Cloud and Radiative Balance Changes in Response to ENSO in Observations and Models

CLAIRE RADLEY AND STEPHAN FUEGLISTALER

Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey

LEO DONNER

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey

(Manuscript received 11 June 2013, in final form 20 December 2013)

ABSTRACT

The authors use observations and four GFDL AGCMs to analyze the relation between variations in spatial patterns and area-averaged quantities in the top-of-the-atmosphere radiative fluxes, cloud amount, and precipitation related to El Niño over the period 1979–2008. El Niño is associated with an increase in tropical average sea surface temperature of order +0.1 K (with a maxima of +0.5 K), large local anomalies of +2 K (maxima +6 K), and tropical tropospheric warming of +0.5 K (maxima +1 K). The authors find that model-to-observation biases in the base state translate into corresponding biases in anomalies in response to El Niño. The pattern and amplitude of model biases in reflected shortwave (SW) and outgoing longwave radiation (OLR) follows expectations based on their biases in cloud amount: models with a positive cloud amount bias, compared to observations, have too strong local responses to El Niño in cloud amount, SW, OLR, and precipitation.

Tropical average OLR increases in response to El Niño in observations and models [correlation coefficients (r) with Niño-3.4 index in the range 0.4–0.6]. Weaker correlations are found for SW (r: −0.6 to 0), cloud amount (r: −0.2 to +0.1), and precipitation (r: −0.2 to 0). Compositing El Niño events over the period 2001–07 yields similar results. These results are consistent with El Niño periods being warmer due to a heat pulse from the ocean, and a weak response in clouds and their radiative effect. These weak responses occur despite a large rearrangement in the spatial structure of the tropical circulation, and despite substantial differences in the mean state of observations and models.

1. Introduction

Cloud feedbacks are known to be one of the largest sources of uncertainty in climate change predictions. A comparison of GCMs indicates that the cloud feedback amplifies the earth's global mean temperature response to forcing by 10%–50%, depending on the GCM (Bony et al. 2006). This spread in cloud feedback is not a new issue, it is a long-standing problem that is crucial to our ability to make accurate climate change predictions (Stephens 2005).

Reliable detection of climate change signals in the observational record is problematic. Hence, in this work we focus on natural variability, in particular the El Niño–Southern Oscillation (ENSO) as the ENSO pattern involves large changes in cloud distribution, precipitation, reflected shortwave (SW), and outgoing longwave radiation (OLR). ENSO is the dominant form of interannual variability in the tropics, occurring on a time scale of 2–7 yr. The short time scale and large amplitude of the signal makes it an ideal natural forcing to focus on as we have observations of multiple events, which can be used to study the observations and models with greater certainty than given by trend analysis over the observational period.

The warm phase of this phenomenon, El Niño, is associated with large sea surface temperature (SST) increases of up to 2°C in the eastern equatorial Pacific. Tropically averaged SSTs rise by as much as 0.5°C during a large El Niño event, thus providing a significant forcing on the atmosphere, which responds in a complex number of ways. The primary atmospheric changes

Corresponding author address: Claire Radley, Princeton University, 300 Forrestal Rd., Princeton, NJ 08540.
E-mail: cradley@princeton.edu

DOI: 10.1175/JCLI-D-13-00338.1

associated with an El Niño event include a weakening of the Walker circulation and an eastward shift of the western Pacific deep convection zone (Collins et al. 2010). In this paper we examine the relationship between radiation, cloud distribution, precipitation, and ENSO by examining a variety of observational datasets and model outputs. The models used are the Geophysical Fluid Dynamics Laboratory (GFDL) atmosphere-only models, AM2 [used in the Coupled Model Intercomparison Project phase 3 (CMIP3)], the latest generation AM3, and the High Resolution Atmospheric Model (HiRAM) in two different horizontal resolutions, c180 and the higher resolution c360. AM3 and HiRAM have both been submitted to CMIP5.

2. Data

The time period of the model runs is 1979–2008 and all models are forced with SSTs from the Hadley Centre Sea Ice and SST (HadISST) dataset (Rayner et al. 2003). AM2 has a resolution of 2° latitude × 2.5° longitude, with 24 vertical levels, and uses the relaxed Arakawa–Schubert convection scheme (GFDL Global Atmospheric Model Development Team 2004). AM3 (described in depth by Donner et al. 2011) uses a cubed sphere implementation of the finite-volume dynamical core. It has the same horizontal resolution as AM2 but with 48 vertical levels, and uses Donner deep convection (Donner 1993; Donner et al. 2001) and Bretherton shallow convection parameterizations (Bretherton et al. 2004). It includes mesoscale updrafts and downdrafts, which leads to more detrainment in the midtroposphere than seen in AM2. The AM3 cloud microphysics are also substantially different as they include aerosol activation and cumulus-scale vertical velocities.

HiRAM c180 and c360 have greater horizontal resolutions, with 180 and 360 latitudinal grid points between the equator and pole, respectively. This corresponds to a resolution of about 0.5° and 0.25°. HiRAM (described in Zhao et al. 2009) also uses a cubed sphere, has 32 vertical levels, and convection based on Bretherton's shallow convection parameterization. The prognostic cloud fraction scheme used in AM2/AM3 has been replaced by a simpler diagnostic scheme assuming a subgrid-scale distribution of total water.

Observations of the top-of-the-atmosphere (TOA) radiation budget were provided by the Earth Radiation Budget Experiment (ERBE; Harrison et al. 1990) and the Clouds and the Earth's Radiant Energy System (CERES; Wielicki and Barkstrom 1996). Global ERBE S-4G data are available from February 1985 to May 1989, while tropical (i.e., 30°S–30°N) ERBE S-4G data are available for the longer period of November 1984 to February 1990. We have used the Energy Balanced and Filled (EBAF) version Ed2.6 of CERES, which accounts for uncertainties in absolute calibration of the measurements and the slight imbalance in net TOA flux of the current climate, using independent observations of changes in ocean heat content. These data cover the period March 2000 to December 2010.

Data from the International Satellite Cloud Climatology Project (ISCCP) were used for the cloud amounts (Rossow and Schiffer 1991). ISCCP D2 data are available from July 1983–June 2008. Following ISCCP, the cloud types are broken into three different layers based on their height. Low refers to clouds below 680 hPa, mid to clouds in the range 680–440 hPa, and high clouds are above 440 hPa. The anomaly is calculated as the absolute increase or decrease (i.e., an increase from 40% to 60% in cloud amount would be shown as an anomaly of 20%).

In addition to the model outputs of cloud amount, AM3 was also run with an ISCCP simulator (Webb et al. 2001), and the output from this is more directly comparable to the ISCCP data provided. The main advantage of the simulator is the exclusion of optically thin clouds, which ISCCP cannot detect, as well as low—midlevel clouds that are obstructed by overlying cloud decks. For the remainder of this paper the AM3 output with the ISCCP simulator applied will be notated as AM3(sim) and the pure model output simply as AM3. The other models, unfortunately, do not include an ISCCP simulator product.

Rainfall data were provided by the Global Precipitation Climatology Project [GPCP from the National Oceanic and Atmospheric Administration/Office of Oceanic and Atmospheric Research/Earth System Research Laboratory/Physical Sciences Division (NOAA/OAR/ESRL PSD), Boulder, Colorado]. [These data are available online at their website http://www.esrl.noaa.gov/psd/ and are described in Adler and Huffman (2003).]

3. Results

a. *Tropically averaged time series*

Figure 1 shows the tropical mean seasonal cycle for OLR, SW, and total cloud amount for both observations and models. The absolute values of the model radiation and cloud amounts are offset from observations by various degrees. In the OLR and SW case, the models are in closer agreement with ERBE observations (dashed red line) than CERES (solid red line), as ERBE observations were used in the development of these models. There are also notable differences in the total cloud amount in Fig. 1c, with AM3 having more, and the remaining models having less than ISCCP.

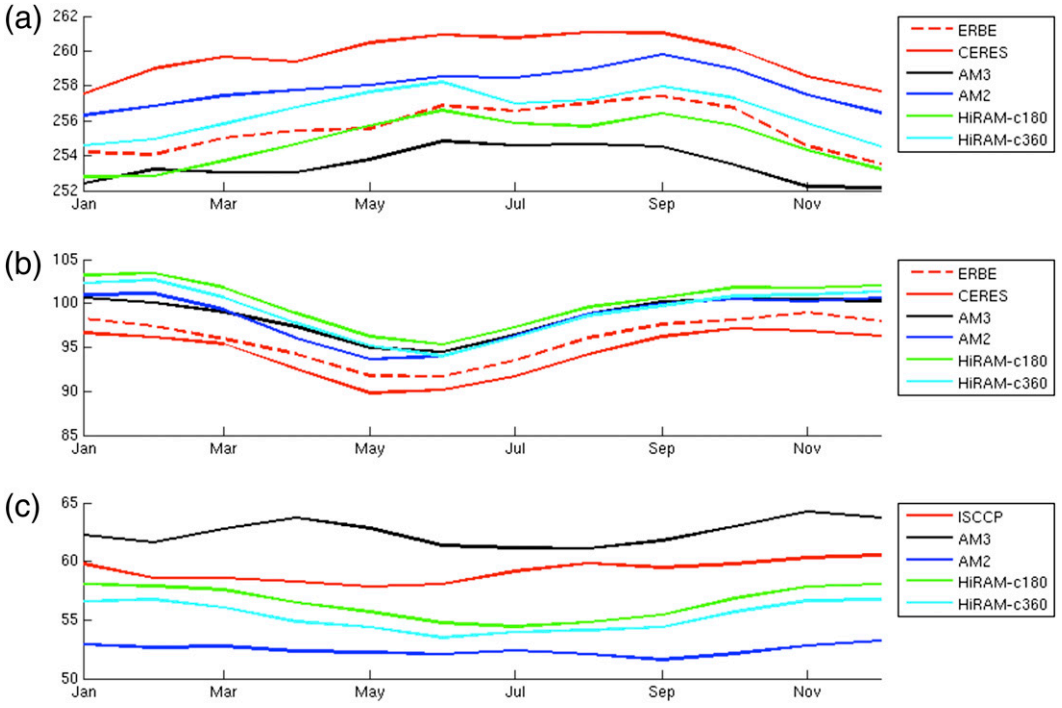

FIG. 1. The seasonal cycle for tropically averaged (a) OLR (W m^{-2}), (b) SW (W m^{-2}), and (c) total cloud amount (%) for observations and models. Values are calculated over the full length of each dataset.

Figure 2 shows the tropically averaged OLR, SW, and total cloud amount anomalies plotted as a function of time. The anomalies are calculated with respect to the climatological averages, and a 5-month running mean is applied. Overall, there is reasonable agreement between the models and observations for OLR and SW (i.e., ERBE and CERES for their respective periods), with the largest differences for OLR over the ERBE period. Wielicki et al. (2002) report a systematic underestimation of model variability on multiyear time scales, a result primarily based on comparison against ERBE nonscanner data in the 1990s. Comparison against the post-2000 period of CERES data does not show such a disagreement. Of particular relevance here is the large discrepancy reported by Wielicki et al. (2002) for the 1997/98 El Niño. For that particular El Niño period, Wielicki et al. (2002) show ERBE nonscanner data to have a maximum OLR anomaly of roughly 8 W m^{-2} (i.e., about 4–8 times larger than the responses to the weaker El Niños observed by CERES). The models shown in their paper have a maximum OLR anomaly around 2 W m^{-2} for this event, similar to the those of the model calculations shown here in Fig. 2.Wong et al. (2006) have published updated ERBE nonscanner OLR anomalies for the 1997/98 El Niño of about 5 W m^{-2}, the reduction being a result of calibration adjustments. Hence, at this point it is not clear to what extent this mismatch is due to model errors and observational uncertainty.

In the post-2000 period examined in this paper, however, there is good agreement between models and CERES observations.

The two areas shaded in gray indicate a 2-yr period after El Chichón and Mt. Pinatubo erupted. As seen in Fig. 2b, the eruptions significantly affect the SW and are therefore left out of the linear regression in section 3b. AM2 is the only model that does not have a strong SW response, as volcanoes were not included in this particular run. Given that we remove these time periods from our results and focus primarily on the ERBE and CERES periods, which are free from eruptions, we believe AM2's lack of eruptive volcanoes will not be problematic.

The light red shaded periods in Fig. 2 indicate the El Niño periods (based on the Niño-3.4 index criterion; Trenberth 1997), with the darker shading indicating the three events during the CERES period that will be used for much of the analyses to follow.

The cloud anomaly time series (Fig. 2c), shows a long-term trend in ISCCP that is absent in the models. This trend is an artifact of the satellite viewing geometry (Evan et al. 2007), and will be discussed further in section 3c.

b. *Tropically/globally averaged linear regression*

In this section we analyze the relationship of tropical (defined as 30°S–30°N) and global mean quantities with the Niño-3.4 index by means of a linear regression. We

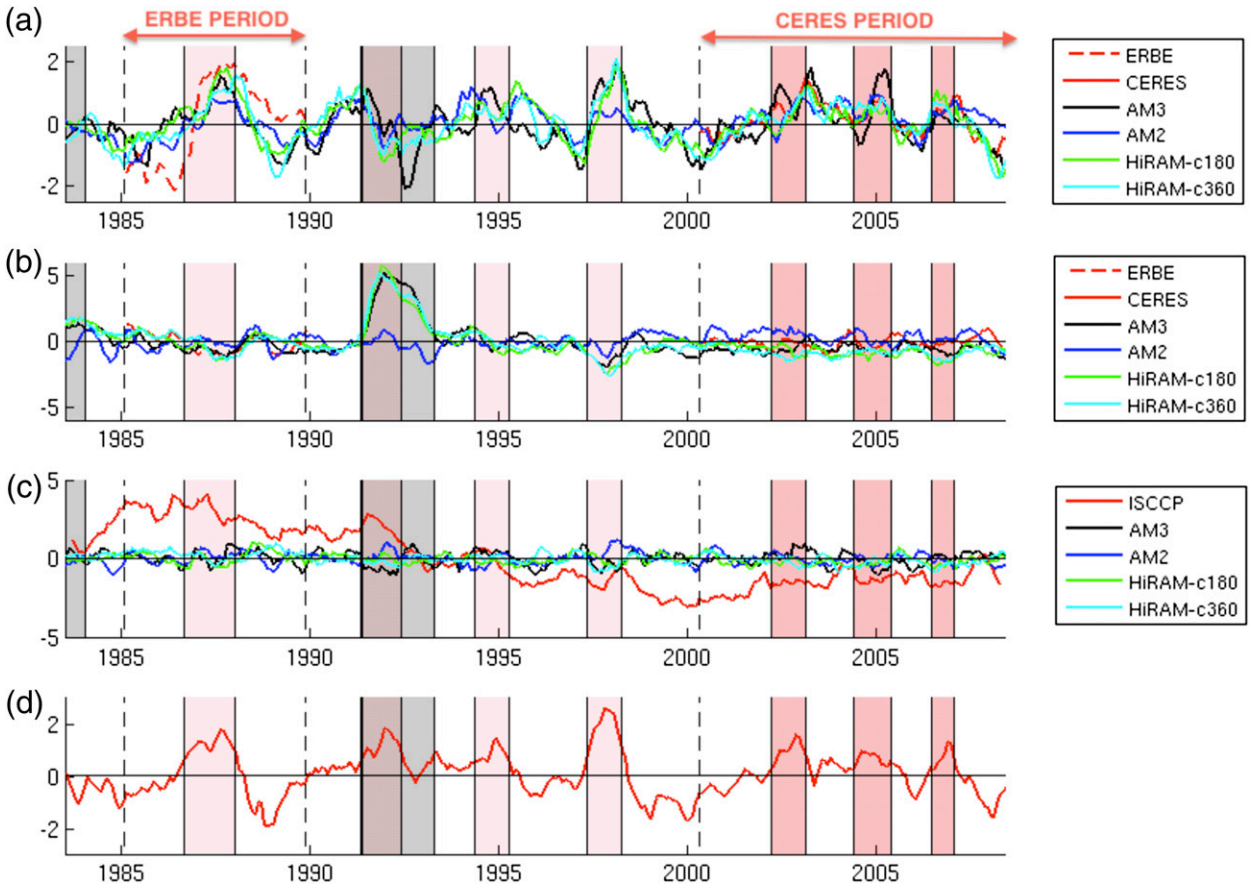

FIG. 2. The time series of tropically averaged anomalies for (a) OLR (W m^{-2}), (b) SW (W m^{-2}), and (c) total cloud amount (%). (d) The Niño-3.4 index (°C). Periods affected by volcanic eruptions are shaded gray and El Niño periods are shaded red.

linearly regress the averaged TOA radiation budget against the Niño-3.4 index, which is defined as the SST anomaly calculated over the region 5°N–5°S, 120°–170°W. The OLR and SW are first de-seasonalized by removing the corresponding monthly-mean value. They are then regressed against the Niño-3.4 index for all months over the various time periods, and the correlation coefficients are calculated. The 2 years of data after the volcanic eruptions, El Chichón and Mt. Pinatubo, have been removed because of their strong effect on the radiation budget. Three time periods are considered. The first is February 1985–May 1989, which corresponds to the ERBE global observation period. The second is January 1979–December 2008, which is the length of the model runs, and the third is March 2000–June 2008, which corresponds to the CERES time period overlap with ISCCP.

The three time periods considered are shown on the x axis in Fig. 3, with the correlation coefficient on the y axis. Correlations that are significant at the 95% level are plotted with larger markers. We find that for data analyzed here, the slope of the linear regression (not shown) is generally linearly related to the correlation coefficient. This result is explained when considering the ratio of the slope (β) to the correlation coefficient (ρ), which is given by

$$\beta = \frac{\sigma_{\text{nino3.4}}}{\sigma_{\text{radiation}}}\rho, \tag{1}$$

where $\sigma_{\text{nino3.4}}$ is the standard deviation of the Niño-3.4 index, and $\sigma_{\text{radiation}}$ is the standard deviation of the radiation. As the monthly Niño-3.4 index is model independent, its variance is constant and the slope is proportional to the correlation coefficient divided by the variance in radiation. For the studies here, $\sigma_{\text{radiation}}$ is similar for different models and observations, which explains the linear relationship between the β and ρ. Quantitatively, $\beta \approx 1\rho$ °C (W m^{-2})$^{-1}$ in the tropics and $\beta \approx 0.5\rho$ °C (W m^{-2})$^{-1}$ for globally averaged quantities.

Figure 3a shows the correlation of the tropical OLR with the Niño-3.4 index to be positive for all datasets and time periods considered. This indicates that during an El Niño event there is an increase in the OLR. The

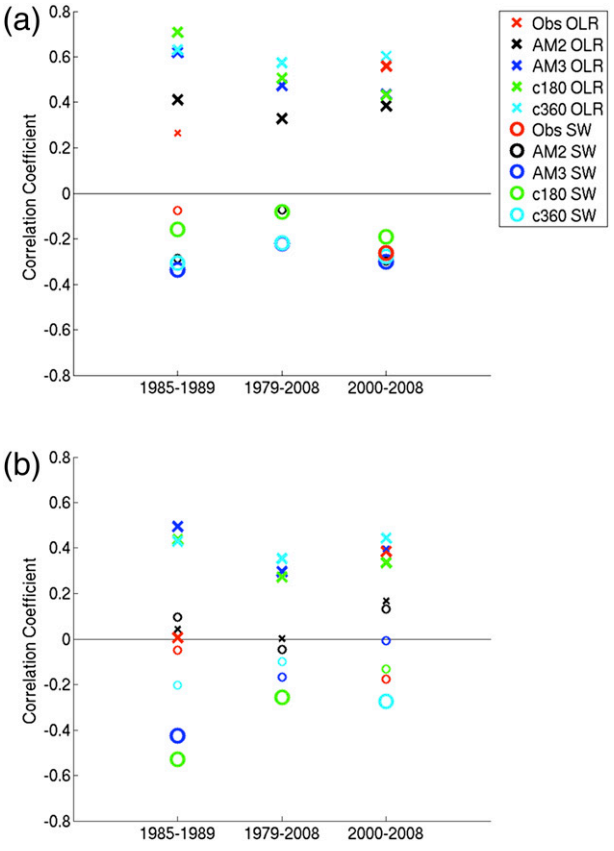

FIG. 3. Correlation coefficient for (a) tropically and (b) globally averaged OLR and SW when linearly regressed against the Niño-3.4 index. The larger markers indicate significance at the 95% level.

magnitude of the correlation is smallest for ERBE and AM2. For the CERES period, the correlations for the other models and the CERES EBAF data are similar, with a value of about 0.6.

In the SW, the models have a larger spread than in the OLR, especially when we consider the 2000–08 time period results. As we are only considering the tropics here, the spread is attributable to differences in clouds. For the 2000–08 period, CERES SW observations show a small correlation of only −0.3 with the Niño-3.4 index, which means that there is only slightly less reflected SW (i.e., an increase in SW absorption) during an El Niño event.

For the case of global OLR regressed against Niño-3.4 (Fig. 3b), there is a weaker relationship across the board. The correlation coefficient is approximately half that of the tropically averaged values, which implies that the extratropical response does not contribute significantly to the global mean. There may also be out-of-phase contributions from the extratropics that reduce the global correlation coefficients.

The SW also has a smaller correlation in the global mean and the spread is centered closer to zero, with very few correlations significant at the 95% level. This indicates that the SW response to an El Niño event is not as robust as that of the OLR. We therefore conclude that from a tropical or global average point of view, ENSO events have a higher correlation and stronger response in the OLR than the SW. As discussed later however, this statement does not hold for local effects, which can be very large in both the OLR and SW due to changes in the locations of deep convection.

Our calculations show the linear regression of global average surface temperature against the Niño-3.4 index reaches maximum correlation at a lag of 3 months, in rough agreement with the 4-month lag found by Trenberth (2002). Similarly, the tropical average surface temperatures reach maximum correlation at a 1-month lag. These lags in maximum correlations are not sharply peaked, however, and make only small differences in either the linear regression slope or the correlation coefficient for TOA radiation.

Given El Niño's seasonal phase locking, where the strongest events occur at the end of the calendar year, we also performed the linear regression for just December–January–February (DJF) months. The tropical correlation coefficient for OLR (for DJF only) increases compared to the full annual calculation, with values of 0.8 in the 2000–08 time period. The same is true for the global correlation coefficient, with values as large as 0.6. This increase can be explained by the variance of the Niño-3.4 index, which is larger in DJF than in the annual mean, and therefore accounts for a larger fraction of the OLR variance in DJF (assuming background noise is roughly constant). This then results in a slightly larger correlation coefficient during DJF than when the regression is performed across all months. There is no consistent change in the correlation coefficients for the SW across the three time periods.

c. *Differences between ERBE and CERES observations*

Given the differences between the ERBE and CERES results in sections 3a and 3b, we investigate these two datasets more thoroughly. Tables 1 and 2 show the tropical averages of the ERBE and CERES periods, for full-year periods in order to eliminate seasonal biases. CERES has a mean tropical OLR 4.5 W m^{-2} larger than that of ERBE, and a mean tropical SW 1.5 W m^{-2} smaller than ERBE. Given the fundamental differences in the ERBE and CERES orbits and postprocessing algorithms, we do not attempt to explain the differences in the absolute values of tropical mean OLR–SW (see Raschke et al. 2012 for a detailed discussion). Given our

TABLE 1. Climatological mean of tropical OLR, LWCRF, SW, and SWCRF ($W m^{-2}$) for the January 1985–December 1989 period.

	OLR	LWCRF	SW	SWCRF
ERBE	255.46	32.38	96.03	−43.90
AM3	253.41	29.88	98.25	−47.75
AM2	257.74	24.36	98.19	−48.84
c180	254.89	27.36	100.07	−49.49
c360	256.37	25.51	99.23	−48.73

TABLE 2. As in Table 1, but for the January 2001–December 2007 period.

	OLR	LWCRF	SW	SWCRF
CERES	259.93	27.42	94.55	−45.33
AM3	253.72	23.29	98.04	−47.51
AM2	258.24	22.72	98.83	−49.48
c180	255.01	22.06	99.40	−48.82
c360	256.76	21.72	98.34	−47.84

interest in spatial patterns related to ENSO, however, the spatial patterns of the differences between ERBE and CERES deserve some attention. Figure 4a shows the CERES minus ERBE field for OLR after removal of the area average offset of $+4.5\,W m^{-2}$, Fig. 4b shows the corresponding difference for SW after removal of the area average offset of $-1.5\,W m^{-2}$.

For large parts of the tropics (e.g., Africa south of the equator, the Amazon region, and the South Pacific convergence zone and ITCZ over the Pacific), there is actually a remarkable anticorrelation between the changes in OLR and SW. Such an anticorrelation is expected for changes in the locations of deep convective clouds (which reduce OLR and increase reflected SW). The OLR–SW changes mentioned are of order $10\,W m^{-2}$. This is considerably larger than the CERES OLR–SW standard deviation of the annual mean, which has maximum values of $1\,W m^{-2}$ in the tropics.

For compactness, the equivalent of Fig. 4 for the four models is not shown. The spatial pattern is remarkably similar to the observations, but there are differences in magnitude. The variance in Fig. 4a for the observed OLR is 8.5 $(W m^{-2})^2$, whereas the variance for AM3, for example, is 18.9 $(W m^{-2})^2$.

Since ISCCP spans both periods, one may wonder to what extent that data can provide additional insights, specifically also whether the changes in OLR and SW pattern are consistent with ISCCP cloud changes.

Figure 4c shows the differences in ISCCP total cloud amount between the two periods, and Fig. 4d shows the differences in ISCCP high cloud amount. Comparison between Figs. 4a,b,d shows that indeed the differences in OLR–SW in the eastern equatorial Pacific sector have the expected counterpart of an increase in high clouds north of the equator. Much of the differences in ISCCP cloud occurrence, however, is quite obviously an artifact

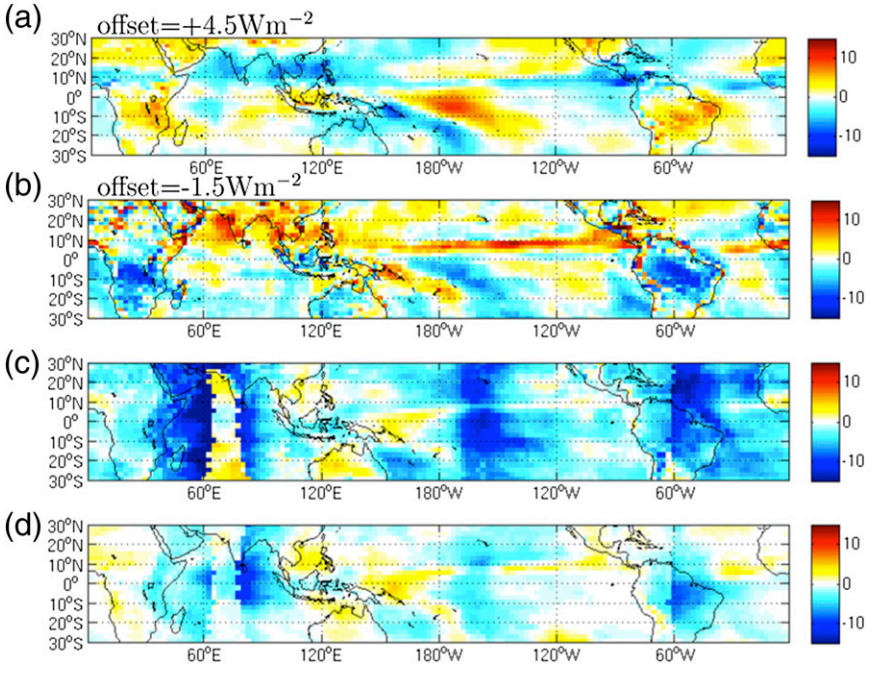

FIG. 4. CERES minus ERBE mean (a) OLR (with $4.5\,W m^{-2}$ offset removed) and (b) SW (with $-1.5\,W m^{-2}$ offset removed). Also shown is the difference in ISCCP mean (c) total and (d) high cloud amount (%) for the CERES minus ERBE. Only full-year periods are used.

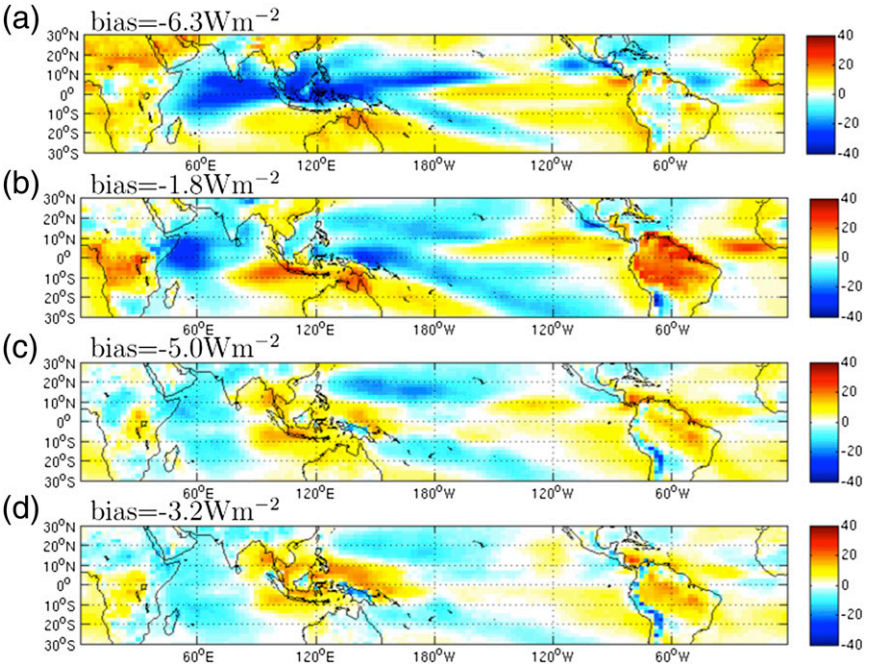

FIG. 5. Climatological mean OLR ($W m^{-2}$) in models minus CERES observations for (a) AM3, (b) AM2, (c) HiRAM-c180, and (d) HiRAM-c360. The biases are shown at the top of each panel.

arising from the major satellite changes, as described by Evan et al. (2007) and Knapp (2008), and prevents any clear conclusions.

d. Base-state regional results

The tropical average base-state OLR and SW are shown in Table 1 and 2 for CERES and the four models. The largest differences between CERES and the models are $6 W m^{-2}$, with smaller differences between the models and ERBE, which is not surprising given that the model development was influenced by the ERBE data. A comparison between unbiased rms differences and correlations against CERES and ERBE show that AM3 agrees better with observations for SW than AM2 does (Donner et al. 2011), while for OLR the two models perform similarly.

Figure 5 shows the regional OLR difference, with the area-averaged bias removed, between the models and the CERES observations over the 2001–07 time period. The bias for each model is calculated from the base-state values in Table 2. Despite the identical SSTs there are strong regional differences in the models. The western warm pool and Amazon region are persistent problem areas, particularly for AM2 and AM3. For example, the bias in the western Pacific in AM3 relative to CERES is as large as $-36 W m^{-2}$. The biases are smaller in HiRAM, with little influence of the models' spatial resolution as HiRAM c180 and c360 are very similar.

The same analysis is performed for the SW in Fig. 6. Similar to OLR, large differences are seen in the western warm pool, but the maxima are found off the western coasts of North America, South America, and Africa. The bias is greatest off the coast of South America in AM2 and AM3, with a magnitude of $-70 W m^{-2}$, and the HiRAM models are only slightly better with biases of order $-60 W m^{-2}$. These regions are dominated by low-level stratus decks. The models' negative bias, relative to CERES, indicates that there are either too few clouds in this region or that they are not sufficiently bright so as to reproduce the observed albedo (see Fig. 6). This strong local SW bias is also the dominant signal when we examine the net TOA radiation budget. This is because the western Pacific has large local biases in both the SW and OLR, which cancel in the net to give reasonable agreement with observations. There is no such local cancellation effect when considering low-level stratus clouds, as they only have a very small influence on the OLR, but a large influence on the SW. The small net SW bias of the tropical average (e.g., $3.5 W m^{-2}$ for AM3), implies that the large local biases are compensated by biases of opposite sign in other parts of the domain.

Analysis of high cloud amount shows that this explains the large radiation anomalies in the western Pacific, as AM3(sim) has roughly 1.7 times the high cloud amount

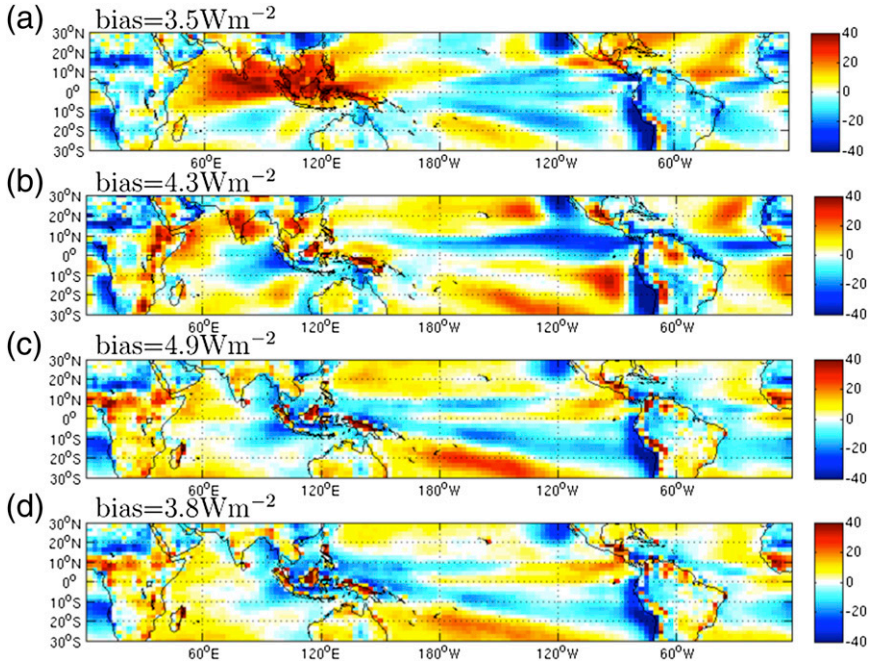

FIG. 6. As in Fig. 5, but for the SW.

reported by ISCCP. The other models also appear to have significantly larger high cloud amounts than ISCCP, as shown in Table 3, although it is difficult to say exactly how much without the use of an ISCCP simulator. Comparison between AM3(sim) and AM3 shows that the ISCCP method fails to detect about 7% of the model clouds. The equivalent number for observations is not known.

It is more difficult to make comparisons between the observations and models for mid- and low-level cloud amounts. Although we have the ISCCP simulator for AM3, the lower level clouds will be partially obscured by AM3's excessive high cloud amount. This can be seen in the differences between AM3(sim) and AM3 for the mid- and low-level cloud amounts (Table 3), where the cloud fraction is significantly reduced when the ISCCP simulator is applied. However, given that all model outputs have smaller values than ISCCP we can state that there are too few midlevel clouds in the models. Exactly what this difference is though is difficult to say. A similar scenario is likely true for the low-level clouds, but of course there is now masking from both mid- and high-level clouds (with the exception of coastal regions where low-level clouds prevail), so no definitive conclusions can be made.

The high cloud base states for ISCCP and the four models are shown in Fig. 7. ISCCP has a maximum high cloud amount in the western Pacific of ~30%, whereas AM3(sim) shows amounts of ~50%. We have broken down the high cloud fraction into three optical depth (τ) categories—thin ($\tau < 3.6$), medium ($3.6 < \tau < 23$), and thick ($\tau > 23$)—as defined by ISCCP, and averaged the cloud amounts over the latitudinal band 30°S–30°N. While AM3(sim) overestimates the high cloud fraction across all optical depths, the difference is largest for the thin clouds with an overestimation of 10%–15% cloud amount. The difference for the high clouds of medium optical thickness is 5%–10% and <5% for the optically thick clouds. Also shown on this figure are the total high cloud amounts for all models, without being run through an ISCCP simulator. The overestimation of high clouds is not isolated to the western Pacific, and we see high cloud amount larger than ISCCP in all four models across the entire tropical Pacific. In the eastern Pacific the effect on the radiation budget, however, is not as large because the absolute values of cloud amount are

TABLE 3. Tropical average cloud base-state amount and anomalies (%) during El Niño events for the January 2001–December 2007 period.

	High	High anom	Mid	Mid anom	Low	Low anom
ISCCP	24.25	0.12	12.24	−0.09	20.87	−0.13
AM3(sim)	40.41	0.05	6.37	−0.08	13.29	−0.07
AM3	47.26	0.06	9.04	−0.11	24.62	−0.07
AM2	31.83	−0.06	10.80	0.01	25.34	−0.00
c180	39.85	−0.05	6.07	−0.06	26.55	−0.07
c360	38.15	−0.12	5.24	−0.01	26.76	−0.06

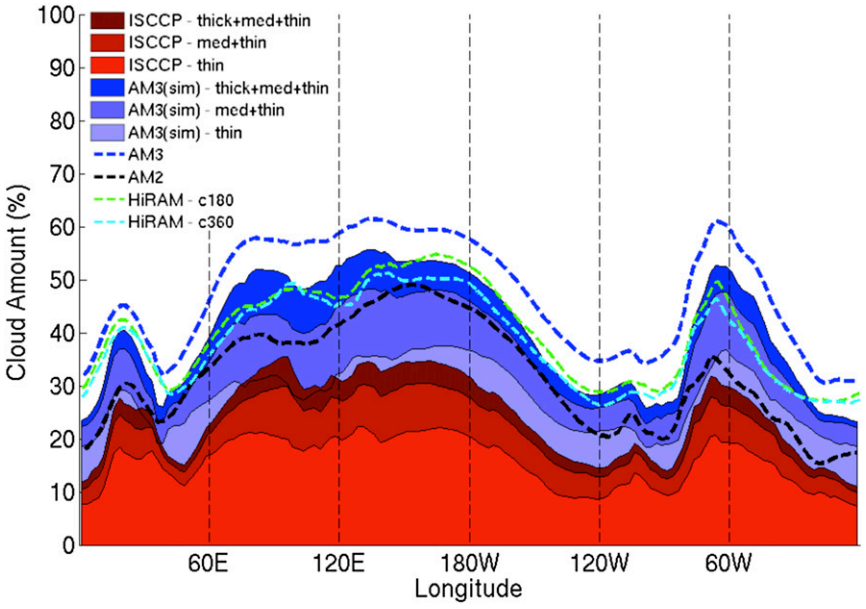

FIG. 7. Climatological mean high cloud amount (%) for ISCCP and AM3 with ISCCP simulator. The values are meridionally averaged from 30°S to 30°N and plotted as a function of longitude for three different optical depth categories: thin, medium, and thick. Dashed lines are total high cloud amounts from AM3, AM2, HiRAM-c180, and HiRAM-c360.

lower in both observations and models. As mentioned above, applying the simulator to AM3 decreases the high cloud amount by roughly 7%, and while AM2 and HiRAM still have more high cloud than ISCCP they are in closer agreement with observations than AM3.

Figure 8 indicates this large high cloud bias of models against ISCCP is partially compensated for by a bias in the midlevel clouds. The ISCCP base state has ~15% midlevel cloud amount in the western Pacific; however, AM3 with the ISCCP simulator shows 10%. Again, this difference is primarily due to thin clouds, followed by clouds of medium optical depth and negligible contribution by the very thick clouds, and extends across the tropical Pacific. When the ISCCP simulator is applied to

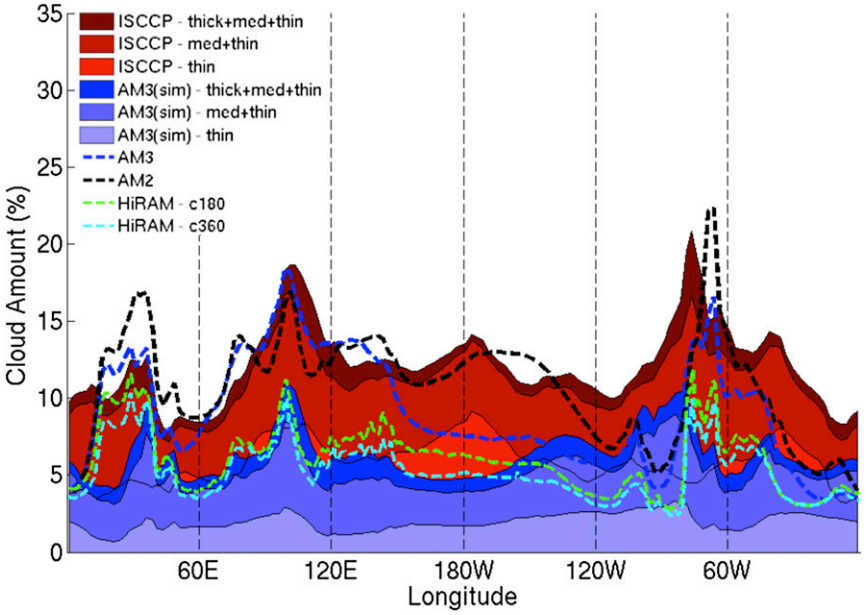

FIG. 8. As in Fig. 7, but for midlevel clouds.

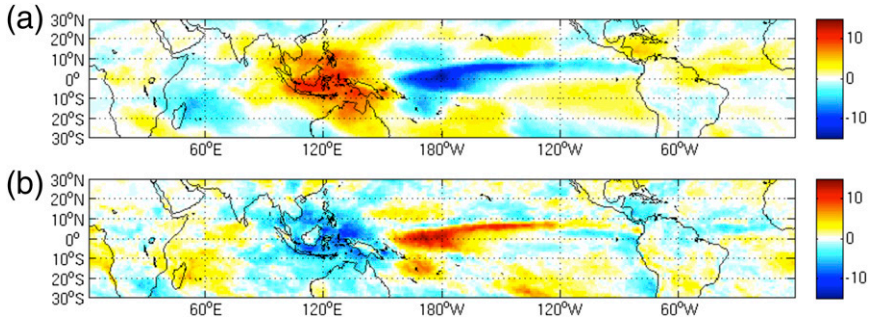

FIG. 9. Composite of CERES (a) OLR and (b) SW anomalies (W m^{-2}) during three El Niño events during 2001–07. Anomalies are defined as the difference from the CERES EBAF 2001–07 climatological mean.

AM3 for midlevel clouds there is a decrease in cloud amount. This difference between the AM3 and AM3 (sim) case is larger than for the high cloud case, and is presumably caused by the masking effect of higher-level clouds.

While these results may suggest AM2 cloud amounts are in better agreement with ISCCP observations, Klein et al. (2013) used an ISCCP simulator to show that AM3 is in fact an improvement over AM2 for both cloud properties and SW-relevant cloud properties.

The large SW bias of models against CERES off the coast of South America is consistent with the bias in the low cloud amounts of models relative to ISCCP. ISCCP observations have roughly 10% more low cloud amount than is shown by AM3 with the ISCCP simulator (figure not shown). ISCCP also has significantly more low cloud amount across the entire tropics than is shown by AM3 with the simulator. The bias is smaller for AM2 and HiRAM, since these models have slightly larger low cloud amount than AM3, but lack of ISCCP simulator data products disallows quantification.

e. *Local TOA radiation anomalies*

In contrast to the tropical average values, local anomalies frequently have a nonmonotonic relationship with the Niño-3.4 index, rendering the linear regression approach inadequate. Instead, periods of El Niño events are examined, which are defined to be when the 5-month running mean of the Niño-3.4 index exceeds a magnitude of 0.5°C for at least 6 months (Trenberth 1997). Because of the incompletely understood differences between ERBE and CERES, discussed in section 3c, we restrict this analysis to CERES data. Figure 9 shows a composite of the OLR/SW anomalies over the three El Niño events of the CERES time period. These events occur during April 2002–March 2003, June 2004–June 2005, and July 2006–February 2007 (see Fig. 2). For these periods, we calculate the PDFs of the anomalies for the quantities of interest. The anomalies are defined as the difference between the monthly-mean outgoing TOA radiation and the climatological mean over the CERES period. There will be some element of a seasonal bias as we are combining odd numbers of winter and summer months. However, since we focus on comparisons of observations to models over identical periods, the preference for certain seasons is not problematic.

Figure 10a shows the probability distributions of local OLR anomalies over the tropics for CERES and the models for all three El Niño events. We note a tendency for the models to have a longer negative tail than the CERES data, at the expense of small anomalies between -10 and $+10$ W m^{-2}, and a longer positive tail for AM3. Inspection of the maps of the anomalies shows that the large tail of the negative anomalies in AM3, for example, arises from the central Pacific where AM3 has local anomalies in OLR of up to -99 compared to -56 W m^{-2} for CERES. The PDF of the clear-sky OLR anomalies (not shown) shows very good agreement between observations and all models, such that the differences in the PDFs of all-sky OLR is a consequence of the models' clouds, in particular, upper-tropospheric convective clouds.

The corresponding SW PDF is shown in Fig. 10b. Even more pronounced than with OLR, the models have larger tails than CERES, especially when we compare the probability of large negative anomalies, which are predominately found in the western Pacific. The large negative anomalies are due to an eastward shift in convection, which decreases the high cloud amount and consequently the reflected SW in the western Pacific. These negative SW anomalies are larger in the models than in observations because the models overestimate the optically thick cloud amount in this region, relative to ISCCP (see Fig. 7). Overall, the variance of the anomalies during El Niño events is larger in models than observations. As discussed in section 3c, a similar result

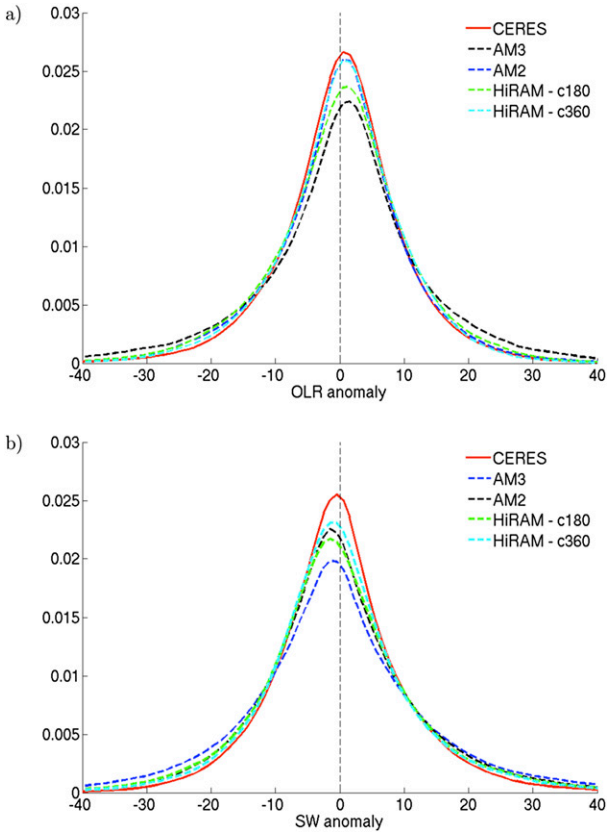

FIG. 10. Probability distribution for (a) OLR and (b) SW anomalies ($W m^{-2}$) for CERES, AM2, AM3, HiRAM-c180, and HiRAM-c360 during 2001–07 El Niño events.

was also found for the variance in the difference between the average ERBE and average CERES period, which was also larger in models than observations.

The tropical average OLR anomalies for El Niño events are shown in Table 4. There is very good agreement between CERES and the models, perhaps with the exception of AM2, which has a weak response. Longwave cloud radiative forcing (LWCRF), defined as the clear-sky minus all-sky values, shows the increase in OLR is no more than 50% due to a reduction in LWCRF. The remaining increase in OLR is therefore due to an increase in the clear-sky OLR. The tropical average SW anomalies are also shown in Table 4. The SW anomalies in models vary in both sign and magnitude as we might expect from the weak correlations shown in Fig. 3, but are within 0.3 $W m^{-2}$ of the CERES SW anomaly.

The generally good agreement of tropical average El Niño response in OLR and SW may seem surprising given the differences in the PDFs of local anomalies. The implication is that apparently the positive and negative *differences* in the PDFs largely cancel, which is

TABLE 4. Tropical average OLR, LWCRF, SW, and SWCRF ($W m^{-2}$) anomalies during El Niño events for the January 2001–December 2007 period.

	OLR	LWCRF	SW	SWCRF
CERES	0.35	−0.20	−0.04	0.10
AM3	0.29	0.05	−0.06	0.05
AM2	0.10	−0.03	−0.01	−0.00
c180	0.18	−0.11	−0.25	0.24
c360	0.23	−0.08	−0.11	0.10

remarkable in particular for the SW where the CERES and models PDFs differ quite substantially. In the following, we analyze this result in more detail.

f. Scaling of anomalies with the base state

We hypothesize that the differences in the PDFs of the anomalies are related to differences in the base state of models and observations. For example, if OLR and SW are linearly related to cloud amount and total cloud amount was conserved, the same local fractional change in cloud amount would result in a stronger OLR and SW response for the case with a larger base-state cloud amount. It also follows that for this scenario, differences in the redistribution of cloud amount in space give differences in the shape of the PDF of the anomalies, but that the sum over all PDFs necessarily evaluates to zero.

In reality, there is not a simple linear relationship between cloud amount and the cloud radiative effect, and explaining the behavior of the SW and OLR anomaly PDFs is not straightforward. However, we find that this reasoning explains the response in high cloud amount remarkably well. Figure 11a shows the probability distribution of high cloud amount during the three El Niño periods for ISCCP, AM3(sim), and the four model outputs. As in the case of SW and OLR, the models' high cloud anomaly PDFs have longer tails. The models also have larger high cloud amounts in the base state (section 3d) and, everything else being equal, one would expect the high cloud anomalies to scale like the ratio of the base-state cloud amount. From Table 3, we calculate the high cloud amount ratio for each model relative to ISCCP and divide the models high cloud anomalies with this factor. For example, for AM3(sim) we get a ratio of 40.41/24.25 ~ 1.7 that we divide with the AM3(sim) high cloud anomalies. The resulting PDFs are shown in Fig. 11b. While not perfect, it is evident that the scaled high cloud PDFs of the models are in much better agreement with the ISCCP anomalies. Quantitatively, the root-mean-square (rms) difference between AM3(sim) and ISCCP decreases from 0.049 (for the PDFs of raw anomalies) to 0.015 (for the PDFs of the scaled anomalies). The rms difference also decreases for

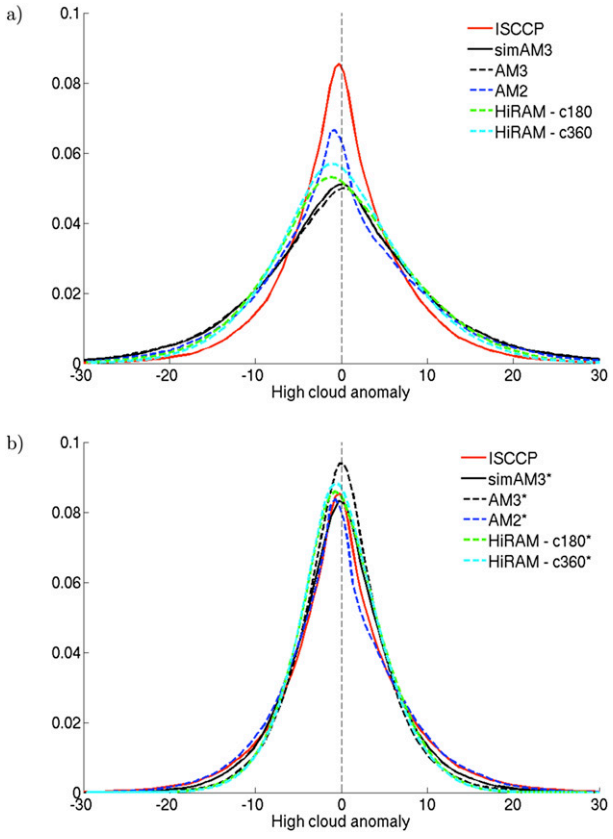

FIG. 11. PDF of (a) raw and (b) scaled high-cloud anomaly (%) during 2001–07 El Niño events for ISCCP, AM3 with and without simulator, AM2, HiRAM-c180, and HiRAM-c360. Scaling factors are taken from the biases in base-state high-cloud amounts in Table 3.

each of the 33 months considered, with the exception of 4 months that have small increases in the rms difference (figure not shown). This result implies that, at least for high cloud amount, the problem is reasonably linear, and that (i) biases in the models' base states are the primary cause of the local biases in the anomalies during El Niño (as opposed to errors in the redistribution of convection during El Niño); and (ii) that the local biases in the anomalies during El Niño necessarily largely cancel in the area average.

This scaling of anomalies cannot be used for the mid- and low-level cloud PDFs because of problems caused by high cloud masking, which make it difficult to compare ISCCP with models.

We may expect the scaling to work for precipitation, however, as convective rainfall and high cloud amount are closely related. The base-state mean precipitation values are shown in Table 5. All of the models have larger mean precipitation than GPCP data, by a factor of approximately 1.2. [It has been suggested that GPCP may underestimate global mean precipitation. Kato et al. (2011) argue that the net surface irradiances imply global mean precipitation about 10%–20% higher than GPCP, and Stephens et al. (2012) argue based on satellite observations that global mean rainfall is likely between the GPCP estimate and 25% higher.]

TABLE 5. Climatological mean precipitation (mm day^{-1}) for the January 2001–December 2007 period.

	Tropical	Global
GPCP	3.01	2.65
AM3	3.68	3.04
AM2	3.64	2.93
c180	3.58	2.97
c360	3.59	2.98

Despite the models base-state bias in precipitation, the PDFs of the models anomalies agree surprisingly well with GPCP observations (figure not shown). Therefore, scaling of the anomalies has only a small effect on the PDFs and does not significantly improve the results.

Given the close relationship between high cloud amount and precipitation, it is perhaps surprising the model precipitation base states agree better with GPCP than the model high cloud amount with ISCCP data. Consequently, the ratio of high cloud amount to precipitation will be different in models than in observations. AM3(sim) has a base-state ratio of optically thick high-level cloud to precipitation of 1.3% (mm day^{-1})$^{-1}$. ISCCP and GPCP observations put this ratio at approximately 0.9% (mm day^{-1})$^{-1}$, a lower cloud fraction per rain rate. During El Niño events the tropical average has perturbations about the base state of order ±1% of the optically thick high cloud and perturbations of order ±0.1% in precipitation. Given the unknown signs of these perturbations, which are dependent on the exact definition of the base-state period, the change in high cloud to precipitation ratio during El Niño events cannot be determined reliably.

The large local cloud anomalies average out to small anomalies in both the tropical and global average. For the high cloud amounts the largest tropical anomaly in models and observations is <0.15% and the largest global anomaly is <0.25% (as shown in Table 3). ISCCP has the smallest mean high cloud amount, but has the largest tropical and global high cloud anomaly. Thus, our scaling argument only works for the shape of the PDFs but does not explain the ratio of the tropical mean anomalies.

It is also of interest to note the tropically and globally averaged values of both LWCRF and SWCRF (Table 4). The tropically averaged LWCRF anomalies for El Niño events are no greater than 1% of the base-state value and are even less for the SWCRF. In the global average, these numbers have decreased further to <0.2%. This is a

surprising result given the large biases in the base-state cloud and radiation fields, as well as the differences in the high cloud anomaly PDF. It appears that, in both models and observations, the cloud radiative effect is rather insensitive to perturbations.

4. Summary

Regressing the observed tropical OLR against the Niño-3.4 index results in a strong positive correlation of 0.6. The corresponding correlation coefficient for the tropical SW case is −0.2. The four models analyzed were found to be in good agreement for the OLR but showed a large spread for the SW.

There are large differences between the ERBE and CERES OLR–SW observations. This is evident in not only the tropically averaged climatological values, but also in the regional differences between the two time periods. The cause of the regional differences between the two time periods is unknown. The four models, however, all exhibit similar regional differences, which suggests they are not an artifact of the satellite calibration and postprocessing, but due to a climatological difference. Given the general uncertainty surrounding cloud and water vapor feedbacks, the regional agreement between the satellite observations and the models, which, among themselves, have substantial differences in their formulation of clouds and convection, is perhaps surprising and encouraging.

Comparison of the four models with CERES and ISCCP data shows large biases in the base states. In the OLR–SW fields, the biases are large on the regional scale but agree well with observations in the tropical mean. The high cloud amount in AM3, for example, is roughly double that of ISCCP observations in the base state, thus causing local OLR biases of up to $-36\,\mathrm{W\,m^{-2}}$, particularly in the western Pacific warm pool. There are also SW biases of up to $-70\,\mathrm{W\,m^{-2}}$ in the stratus deck regions, where all four models appear to have too little low cloud amount. Exact mid–low cloud biases are difficult to determine because of the masking effect of high clouds.

The model high cloud amount bias, relative to ISCCP, is also evident in the anomalies during El Niño events. This is shown by the PDFs of anomalies having much longer tails for the models than for observations. A linear scaling approach is introduced, whereby each gridpoint anomaly is scaled by the model-to-observation tropically averaged base-state ratio. This results in the model PDFs being in much closer agreement with observations. This works for high cloud amount but not for OLR–SW, as the tropical average is in close agreement with observations.

While this scaling argument reasonably explains the differences in the shape of the PDFs, it provides little quantitative insight into the behavior of the tropical average during El Niño. For example, in the base state, AM3 has 1.7 times the high cloud amount observed by ISCCP. AM3's high cloud tropically averaged anomaly during El Niño events, however, is only 0.4 times the corresponding ISCCP high cloud anomaly.

Regardless of the base state, there is a considerable amount of symmetry and, therefore, cancellation between the large positive and negative anomalies. This results in the tropically averaged cloud amount and cloud radiative forcing anomalies being less than 0.5% and 0.2% of the base state, respectively.

As seen in sections 3c, 3d, and 3f, the models consistently have larger regional anomalies than seen in observations, even though they might have a smaller tropically averaged anomaly. This can have important implications for calculated values of ''feedback'' or ''sensitivity,'' as the models have greater sensitivity than observations locally, but smaller sensitivity in the tropical mean.

Acknowledgments. This material is based upon work supported by the Cooperative Institute for Climate Science under Grant 344-6126. SF was supported in part by DOE Grant SC0006841.

REFERENCES

Adler, R., and G. Huffman, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.,* **4,** 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Bony, S., and Coauthors, 2006: How well do we understand and evaluate climate change feedback processes? *J. Climate,* **19,** 3445–3482, doi:10.1175/JCLI3819.1.

Bretherton, C. S., J. R. McCaa, and H. Grenier, 2004: A new parameterization for shallow cumulus convection and its application to marine subtropical cloud-topped boundary layers. Part I: Description and 1D results. *Mon. Wea. Rev.,* **132,** 864–882, doi:10.1175/1520-0493(2004)132<0864:ANPFSC>2.0.CO;2.

Collins, M., and Coauthors, 2010: The impact of global warming on the tropical Pacific Ocean and El Niño. *Nat. Geosci.,* **3,** 391–397, doi:10.1038/ngeo868.

Donner, L. J., 1993: A cumulus parameterization including mass fluxes, vertical momentum dynamics, and mesoscale effects. *J. Atmos. Sci.,* **50,** 889–906, doi:10.1175/1520-0469(1993)050<0889:ACPIMF>2.0.CO;2.

——, C. J. Seman, R. S. Hemler, and S. Fan, 2001: A cumulus parameterization including mass fluxes, convective vertical velocities, and mesoscale effects: Thermodynamic and hydrological aspects in a general circulation model. *J. Climate,* **14,** 3444–3463, doi:10.1175/1520-0442(2001)014<3444:ACPIMF>2.0.CO;2.

——, and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL Global Coupled Model CM3. *J. Climate,* **24,** 3484–3519, doi:10.1175/2011JCLI3955.1.

Evan, A. T., A. K. Heidinger, and D. J. Vimont, 2007: Arguments against a physical long-term trend in global ISCCP cloud amounts. *Geophys. Res. Lett.,* **34,** L04701, doi:10.1029/2006GL028083.

GFDL Global Atmospheric Model Development Team, 2004: The new GFDL global atmosphere and land model AM2–LM2: Evaluation with prescribed SST simulations. *J. Climate,* **17,** 4641–4673, doi:10.1175/JCLI-3223.1.

Harrison, E., P. Minnis, and B. Barkstrom, 1990: Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment. *J. Geophys. Res.,* **95,** 18 687–18 703, doi:10.1029/JD095iD11p18687.

Kato, S., and Coauthors, 2011: Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties. *J. Geophys. Res.,* **116,** D19209, doi:10.1029/2011JD016050.

Klein, S. A., Y. Zhang, M. D. Zelinka, R. Pincus, J. Boyle, and P. J. Gleckler, 2013: Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator. *J. Geophys. Res.,* **118,** 1329–1342, doi:10.1002/jgrd.50141.

Knapp, K. R., 2008: Calibration assessment of ISCCP geostationary infrared observations using HIRS. *J. Atmos. Oceanic Technol.,* **25,** 183–195, doi:10.1175/2007JTECHA910.1.

Raschke, E., S. Kinne, and P. W. Stackhouse, 2012: GEWEX Radiative Flux Assessment (RFA). Vol. 1: Assessment. WCRP Rep. 19/2012, 299 pp. [Available online at http://www.wcrp-climate.org/documents/GEWEX%20RFA-Volume%201-report.pdf.]

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, and D. P. Rowell, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.,* **108,** 4407, doi:10.1029/2002JD002670.

Rossow, W., and R. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 2–20, doi:10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2.

Stephens, G. L., 2005: Cloud feedbacks in the climate system: A critical review. *J. Climate,* **18,** 237–273, doi:10.1175/JCLI-3243.1.

——, and Coauthors, 2012: An update on Earth's energy balance in light of the latest global observations. *Nat. Geosci.,* **5,** 691–696, doi:10.1038/ngeo1580.

Trenberth, K., 1997: The definition of El Niño. *Bull. Amer. Meteor. Soc.,* **78,** 2771–2778, doi:10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2.

——, 2002: Changes in tropical clouds and radiation. *Science,* **296,** 2095, doi:10.1126/science.296.5576.2095a.

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 climate models. *Climate Dyn.,* **17,** 905–922, doi:10.1007/s003820100157.

Wielicki, B., and B. Barkstrom, 1996: Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System experiment. *Bull. Amer. Meteor. Soc.,* **77,** 853–868, doi:10.1175/1520-0477(1996)077<0853:CATERE>2.0.CO;2.

——, and Coauthors, 2002: Evidence for large decadal variability in the tropical mean radiative energy budget. *Science,* **295,** 841–844, doi:10.1126/science.1065837.

Wong, T., B. A. Wielicki, R. B. Lee, G. L. Smith, K. A. Bush, and J. K. Willis, 2006: Reexamination of the observed decadal variability of the earth radiation budget using altitude-corrected ERBE/ERBS nonscanner WFOV data. *J. Climate,* **19,** 4028–4040, doi:10.1175/JCLI3838.1.

Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. *J. Climate,* **22,** 6653–6678, doi:10.1175/2009JCLI3049.1.