Origin of a trend in ECMWF wind analysis and its objective removal

B N GOSWAMI and H ANNAMALAI

Centre for Atmospheric Sciences, Indian Institute of Science, Bangalore 560012, India

Abstract. The suitability of the European Centre for Medium Range Weather Forecasting (ECMWF) operational wind analysis for the period 1980–1991 for studying interannual variability is examined. The changes in the model and the analysis procedure are shown to give rise to a systematic and significant trend in the large scale circulation features. A new method of removing the systematic errors at all levels is presented using multivariate EOF analysis. Objectively detrended analysis of the three-dimensional wind field agrees well with independent Florida State University (FSU) wind analysis at the surface. It is shown that the interannual variations in the detrended surface analysis agree well in amplitude as well as spatial patterns with those of the FSU analysis. Therefore, the detrended analyses at other levels as well are expected to be useful for studies of variability and predictability at interannual time scales. It is demonstrated that this trend in the wind field is due to the shift in the climatologies from the period 1980–1985 to the period 1986–1991.

Keywords. ECMWF objective analysis; interannual variability; combined EOF analysis; removal of unphysical trend.

1. Introduction

The need for a comprehensive, internally consistent, homogeneous, multivariate data set for studies of the mean climate system and its interannual variations has been recognized for sometime (Bengtsson and Shukla 1988). Model-based operational analyses from numerical weather prediction (NWP) centres, such as the National Meteorological Center (NMC) of the United States and the European Center for Medium range Weather Forecasts (ECMWF), have been used for this purpose. Although these data sets have been used in quite a large number of diagnostic studies of the general circulation, they have some serious deficiencies (Trenberth and Olson 1988), specially for studying variability and predictability at interannual time scales. Three major reasons are: (a) gradual improvement of the models' resolution and physical parameterizations, (b) frequent changes in the data assimilation system and (c) unavailability of delayed mode data. For example, the NMC model has evolved from a 12-layer R24 resolution version in 1980 to an 18-layer T106 resolution version in recent times. Similarly, the ECMWF model has grown from a 15-level grid point (1·875°) version in 1980 to a 19-level T213 spectral model in recent times. The analysis system has also gone through gradual changes over the last 15 years. The most significant change is the introduction of multivariate optimum interpolation and diabatic nonlinear normal mode initialization in the mid-eighties. Trenberth and Olson (1988) note that the introduction of a multivariate optimum interpolation scheme to the ECMWF analyses on May 1st, 1985, increased the divergent part of the tropical wind. Trenberth *et al* (1990) noted that the 1000 mb monthly mean wind

convergence near the equator in the ECMWF analysis during 1980–1986 does not agree with that expected from satellite-observed longwave radiation. Hurrell and Trenberth (1992) applied a weighting function for the ECMWF temperature fields for the period 1982–1989. They observed that over the tropics, the analyses have improved, especially after May 1985, when the 1000–200 mb temperatures show more coherence. However, Trenberth and Guillemot (1994) note that only during the most recent period, 1990–1993, the ECMWF analysed mean annual surface pressure is reasonably stable.

As a result of these changes, the archived analysis includes unphysical jumps at different points of time. Therefore, it has become necessary to reanalyse the past data using a recent model and a more recent analysis system (Bengtsson and Shukla 1988). Major reanalysis efforts have been already started in NMC (Kalnay and Jenne 1991), ECMWF and NASA (Schubert *et al* 1993). In spite of the changes noted in the ECMWF analysis, the present analysed circulation data sets have been used increasingly in recent times to study interannual variability (Wang 1992; Murakami and Matsumoto 1994) and predictability (Barnett *et al* 1994). These studies utilize anomalies defined as deviation from a climatology based on the entire time span of the data. The changes in the model-based analysis system give rise to some unphysical interannual variability. The studies that do not attempt to separate this analysis system induced interannual variability from the real interannual variability are expected to arrive at erroneous conclusions. In this study, we investigate whether the analysis system induced interannual variability is systematic. If it is so, could it be separated from real interannual variability? After removal of the analysis system induced variability, is the analysis still useful for interannual variability studies?

While examining the ECMWF analysis from 1980 to 1991 to study interannual variability of the Indian monsoon, we found that the circulation anomalies calculated traditionally have a clear trend, specially over the tropical region. In an attempt to bring out the patterns of the variability related to sea surface temperature variations, we carried out a combined empirical orthogonal function (EOF) analysis of sea surface temperature (SST), outgoing longwave radiation (OLR) and winds at 1000, 850 and 200 mb from ECMWF analysis. We found that some of the leading principal components have a significant trend. As the physical origin of the trend was unclear, we conducted an EOF analysis of the SST, OLR and winds separately. The SST and OLR data do not show any trend. However, EOF analysis of the ECMWF winds at the different levels carried out one at a time or a combined EOF analysis of the winds at all the levels taken together shows a significant trend in the dominant modes. Thus, it was clear that the trend in the combined EOF analysis arose from the trend in the ECMWF wind data. The wind anomalies prior to the mid-eighties tend to be of one sign, while those in the later period tend to be of the opposite sign. Here, we show that this trend arises mainly due to a shift of the climatologies from a state before the mid-eighties to another state after the mid-eighties. This shift in the climatology is due to major changes in the model and analysis system in the mid-eighties. Previous attempts to correct the analysis, compared it with observation at certain points and tried to develop weighting functions to correct the difference between observations and analysis. The limitations of this method are that the weighting functions developed over a particular region may not be applicable over other regions. Moreover, upper air data over the oceanic regions are unavailable to construct these correcting functions. Here, we present an alternative method to remove the systematic errors globally at all levels by comparing with an independent analysis at the surface and using a multivariate EOF technique.

2. Data

The main data set here is the monthly mean operational analysis of the ECMWF from January 1980 to December 1991. Zonal and meridional winds at 1000, 850 and 200 mb are used. The horizontal resolution is $2.5° \times 2.5°$. We have confined our study to the global tropics between 30°S and 30°N. To compare the anomalies with an independent analysis we use the Florida State University (FSU) analysis of the surface wind (Goldenberg and O'Brien 1981). This data set is based on available *in situ* observations and a subjective analysis. This data set covers only the Pacific domain. Updated FSU winds are available with us up to 1991.

3. Results and discussion

3.1 *Anomalies relative to climatology based on the entire period*

First, we constructed the climatology of each field at every grid point for each calendar month averaged over all the 12 years. Anomalies are then calculated by subtracting the climatology corresponding to the given calendar month from the analysis for the particular month. In figure 1(a) we show the time series of anomalies of zonal wind shear $(U_{850}-U_{200})$ averaged over the NINO-4 (5°S–5°N, 160°E–150°W) region of the Pacific. This index is related to the strength of the Walker circulation. Figure 1(b) shows the time series of the broad scale monsoon index (Webster and Yang 1992). This is the zonal wind shear $(U_{850}-U_{200})$ averaged between 5°N–20°N and 40°E–110°E. Similarly, in figure 1(c), we show the vertical shear of the meridional wind anomalies averaged over the equatorial Indian Ocean (10°S–10°N, 40°E–110°E). This represents the strength of the regional Hadley circulation. A clear trend is evident in figures 1(a) and (c). We note that the trend is more pronounced on the time series representing equatorial fields (e.g., figures 1a and c), as compared to fields away from the equator (e.g., figure 1b). This is consistent with the findings of Trenberth *et al* (1990).

It is recognized that the ECMWF analyses have certain inconsistencies in all the fields and at all the levels (Hurrell and Trenberth 1992). However, to obtain a homogeneous data set, it is a formidable task to identify the bias and apply a correction at all grid points and at all levels. For studying variability and predictability over the tropics at interannual time scales, one is interested in examining the large scale patterns of spatial and temporal variability. EOF analysis helps us to identify large scale variability and describes most of the variability in a few leading EOFs. Therefore, systematic errors in these patterns can be corrected more easily by correcting the principal components and then reconstructing the data. To examine how this trend is manifest in the large scale part of the circulation, we carried out a multivariate or combined EOF analysis (Wang 1992; Nigam and Shen 1993). The choice of combined EOF analysis has been adopted for the following reasons. Firstly, a combined EOF analysis provides a more efficient way of compacting a large volume of multifield data. As geophysical fields are not only spatially correlated but also correlated with one another, the combined EOF analysis brings out the patterns that are highly interrelated. This helps us to bring out the three-dimensional patterns of anomalies associated with certain types of variability. We thus get an insight into the dynamics of such variability. More importantly, in the context of our analysis, it is expected to bring out the systematic
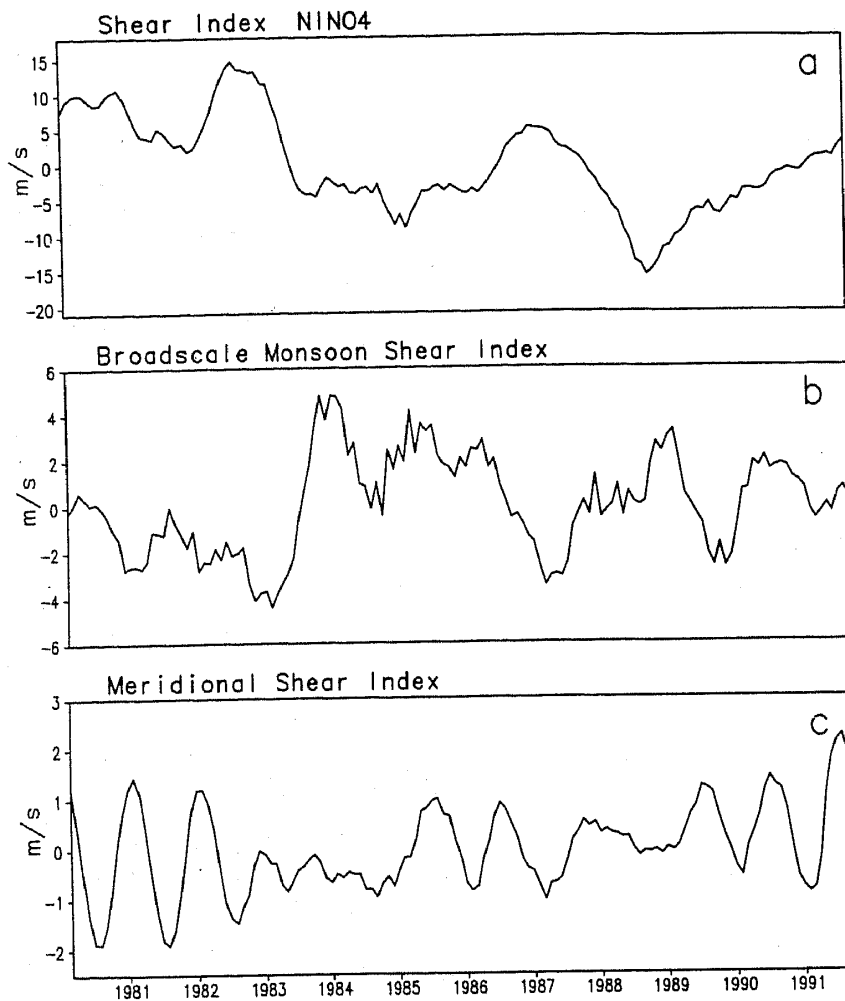
Figure 1. Anomaly time series, based on the 1980–1991 data. (a) The shear index, $(U_{850}-U_{200})$ averaged over the NINO-4 (5°S–5°N, 160°E–150°W) region of the Pacific; (b) Broadscale monsoon index $(U_{850}-U_{200})$ averaged between 40°E–110°E and 5°N–20°N; (c) The meridional shear index, $(V_{850}-V_{200})$ averaged over the equatorial Indian Ocean (10°S–10°N, 40°E–110°E).

errors that are common to all variables in a limited number of EOFs. As we shall show later, this will not only allow us to apply corrective procedures to the data with least amount of effort, it will also provide information about the origin of the errors. As this brings out the patterns that are interrelated amongst different variables, for verification purposes, all that we need is one variable from an independent analysis. This is the most useful feature of this analysis, as we shall show later.

As we are interested in interannual time scales, wind anomalies at all levels are passed through a five-month running mean filter. Unless otherwise stated, all the analyses described below are based on these filtered winds. A combined EOF analysis of winds at surface, 850 and 200 mb was carried out. Figure 2 shows the time series of the first six principal components of the combined EOFs. Figures 3 and 4 show the vector representation of the first four EOFs at the surface and 200 mb respectively. It is clear from figure 2 that the PC1, PC2 and PC5 have significant trends. The anomalies tend to change sign in the mid-eighties. Following Bethea *et al* (1975) the significance of the
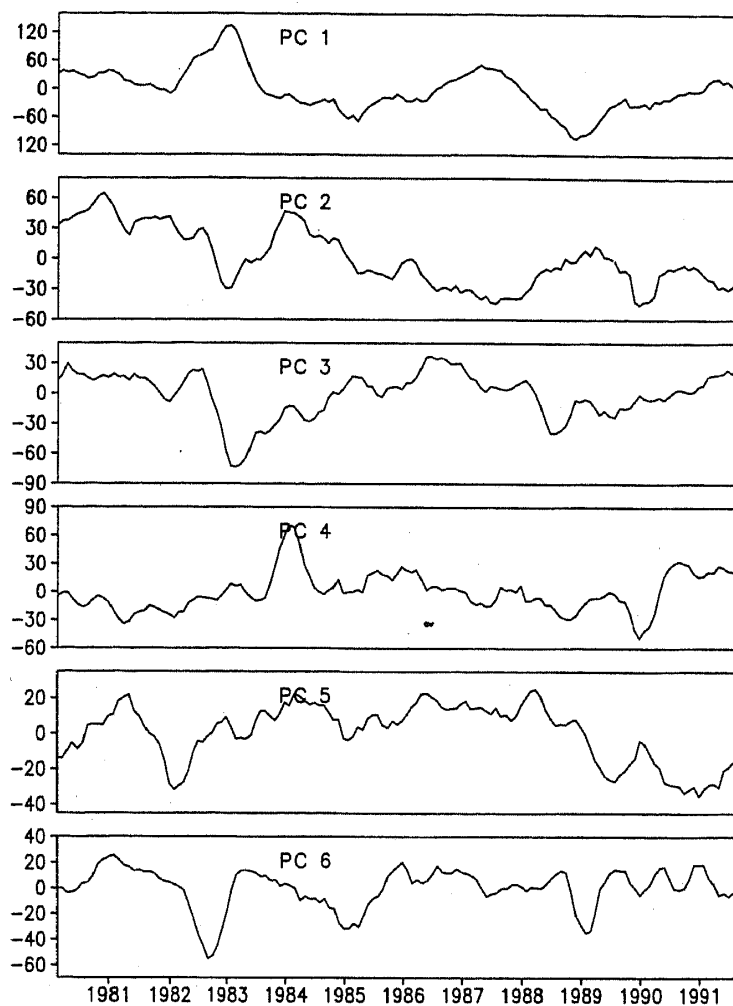
Figure 2. First six principal components (PCs) of the combined EOFs of zonal and meridional winds at 1000, 850 and 200 mb.

slope of the linear regression line was tested by calculating the 't-statistic' of the regression coefficients for all the PCs. It was found that the regression coefficients for the PC1, PC2 and PC5 are significant at 0·005 level. To test if the five-month running mean introduced any spurious low frequency signal, we carried out a similar combined EOF analysis with the unfiltered wind anomalies. It is seen (not shown) that the structure of all the leading EOFs remains unchanged. The PCs also contain exactly the same low frequency variability. The only difference is that some weak high frequency noise is superimposed on the low frequency variability. The PCs have the same trends seen in figure 2. Thus, no additional trend or low frequency variability is introduced by the filter.

This trend in the ECMWF data is related to the change in the 'analysis system' (including model changes and analysis-initialization changes). However, to establish whether this trend is due to model changes or not, we must compare this data with an independent analysis that does not depend on the model. As we have mentioned earlier, because the multivariate EOFs are interrelated, if we have an independent analysis on one variable, comparison between patterns of variability of this variable with those of
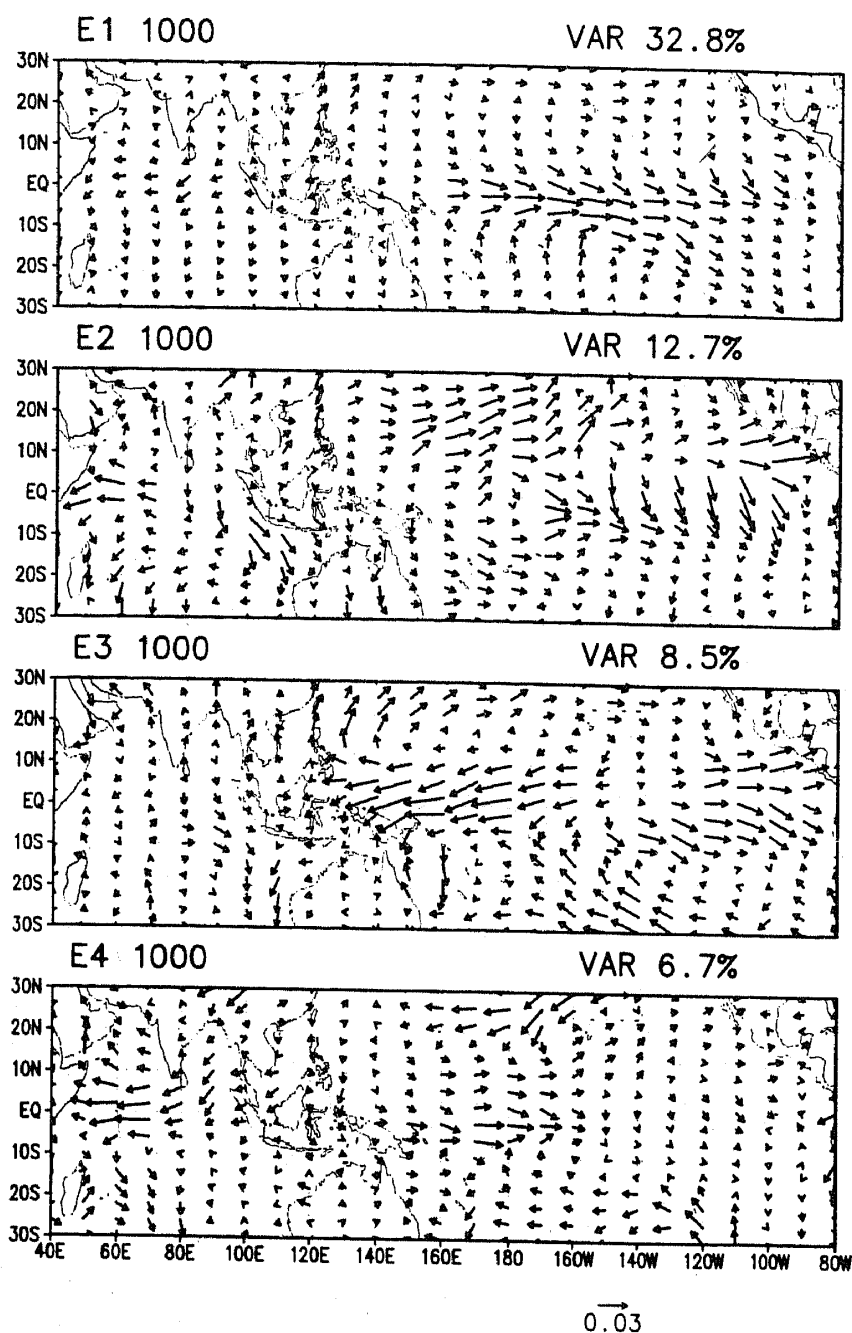
E1 1000   VAR 32.8%

E2 1000   VAR 12.7%

E3 1000   VAR 8.5%

E4 1000   VAR 6.7%

0.03

Figure 3. First four vector EOFs of the winds at 1000 mb. The combined per cent variance explained is also given.

the corresponding variable in multivariate analysis will help us to identify the systematic errors that are common to all variables. FSU analysis of the surface wind provides an alternative to compare the surface wind. Therefore, we carried out a similar EOF analysis of the FSU winds (zonal and meridional winds combined) over the Pacific domain. Figures 5 and 6 show the first four principal components and vector representation of the corresponding EOFs of the FSU winds. We note that the dominant PCs do not show any trend in the FSU winds. The significance of the regression coefficients was carried out as before. It was found that the regression
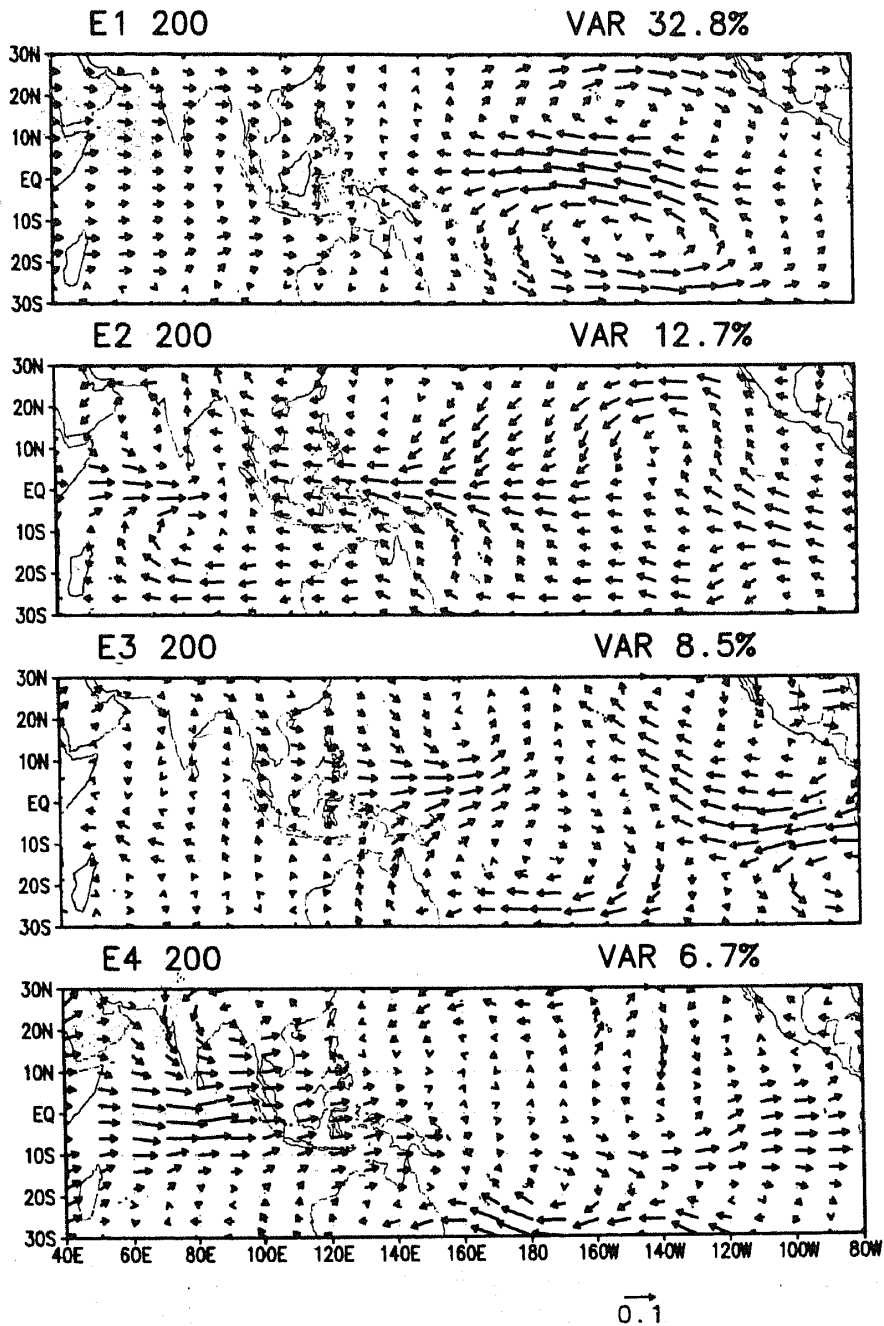
E1 200   VAR 32.8%

E2 200   VAR 12.7%

E3 200   VAR 8.5%

E4 200   VAR 6.7%

$\overrightarrow{0.1}$

Figure 4. Same as figure 3 but for 200 mb.

coefficients of none of the six PCs are significant. Moreover, we note that there is no equivalent of EOF2 of ECMWF winds in the FSU wind. EOF2 of FSU wind qualitatively agrees with EOF3 of the ECMWF surface winds. Thus, we infer that the observed trend in the ECMWF wind analysis is not related to any physical mechanism and the changes in the 'analysis system' seem to have introduced some spurious patterns of variability. Therefore, the use of this data set for interannual variability studies may lead to erroneous conclusions.
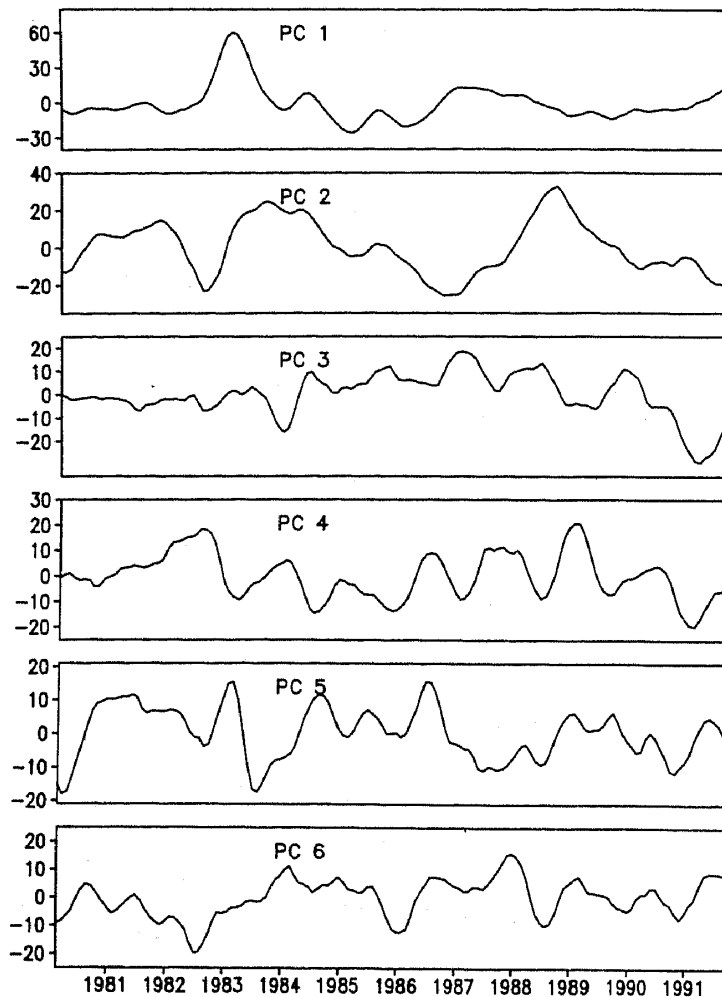
Figure 5. First six principal components of the FSU winds.

3.2 *Origin of the trend*

As documented by Trenberth and Olson (1988) the ECMWF analysis has gone through a number of significant changes (refer to their figure 3). Major changes were introduced in May 1985 when the resolution of the model was increased to T106 (triangular wave 106), and when substantial changes in the physical parameterization affecting the moisture field (clouds, shallow convection and condensation criteria) were introduced. In March 1986 changes in the humidity analyses that had a positive impact in the simulation of tropical rainfall and on latent heat release were also introduced. Prior to this period, the tropical divergent winds were weak, while after this period they were realistic. Thus the mean climatology changed from the period 1980–1985 to the period 1986–1991. Trenberth and Olson (1988) show how the mean Hadley circulation changed around this time. We suggest that the trend seen in the ECMWF analysis is due to a change in the climatological mean from the period before 1985 to the period after 1985. To support this hypothesis, we conduct the two following analyses.

First, we remove the trend from the analysis by objective means. To achieve this we detrended PC1, PC2 and PC5 by fitting a straight line using the least squares method.
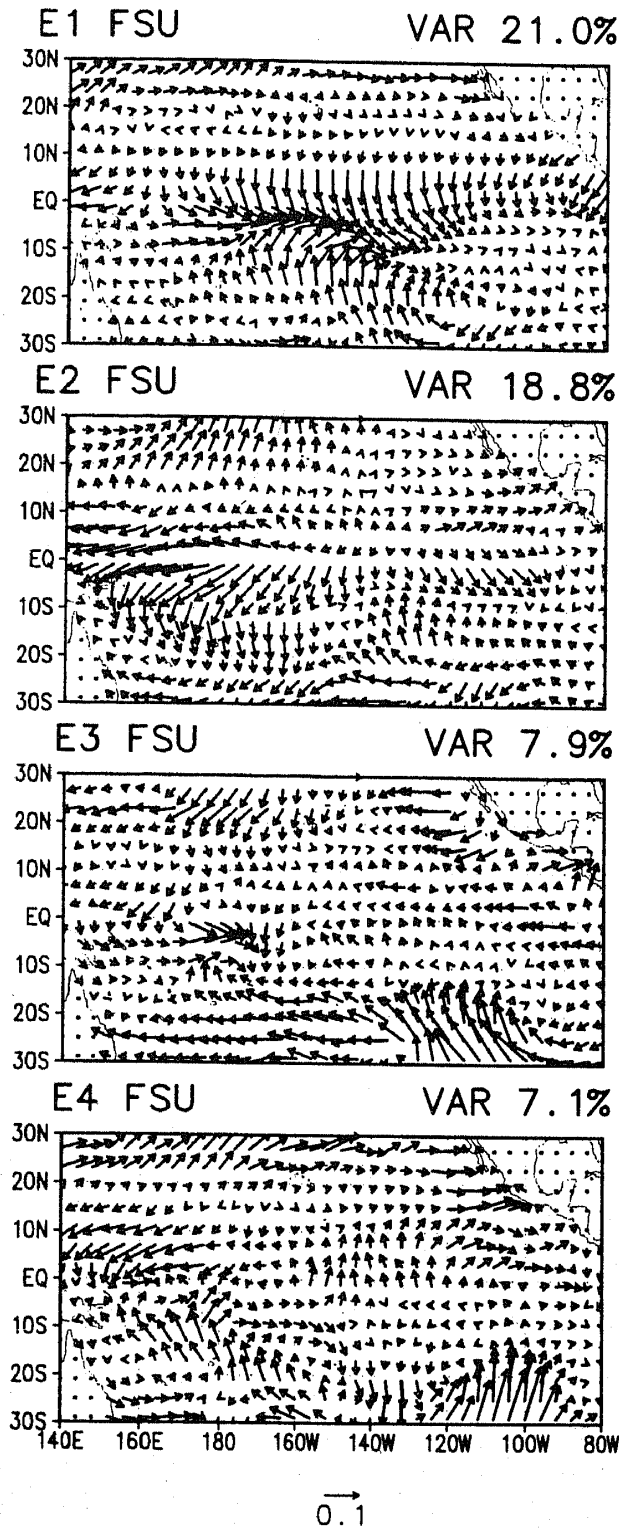
Figure 6. First four vector EOFs of FSU winds. The combined per cent variance explained is also given.

This is shown in figure 7. As mentioned earlier, statistical significance tests indicate that these trends are highly significant. The fields are then reconstructed using the three detrended PCs together with the other PCs without trends. We then conducted an
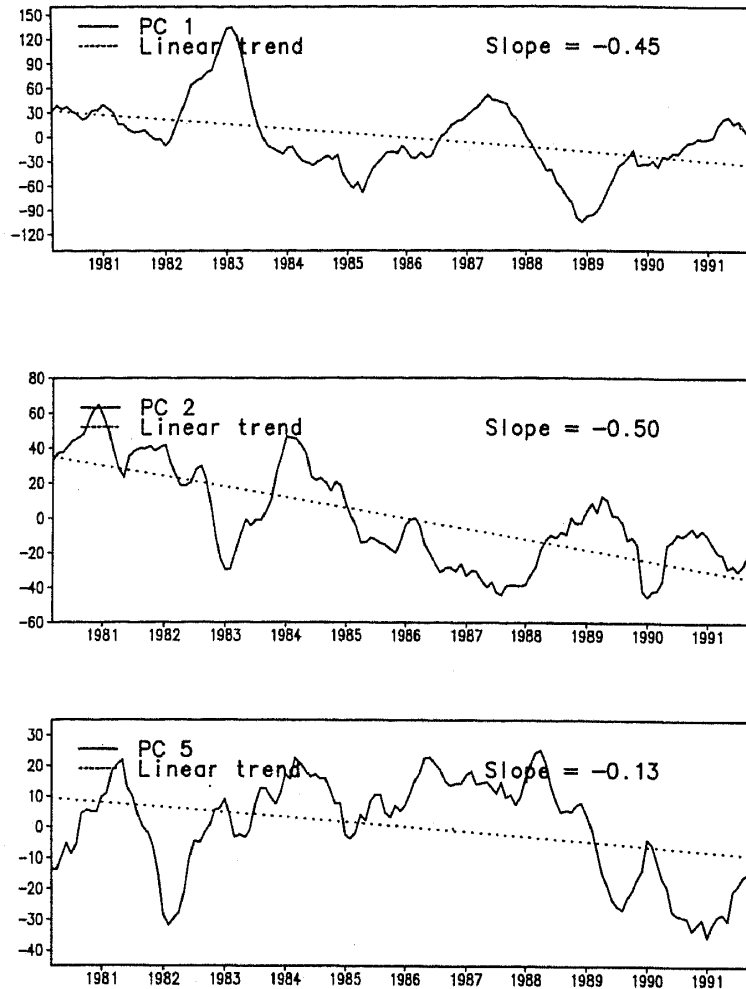
Figure 7. The linear trends obtained for the dominant PCs of figure 2. Shown also is the slope obtained from least squares fit.

EOF analysis as before with the detrended wind analysis. As expected no trend is noted in the PCs of the leading EOFs in this case. The first four vector EOFs of the surface and 200 mb wind of the detrended winds are shown in figures 8 and 9. It is now seen that the first two EOFs of the analysed surface winds agree with the corresponding EOFs of the FSU wind (figure 6). It is interesting to note that the detrending has removed the spurious EOF2, explaining a large amount of variance in the original data (figure 3). Reasonably good agreement between patterns in the detrended surface wind and those of the independent FSU data set over the Pacific indicates that, even though the climatology may have been wrong during the period prior to 1985, the large-scale part of the interannual anomalies during that period may still be reasonable. The time evolution of the first four PCs of the detrended analysis also agrees reasonably well with that of the PCs of the FSU winds. This indicates that major interannual variations are captured in amplitudes as well as in patterns by the detrended analysis. It captures the anomalous surface westerlies and their eastward propagation in the Pacific during 1982/83 and 1987/88 episodes. The first two EOFs represent the low frequency interannual variabilities associated with the El Nino and Southern Oscillation (ENSO). The two anticyclones at 200 mb straddling the equator in the Pacific and their
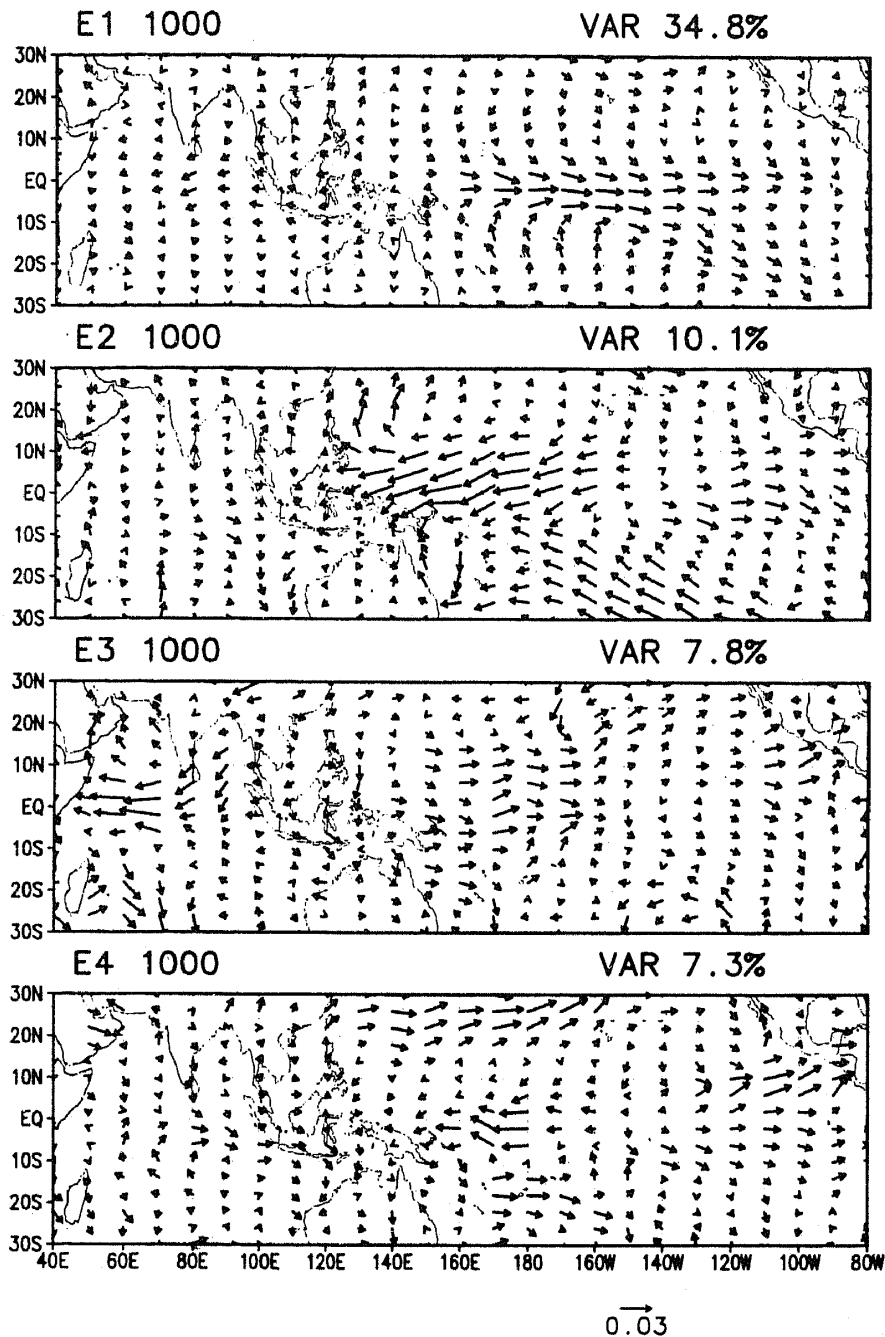
Figure 8. Same as figure 3 but for the detrended data of ECMWF analysis.

eastward migration during major ENSO episodes are also seen. The third EOF has major loading in the Indian Ocean and represents a regional mode of oscillation associated with the Indian monsoon. The actual anomalies are somewhat weaker (about 70% of observed values) during major warm episodes associated with ENSO. This is a general problem with many AGCMs (Goswami *et al* 1995) forced by the observed SST.

Is the detrending of the analysis following the above method equivalent to calculating anomalies with respect to two climatologies; one for the period 1980–1985 and
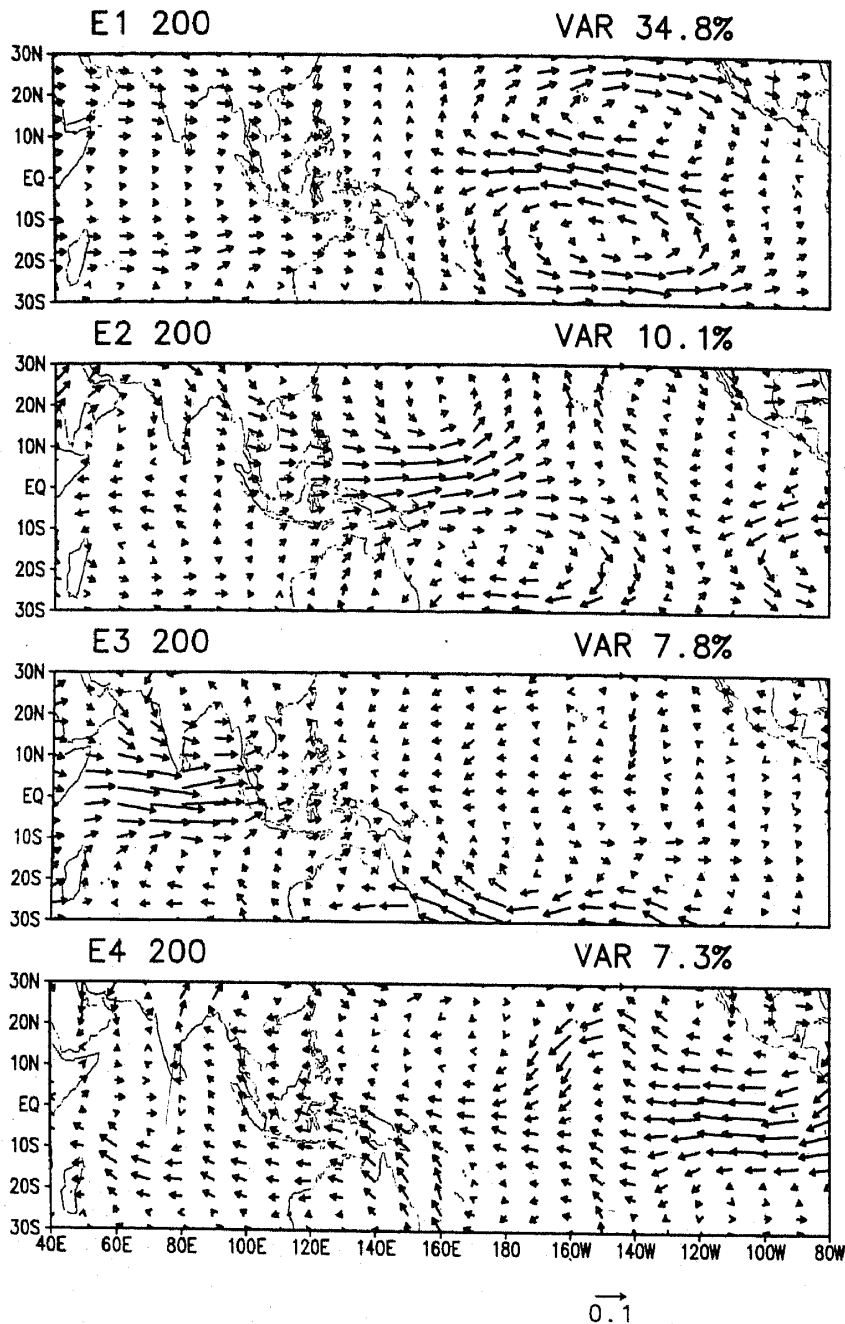
Figure 9. Same as figure 8 but for 200 mb.

another for the period 1986–1991? The delineation of the 12-year period into these two six-year periods is not arbitrary. Even though the changes in the model have been continuous, it is felt that the major model and analysis-initialization changes introduced n mid 1985 had the most significant impact on the tropical analysis. This is why the nalyses before 1985 appears systematically different from that after 1985. Hence the oaration into the two periods. To test this, we calculate climatologies of each field at h grid point separately for the period 1980–1985 and for the period 1986–1991. To strate the differences in the two climatologies, we plot in figure 10, the annual cycle

of the climatological mean zonal wind averaged around the equator (5°S–5°N) corresponding to the two different periods, and their differences at 850 and 200 mb respectively. We note that particularly large changes took place in zonal winds at 850 mb over both the Indian Ocean and Central Pacific sectors. The percentage change from earlier climatology to the later ranges from 75% to 100%. There are also significant changes in the 200 mb zonal wind climatologies for the two periods (figure 10b). However, the mean at this level is higher and the percentage change from the earlier to the later climatologies range from 25% to 50%. The largest changes are in the equatorial region. The anomalies are then calculated with reference to the corresponding climatologies in the two periods. A combined EOF analysis is made with the anomalies of zonal and meridional winds at the surface (1000 mb), 850 and 200 mb. Figure 11 shows the first 4 principal components of the analysis (solid curve) and this is compared with the PCs of the objectively-detrended anomalies (dotted curve). The close matching of the two curves in all the PCs shows that the anomalies calculated by these two different ways are nearly identical. Figure 12 shows the first 4 vector EOFs of the surface wind from the combined EOF analysis of the newly constructed anomalies
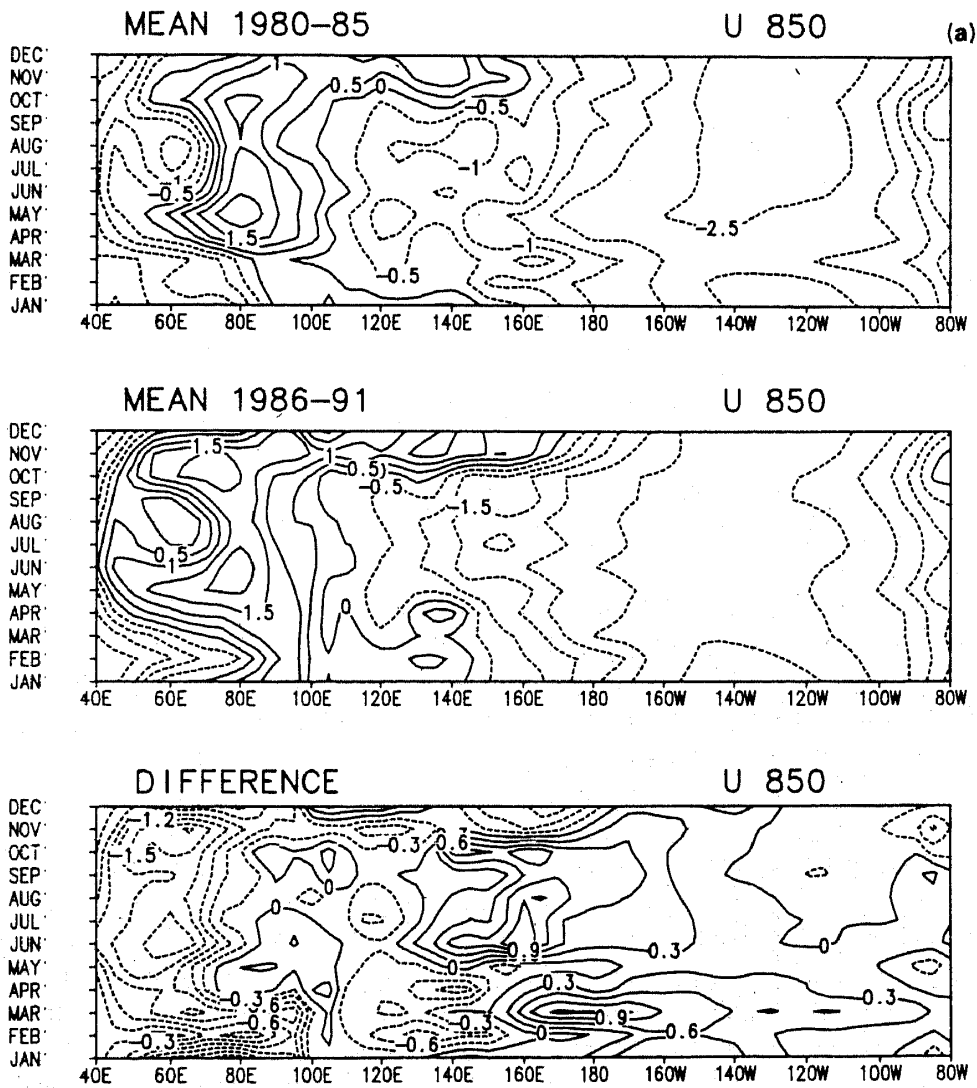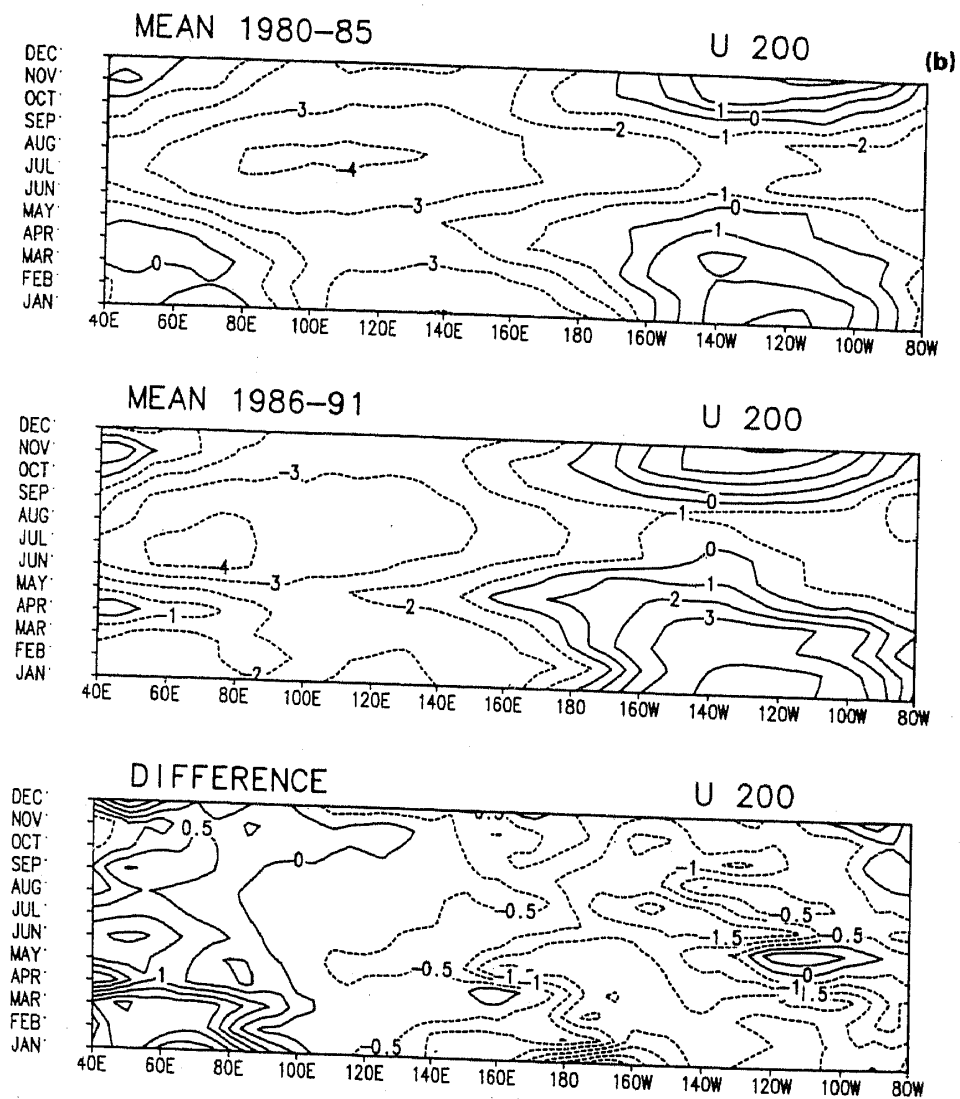
Figure 10. *(Continued)*

MEAN 1980-85   U 200   **(b)**

MEAN 1986-91   U 200

DIFFERENCE   U 200

Figure 10 (a and b).   Annual cycle of climatological zonal wind averaged over 5°S–5°N for the two periods and the difference between these two climatologies at (a) 850 mb and (b) 200 mb.

based on two climatologies. Again we note that they are nearly identical to corresponding EOFs of the objectively-detrended analysis (figure 8). This indicates that the origin of the trend is related to differences in the mean conditions of the model-based analysis system prior to 1985 and in the period after 1985. We also infer that a step-wise or linear way of removing the trends yields the same result.

4. Concluding remark

It has been recognized for some time (Shaw *et al* 1987; Trenberth and Olson 1988) that operational analyses (e.g., the ECMWF analysis) contain discontinuities in time introduced by the changes in the model and the analysis procedure. Efforts are now underway for a reanalysis of the past data (e.g., Kinter and Shukla 1989; Kalnay and Jenne 1991; Schubert *et al* 1993) using a 'frozen' analysis system. As it would still take several years before such a set of homogeneous, consistent data set is available for
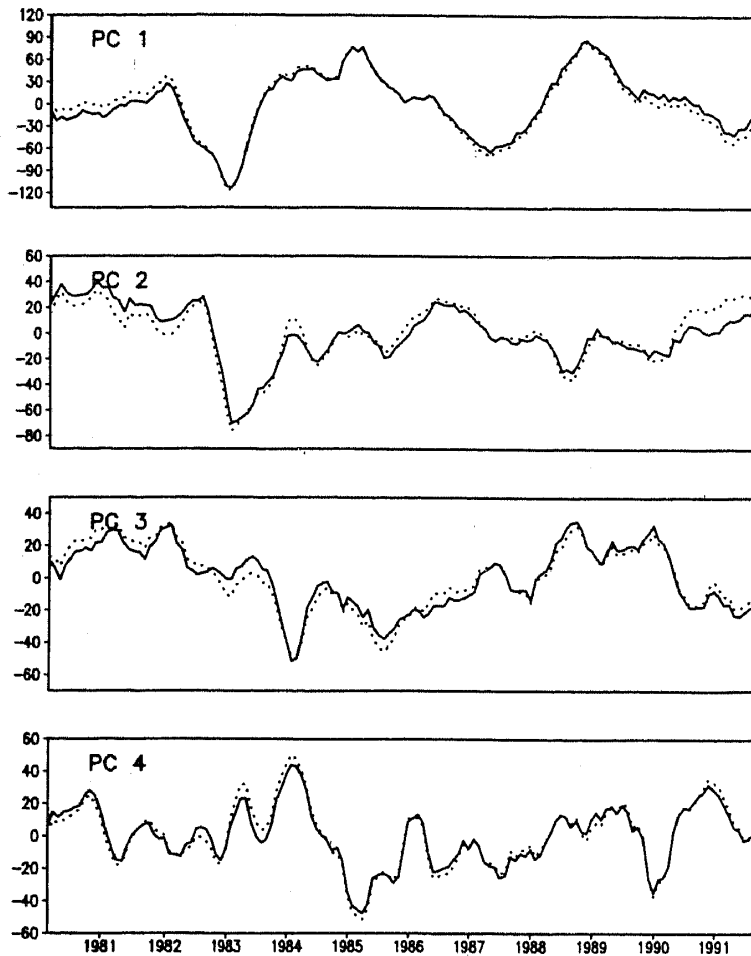
Figure 11. The first four PCs of the combined EOFs of zonal and meridional winds at 1000, 850 and 200 mb but for anomalies formed from the detrended analysis (solid curve) and anomalies formed with respect to the two climatologies (dashed line).

a sufficiently long enough time, some researchers have made use of the existing operational analyses for interannual variability and predictability studies (e.g., Wang 1992; Murakami and Matsumoto 1994; Barnett *et al* 1994). Some other studies (Vincent *et al* 1991; Webster and Yang 1992) have exercised caution and used only the data after 1985.

We show that the changes in the ECMWF analysis system for the period 1980–1991 give rise to a significant trend in the wind analysis. By comparing it with an independent analysis of the surface wind (FSU analysis), we show that this trend is not due to any natural processes. If this trend is not removed, the interannual anomalies prior to 1985 tend to be of one sign while those after it tend to be of the opposite sign. Therefore, erroneous conclusions may be drawn if the trend is not removed. Here we provide a technique of filtering this trend using a multivariate EOF method. The multivariate EOF technique provides a means of identifying the systematic errors that are interrelated at all levels. Then by comparing with an independent analysis at one level (e.g., surface), the systematic errors at all levels may also be corrected. We demonstrated that this trend arises due to changes in the climatology before 1985 and after 1985. We also show that, although the climatology for the period prior to 1985 was weaker and unrealistic, atleast in the tropical region, the interannual anomalies in the
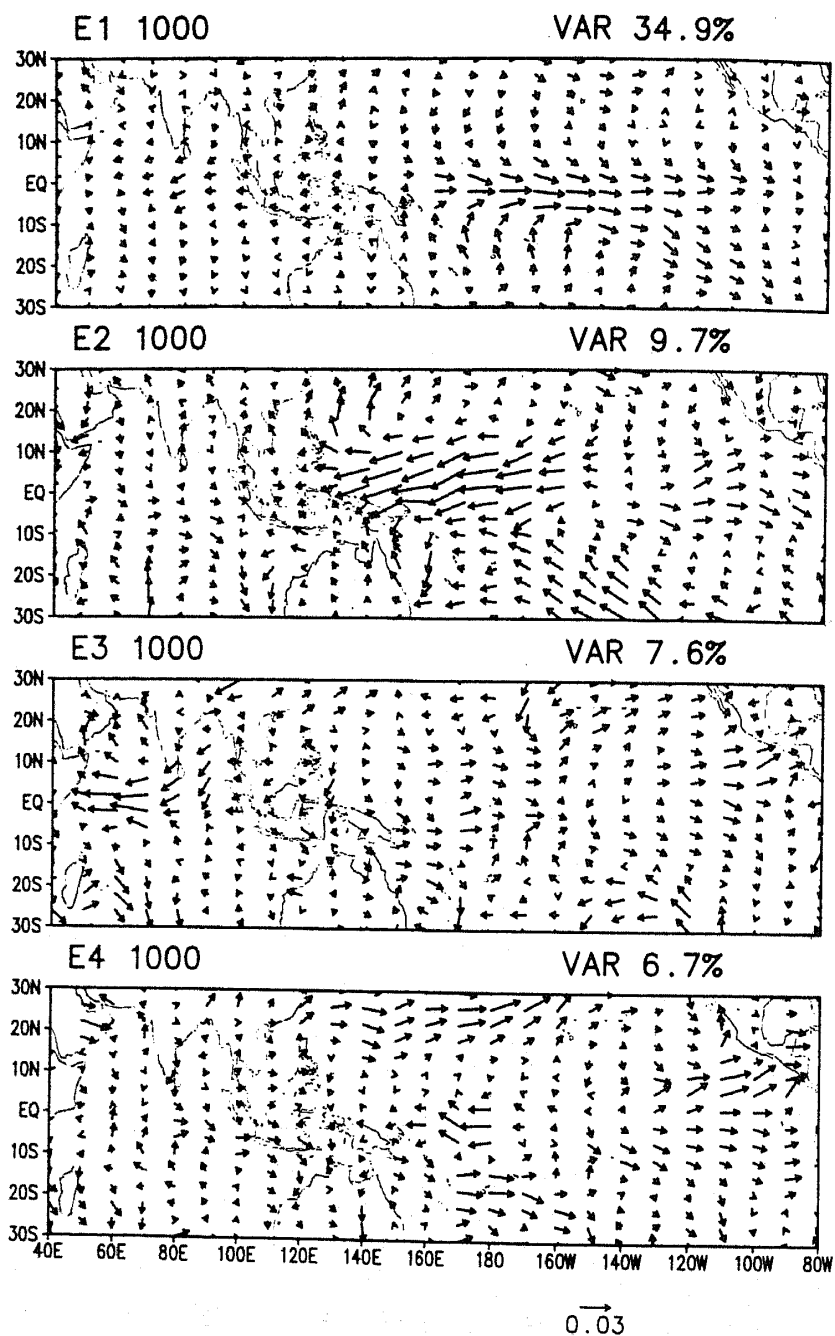
Figure 12. Same as figure 3 but for the anomalies formed with respect to the two climatologies.

detrended analysis are still reasonable. Before the actual reanalysed data are available, the method presented here provides an objective and convenient way of making the data useful for interannual variability studies.

References

Barnett T P, Bengtsson L, Arpe K, Flugel M, Graham N, Latif M, Ritche J, Roeckner E, Schlese U, Schulzweda U and Tyree M 1994 Forecasting global ENSO-related climate anomalies; *Tellus* **A46** 381–397

Bengtsson L and Shukla J 1988 Integration of space and *in situ* observations to study global climate change; *Bull. Am. Meteorol. Soc.* **69** 1130–1143

Bethea R M, Duron B S and Boullio T L 1975 *Statistical methods for engineers and scientists*; (Marcel Pekker Inc.)

Goldenberg S O and O'Brien J J 1981 Time and space variability of tropical Pacific wind stress; *Mon. Weather Rev.* **109** 1190–1207

Goswami B N, Krishnamurthy V and Saji N H 1995 Simulation of ENSO related surface winds in the tropical Pacific by an AGCM forced by observed SST; *Mon. Weather Rev.* **123** 1677–1694

Hurrell J W and Trenberth K E 1992 An evaluation of monthly mean MSU and ECMWF global atmospheric temperatures for monitoring climate; *J. Climate* **5** 1424–1440

Kalnay E and Jenne R 1991 Summary of the NMC/NCAR reanalysis workshop of April 1991; *Bull. Am. Meteorol. Soc.* **72** 1897–1904

Kinter J L and Shukla J 1989 Reanalysis for TOGA; *Bull. Am. Meteorol. Soc.* **70** 1422–1427

Murakami T and Matsumoto J 1994 Summer monsoon over the Asian continent and western north Pacific; *J. Meteorol. Soc. Jpn.* **72** 719–745

Nigam S and Shen H S 1993 Structure of oceanic and atmospheric low-frequency variability over the tropical Pacific and Indian Oceans. Part I: COADS observations; *J. Climate.* **6** 657–676

Schubert S D, Road R B and Pfaendtner J 1993 An assimilated data set for earth science application; *Bull. Am. Meteorol. Soc.* **74** 2331–2342

Shaw D B, Lonnberg P, Hollingsworth A and Unden P 1987 Data assimilation: The 1984/85 revisions of the ECMWF mass and wind analysis; *Q. J. R. Meteorol. Soc.* **113** 533–566

Trenberth K E and Olson J G 1988 An evaluation and intercomparison of global analyses from the National Meteorological Center and the European Center for Medium Range Weather Forecasts; *Bull. Am. Meteorol. Soc.* **69** 1047–1058

Trenberth K E, Large W G and Olson J G 1990 The mean annual cycle in global ocean wind stress; *J. Phys. Oceanogr.* **20** 1742–1760

Trenberth K E and Guillemot C J 1994 The total mass of the atmosphere; *J. Geophys. Res.* **99** 23,079–23,088

Vincent D G, North K H, Velasco R A and Ramsey P G 1991 Precipitation rates in the tropics based on the Q_1-Budget method: 1st June 1984–31st May 1987; *J. Climate* **4** 1070–1086

Wang B 1992 The vertical structure and development of the ENSO anomaly mode during 1979–1989; *J. Atmos. Sci.* **49** 698–712

Webster P J and Yang S 1992 Monsoon and ENSO: Selectively interactive systems; *Q. J. R. Meteorol. Soc.* **118** 877–926