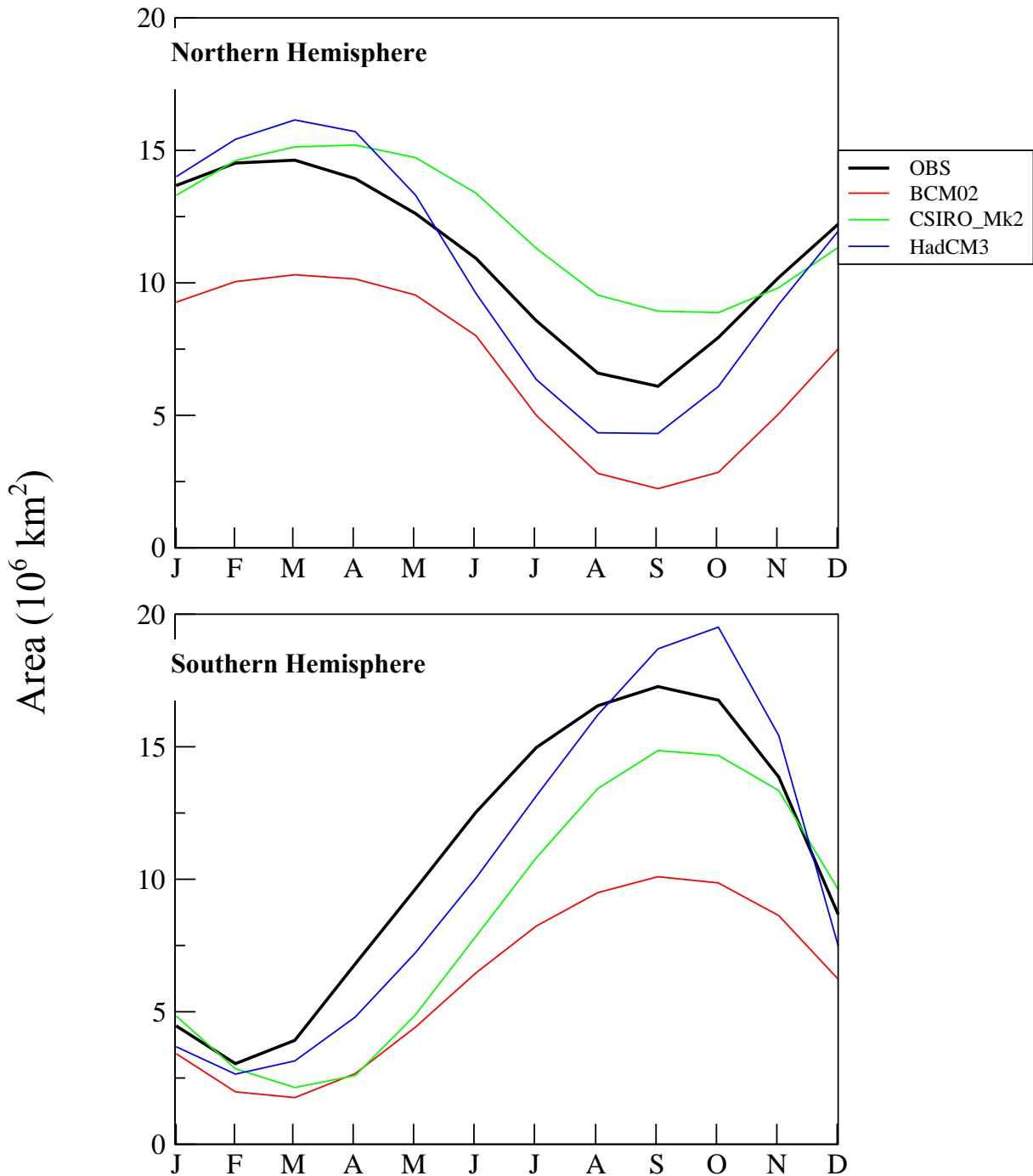


Fig. 5.21. Annual cycle in total sea-ice area ( $10^6 \text{ km}^2$ ) for the 20-year climatologies of the CMIP simulations and observationally based estimates.

# 6 Simulated Variability

*K. Sperber, K. AchutaRao, P. Gleckler, C. Doutriaux*

## 6.1 El Niño/Southern Oscillation

### 6.1.1 Introduction

The El Niño Southern Oscillation (ENSO), a phenomenon driven by coupled processes, is one of the most important sources of interannual variability in the natural climate system. El Niño events, characterized by increased sea surface temperatures in the central and eastern tropical Pacific, have a periodicity of between two and seven years. The early 1990s, characterized by prolonged El Niño conditions, and the strong event of 1997-98, have been unusual compared with past ENSO variability. The extended duration and apparent increase in the frequency of these events have further heightened interest in studying ENSO. The apparent climate shift in the tropical Pacific around 1976 was preceded by a period of reduced ENSO variability between 1920–1960, indicating an interdecadal modulation of ENSO. The effects of ENSO events are felt outside of the tropical Pacific Ocean, including a modulation of temperatures globally. Transient CO<sub>2</sub> CGCM experiments have raised some concerns over the possible changes in ENSO under enhanced greenhouse conditions. Therefore, the role of global warming in altering the strength and periodicity of ENSO and the extent of ENSO's effect on global temperatures are questions of foremost importance. These issues can be addressed with the help of CGCMs, as they provide controlled experiments to assess separately the natural variability of the climate system and the variability caused by anthropogenic influences. The ability of these global models to simulate natural variability like ENSO is therefore of vital importance. Here we address the ENSO characteristics in the control simulations of 10 CGCMs, making use of the complete time series submitted to PCMDI.

### 6.1.2 The Models

The ENSO simulation characteristics for CSIRO\_Mk2, HadCM2, HadCM3, ECHAM4\_OPYC3, and PCM have been discussed in AchutaRao and Sperber (2002) (referred to henceforth as AS2002). Although the current analysis includes additional data from control simulations (primarily because the data available in CMIP2+ are longer time-series for a few models), the essential characteristics are the same. Here, as in AS2002, surface air temperature (TAS) is used interchangeably with sea-surface temperature (SST) since they are highly correlated in observations. Importantly, relative to AS2002, new versions of several models are available, which enables an appraisal of changes in ENSO characteristics due to model development. The new model versions are the CCCma(CGCM2, CCSM2.0, GFDL\_R30\_c, and MRI\_CGCM2.3 models.

### 6.1.3 Seasonal Cycle Phase Locking and Power Spectra

While the amplitudes of El Niño episodes vary, their phasing can be remarkably similar in that the peak SST anomalies occur in the boreal winter. The phase-locking with the seasonal cycle can be seen by plotting the standard deviation of NIÑO3 ( $5^{\circ}\text{S}$ – $5^{\circ}\text{N}$ ,  $150^{\circ}\text{W}$ – $90^{\circ}\text{W}$ ) temperature anomalies for each calendar month. As seen in Fig. 6.1, the observed peak interannual activity occurs during the boreal winter. While most models show the observed peak in the winter months, the CCCma\_CGCM2, GFDL\_R30\_c, and

CSIRO\_Mk2 models are conspicuous by their lower-than-observed variability and a tendency for higher variability in the boreal summer months. The ECHO-G, PCM, and CCSM2.0 models also show high variability in the summer. The high variability of the MRI\_CGCM2.3 and HadCM3 in all the months indicates an overactive ENSO.

An important feature of ENSO is the frequency with which it occurs. The maximum entropy power spectra of the observed and model simulated NIÑO3 temperature anomalies are plotted in Fig. 6.2. The observed NIÑO3 temperature anomalies show a broad peak at time scales of 40–50 months. The ECHO-G and ECHAM4\_OPYC3 models are both dominated by peak power at about 2 years, the low end of the observed periodicity. ECHO-G has a higher amplitude than ECHAM4\_OPYC3. These two models have the same atmosphere component (ECHAM4) and bear out the conclusions of Guilyardi et al. (2003) that the atmospheric component sets the frequency of ENSO in the coupled model, whereas the oceanic component sets the amplitude of the variability. The CCSM2.0 power spectrum shows a peak at a frequency that is higher than observed, and the MRI\_CGCM2.3 model indicates an amplitude that is greater than observed. The CSIRO\_Mk2, CCCma\_CGCM2, and GFDL\_R30\_c models are the only models that do not show peaks in the power spectrum within the observed range of frequencies.

#### **6.1.4 SOI and NIÑO3 Teleconnections**

J. Bjerknes recognized that the Southern Oscillation is connected to the warming in the eastern equatorial Pacific, and he provided the first plausible explanation of the dynamical phenomenon. Since then, numerous studies have shown teleconnections or effects of the tropical Pacific warming in remote areas. The relationship between the pressure variations in the Pacific region and the warming in the NIÑO3 region has been documented in the literature. An important test of models is whether they can simulate ENSO (for the right reasons) and whether they can simulate its influence in regions away from the equatorial Pacific. Figure 6.3 shows the lagged correlation of the Southern Oscillation Index (SOI) and NIÑO3 temperature anomalies. The negative correlation indicates that a negative SOI leads warm sea-surface temperatures by up to 2 months. At time lags of approximately  $\pm 20$  months, the maximum positive correlations occur, indicating a composite life-cycle of  $\sim 40$  months, in agreement with the power spectra in Fig. 6.2. CSIRO\_Mk2, CCCma\_CGCM2, and GFDL\_R30\_c, the models that underestimate temperature variability, lack well-defined seasonality and ENSO periodicity of NIÑO3 temperature, and they do not capture the observed variation in the SOI/NIÑO3 temperature correlation at different lags. They all show negative correlations at all lags with varying degrees of success in capturing the peak anticorrelation. ECHAM4\_OPYC3, ECHO-G, and to a lesser extent CCSM2.0, have peak positive correlations at about  $\pm 12$  months, consistent with a shorter life-cycle in association with their peak power occurring at a higher frequency than observed (Fig. 6.2). The remaining models capture the observed lead lag relation between SOI and NIÑO3 temperature anomalies quite well.

To clearly depict the El Niño and La Niña events, we require a consistent ocean/atmosphere response. Since phase-locking with the seasonal cycle is a basic feature of observed ENSO, we will composite ENSO events based on DJF seasonal anomalies of NIÑO3 temperature and SOI. To be classified as a warm (cold) event, we require that the standardized NIÑO3 temperature anomaly be  $\geq 0.6$  ( $\leq -0.6$ ) and the standardized SOI be  $\leq -0.6$  ( $\geq 0.6$ ). The observed and simulated time evolution of NIÑO3

temperature anomalies are shown in Fig. 6.4. Most of the models show a time evolution that is consistent with the observed progression of a warm event. The higher-than-observed amplitude of the MRI\_CGCM2.3 warm event, and to a lesser extent that of HadCM3 and ECHO-G, and the lower-than-observed amplitudes of the CCCma-CGCM2, GFDL\_R30\_c and CSIRO\_Mk2 models are consistent with evidence presented thus far.

In AS2002 the models were classified into three groups based on their ability to represent spatial patterns of boreal winter temperature (TAS), sea-level pressure (PSL) and precipitation (PR) during composite events. For a composite warm event, Group 1 models have well-defined Walker circulation anomalies, with enhanced temperature anomalies extending from the tropical central Pacific to the west coast of South America, and with the strongest rainfall enhancement near the dateline, consistent with observations. Group 2 models are characterized by a westward displacement of the Walker circulation anomalies, usually with maximum temperature and precipitation anomalies in the tropical western/central Pacific. Group 3 models underestimate the positive PSL anomalies in the western Pacific, have weak temperature anomalies in the central/eastern Pacific, and exhibit maximum positive rainfall at or west of the dateline. Relative to AS2002, the new model versions CCCma\_CGCM2, CCSM2.0, GFDL\_R30\_c, and MRI\_CGCM2.3, all exhibit marked improvements in their ENSO characteristics. Presently, CCCma\_CGCM2, GFDL\_R30\_c, and CCSM2.0 models are categorized as Group 2, an improvement from their previous versions that were in Group 3. The MRI\_CGCM2.3 shows a remarkably improved spatial composite (Fig. 6.5) with the temperature, sea-level pressure and precipitation anomalies in the right location, each with amplitudes comparable to observations. The anomaly patterns are consistent with Group 1 models described in AS2002. This is an improvement from the Group 3 classification of the previous model version MRI-CGCM1 considered in AS2002. ECHO-G, not analyzed in AS2002, is consistent with Group 1 models.

### **6.1.5 Summary**

In summary, a followup intercomparison of the ENSO characteristics of CGCMs strongly indicates that newer versions of models (where available) have improved in their simulation of ENSO. The results seen here point to the usefulness of subjecting the models to standardized intercomparisons as a testbed for evaluating model improvement.

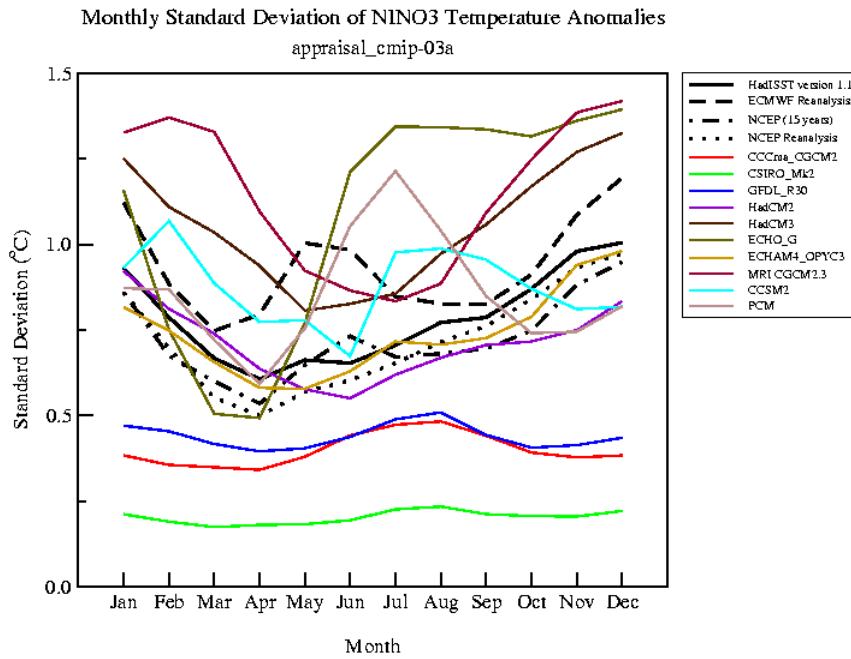


Fig. 6.1. Monthly standard deviations of NIÑO3 surface air temperature anomalies.

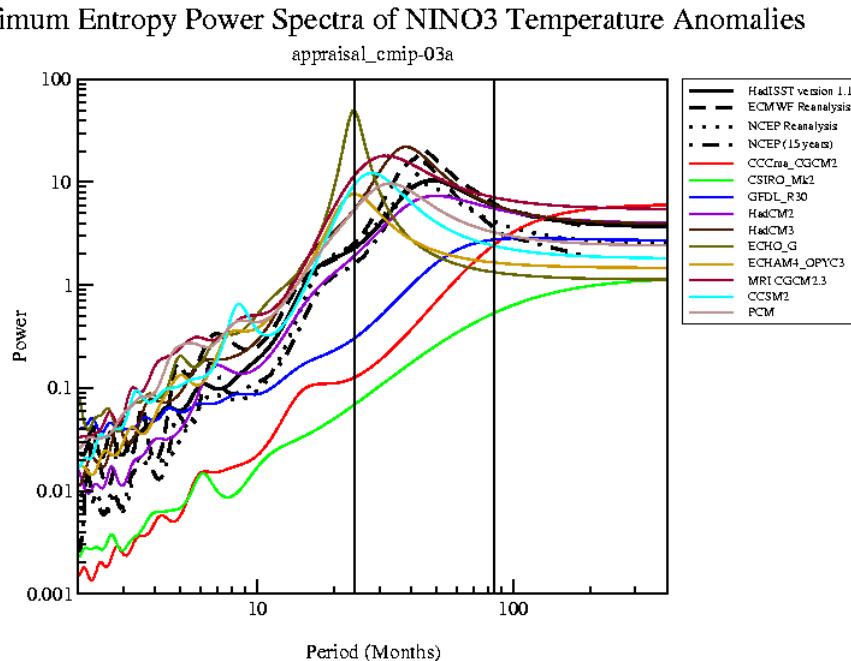


Fig. 6.2. The maximum entropy power calculated for the surface air temperature anomalies (sea surface temperature anomalies from HadISST 1.1) from monthly mean climatology for the NIÑO3 region ( $5^{\circ}\text{S}$ – $5^{\circ}\text{N}$ ,  $150^{\circ}\text{W}$ – $90^{\circ}\text{W}$ ). The vertical lines represent 2- and 7-year periods.

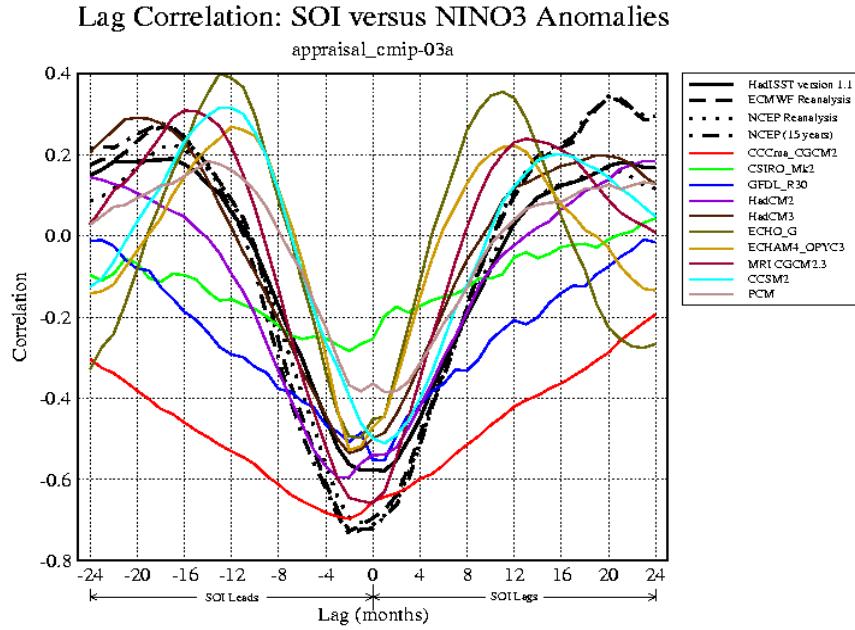


Fig. 6.3. The lag correlation between surface air temperature anomalies in the NIÑO3 region and the SOI as defined above plotted for all models and reanalyses and between the HadISST 1.1 sea surface temperature and CRU SOI datasets.

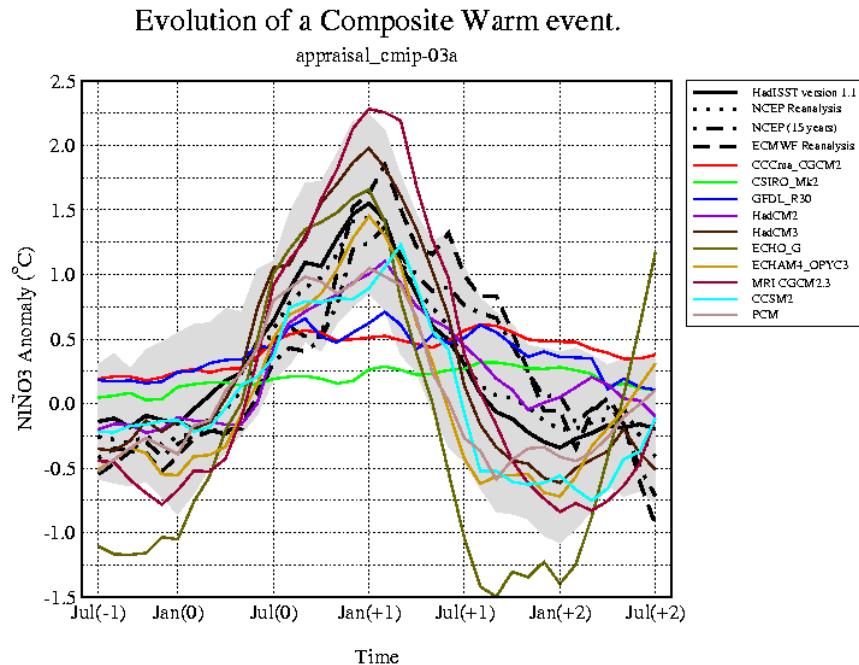


Fig. 6.4. The evolution of the surface air temperature anomaly in the NIÑO3 region for a composite warm event in models, reanalyses and the HadISST 1.1 dataset. The shaded area represents the one standard deviation envelope of the observed NIÑO3 sea surface temperature anomaly for warm events in the HadISST 1.1 dataset.

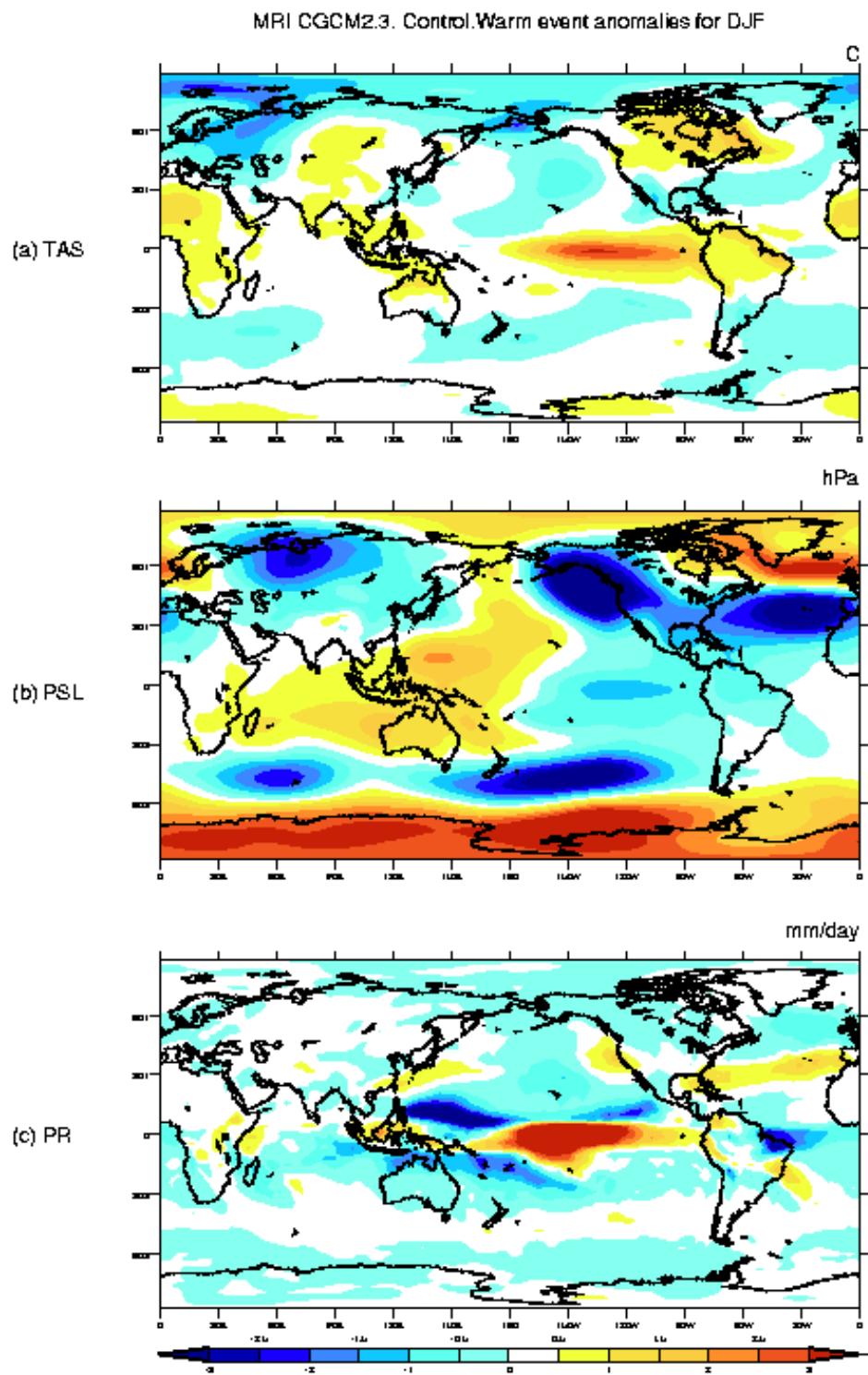


Fig. 6.5. December–February warm event composites of (a) Surface air temperature, (b) Sea-level pressure, and (c) Precipitation from MRI\_CGCM2.3

## 6.2 The North Atlantic Oscillation

### 6.2.1 Introduction

The North Atlantic Oscillation (NAO) is a robust feature of the subtropical and extratropical atmosphere, for example, controlling pressure and temperature variability over much of the Northern Hemisphere. Hurrell et al. (2003) present an overview of the observed NAO, including its influences and seasonality. The simplest index of the NAO is usually based on standardized sea-level pressure anomalies between the Azores and Iceland. This gradient in the sea-level pressure directly affects the strength and penetration of the mid-latitude westerlies into Europe, and is thus related to blocking of synoptic weather systems over Europe. One consequence of the modulation of the storm track is that the NAO can lead to either moderate or colder than normal temperatures over Eurasia. Sir Gilbert Walker recognized that surface temperature is an essential component of the NAO, and included station temperatures in his index of the NAO (Walker and Bliss, 1932). However, to circumvent synoptic conditions (storms etc.) that could contaminate the isolation of the NAO in station records, present techniques of identifying the space–time signature of the NAO usually involve empirical orthogonal function analysis of sea-level pressure over the Atlantic sector. This is the method adopted for the analysis of the NAO in the CMIP2+ and AMIP II models analyzed herein. Since we wish to compare the performance of AMIP II and CMIP2+ runs, we confine our analysis to 20 years of CMIP2+ data to have a comparable record length with the AMIP II models. This would result in a limited number of time points if we were to concentrate our analysis solely on the boreal winter. Though the NAO is most prevalent during the boreal winter, it is present throughout the course of the year (Barnston and Livezey, 1987). Thus, we have chosen to use monthly anomalies for the full calendar year in our analysis. In the future, an analysis using the complete record of simulated data will allow an investigation of the seasonality of the NAO in the models, and test the robustness of the results presented herein. Here, spatial patterns from the CCSM2.0 integration will be presented to illustrate the salient features of the NAO. Additionally, zonal and meridional plots and tabular data will be presented to show results from the suite of coupled models considered. Results from three AMIP II integrations are included in the NAO characteristics table, as these simulations were made using the same atmospheric model as in their CMIP2+ runs, respectively.

The study of the NAO in CMIP has been the subject of intense scrutiny. Stephenson and Pavan (2003) analyzed the NAO in the CMIP I models. They had to base their study on surface air temperature due to the lack of sea-level pressure data in that archive. They found that most models represented the NAO quadrupole pattern in surface air temperature, which for the present models is discussed in Section 6.2.3, and most of the models exhibited persistence of their NAO indices similar to observations. Peer-reviewed manuscripts of the NAO in the CMIP2 models are in preparation (D. Stephenson, 2004, personal communication).

### 6.2.2 Annual Mean Climatology and Variance

#### 6.2.2.1 Sea-level Pressure

Since the primary field through which the NAO is identified is sea-level pressure, the analysis begins with an examination of its annual mean climatology. Given that the NAO

occurs throughout the year, the analysis uses monthly data for the period 1979–95 from observations (the period common to the AMIP II integrations), and from a 20-year segment from each of the CMIP2+ integrations. The climatological sea-level pressures from NCEP/NCAR Reanalysis and CCSM2.0 are given in Figs. 6.6a-b, while the difference (model – reanalysis) is given in Fig. 6.6c. Key features to represent over the Atlantic sector include the location and intensity of high pressure associated with the subtropical anticyclone, and the low pressure to the north extending from Baffin Bay to the Barents Sea. A common feature among the coupled runs and their AMIP II counterparts is the overestimation of sea-level pressure poleward of 60°N over the Atlantic Ocean and north of Russia (Fig. 6.6c).

It is also instructive to examine the variance of the monthly sea-level pressure anomalies (climatological monthly means removed) to ascertain if the model is generating the observed variability maxima north of Siberia, over the Bering Sea, and over the north Atlantic near Iceland (Fig. 6.6d). The variance maximum near Iceland is the northern center of action of the NAO, while the southern center of action tends to be located near the Azores. Systematic error in the PSL variance is also manifest, with an underestimation over the Atlantic Ocean between Greenland and Norway and over western Russia (Fig. 6.6f). This underestimate in CCSM2.0 is not as pronounced as in the other models. However, the overestimate of variance poleward of 70°N in CCSM2.0 is larger than in the other models, and this latter bias is not ubiquitous among these.

### 6.2.2.2 Surface Air Temperature

The climatology and anomaly variance of the surface air temperature (2 m) are given in Fig. 6.7 since the NAO induces large-scale changes in this quantity. In the climatology the main feature of interest is the extension of warm temperatures into the North Atlantic and Barents Sea (Figs. 6.7a-b). In terms of variance (Figs. 6.7d-e), the maxima is located north and east of the 280 K isotherm seen in Figs. 6.7a-b. All of the models qualitatively represent these aspects of the mean state and variance of the surface air temperature. With regard to the time mean error, Fig. 6.7c, there is no compelling evidence of a spatially fixed systematic error among the models. However, from a phenomenological point of view, all models underestimate the variance near the Kara Sea extending to the northwest. The magnitude of the underestimate and the spatial extent of this bias are not consistent from model to model, as they depend on the details of the Gulf Stream extent and its interactions with sea-ice. Comparing the coupled CMIP2+ models and their AMIP II counterparts, respectively, reveals common errors in the mean and variance over land. The exception is HadCM3, which has a cold bias over much of North America compared with HadAM3. However, this may be due to the 2–3°C cold bias over the Pacific Ocean north of 20°N in HadCM3, which may contribute to a colder North America downstream.

### 6.2.3 NAO Teleconnections

To isolate the NAO in the NCEP/NCAR reanalysis, we have subjected the area-weighted monthly anomalies of sea-level pressure over the region 90°W–40°E, 20°N–87.5°N to EOF analysis. EOF-1 isolates the NAO and explains 33% of the total variance. The associated principal component is then regressed against monthly anomalies of sea-level pressure and 850-hPa wind, and the results for a one standard deviation perturbation of PC-1 are plotted in Fig. 6.8a. The data are only plotted at gridpoints at which the

regression is significant at the 5% level that includes a sample size adjustment for estimating the effective degrees of freedom based on the standard error of the slope and critical t-value (Santer et al., 2000). The resulting sea-level pressure and wind anomalies correspond to the positive (high) index phase of the NAO in which stronger than normal mid-latitude westerlies penetrate into Europe.

In order to perform a consistent analysis across all models, the simulated monthly sea-level pressure anomalies from each model are projected onto the observed NAO EOF-1. Thus, for each model we obtain a principal component time series, hereafter known as PC-1. Analogous to the procedure above, for each model, the resulting PC-1 time series is regressed against the simulated monthly anomalies of sea-level pressure, 850-hPa wind, and surface air temperature, and the result for a one standard deviation perturbation of PC-1 is generated. The sea-level pressure and 850-hPa wind (surface air temperature) regressions for CCSM2.0 are presented in Fig. 6.8 (Fig. 6.9). For both sea-level pressure and surface air temperature, the observed patterns of anomalies are consistent with those presented in Hurrell et al. (2003), but the amplitudes are weaker since the present analysis uses all months of data, while Hurrell et al. (2003) emphasized the boreal winter during which the anomalies are strongest. The location of the simulated sea-level pressure dipole centers and their intensity are critical. During the positive (high) index phase of the NAO depicted in Figs. 6.8a-b, the subtropical anticyclone over the Atlantic Ocean is strengthened, and the Icelandic low deepens. The southeasterlies near the east coast of the United States are consistent with the warmer temperatures there due to the enhanced subtropical influence (Figs. 6.9a-b). Similarly, the mid-latitude westerlies result in a warmer Europe, Scandinavia, and Russia. Over Africa and the Middle East the stronger northerlies from the mid-latitudes result in cooler-than-normal temperatures. Greenland and northeastern Canada also experience colder temperatures due to the enhanced polar influence. These features are common among the coupled CMIP2+ models and their AMIP II counterparts, though subtle shifts in the simulated pressure and wind response result in magnitude and spatial differences in the temperature response, particularly over Africa, and to a lesser extent along the east coast of the United States.

The model error in the pressure and wind, and surface air temperature regressions, are given in Figs. 6.8c and 6.9c, respectively. The overestimate of sea-level pressure in the North Atlantic Ocean near 60°N and the associated anticyclonic anomalies result in a cold bias over Eurasia during the positive (high) index phase of the NAO. Conversely, over Greenland and Canada the sea-level pressure anomalies are below normal and the temperature is biased positive. These errors are not only characteristic of CCSM2, but are common among the CMIP2+ models and their AMIP II counterparts. This similarity in the error structure of the coupled and uncoupled integrations is consistent with the NAO being an internal mode of the atmosphere.

The ability of the CMIP2+ models to capture the north-south sea-level pressure anomaly dipole over the Atlantic sector is shown in Fig. 6.10. The data are averaged between 90°W–40°E, and shown from 20°N–87.5°N, which is the domain of the EOF analysis used to isolate the NAO from the NCEP/NCAR reanalysis. The zonal average was taken prior to masking the regressions based on the assessment of 5% significance. The largest uncertainty is in the magnitude of the deepening of the Icelandic low, though the majority of the models are typically within 1 hPa of the observed pressure dipole. The east-west extratropical NAO temperature response, averaged from 50°N–85°N, is shown in Fig. 6.11. The models readily capture the positive temperature gradient from the western

to eastern hemisphere, including the locations of the minima and maxima temperature perturbations. Consistent with systematic model error, the temperature estimates near Greenland are not cold enough, and the warming near Norway and Russia is underestimated.

Comparison statistics for NAO variability are given in Table 6.1. The standard deviation of PC-1 is a direct measure of the strength of the NAO variations, to which the regressions in the afore-mentioned figures are proportional. For most models the PC-1 variability is within 10–15% of the observed estimate from reanalysis. The NAO is highly variable in time, and does not exhibit a preferential periodicity on time scales of weeks to years. As such, the phase of the NAO on these time scales shows little persistence. The lack of month-to-month persistence of the phase of the observed NAO is reflected by the one-month-lag autocorrelation of PC-1 being only ~0.19. This estimate is consistent with that obtained by Stephenson and Pavan (2003) for the year-to-year autocorrelation ( $0.17 \pm 0.10$ ) estimated using a fractionally differenced autoregressive order 1 model. With such a low autocorrelation, the power spectrum of PC-1 is very weakly red (e.g., slightly more power at longer time scales; not shown). In some cases, the models exhibit greater month-to-month persistence (a larger lag-one autocorrelation) in the phase of the NAO compared with that estimated from the reanalysis. In such cases the NAO tends to remain in one phase or the other for longer than is observed, on average. Thus, over a given period of time, the model will have fewer transitions between the opposing phases of the NAO, which is then related to the frequency and duration of blocked and unblocked conditions. However, additional observed and simulated data need to be analyzed to better characterize the interdecadal variability of the month-to-month persistence of the NAO.

Examination of the data in Table 6.1 reveals an interesting relationship between the variability and the lag-one autocorrelation of the principal components. As seen in Fig. 6.12, the lag-one autocorrelation tends to increase with the standard deviation of the NAO PC. The result is robust between the CMIP2+ and the (full suite of) AMIP II models. Essentially, it takes a longer period of time for the transition to occur between the high- and low-index phases of the NAO when the amplitude of the NAO excursions is large.

#### 6.2.4 Summary

The good overall simulation of the NAO is represented by the low root-mean-squared-error (RMSE) and high pattern correlations between the observed and simulated sea-level pressure (Table 6.1). While the AMIP II models each have lower RMS error than their respective coupled counterparts, this may be due to sampling. The analysis of the simulated NAO during other decades will enable an assessment of the robustness of the results presented herein, and with a longer record the seasonality of the NAO can be explored. While the NAO teleconnection pattern with surface air temperature is well represented by all of the models, there is systematic error in the magnitude of the anomalies (e.g., Eurasia too cold during the positive phase of the NAO, Fig. 6.9c) which is related to the systematic model error in the sea-level pressure (Fig. 6.8c). An improved representation of the extra-tropical flow should reduce these systematic errors.

Table 6.1. Reanalyzed, CMIP2+, and AMIP II (the last 3 rows) NAO characteristics. The columns give the observation/model designation, the standard deviation of PC-1 (hPa), the one-month lag correlation of PC-1, RL1, and the root-mean-squared error, RMSE (hPa), and pattern correlation, Rpat, of the linear regression of PC-1 with sea-level pressure (scaled by a one-standard-deviation perturbation of PC-1). Shaded models used the same atmospheric component in their CMIP2+ and AMIP II simulations.

<b>Model</b>	<b>PC-1</b>	<b>RL1</b>	<b>RMSE</b>	<b>Rpat</b>
NCEP/NCAR	42.2	0.19	0.00	1.00
CCCMa_CGCM2	34.3	0.25	0.37	0.97
CCSM2	44.5	0.20	0.45	0.93
CSIRO_Mk2	31.6	0.19	0.36	0.96
ECHAM4_OPYC3	35.9	0.14	0.30	0.97
ECHO-G	36.3	0.24	0.36	0.95
GFDL_R30_c	38.1	0.22	0.29	0.97
HadCM2	43.4	0.21	0.35	0.96
HadCM3	43.8	0.30	0.32	0.97
MRI_CGCM2.3	48.0	0.38	0.36	0.97
PCM	39.9	0.40	0.38	0.95
CAM2.0	44.3	0.33	0.36	0.96
ECHAM4	41.1	0.18	0.29	0.97
HadAM3	36.4	0.24	0.29	0.97

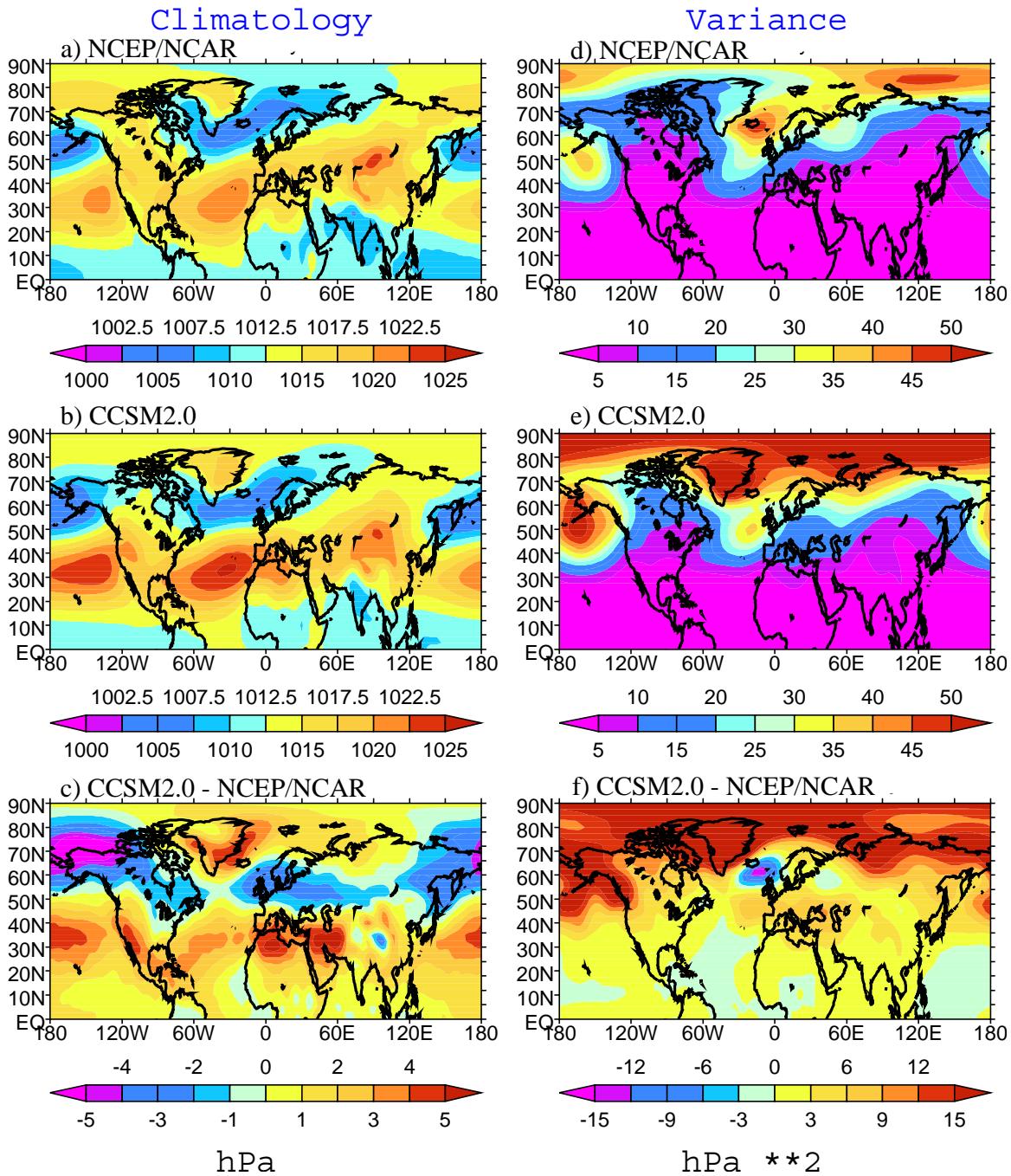


Fig. 6.6. Annual mean sea-level pressure climatology (hPa): (a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) model–NCEP/NCAR reanalysis. Variance of monthly mean sea-level pressure anomalies (hPa<sup>2</sup>), (d) NCEP/NCAR reanalysis (1979–95), (e) 20 years of CCSM2.0, (f) model–NCEP/NCAR reanalysis.

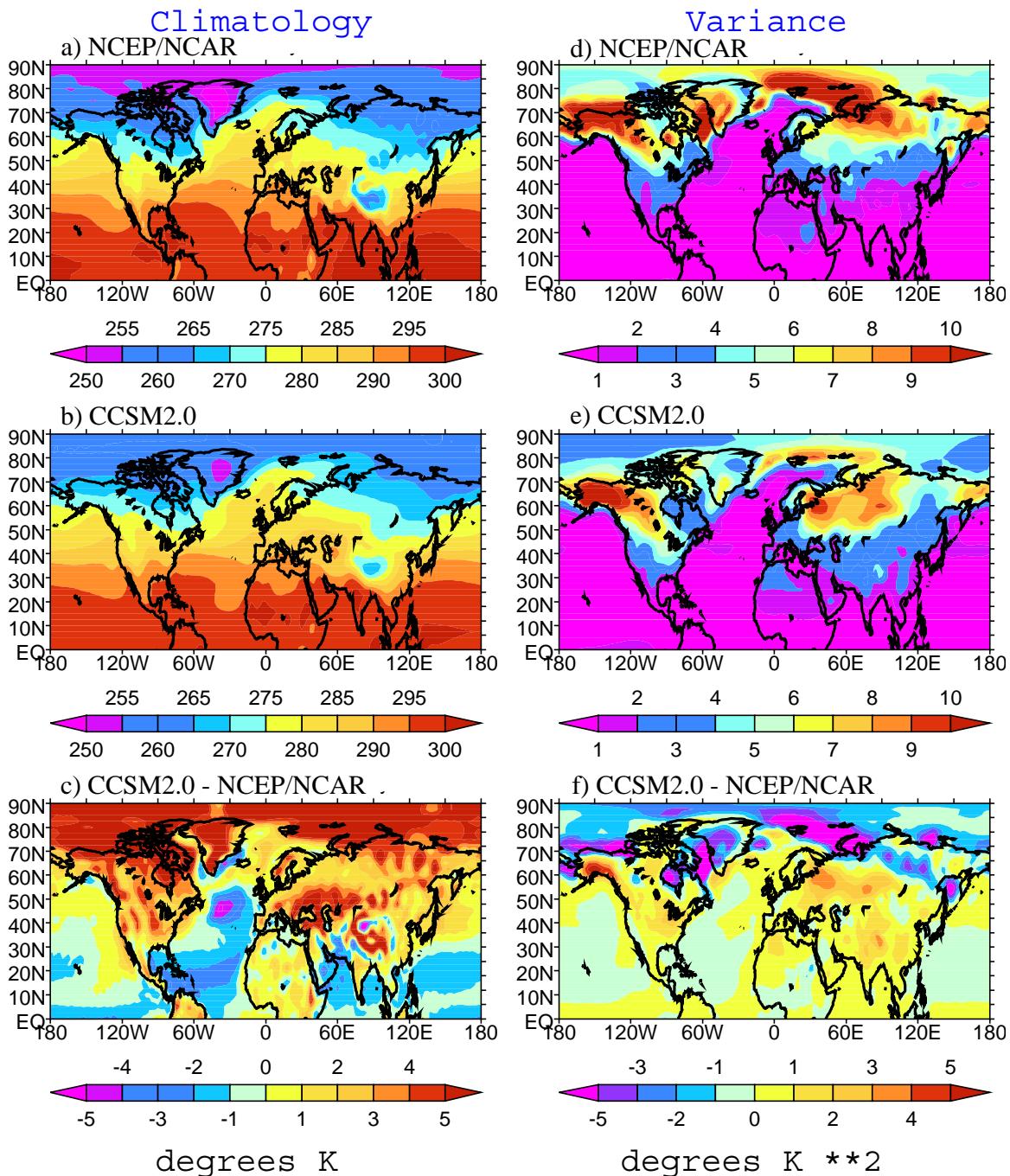


Fig. 6.7. Annual mean surface air temperature climatology (K): (a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) model–NECP/NCAR reanalysis. Variance of monthly mean surface air temperature anomalies ( $K^2$ ), (d) NCEP/NCAR reanalysis (1979–95), 20 years of CCSM2.0, (f) model–NCEP/NCAR reanalysis.

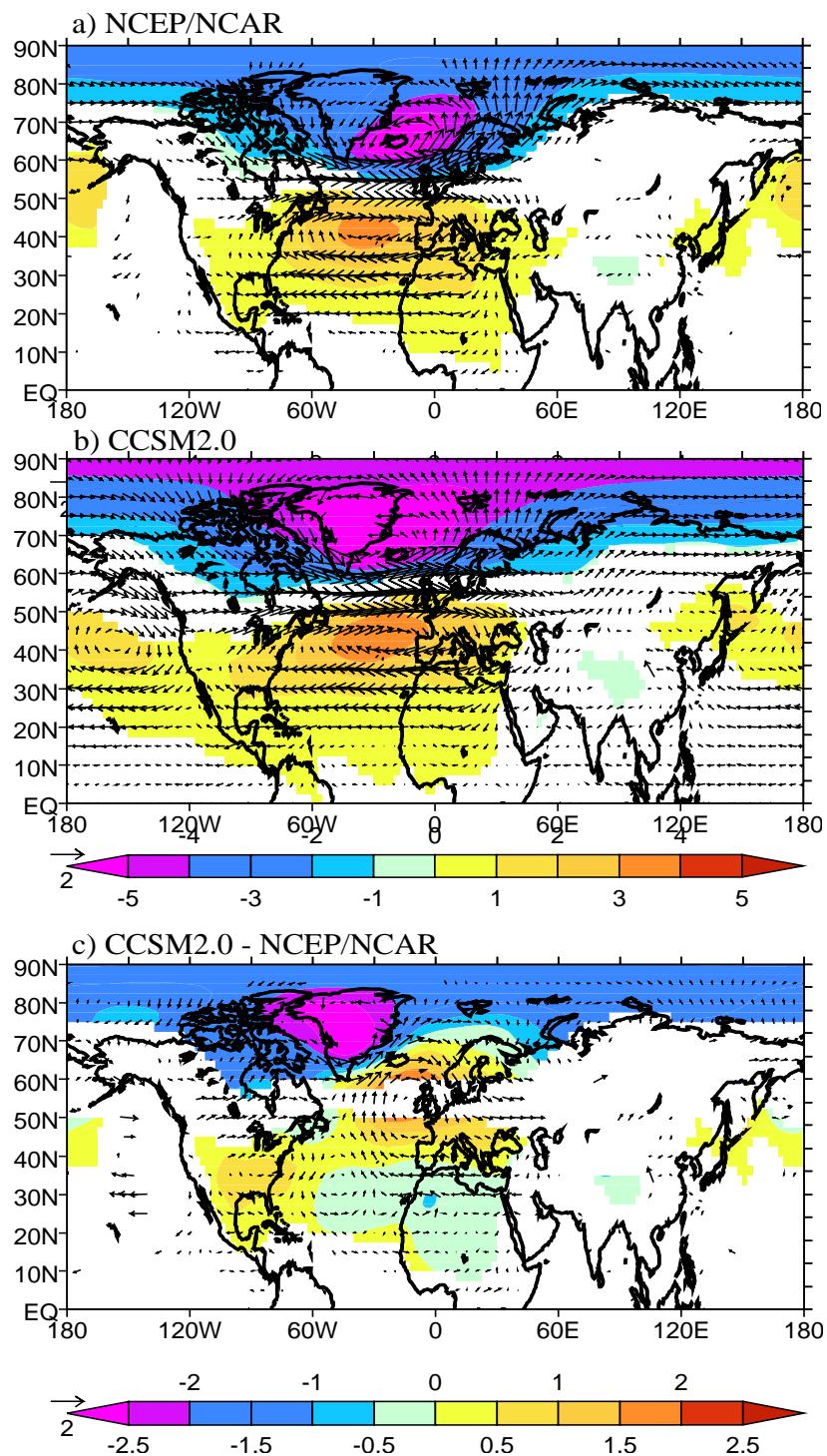


Fig. 6.8. Linear regression of the NAO principal component with monthly anomalies of sea-level pressure (hPa) and 850-hPa wind ( $\text{ms}^{-1}$ ):  
(a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) CCSM2.0–NCEP/NCAR reanalysis. The regressions have been scaled by a one-standard-deviation perturbation of the respective principal components, and plotted where they are at least 5% significant.

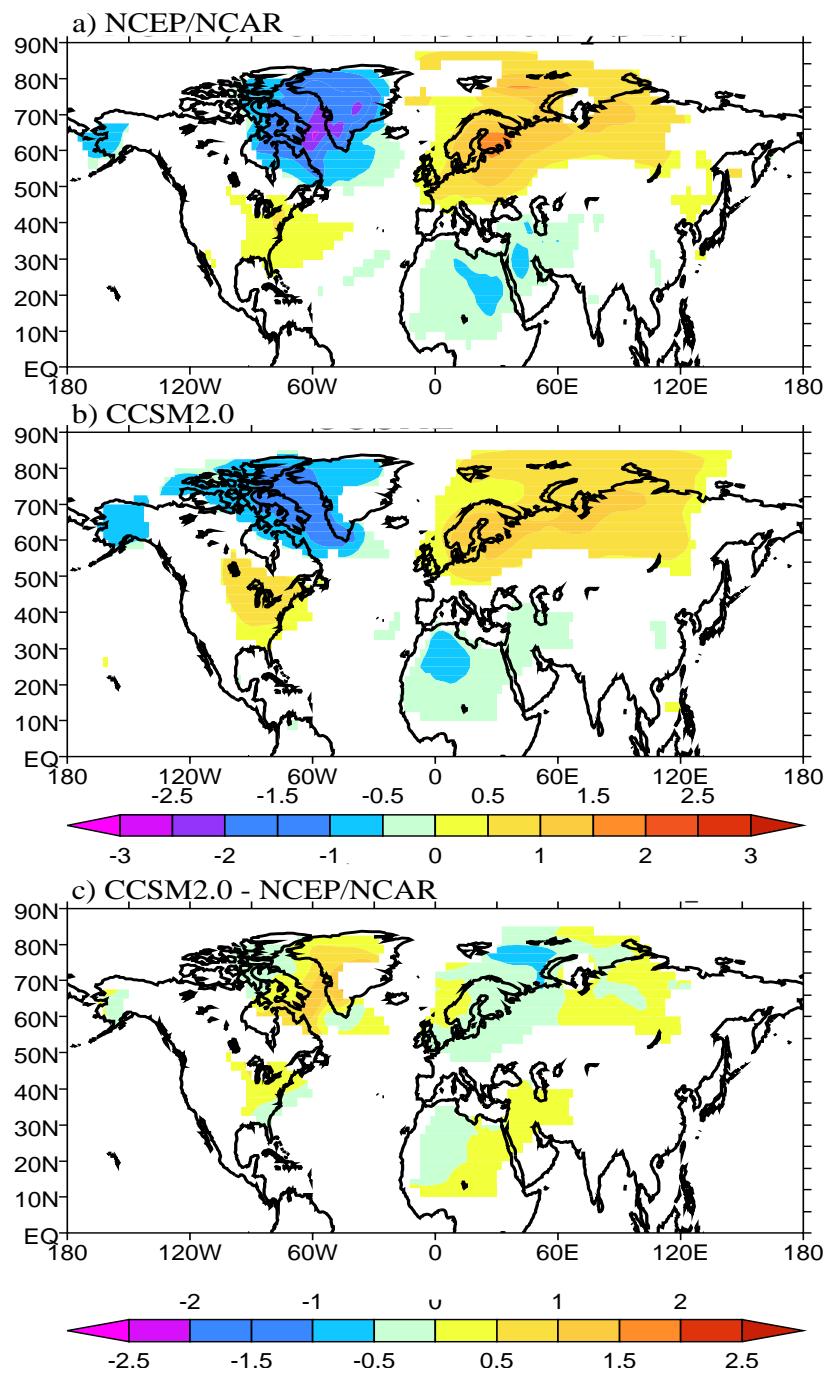


Fig. 6.9. Linear regression of the NAO principal component with monthly anomalies of surface air temperature (K): (a) NCEP/NCAR reanalysis (1979–95), (b) 20 years of CCSM2.0, (c) CCSM2.0–NCEP/NCAR reanalysis. The regressions has been scaled by a one-standard-deviation perturbation of the respective principal components, and plotted where they are at least 5% significant.

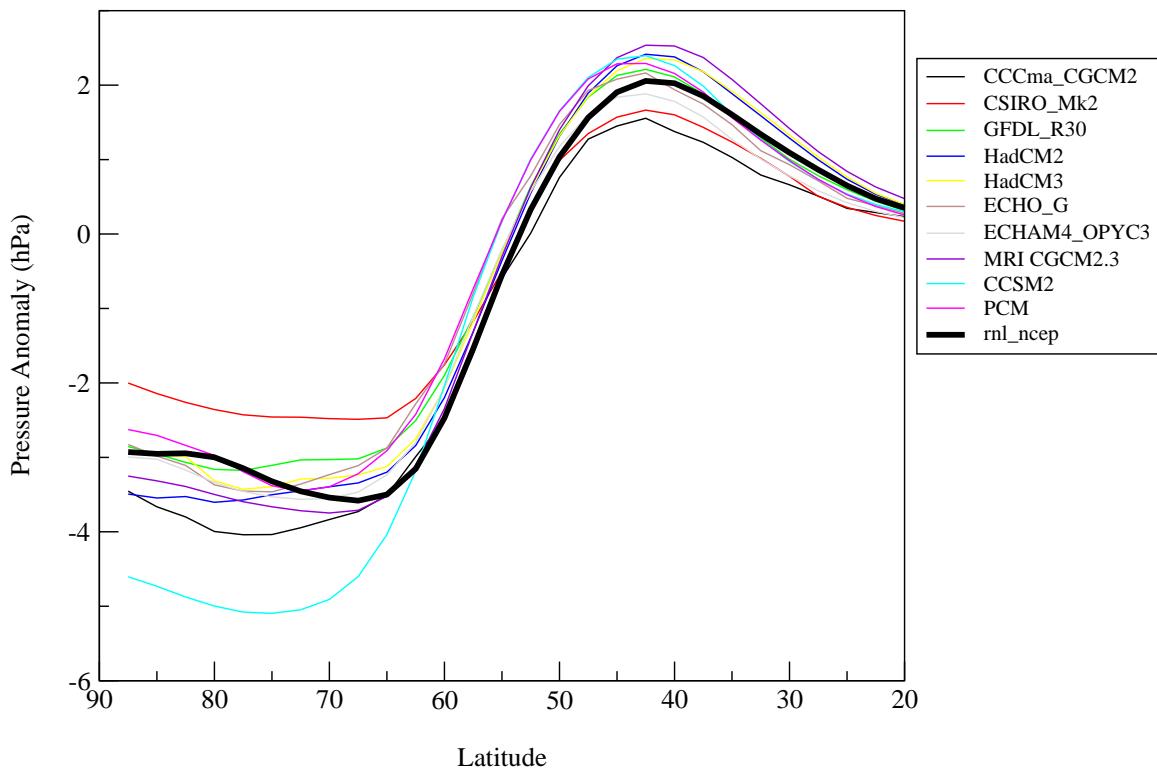


Fig. 6.10. Zonal mean ( $90^{\circ}\text{W}$ – $40^{\circ}\text{E}$ ) of the regression of the observed and simulated NAO principal components with monthly anomalies of sea-level pressure (hPa). The regressions have been scaled by a one-standard-deviation perturbation of the respective principal components. The zonal averaging has been done prior to masking for statistical significance.

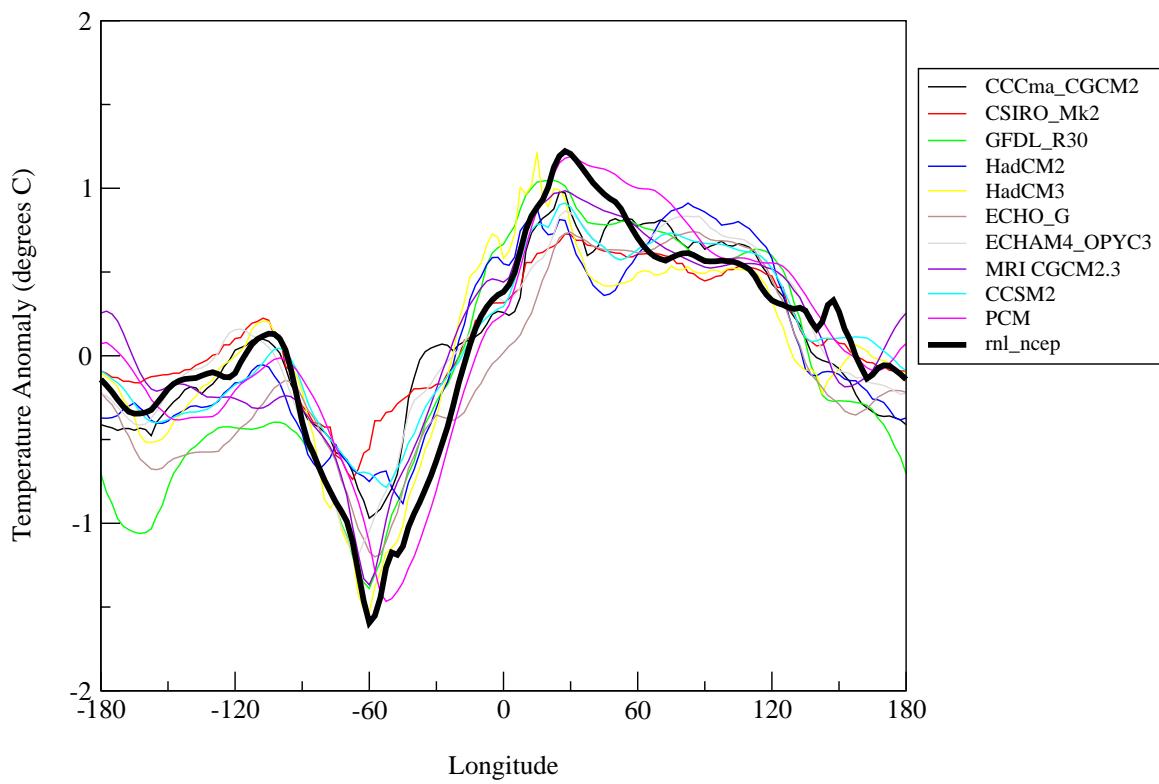


Fig. 6.11. Meridional mean ( $50^{\circ}\text{N}$ – $85^{\circ}\text{N}$ ) of the regression of the observed and simulated NAO principal components with monthly anomalies of surface air temperature (K). The regressions have been scaled by a one-standard-deviation perturbation of the respective principal components. The zonal averaging has been done prior to masking for statistical significance.

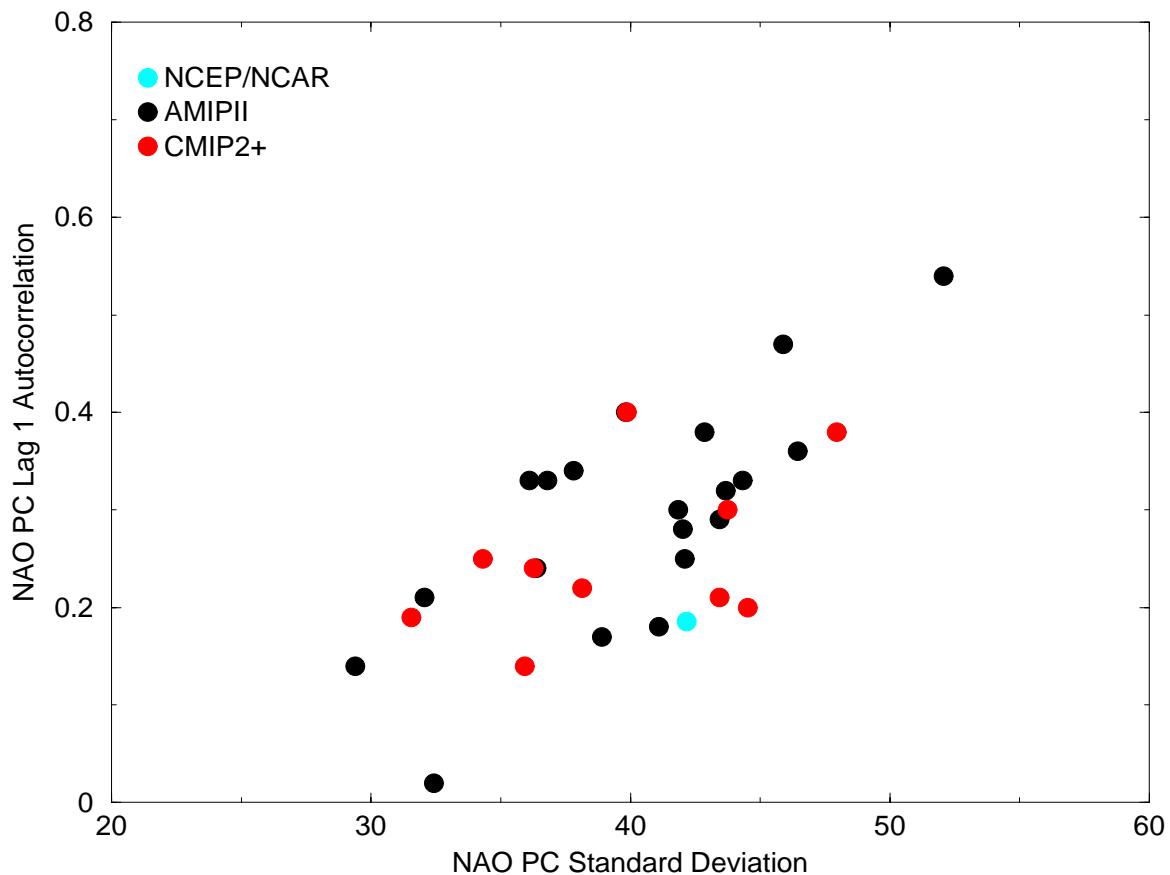


Fig. 6.12. Scatter-plot of the standard deviation of the NAO principal components with respect to their lag-1 autocorrelation. The result(s) from NCEP/NCAR reanalysis is given in blue; those from the coupled CMIP2+ and AMIP II models are given in red and black, respectively.

### 6.3 Tropical Wavenumber–Frequency Spectra

Wheeler and Kiladis (1999) showed that frequency–wavenumber power spectra of outgoing longwave radiation (OLR) in the tropics correspond closely to equatorially trapped waves from shallow water theory (Matsuno, 1966; Lindzen, 1967). These convectively coupled waves were later identified in ECMWF Reanalysis (ERA-15) and brightness temperature derived from the International Satellite Cloud Climatology Project B3 data using a different approach (Yang et al., 2003). Similar to Wheeler and Kiladis (1999), observed and simulated OLR is decomposed into eastward and westward propagating components as a function of wavenumber after removal of the climatological daily means, and the power spectra for the symmetric and antisymmetric components for data between 15°N–15°S are presented. The sum of the symmetric and antisymmetric components gives the total power. The average power spectrum is based on the analysis of successive 96-day segments that overlap by 2 months for the duration of each dataset. Since we use daily data, periods of 2 to 96 days are resolved. Based on the total spectrum, a smoothed background spectrum is calculated, which when divided into the symmetric and antisymmetric spectra reveals the equatorial wave modes.

Consistent with the results of Wheeler and Kiladis (1999), the analysis of AVHRR OLR resolves mixed Rossby-gravity waves and  $n = 0$  eastward inertio-gravity waves in the antisymmetric spectrum (Fig. 6.13a), and  $n = -1$  Kelvin waves,  $n = 1$  equatorial Rossby waves,  $n = 1$  westward inertio-gravity waves, and the MJO, which dominates eastward variance with a period of 48 days, are identified in the symmetric spectrum (Fig. 6.13b). Due to the Nyquist frequency cutoff,  $n = 2$  westward inertio-gravity waves are not resolved. With the exception of the MJO, which does not fall on the dispersion curves and is distinct from the Kelvin waves, all of the aforementioned waves typically occur for equivalent depths of approximately 12–50 m.

Figures 6.13c-d indicate that the NCEP/NCAR reanalysis simulates these modes of variability, though with reduced variance compared with the AVHRR OLR. However, the  $n = -1$  Kelvin waves,  $n = 0$  equatorial inertio-gravity waves, and the mixed Rossby-gravity waves tend to have variance maxima at greater equivalent depths than observed. For the former two modes (latter mode) this corresponds to the variability occurring at smaller (larger) zonal wavenumbers compared with the AVHRR OLR. The equivalent depth is sensitive to the static stability (Wheeler and Kiladis, 1999), and the theoretical dispersion curves are calculated assuming a resting atmosphere. As discussed in Yang et al. (2003), doppler shifting can strongly affect the slopes of the dispersion curves in non-resting conditions.

Common features of the models (Figs. 6.13e-r) include an underestimate of variance at time scales of less than about 10 days in both the antisymmetric and the symmetric components compared with the AVHRR OLR. The  $n = -1$  Kelvin waves, the mixed Rossby-gravity waves, and the  $n = 0$  eastward inertio-gravity waves are particularly poorly represented in most models, and similar to the NCEP/NCAR reanalysis OLR, equatorial waves at greater than observed equivalent depths can occur.

An interesting feature is that the models typically have power in the observed MJO band (symmetric spectrum, eastward wavenumbers 1–3, period = 48 days). However, in most cases they also have substantial power at the same westward propagating wavenumbers at the same time scale, suggesting that the simulated intraseasonal variability may be more

akin to a standing mode depending upon their phasing at any given time. For clarity, the observed westward propagating  $n = 1$  equatorial Rossby waves have dominant power at 30 days at wavenumbers 3–4 in contrast to the MJO where the dominant power is at 48 days in wavenumbers 1–3. This example serves to illustrate that critical examination of frequency–wavenumber plots is warranted. Additionally, one should not extrapolate frequency–wavenumber spectral peaks to indicate that a model successfully simulates the observed space–time variability of particular modes. Rather it is imperative to perform process studies of the individual modes to ascertain the fidelity with which the observed space–time characteristics are simulated. Such is the case for the MJO, which is analyzed in the next subsection.

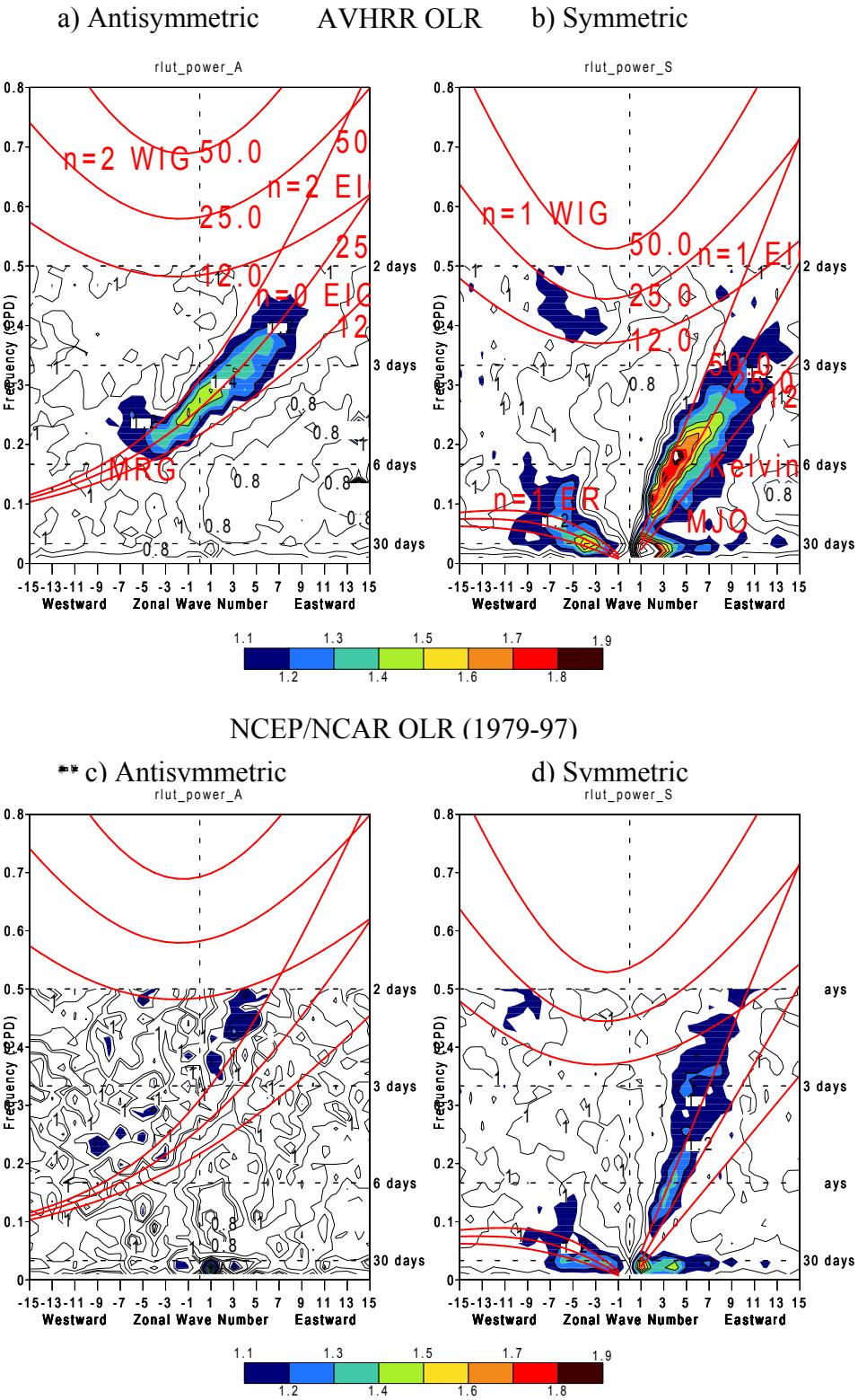


Fig. 6.13. Frequency–wavenumber plots of OLR between  $15^{\circ}\text{N}$ – $15^{\circ}\text{S}$  decomposed into eastward- and westward-propagating wavenumbers for the antisymmetric and symmetric spectra: (a–b) AVHRR OLR, (c–d) NCEP/NCAR Reanalysis (1979–97) OLR, (e–f) CCSM2.0 OLR, (g–h) CSIRO\_Mk2 OLR, (i–j) , (k–l) ECHO-G OLR, (m–n) GFDL\_R30\_c OLR, (o–p) HadCM3 OLR, and (q–r) PCM OLR.

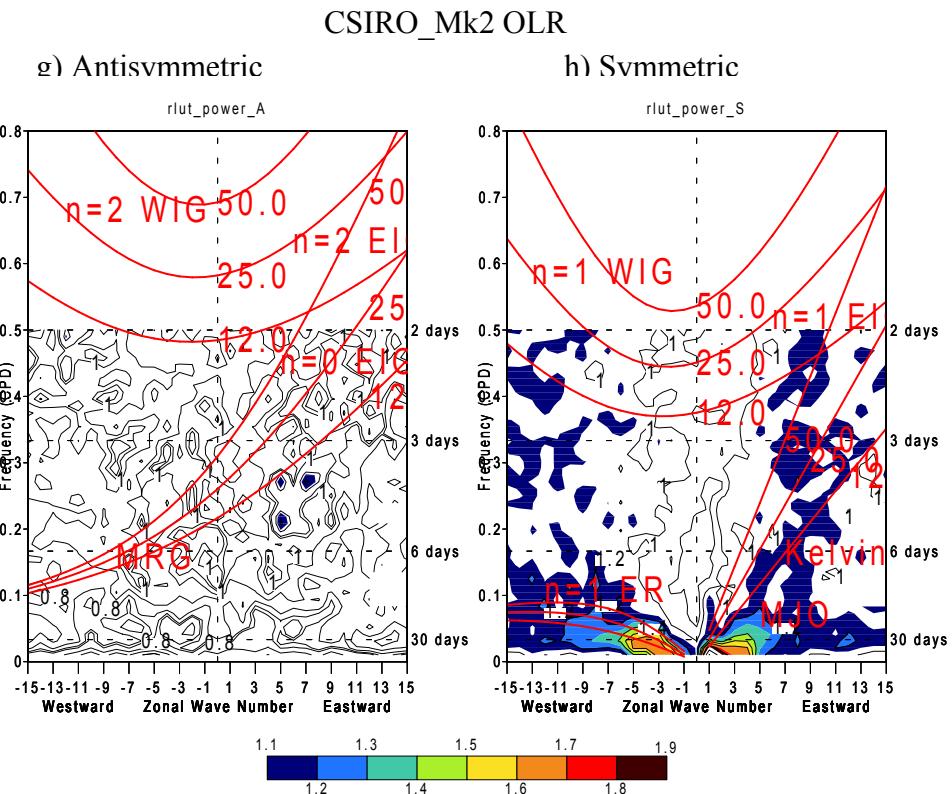
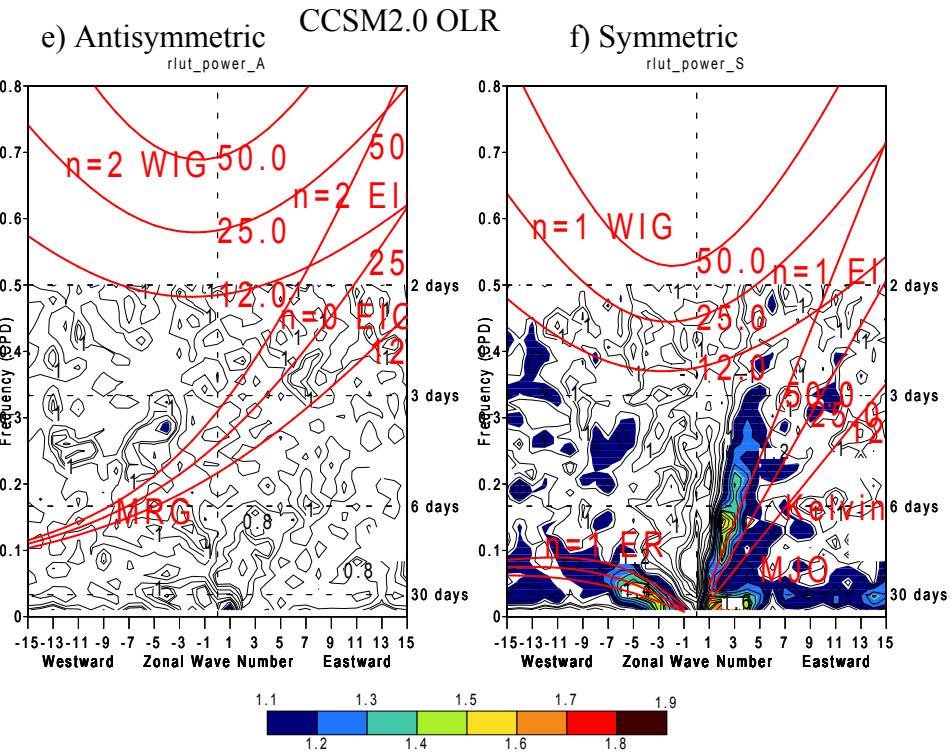
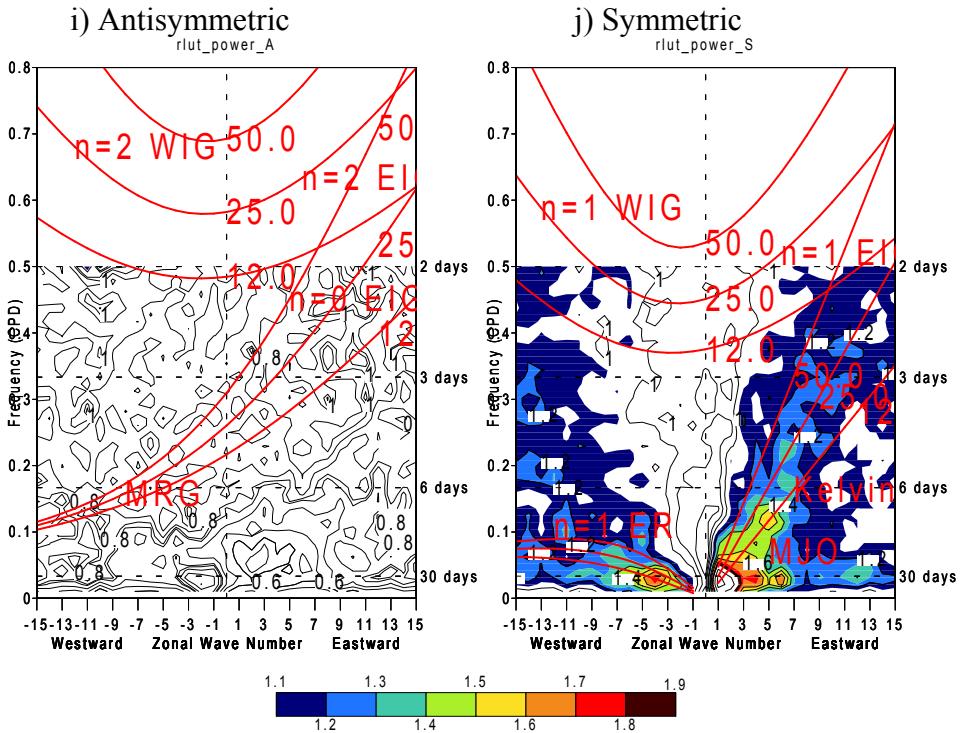


Fig. 6.13. (continued).

### ECHAM4 OPYC3 OLR



### ECHO-G OLR

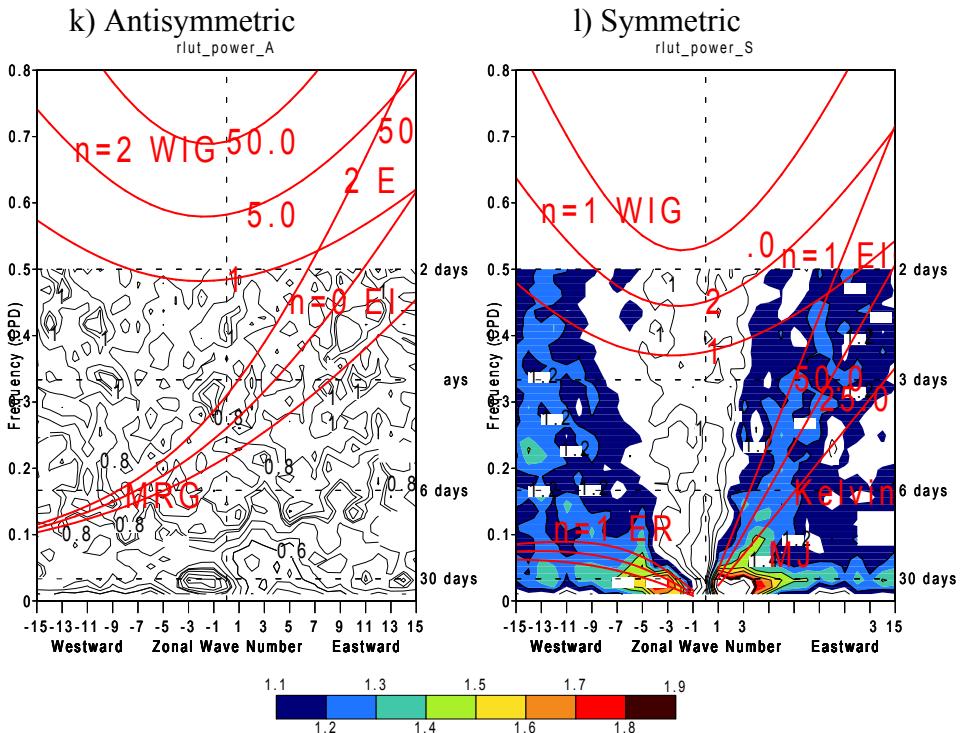
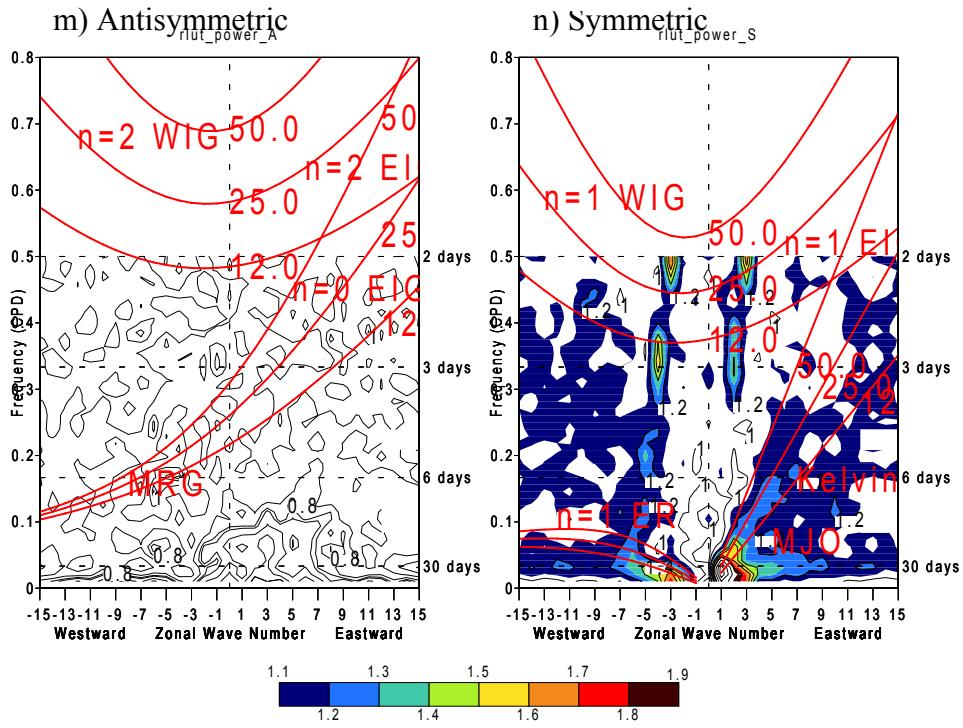


Fig. 6.13. (continued).

### GFDL\_R30\_c OLR



### HadCM3 OLR

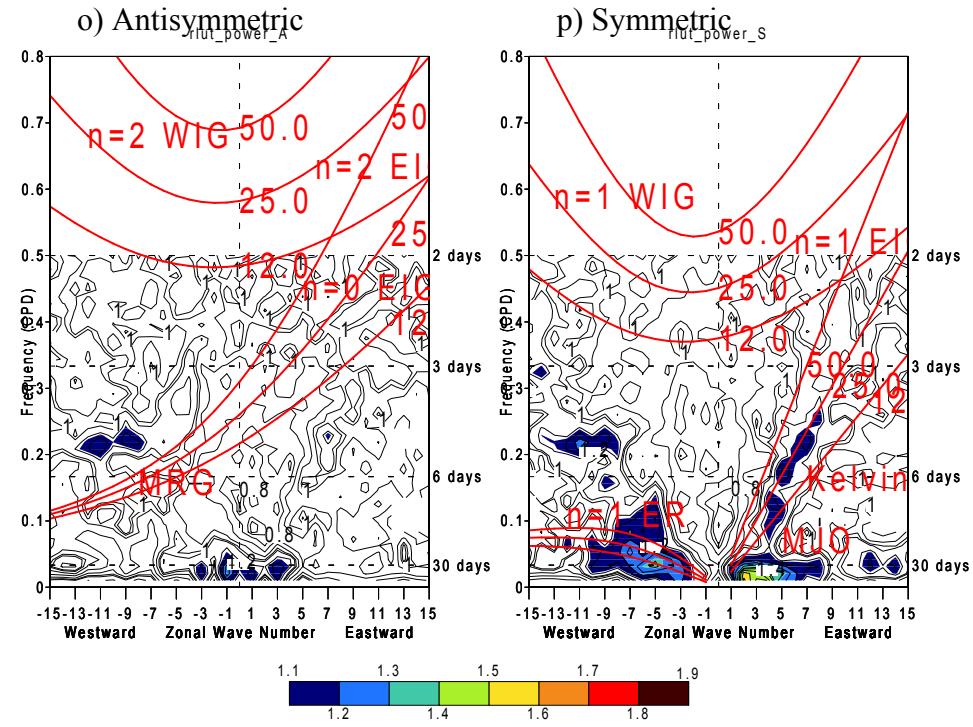
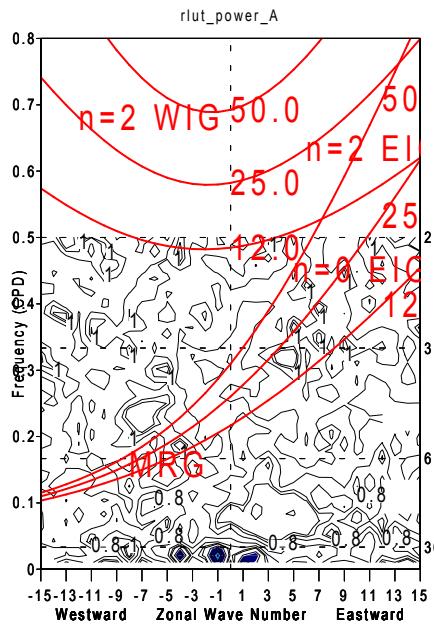


Fig. 6.13. (continued).

### PCM OLR

q) Antisymmetric



r) Symmetric

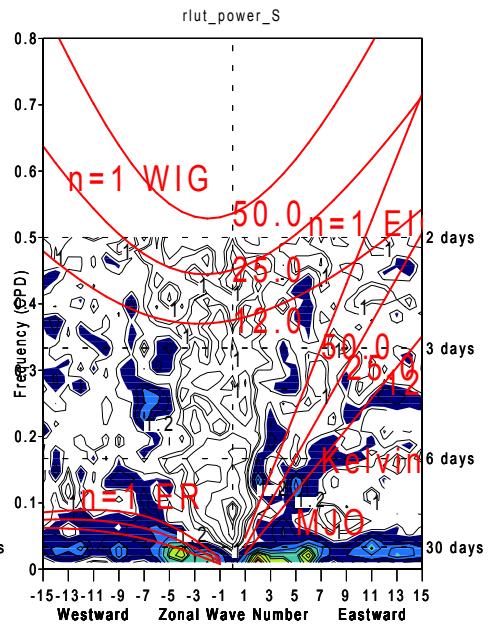


Fig. 6.13. (continued).

## 6.4 The Madden–Julian Oscillation

### 6.4.1 Introduction

The Madden–Julian Oscillation (MJO) dominates tropical variability on time scales of ~30–70 days (Madden and Julian, 1971, 1972). It is manifested through large-scale circulation anomalies in conjunction with eastward propagating convective anomalies over the eastern hemisphere, and is strongest during the boreal winter/spring. The convective anomalies are of sufficient spatial extent and duration to result in extratropical teleconnections (Weickmann et al., 1985; Murikami, 1988), including a link to rainfall over the western United States (Mo and Higgins, 1998), and an improvement in extratropical skill in medium- and extended-range numerical weather predictions (Ferranti et al., 1990).

Near the equator, low-level moisture convergence is the mechanism through which the eastward propagation of the MJO is maintained. Enhanced convergence (Rui and Wang, 1990) and boundary layer moisture (Hendon and Salby, 1994; Jones and Weare, 1996; Maloney and Hartmann, 1998) destabilize the atmosphere in advance of the main center of convection. Additionally, Sperber (2003) notes that free-tropospheric interactions impact the life-cycle of the MJO. Off the equator, Kemball-Cook and Weare (2001) find that the pre-moistening of the boundary layer is not due to low-level convergence; rather local thermodynamic processes govern the development of convective instability, consistent with the “discharge-recharge” mechanism of Blade and Hartmann (1993). This wide range of interactions, and the possible influence of the extratropics (Hsu et al., 1990), attest to the complexity of the MJO, which has proven to be a challenging test of a model’s ability to simulate the tropics (Hayaashi and Golder, 1986, 1988; Park et al., 1990; Slingo and Madden, 1991; Hayashi and Golder, 1993; Slingo et al., 1996; Sperber et al., 1997).

Here we utilize a methodology to analyze the Madden–Julian Oscillation (MJO) in general circulation models in a common fashion. We begin by projecting 20–100-day bandpass-filtered outgoing longwave radiation (OLR) from the models onto the two leading empirical orthogonal functions (EOFs) of observed OLR that characterize the propagation of MJO convection from the Indian Ocean to the central Pacific Ocean. The resulting principal component (PC) time series are then screened to isolate boreal winters during which they exhibit a lead-lag relationship consistent with observations. This PC subset is used for linear regression to determine the ability of the models to simulate the observed space-time variability of the MJO.

### 6.4.2 The Models and Validation Data

Here we analyze Madden–Julian variability in the CMIP2+ models and AMIP II models to determine the extent to which the MJO is simulated, and the influence that air–sea interaction has on the representation of the MJO. The CMIP2+ analysis is based on the last 20 years of data submitted to PCMDI. The exceptions are NCAR CCSM2 and HADCM3 (Reading University integration from P. Inness), where data were provided directly to K.R. Sperber. The results are compared against National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al., 1996) and the Advanced Very High Resolution Radiometer outgoing longwave radiation (AVHRR OLR; Liebmann and Smith, 1996) for 1979/80–

1994/95. This observed period also corresponds to that analyzed from the AMIP II integrations. All data are bandpassed with a 20–100-day (Lanczos) filter.

Results from the European Centre Hamburg version 4/Hamburg Ocean Primitive Equation model (known as ECHO-G; Legutke and Maier-Reimer, 1999; Legutke and Voss, 1999) are highlighted due to the plethora of model output available.

#### 6.4.3 OLR and the MJO

Figures 6.13a-b show the observed and simulated OLR climatology. The observed features of import include the convective maximum extending from the Indian Ocean into the western Pacific and South Pacific Convergence Zone (SPCZ). Figure 6.13c shows the daily variance of the OLR from observations which is overestimated by the model, Fig. 6.13d. Importantly, the variability tends to be largest where the convection is strongest. Figures 6.13e-f indicate that the observed and simulated 20–100-day variance is largest over the Indian Ocean, the Banda Sea, and the western/central Pacific Ocean. The percent of the total variance explained by 20–100-day time scales is largest in the central/eastern Indian Ocean and the western Pacific (Figs. 6.13g-h). Though the model has overestimated the intraseasonal variance, importantly, the spatial pattern of intraseasonal variance is well represented.

Sperber (2003) identified seven years when the boreal winter MJO was notably active as a well-defined eastward-propagating mode. Using these periods, the eastward propagation of convection was isolated via EOF analysis of filtered AVHRR OLR. In the present study, filtered OLR from satellite data and the models is projected onto the aforementioned EOFs, yielding principal component time series (PCs). Thus, all models are evaluated relative to a common metric, assuring that the model–model and model–data comparisons are consistent. The analysis is confined to the months November–March, for 1979/80–1994/95 for the observations and the AMIP II models, and for 9–19 winters from the coupled models. November–March is used since this is when the boreal winter eastward-propagating MJO dominates the tropics. In the subsequent regression analysis, data are plotted when the regression is 5% significant or better, assuming each pentad is independent as in Sperber et al. (1997) and Sperber (2003).

The observed PCs exhibit a characteristic lead/lag structure (Fig. 6.14a). From the average of all winters (the thick black line), the maximum positive correlation indicates that PC-2 leads PC-1 with a time scale of 12 days. Regression of the PCs with 20–100-day filtered OLR is presented in Figs. 6.15a and Figs. 6.15c. For a one-standard-deviation perturbation of the PCs, the strongest convective anomalies are  $\sim 20 \text{ Wm}^{-2}$ , with the convection over the Indian Ocean leading that over the west Pacific Ocean.

The maximum positive correlations as a function of time lag for the individual winters are plotted in Fig. 6.14c, where the upper-right quadrant is taken to be the observational phase-space. The lead-lag structure of the PCs from ECHO-G are given in Fig. 6.14b. Sixteen of nineteen winters have a maximum positive correlation vs. time lag that falls within the observational phase-space (Fig. 6.14c), and only these winters are used for the subsequent regressions to best isolate the eastward propagating convective anomalies. The average lead-lag structure of these 16 years, the thick dashed black line in Fig. 6.14b, indicates that PC-2 leads PC-1 by 12 days, the maximum positive correlation being 0.68 (see Table 6.2). This is comparable to the observed lead-lag structure, and is an

improvement relative to the average over all years, the thick solid line in Fig. 6.14b. For the ECHO-G model, the regressions of the PCs with filtered OLR in Figs. 6.15b and d agree well with those observed. However, the maximum convective anomalies exceed those from the AVHRR OLR, consistent with the overestimate of intraseasonal variance of OLR over the tropical eastern hemisphere (Figs 6.13f and h).

Averaging data between  $5^{\circ}\text{N}$ - $5^{\circ}\text{S}$ , and plotting lagged regressions as a function of longitude, succinctly captures MJO propagation. The ECHO-G model represents well the eastward propagation of convection from the Indian Ocean into the central Pacific (Figs. 6.16a-b). To the east (west) of the convection the sea-surface temperature (SST) is above (below) normal (Figs. 6.16c-d). The above (below) normal SST occurs in the presence of easterly (westerly) wind anomalies that oppose (enhance) the climatological flow over the eastern hemisphere (Figs. 6.16e-f). The wind anomalies result in reduced evaporative cooling to the east of the convection, and enhanced evaporative cooling to the west (Fig. 6.17a). These characteristics are also present in the European Centre for Medium-Range Weather Forecasts 15-year reanalysis (Woolnough et al., 2000). The model does not capture the latent heat flux variations, particularly to the west of the convection (Fig. 6.17b), even though the wind anomalies are realistic (Fig. 6.16f). Other aspects of the surface energy balance need to be analyzed to understand the mechanism by which the model simulates the realistic SST anomalies (Fig. 6.16d). In observations the latent heat flux is the dominant term in the net surface heat flux during the MJO life-cycle (Sperber, 2003).

The eastward propagation of the MJO is associated with low-level moisture convergence. As seen in Figs. 6.17c and Figs. 6.17e, 1000-hPa convergence anomalies and enhanced moisture lead the convection, features also captured by the model (Figs. 6.17d and 6.17f). At time lag 0, this is manifested as a westward vertical tilt in the filtered divergence and specific humidity (Figs. 6.18a-d). As in Sperber (2003), the vertical profile of the zonal wind and vertical velocity indicates that free-tropospheric interactions also contribute to the life-cycle of the MJO (Fig. 6.18e). The dominant upward vertical velocity is strongest in the lower and upper troposphere, especially to the east of the center of the convection. Here, the upward motion and enhanced moisture help fuel the convection, while farther west the westerly anomalies and below-normal moisture erode the western limit of the convective complex. Figure 6.18f indicates that the vertical profile of the zonal wind from the model is asymmetric compared to the reanalysis, and Fig. 6.18d indicates that the enhanced moisture occurs higher up in the atmosphere than suggested by the reanalysis. Even so, the model readily captures the observed westward vertical tilt of divergence and moisture.

Another view of the developing conditions is given in Fig. 6.19 in which the divergence and specific humidity anomalies are shown as a function of time lag and pressure at  $\sim 125^{\circ}\text{E}$ , where the strongest convective anomalies occur at time lag 0 for PC-1 (Fig. 6.15a). At day -25 divergence anomalies and below-normal moisture predominate at this location. This is the inactive phase of the MJO during which convection is suppressed. At about day -20, convergence anomalies develop at 1000 hPa (Fig. 6.19a), and subsequently enhanced moisture occurs near the surface (Fig. 6.19c). With time the convergence anomalies deepen, and the moisture enhances further as the destabilization of the atmosphere intensifies. The inactive phase of the MJO develops in the same fashion, first being evident at the surface, and then dominating the atmospheric column. As seen in Figs. 6.19b and d, the model captures these features well, especially for the

active phase of the MJO. However, as mentioned earlier, the largest moisture enhancement occurs higher in the atmosphere than in the reanalysis. Additionally, the suppressed moisture at 1000 hPa does not lead the drying at altitude as the inactive phase of the MJO initiates.

The eastward propagation of convection from the additional CMIP2+ models is presented in Fig. 6.20. The ECHAM4\_OPYC model uses essentially the same atmospheric model as the ECHO-G simulation. The results from Fig. 6.20c and the vertical structure of the ECHAM4\_OPYC divergence and moisture (not shown) indicate that for the family of coupled models using the ECHAM4 atmospheric model, the representation of the MJO is robust independent of the ocean model used. This is further supported by results from the SINTEX model (Gualdi, 2003), which also uses ECHAM4, but with the OPA8.1 ocean model. CSIRO\_Mk2, GFDL\_R30\_c, and HadCM3 best represent the eastward propagation of convection from the Indian Ocean to Indonesia, though the amplitude of the convective signals varies substantially.

#### **6.4.4 MJO Convection: CMIP2+ vs. AMIPII**

Table 6.2 shows characteristics of the PCs that pertain to the propagation and amplitude of MJO convection. The standard deviations of the PCs from the NCEP/NCAR reanalysis OLR are weaker than from the AVHRR observations, as is the maximum positive correlation between the PCs. Additionally, not all years had a lead-lag structure comparable to the AVHRR data. Even so, the correlation of PC-1 (PC-2) between AVHRR OLR and NCEP/NCAR OLR is 0.89 (0.90), indicating strong agreement between their MJO OLR variations.

Four of seven CMIP2+ models have weaker PC variability than observed, while 5 of 7 have maximum positive correlations weaker than observed. The average time for the convection to transition from the Indian Ocean to the western Pacific varies by a factor of 2, and the models show a wide-ranging ability to represent the dominance of the eastward propagation.

Compared with their respective AMIPII counterparts integrated with observed monthly mean SST, the coupled versions all have larger maximum positive correlations. This indicates an improvement in the propagation of convection from the Indian Ocean to the western Pacific in the presence of air-sea interaction. Additionally, in coupled mode a greater fraction of years analyzed were dominated by eastward propagation. Waliser et al. (1999), Inness and Slingo (2003), Inness et al. (2003), and Sperber (2004) have all reported improvement in the MJO in coupled ocean-atmosphere simulations relative to their AMIP counterparts.

Table 6.2. Observed, reanalyzed, CMIP2+, and AMIP II (the last 4 rows) model characteristics of the Madden–Julian Oscillation for years when eastward propagation of convection dominates (see Fig. 6.14). Given are the standard deviations of PC-1 and PC-2, the maximum positive correlation of PC-1 relative to PC-2, R, the time lag (days) at which it occurred, and the fraction of years for which the PCs had a lead-lag relationship consistent with the observations. Shaded models used the same atmospheric component in their CMIP2+ and AMIP II simulations.

Model	PC-1	PC-2	R	Lag (days) PC-2 leads PC-1 (positive)	No. Years Eastward/Total
AVHRR	211.3	205.6	0.67	12	16/16
NCEP/NCAR	119.4	103.4	0.60	12	14/16
CCSM2.0	103.6	119.8	0.40	16	5/9
CSIRO_Mk2	113.6	165.7	0.19	16	13/19
ECHAM4_OPYC3	215.8	217.9	0.71	11	19/19
ECHO-G	293.8	267.1	0.68	12	16/19
GFDL_R30_c	221.1	198.9	0.18	10	16/19
HadCM3 (1_30)	105.5	99.8	0.51	8	11/19
PCM	109.1	91.9	0.12	15	10/15
CAM2.0	93.8	97.2	0.17	25	6/16
ECHAM4	221.2	232.2	0.43	12	11/16
HadAM3 (1.58)	125.4	99.0	0.42	13	9/16
CCM3	83.7	32.8	0.37	16	7/16

#### 6.4.5 Summary

The simulation of the MJO remains a critical test of a model’s ability to simulate the tropics. The majority of models underestimate the amplitude of the MJO convective anomalies by a factor of 2 or more, and the eastward propagation of convection is typically less coherent than observed. For a given family of atmospheric models, coupling to an ocean leads to better organization of the large-scale convection. The low-level moisture convergence mechanism for eastward propagation is represented in limited cases, as is the vertical structure of the MJO.

Additional regressions and examination of space–time spectra indicate (1) the models typically fail to represent the intraseasonal dominance of the large-scale circulation, (2) within a family of models ocean–atmosphere coupling leads to an improved lag/lead MJO structure, and (3) eastward propagation is limited by systematic error of the mean state. In particular, the unrealistic extension of low-level tropical easterlies west of the date line suppresses MJO convection, as per Inness and Slingo (2003) and Inness et al. (2003). Other variables are being analyzed to examine the mechanism of propagation in the models, and a more comprehensive peer-reviewed journal article is in preparation.

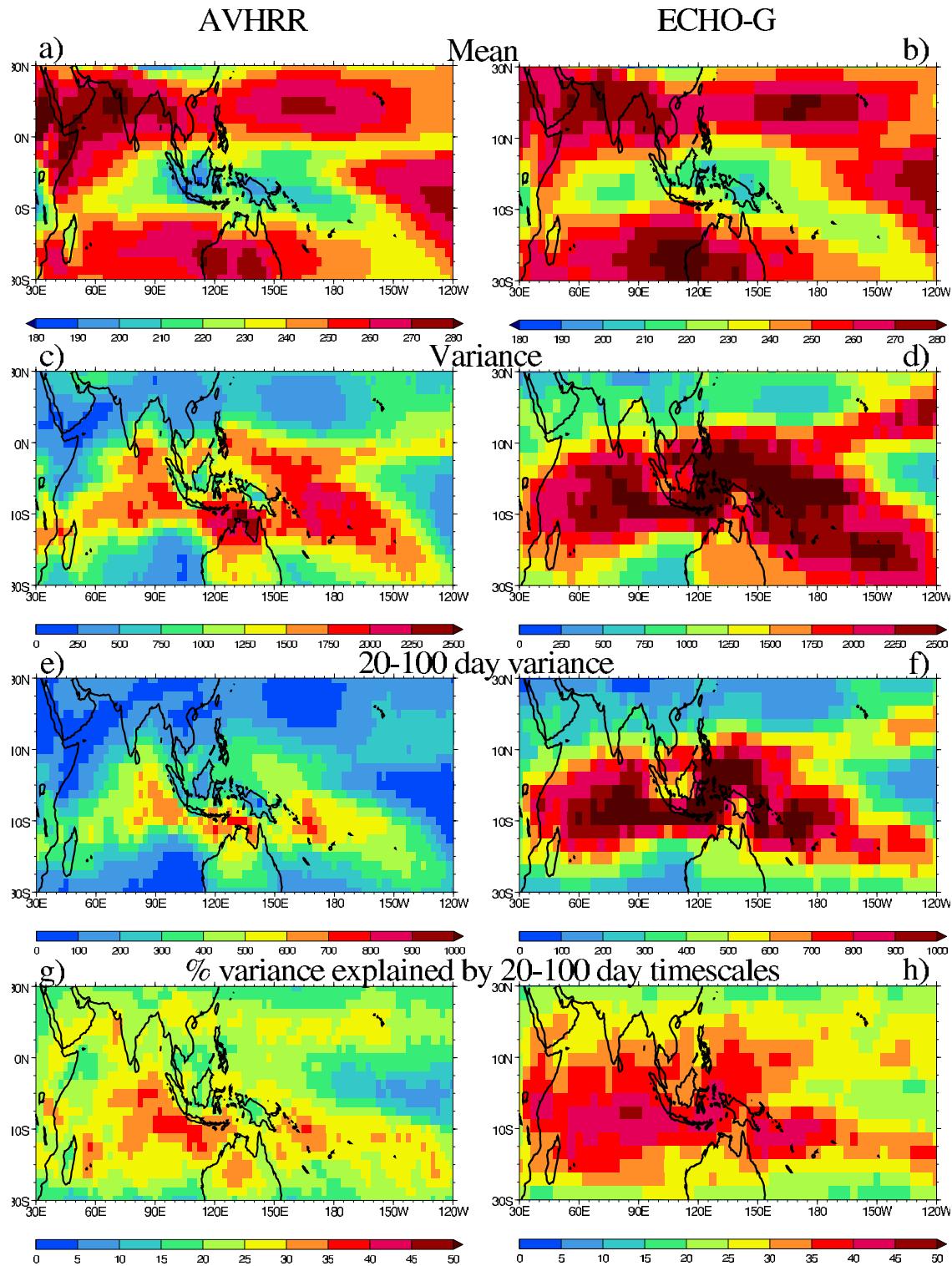


Fig. 6.14. November–March outgoing longwave radiation (OLR): climatology of (a) AVHRR and (b) ECHO-G; daily variance of (c) AVHRR, and (d) ECHO-G; 20–100-day bandpass-filtered variance (e) AVHRR, (f) ECHO-G, and percent of daily variance explained by the 20–100-day bandpass-filtered OLR (g) AVHRR and (h) ECHO-G.

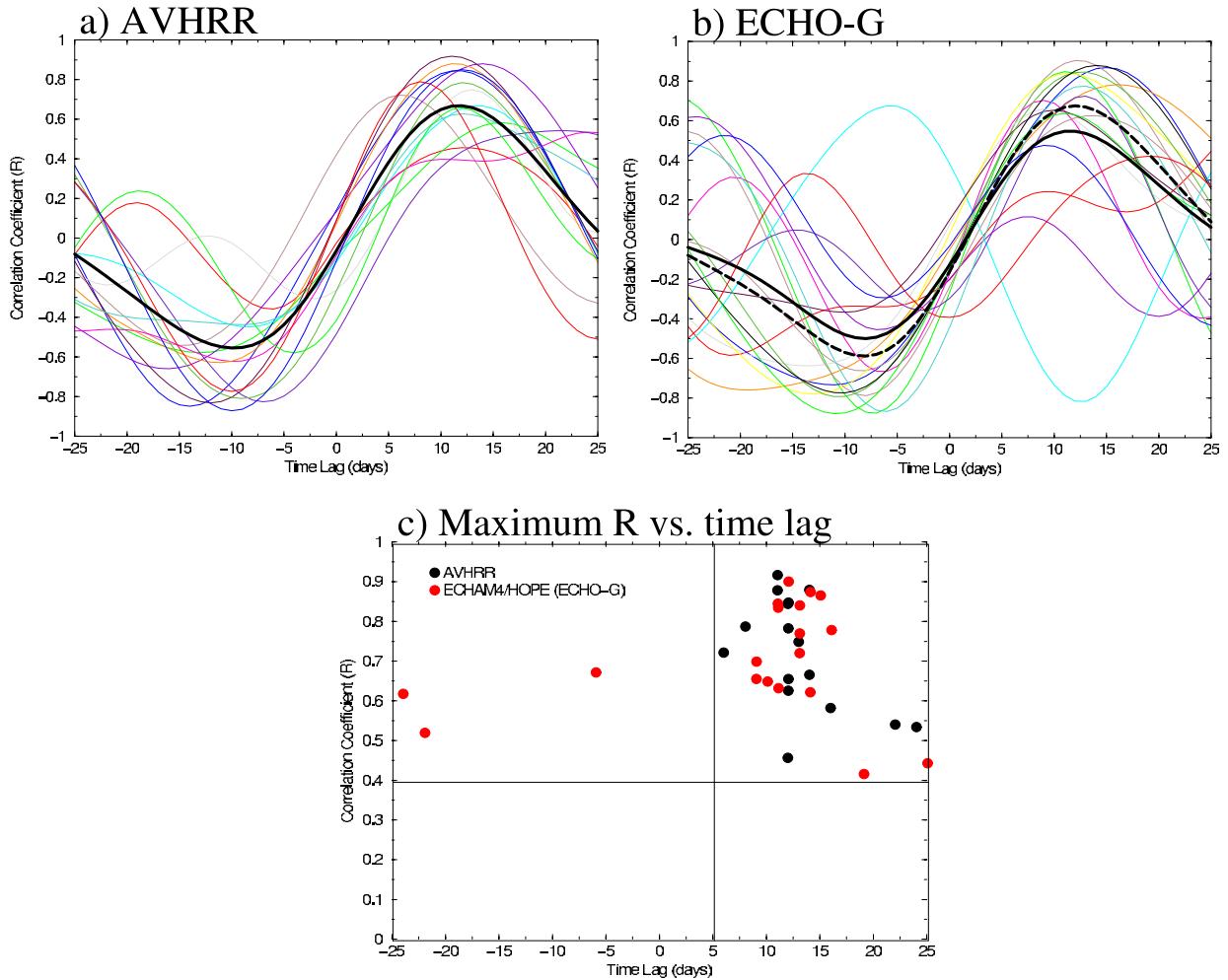


Fig. 6.15. Lead-lag correlation between PC-1 and PC-2 for each winter. Positive correlations at positive time lags indicate that convection over the Indian Ocean leads that over the western Pacific Ocean. The solid black curve is the average over all years of data. For the model, the dashed black curve is the average for years that lie in the observed phase-space (upper-right quadrant) of Fig. 6.14c. (a) AVHRR, (b) ECHO-G, (c) Phase-space of the maximum positive correlation and its associated time lag for each year of data.

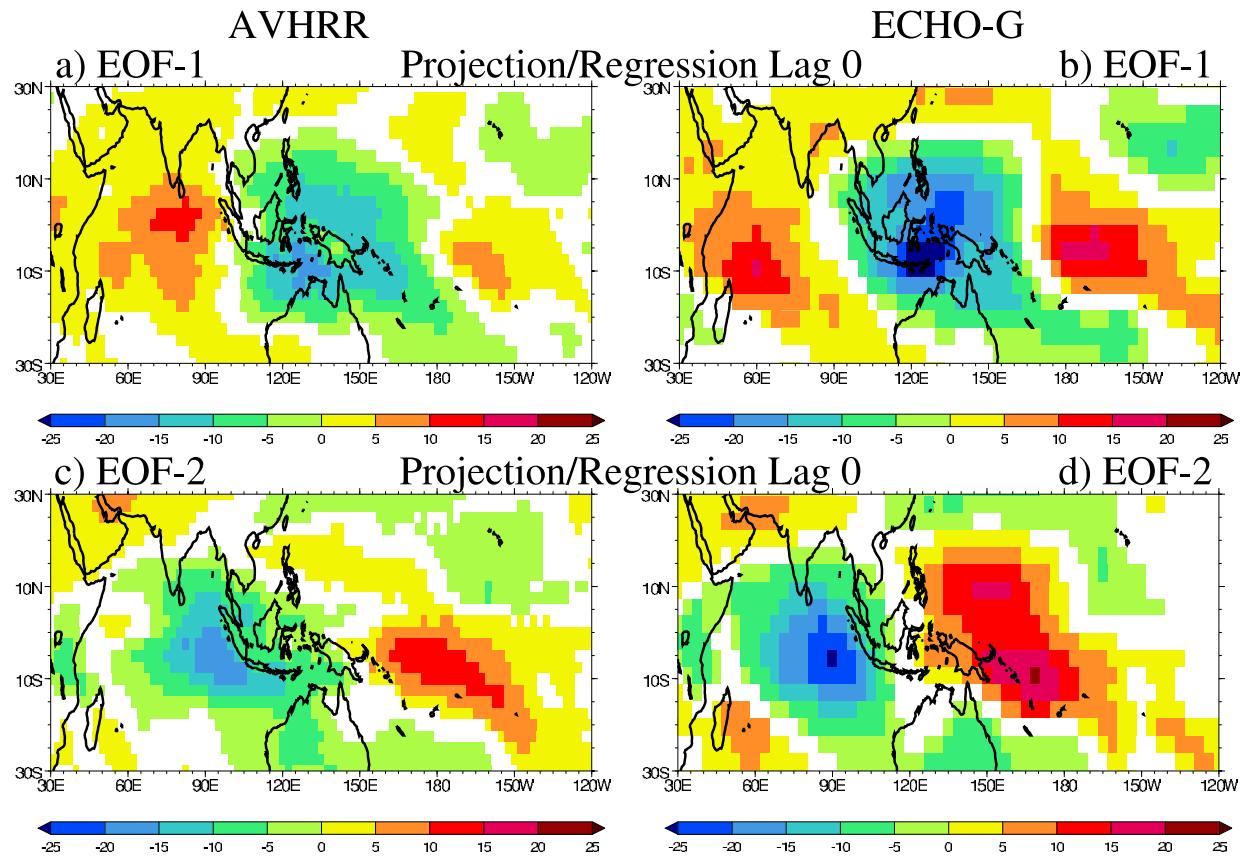
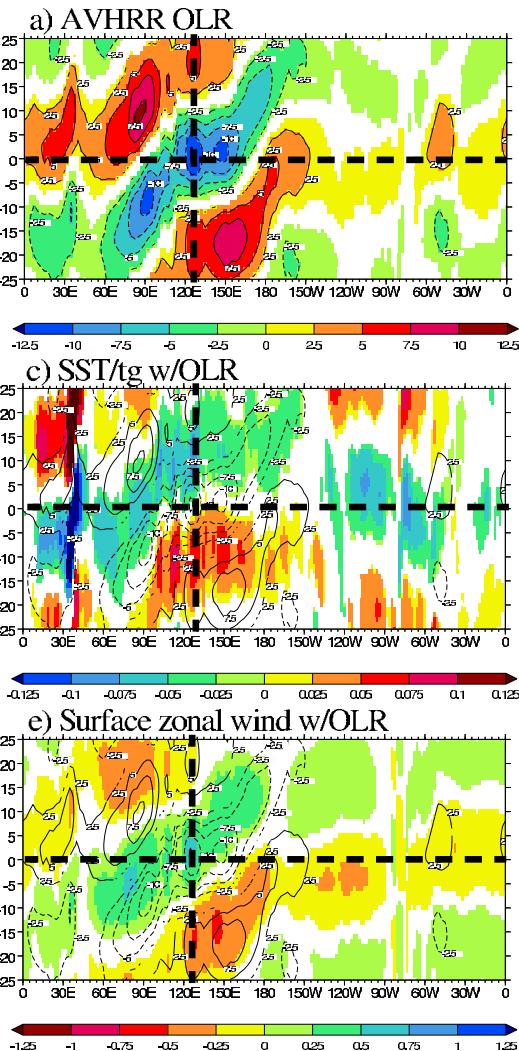


Fig. 6.16. Lag-0 linear regressions of PC-1 with 20–100-day filtered OLR ( $\text{Wm}^{-2}$ ): (a) AVHRR OLR, (b) ECHO-G. Panels (c) and (d) are as (a) and (b), but for regressions using PC-2. Data are plotted for a one-standard-deviation perturbation of the respective principal components where the fit is 5% significant or better, assuming each pentad is independent.

## Observations/NCEP/NCAR



## ECHO-G

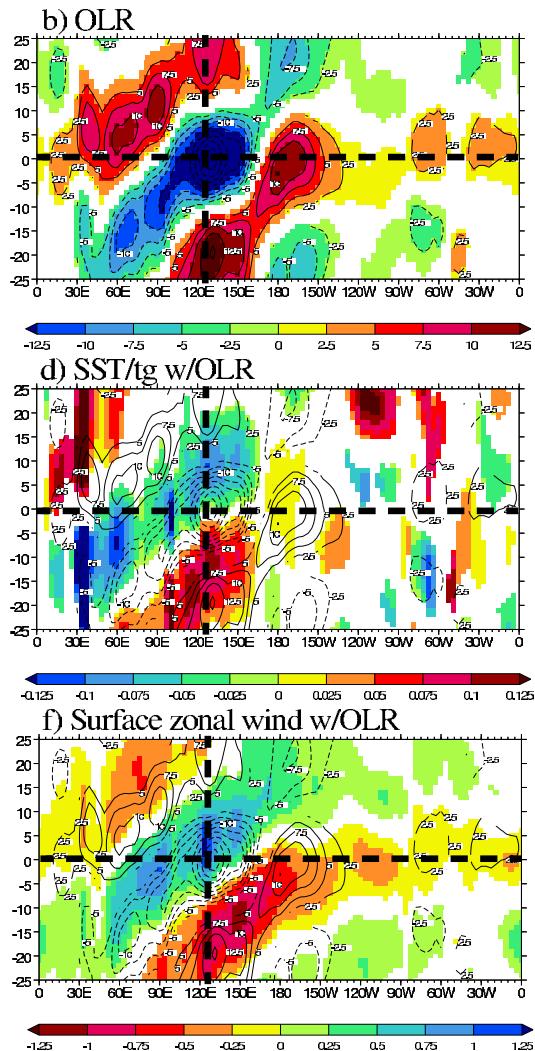
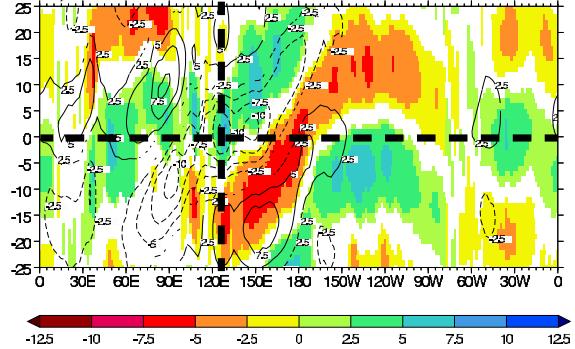


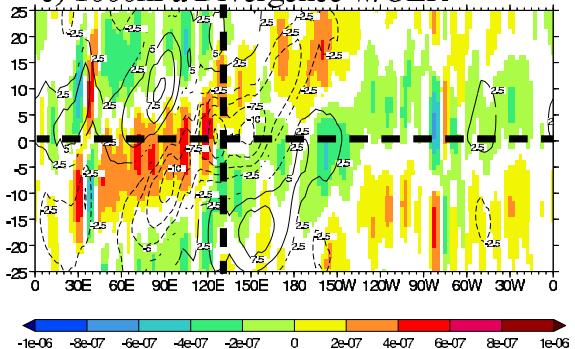
Fig. 6.17. Longitude–time lag plots of the linear regression of PC-1 with  $5^{\circ}\text{N}$ – $5^{\circ}\text{S}$  averaged 20–100-day bandpass-filtered (a) AVHRR OLR ( $\text{Wm}^{-2}$ ), (c) SST and ground temperature (K), and (e) 10-m zonal wind ( $\text{ms}^{-1}$ ). Contours of the OLR regression are plotted on each panel in increments of  $2.5 \text{ Wm}^{-2}$ . Panels (b), (d), and (f) are as (a), (c), and (e) but for ECHO-G. Time lags run from -25 to 25 days. The vertical dashed line gives the longitude of strongest convection in EOF-1 (Fig. 6.15a), and the horizontal dashed line corresponds to zero time lag. Data are plotted for a one-standard-deviation perturbation of the respective principal components where the fit is 5% significant or better assuming each pentad is independent. On each plot isolines of the OLR regression are plotted (negative values correspond to enhanced convection).

## Observations/NCEP/NCAR

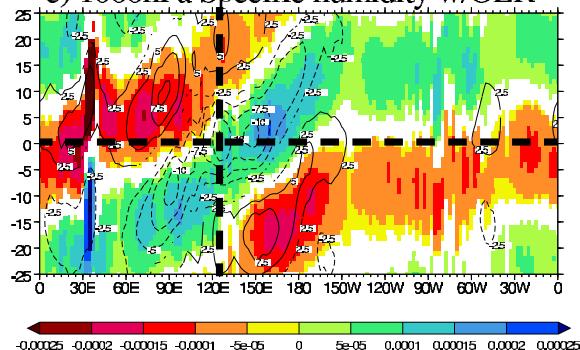
a) Latent heat flux w/OLR



c) 1000hPa Divergence w/OLR

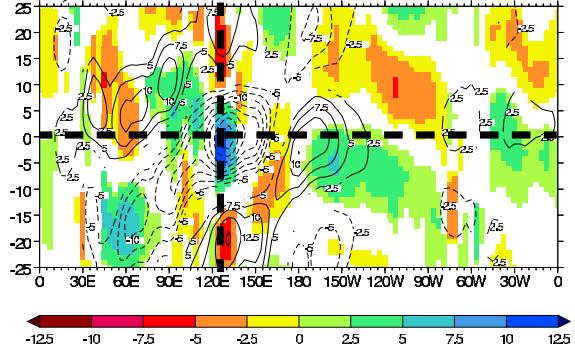


e) 1000hPa Specific humidity w/OLR

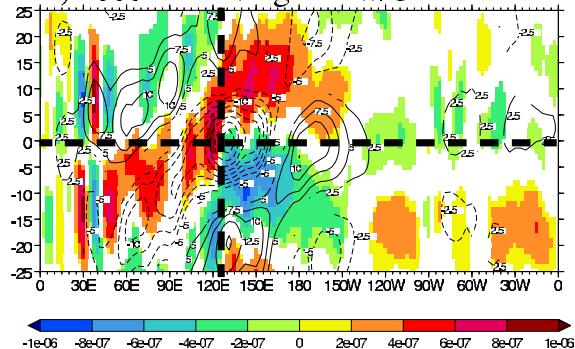


## ECHO-G

b) Latent heat flux w/OLR



d) 1000hPa Divergence w/OLR



f) 1000hPa Specific humidity w/OLR

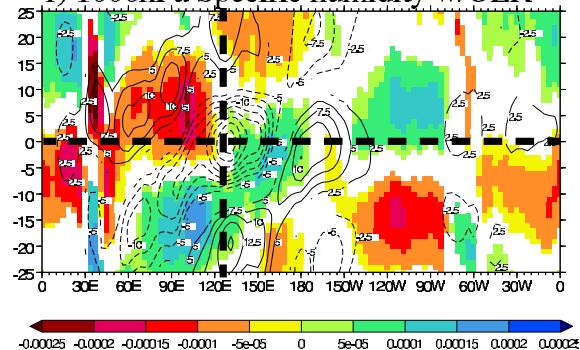


Fig. 6.18. As Fig. 6.16 but for (a-b) latent heat flux ( $\text{Wm}^{-2}$ ), (c-d) 1000-hPa divergence ( $\text{s}^{-1}$ ), and (e-f) 1000-hPa specific humidity ( $\text{kg kg}^{-1}$ ).

## NCEP/NCAR

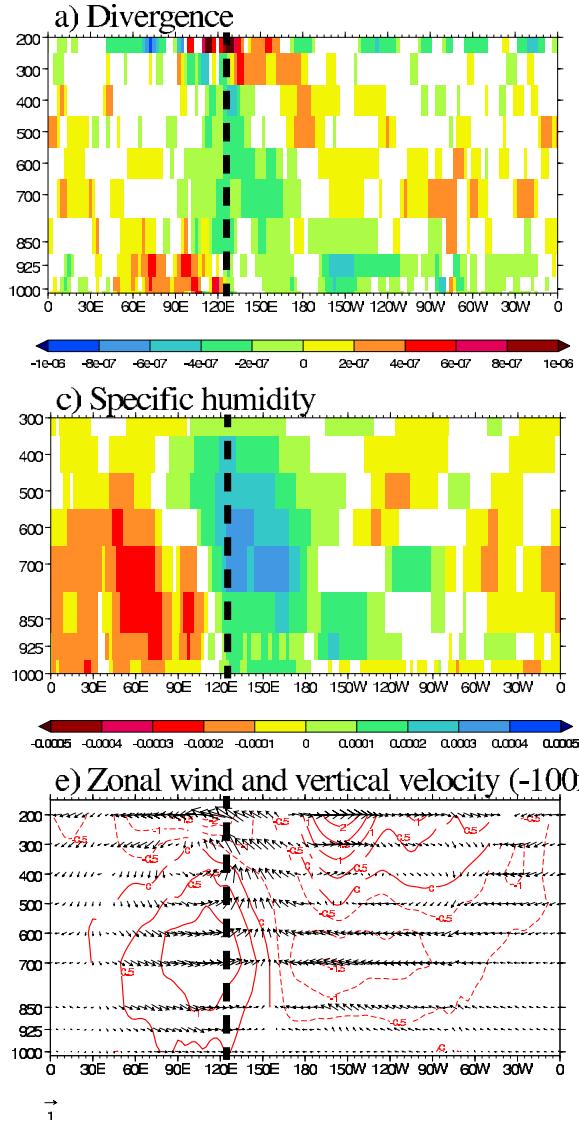


Fig. 6.19. Longitude–height cross sections of zero-time-lag linear regressions of PC-1 with  $5^{\circ}\text{N}$ – $5^{\circ}\text{S}$  averaged 20–100-day bandpass-filtered (a) divergence ( $\text{s}^{-1}$ ), (c) specific humidity ( $\text{kg kg}^{-1}$ ), and (e) zonal wind/vertical velocity vectors [note: the vertical velocity ( $\text{Pa s}^{-1}$ ) has been multiplied by  $-100$  to give scaling compatible with the u-wind ( $\text{ms}^{-1}$ )] and contours of the u-wind in increments of  $0.5 \text{ ms}^{-1}$ . Panels (b), (d), and (f) are as (a), (c), and (e), but for ECHO-G. Note: in (f) the vertical velocity was unavailable, so the vectors are omitted. The vertical dashed line at  $125^{\circ}\text{E}$  is the longitude of strongest convection in Fig. 6.15a. Wind vectors are plotted at every other grid point for clarity. Data are plotted for a one-standard-deviation perturbation of PC-1 where the fit is 5% significant or better, assuming each pentad is independent.

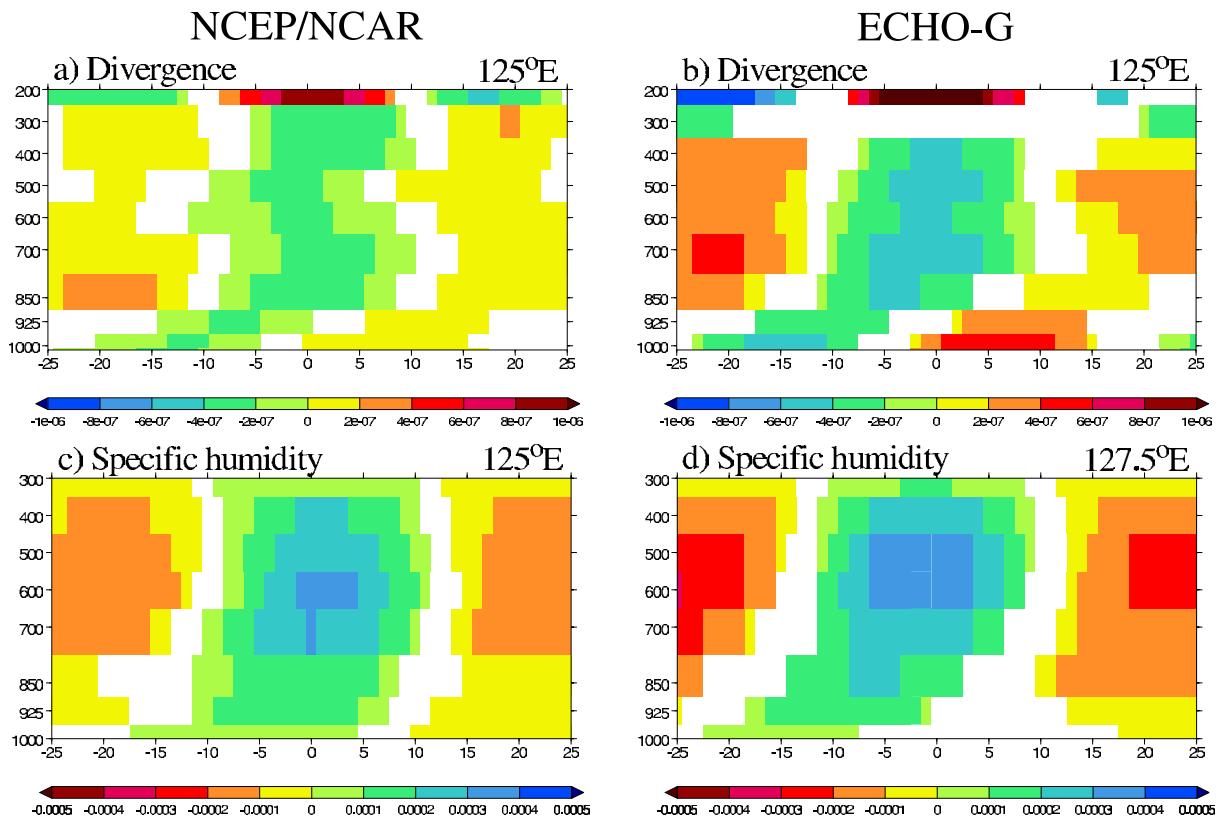


Fig. 6.20. Time-lag versus height plots of linear regressions of PC-1 at 125°E (5°N–5°S averaged) of 20–100-day bandpass-filtered (a) divergence (s<sup>-1</sup>), and (c) specific humidity (kg kg<sup>-1</sup>). Panels (b) and (d) are as (a) and (c), but for ECHO-G. Data are plotted for a one-standard-deviation perturbation of PC-1 where the fit is 5% significant or better assuming each pentad is independent.

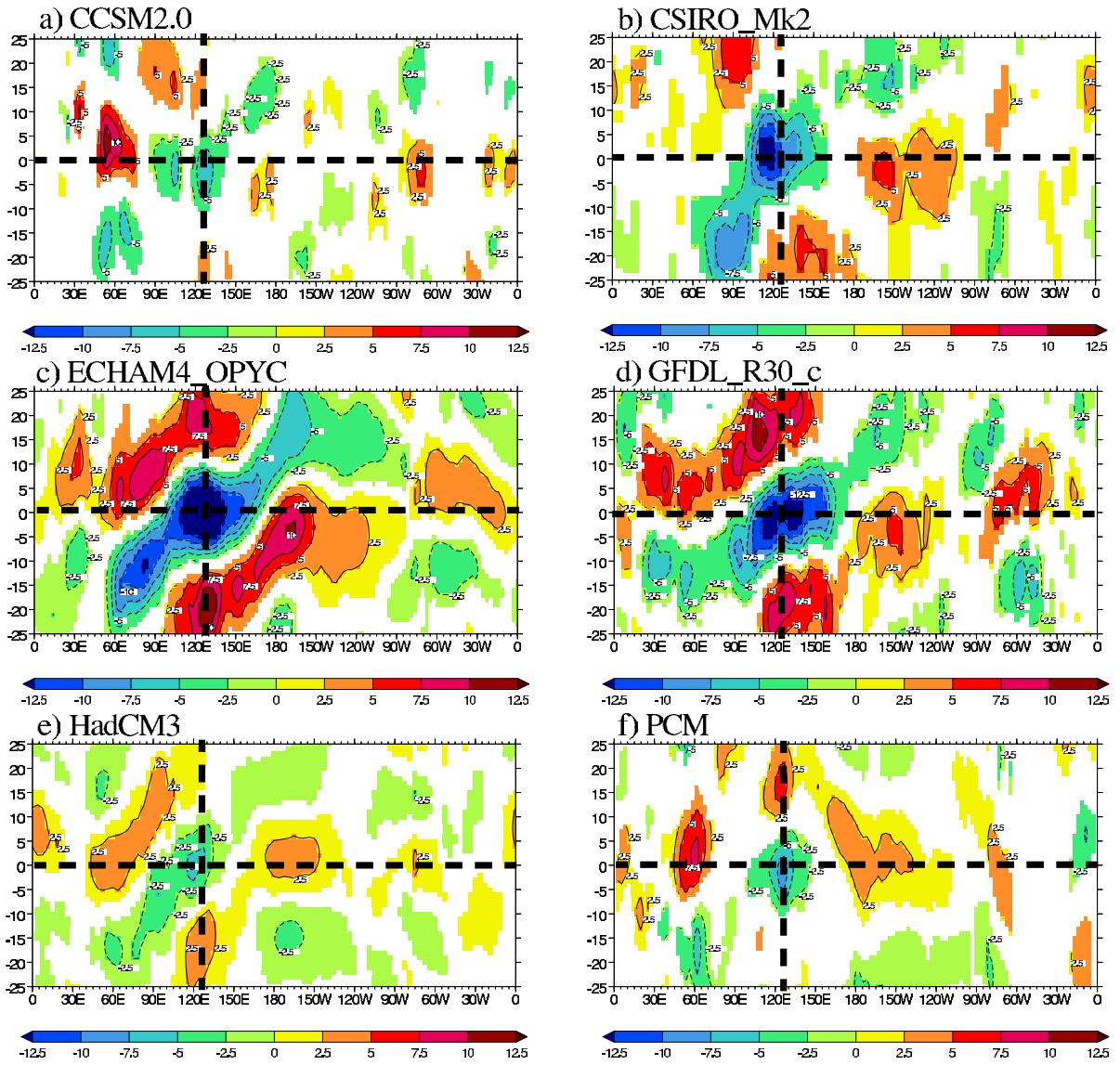


Fig. 6.21. As Fig. 6.16, but for longitude–time lag plots of the linear regression of PC-1 with  $5^{\circ}\text{N}$ – $5^{\circ}\text{S}$  averaged 20–100-day bandpass-filtered OLR ( $\text{Wm}^{-2}$ ) from the additional CMIP2+ models  
(a) CCSM2.0, (b) CSIRO\_Mk2, (c) ECHAM4\_OPYC, (d) GFDL\_R30\_c, (e) HadCM3, and  
(f) PCM.

## 7 Appraisal Summary

### Features of Appraisal Models (*Chapter 2*)

The climate simulations analyzed in this PCMDI appraisal of coupled ocean–atmosphere models are drawn from submissions to the second phase of the Coupled Model Intercomparison Project, designated as CMIP2+. The CMIP2+ comprises paired runs from each participating coupled model: 1) a control simulation of present-day climate and 2) a simulation of a perturbed climate resulting from increasing the concentration of atmospheric carbon dioxide (CO<sub>2</sub>) at a rate of 1 percent per year for a minimum period of 80 years. The emphasis of this appraisal is on the present-day control simulations, but the idealized increasing CO<sub>2</sub> experiments are examined in a few instances. There are also example comparisons of the coupled control runs with atmospheric-only runs from the Atmospheric Model Intercomparison Project (AMIP).

We have worked to synthesize key model and experimental features such as the resolution of the ocean and atmospheric component models, the spin-up procedure form of flux adjustment (if any), and the basic characteristics of the land and sea-ice models. A thorough documentation of any model is an arduous undertaking; describing a suite of models is even more challenging. We have provided a collection of references to facilitate efforts to acquire more in-depth understanding of particular model features. While this information is not comprehensive, we find that it is already proving to be valuable.

### Climate Trends in Unforced and Forced Simulations (*Chapter 3*)

We have examined the overall climate drift of the CMIP2+ control simulations by looking at global and hemispheric secular trends in surface air temperature, sea surface temperature, total sea ice extent, and the area-averaged ocean temperature and salinity (for both the global ocean and Arctic Basin as a function of depth). We also briefly examined differences between the CMIP2+ control and 1%/year CO<sub>2</sub> simulations for both surface air temperature and precipitation.

Results suggest that these models are relatively drift-free, including those without flux adjustment. One difficulty encountered in working to compare these simulations is the varied use of ocean spin-up procedures and the differing subset of each run provided to PCMDI, both of which make model-to-model comparisons problematic. However, our somewhat arbitrary choice of defining a climatology period was reasonable since in a global averaged sense the models appear to have reached quasi-equilibrium in the control simulations.

### Atmospheric Climatology and Annual Cycle (*Chapter 4*)

We have summarized some basic features of the simulated climate by analyzing 20-year climatologies of the atmosphere and ocean from each control run in the CMIP2+ database. The mean state defined with these climatologies can depend on differences in ocean-initialization, spin-up procedure and length of the integration. An additional caveat is that some models are flux-adjusted whereas others are not.

The simulated atmosphere is evaluated by comparing it to observationally based references of surface air temperature, precipitation, winds, outgoing longwave radiation, surface pressure and cloud cover—the subset of fields that were provided for most simulations. This collection of coupled models compares well with observations, illustrating strengths and weaknesses consistent with previous CMIP comparisons (e.g., Covey et al., 2003). Making use of statistical metrics developed at PCMDI (e.g., the Taylor Diagram), we find the multi-model ensemble mean compares better with observations than any individual model. With the same statistical metrics, we find these models compare well with atmospheric-only AMIP simulations. Using harmonic analysis, we have compared the annual cycle amplitude and phase of these models with available observations. Apart from several important biases, we find that most large-scale features of the annual cycle are captured by the multi-model ensemble mean. As expected, on larger regional scales there are more discrepancies.

Biases in precipitation, SST, and wind stress over the tropics have also been examined. All of the coupled models exhibit a “split” intertropical convergence zone (ITCZ) in the Pacific Ocean. This typically occurs in conjunction with a westward extension of the equatorial cold tongue and too-strong easterlies that extend into the western Pacific. Comparison with AMIP integrations indicates that the “split” ITCZ is also apparent, even when observed SST was prescribed as the surface boundary condition. This suggests that air-sea interaction may not be the root cause of the “split” ITCZ, but rather has exacerbated a pre-existing shortcoming in the atmospheric models. Systematic error in the coupled model is also present in the Atlantic basin, with the Southern Hemisphere anticyclone being too weak.

We have applied an extended version of the Köppen climate (KC) classification system to characterize the mean annual-cycle “climate” of 11 CMIP2+ models in geographical terms relevant to both the GCM modeling community and non-modelers. As a group, the models successfully simulated the five major KC classes (tropical, desert/steppe, temperate, snowy and polar) over ~80% of the land surface area compared with current-climate observations. However, the “mean” model (ensemble mean of the 11 models) showed the greatest accuracy, suggesting a compensation of errors among the models and that the mean model would be a good benchmark for measuring GCM improvement.

The KC scheme was also used to measure climate change by comparing the control (fixed CO<sub>2</sub>) and climate-change (1%/y CO<sub>2</sub> increase) experiments. Both the difference class, a unique output of the KC scheme, and shifts in area covered by the major KC classes showed a consistent response of the models to increasing CO<sub>2</sub>—warming with little change in precipitation. However, the net area of substantial warming varied among the models by a factor of 3.

### **The Mean State of the Ocean (*Chapter 5*)**

Our focus on the mean state of temperature and salinity in the major ocean basins represents a more in-depth examination of the deep ocean in these models than previous studies. Our qualitative comparison with the Levitus climatology is rather encouraging, although it must be pointed out that departures from the initial state in the deep ocean occur on time scales much longer than most of the simulations examined here. Several sample cross sections illustrate the currents at depth in the subtropical Atlantic (meridional), the tropical Pacific (zonal), and near the Fram Strait (meridional) and

Drakes Passage (zonal). The discrepancy among models and the Levitus climatolgies are most prevalent in the Arctic Basin.

### **Simulated Variability (*Chapter 6*)**

Select aspects of simulated climate variability have been presented in the appraisal. These include the El Nino/Southern Oscillation and the Madden–Julian Oscillation, the most important modes of tropical variability on interannual and intraseasonal time scales, respectively. Importantly, both of these phenomena force extratropical variability, and are therefore of global importance. To characterize the space–time variability of the tropics, frequency–wavenumber plots are presented. These show the ability of the models to represent equatorially trapped waves, originally derived from shallow water theory, which have recently been analyzed in detail in observations. From a hemispheric point of view, the North Atlantic Oscillation is associated with weather changes over portions of North America and Eurasia on a multitude of time scales.

With respect to the El Nino/Southern Oscillation, numerous intercomparison studies, with increasing complexity, have been prepared over the last decade. The appraisal most directly benefited from the study of AchutaRao and Sperber (2002), which provided a standard methodology for analysis, and importantly it served as a benchmark for several revised models in the appraisal. While difficulties remain in the simulation of ENSO (e.g., amplitude, seasonality, and periodicity), our results indicate that improvement in the simulation of ENSO has been realized through model development.

The Madden–Julian Oscillation is among the most challenging modes of variability to simulate. The realization that air–sea interaction is important in organizing the large-scale convection associated with the MJO is borne out through the comparison of AMIP and CMIP models herein, as is the detrimental influence of systematic error. These results confirm and extend previously reported findings from individual modelling centers, and suggest that for the MJO these sensitivities may be more broadly applicable. Findings also indicate that a reputable simulation of the MJO is presently possible, including its 3-dimensional space–time evolution.

Tropical wave spectra further confirm the difficulty the models have in representing tropical interactions. The models have difficulty in generating the observed level of variability as a function of space and time scale. This is particularly apparent for synoptic timescales (2–6 days) and higher-order wave numbers (5–15).

The North Atlantic Oscillation is well represented by the appraisal models. The spatial pattern of the large-scale surface air temperature response to the sea-level pressure perturbation over the Atlantic is akin to that observed, with the spatial error indicating the model response was not as strong as observed. This shortcoming was systematic across the models analyzed, and it is suggested that improvement in the extratropical flow will improve this aspect of the NAO response.

The variability analyzed here is a small, but important, subset of variations that models should strive to represent. In future appraisals the plan is to extend this section to include additional measures of climate variability in order to provide a more comprehensive suite of diagnostics for model validation.

## **Future Appraisals**

This appraisal focuses on the strengths and weaknesses of coupled ocean-atmospheric model control runs. Our priorities for future appraisals include expanding our suite of diagnostics, enhancements to our performance summary statistics, and an extension of the current emphasis on control runs to include simulations of various climate change scenarios. Importantly, the simulations used in this appraisal (CMIP2+) will provide a benchmark for the evaluation of future model versions.

## Acknowledgments

This work was supported by the U.S. Dept. of Energy, Office of Science, Climate Change Prediction Program at the University of California Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

The entire PCMDI staff deserves thanks for helping to make this report possible. We are especially grateful to the computer scientists at PCMDI for our core diagnostic capability (the Climate Data Analysis Tools, CDAT), which has revolutionized our efficiency, and thus enabled us to focus more on the scientific aspects of our work. Virtually all of the analysis in this report is based on CDAT. We are also thankful for the continued insight and support of PCMDI's founding director, W. Lawrence Gates.

The Naval Research Laboratory's Köppen Climate Classification application was kindly provided by Dr. Tim Hogan of NRL, Monterey. We gratefully acknowledge his assistance in implementing his well-designed code at PCMDI. Dr. M. Wheeler (BMRC) is thanked for his assistance in implementing the space-time spectral analysis.

Special thanks to Drs. T. Delworth (GFDL), G. Flato (CCCma), J. Gregory (Met Office), S. Legutke (M&D), G. A. Meehl (NCAR), S.-K. Min (University of Bonn), and R. Stouffer (GFDL) for providing helpful comments on the first draft of the appraisal. Insightful advice was provided by Dr. B. Semtner (Naval Postgraduate School) regarding the ocean and cryosphere chapter, and by Dr. D. Stephenson (University of Reading) regarding the North Atlantic Oscillation.

Most importantly, we are grateful to our colleagues who have graciously submitted their simulations for use in this report.



## References

- AchutaRao, K. and K. R. Sperber, 2002: Simulation of the El Niño southern oscillation: Results from the coupled model intercomparison project. *Climate Dyn.*, **19**, 191-209.
- Barnett, T. P., D. W. Pierce, and R. Schnur (2001): Detection of anthropogenic climate change in the world's oceans, *Science*, **292**, 270-274.
- Barnston, A. G., and R. E. Livezey (1987): Classification, seasonality and persistence of low frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083-1126.
- Blade, I. and D. L. Hartmann, 1993: Tropical intraseasonal oscillations in a simple nonlinear model. *J. Atmos. Sci.*, **50**, 2922-2939.
- Bleck, R., C. Rooth, D. Hu, and L.T. Smith, 1992: Salinity-driven thermocline transients in a wind- and thermohaline-forced isopycnic coordinate model of the North Atlantic. *J. Phys. Oceanogr.*, **22**, 1486-1505.
- Bonan, G.B., 1998: The land surface climatology of the NCAR land surface model (LSM 1.0) coupled to the NCAR Community Climate Model (CCM3). *J. Climate*, **11**, 1307-1326.
- Bonan, G.B., K.W. Oleson, M. Vertenstein, and S. Levis, 2002: the land surface climatology of the Community Land Model coupled to the NCAR Community Climate Model. *J. Climate*, **15**, 3123-3149.
- Branstetter, M.L., and D.J. Erickson III, 2003: Continental runoff dynamics in the Community Climate System Model (CCSM2) control simulation. *J. Geophys. Res.*, **108** (D17), 4550, doi:10.1029/2002JD003212.
- Briegleb, B.P., C.M. Bitz, E.C. Hunke, W.H. Lipscomb, and J.L. Schramm, 2002: Description of the Community Climate System Model Version 2 Sea Ice Model. Accessible online at [http://www.cgd.ucar.edu/csm/models/ccsm2.0/csims/SciGuide/ice\\_scidoc.pdf](http://www.cgd.ucar.edu/csm/models/ccsm2.0/csims/SciGuide/ice_scidoc.pdf).
- Bryan, F. O., 1998: Climate drift in a multicentury integration of the NCAR Climate System Model, *J. Climate*, **11**, 1455-1471.
- Bryan, K., 1969: Climate and the ocean circulation. Part III: The ocean model. *Mon. Wea. Rev.*, **97**, 806-827.
- Carton, J.A., G. Chepurin, X. Cao, and B.S. Giese, 2000a: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950-1995, Part 1: methodology, *J. Phys. Oceanogr.*, **30**, 294-309.
- Carton, J.A., G. Chepurin, and X. Cao, 2000b: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950-1995 Part 2: results, *J. Phys. Oceanogr.*, **30**, 310-326.

Cattle, H., and J. Crossley, 1995: Modelling Arctic climate change. *Philos. Trans. R. Soc. London A***352**, 201-213.

Collins, W.D., J.J. Hack, B.A. Boville, P.J. Rasch, D.L. Williamson, J.T. Kiehl, B. Briegleb, and J.R. McCaa, 2003: Description of the NCAR Community Atmosphere Model (CAM2). Accessible online at <http://www.ccsm.ucar.edu/models/atm-cam/index.html>.

Covey, C., K.M. AchutaRao, U. Cubasch, P. Jones, S.J. Lambert, M.E. Mann, T.J. Phillips and K.E. Taylor, 2003: An overview of results from the Coupled Model Intercomparison Project, *Glob. and Planet. Change*, **37**, 103-133.

Cox, M.D., 1984: A primitive equation, three-dimensional model of the ocean. GFDL Ocean Group Technical Report No. 1, Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, 141 pp.

Cox, P.M., R.A. Betts, C.B. Bunton, R.L.H. Essery, P.R. Rowntree, and J. Smith, 1999: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. *Climate Dyn.*, **15**, 183-203.

Cullen, M.J.P., 1993: The unified forecast/climate model. *Meteor. Mag.*, **122**, 81-94.

Cubasch, U., et al., 2001: Projections of future climate change. In *Climate Change 2001: The Scientific Basis*, J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.), Cambridge University Press, pp. 525-582.

Delworth, T.L., R.S. Stouffer, K.W. Dixon, M.J. Spelman, T.R. Knutson, A.J. Broccoli, P.J. Kushner, and R.T. Wetherald, 2002: Simulation of climate variability and change by the GFDL R30 coupled model. *Climate Dyn.*, **19**, 555-574.

Deque, M., C. Dreveton, A. Braun, and D. Cariolle, 1994: The ARPEGE/IFS atmosphere model: A contribution to the French community climate modeling. *Climate Dyn.*, **10**, 249-266.

Douville, H., J.-F. Royer, and J.-F. Mahfouf, 1995: A new snow parameterization for the Meteo-France climate model. *Climate Dyn.*, **12**, 21-35.

Drange, H. and K. Simonsen, 1996: Formulation of air-sea fluxes in the ESOP2 version of MICOM. Nansen Environmental and Remote Sensing Center (NERSC) Report 125. Available from Edv. Griegsv. 3A, N-5037 Solheimsviken, Norway.

Ferranti, L., T. N. Palmer, F. Molteni, and E. Klinker, 1990: Tropical-extratropical interaction associated with the 30-60 day oscillation and its impact on medium and extended range prediction. *J. Atmos. Sci.*, **47**, 2177-2199.

Flato, G.M., and G.J. Boer, 2001: Warming asymmetry in climate change simulations. *Geophys. Res. Letters*, **28**, 195-198.

Flato, G.M., G.J. Boer, W.G. Lee, N.A. McFarlane, D. Ramsden, M.C. Reader, and A.J. Weaver, 2000: The Canadian Centre for Climate Modelling and Analysis global coupled model and its climate. *Climate Dyn.*, **16**, 451-467.

Flato, G.M., and W.D. Hibler, 1990: On a simple sea-ice dynamics model for climate studies. *Ann. Glaciol.*, **14**, 72-77.

Furevik, T., B. M. Bentsen, H. Drange, I.K.T. Kindem, N.G. Kvamstø, and A. Sorteberg, 2003: Description and validation of the Bergen Climate Model: ARPEGE coupled with MICOM. *Climate Dyn.*, **21**, 27-51.

Gates, W.L., U. Cubasch, G.A. Meehl, J.F.B Mitchell and R.J. Stouffer, 1993: An intercomparison of selected features of the control climates simulated by coupled ocean-atmosphere general circulation models. *World Climate Research Programme WCRP-82*, WMO/TD No. 574, World Meteorological Organization, Geneva, 46pp.

Gates, W. L., and Coauthors, 1996: Climate models—evaluation. In *Climate Climate 1995: The Science of Climate Change*, J. T. Houghton et al. (eds.), Cambridge University Press, 229-284.

Gates, W.L., and Coauthoers, 1999: An overivew of results from the Atmospheric Model Intercomparison Project (AMIP1). *Bull. Amer. Meteorol. Soc.*, **80**, 29-55.

Geiger, R. and W. Pohl, 1954: Eine neue Wandkarte der Klimagebiete der Erde nach W. Köppen Klassifikation. *Erdknunde VIII*, 58-60.

Gordon, C., C. Cooper, C.A. Senior, H.T. Banks, J.M. Gregory, T.C. Johns, J.F.B. Mitchell, and R.A. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.*, **16**, 147-168.

Gordon, H.B., and S.P. O'Farrell, 1997: Transient climate change in the CSIRO coupled model with dynamic sea ice. *Mon. Wea. Rev.*, **125**, 875-907.

Gregory, D., and R.N.B. Smith, 1990: Canopy, surface, and soil hydrology. Unified Model Documentation Paper No. 25, United Kingdom Meteorological Office, Bracknell, Berkshire RG12 2SZ, UK.

Gualdi, S., 2003: Simulation of the tropical intraseasonal oscillation with a coupled GCM. Presented at “The ECMWF/CLIVAR Workshop on Simulation and Prediction of Intra-seasonal Variability with Emphasis on the MJO”, 3-6 November 2003, ECMWF, Reading, UK. Accessible online at [http://www.ecmwf.int/newsevents/meetings/workshops/Intra-seasonal\\_variability/presentations/gualdi.pdf](http://www.ecmwf.int/newsevents/meetings/workshops/Intra-seasonal_variability/presentations/gualdi.pdf).

Guilyardi, E, S. Gualdi, J. Slingo, A. Navarra, P. Delecluse, J. Cole, G. Madec, M. Roberts, M. Latif, & L. Terray, 2003: Representing El Niño in coupled ocean-atmosphere GCMs: the dominant role of the atmosphere. *J. Climate* (submitted).

Hayashi, Y. and D. G. Golder, 1993: Tropical 40-50 and 25-30 day oscillations appearing in realistic and idealized climate models and the ECMWF dataset. *J. Atmos. Sci.*, **50**, 464-494.

Hayashi, Y. and D. G. Golder, 1986: Tropical intraseasonal oscillations appearing in a GFDL general circulation model and FGGE data. Part I: phase propagation. *J. Atmos. Sci.*, **43**, 3058-3067.

Hayashi, Y. and D. G. Golder, 1988: Tropical intraseasonal oscillations appearing in a GFDL general circulation model and FGGE data. Part II: structure. *J. Atmos. Sci.*, **45**, 3017-3033.

Hendon, H. H., and M. L. Salby, 1994: The life cycle of the Madden-Julian Oscillation. *J. Atmos. Sci.*, **51**, 2225-2237.

Hewitt, C.D., and J.F.B. Mitchell, 1996: GCM simulations of the climate of 6 kyr BP: Mean changes and interdecadal variability. *J. Climate*, **9**, 3505-3529.

Hibler, W.D., 1979: A dynamic-thermodynamic sea ice model. *J. Phys. Oceanogr.*, **9**, 817-846.

Hirst, A.C., S.P. O'Farrell, and H.B. Gordon, 2000: Comparison of a coupled ocean-atmosphere model with and without oceanic eddy-induced advection. 1. Ocean spin-up and control integrations. *J. Climate*, **13**, 139-163.

Hogan, T.F., 1995: Identifying sensitivities in GCM parameterizations using the Köppen climate classification. Proceedings of the First International AMIP Scientific Conference, Monterey, California, USA 15-19 May 1995. World Meteorological Organization Report WMO/TD-92; WCRP-92. 532 p.

Hsu, C.F., and J.M. Wallace, 1976: The global distribution of the annual and semiannual cycles in precipitation. *Mon. Wea. Rev.*, **104**, 1093-1101

Hsu, H., B. J. Hoskins, and F. Jin, 1990: The 1985/86 intraseasonal oscillation and the role of the extratropics. *J. Atmos. Sci.*, **47**, 823-839.

Huffman, G. J., R. F. Adler, B. Rudolf, U. Schneider, and P. Keehn, 1995: Global precipitation estimates based on a technique for combining satellite based estimates, rain gauge analysis, and NWP model precipitation information. *J. Climate*, **8**, 1284-1295.

Huffman, G. J., and Coauthors, 1997: The Global precipitation climatology project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5-20.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2003: An overview of the North Atlantic Oscillation. In *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, Ed. J. W. Hurrell, Y. Kushnir, G. Ottersen, and M. Visbeck, *Geophysical Monograph Series*, **134**, 279 pp.

- Inness, P. M., and J. M. Slingo, 2003: Simulation of the Madden-Julian Oscillation in a coupled general circulation model. Part I: Comparison with observations and an atmosphere-only GCM. *J. Climate*, **16**, 345-364.
- Inness, P. M., J. M. Slingo, E. Guilyardi, and J. Cole, 2003: Simulation of the Madden-Julian Oscillation in a coupled general circulation model. Part II: The role of the basic state. *J. Climate*, **16**, 365-382.
- IPCC, 2001: *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881 pp.
- Jenkins, G. M., and D. G. Watts, 1968: *Spectral Analysis and its Applications*, Holden-Day, pp. 310-311.
- Johns, T.C., R.E. Carnell, J.F. Crossley, J.M. Gregory, J.F.B. Mitchell, C.A. Senior, S.F.B. Tett, and R.A. Wood, 1997: The second Hadley Centre coupled ocean-atmosphere GCM: Model description, spinup and validation. *Climate Dyn.*, **13**, 103-134.
- Jones, C. and B. C. Weare, 1996: The role of low-level moisture convergence and ocean latent heat fluxes in the Madden and Julian Oscillation. *J. Climate*, **9**, 3086-3140.
- Kalnay, E. M. and co-authors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.
- Kemball-Cook, S. R., and B. C. Weare, 2001: The onset of convection in the Madden-Julian Oscillation. *J. Climate*, **14**, 780-793.
- Kiehl, J. T., and P. R. Gent, 2004: The Community Climate System Model, Version Two, *J. Climate* (submitted).
- Kiehl, J.T., J.J. Hack, G. Bonan, B.A. Boville, D. Williamson, and P. Rasch, 1998: The National Center for Atmospheric Research Community Climate Model: CCM3. *J. Climate*, **11**, 1131-1149.
- Kim, S.-J., G.M. Flato, G.J. Boer, and N.A. McFarlane, 2002: A coupled climate model simulation of the Last Glacial Maximum, Part 1: Transient multi-decadal response. *Climate Dyn.*, **19**, 515-537.
- Kirkyla, K.I., and S. Hameed, 1989: Harmonic analysis of the seasonal cycle in precipitation over the United States: A comparison between observations and a general circulation model. *J. Climate*, **2**, 1463-1475.
- Klein, S. A. and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Monthly Weather Review*, **127**, 2514-2531.
- Köppen, W., 1923: *Die Klimate der Erde*. Wlater de Gruyter, Berlin, 369 pp.

Kowalczyk, E.A., J.R. Garratt, and P.B. Krummel, 1991: A soil-canopy scheme for use in a numerical model of the atmosphere—1 D stand-alone model. CSIRO Division of Atmospheric Research Technical Paper No. 23, 56 pp.

Kowalczyk, E.A., J.R. Garratt, and P.B. Krummel, 1994: Implementation of a soil-canopy scheme into the CSIRO GCM-Regional aspects of the model response. CSIRO Division of Atmospheric Research Technical Paper No. 32, 59 pp.

Legates, D. R. and C. J. Willmott, 1990a: Mean seasonal and spatial variability of global surface air temperature. *Theor. and Appl. Climatol.*, **41**, 11-21.

Legates, D. R. and C. J. Willmott, 1990b: Mean seasonal and spatial variability in gauge-corrected global precipitation. *Int. J. Climatol.*, **10**, 111-127.

Legutke, S. and Maier-Reimer, 1999: Climatology of the HOPE-G Global Ocean General Circulation Model. Technical Report, No. 21, German Climate Computer Centre (DKRZ), Hamburg, 90 pp. Also online at <http://www.mad.zmaw.de/Pingo/reports/ReportNo.21.pdf>.

Legutke, S. and R. Voss, 1999: The Hamburg Atmosphere-Ocean Coupled Circulation Model. DKRZ Technical Report No. 18, Deutsches KlimaRechenZentrum, Hamburg, Germany, 61 pp. Also online at <http://mad.zmaw.de/Pingo/reports/ReportNo.18.pdf>

Levitus, S., 1982: Climatological atlas of the world's oceans. NOAA Professional Paper 13, 173 pp.

Levitus, S., 1998: Data set available online at  
<http://www.nodc.noaa.gov/OC5/docwoa98.html>.

Levitus, S. and T.P. Boyer, 1994: World ocean atlas 1994, Vol. 4: Temperature. NOAA/NESDIS E/OC21, U.S. Department of Commerce, Washington, D.C., 117 pp.

Levitus, S., R. Burgett, and T.P. Boyer, 1995: World ocean atlas 1994, Vol. 3: Salinity. NOAA/NESDIS E/OC21, U.S. Department of Commerce, Washington, D.C., 99 pp.

Liebmann, B. and C. A. Smith, 1996: Description of a complete (interpolated) OLR dataset. *Bull. Amer. Meteor. Soc.*, **77**, 1275-1277.

Lindzen, R. S., 1967: Planetary waves on beta planes. *Mon. Wea. Rev.*, **95**, 441-451.

Lohmann, U., R. Sausen, L. Bengtsson, U. Cubasch, J. Perlitz, E. Roeckner, 1992: The Köppen climate classification as a diagnostic tool for general circulation models. Report No. 92, Max-Plank-Institut für Meteorologie, Hamburg, Germany, 22 pp.

Madden, R. A. and P. R. Julian, 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702-708.

Madden, R. A. and P. R. Julian, 1972: Description of global-scale circulation cells in the tropics with a 40-50 day period. *J. Atmos. Sci.*, **29**, 1109-1123.

- Mahfouf, J.-F., A.O. Manzi, J. Noilhan, H. Giordani, and M. Deque, 1995: The land surface scheme ISBA within the Meteo-France climate model ARPEGGE. Part 1: Implementation and preliminary results. *J. Climate*, **8**, 2039-2057.
- Maloney, E. D. and D. L. Hartmann, 1998: Frictional moisture convergence in a composite life cycle of the Madden-Julian Oscillation. *J. Climate*, **11**, 2387-2403.
- Maltrud, M.E., R.D. Smith, A.J. Semtner, and R.C. Malone, 1998: Global eddy-resolving ocean simulations driven by 1985-1995 atmospheric winds. *J. Geophys. Res.*, **103**, 30,825-30,853.
- Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteorol. Soc. Japan*, **44**, 25-43.
- McAvaney, B. J., and Coauthors, 2001: Model Evaluation. In *Climate Change 2001: The Scientific Basis*, J. T. Houghton, et al., Eds., Cambridge University Press, 471-521.
- McFarlane, N.A., G.J. Boer, J.-P. Blanchet, and M. Lazare, 1992: The Canadian Climate Centre second-generation general circulation model and its equilibrium climate. *J. Climate*, **5**, 1013-1044.
- McGregor, J.L., H.B. Gordon, I.G. Watterson, M.R. Dix, and L.D. Rotstayn, 1993: The CSIRO 9-level atmospheric general circulation model. CSIRO Divsion of Atmospheric Research Technical Paper No. 26, CSIRO, Mordialloc, Australia, 89 pp.
- Mellor, G.L. and L. Kantha, 1989: An ice-ocean coupled model. *J. Geophys. Res.*, **94**, 10,937-10,954.
- Min, S-K., S. Legutke, A. Hense, and W-T. Kwon, 2004: Climatology and internal variability in a 1000-year control simulation with the coupled climate model ECHO-G. M&D Technical Report No. 2, Modelle & Daten, Hamburg, Germany, 67 pp. Accessible online at [http://mad.zmaw.de/Pingo/reports/TeReport\\_Web02.pdf](http://mad.zmaw.de/Pingo/reports/TeReport_Web02.pdf). Also submitted to *Climate Dynamics*.
- Milly, P.C.D., 1992: Potential evaporation and soil moisture in general circulation models. *J. Climate*, **5**, 209-226.
- Mo, K. C. and R. W. Higgins, 1998: Tropical influences on California precipitation. *J. Climate*, **11**, 412-430.
- Murakami, T., 1988: Intraseasonal atmospheric teleconnection patterns during Northern Hemisphere winter. *J. Climate*, **1**, 117-131.
- Noda, A., S. Yukimoto, S. Maeda, T. Uchiyama, K. Shibata, and S. Yamaki, 2001: A new Meteorological Research Institute coupled GCM (MRI-CGCM2)—Transient response to greenhouse gas and aerosol scenarios. CGER's Supercomputer Monograph Report Vol. 7, Center for Global Environmental Research, National Institute for Environmental Studies, Ibaraki 305-0053, Japan, 63 pp.

- Oberhuber, J. M., 1993: Simulation of the Atlantic circulation with a coupled sea-ice-mixed layer-isopycnical general circulation model. Part I: model description. *J. Phys. Oceanogr.*, **23**, 808-829.
- O'Farrell, S.P., 1998: Investigation of the dynamic sea ice component of a coupled atmosphere sea-ice general circulation model. *J. Geophys. Res. (Oceans)*, **103**, 15751-15782.
- Pacanowski, R.C., K. Dixon, and A. Rosati, 1993: The GFDL modular ocean model users guide, version 1.0. GFDL Ocean Group Technical Report No. 2, Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey.
- Park, C.-K., D. M. Straus, and K.-M. Lau, 1990: An evaluation of the structure of the tropical intraseasonal oscillations in three general circulation models. *J. Meteor. Soc. Japan*, **68**, 403-417.
- Pope, V.D., M.L. Gallani, P.R. Rowntree, and R.A. Stratton, 2000: The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dyn.*, **16**, 123-146.
- Rayner, N. A., D. E. Parker, P. Frisch, E. B. Horton, C. K. Folland, and L. V. Alexander, 2000: SST and sea-ice fields for ERA40. In Proceedings of the Second WCRP International Conference on Reanalyses, Wokefield Park, Reading, UK, 23-27 August 1999, WCRP-109, WMO/TD-985.
- Rayner, N.A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: "Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century". *Journal of Geophysical Research*, Vol. 108, No. D14, 4407, doi:10.1029/2002JD002670, 2003
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimal interpolation. *J. Climate*, **7**, 929-948.
- Roeckner, E., J.M. Oberhuber, A Bacher, M. Christoph, and I. Kirchner, 1996a: ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM. *Climate Dyn.*, **12**, 737-754.
- Roeckner, E., K. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schlese, and U. Schulzweida, 1996b: The atmospheric general circulation model ECHAM4: Model description and simulation of present-day climate. MPI Report No. 218, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 90 pp.
- Rui, H. and B. Wang, 1990: Development characteristics and dynamic structure of tropical intraseasonal convective anomalies. *J. Atmos. Sci.*, **47**, 357-379.
- Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, 2000: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J. Geophys. Res.*, **105(D6)**, 7337-7356.

Sato, N., P.J. Sellers, D.A. Randall, E.K. Schneider, J. Shukla, J.L. Kinter III, Y-T. Hou, and E. Albertazzi, 1989: Implementing the simple biosphere model in a general circulation model. *J. Atmos. Sci.*, **46**, 2757-2782.

Sellers, P.J., Y. Mintz, Y.C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.*, **43**, 505-531.

Simmons, A.J. and J.K. Gibson, 2000: The ERA-40 project plan. ERA-40 Project Report Series No. 1, European Centre for Medium-Range Weather Forecasts, Reading, UK, 62 pp. Accessible online at [http://www.ecmwf.int/publications/library/ecpublications/pdf/ERA40\\_PRS\\_1.pdf](http://www.ecmwf.int/publications/library/ecpublications/pdf/ERA40_PRS_1.pdf).

Slingo, J. M., and R. A. Madden, 1991: Characteristics of the tropical intraseasonal oscillation in the NCAR Community Climate Model. *Q. J. Roy. Meteor. Soc.*, **117**, 1129-1169.

Slingo, J. M., and Co-authors, 1996: Intraseasonal oscillations in 15 atmospheric general circulation models: Results from an AMIP diagnostic subproject. *Clim. Dynam.*, **12**, 325-357.

Smith, R., and P. Gent (eds.), 2002: Reference manual for the Parallel Ocean Program (POP): Ocean component of the Community Climate System Model (CCSM2.0). Accessible online at [http://www.cgd.ucar.edu/csm/models/ccsm2.0.1/pop/doc/sci\\_ref\\_manual.pdf](http://www.cgd.ucar.edu/csm/models/ccsm2.0.1/pop/doc/sci_ref_manual.pdf).

Sperber, K. R., 2003: Propagation and the vertical structure of the Madden-Julian Oscillation. *Mon. Wea. Rev.*, **131**, 3018-3037.

Sperber, K. R., 2004: Madden-Julian variability in NCAR CAM2 and CCSM2. *Climate Dyn.*, (submitted, Oct. 2003).

Sperber, K. R., J. M. Slingo, P. M. Inness, and W. K.-M. Lau, 1997: On the maintenance and initiation of the intraseasonal oscillation in the NCEP/NCAR reanalysis and in the GLA and UKMO AMIP simulations. *Climate Dyn.*, **13**, 769-795.

Stephenson, D. B., and V. Pavan, 2003: The North Atlantic oscillation in coupled models: a CMIP1 evaluation. *Climate Dyn.*, **20**, 381-399.

Taylor, K.E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183-7192. Also see <http://www-pcmdi.llnl.gov/pcmi/pubs/ab55.html>.

Walker, G. T., and E. W. Bliss, 1932: World weather. *V. Mem. Roy. Meteor. Soc.*, **4**, 53-84.

Waliser, D. E., K. M. Lau, and J.-H. Kim, 1999: The influence of coupled sea surface temperatures on the Madden-Julian Oscillation: a model perturbation experiment. *J. Atmos. Sci.*, **56**, 333-358.

Warrilow, D.A., A.B. Sangster, and A. Slingo, 1986: Modelling of land surface processes and their influence on European climate. DCTN 38, Dynamical Climatology Branch, United Kingdom Meteorological Office, Bracknell, Berkshire RG12 2SZ, UK.

Washington, W.M., J.M. Weatherly, G.A. Meehl, A.J. Semtner, Jr., T.W. Bettge, A.P. Craig, W.G. Strand, J. Arblaster, V.B. Wayland, R. James, and Y. Zhang, 2000: Parallel Climate Model (PCM) control and transient simulations. *Climate Dyn.*, **16**, 755-774.

Webb, M., C. Senior, S. Bony and J. J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF, and LMD atmospheric models. *Climate Dyn.*, **17**, 905-922.

Weickmann, K.M., G.R. Lussky, and J.E. Kutzbach, 1985: Intraseasonal (30-60 day) fluctuations of outgoing longwave radiation and 250mb streamfunction during northern winter. *Mon. Wea. Rev.*, **113**, 941-961.

Wheeler, M, and G.N. Kiladis, 1999: Convectively coupled waves: Analysis of clouds and temperature in the wavenumber-frequency domain. *J. Atmos. Sci.*, **56**, 374-399.

Willmott, C.J. and K. Matsuura, 2000: *Terrestrial Air Temperature and Precipitation: Monthly and Annual Climatologies*. Accessible online at [http://climate.geog.udel.edu/~climate/html\\_pages/download.html#ghcn\\_T\\_P\\_climo](http://climate.geog.udel.edu/~climate/html_pages/download.html#ghcn_T_P_climo).

Wolff, J-O., E. Maier-Reimer, and S. Lebutke, 1997: The Hamburg ocean primitive equation model. DKRZ Technical Report No. 13, Deutsches KlimaRechenZentrum, Hamburg, Germany, 100 pp. Also at <http://www.mad.zmaw.de/Pingo/reports/ReportNo.13.pdf>.

Woolnough, S. J., J. M. Slingo, and B. J. Hoskins, 2000: The relationship between convection and sea surface temperature on intraseasonal timescales. *J. Climate*, **13**, 2086-2104.

Zhang, M.-H., and co-authors, 2004: Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *J. Geophysical Research*, submitted.

Xie, P., and P. Arkin, 1996: analysis of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate*, **9**, 840-858. Also see <http://www.cdc.noaa.gov/cdc/data.cmap.html> .

Xie, P., and P. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539-2558.

Yang, G.-Y, B. Hoskins, J. Slingo, 2003: Convectively coupled equatorial waves: A new methodology for identifying wave structures in observational data. *J. Atmos. Sci.*, **60**, 1637-1654.

Yukimoto, S., A. Noda, A. Kitoh, M. Sugi, Y. Kitamura, M. Hosaka, K. Shibata, S. Maeda, and T. Uchiyama, 2001: The new Meteorological Research Institute global ocean-atmosphere coupled GCM (MRI-CGCM2)--Model climate and variability. *Papers Meteor. and Geophys.*, **51**, 47-88.

Yukimoto, S., and A. Noda, 2003: Improvements of the Meteorological Research Institute global ocean-atmosphere coupled GCM (MRI-GCM2) and its climate sensitivity. CGER's Supercomputing Activity Report, National Institute for Environmental Studies, Tsukuba, Japan.

Zhang, Y., W. Maslowski, and A.J. Semtner, 1999: Impacts of mesoscale ocean currents on sea ice in high-resolution Arctic ice and ocean simulations. *J. Geophys. Res. (Oceans)*, **104(C8)**, 18409-18429.



## Appendix 1: Observationally Based Resources

Variable	Observationally based reference data set	Comment
Surface Air Temperature	Climate Research Unit ( <a href="#">Jones</a> )	surface instrumental record
Precipitation	CPC Merged Analysis of Precipitation ( <a href="#">CMAP</a> )	a merged analysis of observations from rain gauges and estimates from satellite-based algorithms
shortwave (SW) and longwave (LW) fluxes SW & LW clear-sky fluxes SW & LW cloud radiative forcing	Earth Radiation Budget Experiment ( <a href="#">ERBE</a> )	remotely sensed from satellites
Surface Sensible Heat Flux (ocean only) Surface Latent Heat Flux (ocean only) Eastward & Northward Surface Wind Stress (ocean only)	Southampton Oceanography Centre Atlas ( <a href="#">SOC</a> )	bulk parameterizations based on surface instrumental records
500 hPa Geopotential Height 200 & 850 hPa Zonal and Meridional Wind 200 & 850 hPa Temperature Mean Sea Level Pressure (ocean only)	European Centre Reanalysis ( <a href="#">ERA15</a> )	reanalysis product
Precipitable Water	NASA Water Vapor Project ( <a href="#">NVAP</a> )	a blended analysis of global water vapor from ground-based radiosondes and satellite instruments
Surface Zonal & Meridional Wind	NCEP/NCAR Reanalysis ( <a href="#">NCEP/NCAR</a> )	reanalysis product
Total Cloud Fraction	International Satellite Cloud Climatology Project ( <a href="#">ISCCP</a> )	remotely sensed from satellites



## Appendix 2: Interpretation of the Statistical Taylor Diagram

Taylor diagrams (Taylor, 2001) provide a way of graphically summarizing how closely a pattern (or a set of patterns) match observations. The similarity between two patterns is quantified in terms of their correlation, their root-mean-square difference, and the amplitude of their variations (represented by their standard deviations). These diagrams are especially useful in evaluating multiple aspects of complex models or in gauging the relative skill of many different models (e.g., IPCC, 2001).

Figure A.2.1 is a sample Taylor diagram that shows how it can be used to summarize the relative skill with which several global climate models simulate the spatial pattern of mean precipitation (averaged over the months of June, July, and August). Statistics for eight models were computed, and a letter was assigned to each model considered. The position of each letter appearing on the plot quantifies how closely that model's simulated precipitation pattern matches observations. Consider model F, for example. Its pattern correlation with observations is about 0.65. The centered root-mean-square (RMS) difference between the simulated and observed patterns is proportional to the distance to the point on the x-axis identified as "observed." The green contours indicate the RMS values and it can be seen that in the case of model F the centered RMS error is about 2.6 mm/day. The standard deviation of the simulated pattern is proportional to the radial distance from the origin. For model F the standard deviation of the simulated field (about 3.3 mm/day) is clearly greater than the observed standard deviation, which is indicated by the dashed arc at the observed value of 2.9 mm/day.

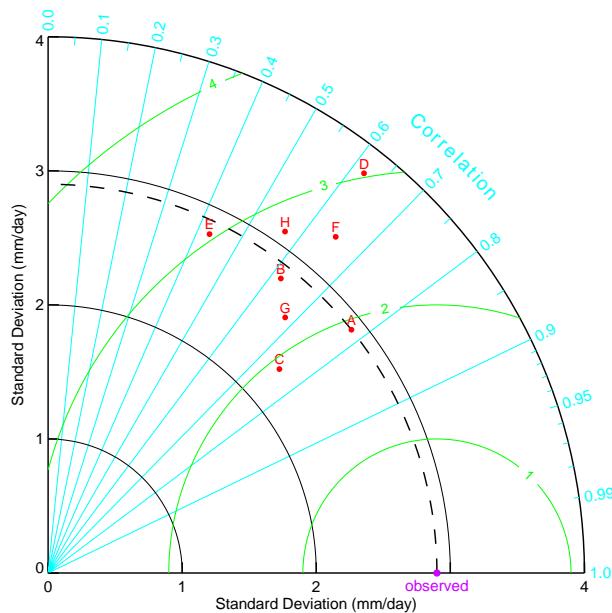


Figure A.2.1. Sample Taylor diagram displaying a statistical comparison between eight model estimates of the June/July/August precipitation fields and observations.

The relative merits of various models can be inferred from Fig. A.2.1. Simulated patterns that agree well with observations will lie nearest the point marked “observed” on the x-axis. These models will have relatively high correlation and low RMS errors. Models lying on the dashed arc will have the correct standard deviation (which indicates that the pattern variations are of the right amplitude). In Fig. A.2.1 it can be seen that models A and C generally agree best with observations, each with about the same RMS error. Model A, however, has a slightly higher correlation with observations and has the same standard deviation as the observed, whereas model C has too little spatial variability (with a standard deviation of 2.3 mm/day compared to the observed value of 2.9 mm/day). Of the poorer performing models, model E has a low pattern correlation, while model D has variations that are much larger than observed, in both cases resulting in a relatively large ( $\sim$ 3 mm/day) centered RMS error in the precipitation fields. Note also that although models D and B have about the same correlation with observations, model B simulates the amplitude of the variations (i.e., the standard deviation) much better than model D, and this results in a smaller RMS error.

In general, the Taylor diagram characterizes the statistical relationship between two fields, a “test” field (often representing a field simulated by a model) and a “reference” field (usually representing “truth,” based on observations). Note that the means of the fields are subtracted out before computing their second-order statistics, so the diagram does not provide information about overall biases, but solely characterizes the *centered* pattern error.

The reason that each point in the two-dimensional space of the Taylor diagram can represent three different statistics simultaneously (i.e., the centered RMS difference, the correlation, and the standard deviation) is that these statistics are related by the following formula:

$$E'^2 = \sigma_f^2 + \sigma_r^2 - 2\sigma_f\sigma_r R,$$

where  $R$  is the correlation coefficient between the test and reference fields,  $E'$  is the centered RMS difference between the fields, and  $\sigma_f^2$  and  $\sigma_r^2$  are the variances of the test and reference fields, respectively. (The formulas for calculating these second-order statistics are provided at the end of this document.) The construction of the diagram (with the correlation given by the cosine of the azimuthal angle) is based on the similarity of the above equation and the Law of Cosines:

$$c^2 = a^2 + b^2 - 2ab \cos \phi$$

There are several minor variations on the diagram that have been found useful for various purposes (see Taylor 2001). For example,

- The diagram can be extended to a second “quadrant” (to the left) to allow for negative correlations.

- The statistics can be normalized (and non-dimensionalized), dividing both the RMS difference and the standard deviation of the “test” field by the standard deviation of the observations. In this case the “observed” point is plotted on the x-axis at unit distance from the origin. This makes it possible to plot statistics for different fields (with different units) on the same plot.
- The isolines drawn on the sample plot above are often omitted to make it easier to see the plotted points.
- When comparing fields simulated by two different versions of a model, the two points on the graph representing those fields are often connected by an arrow to indicate more clearly whether or not the model is moving toward “truth,” as defined by observations.

**Further notes:**

Given a “test” field ( $f$ ) and a reference field ( $r$ ), the formulas for calculating the correlation coefficient ( $R$ ), the centered RMS difference ( $E'$ ), and the standard deviations of the “test” field ( $\sigma_f$ ) and the reference field ( $\sigma_r$ ) are given below:

$$R = \frac{\frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})(r_n - \bar{r})}{\sigma_f \sigma_r}$$

$$E'^2 = \frac{1}{N} \sum_{n=1}^N [(f_n - \bar{f}) - (r_n - \bar{r})]^2$$

$$\sigma_f^2 = \frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})^2$$

$$\sigma_r^2 = \frac{1}{N} \sum_{n=1}^N (r_n - \bar{r})^2$$

where the overall mean of a field is indicated by an overbar. In the case of a time-independent field, the sum is computed over all grid cells. For the typical spatial grid, the grid cell area is not uniform, so each grid cell must be weighted by the fraction of the total area represented by that grid cell. In the case of a time-varying field, the sum is a double-sum computed over all grid cells and all time samples.



## Appendix 3: Acronyms

See Table 2.1 for acronyms that identify the models used in this report.

AGCM	Atmospheric General Circulation Model
AMIP	Atmospheric Model Intercomparison Project
AOGCM	Atmospheric-Ocean General Circulation Model (can be synonymous with CGCM)
AVHRR	Advanced Very High Resolution Radiometer
CGCM	Coupled General Circulation Model (can be synonymous with AOGCM)
CMAP	CPC Merged Analysis of Precipitation
CMIP	Coupled Model Intercomparison Project
DJF	December-January-February
EOF	Empirical Orthogonal Function
ENSO	El Nino Southern Oscillation
ERA15	ECMWF Re-Analysis ERA-15
ERA40	ECMWF Re-Analysis ERA-40
IPCC	Intergovernmental Panel on Climate Change
ISCCP	International Satellite Cloud Climatology Project
JJA	June-July-August
KC	Köppen Climate Classification
MJO	Madden Julian Oscillation
NAO	North Atlantic Oscillation
NH	Northern Hemisphere
NRL	Naval Research Laboratory
OLR	Outgoing Longwave Radiation
RMS	Root-Mean-Square (difference)
PC	Principal Component
PSL	Mean Sea-Level Pressure
SPCZ	South Pacific Convergence Zone
SH	Southern Hemisphere
SOI	Southern Oscillation Index
SST	Sea Surface Temperature
TAR	IPCC Third Assessment Report
TAS	Surface (2m) Air Temperature
WCRP	World Climate Research Programme
WGCM	Working Group on Coupled Models
WGNE	Working Group on Numerical Experimentation
WOCE	World Ocean Circulation Experiment