The Vertical Structures of Atmospheric Temperature Anomalies Associated with Two Flavors of El Niño Simulated by AMIP II Models

TIANJUN ZHOU
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

JIE ZHANG
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, and Graduate University of Chinese Academy of Sciences, Beijing, China

(Manuscript received 5 November 2009, in final form 10 October 2010)

ABSTRACT

Recent studies have identified different modes associated with two flavors of El Niño in terms of the three-dimensional structure of atmospheric temperature. The first is a deep-warm mode, which features a coherent zonal mean warming throughout the troposphere from 30°N to 30°S with cooling aloft. The second is a shallow-warm mode, which features strong wave signatures in the troposphere with warmth (coolness) over the central Pacific (western Pacific). The ability to simulate these two modes is a useful metric for evaluating climate models. To understand the reproducibility of these two modes, the authors analyzed the multimodel ensemble mean (MMEM) of 11 atmospheric general circulation models (AGCMs) that participated in the second phase of the Atmospheric Model Intercomparison Project (AMIP II). Each model was run in an AGCM-alone mode forced by historical sea surface temperatures covering the period 1980–99. The authors find that atmospheric temperature variability is generally well captured in the MMEM of AMIP II models, demonstrating that the observational changes documented here are driven by SST changes during the El Niño events and the variety of vertical temperature structures associated with two flavors of El Niño are highly reproducible. The model skill for the first mode is slightly higher than the second mode. The skill in the upper troposphere–lower stratosphere is lower than for the tropospheric counterpart, especially at high latitudes. The performances of individual models are also assessed. The authors also show some differences from previous data analyses, including the variance accounted for by the two modes, as well as the lead–lag relationship of the shallow-warm mode with the Niño-3.4 index.

1. Introduction

The El Niño–Southern Oscillation (ENSO) is the main interannual variation in the global climate system and is regarded as one of the most prominent sources of short-term climate prediction (Rasmusson and Carpenter 1982; Goddard et al. 2001). An understanding of atmospheric changes associated with ENSO is fundamental to the practice of short-term climate prediction efforts. Global teleconnections forced by ENSO and the fingerprints of ENSO in the evolution of global climate variability have been extensively studied (Lau 1997; Hoerling et al. 1997; Trenberth et al. 1998, 2002; Giannini et al. 2001a,b; Alexander et al. 2002; Diaz et al. 2002; Hoerling and Kumar 2002).

To reveal the differences in character among the different El Niño events, empirical orthogonal function (EOF) analysis of sea surface temperatures (SSTs) is often used. The first EOF of SSTs is well measured by the Niño-3.4 index, while the second EOF of SSTs is characterized as the trans-Niño index (Trenberth and Stepaniak 2001). The second EOF is also referred to as El Niño Modoki (Ashok et al. 2007), central Pacific ENSO (Kao and Yu 2009), or warm-pool El Niño (Kug et al. 2009). These previous studies on the climate impact of ENSO have mainly focused on surface climate and specific levels, and less effort has been devoted to vertical structure. Recently, by analyzing the full three-dimensional temperature structure of the atmosphere...
within the World Climate Research Program (WCRP), various projects evaluated the performance of climate models when driven by the same boundary conditions or forcing scenarios. One of these projects is the Atmospheric Model Intercomparison Project (AMIP) (Gates et al. 1999), which is currently in its second phase of operation (AMIP II). The performances of AMIP II models in simulating the cloud and radiation fields, solar insolation, and the leading interannual variability modes of the Asian–Australian monsoon have been discussed by Potter and Cess (2004), Wild (2005), Raschke et al. (2005), and Zhou et al. (2009a,b), respectively. The vertical structure of temperature related to different flavors of El Niño has not yet been discussed, mostly because this is a relatively new finding. The main purpose of this study is to assess the ability of AMIP II models to simulate the 3D structure of the atmospheric temperature fields associated with different flavors of El Niño, based on the metrics identified in TS2006 and TS2009. The use of a multimodel ensemble mean (MME) allows us to identify atmospheric responses to different flavors of El Niño by reducing uncertainties influenced by specific model biases. Assessing the ability of climate models will indicate whether the observed variability is driven by SST changes. This will improve our understanding of the driving mechanisms of 3D atmospheric temperature changes.

The rest of the paper is organized as follows. Section 2 details the methods and model data used, and section 3 presents the results. The conclusions are given in section 4.
of corresponding ocean–atmosphere coupled models. For brevity, we use the names of coupled models to represent the AGCM components in the following discussion.

The verification datasets used in this study include the following: 1) ERA-40 data (Uppala et al. 2005) were used to obtain the temperature at different vertical levels, sea level pressure, and geopotential height; all the data are available on a 2.5° × 2.5° grid. 2) The Niño-3.4 SST index, which is the area-averaged SST anomalies over 5°N–5°S, 170°–120°W, was obtained from the Climate Prediction Center of the National Centers for Environmental Prediction (NCEP; available online at http://www.cpc.ncep.noaa.gov/data/indices/). (3) The monthly historical SST dataset (Rayner et al. 2003) was used. The monthly annual cycle was removed from all fields using a 1980–99 climatology. The anomalies in the modeled monthly annual cycle was removed from all fields using the historical SST dataset (Rayner et al. 2003). The results were then projected onto all levels and the global domain using regression/correlation. Before performing EOF analysis, the values were weighted by the square root of the cosine of the latitude to account for the convergence of meridians. To emphasize covariability, the EOF analysis was performed with a correlation matrix, which effectively normalizes each grid point by the standard deviation of the temporal variability. When presenting figures in the following discussion, we show both correlation and regression maps. While the correlation map emphasizes the statistical connection, the regression map gives the actual anomaly patterns. All of the models were analyzed separately as well as the MMEM.

It should be noted that in TS2006, the EOF2 derived from ERA-40 had a complex vertical structure and has been demonstrated to be largely spurious and is associated with problems in assimilating satellite data (TS2006, TS2009). Hence, the observational metrics used in our model evaluation include the following: 1) EOF1 highlighting the 1997–98 El Niño, along with El Niño events in 1982–83 and 1986–87, and the 1988–89 La Niña; we will refer to EOF1 as the deep-warm mode. 2) EOF3 highlights shallow-warm mode.

In our analysis, we compare the model simulation against the “observational metrics” derived from reanalysis. A high reproducibility refers to a good resemblance. We note that in nature the observed variability includes both the SST-forced part and internal variability. Given the SSTs,
the performance of models has two components. The first is how good the model is, and the second relates to reproducibility. If the response of the real world is largely deterministic, then the result is reproducible. It is likely that some aspects are not deterministic, so are poorly reproduced in models. It is also likely that none of the models are perfect and model flaws affect the reproducibility. In our discussion, while we mainly focus on the MMEM, we also comment on individual model performances based on objective criteria.

3. Results

a. Zonal mean vertical structure of EOF modes

For both deep-warm and shallow-warm modes, we present the zonal mean temperature anomaly structure given by regression (Fig. 1) and corresponding correlation coefficients (Fig. 2). The percentage variances accounted for by the EOF modes are shown in Fig. 3. The principal component (PC) time series of the two modes are shown in Figs. 4a and 4b. The results derived from the MMEM are given. The observational features of the modes have been well documented by TS2006 and TS2009. To facilitate the model evaluation, we summarize the main points of the observations below.

1) DEEP-WARM MODE

In the reanalysis, the first EOF (Fig. 1a) accounts for 43.5% of tropical temperature variance (Fig. 3a) and features the 1982–83, 1986–87, and 1997–98 El Niño events in its time series (Fig. 4a). A coherent zonal mean warming is clear throughout the tropical troposphere from 30°N to 30°S. The magnitude increases with height to about 300 hPa, with a reversed sign in the lower stratosphere centered at 70 hPa. At higher latitudes in both hemispheres, there is modest warming in the lower troposphere and cooling (warming) aloft in the Southern Hemisphere (Northern Hemisphere) (Fig. 1a). The correlation coefficients exceed 0.6 throughout the tropical troposphere from 200 to 500 hPa. At high latitudes, the correlation coefficients are insignificant (Fig. 2a). Our model evaluation focuses on the tropical signals.

2) SHALLOW-WARM MODE

In the reanalysis, the third EOF (Fig. 1b) accounts for 9.9% of the variance (Fig. 3a). The time series of this
EOF features the prolonged and unusually weak El Niño events in the early 1990s along with the 1982–83 event (Fig. 4b). So, the two modes sometimes occur at the same time. The shallow-warm mode has a shallow zonal-mean warming in the tropics from the surface to 400 hPa, a weak cooling aloft to 100 hPa, and then a peak warming around 30 hPa that extends into both hemispheres. The low-level tropical warming is surrounded by subtropical cooling in both hemispheres. There are peak negative correlation coefficients at 30° latitude in both hemispheres (Fig. 2b). In addition, large differences are seen in the regressed vertical structures at high latitudes in Figs. 1b and 1d: the physical reason is that the shallow-warm mode is more important for weak El Niño events (TS2006).

While the spatial structures of the two modes generally coincide well with TS2006, there are some differences: first, the deep-warm mode explains a larger variance than the deep-warm mode identified by TS2006 (43.5% versus 22.2%) and, second, the shallow-warm mode exhibits a slightly different structure above 100 hPa that is spatially amorphous in polarity. The corresponding pattern of TS2006 exhibits the same polarity throughout the tropical–subtropical domain. How is the difference between this study and TS2006 explained? Although the time period of data coverage is slightly different, with 1980–99 versus 1979–2001, the main reason for this different structure may be due to the sample size of the grids. Our analysis is done on a regular 2.5° × 2.5° grid, while TS2006 was performed on a Gaussian grid at T63 resolution and using every second grid point in longitude (TS2006, TS2009). We also repeat TS2009 by using the new Japanese reanalysis data but do the analysis on a regular 2.5° × 2.5° grid. While the spatial structures of two modes are the same as TS2009 (figures not shown here), the explained variances are 50.7% for EOF1 (i.e., deep-warm mode) and 10.4% for EOF2 (i.e., shallow-warm mode)—both higher than TS2009, indicating the impact of grid size.

In the MMEM of AMIP, the first EOF mode relates to the well-documented warm and cold ENSO events during this period, as shown by the PC time series (Fig. 4a), which has a simultaneous correlation coefficient 0.79 with its observational counterpart. It is significantly separated from the other modes based on the rule of North et al. (1982). The deep-warm mode accounts for 69.5% of tropical temperature variance, larger than its counterpart in the reanalysis, at 43.5% (Fig. 3). We may consider the observation as one member because the MMEM does not contain as much internal variability as the observations: the two leading EOF modes in the MMEM due to external
SST forcing should explain more variance than the observation and a higher variance is expected.

The deep-warm mode in the MMEM bears close similarity to the corresponding mode in the reanalysis of the deep vertical structure, with a strong tropospheric warming extending from the surface to 100 Pa and cooling above 70 hPa (Fig. 1c). The spatial correlation coefficient between the simulation and the reanalysis pattern within 30°N–30°S and 1000–100 hPa is 0.92, which is statistically significant at the 1% level. The midtropospheric cooling signals at high latitudes in the reanalysis are absent in the regression pattern (Fig. 1c) and very
weak in the correlation pattern (Fig. 2c). In the MMEM, the tropospheric warming spreads from the equatorial zone toward the northern and southern subtropics above 100 hPa.

In the MMEM of AMIP, the second EOF mode is well distinguished from the other modes in terms of the sampling error bars based on North et al. (1982) and, thus, is statistically significant. It highlights the prolonged sequence of three successive El Niño events in the early 1990s as well as the 1982–83 event (Fig. 4b). The PC2 time series of the MMEM is highly consistent with the PC3 reanalysis time series, having a correlation coefficient of 0.71. Hence, the EOF2 mode derived from the MMEM corresponds to that of EOF3 derived from ERA-40. TS2006 found that the EOF2 mode derived from ERA-40 is spurious. Thereby, it cannot be reproduced with specified historical SST forcing. The shallow-warm mode explains 11.1% of the tropical variance, slightly higher than 9.9% of its counterpart in the reanalysis (Fig. 3).

The vertical structure of temperature anomalies associated with the shallow-warm mode bears similarity to the corresponding structure in the reanalysis in the troposphere below 150 hPa; both feature a tropical warming surrounded by subtropical cooling in both hemispheres (Fig. 1d). The pattern correlation coefficient between the simulation and the reanalysis for the shallow-warm mode within the tropical troposphere (30°N–30°S and 1000–100 hPa) is 0.83, which is statistically significant at the 1% level. Two negative correlation minima occur in the subtropics of both hemispheres around 200–300 hPa. The subtropical cooling straddling 30° in both hemispheres is stronger than in the reanalysis (Fig. 2d). One difference appears in the tropics above 100 hPa: there a cooling center exists in the MMEM, whereas the reanalysis shows weak warming.

b. Relationship with the Niño-3.4 index

The time series of the deep-warm mode, that is, EOF1, clearly features the well-documented El Niño events (Fig. 4a). We further examine the correlation of the PCI time series with the Niño-3.4 index at various leads and lags. The correlation at zero lag is 0.47 (Fig. 4c). The maximum correlation coefficient of 0.70 occurs when the deep-warm mode lags the Niño-3.4 index by 4–5 months. The correlation coefficients presented here are generally lower than those of TS2006 and TS2009. Power spectra analysis of the time series of the deep-warm mode reveals a major significant spectral peak at 2–3 yr and a secondary insignificant peak around 3–4 yr (Fig. 5a).

The relationship between the deep-warm mode and the Niño-3.4 index is reasonably simulated in the MMEM, also showing a maximum correlation coefficient when the Niño-3.4 index leads by about 4–5 months (Fig. 4c). The results of individual models are close to the MMEM. The power spectra of PC1 time series of the MMEM are, however, slightly different from the reanalysis: the major spectral peak of the time series is at 3–4 yr and the secondary peak at 2–3 yr (Fig. 5c).

The time series of the shallow-warm mode, that is, EOF3 in ERA-40 and EOF2 in MMEM, features the 1992–95 prolonged event (or series of events, Trenberth and Hoar 1996), along with the 1982–83 El Niño event (Fig. 4b). The time series of the shallow-warm mode is correlated at zero lag with the Niño-3.4 index at 0.60 (Fig. 4d), consistent with TS2006. However, TS2006 found that their EOF3 led the ENSO index by 1–2 months, whereas the maximum correlation between the time series of the shallow-warm mode and the Niño-3.4 index occurs at zero lag in our analysis (Fig. 4d). Our result is supported by TS2009, who found that the maximum correlation of the shallow-warm mode with the Southern Oscillation index is at zero lag. Power spectra analysis of the time series of the shallow-warm mode reveals a major peak at 2.5 yr and a secondary peak at 5 yr, both statistically significant at the 5% level (Fig. 5b).

The relationship of the shallow-warm mode with the Niño-3.4 index is overestimated in the MMEM, as indicated by the higher correlation coefficient (Fig. 4d). The MMEM is the average of the models, so the noise cancels. This reduces the noise variance in the MMEM and accounts for why the mean of the models in Fig. 4d is much lower than the MMEM result. The time series of the shallow-warm mode in the MMEM also has a maximum correlation coefficient with the Niño-3.4 index at zero lag; the correlation coefficient of 0.75 is stronger than that of 0.60 for the reanalysis. Particularly noteworthy is the spread among the models in the correlation coefficients and the lead–lag relationships (Fig. 4d). This is in contrast to the convergence of all model responses in simulating the deep-warm mode (Fig. 4c). The power spectra of the MMEM resembles the reanalysis power spectrum, but only the major spectral peak at 2.5 yr is statistically significant (Fig. 5d).

c. Spatial structure of the EOF modes

The zonal mean patterns presented in Fig. 1 are not enough to clarify the spatial structures of temperature anomalies associated with different flavors of El Niño because some important signals may be cancelled out in the zonal mean owing to the existence of wave structures. The spatial structures in the MMEM at 1000, 850, 700, 500, 300, 200, 100, and 50 hPa are illustrated in Figs. 6 and 7. The corresponding structures derived from the reanalysis bear similarities to those given in TS2006 (see Figs. 6 and 8 of TS2006). For the sake of brevity, the
results derived from the reanalysis are not shown here. Pattern correlation coefficients between the simulation and the reanalysis for each mode at various levels are shown at the top-right corner of each panel of Figs. 6 and 7.

For the deep-warm mode or EOF1 (Fig. 6), the simulation closely resembles the observations: the near surface (1000–850 hPa) warm temperatures in the central-eastern Pacific are surrounded by a cold region, which extends from the western tropical Pacific into the subtropics of the Pacific in both hemispheres. Both the tropical Indian Ocean and the western Atlantic exhibit a comparable warming. The teleconnection patterns of Pacific–North America (PNA) and Pacific–South America (PSA), as defined in Trenberth et al. (1998), are evident in extratropics. The sign of the temperature anomaly across the midlatitudes is reversed over the ocean and the land, and this wave structure cancellation explains the absence of zonal mean signal at high latitudes (Fig. 1). The temperature structures at 500 and 700 hPa are more zonal, with maximum centers over the eastern Pacific and tropical Indian Ocean at both levels.

In the MMEM, the tropical structure at 200–300 hPa bears similarities to the structure at 500–700 hPa. However, in the reanalysis shown by TS2006 the tropical temperature anomalies at 200–300 hPa are characterized by two warm centers over the eastern Pacific and Indian Ocean. Another deficiency of the simulation is the weak cooling anomalies over extratropical East Asia and North America at 200–300 hPa. This explains the weak cooling centers located around 200 hPa in the zonal mean plot (Fig. 2c). The pattern at 100 hPa features a characteristic structure of off-equatorial Rossby waves, with two dumbbell-shaped cooling centers in the eastern Pacific and two reversed polarity warming centers over the western Pacific. By 50 hPa, the structure of the temperature anomalies is more zonal, with uniform cooling in the tropics and weak warming in the extratropics.

The patterns of temperature anomalies at the above selected levels are reasonably represented by the MMEM. Pattern correlation coefficients between the simulation and the reanalysis all exceed 0.74 from 1000 to 50 hPa. The highest correlation coefficient of 0.93 is seen at 500 and 700 hPa, indicating a high level of SST forcing in the vertical temperature changes. In terms of pattern correlation coefficient SST forcing at higher levels, such as 100–50 hPa, is less than that at lower levels.

The corresponding results of the shallow-warm mode are shown in Fig. 7. The surface temperature patterns at
1000–850 hPa associated with the shallow-warm mode (i.e., EOF3 in ERA-40) in the reanalysis of TS2006 are characterized by a warming in the tropical eastern Pacific and cooling in the western Pacific. This bears a slight similarity to the deep-warm mode. However, the shallow-warm mode exhibits features that differ from the deep-warm mode: 1) the tropical Pacific warming is broader in meridional extent than the deep-warm mode.
and extends northward and southward along the coasts of the Americas; 2) the western Pacific cooling is stronger than the deep-warm mode in amplitude; and 3) cool conditions prevail over the eastern tropical Indian Ocean, the South Asian continent, and in the tropical Atlantic.

These pronounced features are well reproduced in the MME (cf. Fig. 7 with Fig. 8 of TS2006), except for the surface cooling over South Asia. From 700 to 200 hPa the tropical temperature anomalies feature a wave structure straddling the equator. Two cool centers over South
Asia and the Indian Ocean and corresponding warm centers over the central Pacific are evident in the simulation. The pattern reverses sign at 100 hPa, with cold temperatures over the central Pacific and warm centers over the Maritime Continent. The sign reversal of the wave structure from 200 to 100 hPa is reasonably simulated. From 1000 to 100 hPa the MMEM shows comparable skill in terms of pattern correlation coefficient.

One difference of the simulation is seen over tropical Africa at 100 hPa where the dominant feature is a cooling in the MMEM, but a warming in TS2006. This difference may be related to uncertainty of the reanalysis; the result of TS2009 based on Japanese reanalysis data also exhibits a weak cooling there (cf. Fig. 5 of TS2009 with Fig. 7 of this paper). The largest difference between the reanalysis and the MMEM appears at 50 hPa where the dominant feature of the tropics is warming in the reanalysis, but cooling in the simulation.

From 1000 to 100 hPa the pattern correlation coefficients between the reanalysis and the MMEM range from 0.52 to 0.79, all are statistically significant at the 1% level. This indicates that the temperature anomalies at tropospheric levels associated with the shallow-warm mode are also strongly SST forced. The skill of the shallow-warm mode is generally lower than the deep-warm mode in terms of pattern correlation coefficient. A negative pattern correlation is seen for 50 hPa. The following reason may explain the low skills of AMIP models at 50 hPa: first, the SST forcing associated the shallow-warm mode during ENSO events is weaker than the deep-warm mode and, second, the AMIP models generally employ low vertical resolution and their description of the stratosphere is far from perfect. The problems at 50 hPa are likely related to model deficiencies. We have examined the results of individual models, the vertical levels of AMIP models range from 20 to 56 (see Table 1), but a greater vertical resolution is not followed by a higher pattern correlation coefficient (figures not shown). Whether an increase in vertical resolution of a specific model could improve the model skill in the low stratosphere deserves further study.

d. Multimodel intercomparisons

Analysis of the MMEM reveals reasonable agreement between the modeled and observed vertical temperature anomalies. Here we will examine the performance of individual models. We again use correlation analysis to assess the level of agreement between the observation and individual models in PC time series (Fig. 8a) and the spatial patterns (Fig. 8b) of the two EOF modes. In Fig. 8, the EOF1 (EOF3) mode derived from ERA-40 is used as ground truth for the deep- (shallow-) warm mode for model comparison. All models show high temporal correlation coefficients with the reanalysis in reproducing variations of the PC time series (Fig. 8a), confirming the results based on visual inspection of Figs. 4a and 4b. The correlation coefficients for the deep-warm mode are generally higher than the shallow-warm mode: those for the former range from 0.71 to 0.79, while for the latter range from 0.48 to 0.71. While all 11 analyzed models show comparable skill, the MMEM has the highest skill. The NCAR and MRI models are generally better than the others in terms of correlation coefficients.

In terms of spatial patterns represented by the regression coefficient, the AMIP models show high skill in reproducing the zonal-mean vertical structures of temperature anomalies associated with the two different flavors of El Niño. Most models have pattern correlation coefficients higher than 0.86 (0.63) for the deep (shallow)-warm mode. For most models, the skill in reproducing the deep-warm mode is generally better than the skill in reproducing the shallow-warm mode. This also applies to the MMEM (0.93 for the deep-warm mode versus 0.83 for the shallow-warm mode). Because the shallow-warm mode is likely more important for weak El Niño events (TS2006), the skill difference is consistent with Kumar and Hoerling (1997), who found that the model response increases almost linearly with the strength of SST warming, although their conclusion was based on the amplitude of the signal in the extratropics. The MRI, IPSL, and UKMO models show higher skill in simulating the deep-warm mode, with correlation coefficients higher than 0.92. The MPI and MRI models also exhibit high skill in reproducing the shallow-warm mode, with correlation coefficients exceeding 0.85.

The spatial pattern is also assessed by rms deviation (RMSD), which is frequently used to measure the difference between the values simulated by a model and those actually observed. As shown in Fig. 8c, the minimum bias is seen in the MMEM. The bias for the deep-warm mode is generally less than that of the shallow-warm mode. The skill of the MRI model is close to the MMEM. The model performance measured with RMSD is consistent with model performance measured with pattern correlation.

A high pattern correlation coefficient (or low RMSD) does not indicate well-simulated amplitude. The variance of the simulated temperature anomalies associated with the two modes compared to the reanalysis are examined in Fig. 8d, which shows the rms of temperature anomalies over the target domain normalized by the “observed” rms represented by the reanalysis. If the normalized rms is smaller (larger) than 1, it implies that the mean amplitude of the simulated anomaly is weaker (stronger) than the reanalysis. Most models overestimate
the amplitude of the deep-warm mode, but underestimate that of the shallow-warm mode. The amplitude of the MMEM is comparable to (weaker than) the reanalysis for the deep (shallow)-warm mode. The use of a multimodel ensemble technique generally improves both spatial and temporal correlation coefficients with no, or only a small, reduction in the amplitudes of anomalies. This suggests that the zonal average anomalies are very largely SST driven.

The statistics shown above are only for the zonal mean temperature at 30°S–30°N and between 1000 and 100 hPa. We also investigate the performances of individual models in reproducing the horizontal spatial patterns of temperature anomalies. An evaluation of the spatial structures at selected levels shows a reasonable performance of the MMEM (see Figs. 6 and 7). We further examine the skill of individual models in Fig. 9, which is the same as in Fig. 8 except for the global temperature anomalies at 500 hPa. The skill for the deep-warm mode is generally higher than that for the shallow-warm mode. The MMEM has the highest spatial correlation coefficient for the deep-warm mode, but
not for the shallow-warm mode: four models show comparable or better skill than the MMEM in terms of spatial correlation coefficients (Fig. 9a). The MMEM exhibits the minimum RMSD for the deep-warm mode (Fig. 9b). In terms of the amplitude of the anomalies, five models show results comparable to the reanalysis in the deep-warm mode, and the corresponding result of the MMEM is about 82% of the reanalysis, mainly due to the cancellation of individual model skill over middle and high latitudes. Most models underestimate the rms of the shallow-warm mode; the amplitude of the MMEM is about 77% of the reanalysis (Fig. 9c).

One deficiency of the MMEM is that the percentage variance accounted for by the deep-warm mode is higher than for the reanalysis. This is, at least in part, due to the ensemble technique, which suppresses the internal variability. Given the noise cancellation in the MMEM, one expects up front that the MMEM variance explained should be higher. To make the simulation and the reanalysis comparable, that is, each having only one ensemble member, we calculated the fractional variances explained by the two major modes in each individual model. Table 2 shows how the percent variances explained by the two modes vary among different models. The fractional variance for individual models is, as expected, lower than the fractional variance for the MMEM. While the fractional variances for the shallow-warm mode accounted for by several models are comparable to the fractional variance derived from the reanalysis, those for the deep-warm mode are all larger than the latter. It seems that all models put more variability into EOF1 than they should. This may be related to the way convection is parameterized in models and is linked to the unrealistically strong coupling of tropical convection to local SST. So, when triggered, deep convection occurs

![Fig. 9. As Figs. 8(b)–(d) but for the 500-hPa global temperature anomalies depicted by regression coefficients.](image-url)
much more in models than in the real world (Dai 2006; Lin et al. 2006).

e. Related sea level pressure and geopotential height patterns

In TS2006, although the vertical temperature structures associated with the two flavors of El Niño differ considerably, the sea level pressure patterns are similar. Both feature SO-like patterns with slight regional differences. To examine whether this is also true for the MMEM, we regressed the time series for the two flavors of El Niño onto mean sea level pressure and 300-hPa geopotential height fields. The results are shown in Fig. 10. Corresponding maps for correlation are presented in Fig. 11. A comparison of the model and the reanalysis results presented in TS2006 indicates that the MMEM is highly consistent with the reanalysis in reproducing the circulation fields, especially for the deep-warm mode (Figs. 10a,b and 11a,b). The classic Southern Oscillation structure appears in the shallow-warm mode (Figs. 10b and 11b), with maximum correlation coefficients off the equator in the eastern Pacific and high correlation coefficients in the South Pacific, like the observations (Trenberth and Caron 2000). The PNA spiral teleconnection pattern is clear in both cases, but the shallow-warm mode features stronger amplitude, especially for the low pressure over the North Pacific. The modeled spiral teleconnection patterns in the tropical Pacific and North Pacific bear close similarity to that in the reanalysis, but the center of PNA over North America is weaker than the reanalysis in amplitude.

In the upper troposphere, at 300 hPa, the tropical signal is stronger in the deep-warm mode, while the extratropical teleconnection is stronger in the shallow-warm mode (Figs. 10c,d and 11c,d). The former features a zonal structure in the tropics with the maximum in the eastern Pacific straddling the equator. The latter has major changes over the extratropical Pacific in both hemispheres, with a fairly weak dumbbell-shaped positive signal in the central Pacific straddling the equator. These features are also clear in the reanalysis presented in TS2006. The skill is also confirmed by the significant
pattern correlation coefficients marked at the top-right corner of each panel of Figs. 10 and 11. One weakness of the simulation is the weaker than observed positive height anomaly over the North American region associated with the shallow-warm mode. This indicates a weak response of the AMIP models in simulating the teleconnection pattern extending from the North Pacific to high latitudes.

To provide a quantitative measure of the simulation of the teleconnection pattern, following Kumar and Hoerling (1997), we computed pattern correlation coefficients, RMSD, and rms for the regression pattern of $Z_{300}$ over the PNA region, defined as the area within 20°–70°N, 180°–60°W (Fig. 12). In Fig. 12a, the GISS, NCAR, and the MIROC models are inferior but the others are all comparable, with correlation coefficients for the deep-warm mode varying from 0.40 to 0.95. The correlation coefficients for the shallow-warm mode are generally lower than for the deep-warm mode but still higher than 0.40 except for the MIROC and NCAR models. The correlation coefficient of the MMEM is 0.95 for the deep-warm mode but only 0.40 for the shallow-warm mode (Fig. 12a). Many models show skill higher than the MMEM in simulating the shallow-warm mode. The bias is large for both modes in terms of RMSD, but the MMEM clearly has a smaller bias. The amplitude of the height anomalies over the PNA region in the MMEM is weakly simulated for the deep-warm mode, that is, only 67% of the reanalysis. The amplitude of the shallow-warm mode is comparable to or slightly stronger than the reanalysis, as evidenced by the rms value of 1.1 (Fig. 12c). A further comparison with Fig. 10d indicates that the skill is dominated by the North Pacific part of the PNA spiral pattern.

**f. Difference in SST forcing fields**

In conjunction with TS2006 and TS2009, the evidence presented above indicates that the deep-warm mode and shallow-warm mode differ considerably in the vertical structures and extratropical teleconnections but bear similarity in terms of sea level pressure. Both the difference and the similarity can be reasonably reproduced in the AMIP MMEM. TS2009 explains the difference between the two modes in the reanalysis in terms of the atmospheric response to nearly similar diabatic heating in different seasons. Because the only source of forcing in the AMIP experiment is from the sea surface temperature, we have examined the SST anomalies associated with two modes (figures not shown). The patterns are similar to the temperature at 1000 hPa. While a warmer central and eastern Pacific and a reversed-sign temperature anomaly in the western Pacific are still evident, significant differences are seen between the two
SST forcing fields. First, the cooling in the western Pacific is much stronger for the shallow-warm mode. Second, a prominent tropical Indian Ocean warming is seen in the deep-warm mode but not in the shallow-warm mode, especially in the eastern tropical Indian Ocean, associated with the shallow-warm mode is an Indian Ocean dipole pattern (Saji et al. 1999; Webster et al. 1999).

Third, both the tropical and the extratropical North Atlantic are dominated by warming SST anomalies in the deep-warm mode but not in the shallow-warm mode. At least in the model world, these are the three aspects of difference that result in the different vertical temperature structures.

4. Conclusions

To examine whether the vertical temperature anomalies related to two different flavors of El Niño can be reproduced by AGCMs with prescribed SST forcings, we analyzed 11 AGCMs that participated in the second phase of the Atmospheric Model Intercomparison Project, which used historical SST and sea ice to drive AGCMs for the period 1980–99. The simulations are compared with observations. The results demonstrate high skill in reproducing vertical temperature structures associated with two flavors of El Niño. The main findings of the study are summarized below.

1) The different flavors of El Niño considered with respect to the full three-dimensional structure of the atmospheric circulation and temperature field are well simulated by AGCMs driven by observational sea surface temperature. The deep-warm mode exhibits a highly coherent zonal mean warming throughout the troposphere from 30°N to 30°S with cooling aloft. The shallow-warm mode exhibits a warming over the central Pacific and cooling over the western Pacific in the low troposphere. This pattern reverses sign at 100 hPa. Major features of the atmospheric changes associated with two flavors of El Niño are reasonably reproduced in the MMEM of AMIP simulations (i.e., including the three-dimensional spatial patterns, temporal variations, and the relationship with the Niño-3.4 index). The tropical–subtropical tropospheric (30°N–30°S, 1000–100-hPa) temperature anomalies associated with two flavors of El Niño are all also well-simulated, with a pattern correlation coefficient of 0.92 between the zonal mean structure in the MMEM and the reanalysis for the deep-warm mode and 0.83 for the shallow-warm mode.

2) The skill in simulating upper troposphere–lower stratosphere temperature anomalies is lower than the tropospheric counterpart, especially over the high latitudes. For the deep-warm mode, an artificial warming spreading from the tropics to the high latitudes of both hemispheres is evident in the MMEM. For the shallow-warm mode, cooling (rather than the warming seen in the reanalysis counterpart) dominates the lower stratosphere over the tropics. This deficiency is most evident at 50 hPa.
3) The skill of the model in simulating extratropical tropospheric temperature anomalies associated with two flavors of El Niño is lower than the skill associated with reproducing the tropical anomalies. The PNA teleconnection spiral pattern associated with the deep-warm mode is reasonably simulated, but its amplitude is underestimated. The spiral pattern of the PNA teleconnection associated with the shallow-warm mode is poorly simulated at 300 hPa.

4) Performance of individual models in simulating the vertical structure of temperature anomalies associated with two flavors of El Niño is assessed. The MMEM shows the highest skill in simulating the tropical vertical structure in terms of temporal and spatial correlation coefficients, root-mean-square deviation, and root-mean-square value. The spread in skill among individual models in simulating the deep-warm mode is far smaller than for the shallow-warm mode. The MMEM shows the highest skill in simulating the spatial pattern of the PNA associated with the deep-warm mode, but the simulated amplitude is weaker than the amplitude in both the reanalysis and some individual models. The skill of some individual models is higher than that of the MMEM in simulating the PNA teleconnection associated with the shallow-warm mode, both in spatial pattern and in amplitude.

5) In the model world, robust different vertical temperatures associated with two different flavors of El Niño result from the regional difference in SST forcing fields, including a stronger cooling in the western Pacific for the shallow-warm mode and a prominent warming in the tropical Indian Ocean and the tropical and extratropical North Atlantic for the deep-warm mode. An Indian Ocean dipole pattern is associated with the shallow-warm mode.

6) Analysis of ERA-40 data also reveals some features that are different from TS2006. The variance accounted for by the deep-warm mode (43.5%) is larger than in TS2006 (22.2%). The surface tropospheric warming associated with the shallow-warm mode is deeper in our analysis than in TS2006 in the vertical direction, and the temperature anomalies above 100 hPa also differ. In TS2006, the shallow-warm mode leads the Niño-3.4 index by 1–2 months, whereas the maximum correlation in our analysis is found at zero lag.

In summary, we have presented evidence that the observational changes documented in TS2006 and TS2009 are driven by SST changes during El Niño events. The 3D structure of atmospheric temperature variability is generally well captured in the multimodel ensemble mean of the AMIP models. This study also indicates that the findings of TS2006 and TS2009 can serve as useful observational metrics for evaluating climate models.

Acknowledgments. This work was supported by the Research and Development Special Fund for Public Welfare Industry (meteorology) (GYHY200806010), the National Key Technologies Research and Development Program (2007BAC29B03), and the National Natural Science Foundation of China (40890054, 40821092). Helpful review comments from Drs. Kevin Trenberth, Scott Power, and one anonymous reviewer, and the editor, Dr. James Renwick, are gratefully acknowledged.

REFERENCES


