Q. J. R. Meteorol. Soc. (2000), **126**, pp. 2265–2291

Dynamical seasonal predictions with the COLA atmospheric model

By J. SHUKLA*, D. A. PAOLINO, D. M. STRAUS, D. DE WITT, M. FENNESSY, J. L. KINTER, L. MARX and R. MO

Center for Ocean–Land–Atmosphere Studies, USA

(Received 1 March 1999; revised 9 August 1999)

SUMMARY

Seasonal hindcasts were made using the COLA (Center for Ocean–Land–Atmosphere Studies) model for 16 winter seasons (mid-December through March for 1981/82 to 1996/97). For each season, a nine-member ensemble was generated using observed initial conditions in mid-December and observed global sea surface temperatures (SSTs). It is found that in the presence of large tropical SST anomalies, the model is quite successful in simulating seasonal-mean height anomalies over the Pacific–North America (PacNA) region.

A local spatial pattern correlation field is computed for the ensemble seasonal mean of 500 hPa height from the General Circulation Model (GCM) and the seasonal means from the re-analyses of the National Centers for Environmental Prediction, hereafter 'observations'. This field exceeds 0.6 over the eastern tropical Pacific and western North America; the maximum values are greatly enhanced during El Niño Southern Oscillation (ENSO) years. Similar results are obtained for 200 hPa winds. The observed enhancement of intraseasonal low-pass (10–90 day) variability of 500 hPa height during cold events is simulated, as is the shift in the storm tracks (2–10 day variability of 850 hPa meridional heat flux) during warm events.

Empirical orthogonal function (EOF) analysis is applied in the PacNA region to the ensemble seasonal means of GCM 500 hPa height and the corresponding observed seasonal means; the global combined Arctic/North Atlantic Oscillation is removed from both GCM and observations. The leading EOF mode explains about 50% of the variance for both GCM and observations; the two patterns are nearly identical. Singular-value decomposition (SVD) analysis between the tropical Pacific SST and 500 hPa height in the PacNA region shows that the nature of the coupling between SST and height is nearly the same in the GCM and observations. The great similarity between the height patterns in the first SVD mode and the corresponding leading mode EOF pattern indicates that the leading height variations are forced by the SST.

The SST-forced variance of height was also estimated by regression analysis of (ensemble) seasonal means for the GCM and observations for the 16 years onto an index of tropical Pacific SSTs derived from SVD analysis of a long (30-year) record of observed heights and SSTs. The pattern of percentage variance explained in the GCM and in the observations are very similar to each other (and to the EOF described above). The higher absolute values in the GCM case reflect the effectiveness of the ensemble in filtering out variability unrelated to SST forcing.

SVD analysis was applied to 100 GCM 'samples' coupled with the observations; a sample is defined as a single 16-year record obtained by picking one ensemble member randomly for each of the 16 years. Probability distribution functions (p.d.f.s) of the pattern correlation for the leading SVD patterns, the percentage explained squared covariance, and the time series correlations all indicated sharp peaks at values of 0.87, 87%, and 0.82 respectively.

The p.d.f. of the projection of individual 5-day means onto the leading EOF described above is quite dramatically shifted during strong warm and cold tropical SST events; the warm (cold) event p.d.f. has almost all its weight in the negative (positive) EOF region. GCM and observations agree well.

The intra-ensemble spread was estimated by computing the PacNA anomaly correlation coefficient (ACC) for each of 36 possible intra-ensemble pairs. For eight years in which histograms of the ACC indicate predominantly positive values, the ACC of the ensemble mean with the observed seasonal mean is also relatively high.

Brier skill scores and reliability diagrams were computed for the 'event' of the 500 hPa height being one standard deviation above (or below) the normal, with all such events pooled over the entire northern hemisphere or North America only. All skill scores are positive and statistically significant at the 99% level. The North American scores are higher than the whole hemisphere scores; the ENSO year scores are higher than those for all years.

KEYWORDS: Ensemble prediction Predictability Statistical techniques

1. INTRODUCTION

Nearly 20 years ago, it was suggested that the year-to-year variations of the seasonal mean atmospheric circulation and rainfall in the tropical regions are largely determined by slowly varying boundary conditions at the earth's surface, notably sea surface temperature (SST) and land surface conditions (Charney and Shukla 1981; Shukla 1981a). This suggestion was based on sensitivity and predictability calculations using atmospheric

* Corresponding author: Center for Ocean–Land–Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705-3106, USA. e-mail: shukla@cola.iges.org

general circulation models (AGCMs) in which the model-simulated variances of circulation and rainfall with interannually varying SST (hereafter referred to as the signal) were compared to the variances generated by the internal dynamics of the atmosphere alone (to be referred to as noise). The noise was estimated by integrating an AGCM with climatological SSTs. Such calculations were repeated using several AGCMs and, without exception, all models showed that the signal in the tropical regions was far larger than the noise, at a very high level of statistical significance. The evidence for the existence of extratropical predictability had been difficult to find. Lau (1981) showed that model simulations with climatological SSTs could capture 'the most prevalent anomaly patterns observed in the atmosphere'. Estimates of predictability based on model-simulated signal and noise gave a generally pessimistic view about the prospects for dynamical seasonal prediction in mid latitudes (Chervin 1986). The pioneering monthly-prediction experiments by Miyakoda *et al.* (1983, 1986) had shown the feasibility of dynamical monthly prediction, and studies on the dynamical predictability of monthly means had shown a scientific basis for predictability (Shukla 1981b). However, it was clearly understood that it was the deterministic predictability of the first ten days that contributed to the dynamical predictability of the monthly mean. It was shown unambiguously that, in the absence of the possible influence of boundary conditions, there was no prospect for predicting monthly means even for the second month. These results further contributed to the scepticism about the prospects for dynamical seasonal prediction in extratropical regions.

This impression was not consistent with the fact that the observed correlation coefficient (CC) between tropical SST anomalies and mid-latitude circulation anomalies, especially over the Pacific–North America (PacNA) region, was 0.6 or higher (Horel and Wallace 1981). It was also found that the amplitude of the extratropical height anomalies simulated by some of the early models were too small by a factor of two or three compared with the observed anomalies (see Lau 1997, for a review). For some time it remained an unresolved question as to whether the predictability of seasonal averages in the mid latitudes is limited by the chaotic nature of the internal dynamics, or whether the deficiency of the early AGCMs produced such a weak boundary-forced response that this response was overwhelmed by the noise associated with internal dynamics (Shukla 1998).

Recently, several modelling studies have suggested that the deficiencies of AGCMs could be primarily responsible for unrealistically low values of simulated interannual variability and the correspondingly low estimates of mid-latitude predictability. Kumar *et al.* (1996) showed that a rather small change in the treatment of the physical parametrizations made a significant difference to the amplitude of model-simulated, SST-forced mid-latitude height anomalies. In particular, the amplitude of the SST-forced height anomalies in the revised model nearly doubled over the PacNA region. Bengtsson *et al.* (1993) and Branković *et al.* (1994) showed that the height anomalies over the PacNA region using two different models were quite well-simulated. DeWitt (1996) showed that by replacing the Kuo convection scheme with the relaxed Arakawa–Schubert, or RAS, scheme (Moorthi and Suarez 1992) in the COLA AGCM, the amplitude of simulated mid-latitude height anomalies became comparable to the observed height anomalies, especially for cases with large tropical SST anomalies.

It is in this context that we proposed to conduct a well co-ordinated study to estimate the predictability of seasonal-mean anomalies over the PacNA region, which is the region of largest interannual variability of height anomalies*. We proposed a

* Calculation of the interannual variance of 500 hPa height using 16 winters of National Centers for Environmental Prediction re-analysis shows a maximum over the PacNA region.

multi-institution joint research project called Dynamical Seasonal Prediction (DSP) in the USA (Shukla *et al.* 2000). The DSP, and a similar European project called PROVOST (PRediction Of climate Variations On Seasonal to interannual Time-scales), also contribute to an international project called SMIP (Seasonal prediction Model Intercomparison Project) coordinated by the CLIVAR (Climate Variability) committee of the World Climate Research Program (WCRP).

There have been several simulations made using observed SSTs, conducted in a manner similar to that of the Atmospheric Model Intercomparison Project (AMIP; Gates *et al.* 1999). Some of these studies have carried out detailed and insightful analysis of seasonal predictability (Bengtsson *et al.* 1996; Kumar and Hoerling 1997). The motivation for the DSP project is different in one important respect: it has been proposed that, in addition to identical global SSTs and sea ice, all models will utilize the same initial conditions of the global atmosphere and, as far as possible, identical land surface conditions. Since in reality the SST itself must also be predicted, the idea here is to estimate the upper limit of seasonal predictability with the current generation of atmospheric models. These results can then be compared with the AMIP-type simulations, for which the initial conditions of the atmosphere and the land surface can differ greatly from observations. A systematic comparison of DSP integrations and AMIP integrations will help us to estimate the importance of using the most accurate initial conditions of atmosphere and land surface properties for seasonal prediction. The DSP study described here was unable to use the observed initial conditions of soil wetness and other land surface properties, because such a dataset does not yet exist. However, all participating models used observed global atmospheric initial conditions for seasonal integrations.

This paper describes the results of seasonal hindcasts made using the COLA model for 16 winter seasons (1981/82 through 1996/97). For each winter season, the COLA atmospheric model was integrated from the observed global atmospheric initial conditions produced by the National Centers for Environmental Prediction (NCEP) in mid-December, and the integration was carried through to 0000 UTC 1 April of the following year. Section 2 describes the COLA atmospheric model and the initial and boundary datasets. The mean model error is given in section 3. The ensemble-mean anomalies and low- and band-pass variance patterns simulated by the model are compared to those of the NCEP re-analyses in section 4. In addition, section 4 describes empirical orthogonal function (EOF) analysis of the 500 hPa height, and singular-value decomposition (SVD) analysis linking the height field with the tropical Pacific diabatic heating. The statistical reliability and reproductibility of the results using a variety of methods is described in section 5, while a summary and discussion is given in section 6.

2. COLA MODEL, INITIAL AND BOUNDARY DATASETS

(*a*) *COLA model*

The COLA AGCM is a global spectral model which is truncated rhomboidally at zonal wave number 40 (Kinter *et al.* 1997). The associated Gaussian grid on which the physical parametrizations are calculated has 128 points in the longitudinal direction and 102 points in the latitudinal direction. The vertical structure of the model is represented by 18 unevenly spaced levels using sigma as the vertical coordinate. The dynamical core of the model is based on a former version of the operational medium-range forecast model at NCEP (NMC 1988). The subgrid-scale physical parametrizations have been modified from those which were implemented in the original NCEP code.

This version of the model includes parametrizations of solar radiative heating, terrestrial radiative heating, cloud–radiation interaction, deep convection (RAS scheme), large-scale condensation, shallow convection, a turbulence closure scheme for subgrid-scale exchange of heat, momentum and moisture, and biophysically controlled interaction between the land surface with vegetation and the atmosphere (the simplified biosphere model or SiB). The simulated time mean and anomalous wind and geopotential height fields were found to be more realistic when using the RAS scheme compared with the previously used Kuo scheme (DeWitt 1996).

(*b*) *Experiments, initial conditions, and boundary conditions*

For each winter season from 1981/82 through 1996/97 the COLA atmospheric model was integrated from the observed global atmospheric initial conditions produced by the NCEP in mid-December, and the integration was carried through to 0000 UTC 1 April of the following year. The period January, February, March (JFM) is considered for most of the COLA model results presented here; the winter is labelled by the year corresponding to JFM. More specifically, for each winter season, a nine-member ensemble was generated using observed initial conditions for the nine consecutive 12-hour periods beginning at 0000 UTC 13 December.

The SST and sea ice data used in this study are based on Climate Prediction Center (CPC) weekly $1° \times 1°$ Optimum Interpolation Sea Surface Temperature (OISST) analyses (Reynolds and Smith 1994). The data were spatially interpolated by area-weighted averaging to the model Gaussian grid of (x, y) dimension (128, 102). Furthermore, the SSTs were adjusted to be consistent with the gridded, non-zero, ocean surface topography resulting from the transform of the surface spectral topography used in the model. The adjustment used was -6.5 degC km^{-1}. This gives a closer representation of the corresponding atmospheric temperature in the surface boundary layer. The sea ice locations are based on the lowest OISST value. The temperature over sea ice points undergoes no height correction. Due to bias corrections added to the OISST, the data were subsequently filtered in time using a 1-2-1 filter (ignoring sea ice points). This time-filtered dataset is used by the model, and is linearly interpolated in time to twice-daily values in all the integrations. All points above -2 °C are considered to be open-water points, and the value of SST is used for heat flux calculations. Points below -2 °C are treated as sea ice points with sea ice assumed to extend over the entire Gaussian grid box. The sea ice temperature computation is based on an energy balance calculation, and is carried forward in time until the SST rises above -2 °C.

The initial soil wetness was obtained from a climatology calculated from the soil wetness analysis of the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis–forecast system for 1987–1993 (see Fennessy and Shukla 1999). After initialization, the COLA model determines the evolution of the soil wetness fields. The operational version of the ECMWF model for that period used three soil layers: the top two layers were prognostic and the bottom layer was prescribed according to time-varying climatological values. Both ECMWF prognostic soil moisture layers, the surface layer with a maximum capacity of 20 mm of liquid water and the deep soil layer with a maximum capacity of 120 mm of liquid water, were used for initialization. The bottom ECMWF soil moisture layer was not used. Since the land surface parametrization used in the ECMWF model is different from the SiB used in the COLA model, the ECMWF soil moisture cannot be used directly. Sato *et al.* (1989) developed a method to transform soil moisture calculated by other land models to be consistent with the SiB.

Essentially, the method calculates the time-integral of evaporative demand to which a model grid area would have to be exposed in order to dry from saturation to a specified level. This same time-integral is then applied to a greatly reduced set of SiB energy balance equations to calculate an equivalent SiB soil moisture level. Soil wetness is the soil moisture content expressed as a fraction of the maximum liquid water capacity for a given grid area. Since the SiB carries three prognostic soil wetness variables (the surface layer, the root zone and the gravitational drainage zone), the procedure outlined by Fennessy *et al.* (1994) is used to transform the two combined ECMWF prognostic soil moisture layers into a soil wetness; this is used to intialize all three SiB prognostic soil wetness layers.

The definition of the initial snow mass on land points is based on the values of albedo. For albedo values in the range 69–75%, 49–69% and 40–49% the initial snow mass is prescribed to be 20, 10 and 5 kg m^{-2}, respectively. The initial snow depth produced by this procedure is far less than the observed snow depth values. For land points with albedo greater than 75% and land ice points, the initial snow mass is prescribed to be 3×10^3 kg m^{-2}.

3. MEAN MODEL ERROR

The seasonal means for JFM for all 16 years (1982–1997) were constructed from the twice daily (0000 and 1200 UTC) values from all 144 integrations, and compared with observations from NCEP re-analyses for the corresponding 16-year period.

Figure 1(a) shows the JFM mean model error (generally referred to as the systematic error) for geopotential height at 500 hPa. The largest errors are found in the polar regions of the southern hemisphere where the model tends to produce an anomalous zonally symmetric gradient of geopotential between 40°S and the South Pole. The largest errors over the northern hemisphere (NH) are found over the European region. The zonally averaged mean error for height and temperature (not shown) retains the same sign for the entire depth of the troposphere and, for levels above 500 hPa, the amplitude of error increases with height. The model tends to be much too cold over the poles in the upper troposphere and, in general, most of the model atmosphere is colder than the NCEP re-analysis. Figure 1(b) shows the JFM mean error for zonal wind at 200 hPa. The most conspicuous feature is a large mean error of 15 m s^{-1} in the eastern equatorial Pacific area and the north-west part of South America. The magnitude of this error is so large that the observed weak easterlies over equatorial South America are replaced by spurious westerlies. The main features of the zonal wind error over the northern mid latitudes reflect a northward displacement of the jet stream in the model compared with the observations. The model-simulated core of the jet is also displaced eastward.

Figures 2(a) and (b) show the model mean JFM error for precipitation and surface (~2 m) temperature over land. The observed fields are taken from Xie and Arkin (1996) for precipitation, and Ropelewski *et al.* (1985) for temperature. The largest precipitation errors (2–5 mm day^{-1}) are found over the regions of largest climatological mean precipitation (equatorial Pacific, South America and southern Africa). Significant errors are also found in the Pacific and the Atlantic storm-track regions where the model tends to overestimate precipitation. The model has a cold bias over most of the land areas with the exception of north-west North America and western Eurasia between 40 and 55°N, where the model can be 2–5 degC warmer than the observations.

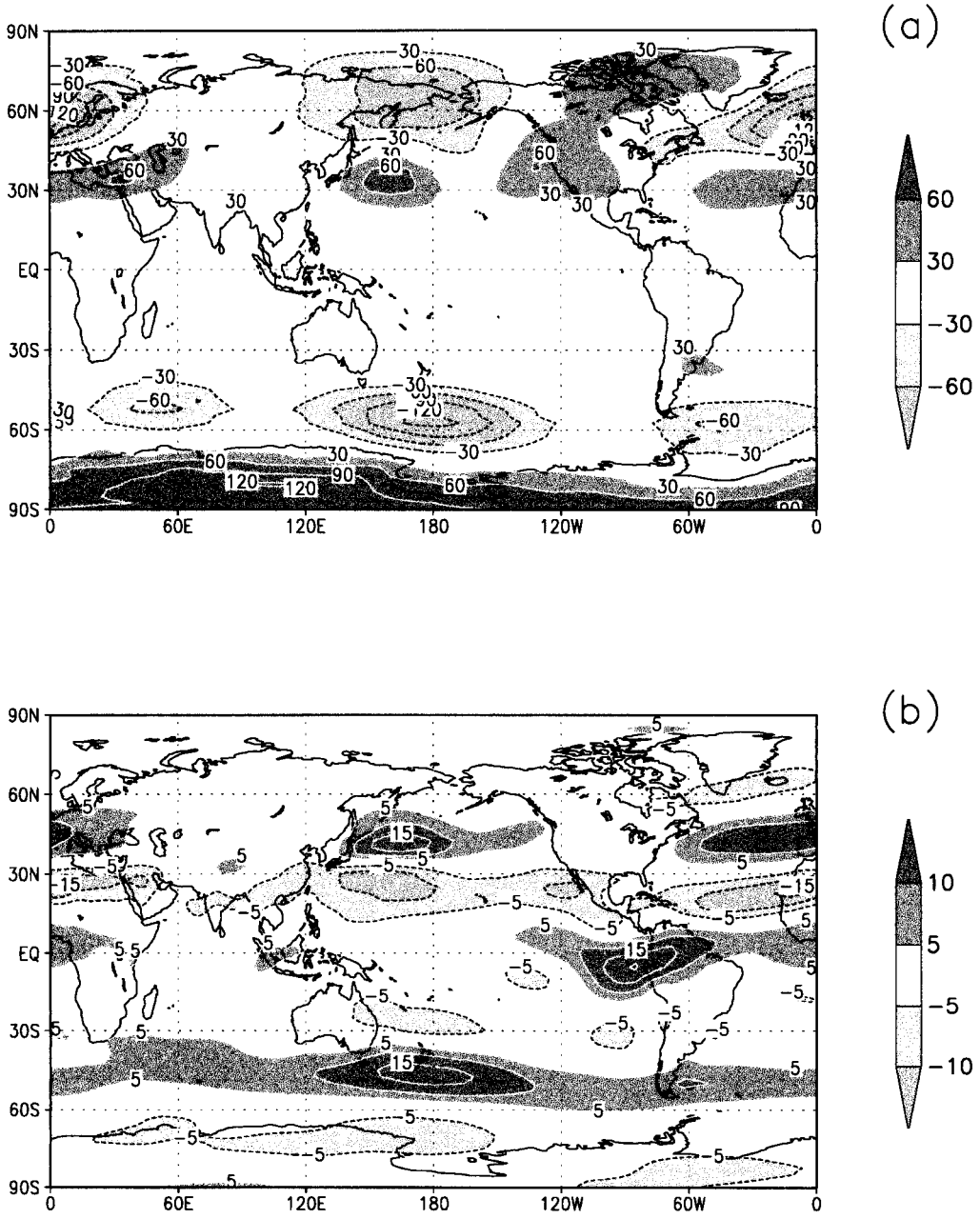

Figure 1. Mean difference between COLA DSP model forecasts for January to March from 144 integrations and NCEP re-analyses, over the years 1982 to 1997. This mean model error is generally referred to as the systematic error. (a) 500 hPa height (m); (b) 200 hPa zonal wind (m s^{-1}). See text for further details.

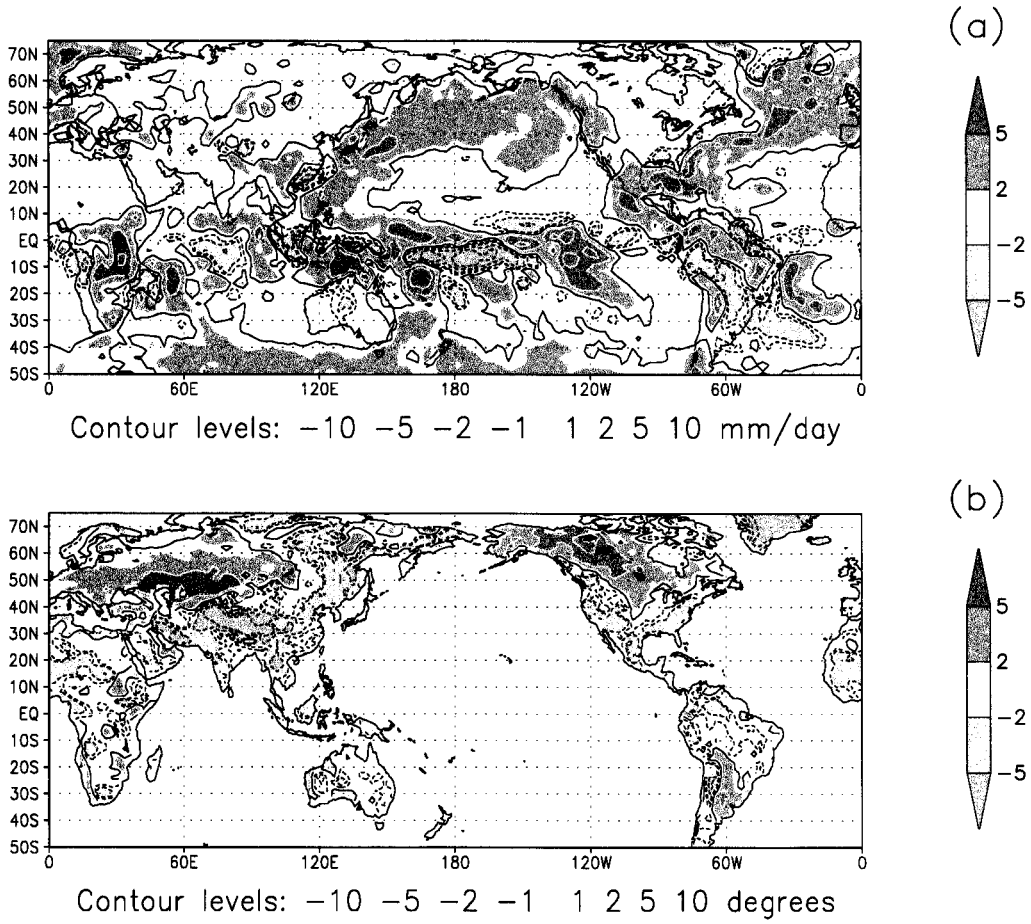

Figure 2. As Fig. 1 but observed fields from Xie and Arkin (1996), and for (a) precipitation (mm day^{-1}) and (b) surface (2 m) temperature (°C). Contours are at ± (1, 2, 5 and 10) in each frame.

4. Ensemble mean and intra-ensemble results

(*a*) *Height and temperature anomaly*

Figure 3 shows the composite ensemble-mean 500 hPa height difference (m) between the averages of three warm ENSO events (1983, 1987, and 1992) and the average of two cold ENSO events (1985, 1989) which occurred during the 16 winter periods, for the model (top) and the observations (bottom). The model simulates the observed anomalies quite well, with a pattern correlation of 0.98. There are several other similar model results presented in this volume which clearly show that models have generally improved to such an extent that they can now realistically simulate both the magnitude and the pattern of observed height anomalies. This is in sharp contrast to earlier models (see Figs. 2 and 4 of Lau 1997) for which model-simulated anomalies were weaker than the observed anomalies by factors of two or three. Figure 4 shows the composite 850 hPa temperature anomaly difference for the same three warm and cold events as shown in Fig. 3. Both the structure and amplitude of the temperature anomalies are simulated very well (pattern correlation of 0.85). However, the GCM has a tendency to over-predict the warm anomalies over north-east Canada.

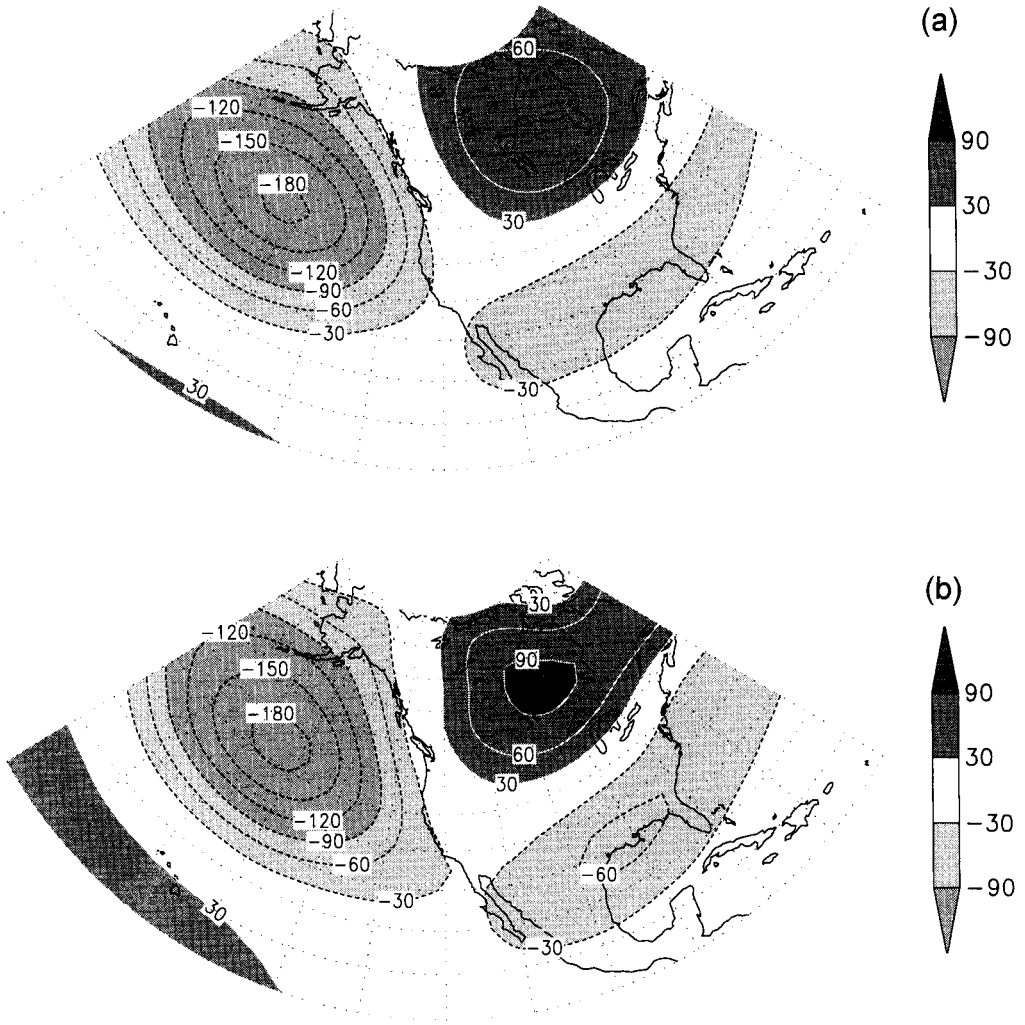

Figure 3. Ensemble-mean January to March 500 hPa height differences (m) between the average of warm ENSO years, 1983, 1987 and 1992, and the average of cold years 1985 and 1989: (a) COLA GCM difference; (b) Observed difference. The anomaly correlation coefficient of the patterns is 0.98. The contour intervals are 30 m. See text for details.

(*b*) *Anomaly correlation coefficients*

Anomaly correlation coefficients (ACCs) between ensemble-mean model anomalies and observed anomalies were computed for every year, to verify the model's ability to simulate seasonal atmospheric circulation anomalies. The model anomaly for any given year is calculated with respect to the model climatology, with the given year excluded. The same is true for the calculation of the observed anomalies. This procedure is equivalent to removing the systematic error of the model. The ACCs are the area-weighted correlations between observed and simulated JFM mean anomaly fields over a specific area. The ACCs for 500 hPa geopotential height are presented in Table 1 for several different regions, some of which overlap. The model shows skill in the simulation of the circulation over the PacNA region (defined as 15–70°N, 180–60°W). For the five ENSO events during this period (1983, 1985, 1987, 1989, 1992), the

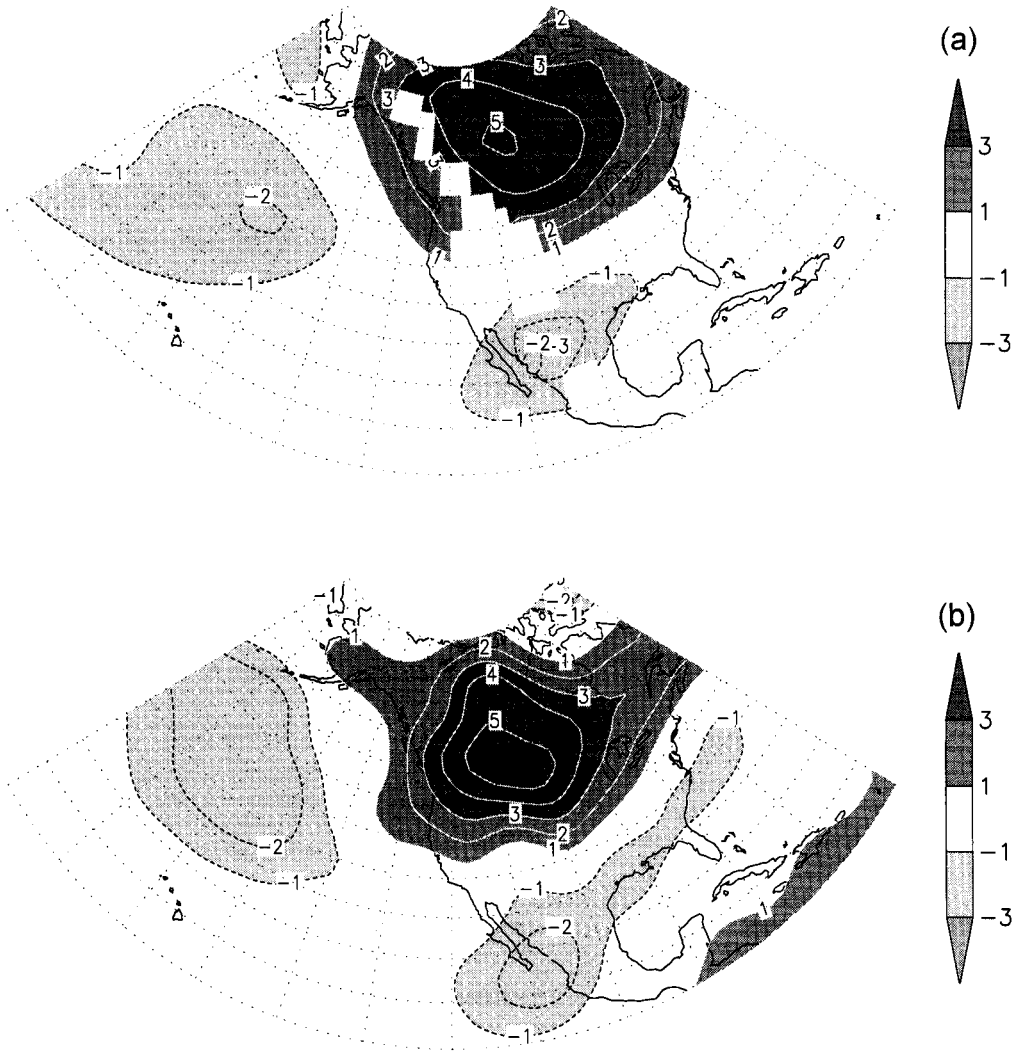

Figure 4. As Fig. 3 but for 850 hPa temperature differences (degC). The anomaly correlation coefficient of the patterns is 0.85.

ACC values for the PacNA region are 0.91, 0.49, 0.74, 0.84 and 0.64, respectively. Significance levels at five percent for these ACCs (shown at the bottom of Table 1) were estimated by a Monte Carlo test, in which forecasts were paired with observations for all years other than the forecast year. For the PacNA region a value of 0.64 for the ACC was found to be significant at the five percent level. There is only one non-ENSO year (1996) for which the PacNA ACC is significant.

In order to determine the geographic dependence of the ability of the ensemble-mean seasonal-mean simulations to capture the observed anomaly pattern, we introduce the pattern correlation maps (Straus and Shukla 1997) shown in Figs. 5 and 6. *At each point* the pattern correlation between the anomaly of the ensemble-mean JFM simulation and the observed winter average anomaly is calculated over a circular region, centred at the base point and having a radius of 2000 km. The resulting map is averaged for

TABLE 1. JANUARY TO MARCH ANOMALY CORRELATION COEFFICIENTS FOR 500 hPa HEIGHT FOR MODEL VERSUS OBSERVATIONS

Hindcast	SH	TR	TrPac	NH	PNA	NAm	EUR
1982	.32	−.08	−.12	−.13	−.09	.18	−.29
1983	.57	.81	.82	.75	.91	.94	.69
1984	.59	.22	.23	.40	−.10	.11	.86
1985	.28	.52	.72	.08	.49	.26	−.08
1986	.22	.15	.37	.08	.02	−.31	.23
1987	.20	.23	.44	.51	.74	.75	.30
1988	.19	.32	.57	.37	.53	.30	.39
1989	.57	.70	.86	.72	.84	.79	.54
1990	.29	.42	.26	.52	.60	.63	.76
1991	.42	.37	.52	.40	.18	.23	.74
1992	.33	.54	.72	.55	.64	.53	.77
1993	−.42	.31	.43	.24	−.06	.04	.26
1994	.66	.22	.05	−.23	−.56	−.16	−.24
1995	.56	.62	.75	.45	.46	.18	−.05
1996	.31	.32	.48	−.22	.67	.59	−.82
1997	.39	.29	.65	−.23	.15	−.12	−.64
5% Conf.	.58	.40	.53	.43	.65	.69	.74

SH, southern hemisphere (90°–20°S); TR, tropics (30°S–30°N); TrPac, tropical Pacific (150°–270°E, 30°S–30°N); NH, northern hemisphere (20°–90°N); PNA, Pacific–North America (15°–70°N, 180°–60°W); NAm, North America (25°–70°N, 150°–60°W); EUR, Europe (35°–75°N, 12.5°W–42.5°E).
5% Confidence limits were evaluated by Monte Carlo test for non-ENSO years.

the 16 winters using the Fischer Z-transform, and the result shown by the shading, for 500 hPa height (Fig. 5(a)) and 200 hPa zonal wind (Fig. 5(b)). The contours in these figures indicate the 16-winter average covariance between simulated and observed anomalies. Pattern correlations exceeding 0.6 (indicative of forecast skill) are seen only in the tropical Pacific, and for the upper-level winds in the subtropics off the west coast of North America. That this skill is related to ENSO is confirmed in Fig. 6(a) and (b), which shows the same correlation maps but averaged only over the strong ENSO winters during the DSP period: 1983, 1987, 1992 (warm) and 1985, 1989 (cold). Here correlation values above 0.6 (and for height, above 0.8) are seen not only in the eastern tropical Pacific but also over much of the PacNA region, with clearly enhanced values of the covariances.

(*c*) *Low-pass and band-pass variance*

The ability of the ensemble mean to reproduce the observed anomaly pattern in the PacNA region, suggests that the GCM is also able to reproduce shifts in the transient fluctuations associated with ENSO. Figure 7 compares the low-pass variance of 500 hPa height averaged over the three warm events (1983, 1987, 1992) from the high-resolution operational ECMWF analyses* (Fig. 7(a)) to the same quantity averaged over all ensemble members for the same warm winters (Fig. 7(c)). The cold event low-pass variance (averaged over the winters of 1985 and 1989) is given in the same manner in Figs. 7(b) and (d). (The low-pass variance captures fluctuations with periods of about 10–90 days, independent of the annual cycle, which is estimated as in Straus 1983. The low-pass filtering was accomplished by applying a digital filter after removal of the annual cycle.) For purposes of this calculation, the winter season is defined as the

* In general, mid-latitude intraseasonal variances are quite similar in the ECMWF operational analyses and the NCEP re-analyses (Straus and Yang 1997).

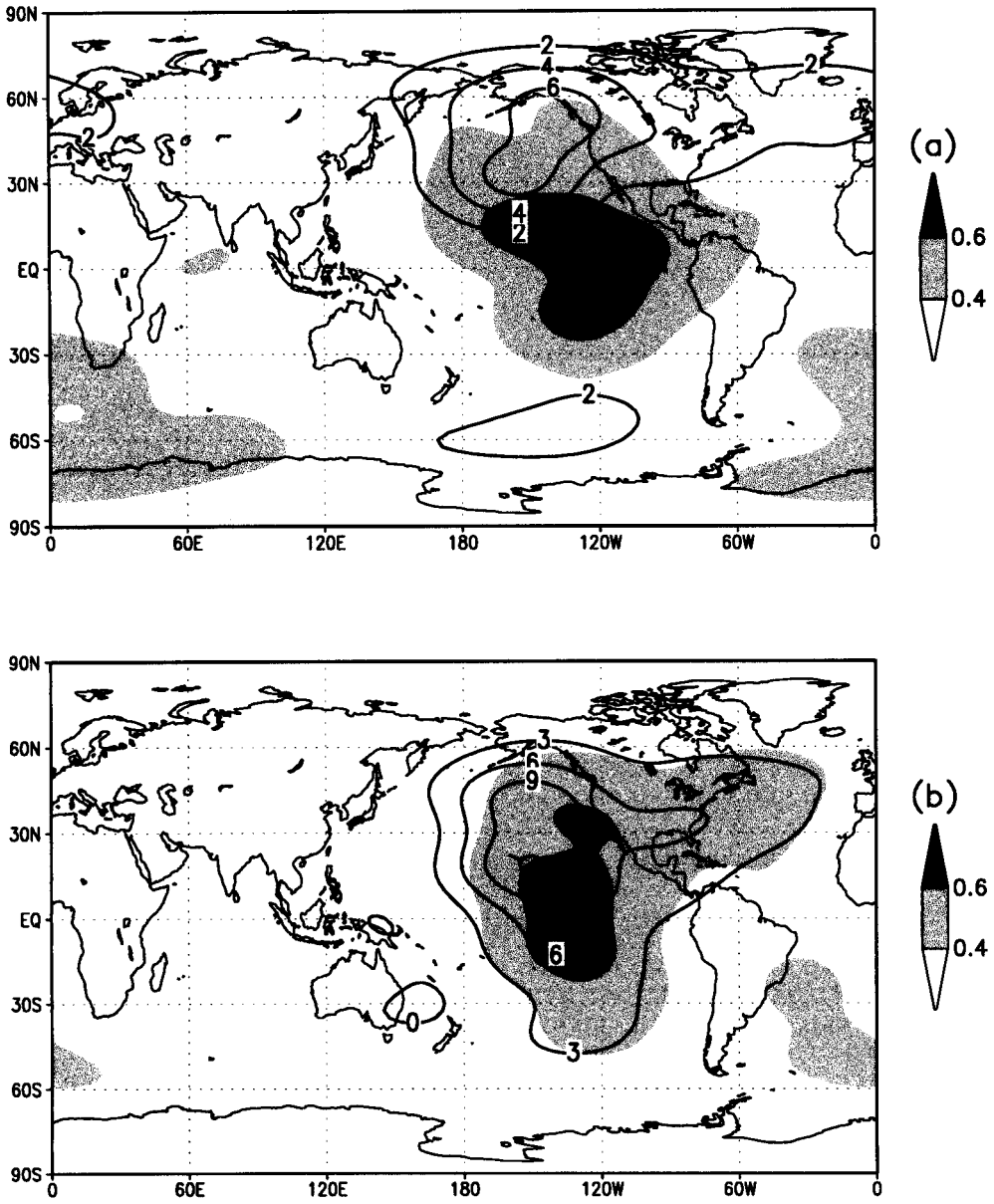

Figure 5. (a) Spatial correlation and covariance between the January to March mean observed and ensemble-mean COLA GCM 500 hPa height for 16 winters (1982–1997). The correlation (covariance) map gives a sliding spatial correlation (covariance) between two fields over a region of radius 2000 km centred at the given point. Spatial correlation (covariance) is given by the shading (contour interval 10^2 m^2). (b) As (a) but for 200 hPa zonal wind (contour interval 3 m^2s^{-2}). See text for further details.

100-day period starting on 20 December. The clear enhancement of the variance seen in the ECMWF analyses in the north-eastern Pacific during cold events (as in Palmer 1988, and Chen and Van den Dool 1997) is also noted in the GCM results, although it is not as dramatic. However, since the GCM variance has been averaged over all ensemble members for each year, it is expected that the individual maxima would be smoothed. Figure 8 shows the corresponding results for the transient meridional heat

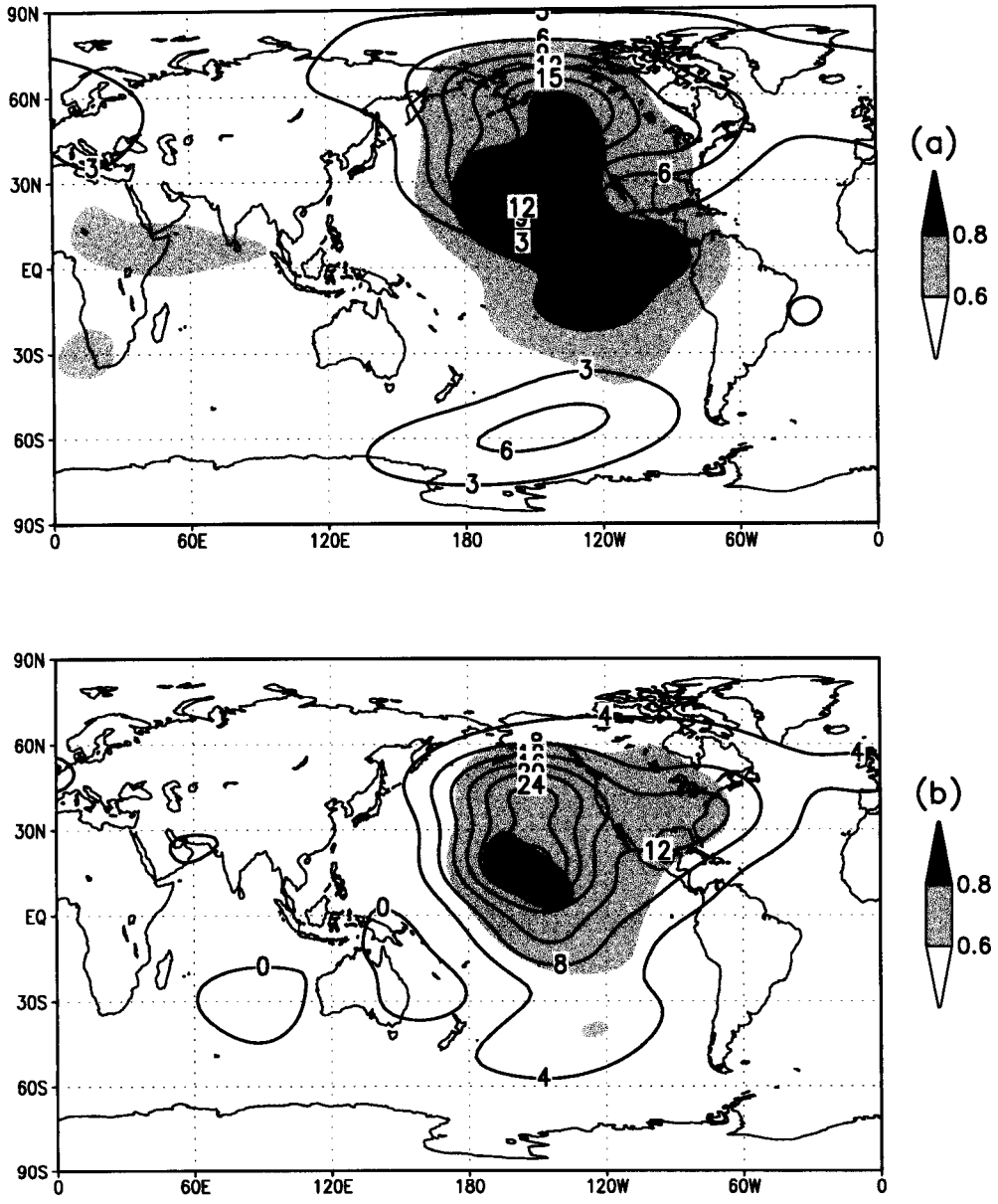

Figure 6. As Fig. 5 but for the strong ENSO winters of 1983, 1985, 1987, 1989 and 1992.

flux on time-scales of about 2–10 days (band-pass), except that here it is the anomaly from the 16-winter climatology that is shown. (Again, the filtering was accomplished by applying a digital filter after removal of the annual cycle.) During warm events there is an equatorward shift of the Pacific storm-tracks in both ECMWF analyses and in the GCM (as in Trenberth and Hurrell 1994; and Straus and Shukla 1997). The opposite shift is consistently seen in the Atlantic. A poleward shift of the Pacific storm-track is seen to occur during cold events in the analyses, and again this is captured by the GCM.

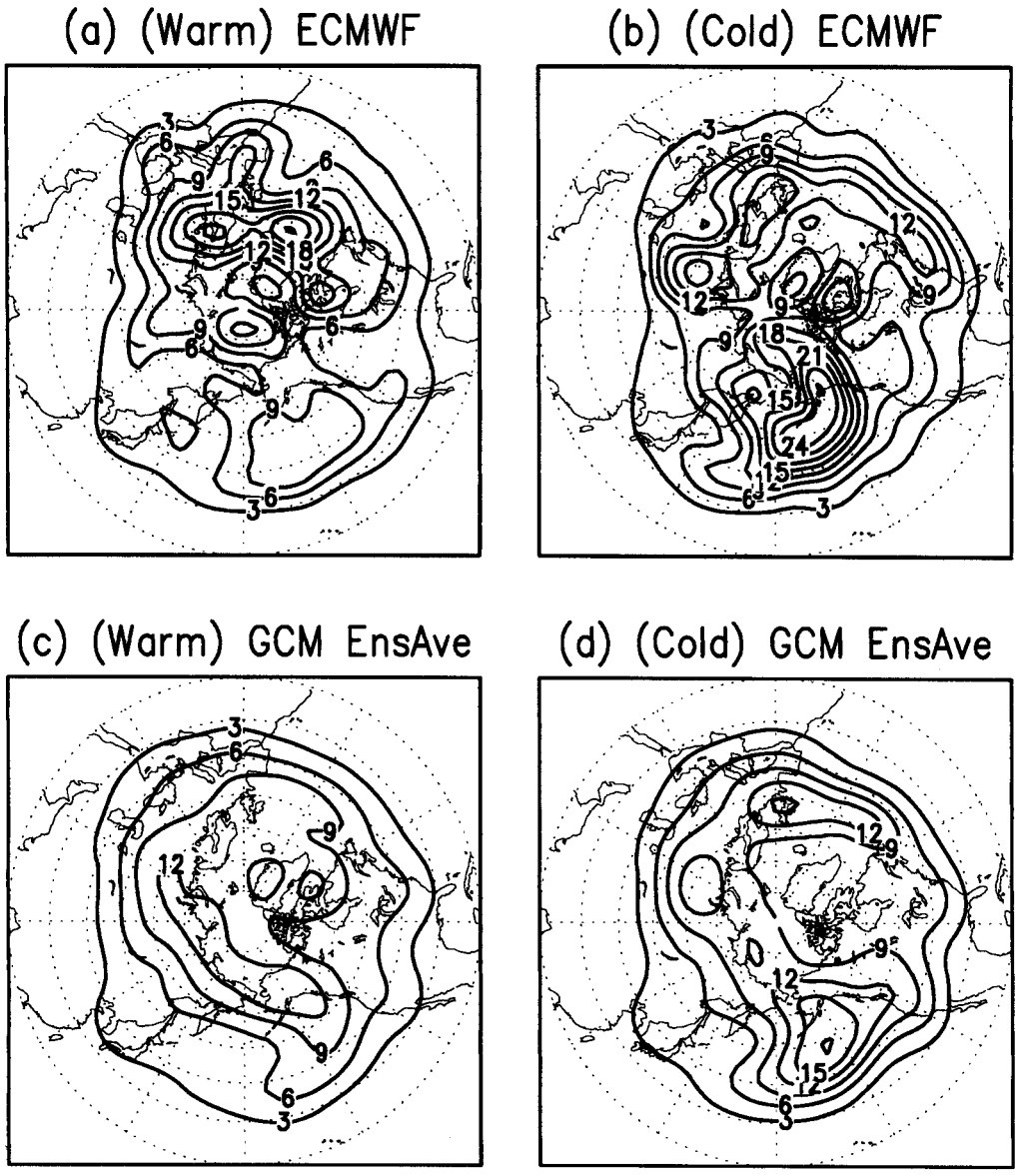

Figure 7. Low-pass variance of 500 hPa height for: (a) warm ENSO winters (1983, 1987 and 1992) from ECMWF analyses; (b) cold ENSO winters (1985 and 1989) from ECMWF analyses; (c) warm ENSO winters from COLA GCM ensemble average; (d) cold ENSO winters from COLA GCM ensemble average. Contour intervals are 3×10^3 m^2. See text for details.

(*d*) *EOF analysis*

A pattern which is characteristic of the 16-winter period is obtained by performing an EOF analysis (Bretherton *et al.* 1992) of winter mean 500 hPa height for observations (NCEP re-analyses) and the ensemble mean from the GCM integrations for the 16 winters of 1982–1997. The EOF analysis was carried out for the region shown in Fig. 9 (150°E–30°W; 20°–90°N) after the combined hemispheric North Atlantic/Arctic oscillation had been removed (see the companion paper, Straus and Shukla 2000). The

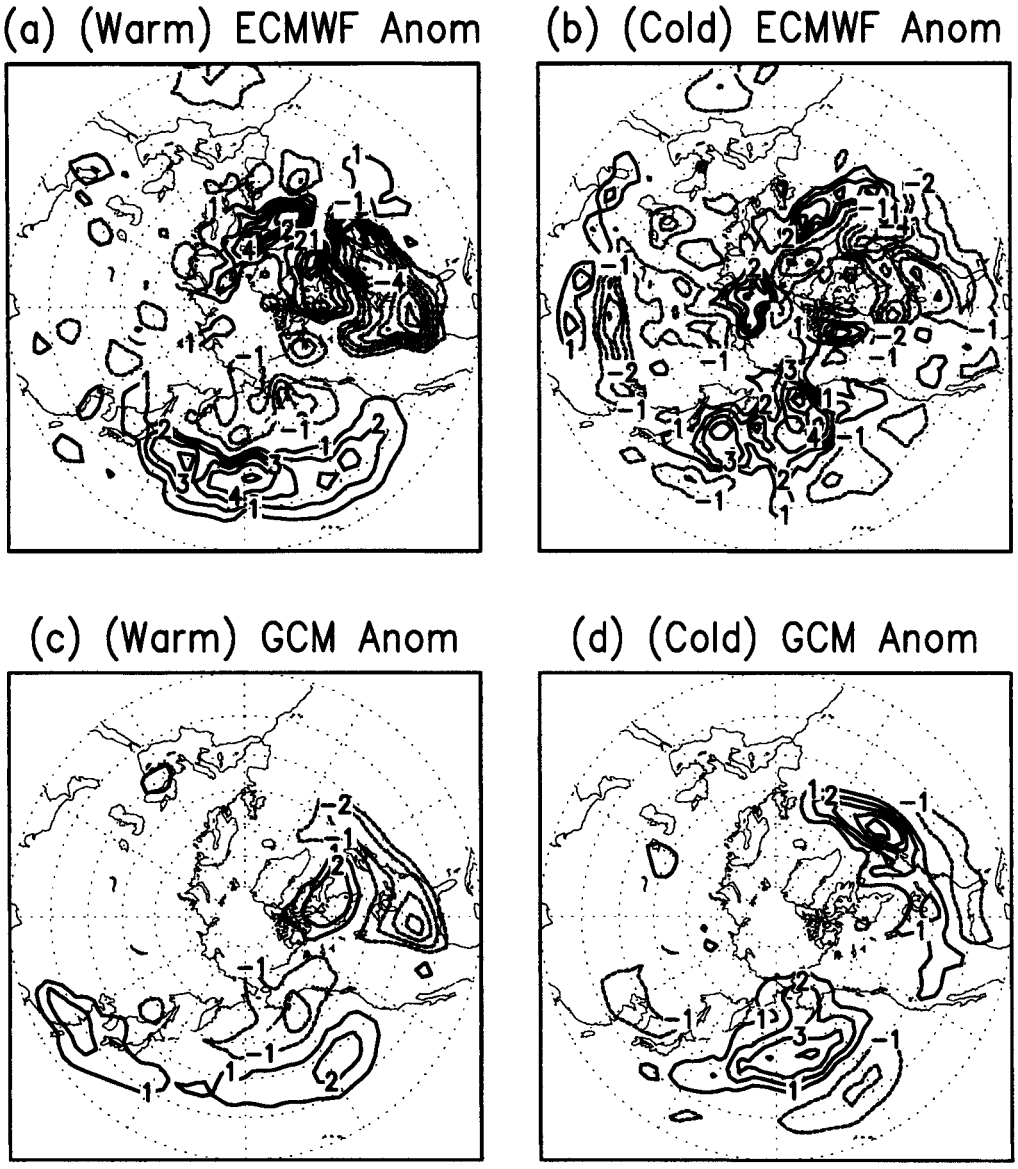

Figure 8. Band-pass variance anomalies of 850 hPa transient meridional heat flux for: (a) warm ENSO winters (1983, 1987 and 1992) from ECMWF analyses; (b) cold ENSO winters (1985 and 1989) from ECMWF analyses; (c) warm ENSO winters from COLA GCM ensemble average; (d) cold ENSO winters from COLA GCM ensemble average. Contour intervals are 1 K m s^{-1}. See text for details.

leading EOF mode explains 47% and 50% of the variance, for model and observations, respectively, which is far greater than the variance explained by any other mode. The patterns, shown in Fig. 9(a) and (b) are in generally close agreement, as are the associated principal components (time series) given in Fig. 9(c). Minor exceptions to this agreement include the small region off the coast of Newfoundland, which is involved in the first mode from the observations but not in the GCM's first mode. Also the winter of

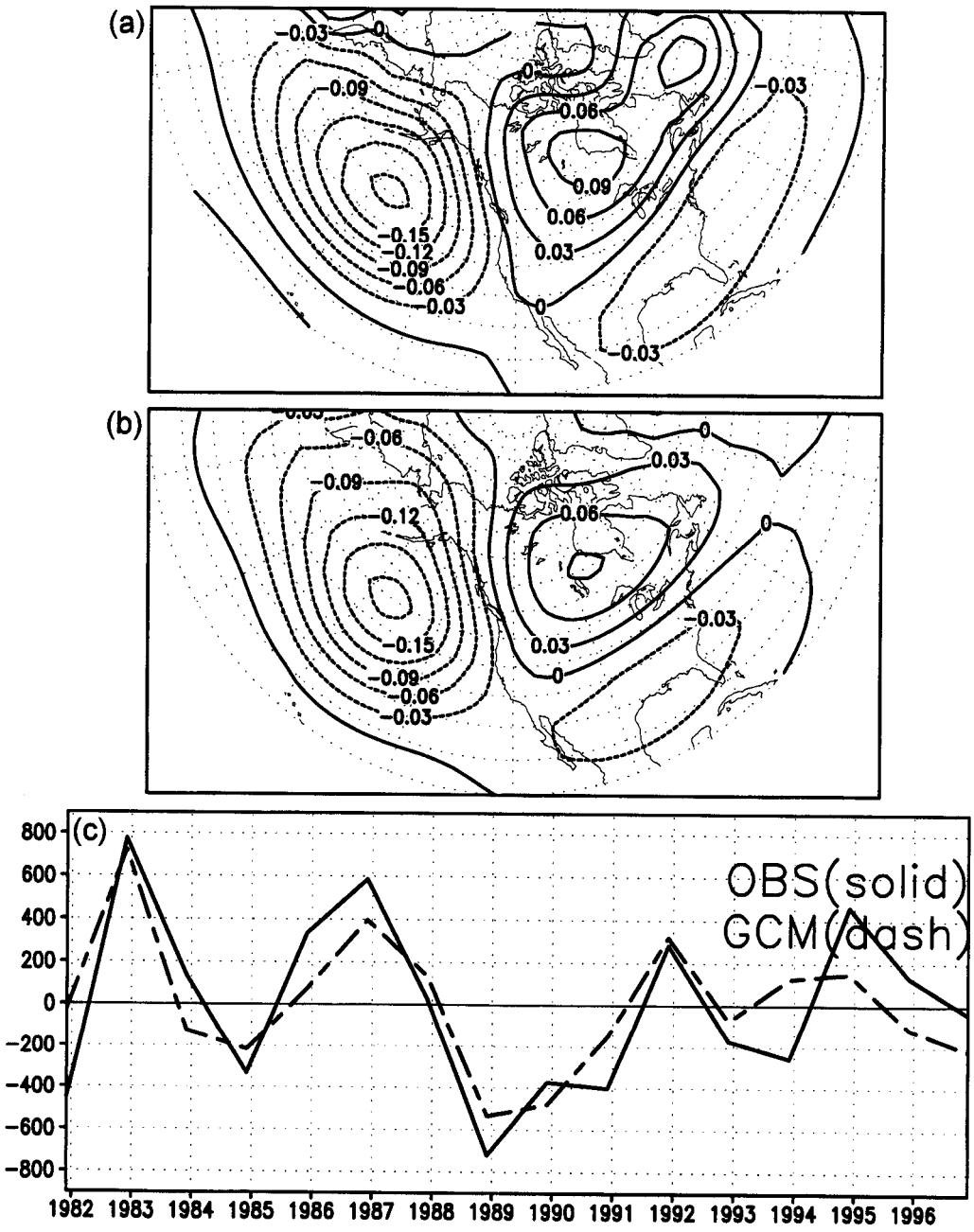

Figure 9. Leading empirical orthogonal functions (EOF-1) of 500 hPa height (in the region shown) for 16 winters for: (a) seasonal-mean observations (NCEP re-analyses); (b) COLA GCM ensemble seasonal mean; (c) the principal-component coefficients associated with the patterns in (a) and (b) as a function of time. For any winter, using the coefficient to scale the pattern gives the height in metres. In (a) and (b) the percentages of the variance explained by EOF-1 are 50 and 47 respectively. See text for further details.

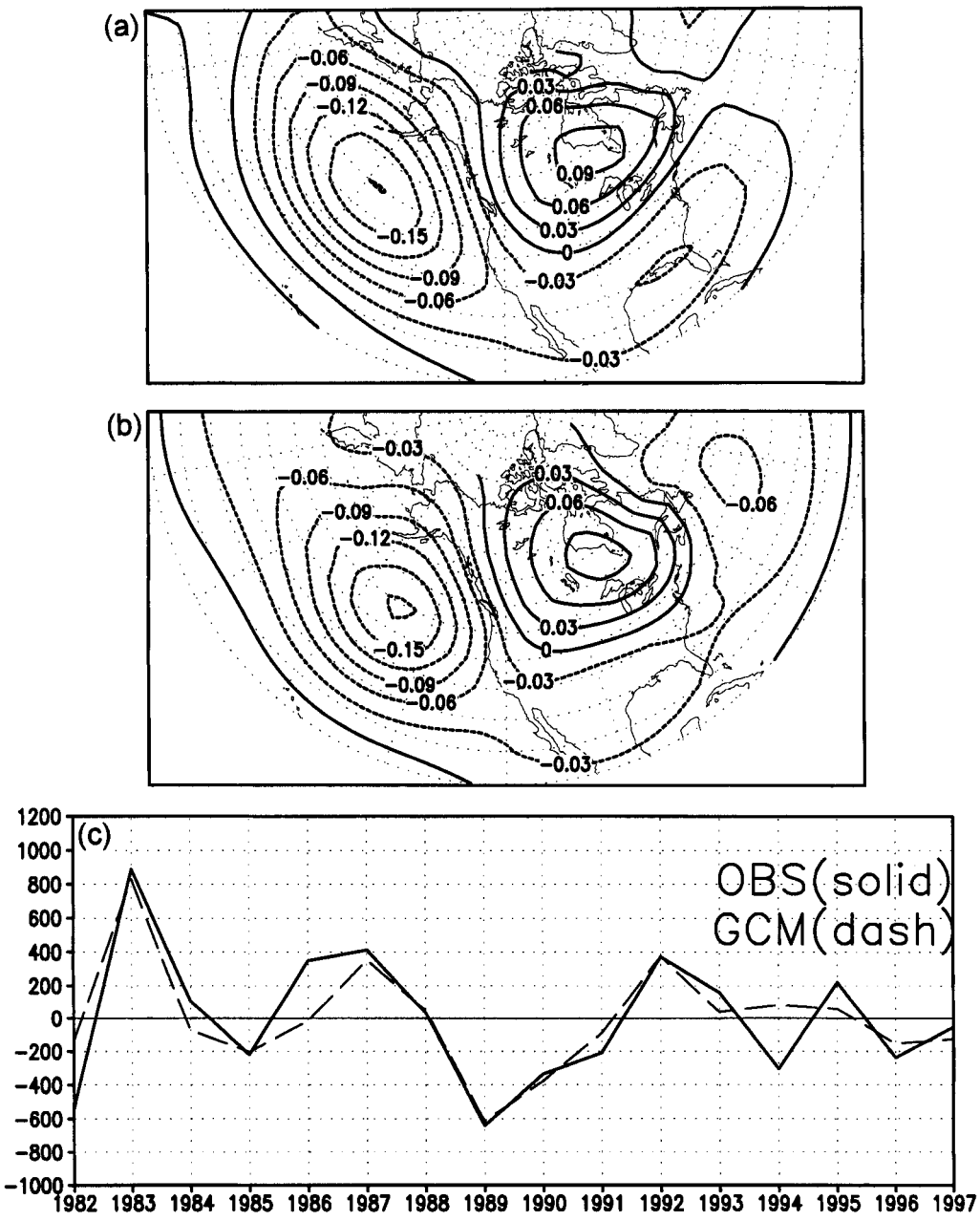

Figure 10. As Fig. 9 but for the leading mode of singular-value decomposition (SVD-1) of 500 hPa height in the region shown with tropical Pacific SST. SVD-1 explains 92% of the squared covariance both for observations and GCM forecasts.

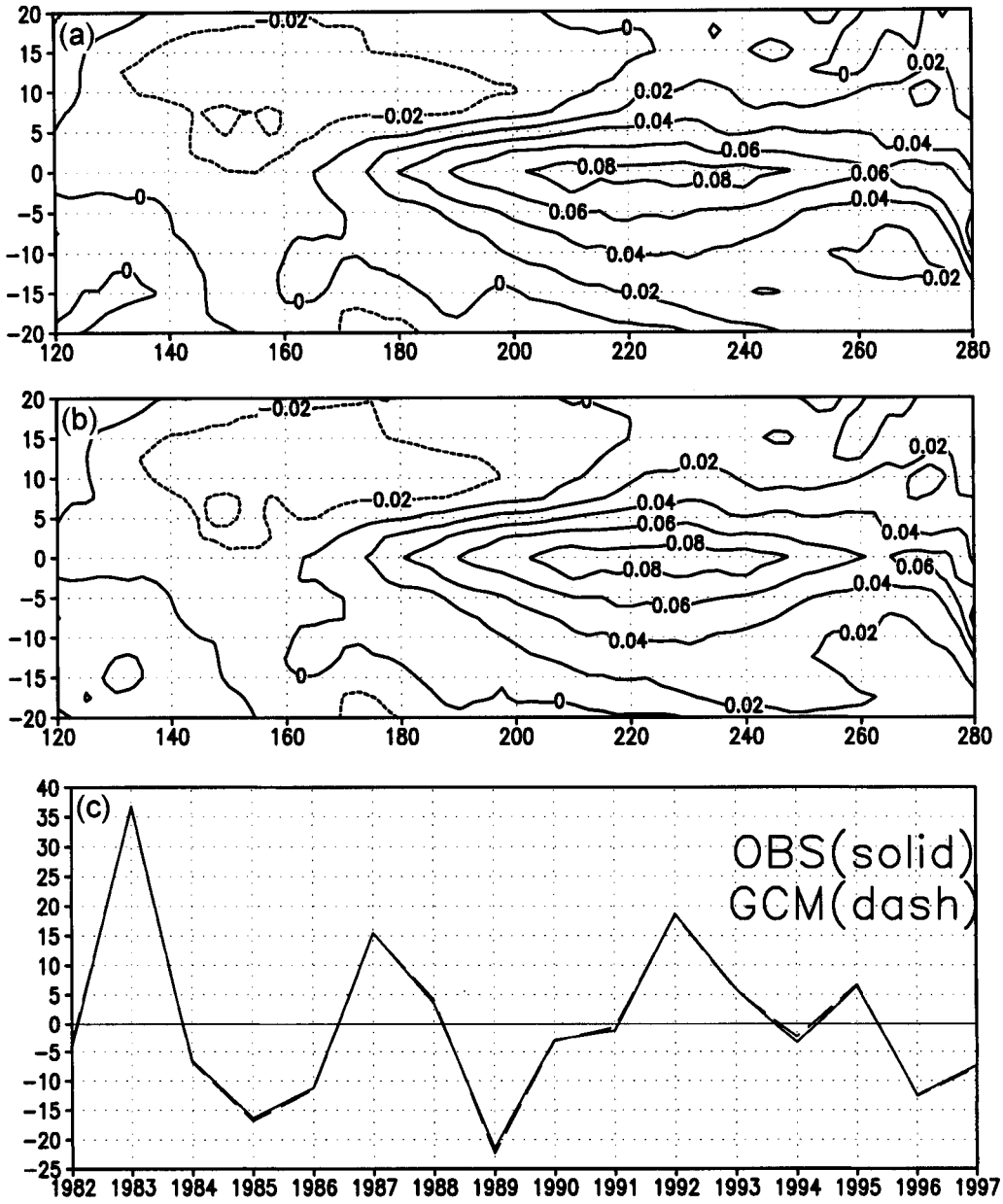

Figure 11. Leading mode of singular-value decomposition (SVD-1) of tropical Pacific SST in the region shown with 500 hPa height in the Pacific–North America region for 16 winters for: (a) seasonal-mean observations (NCEP re-analysis); and (b) COLA GCM ensemble seasonal mean. (c) The first mode coefficients associated with the patterns in (a) and (b) as a function of time (the two lines coincide). SVD-1 explains 92% of the squared covariance both for observations and GCM forecasts. For any winter, using the coefficient to scale the pattern gives the SST in degC. See text for further details.

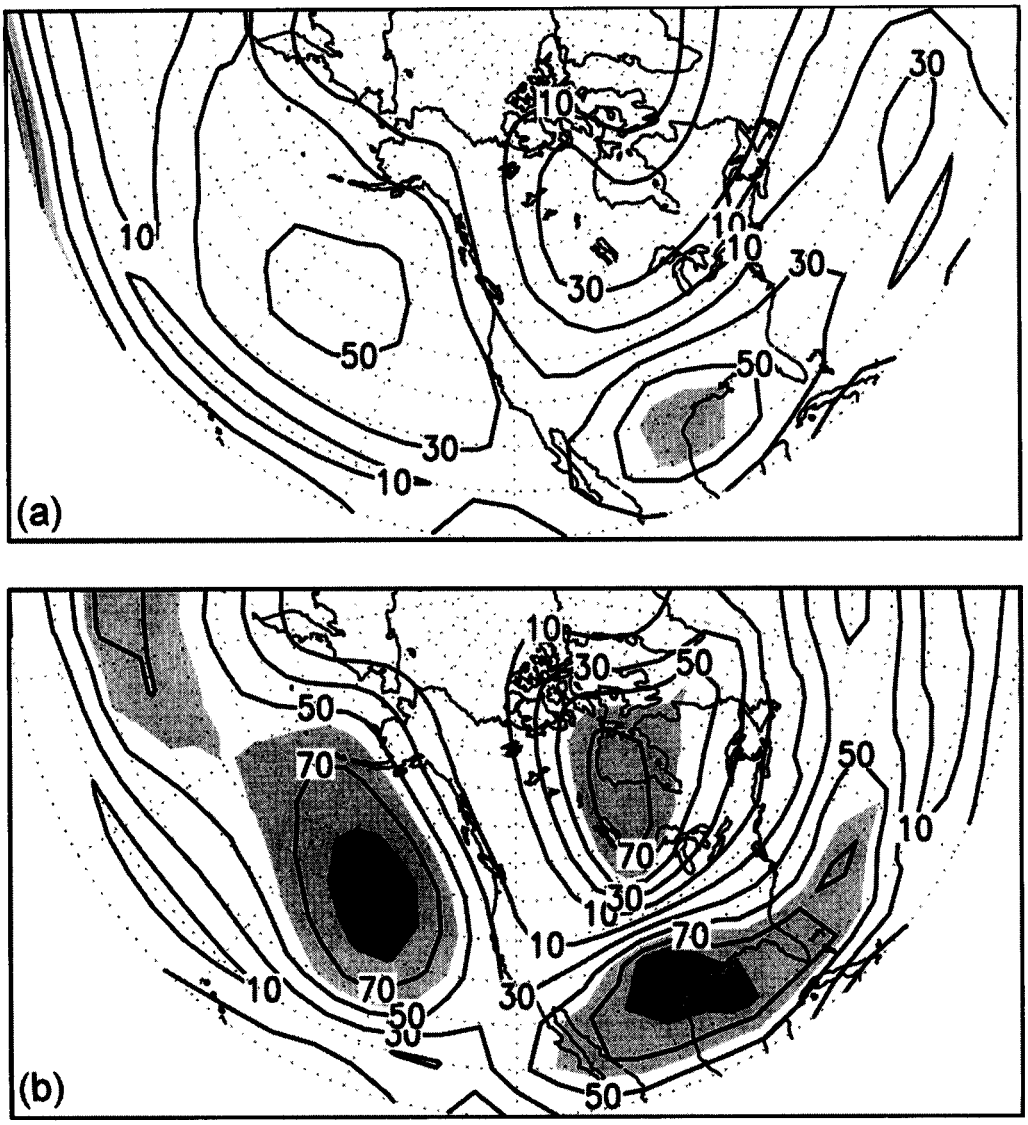

Figure 12. Regression of seasonal-mean (January to March) 500 hPa height for 16 winters (1982–1997) on time series of tropical Pacific SST. The time series was derived from singular-value decomposition of tropical Pacific SST with 500 hPa height for 30 winters (1968–1997). (a) Variance explained for observations (NCEP re-analyses); (b) variance explained by the GCM ensemble seasonal means. Light (dark) shading is for values above 60% (80%). Contours are at intervals of 20%. See text for further details.

1982 is characterized by a large value of the principal components in the observations, but not in the GCM.

(*e*) *SVD for tropical SST and mid-latitude height*

To establish that this characteristic pattern is related to the ENSO tropical Pacific SST forcing, an SVD analysis (Bretherton *et al.* 1992) was performed, linking the 500 hPa height field over the mid-latitude region (150°E–30°W; 20–90°N) with the tropical Pacific SST field in the region (120°E–80°W; 20°S–20°N). Again seasonal (seasonal-ensemble) means were used for observations (the GCM). Figure 10 shows

the 500 hPa height field patterns and time series associated with the first mode in the same form as in Fig. 9: with NCEP re-analysis seasonal mean observations in Fig. 10(a); the COLA GCM ensemble seasonal mean in Fig. 10(b); and in Fig. 10(c) the first-mode coefficients associated with (a) and (b) as functions of time. Figure 11 shows the corresponding SST field patterns and time series. The first mode explains 92% of the squared covariance between the height and SST fields both for observations and for the GCM. The fact that the two SVD SST patterns are nearly identical, suggests that the nature of the coupling between SST and height is nearly the same in observations and model simulations. The strong resemblance between the first-mode pattern for SST (Fig. 11(a) and (b)) and the first EOF of tropical Pacific SST (not shown), as well as the identity of the extrema in the time series (Fig. 11(c)) which match the designated SST strong warm and cold events, establish that the SST variation captured by the first SVD mode is that of ENSO. Further, the similarity between the height patterns of the first SVD mode (Fig. 10(a) and (b)) and that of the first EOF mode (Fig. 9(a) and (b)) for both observations and GCM, indicate that this dominant pattern is linked to the ENSO SST forcing. The time series in Figs. 9 and 10 are also quite similar. The time series of the SST component of the first SVD mode from the GCM and the NCEP re-analyses are nearly identical (Fig. 11(c)). In a companion paper by Straus and Shukla (2000), it is shown that the SST-forced pattern (shown in Figs. 9(a), 9(b), 10(a), 10(b)) is distinctly different from the pattern that occurs in the absence of ENSO-related SST anomalies.

(*f*) *Explained variance*

The amount of height variance which can be explained by the ENSO signal can be assessed from the heterogeneous correlation pattern for the first SVD mode for the observations and (separately) the GCM. An alternative approach is to calculate the explained variance on the basis of a time series of SST which is based on an observed record longer than 16 years. This procedure will test whether the height response is sensitive to the choice of period used to define ENSO. Figure 12 shows the percentage of seasonal-mean variance explained for the 16-year record of observations and the ensemble-mean GCM record, from a regression on a time series of SST which emerges from the SVD analysis of a 30-year record (winters from 1958 to 1997). (Clearly, only the appropriate part of the latter time series was used in the regression.) Locally up to 60% of the variance is explained for the observations (80% for the ensemble-mean GCM), and the patterns are not only very similar to each other but also to those in Figs. 9 and 10.

While these results give a strong indication that the ensemble mean of the GCM integrations captures the ENSO SST-forced signal in both pattern and magnitude, it must be recognized that the ensemble mean filters out variations not forced by SST which are due to the sensitivity to initial conditions. That is why the percentage variance explained is higher in the GCM than in observations (Fig. 12). These chaotic fluctuations are, of course, still present in the observations. It may be that individual realizations of the GCM have a larger level of chaotic fluctuations compared to the ENSO SST-forced signal than is realistic, and that only by ensemble averaging can we distinguish this signal so clearly. If the calculation is repeated many times using randomly chosen individual ensemble members (one per calendar year), the explained model variance is, on average, lower, and in fact slightly lower than the observations (not shown).

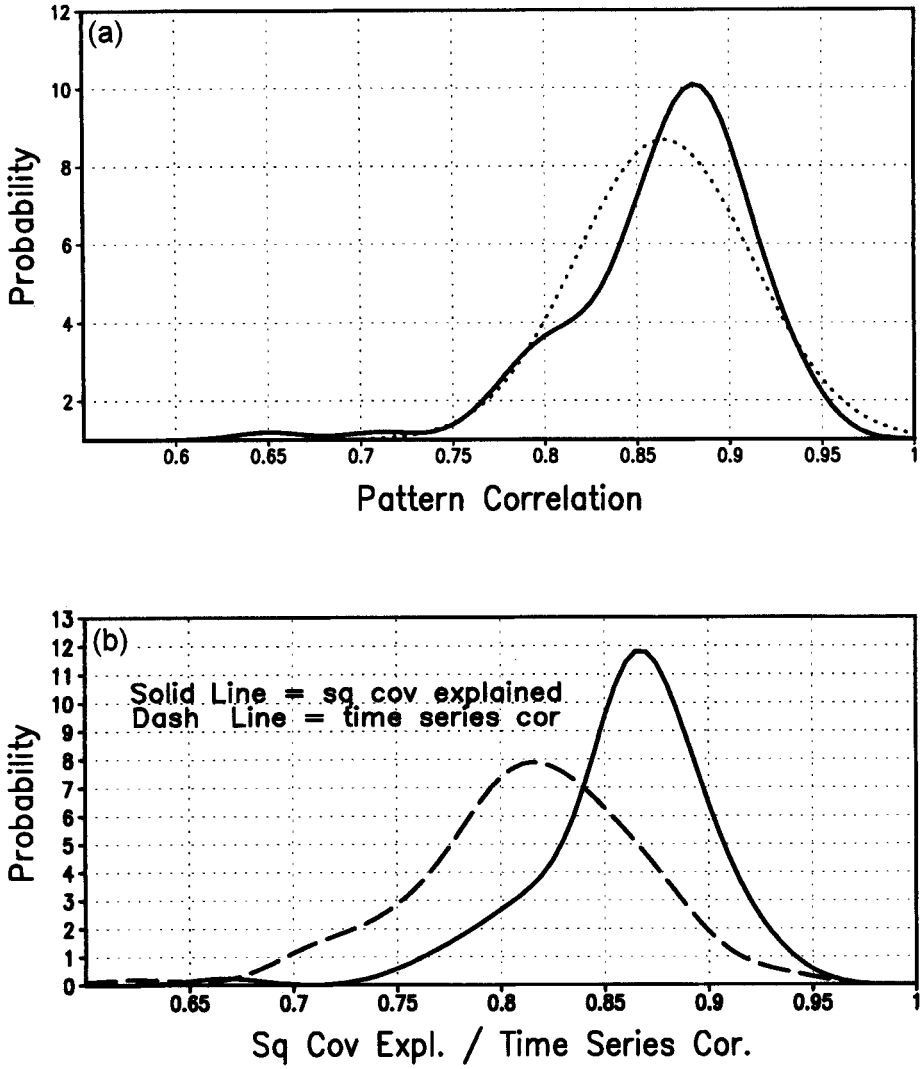

Figure 13. (a) Probability distribution functions (p.d.f.s) of pattern correlation (solid line) between leading singular-value decomposition (SVD) patterns for observations (NCEP re-analyses) and 100 GCM samples, each sample being obtained by randomly choosing one GCM integration per calendar winter. The dotted line gives the equivalent normal distribution with the same variance. The SVD analysis used seasonal-mean (January–March) tropical Pacific SST and 500 hPa height in the Pacific/America region for 16 winters (1982–1997). (b) P.d.f.s for the GCM squared covariance explained (solid line) and correlation between time series of height and SST for the first SVD mode (dashed line) for 100 GCM samples. See text for further details.

5. Reliability and reproducibility

(*a*) *Comparison between model ensembles and a single observed realization*

In order to assess the strength of the ENSO signal in single records of 16 seasonal GCM integrations, comparable to the single record available from observations, we have prepared 'samples' of GCM integrations. Each sample consists of one integration per calendar year, chosen randomly out of the ensemble for each year. (Note that for an ensemble size of N, over 16 years there are N^{16} possible samples.) We can then repeat the 16-year SVD analysis of 500 hPa height and tropical Pacific SST on each such

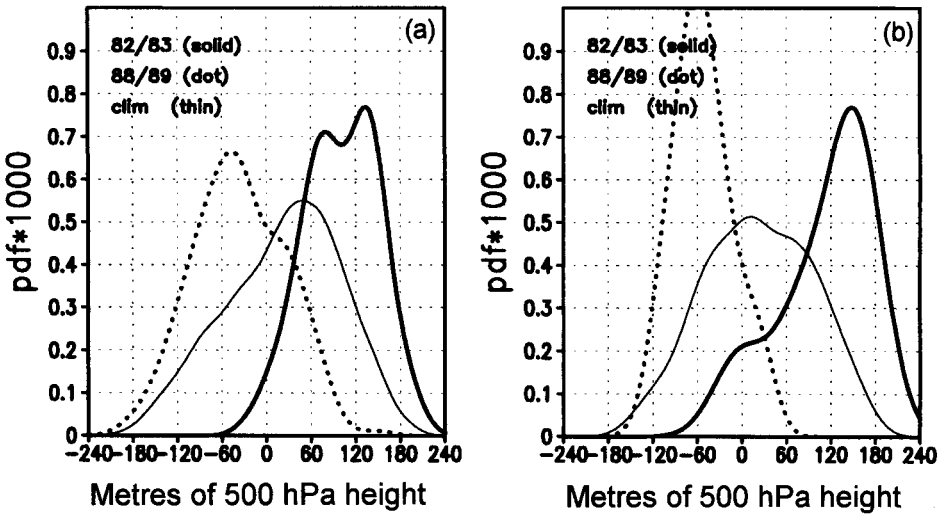

Figure 14. Probability distribution functions (p.d.f.s) for projection of pentad means onto the leading empirical orthogonal function (EOF-1) of: (a) ensemble seasonal means for the COLA GCM; and (b) seasonal means for observations (NCEP re-analyses) for 16 winters. The thin lines are the p.d.f.s of all pentad means for 16 winters, the heavy (dotted) lines are the p.d.f.s for the winter of 1983 (1989). See text for further details.

sample. Figure 13(a) shows the distribution of pattern correlations for the first SVD mode height field between each of a set of 100 samples and the observations. The distribution, presented as a probability distribution function (p.d.f.) using a Gaussian kernel estimation method (Silverman 1986), is strongly peaked at about 0.88, and indicates little probability of the GCM/observed pattern correlation falling below 0.80. Similarly, Fig. 13(b) shows the p.d.f.s for the squared covariance (height/SST) explained by the first SVD mode obtained from the 100 GCM samples (solid line), and for the correlation between the first-mode time series of SST and height. The peak in squared covariance explained is 87%, only slightly lower than the 92% explained by GCM ensemble means, with a very small probability of a sample indicating less than 80% explained squared covariance. Although the time series correlations are somewhat lower, very few samples give a correlation of less than 0.75.

While the occurrence of ENSO clearly affects intraseasonal fluctuations (as in Figs. 7 and 8), it is not clear whether the characteristic seasonal mean ENSO pattern defined above leaves a consistent imprint on intraseasonal fluctuations. Put another way, is the projection of individual 5-day means (pentads) on the SST-forced pattern of Fig. 9 predominantly of one sign during a strong ENSO winter, or is that pattern a statistical residue of large episodes of both signs? Figure 14 shows the p.d.f.s for these projections for all GCM pentads (Fig. 14(a)) and all observed pentads (Fig. 14(b)). In each case the p.d.f. for all pentads is given by the thin line, while those for the winter of 1983 (1989) are given by the thick solid (dotted) line. Clearly; the projections for the extreme warm (cold) ENSO event are predominantly positive (negative). The positive (negative) seasonal mean coefficient for 1983 (1989) in Fig. 9(c) is thus consistently seen in nearly all pentads. The shifts in the p.d.f. for warm and cold events is less in calculations by Renshaw *et al.* (1998), perhaps because their EOF was calculated using low-frequency filtered daily data.

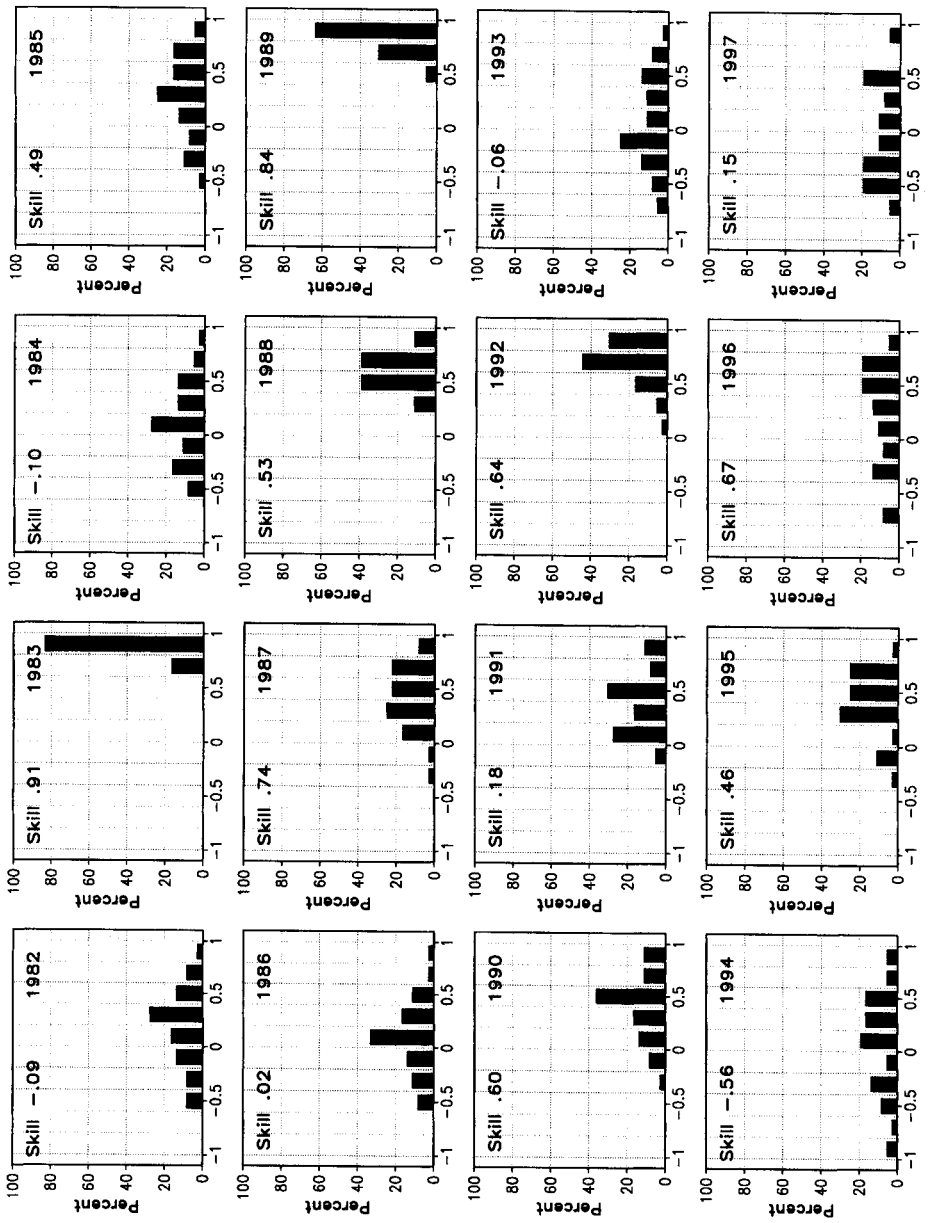

Figure 15. Anomaly correlation coefficients (ACCs) along the x-axes among the 36 pairs from nine ensemble members for 500 hPa height for each year from 1982 to 1997. Percentage of total pairs is along the y-axis. The ACC between the ensemble-mean model and observed 500 hPa height anomalies at the top left of each box is denoted as 'skill'. See text for further details.

(*b*) *Spread within an ensemble*

In order to make a simple estimate of spread among each of the nine members of the ensemble for each year, the ACC over the PacNA region for 500 hPa height anomalies was calculated among all 36 possible pairs of ensemble members. Figure 15 shows the frequency distribution of these 36 values of ACC for each year separately. Each panel also contains the ACC between the ensemble-mean height anomalies and the corresponding observations (at the top, designated 'skill'). It is seen that for eight cases whose intra-ensemble ACC is predominately positive (1983, 1985, 1987, 1988, 1989, 1990, 1992, 1995), the ensemble mean has a high value of ACC with respect to observations. This implies that the spread among the ensemble members is a good predictor of the skill of the ensemble mean. However, there are a few notable exceptions. For example, for 1991 most of the members of the ensemble have high ACC values among themselves, but the ensemble mean has a low ACC with respect to the observations. In 1996 the intra-ensemble ACC is low, but the ACC with respect to observations is high. Apart from the five ENSO years (1983, 1985, 1987, 1989 and 1992), the largest NINO3 SST anomalies occurred for the years 1988 and 1995, which are marked by consistency among the ensemble members. We do not know the reason for either the relatively high accuracy of 1990 compared to the observations, or for the high consistency of the ensemble members. A comparison of individual ensemble members with observations reveals that, even when the ensemble mean has a low ACC, some members of the ensemble have high ACC. What needs to be further explored is the question of whether increasing the ensemble size (either by making a larger number of integrations with the same model or by combining integrations from several models) would lead to a better estimate of the probability distribution of predicted states.

(*c*) *Brier score and reliability diagram*

In the context of ensemble predictions, it is of interest to consider skill scores which measure the occurrence (or non-occurrence) of an 'event', e.g. seasonal-mean 500 hPa geopotential height, Z, being one standard deviation greater (or less) than normal. The probability forecast for this happening at each grid point in the GCM can simply be given by the fraction of ensemble integrations, while for observations such an event either happens or does not happen (probability of 1 or 0). The observed frequency of events over NH is plotted against the forecast probability in Fig. 16(a) ($\Delta Z_{<(-1)}$; height below minus one standard deviation) and Fig. 16(b) ($\Delta Z_{>(+1)}$; height above one standard deviation). For these calculations 17 years' (1982–1998) data were used. Similar diagrams for events over North America are plotted in Fig. 16(c) and (d). Such 'reliability diagrams' (see Stanski *et al.* 1989 for details) are used to give a graphical depiction of probability forecast performance. All but one of the corresponding Brier skill scores* (Stanski *et al.* 1989; Palmer *et al.* 2000) are positive and statistically significant at the 99% level (based on a Monte Carlo approach), suggesting that the skill of ensemble forecasts is higher than that of a climatological probability forecast. When all grid points in NH are counted, the below-normal events are significantly over-forecast (Fig. 16(a)), leading to a lower Brier skill score than that of the above-normal events (Fig. 16(b)). Over North America, the reliability curves lie closer to the 'perfect reliability' 45° line, with correspondingly higher Brier skill scores than their

* The Brier skill score is defined as $B = (b_{\mathrm{cli}} - b)/b_{\mathrm{cli}}$, where b is the Brier score (Brier 1950) and b_{cli} is the Brier score for a climatological forecast. We define the climatological probability of an event at a grid point for each year as the climatological frequency of occurrence of the event, calculated from the observations of all *other* years at the *same* grid point, and a climatological forecast of the event as the forecast in which the climatological probability is predicted.

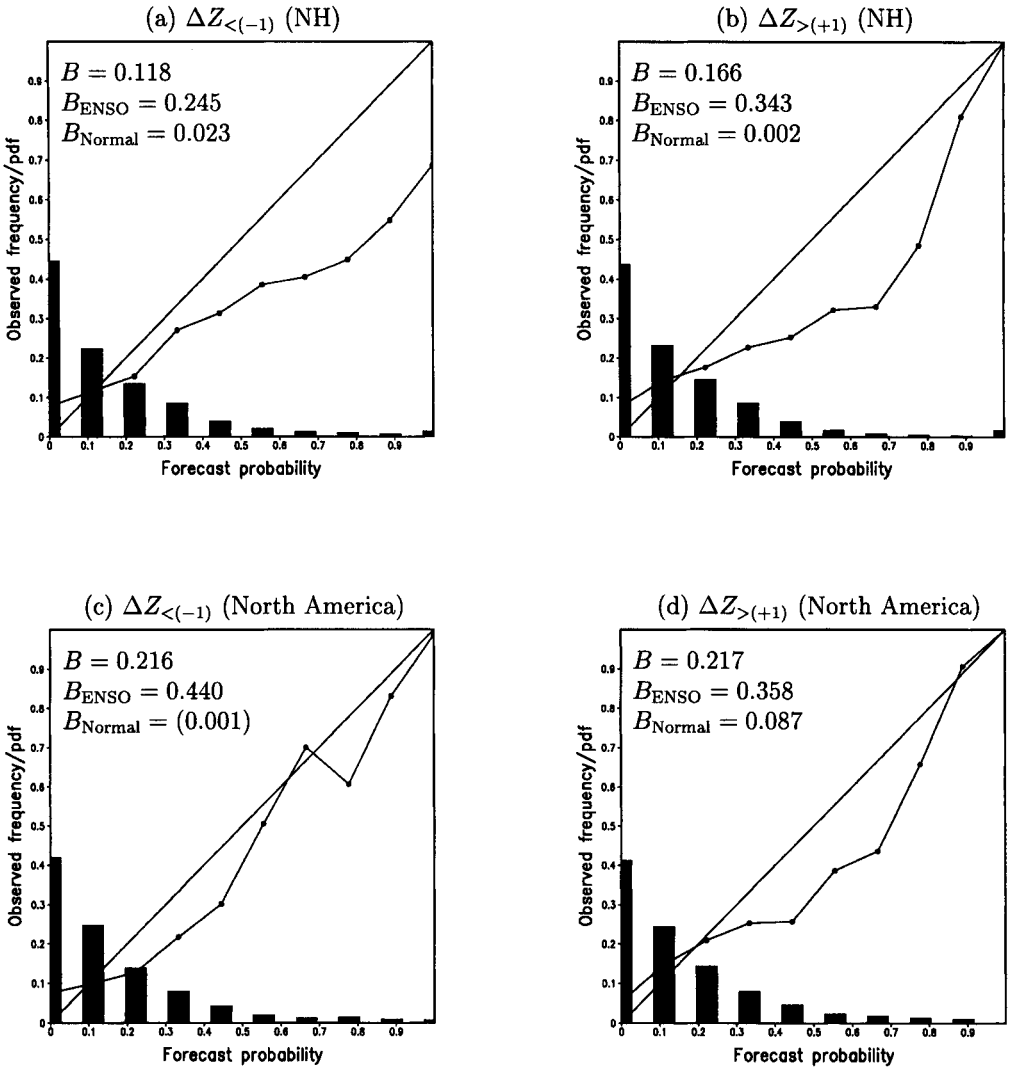

Figure 16. Reliability diagram and frequency distribution for the COLA DSP 17-year ensemble forecast of the seasonal mean (January–March) 500 hPa height, Z. The horizontal axis gives the forecast probability and the vertical axis gives the observed frequency. The probability density function is plotted as histograms. Two events are forecast: (i) Z is one standard deviation lower than normal, denoted by $\Delta Z_{<(-1)}$; and (ii) Z is one standard deviation higher than normal, denoted by $\Delta Z_{>(+1)}$. The upper panels show results derived from all gridpoints in the northern hemisphere (NH) and the lower panels from gridpoints in North America (20–70°N, 125–65°W). Also shown in the top left-hand corner of each diagram is the Brier skill score, B, B_{ENSO} for ENSO years and B_{Normal} for normal years; the score in parenthesis is not significant at the 99% level. See text for details.

NH counterparts. It thus suggests that the model performs better over North America than it does over the whole NH domain. By comparing the Brier skill scores for the ENSO years (i.e. 1983, 1987, 1992, 1998 for warm events and 1985, 1989 for cold events) with the normal years, it is seen that the model performs better during ENSO. The high skill scores over North America, therefore, are likely to be attributed to the ENSO SST forcing.

On the other hand, the model forecast over the European region (not shown) is very unreliable. The corresponding Brier skill scores for the 17-years of data are negative; this suggests that the skill of probability forecasts for the model is no better, if not worse, than the skill of the climatological forecast. The model performance over the European region is slightly improved during ENSO years.

6. SUMMARY AND DISCUSSION

The COLA model is successful in simulating seasonal-mean height anomalies over the PacNA region in the presence of large tropical SST anomalies. This enhanced predictability is not a unique property of this model, but is a property of the atmosphere, because as shown by Shukla *et al.* (2000) and others several models have produced similar simulations. However, in the absence of significant SST anomalies there is a large spread in the values of ACC among the members of an ensemble, and there is little, if any, predictability of extratropical seasonal-mean anomalies. The SVD analysis linking the 500 hPa height field over the PacNA region and the tropical SST, clearly suggests the predominant role of tropical forcing in producing mid-latitude circulation anomalies.

In the presence of large SST anomalies, the spread in the values of ACC among the members of each ensemble is small suggesting, at least in this limited number of cases, that if SST were predicted accurately the winter seasonal mean over the PacNA region could be predicted with a high degree of confidence. The prospects for prediction should not be considered entirely hopeless even for the years when SST anomalies are not large, because models still have quite large systematic errors which need to be reduced, and the probability distributions of predicted states have not been examined in large ensembles (50–100). The possible influence of land surface boundary conditions in improving seasonal predictability also needs to be investigated.

While the diagnostics we have presented emphasize the linear part of the height response to ENSO-related SST variations, it is clear from the ACCs in Fig. 15 that the GCM accurately responds to individual warm and cold events. Examination of differences between individual events shows some promising results. For example, the ACC between the observed and simulated difference between 1983 and 1992 warm anomalies was 0.85, while the ACC for the cold event difference 1989 minus 1985 was 0.55. The enhancement of low-pass kinetic energy during cold events only, is also testimony to the distinction between warm and cold events in both the simulations and in nature. Further work on the asymmetries between warm and cold events, and on differences between individual events, is a future goal. It will require a larger set of integrations, encompassing more years and also a larger ensemble size.

A preliminary analysis of summer season hindcasts (not shown) with this model has revealed that the mean model errors in rainfall are far larger than the observed anomalies of rainfall especially over the land areas. In fact, even with prescribed global observed SST, the model could not simulate the rainfall difference over the USA even for two years of opposite extreme rainfall anomalies: the drought in 1988 and the floods in 1993. There is a clear need to improve models to simulate mean summer rainfall over land areas, and only then it can be determined whether summer rainfall anomalies have any predictability.

ACKNOWLEDGEMENTS

This research was supported by NSF grant ATM-93-21354, NOAA grant NA-76-GP0258 and NASA grant NAGW-5213. We gratefully acknowledge the computing resources provided by the Climate Simulation Laboratory.

REFERENCES

Bengtsson, L., Schlese, U., Roeckner, E., Latif, M., Barnett, T. P. and Graham, N. 1993 A two-tiered approach to long range climate forecasting. *Science*, **261,** 1026–1029

Bengtsson, L., Arpe, K., Roeckner, E. and Schulzweida, U. 1996 Climate predictability experiments with a general circulation model. *Clim. Dyn.*, **12,** 261–278

Branković, Č., Palmer, T. N. and Ferranti, L. 1994 Predictability of seasonal atmospheric variations. *J. Climate*, **7,** 217–237

Bretherton, C. S., Smith, C. and Wallace, J. M. 1992 An intercomparison of methods for finding coupled patterns in climate data. *J. Climate*, **5,** 541–560

Brier, G. W. 1950 Verification of forecasts expressed in terms of probabilities. *Mon. Weather Rev.*, **78,** 1–3

Charney, J. G. and Shukla, J. 1981 'Predictability of monsoons'. Pp. 99–109 in Monsoon dynamics: proceedings of the symposium on monsoon dynamics, New Delhi, December 1977. Eds. Sir James Lighthill and R. P. Pearce. Cambridge University Press, Cambridge, UK

Chen, W. Y. and Van den Dool, H. M. 1997 Asymmetric impact of tropical SST anomalies on atmospheric internal variability over the North Pacific. *J. Atmos. Sci.*, **54,** 725–740

Chervin, R. M. 1986 Interannual variability and seasonal climate variability. *J. Atmos. Sci.*, **43,** 233–251

DeWitt, D. G. 1996 'The effect of cumulus convection on the climate of the COLA general circulation model'. COLA Tech. Report #27. COLA, 4041 Powder Mill Road, Suite 302, Calverton, MD, USA

Fennessy, M. J. and Shukla, J. 1999 Impact of initial soil wetness on seasonal atmospheric prediction. *J. Climate*, **12,** 3167–3180

Fennessy, M. J., Kinter III, J. L., Marx, L., Schneider, E., Sellers, P. J. and Shukla, J. 1994 GCM simulations of the life cycles of the 1988 drought and heat wave. COLA Tech. Report 6. COLA, 4041 Powder Mill Road, Suite 302, Calverton, MD, 20705

Gates, L., Boyle, J., Covey, C., Dease, C., Doutriaux, C., Drach, R., Fiorino, M., Gleckler,P., Hnilo, J., Marlais, S., Phillips, T., Potter, G., Santer, B., Sperber, K., Taylor, K. and Tribbia, J. 1999 An overview of the results of the atmospheric model intercomparison project (AMIP I). *Bull. Am. Meteorol. Soc.*, **80,** 29–55

Horel, J. D. and Wallace, J. M. 1981 Planetary-scale phenomena associated with the Southern Oscillation. *Mon. Weather Rev.*, **109,** 813–829

Kinter III, J. L., DeWitt, D. G., Dirmeyer, P. A., Fennessy, M. J., Kirtman, B. P., Marx, L., Schneider, E. K., Shukla, J. and Straus, D. M. 1997 'The COLA atmosphere–biosphere general circulation model. Volume 1: formulation'. COLA Tech. Report, 51. COLA, 4041 Powder Mill Road, Suite 302, Calverton, MD, USA

Kumar, A. and Hoerling, M. P. 1997 Interpretation and implications of the observed inter-El Niño variability. *J. Climate*, **10,** 83–91

Kumar, A., Hoerling, M. P., Ji, M., Leetma, A. and Sardeshmukh, P. D. 1996 Assessing a GCM's suitability for making seasonal predictions. *J. Climate*, **9,** 115–129

Lau, N. C. 1981 A diagnostic study of recurrent meteorological anomalies appearing in a 15-year simulation with a GFDL general circulation model. *Mon. Weather Rev.*, **109,** 2287–2311

1997 Interactions between global SST anomalies and the mid-latitude atmospheric circulation. *Bull. Am. Meteorol. Soc.*, **78,** 21–33

Miyakoda, K., Sirutis, J. and Ploshay, J. 1986 One month forecast experiments without boundary forcings. *Mon. Weather Rev.*, **114,** 2363–2401

Miyakoda, K., Gordon, T., Caverly, R., Stern, W., Sirutis, J. and Bourke, W. 1983 Simulation of a blocking event in January 1977. *Mon. Weather Rev.*, **111,** 846–869

Moorthi, S. and Suarez, M. J. 1992 Relaxed Arakawa–Schubert: A parameterization of moist convection for general circulation models. *Mon. Weather Rev.*, **120,** 978–1002

NMC 1988 'Research version of the medium range forecast model'. NMC Documentation Series #1. Development Division, National Meteorological Center, Washington, USA

Palmer, T. N. 1988 Medium and extended-range predictability and the stability of the Pacific–North America mode. *Q. J. R. Meteorol. Soc.*, **114,** 691–713

Palmer, T. N., Branković, Č. and Richardson, D. S. 2000 A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. *Q. J. R. Meteorol. Soc.*, **126,** 2013–2033

Reynolds, R. W. and Smith, T. M. 1994 Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7,** 929–948

Ropelewski, C. F., Janowiak, J. E. and Halpert, M. F. 1985 The analysis and display of real time surface climate data. *Mon. Weather Rev.*, **113,** 1101–1107

Sato, N., Sellers, P. J., Randall, D. A., Schneider, E. K., Shukla, J., Kinter III, J. L., Hou, Y.-T. and Albertazzi, E. 1989 Effects of implementing the Simple Biosphere Model in a general circulation model. *J. Atmos. Sci.*, **46,** 2757–2782

Shukla, J. 1981a 'Predictability of the tropical atmosphere'. NASA Tech. Memo. #83829, NASA Goddard Space Flight Center, Greenbelt, MD, USA

1981b Dynamical predictability of monthly means. *J. Atmos. Sci.*, **38,** 2547–2572

1998 Predictability in the midst of chaos: a scientific basis for climate forecasting. *Science*, **282,** 728–731

Shukla, J., Anderson, J., Baumhefner, D., Branković, Č., Chang, Y., Kalnay, E., Marx, L., Palmer, T. N., Paolino, D., Ploshay, J., Schubert, S., Straus, D., Suarez, N. and Tribbia, J. 2000 Dynamical seasonal prediction. *Bull. Am. Meteorol. Soc.*, **81,** in press

Silverman, B. W. 1986 *Density estimation for statistics and data analysis*. Chapman and Hall, London, UK

Stanski, H. R., Wilson, L. J. and Burrows, W. R. 1989 'Survey of common verification methods in meteorology'. World Weather Watch Tech. Report #8. WMO/TD, #358, World Meteorological Organisation, Geneva, Switzerland

Straus, D. M. 1983 On the role of the seasonal cycle. *J. Atmos. Sci.*, **40,** 303–313

Straus, D. M. and Shukla, J. 1997 Variations of midlatitude transient dynamics associated with ESNO. *J. Atmos. Sci.*, **54,** 777–790

2000 Distinguishing between the SST-forced variability and internal variability in mid latitudes: Analysis of observations and GCM simulations. *Q. J. R. Meteorol. Soc.*, **126,** 2323–2350

Straus, D. M. and Yang, Q. 1997 Vertical structure and dominant horizontal scales of baroclinic waves in the NASA DAO and NCEP reanalyses. *Mon. Weather Rev.*, **125,** 3266–3278

Trenberth, K. E. and Hurrell, J. W. 1994 Decadal atmosphere–ocean variations in the Pacific. *Clim. Dyn.*, **9,** 303–319

Xie, P. and Arkin, P. 1996 Analysis of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate*, **9,** 840–858