A Comparison of Some Tropical Ocean Models: Hindcast Skill and El Niño Evolution

ARTHUR J. MILLER, TIM P. BARNETT, AND NICHOLAS E. GRAHAM

Scripps Institution of Oceanography, La Jolla, California

4 May 1992 and 10 September 1992

ABSTRACT

Tropical Pacific SST hindcasts are examined in the Zebiak and Cane (Lamont), Latif (MPIZ), Oberhuber (OPYC), and GFDL ocean models, each forced by the same wind-stress fields over the 1970–85 time interval. Skill scores reveal that, although all the models exhibit significant skill, the regions where the skill is maximized differ from model to model. The simplest model (Lamont) has maximum skills in the eastern basin near the boundary while the three GCMs have maxima in central Pacific regions. We also examine, via canonical correlation analysis (CCA), the heat budgets of the surface layers of the Lamont, MPIZ, and OPYC models. We find that although similar spatial relationships exist for the mechanisms that excite SST anomalies (i.e., zonal advection, meridional advection, and vertical advection/mixing), the balance of the strength of these terms is different for each model. Vertical advection tends to control the large-scale structure of SST in the Lamont model, meridional advection provides the dominant large-scale forcing for SST anomalies in the MPIZ model, and all three terms are important in the region of developing SST in the OPYC model. CCA reconstructions of the El Niño events of 1972–73 and 1982–83 reveal that the Lamont model does not exhibit any clear eastward propagation of SST; the MPIZ model propagates SST anomalies eastward for both the 1972–73 and 1982–83 El Niño events while the OPYC model propagates SST eastward for the 1982–83 El Niño and develops SST in place for the 1972–73 El Niño.

1. Introduction

Much of our understanding of the processes important in tropical ocean variability has been gained from analysis of the equations of motion (e.g., Philander 1990). Numerical simulations in particular have provided comprehensive views of El Niño evolution that are impossible to attain with limited arrays of instrumentation (e.g., Busalacchi and O'Brien 1981; Philander and Siegel 1985; Zebiak and Cane 1987). Although extensive model–data intercomparisons have helped to verify the efficacy of numerical models in representing oceanic flux processes (e.g., Latif 1987; Philander et al. 1987; Battisti 1988; Seager 1989; Harrison et al. 1990; Barnett et al. 1991; Harrison 1992) it is still sometimes difficult to understand how the results depend upon the model or the chosen forcing. It is therefore compelling to take an alternative step toward model validation by intercomparing different models themselves to determine whether they are consistently representing upper-ocean variability during El Niño events. [This approach is complementary to that of Neelin and Latif et al. (1992), who discussed a suite of *coupled* ocean–atmosphere models of El Niño and their commensurate intrinsic variability].

Several different numerical models have been used routinely for hindcasting (and recently forecasting) tropical ocean variability. Many factors contribute to differences between model simulations, from differences in resolution, coordinate systems, and wind-stress forcing to more subtle differences like the influence of vertical mixing and the damping/forcing effects of surface heat flux parameterizations. In attempting to understand the major differences in model El Niño evolution among some of the most widely implemented models, we have asked the following questions. Do the models hindcast SST anomalies with similar fidelity? Are the regions of highest hindcast skill similar from model to model? Are the structures and timing of simulated El Niño events consistent among the models? Do the flux processes that control SST have comparable balances of terms in the heat budgets for the different simulations?

To find preliminary answers to these questions, we examine here the similarities and differences in model El Niño development and decay in a subset of the ocean models that today form a primary basis for our understanding of tropical ocean response to wind-stress variability. We force all the models considered here

Corresponding author address: Dr. Arthur J. Miller, Applied Oceanography Group, SACLANT Undersea Research Centre, Viale San Bartolomeo 400, 19138 La Spezia, Italy.

with the same winds (FSU; Goldenberg and O'Brien 1981), with identical drag coefficient ($c_d = 1.5 \times 10^{-3}$), over the same time interval (1970–85). All the model SST anomalies are damped by some form of Newtonian relaxation, although the details of the damping parameterization differ from model to model. The ocean models that we study are the Lamont model (Zebiak and Cane 1987), the MPIZ model (Latif 1987; Barnett et al. 1991), and the OPYC model (Oberhuber 1993; Miller et al. 1992). Hindcast skill of SST output (courtesy of B. Giese) for the GFDL Modular Ocean Model (Pacanowski et al. 1991; Philander 1990) is also discussed. The Lamont ocean model consists of a shallow water layer (which predicts currents and thermocline depth) that contains a constant-depth, frictional, surface layer in which temperature is computed (i.e., a passive mixed layer with constant depth). The MPIZ ocean model is similar to the GFDL model being set in z coordinates with Richardson number–dependent vertical mixing, with the only major difference with the GFDL model being that the MPIZ model has a prognostic sea level. The OPYC model is formed from eight isopycnal-coordinate interior layers that are fully coupled to a surface turbulent bulk mixed-layer model.

Note that each model is constructed on a different grid and has a different parameterization for heat-flux damping. The Lamont model has a 2° longitude and 0.5° latitude resolution throughout the domain, with a constant Newtonian damping coefficient of (150 days)$^{-1}$ for SST anomalies. The MPIZ model has a 6° east–west resolution, with 0.5° north–south resolution at the equator, gradually stretching to 4° at 30°N. The MPIZ damping is via a Haney (1971)-type relaxation to a seasonally and spatially dependent reference temperature, with a damping coefficient of (30 days)$^{-1}$. The OPYC model has 4° east–west resolution in the interior basin, with gradual enhancement to 1.6° near the eastern and western boundaries, and 0.5° north–south resolution at the equator, gradually expanding to 4° in the midlatitudes. For the OPYC model, a Newtonian scheme is invoked that damps to the model climatology, with the damping time being proportional to the depth of the mixed layer as well as dependent on the sensitivity of the total heat flux to SST. The GFDL model has 2.5° east–west resolution and 0.5° north–south resolution at the equator, gradually expanding in the poleward directions, with a standard heat flux computation scheme to damp the model SST to its own climatology. The different features of the models can lead to significant differences in model response characteristics. For example, discretization errors in, say, resolving the gradients required for evaluating the transport or width of a given current depend upon the model layout. A coarse-resolution model may favor vertical mixing processes in a heat budget, while a high-resolution model may allow horizontal transports to be more important in heat balances. As another example, differences in heat flux damping coefficients will influence the size and duration of SST anomalies. These and other considerations should be borne in mind while examining our results.

Note, furthermore, that each of the models herein discussed has been undergoing continual change in efforts for improvement. For example, Seager et al. (1988) have developed a model for computing total SST using a realistic heat flux formulation, which allows the interesting possibility for surface heat fluxes to drive SST anomalies, in contrast to the pure anomaly model of Zebiak and Cane (1987) that is discussed in this paper. In the case of the GFDL model, many researchers (e.g., Harrison 1992) have been testing improvements in such things as the heat flux formulations and vertical mixing schemes. Oberhuber (1993) and Miller et al. (1992) discuss improvements in horizontal discretization schemes and cross-isopycnal mixing parameterizations that have been developed for the OPYC model but not yet applied to the tropical Pacific. Latif (1992, private communication) has developed a new version of his tropical ocean model that is presently being tested in various atmospheric coupling and forecasting scenarios. Thus, the results discussed in this paper, particularly the skill scores, should be viewed with these on going developments in mind.

Our primary variable for intercomparison is SST, since it is the most important variable affecting the ocean–atmosphere flux processes. For the model-observation intercomparisons, we employ standard statistical-analysis techniques. For the model–model intercomparisons, we use canonical correlation analysis (CCA) (Barnett and Preisendorfer 1987) to identify large-scale patterns of variability among model variables. We examine in detail the terms in the surface layer temperature equation that influence the variations of SST, namely, zonal and meridional advection, vertical advection/mixing processes, surface heat fluxes, and model-parameterized horizontal eddy diffusion (which is usually quite small). The primary motivation for these analyses is to identify model strengths, which can be exploited, and weaknesses, which can be remedied, to enhance the continuing development of realistic tropical ocean–atmosphere simulations.

Because of the evident complexity of tropical ocean model dynamic and thermodynamic adjustment to wind fluctuations (e.g., Battisti 1988) and the ominous volume of numerical output involved, we will not describe the results in excessive detail. Rather we will discuss the model response with as much generality as possible to increase readability without oversimplifying the presentation. For example, part of the differences in model responses can be attributed to the differences in background mean conditions upon which the anomalous response forms. The Lamont model has its climatology prescribed from observations, while the MPIZ and OPYC models must develop their own climatologies that will differ from each other and the observed. Only when it is clear that the anomalous re-

sponse has been affected by, for example, the background state, is such a difference pointed out.

2. Overall model performance

a. The Niño regions

As one measure of overall performance, Fig. 1 shows time series of SST for the four models and COADS observations in the so-called Niño-4, 3, and 2 regions of the tropical Pacific. Table 1 shows the analogous values of the anomaly correlation for each model and region. Model anomalies are defined with respect to its respective monthly mean climatology for the 1970–85 time interval. In the western Pacific (Niño-4), the GCMs (MPIZ, GFDL, OPYC) do a much better job of reproducing the observations than the simple model. In this region, the oceanic response is mostly locally forced by the wind. This also results in a very strong comparability of SST response among the GCMs in this region. For example, the correlation in SST between OPYC and GFDL is 0.87 in this region, while GFDL–MPIZ correlation is 0.72 and MPIZ–OPYC correlation is 0.74. A delay with respect to the observed onset time of the 1982–83 El Niño occurs in each of the three GCMs.

In the eastern Pacific (Niño-3), where equatorial wave propagation has much more influence on the SST (McCreary 1976; Hurlburt et al. 1976), similar results hold, although the simplest model's skill is comparable to the others in this region. There is also more disparity among the responses of the GCMs. Although OPYC and GFDL SST are well correlated (0.90) there, the SST anomaly correlation between MPIZ–GFDL drops to 0.65 and MPIZ–OPYC drops to 0.57. The three GCMs each produce a similar 1982–83 event, with an initial peak that is too weak and a follow-up peak that is too strong with respect to the observations. For the cold events of the 1970s, the MPIZ model appears to do the best job.

In the Niño-2 region, where variations in upwelling and vertical mixing are thought to predominantly control SST variability (Battisti 1988; Seager 1989; cf. Gill 1983), the Lamont model outperforms the GCMs. The stronger correlation with observations is visually evident in Fig. 1c, and the three El Niño events of this time interval are fairly well reproduced by the Lamont model, although with weaker amplitude than observed. Inspection of the three GCM time series reveals little "stand-out" evidence for El Niños except for the 1982–83 events. The three GCMs exhibit phase and amplitude discrepancies among themselves in their representations of the peak of the 1982–83 event, evidently a consequence of different wave propagation and vertical mixing characteristics among the models.

To summarize, for these large-scale regional averages all the models perform respectably in modeling COADS observations in the open-ocean regions (Niño-

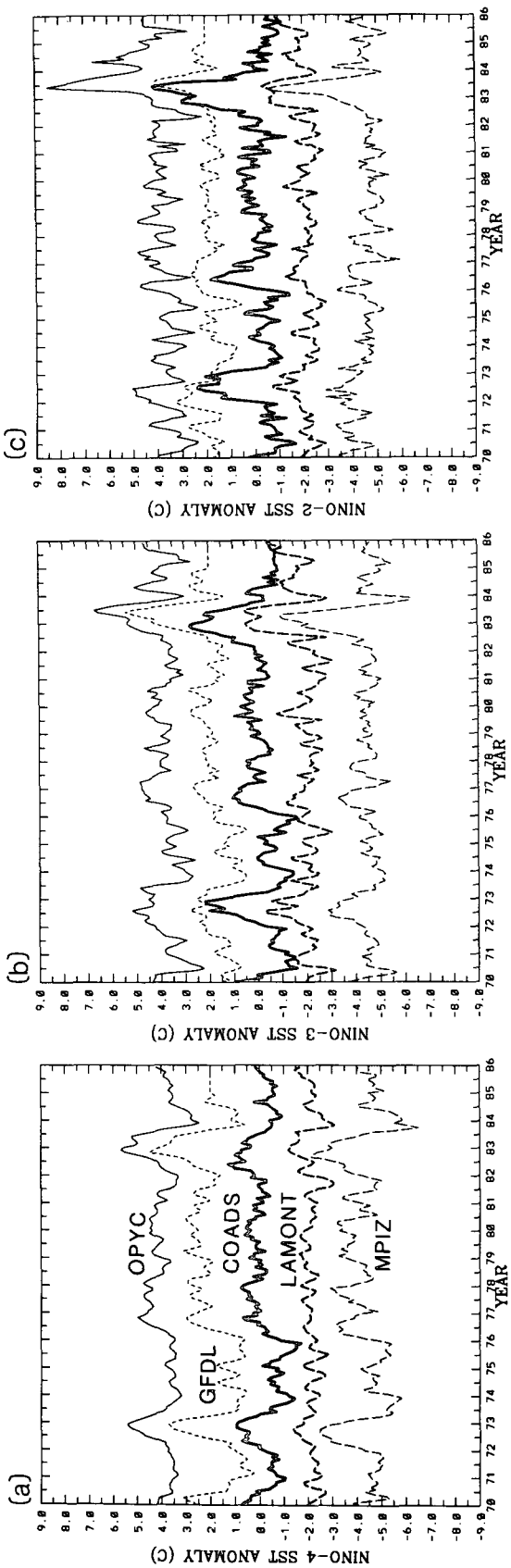


FIG. 1. Time series of SST anomalies averaged over the (a) Niño-4 region (160°E–150°W; 5°N–5°S), (b) Niño-3 region (150°W–90°W; 5°N–5°S), and (c) Niño-2 region (90°W–coast; 0°N–5°S). From top to bottom, OPYC (thin), GFDL (short dash), COADS (thick), Lamont (thick dash), and MPIZ (long dash) are plotted and offset by 2°C.

TABLE 1. Anomaly correlation: Models vs COADS (1970–85).

	Niño-2	Niño-3	Niño-4
Lamont	.68	.60	.46
MPI	.59	.74	.76
OPYC	.46	.63	.72
GFDL	.57	.69	.81

4 and 3) where the anomaly correlations are near 0.7. Along the coast of South America, the GCMs correlate less favorably with observations with values dropping into the 0.5 range. The Lamont model tends to have larger correlations with observations in the eastern Pacific than in the west, which is the opposite tendency of the GCMs. All three GCMs produce a delayed response during the 1982–83 event, particularly evident in the Niño-4 region.

b. *Pointwise comparisons*

Although the model variability when averaged over large regions compares rather well with observations (and fairly well among the models), the smaller scale structure of the model responses differs considerably from observations (and model to model) as can be seen in pointwise comparisons of model and observations. For example, consider Fig. 2, which shows rms error, $\sqrt{\langle \epsilon^2 \rangle}$, where $\epsilon \equiv T_0 - T_m$ and the angle brackets indicate averaging over the record length, for the models relative to COADS observations; all time series are smoothed by a 5-month running mean filter. The Lamont model response has a maximum of rms error of 0.9°C near 115°W, while in the remainder of the basin the error tends to vary between 0.4° and 0.7°C. The rms error plots of the MPIZ, GFDL, and OPYC models each reveal twin peaks, one in the east and one in the west Pacific. The rms error of the GCMs in the western Pacific, where SST is observed to vary relatively weakly, indicates that excessive SST variability is being generated there by the GCMs.

Note that, even though the rms error for the Lamont model is much lower in the western Pacific (due simply to its SST response being weaker there), the correlations with observations seen in Table 1 for the west Pacific are higher for the three GCMs relative to the Lamont model. The primary reasons that the GCMs have excessive SST variability in the western Pacific are because the GCM mean thermoclines are too diffuse there and the GCM mean cold tongues are too strong, penetrating into the warm pool region. The result is that the surface heat budget is affected by excessive variations in vertical mixing (by the first effect) and excessive anomalous horizontal advection of SST (by the second effect). These well-known problems are being addressed in many modeling efforts around the world. The peaks in rms error in the east Pacific, on the other hand, where SST variability is observed to be largest, reveal where clear improvements can be made in GCM model performance, since as we shall see throughout this paper, the much simpler Lamont model often performs better than the GCMs in this region.

In order to quantify the skill associated with the pointwise variability in rms error, we compute a skill score, $S = 1 - \langle \epsilon^2 \rangle / \langle T_0^2 \rangle$, which is a measure of the variance explained by the model hindcasts. A value of $S = 1$ corresponds to perfect skill, while $S \leq 0$ implies excessively poor skill. Figure 3 shows the skill scores for SST fields that have been smoothed in time by a 5-month running mean. All three GCMs have a peak in skill in the eastern-central Pacific. The Lamont model, in contrast, has a skill maximum that hugs the eastern boundary and is, indeed, the strongest among all the models with values (when no spatial smoothing is done for plotting purposes; see figure caption) in the 0.60s south of the equator there. The MPIZ model produces a wide horseshoe-shaped skill maximum (values in the range 0.40–0.50) with significant skill north and south of the equator in the central Pacific and along the equator in the eastern central Pacific, but dropping off near the eastern boundary. OPYC exhibits similar skill scores to the MPIZ model, but the peak is located more in the central Pacific with off-equatorial tongues of skill extending eastward from the central Pacific maximum. The GFDL skill tends to be smaller than the other two GCMs both in areal extent and amplitude. In the western Pacific, all three GCMs have strongly negative skill associated with the excessive SST variability they each generate there. The Lamont model, unlike the GCMs, produces an additional peak in skill near the date line, just north of the equator. Note that since this definition of skill accentuates the correctness of magnitude rather than phase of hindcasted SST, the Lamont model's performance quality in the western Pacific is somewhat exaggerated relative to the GCMs. Indeed, the correlation between model and observed SST for the GCMs tends to be higher there than for the Lamont model, as can be seen in Table 1 and in the next paragraph.

A second quantitative measure of model verisimilitude, the anomaly correlation, $r \equiv \langle T_m T_0 \rangle / (\langle T_m^2 \rangle \langle T_0^2 \rangle)^{1/2}$, measures the phasing of the SST time variations rather than the absolute variance levels of S. Plotted in Fig. 4 is the pointwise anomaly correlation with the COADS observations for the four models. The results are fairly similar to the skill scores, although the areal extent of what might be considered to be useful results appears to be larger in the correlation maps than the skill maps. The three GCMs have maxima in the central Pacific while the Lamont model maximum occurs in the eastern basin and the along the South American coast. The largest correlations, more than 0.8, are found in the GFDL model in the central Pacific and in the MPIZ model in the central equatorial Pacific between 130°W and 120°W. All four

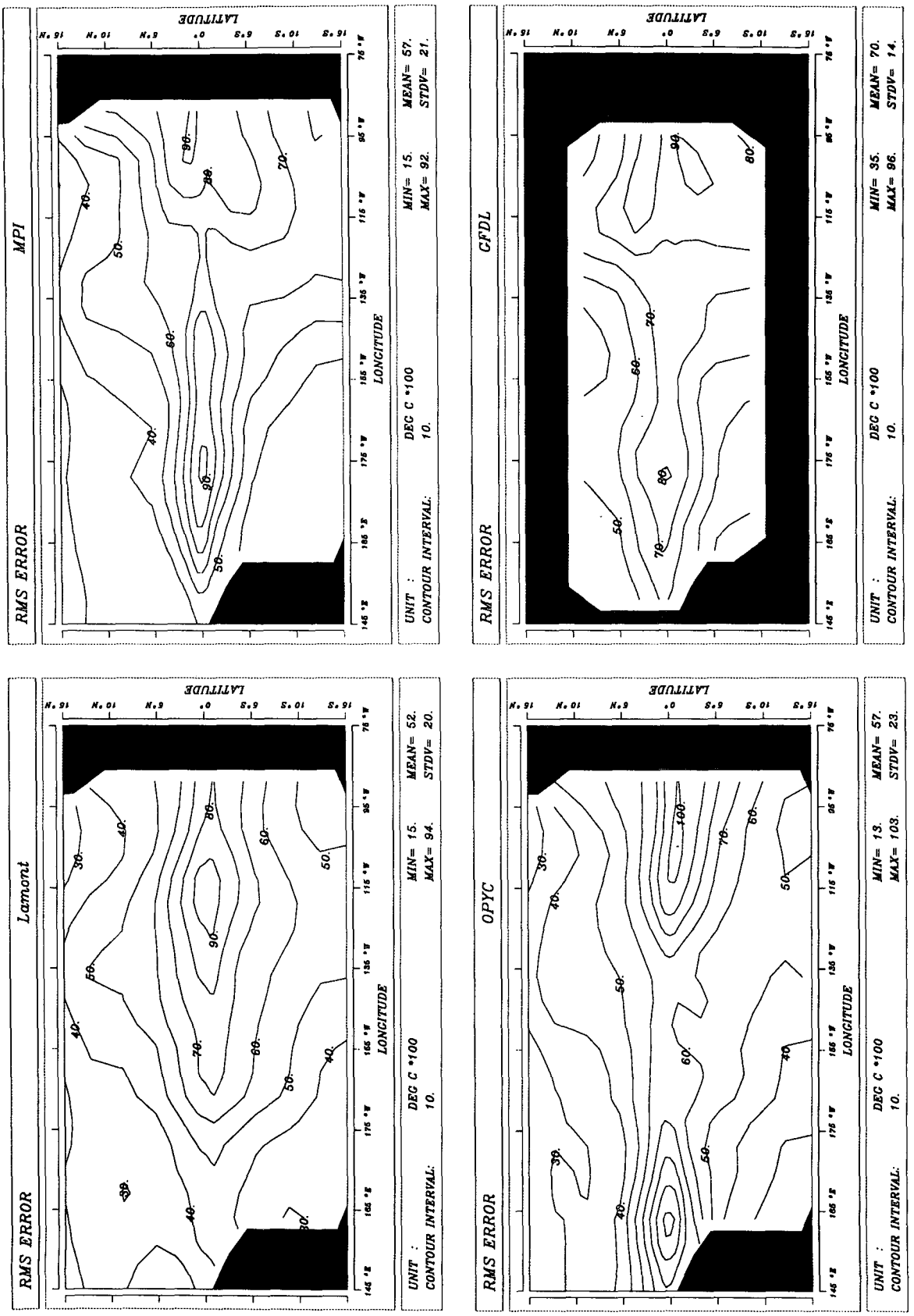


FIG. 2. Rms error (relative to COADS observations) of SST (°C × 100) for 1970–85 interval for (a) Lamont model (b) MPIZ model, (c) OPYC model, and (d) GFDL model. To facilitate implementation of the statistical analysis programs, the observed and model SST anomaly fields from the Lamont, OPYC, and GFDL models are each interpolated to the MPIZ grid before analysis. The time series at each point have been smoothed by a five-month running mean filter. In addition, the plotted field has been lightly smoothed in space for ease in viewing the large-scale features of the results (contour interval = 10).

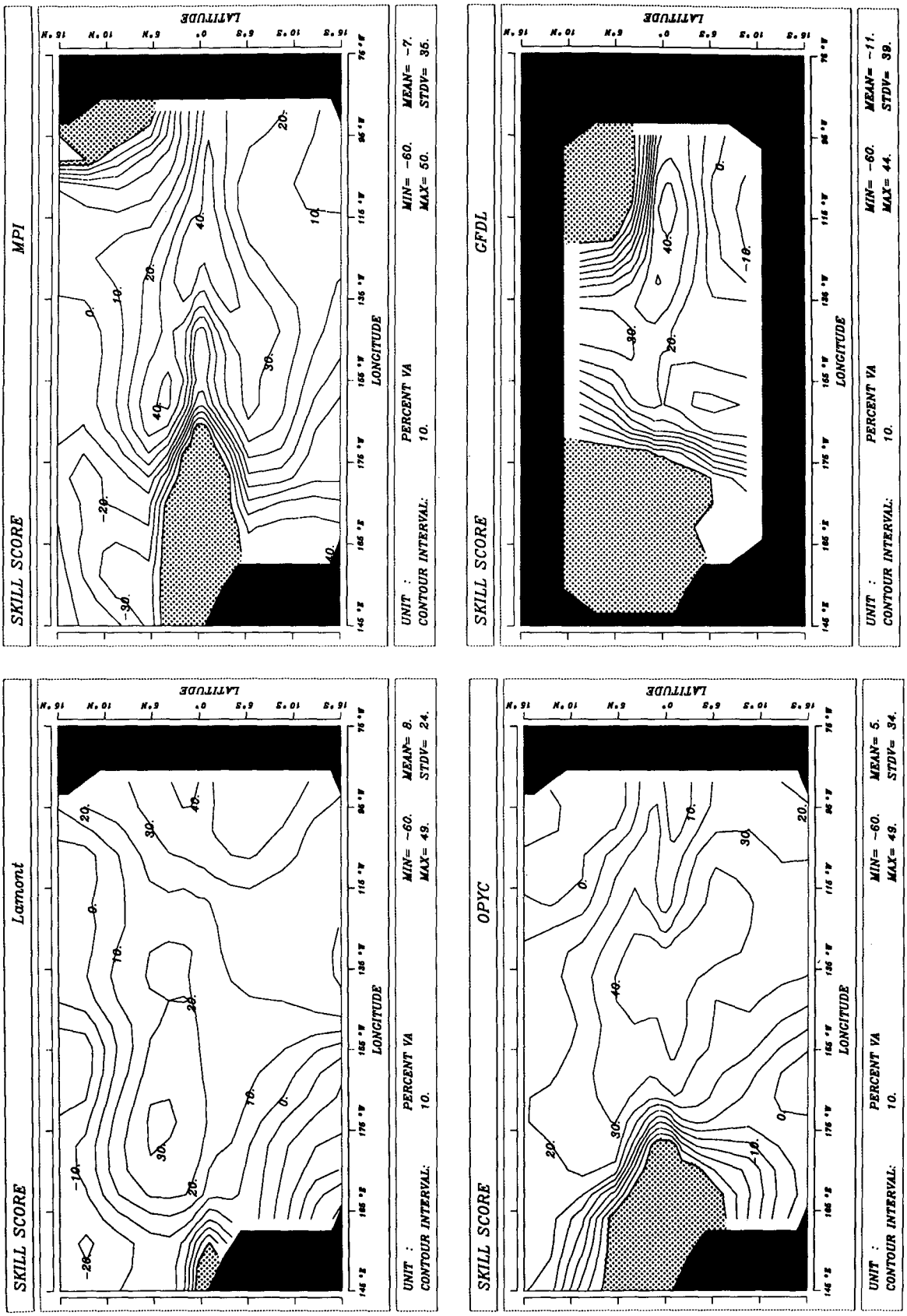


FIG. 3. As in Fig. 2 but for skill score (×100) as defined in the text ($CI = 10$). Regions of skill less than −0.50 are shaded.

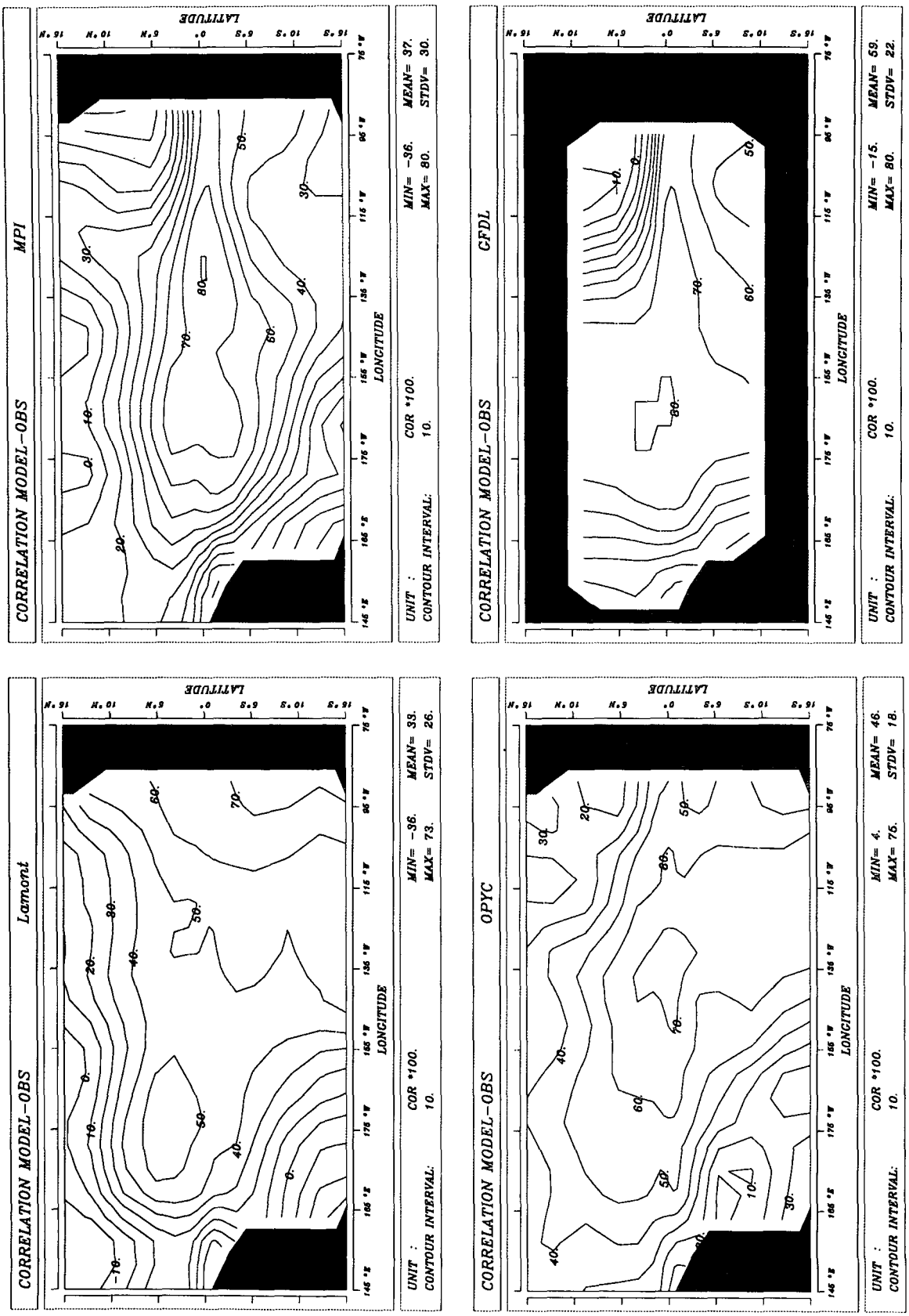


FIG. 4. As in Fig. 2 but for correlation (×100) between model SST and COADS SST ($CI = 10$).

models have correlations greater than 0.5 over sizable areas of the domain, though OPYC also exhibits some 0.5 regions farther poleward than do the other two GCMs. The correlation results suggest that the four models are all producing significant results, but the four models yield their best results in different parts of the basin. This suggests that improvements in any one model may be made by identifying the reasons for the strengths of the other models.

It is of interest to consider the relationship of the models' anomalous SST error to the model SST anomalies themselves. This is because if there is a systematic relationship between model SST and SST error, the possibility exists (if the model response is nearly linear) that a large part of the error can be removed from the model runs by invoking a statistical error correction scheme to improve model performance. For example, if the model SST anomalies in a certain region are consistently too large by a factor of 2, one can increase the strength of SST damping by including a Newtonian damping coefficient related to the error. If, however, the error is unrelated to the model SST, this formalism fails. Consider, therefore, the correlation between error and SST, $\beta \equiv \langle \epsilon T_m \rangle / \langle T_m^2 \rangle$, for the four simulations (Fig. 5). The three GCMs exhibit high negative values of β along the equator with strongest negative values in the regions where the skill is lower. One can immediately see that in the western Pacific, where the skill scores, S, are the worst for all four models, β is highly negative for all four models, surpassing -0.8 for the three GMCs and -0.6 for the Lamont model. This implies that the error can be considerably reduced in the west Pacific in all four models and that significant improvements in modeling SST variability along the equator can be anticipated for the GCMs.

In summary, the simple model (Lamont) hindcasts SST best in the region close to the eastern equatorial boundary, while the three GCMs tend to do best over differing regions of the central tropical Pacific strip. All three GCMs have exaggerated SST variability in the western tropical Pacific owing to the model thermoclines being too diffuse and/or the cold tongues being too strong. It thus appears clear that, among this subset of tropical models, the greater complexity of the GCMs does succeed in adding hindcast skill to a larger region of the basin. However, improvements must still be made in the GCMs to increase skill near the eastern boundary and to reduce the excessive SST variations in the western basin. In the next section, we seek to isolate what mechanisms contribute to the differing SST variability distributions among the models.

3. Large-scale pattern relationships: CCA modes

Canonical correlation analysis (CCA) is a statistical technique for identifying relationships between different fields as they evolve over time (e.g., Graham et al. 1987; Barnett and Preisendorfer 1987). We computed the CCA modes for the relation between monthly mean fields of SST and the terms of the surface-layer heat budget, namely, the zonal advection, meridional advection, vertical advection/mixing and surface heat flux terms. The rather small effects of horizontal diffusion were ignored here. The CCA applies to the time interval 1970–85 for the Lamont, MPIZ, and OPYC models; the GFDL model heat budget was not available for this study. The heat-budget terms correspond to (i) the SST component of the Lamont model, (ii) the top two grid points of the MPIZ model, and (iii) the variable depth surface mixed layer of the OPYC model. For the Lamont and MPIZ models, the vertical advection term explicitly corresponds to variations in upwelling [i.e., $w(\partial T/\partial z)$]. However, since the temperature in the variable depth mixed layer of OPYC is not directly changed by upwelling, variations in vertical mixing [i.e., $w_e(T_m - T_0)/H$, where w_e is the entrainment velocity, T_m the mixed-layer temperature, T_0 the underlying ocean temperature, and H the mixed-layer depth] are the proper interpretation of that term. The CCA modes were computed based on an expansion of the top ten empirical orthogonal functions of the fields in question, which explained a very large fraction of the total variance of those fields; the amount of total variance explained by each CCA mode is indicated in the relevant figure for each field. Since all the models used some form of Newtonian damping, the CCA-mode heat fluxes generally acted in opposition to CCA-mode SST anomalies, so we have not included these fields in any of the figures, although they are continuously in action. In reality, heat fluxes may have the more interesting effect of exciting SST anomalies in some regions of the tropics (Seager 1989; Liu 1988; McPhaden and Hayes 1991).

Note that although the CCA corresponds to all the important terms of the heat budget, the fields of CCA modes are not constrained to satisfy the heat budget on a point-by-point basis. Also, if SST tendency (proportional to subsequent month minus antecedent month) is substituted for SST in the CCA, the results are very similar, except for a phase shift of the time variability of SST with respect to SST tendency. Last, recall that the CCA modes are ordered by the strength of the correlation among the fields rather than by the total variance explained.

In the Lamont model, the dominant CCA mode is associated with 52% of the SST variability in the basin (Fig. 6). In line with what we have seen from the skill maps, the strongest signal in SST is near the coast of South America, and the amplitude drops off in the western basin. The pattern has one sign across the entire equatorial Pacific. In the eastern basin, where the mode is strongest, the dominant processes that establish this SST mode are upwelling and, to a lesser degree, meridional advection (north of the equator in the eastern basin). Zonal advection (in the west-central Pacific) also contributes to the excitation of this mode. These

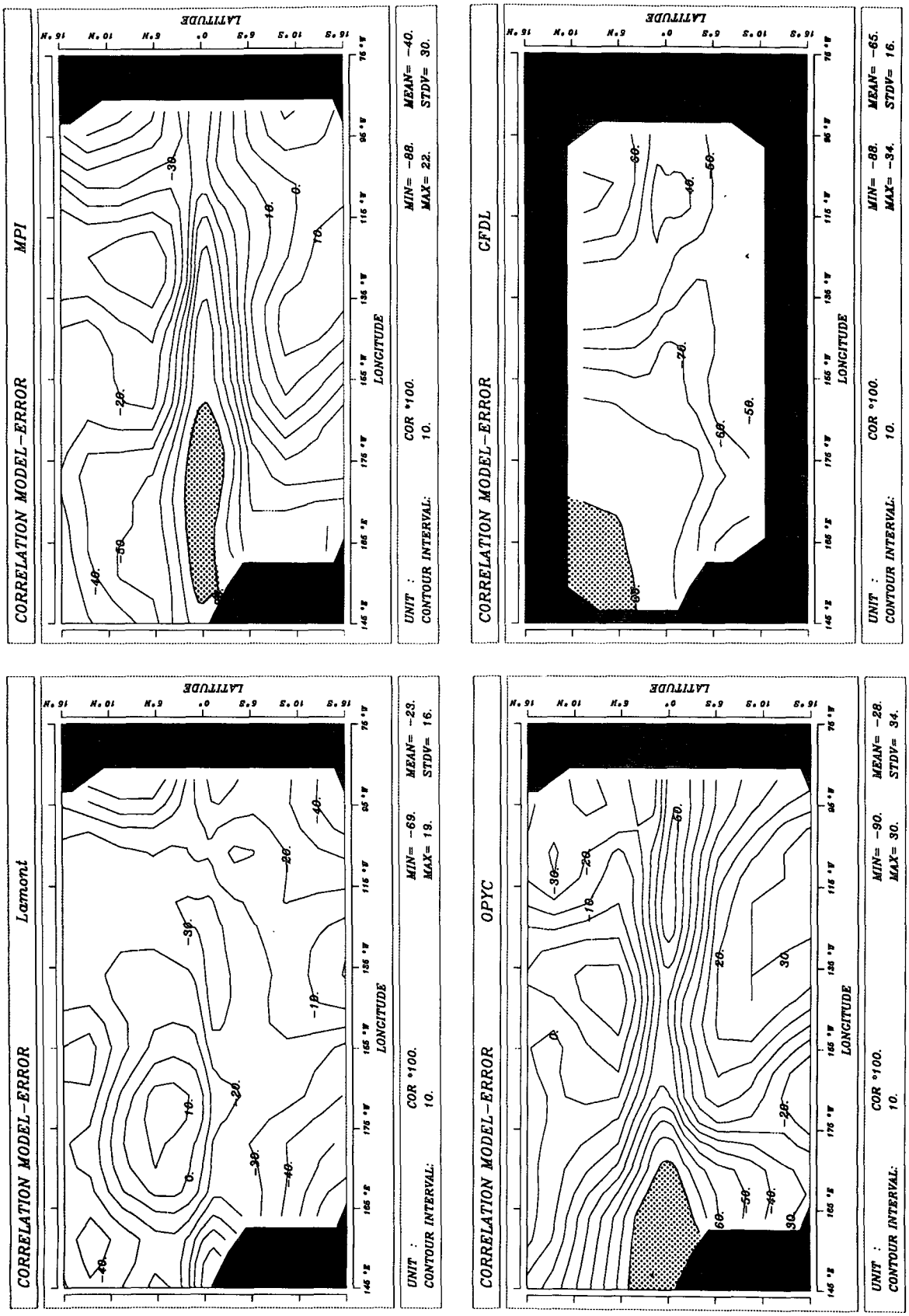
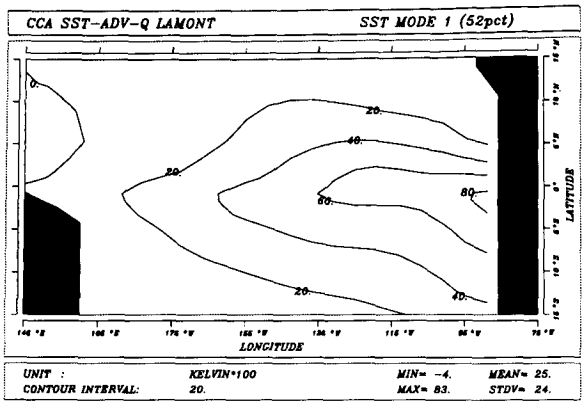
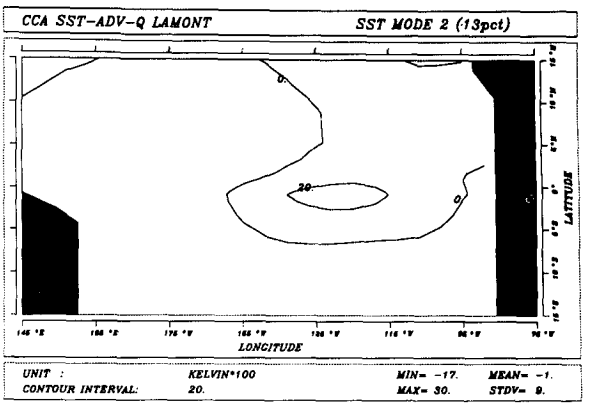
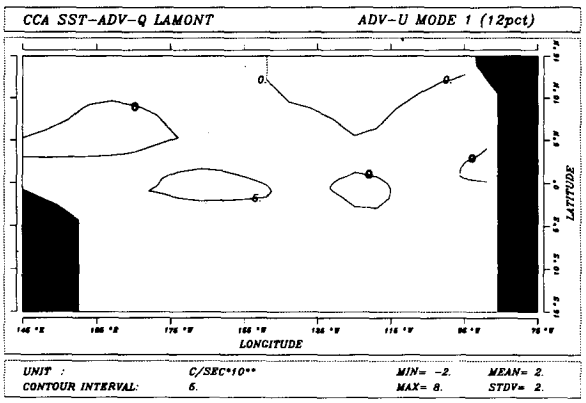
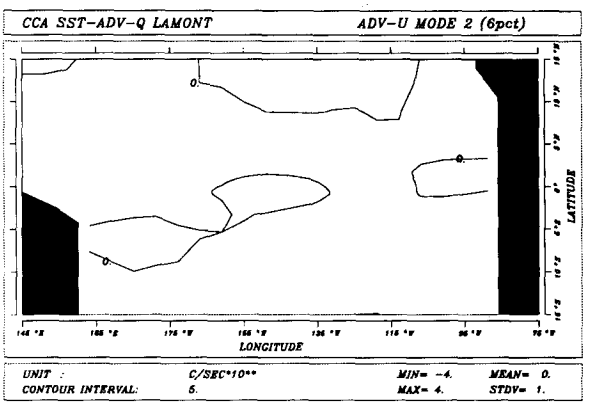


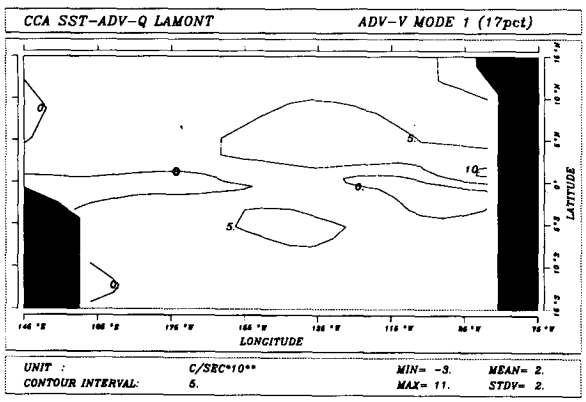
FIG. 5. As in Fig. 2 but for correlation ($\times 100$) between model SST and SST error ($CI = 10$). Regions of correlation less than -0.80 are shaded.

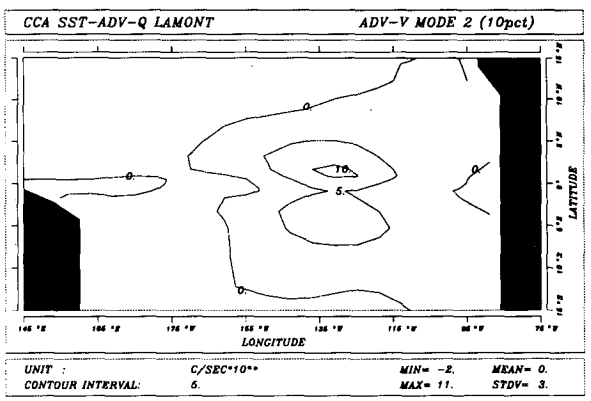

CCA SST-ADV-Q LAMONT
SST MODE 1 (52pct)
LONGITUDE
LATITUDE
UNIT : KELVIN*100
CONTOUR INTERVAL: 20.
MIN= -4. MEAN= 25.
MAX= 83. STDV= 24.

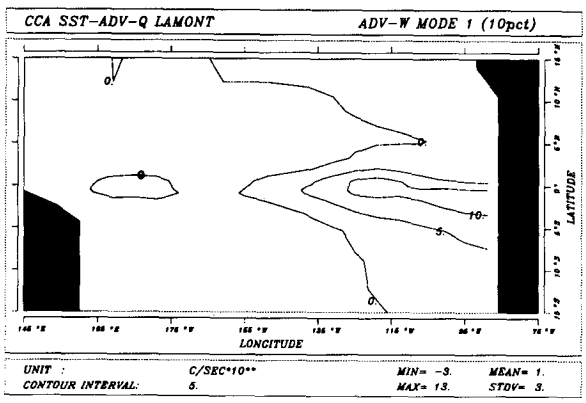

CCA SST-ADV-Q LAMONT
SST MODE 2 (13pct)
LONGITUDE
LATITUDE
UNIT : KELVIN*100
CONTOUR INTERVAL: 20.
MIN= -17. MEAN= -1.
MAX= 30. STDV= 9.

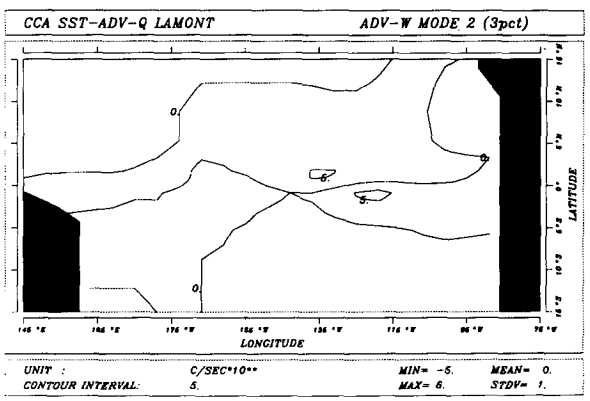

CCA SST-ADV-Q LAMONT
ADV-U MODE 1 (12pct)
LONGITUDE
LATITUDE
UNIT : C/SEC*10**
CONTOUR INTERVAL: 5.
MIN= -2. MEAN= 2.
MAX= 8. STDV= 2.


CCA SST-ADV-Q LAMONT
ADV-U MODE 2 (6pct)
LONGITUDE
LATITUDE
UNIT : C/SEC*10**
CONTOUR INTERVAL: 5.
MIN= -4. MEAN= 0.
MAX= 4. STDV= 1.


CCA SST-ADV-Q LAMONT
ADV-V MODE 1 (17pct)
LONGITUDE
LATITUDE
UNIT : C/SEC*10**
CONTOUR INTERVAL: 5.
MIN= -3. MEAN= 2.
MAX= 11. STDV= 2.


CCA SST-ADV-Q LAMONT
ADV-V MODE 2 (10pct)
LONGITUDE
LATITUDE
UNIT : C/SEC*10**
CONTOUR INTERVAL: 5.
MIN= -2. MEAN= 0.
MAX= 11. STDV= 3.


CCA SST-ADV-Q LAMONT
ADV-W MODE 1 (10pct)
LONGITUDE
LATITUDE
UNIT : C/SEC*10**
CONTOUR INTERVAL: 5.
MIN= -9. MEAN= 1.
MAX= 13. STDV= 3.


CCA SST-ADV-Q LAMONT
ADV-W MODE 2 (3pct)
LONGITUDE
LATITUDE
UNIT : C/SEC*10**
CONTOUR INTERVAL: 5.
MIN= -5. MEAN= 0.
MAX= 6. STDV= 1.

results are generally consistent with those of Battisti (1988) who analyzed the heat budget of the ocean of a coupled version of the Zebiak and Cane (1987) model and those of Goswami and Shukla (1991) who computed the empirical orthogonal function of the Lamont model SST for this same time interval.

The second CCA mode in the Lamont model (Fig. 6, 13% of the SST variance) has a small peak in the central-eastern tropical Pacific. Both upwelling and meridional advection contribute significantly to driving this mode. The third CCA mode of the Lamont model (not shown, 8% of the SST variance) exhibits a peak in SST centered along the equator near 140°W, surrounded by weaker SST of opposite sign, with zonal advection appearing to contribute most strongly to its excitation.

For the MPIZ model, we find that the first CCA mode explains 26% of the basinwide SST variance. This mode (Fig. 7) has one sign across the equatorial Pacific, with maximum amplitude near 160°W. The CCA mode reveals that all three advection components act to drive this SST mode. However, near-equatorial meridional advection appears to control the strong basinwide signature of this mode, in particular the location of the maximum SST. Zonal advection (in the west Pacific) and vertical advection (in the eastern Pacific) contribute somewhat weaker driving affects along the equator. This is in agreement with the results obtained by Barnett et al. (1991). The large amplitude of the meridional advection term is likely a consequence of strong meridional gradients in the model mean SST due to a too strong cold tongue.

The second CCA mode of the MPIZ model (Fig. 7; 10% of the total variance of SST) has a nodal line near 160°W with SST anomalies of opposing sign over each side of the basin. With dominant forcing by meridional and zonal advection, its time variability (not shown) tends to be 90 degrees out of phase with the first CCA mode, indicating cool SST in the eastern basin preceding El Niño and warm SST in the eastern basin following the development of the basinwide warm anomaly of CCA mode 1. The net effect of the first and second CCA modes thus suggests an eastward propagation of the warm events (as described by Barnett et al. 1991). The third CCA mode (not shown; 11% of the variance of basinwide SST) has a maximum near 180°W, where we have already described the occurrence of overestimated variability in SST.

The leading CCA modes of the OPYC model are less distinct than in either the simple Lamont model or the MPIZ GCM. The first mode (Fig. 8; 9% of the total SST variance) has broad equatorial scale with opposite-signed maxima in the central and eastern basin combined with another (clearly dubious) maximum in the western basin. In the eastern basin, this mode is excited by both variations in vertical mixing and by meridional advection. Zonal advection (eastward currents flowing up the mean SST gradient) serves nearly as strongly as a damping mechanism on the eastern side of the central Pacific SST peak. The second mode in SST (not shown; 20% of the total SST variance) also has broad equatorial scale, but has a strong maximum in the western basin and is, therefore, associated with the excessive variability there.

The third mode (11% of total SST variance) is more localized in the eastern basin with peak amplitude in SST near 100°W. Akin to the first mode, the balance of terms that affects SST tends to be excitation by vertical mixing effects and meridional advection combined with damping (forcing) effects of zonal advection on the eastern (western) flank of the SST maximum.

In summary, the canonical correlation analysis of the models has revealed that the models are similar in that zonal advection tends to excite the SST anomaly on its western flank, meridional advection tends to excite the anomaly north and south of the SST peak, and upwelling/mixing effects tend to drive the SST anomaly on its eastern flank. However, the CCA also shows that the balances of these terms are different in each of the three models. In the Lamont model, vertical advection has primary control over the dominant CCA mode. In the MPIZ model, meridional advection, supplemented by weaker effects of zonal and vertical advection, controls the structure of SST response. In the eastern basin of OPYC, all three terms have relatively important effects in the development of the anomalous SST. We have no way to tell which, if any, of the models is more correct physically than the others, although readers might make their own judgements based on the performance results show previously in Figs. 1–5. One feature that stands out from the maps of SST, however, is that the dominant Lamont model CCA mode 1 has two features that none of the GCM CCA modes possess, yet which appear in well-known maps of the first empirical orthogonal function of observed SST. These are an eastern intensification of the SST pattern and a rather broad latitudinal scale. Thus, in terms of pattern simulation, the Lamont model CCA mode 1 exhibits more realistic characteristics than either GCM.

FIG. 6. Canonical correlation analysis results for (left) mode 1 and (right) mode 2 of Lamont model. From top to bottom, SST (°C × 100.; $CI = 20$), zonal advection (°C $s^{-1} \times 10^8$; $CI = 5$), meridional advection (°C $s^{-1} \times 10^8$; $CI = 5$), and vertical advection/mixing (°C $s^{-1} \times 10^8$; $CI = 5$). Since the heat fluxes are parameterized in a Newtonian fashion, they are generally in opposition to the plotted SST anomaly and so are not shown. To facilitate implementation of the statistical analysis programs, the model anomaly fields from the Lamont, OPYC, and GFDL models are each interpolated to the MPIZ grid before analysis. The plotted fields have been lightly smoothed in space for ease in viewing the large-scale features of the results. The percent of the total variance explained by the CCA mode field is indicated in each frame.

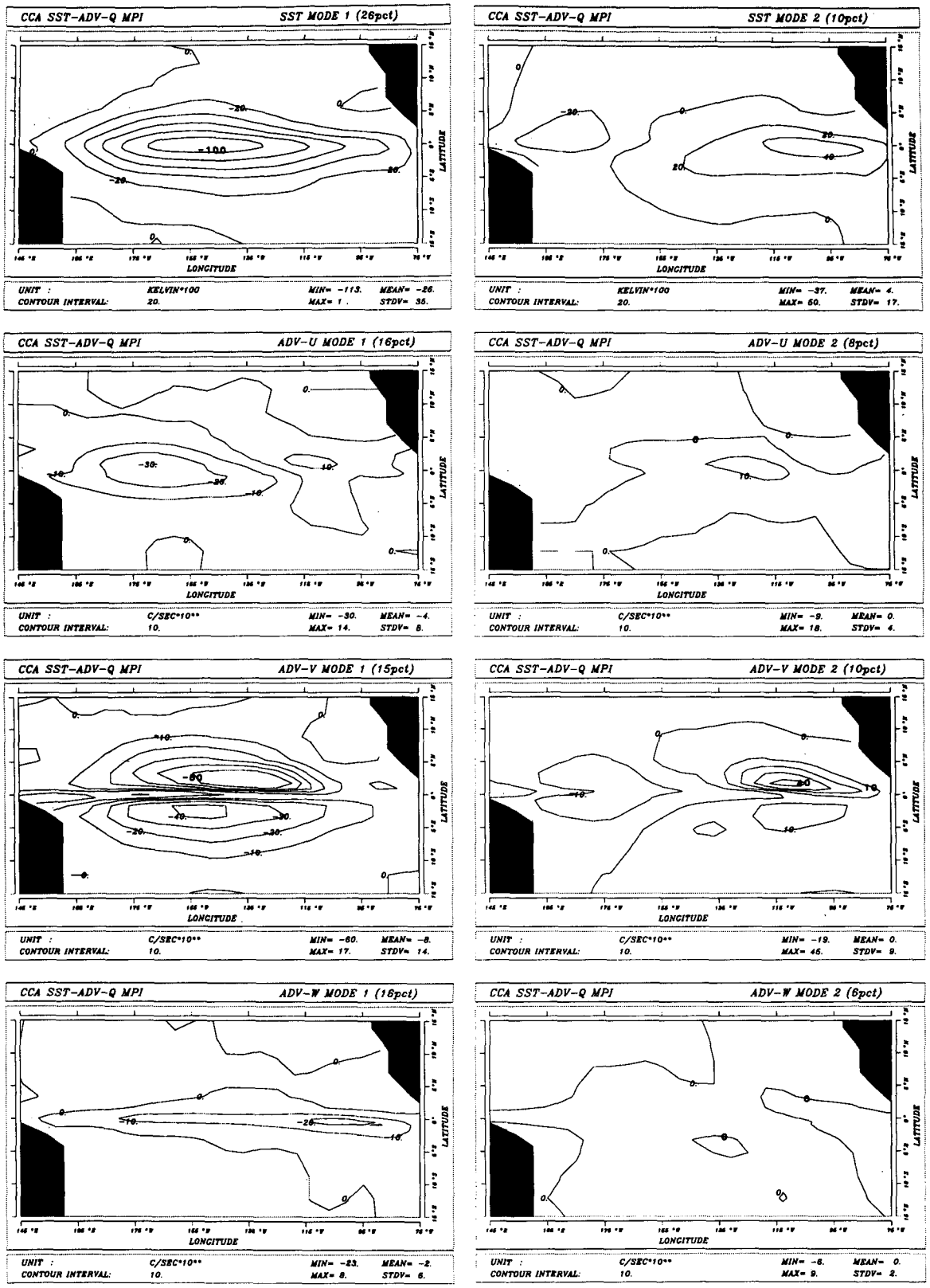


FIG. 7. As in Fig. 6 but for MPIZ model. (CI = 10 for zonal, meridional, and vertical advection fields.)

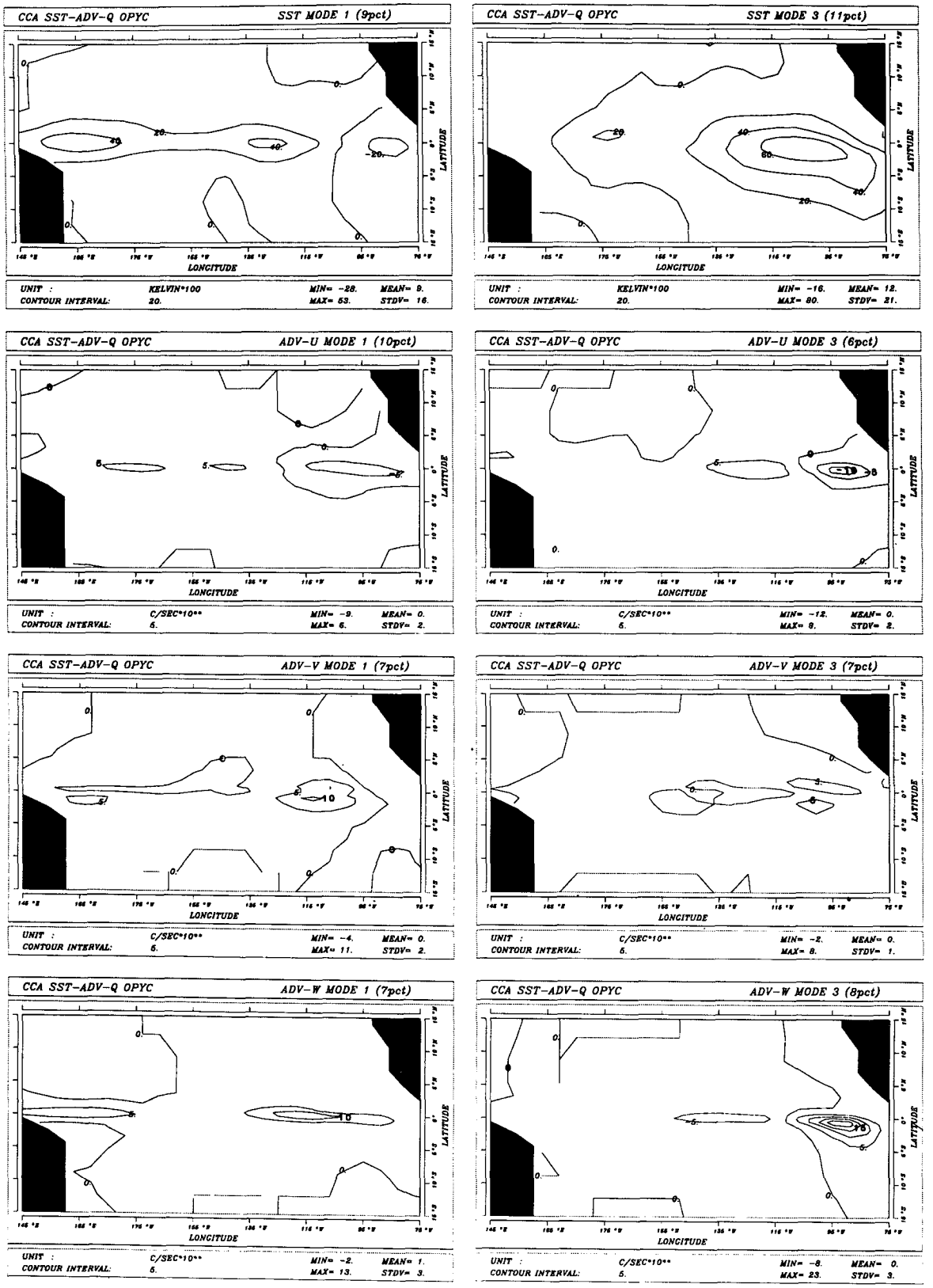


FIG. 8. As in Fig. 6 but for (left) mode 1 and (right) mode 3 of the OPYC model.

4. CCA reconstructions of El Niño events

In this section, we discuss reconstructions of the El Niño events from 1972–73 and 1982–83 based on the first three canonical modes discussed in section 3, modulated by their time variability, for each of the three models. We emphasize that these reconstructions represent only that portion of the total response of the models that has organized large-scale behavior; the total SST anomaly fields for these models, which were used in Figs. 1–5, are not shown here. Our intent is to distinguish the differences in El Niño development among the three models under consideration. (Since the model heat fluxes always serve to damp these model SST anomalies, in proportion to their amplitude, they are left out of the following discussion.)

a. The 1972–73 El Niño

During 1972–73, all three ocean model SST fields develop during spring 1972 and last through winter 1973. For the Lamont model, vertical advection, and to a lesser degree meridional advection, contributes to the establishment of the event in the spring, summer (Fig. 9, left) and fall of 1972, which is consistent with the first CCA mode excitation (Fig. 6, left). By the following summer, some evidence for the excitation of the second CCA mode, namely weak activity near 130 W, is present. For the MPIZ model, the dominance of meridional advection in controlling the broad spatial scale of the event is clear in Fig. 10. Anomalously weak upwelling, combined with an eastward shift in the maximum of forcing due to meridional advection, causes the MPIZ El Niño to spread eastward toward the coast of South America in the summer and fall of 1972. Zonal advection acts mainly in the western Pacific, during the development time interval, resulting in the amplitude of SST remaining rather strong there. Thus, the MPIZ model appears to exhibit a shift in emphasis from the first CCA mode (Fig. 7, left) to the second CCA mode (Fig. 7, right) over the course of the development of the El Niño. For OPYC, a similar result holds, namely, the first CCA mode (Fig. 8, left) appears strongest near the beginning of the event (Fig. 11, left) followed by the excitation of the third CCA mode (Fig. 8, right) by winter 1973 (Fig. 11, right). As found in the CCA analysis, all three forcing effects contribute to the OPYC model's 1972–73 El Niño, with zonal advection acting to cool the SST near the coast of South America [similar to the effects of zonal advection observed at 110°W by Hayes et al. (1991)].

As might be expected, the reverse of the effects that excite the model El Niños tends to help extinguish them, along with the damping effects of the heat flux. In the spring of 1973, vertical advection and zonal advection cause the Lamont model El Niño to lose most of its strength. For the MPIZ model, meridional advection serves as the dominant dynamic damping mechanism during the spring and summer of 1973. All three dynamical effects act to cool the waning El Niño in the OPYC model during summer 1973.

b. The 1982–83 El Niño

Casual inspection of Figs. 12–14 (July 1972–January 1973) compared with Figs. 9–11 (October 1982–April 1983) reveals that a surprisingly similar sequence of events transpires in the development of the 1972–73 and 1982–83 model El Niños, except that the 1982–83 event developed three months later and with larger amplitude. The Lamont and OPYC model El Niños begin in fall 1982, while the MPIZ model El Niño begins slightly earlier in the summer of 1982. The Lamont and MPIZ model El Niños last through summer 1983, while the OPYC model's lasts through fall 1983 near the eastern boundary.

For the Lamont model, both vertical and meridional advection contribute strongly to the growth and maintainence of the event during the fall, winter, and spring periods of 1982–83. Zonal advection plays only a minor role in the Lamont model SST excitation, except during the fall of 1982. (Fig. 12, right, clearly shows the excitation of the second CCA mode in the Lamont model) All three dynamical mechanisms, along with the heat fluxes, contribute to the demise of the Lamont event in fall 1983. Note that the Lamont model does not propagate the event eastward as was observed, which seems to be a general property of that model.

In the MPIZ model El Niño, meridional advection again dominates the large-scale development of the event, supplemented by weaker effects of zonal advection to the west of the anomaly and vertical advection to the east. The phasing of the first two CCA modes is again evident in Fig. 13 and represents the eastward propagation of the SST anomaly.

For the OPYC model El Niño, we again find vertical mixing effects and meridional advection exciting the SST with additional forcing (damping) by zonal advection to the west (east) of the SST anomaly (Fig. 14). Eastward propagation of the developing El Niño, in line with excitation of the first and third CCA modes, is clearly evident. Again, for the three models, the decrease in SST during fall 1983 tends to be the reciprocal of the processes that warmed the SST during the beginning of the event.

c. SST propagation during the 1982–83 event

In order to better clarify the degree to which model reconstructions of SST propagate vis-a-vis observed SST, we present Fig. 15. During the 1982–83 event warm SST anomalies were observed to begin to develop in July 1982 in the central equatorial Pacific near 120° W. During the subsequent fall and winter seasons the observed anomalies grew in strength and spread eastward toward the coast. In the spring of 1983 the ob-

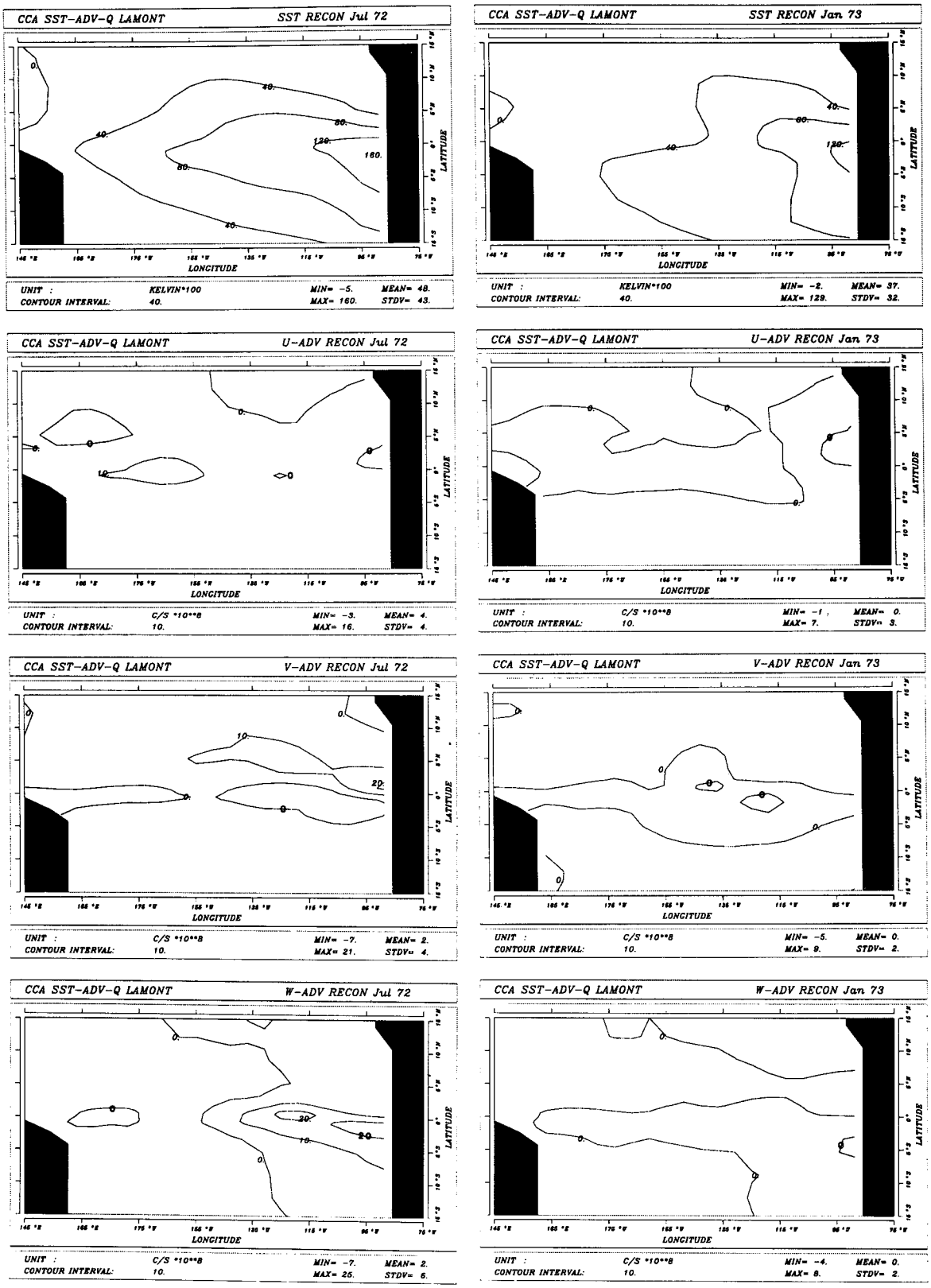


FIG. 9. As in Fig. 6 but for reconstructions of the 1972–73 El Niño, for (left) July 1972 and (right) January 1973 for the Lamont model. (CI = 40 for SST and CI = 10 for zonal, meridional, and vertical advection fields.)

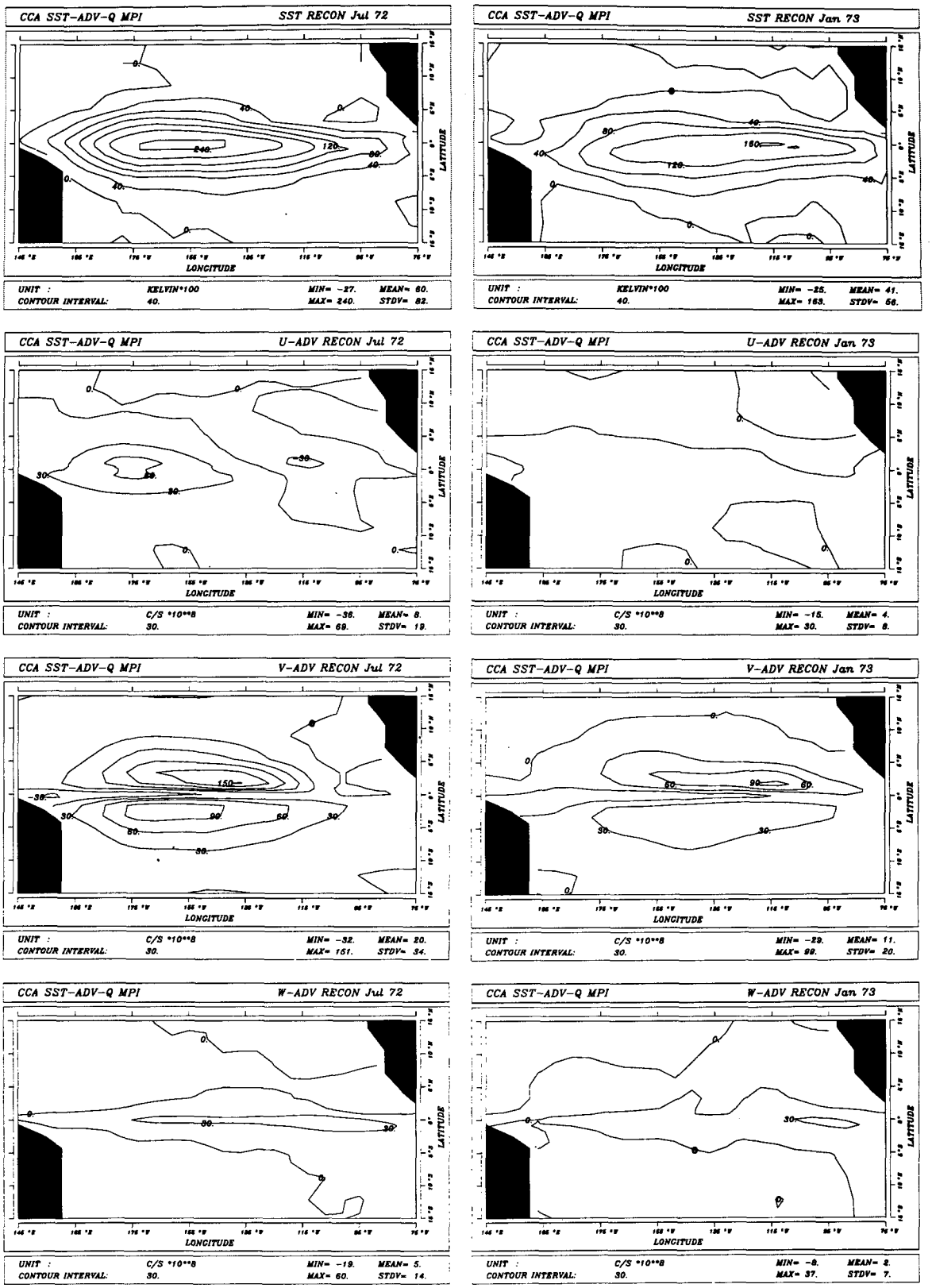


FIG. 10. As in Fig. 9 but for MPIZ model. (CI = 30 for zonal, meridional, and vertical advection fields.)

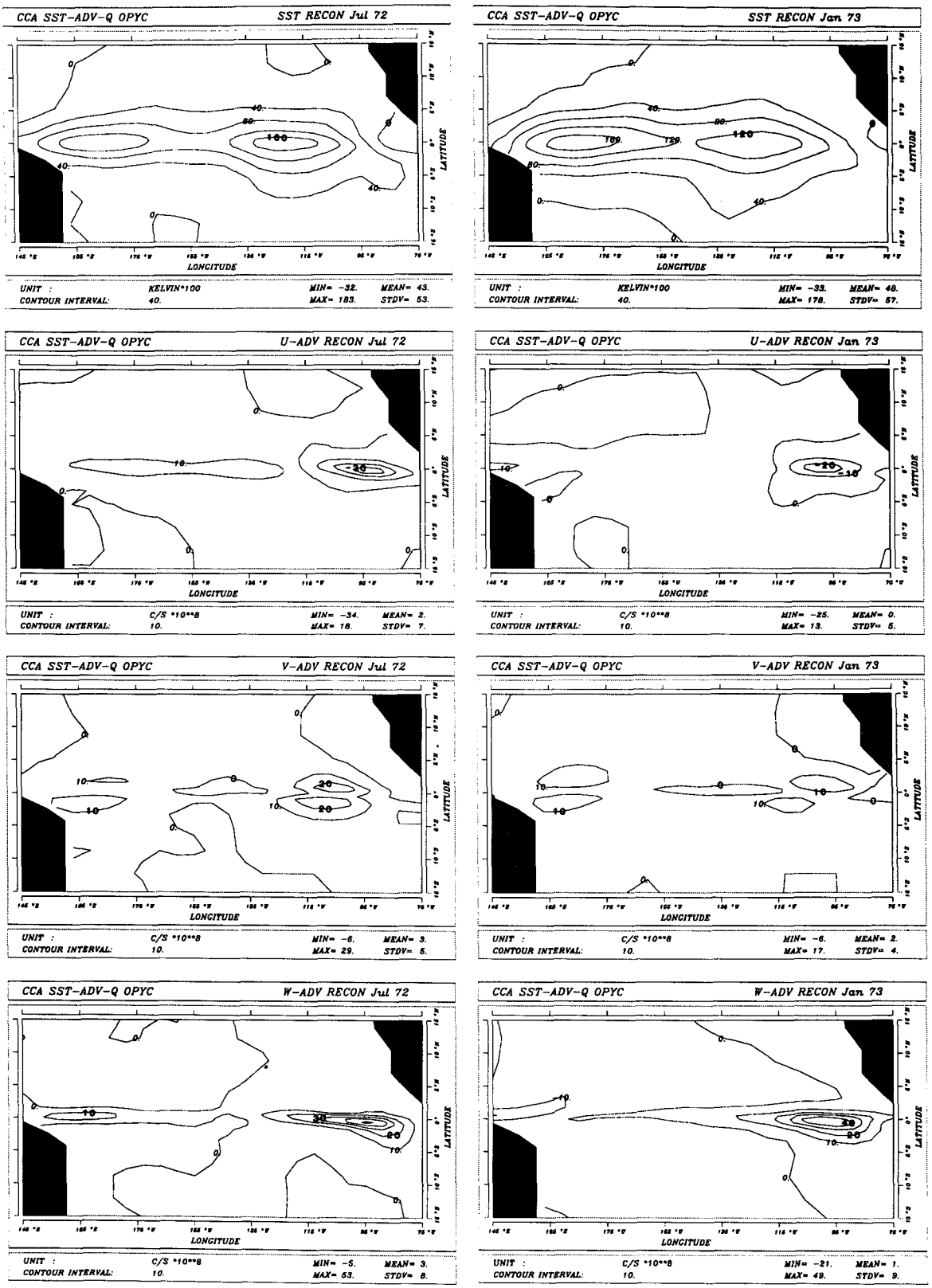


FIG. 11. As in Fig. 9 but for OPYC model.

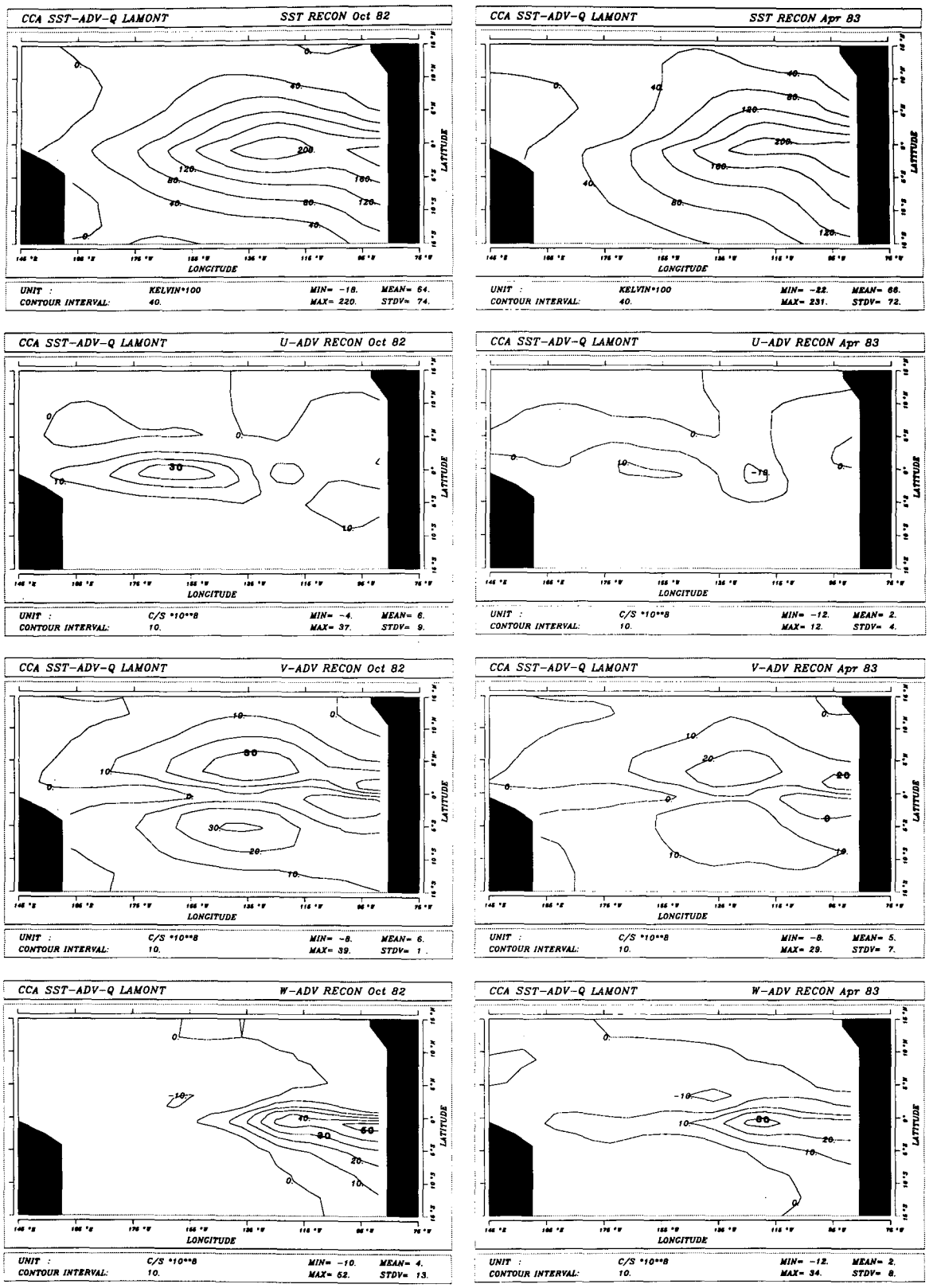


FIG. 12. As in Fig. 6 but for reconstructions of the 1982–83 El Niño, for (left) October 1982 and (right) April 1983 for the Lamont model.

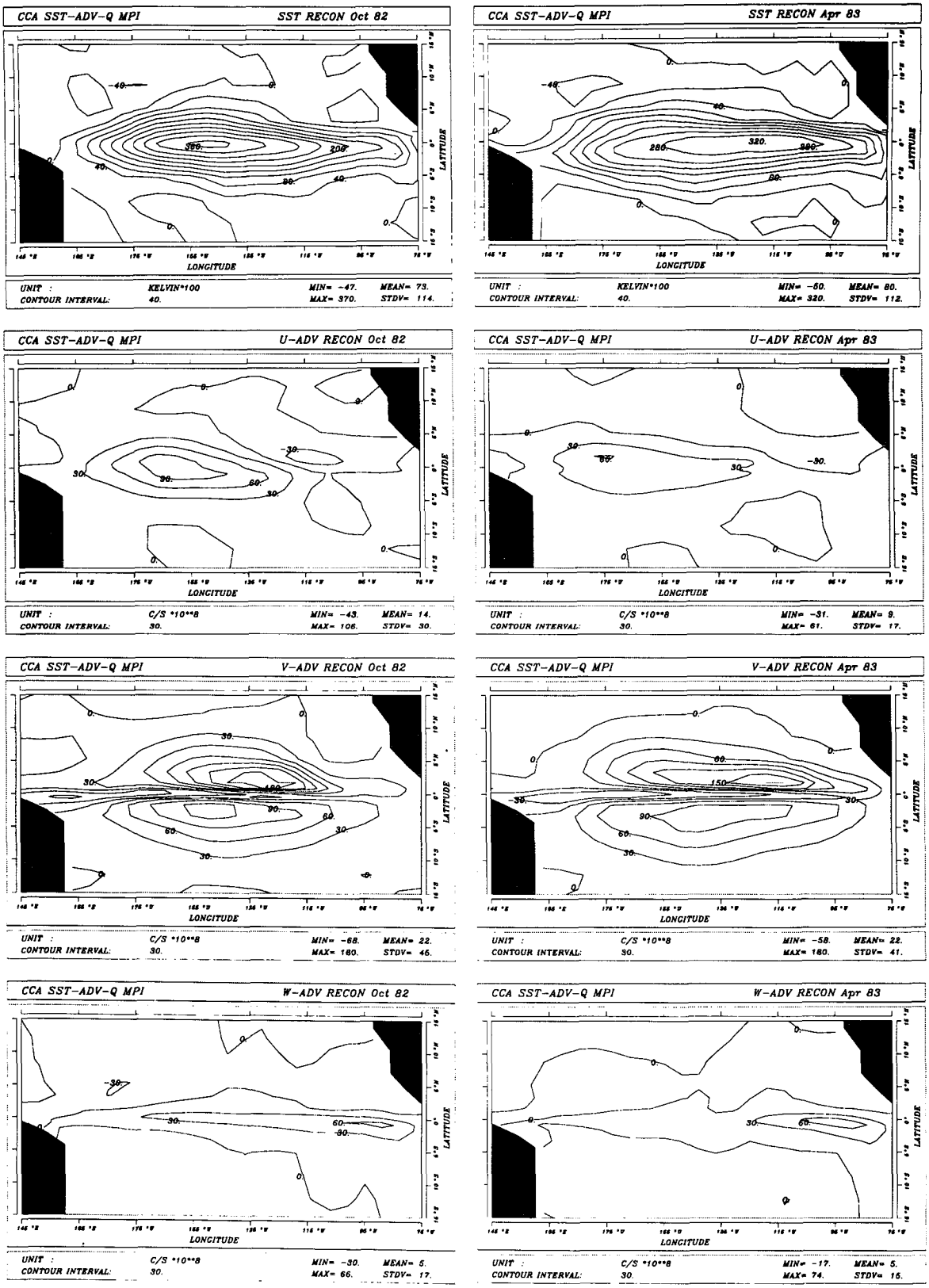


FIG. 13. As in Fig. 12 but for MPIZ model. (CI = 30 for zonal, meridional, and vertical advection fields.)

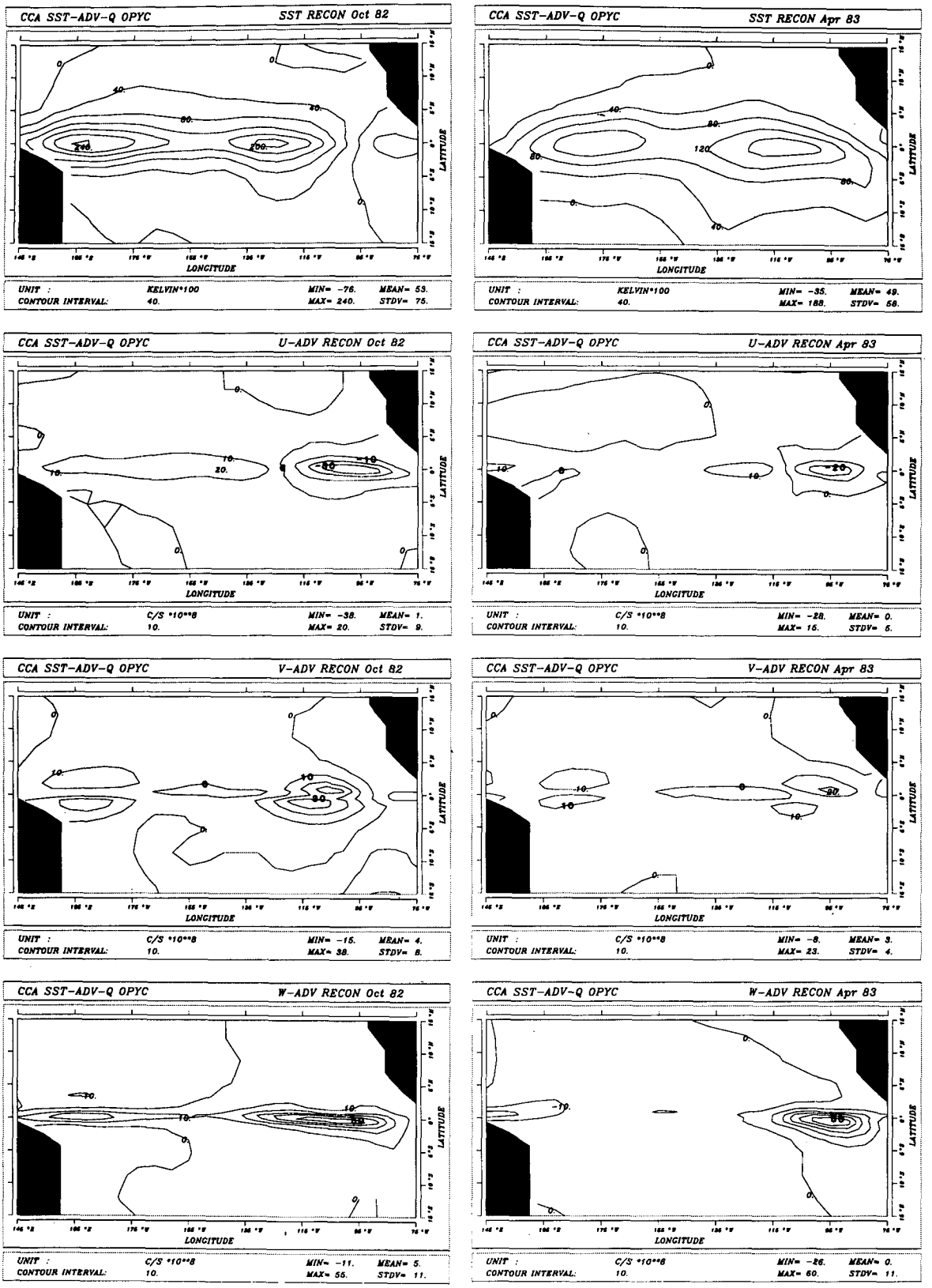


FIG. 14. As in Fig. 12 but for OPYC model.

served anomalies intensified strongly near the coast and decreased in amplitude in the central basin and continued this behavior into the summer of 1983. This El Niño development pattern is unusual in that eastern equatorial Pacific warming usually precedes central tropical Pacific warming (e.g., Harrison and Schopf 1984; Philander 1990).

For the Lamont model, the reconstruction reveals that a strong SST anomaly developed essentially in place in the eastern basin near 120°W during fall 1982 and retained that spatial pattern during the subsequent winter and spring. Though only a very weak eastward shift in the peak SST anomaly is seen in the transition from fall to winter in Fig. 15, it should again be noted that the broad latitudinal scale of the response and general pattern of SST seen in the Lamont model is quite similar to the observed SST in the eastern basin.

The MPIZ model reconstruction shows a warm SST anomaly that develops in the central Pacific during July 1982 and then grows in amplitude and spreads eastward in the fall and winter seasons. During spring 1983, the anomaly subsides without changing shape. A clear eastward propagation of SST is, therefore, evident in Fig. 15, although the anomaly does not intensify strongly along the eastern boundary as was observed. Furthermore, the MPIZ model anomaly is too narrowly confined to the equator and has a zonal scale that is too broad compared to the observed.

The reconstruction from the OPYC model begins to develop a warm anomaly near 130°W in July 1982, but a rather cold anomaly persists from the antecedent spring to the east. During the next two seasons the anomaly grows and spreads eastward, with the SST maximum continuing to migrate eastward through the fall of 1983. However, the model SST anomaly fails to develop a sizable maximum along the eastern boundary until summer 1983, several months later and with greater intensity than observed. The OPYC model SST anomaly is narrower in meridional scale and slightly greater in latitudinal extent than the MPIZ model anomaly. Thus, the OPYC model, like the MPIZ model, clearly exhibits an eastward migration of SST though the model response develops more slowly and is less coastally intensified than the observed.

In contrast to the 1982–83 event, the observed 1972–73 El Niño (not shown) developed most strongly in the eastern basin around 100°W during the spring and summer of 1972, then intensified in the central Pacific during the fall and remained strong until the spring of 1973. As pointed out previously, the three model SST reconstructions tend to develop similarly for the 1972–73 and 1982–83 events, but with an approximately three-month phase difference and, of course, different amplitudes. Therefore, the Lamont model in this case, which we have just described as developing SST essentially in place near the eastern boundary, more closely conforms to the observed development of the 1972–73 SST anomaly than the two GCMs. The MPIZ model SST reconstruction develops essentially as described for the 1982–83 event, while the OPYC model, 1972–73 anomaly reconstruction differs slightly in that it tends to develop in place, centered on 120°W, without propagation during the event.

d. Summary of CCA reconstructions

The primary result from this analysis is that the three models differ in the detailed manner that the reconstructed El Niños develop. However, in any one model, both the 1972–73 and 1982–83 events develop similarly. Although the timing and amplitude of the SST development is different for each event, for both the 1972–73 and 1982–83 El Niños, the Lamont model SST field develops in the eastern basin and persists there. It demonstrates no clear eastward propagation of information. In contrast, the MPIZ model develops warm SST anomalies, first near 160°W, which then grow and, after several months, spread eastward across the basin (Barnett et al. 1991) for both El Niño events. The OPYC model response (ignoring the excessive SST variability in the western basin) is somewhat similar to the MPIZ model except that the initial SST development occurs farther east, near 120°W, before developing and spreading eastward for the 1982–83 event (or developing in place for the 1972–73 event).

Note that the GCMs must predict the mean SST, currents, and vertical structure fields, as well as the anomalous fields that interact with the mean. The simple model uses mean fields specified from observations and predicts only the anomalies. Since the mean fields can vary strongly throughout the seasonal cycle, their resultant effects on the timing and strength of the terms in the heat budget should be kept in mind (see, e.g., Harrison and Schopf 1984).

A final point concerns the visual correspondence between the model and observed SST anomalies (Fig. 15) in the light of the hindcast scores presented in Figs. 1–5. Casual inspection of Fig. 15 leads one to assume that the Lamont model is superior to the GCMs because the model SST anomaly is intensified in the eastern basin and possesses a broad latitudinal scale, both features of which are observed but are superficially absent from the GCMs. In contrast, the hindcast scores suggest that the Lamont model is superior to the GCMs only in the far eastern basin. The rationalization of this seeming inconsistency lies in the fact that Fig. 15 shows only four months during the 1982–83 event, while the hindcast scores include 16 years worth of variability. This leads to questions of how to weight a given model's ability to reproduce the spatial pattern of El Niño versus its ability to simulate the timing of the event, or how to rate the relative virtues of high correlation and low rms error. Furthermore, it may be more or less desirable to generate a "great" hindcast for a key event, such as the 1982–83 El Niño, rather than to generate a long period of "decent" hindcasts for less celebrated time

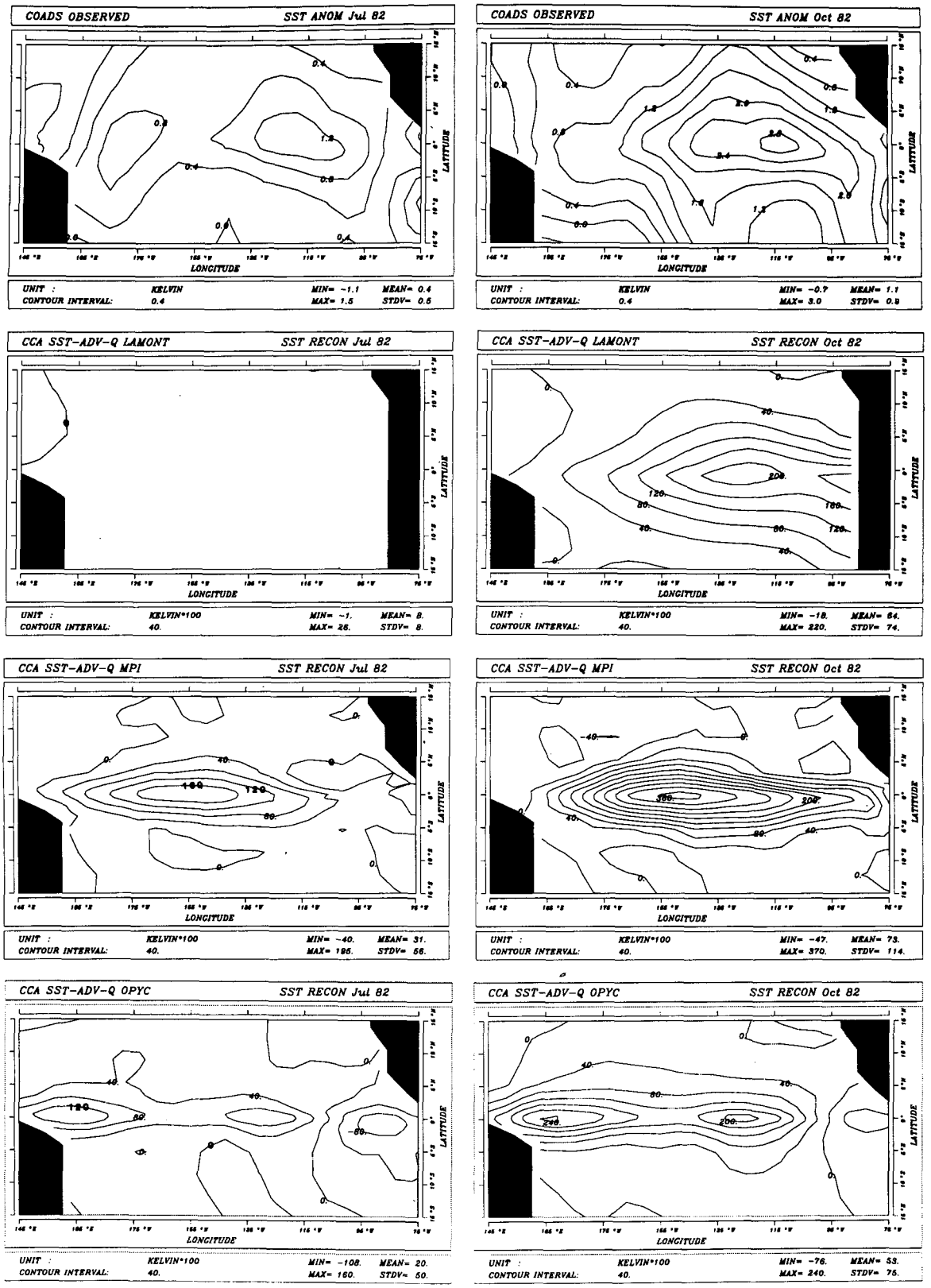


FIG. 15. SST anomalies for (left to right) July 1982, October 1982, January 1983, April 1983 for (top to bottom) COADS observations (SST in °C; $CI = 0.4$) and CCA reconstructions (SST in °C × 100; $CI = 40$) from Lamont, MPIZ, and OPYC models.

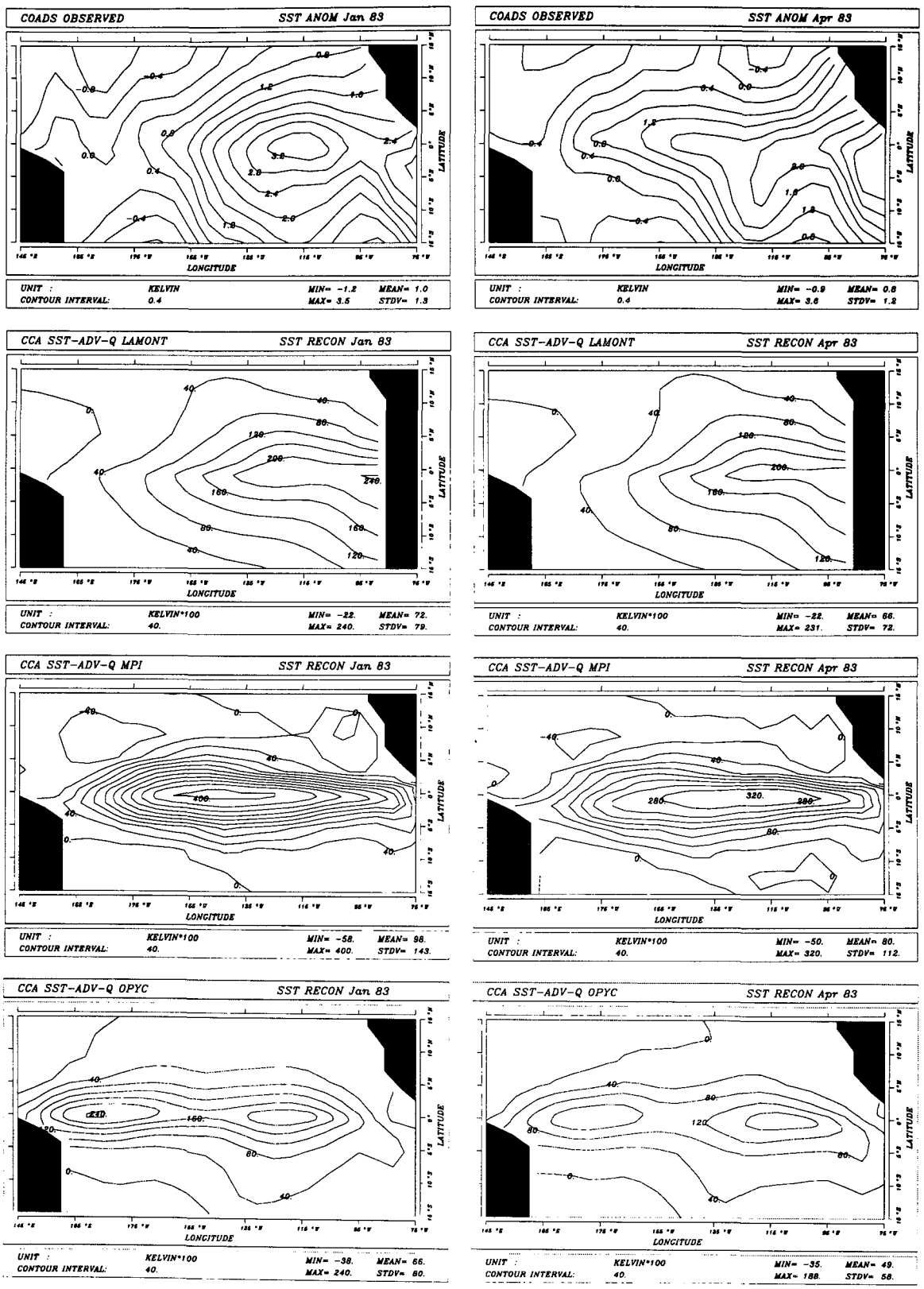


FIG. 15. (*Continued*)

intervals. The relative merits of these types of model characteristics depend, of course, on the intended application.

5. Summary and discussion

The hindcast skill analysis has shown that each of the four models considered here exhibits significant skill in some region of the basin. The four models, however, tend to have different regions wherein the skill is maximized, suggesting that further improvements in model skill are indeed achievable. The Lamont model skill is maximized in the eastern basin, particularly near the South American coast, a region where the three GCMs are less skillful presumably due to the dominance of vertical mixing mechanisms there (which are the most questionable aspects of the GCMs). The three GCMs, in contrast, exhibit broad regions of skill in various parts of the central Pacific (the key region for atmospheric sensitivity), but tend to poorly represent the SST in the far western basin (primarily due to overestimated SST variability there).

The canonical correlation analysis showed that, among the three models considered, similar spatial relationships exist for the mechanisms that excite SST (zonal advection, meridional advection, and vertical advection/mixing), but the balance of the strength of these terms was different for each model. Vertical advection (variations in upwelling) tended to control the large-scale structure of SST in the Lamont model (cf. Battisti 1988; Seager 1989). In the MPIZ model, meridional advection provided the dominant large-scale forcing for SST anomalies (cf. Barnett et al. 1991). For the OPYC model, all three terms were important in the region of developing SST (cf. Miller et al. 1992). This result was also evident in the CCA reconstructions of the El Niño events of 1972–73 and 1982–83. In essence, all three models seem to be working but for different physical reasons.

Differences in the spatial development of the model SST anomalies were also revealed in the CCA analysis. The Lamont model did not exhibit any clear eastward propagation of SST, the MPIZ model propagated SST anomalies eastward for both the 1972–73 and 1982–83 El Niño events, while the OPYC model propagated SST eastward for the 1982–83 El Niño and developed SST in place for the 1972–73 El Niño. The observed SST for the 1982–83 event exhibited an eastward propagation, while the 1972–73 El Niño developed first in the eastern basin, then in the central basin. In terms of the spatial patterns of SST, however, the Lamont model exhibits eastern intensification and broad latitudinal scale, which is more similar to the observed patterns of SST during El Niño than are the patterns of the GCMs.

Collectively, our results suggest that further improvements in these models are accessible since each model tends to maximize skill in different regions of the basin. The eventual improvements will then help us both to better predict El Niños and to better diagnose the relationship of developing SST anomalies to the oceanic flux processes that drive and dampen them. Other physics not included in the models presented here, such as the effects of surface heat fluxes exciting tropical SST anomalies (e.g., Seager 1989), must additionally be considered. It is also essential that existing (or new) observations be analyzed to determine which set of model physics is most applicable to the real ocean.

Last, the reviewers raised some important, specific questions that we found difficult to conclusively answer within the scope of our study, for instance, What is the source of the problems that the GCMs have in the western Pacific? Can runs be made with the GCMs in which a modified heat flux parameterization is applied to remove the bias in the west Pacific? Why is meridional advection so important in the MPIZ model relative to the OPYC model and the Lamont model? Would an extended CCA analysis better distinguish the propagation characteristics of SST (and its forcing terms) among the models? If a model gets the spatial distribution correct but always gets the phase wrong, is it worse or better than a model that get the patterns wrong but the phases correct? Although we have included some partial and some speculative answers to these questions throughout the paper, we hope the reviewers' questions will help to provoke further and more definitive studies of these issues.

Acknowledgments. Support was provided by NOAA Grants NA90-AA-D-CP526 and NA16RC0076-01, NASA Grant NAG5-236, NSF Grant ATM88-14571-03, the G. Unger Vetlesen Foundation and the University of California INCOR program for Global Climate Change. AJM also acknowledges the hospitality and support of the Institute of Geophysics and Planetary Physics of the Los Alamos National Laboratory while he was an Orson Anderson Visiting Scholar. Supercomputing resources were provided by the Department of Energy through Lawrence Livermore National Laboratory and by the National Science Foundation through the San Diego Supercomputer Center and the National Center for Atmospheric Research. Tony Tubbs and Jack Ritchie of SIO executed much of the computational analysis of the data from the models and Marguerette Schultz drafted the figures. We thank Ben Giese for supplying the GFDL model SST fields and Josef Oberhuber for insightful comments on the manuscript. As referees, Richard Seager and Mojib Latif provided exceptionally helpful detailed suggestions for improving the clarity and interpretation of our results.

REFERENCES

Barnett, T. P., and R. Preisendorfer, 1987: Origins and level of monthly and seasonal forecast skill for the United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.*, **115**, 1825–1850.

——, M. Latif, E. Kirk, and E. Roeckner, 1991: On ENSO Physics. *J. Climate,* **4,** 487–515.

Battisti, D. S., 1988: Dynamics and thermodynamics of a warming event in a coupled tropical atmosphere–ocean model. *J. Atmos. Sci.,* **45,** 2889–2919.

Busalacchi, A. J., and J. J. O'Brien, 1981: Interannual variability in the equatorial Pacific in the 1960's. *J. Geophys. Res.,* **86,** 10 901–10 907.

Gill, A. E., 1983: An estimation of sea level and surface current anomalies during the 1972 El Niño and consequent thermal effects. *J. Phys. Oceanogr.,* **13,** 586–606.

Goldenberg, S. B., and J. J. O'Brien, 1981: Time and space variability of tropical Pacific wind stress. *J. Phys. Oceanogr.,* **11,** 1190–1207.

Goswami, B. N., and J. Shukla, 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate,* **4,** 3–22.

Graham, N. E., J. Michaelson, and T. P. Barnett, 1987: An investigation of El Niño–Southern Oscillation cycle with statistical models. *J. Geophys. Res.,* **92,** 14 251–14 270.

Haney, R. H., 1971: Surface thermal boundary conditions for ocean circulation models. *J. Phys. Oceanogr.,* **1,** 241–248.

Harrison, D. E., 1992: The tropical Pacific seasonal cycle of SST: An ocean circulation model simulation and its thermal balance. *J. Climate,* submitted.

——, and P. S. Schopf, 1984: Kelvin-wave-induced anomalous advection and the onset of surface warming in El Niño events. *Mon. Wea. Rev.,* **112,** 923–933.

——, B. S. Giese, and E. S. Sarachik, 1990: Mechanisms of SST change in the equatorial waveguide during the 1982–83 ENSO. *J. Climate,* **3,** 173–188.

Hayes, S. P., P. Chang, and M. J. McPhaden, 1991: Variability of the sea surface temperature in the eastern equatorial Pacific. *J. Geophys. Res.,* **96,** 10 553–10 566.

Hurlburt, H. E., J. C. Kindle, and J. J. O'Brien, 1976: A numerical simulation of the onset of El Niño. *J. Phys. Oceanogr.,* **6,** 621–631.

Latif, M., 1987: Tropical ocean circulation experiments. *J. Phys. Oceanogr.,* **17,** 246–263.

Liu, W. T., 1988: Moisture and latent heat flux variabilities in the tropical Pacific derived from satellite data. *J. Geophys. Res.,* **93,** 6749–6760.

McCreary, J. P., 1976: Eastern tropical ocean response to changing wind systems with application to El Niño. *J. Phys. Oceanogr.,* **6,** 632–645.

McPhaden, M. J., and S. P. Hayes, 1991: On the variability of winds, sea surface temperature, and surface layer heat content in the western equatorial Pacific. *J. Geophys. Res.,* **96,** 3331–3342.

Miller, A. J., J. M. Oberhuber, N. E. Graham, and T. P. Barnett, 1992: Tropical Pacific Ocean response to observed winds in a layered general circulation model. *J. Geophys. Res.,* **97,** 7317–7340.

Neelin, J. D., and coauthors, 1992: Tropical air–sea interaction in general circulation models. *Climate Dyn.,* **7,** 73–104.

Oberhuber, J. M., 1993: Simulation of the Atlantic circulation with a coupled sea ice–mixed layer–isopycnal general circulation model. Part I: Model Description. *J. Phys. Oceanogr.,* **23,** 808–829.

Pacanowski, R., K. Dixon, and A. Rosati, 1991: User's Guide to GFDL Modular Ocean Model. Geophysical Fluid Dynamics Laboratory, Princeton University, 100 pp.

Philander, S. G. H., 1990: *El Niño, La Niña, and the Southern Oscillation,* Academic Press, 293 pp.

——, and A. D. Siegel, 1985: Simulation of El Niño of 1982–1983. *Coupled Ocean-Atmosphere Models.* J. Nihoul, Ed., Elsevier, 517–541.

——, W. J. Hurlin, and A. D. Siegel, 1987: Simulation of the seasonal cycle of the tropical Pacific Ocean. *J. Phys. Oceanogr.,* **17,** 1986–2002.

Seager, R., 1989: Modeling tropical Pacific sea surface temperature: 1970–87. *J. Phys. Oceanogr.,* **19,** 419–434.

——, S. E. Zebiak, and M. A. Cane, 1988: A model of the tropical Pacific sea surface temperature climatology. *J. Geophys. Res.,* **93,** 1265–1280.

Zebiak, S. E., and M. A. Cane, 1987: A model ENSO. *Mon. Wea. Rev.,* **115,** 2262–2278.