Estimating Subpycnocline Density Fluctuations in the California Current Region from Upper Ocean Observations

ROBERT L. HANEY AND ROBERT A. HALE

Department of Meteorology, Naval Postgraduate School, Monterey, California

CURTIS A. COLLINS

Department of Oceanography, Naval Postgraduate School, Monterey, California

(Manuscript received 1 March 1994, in final form 22 August 1994)

ABSTRACT

A method for extending upper ocean density observations to the deep ocean is tested using a large number of deep CTD (conductivity–temperature–depth) stations in the California Current. The specific problem considered is that of constructing the best estimate for the density profile below a certain depth D given an observed profile above that depth. For this purpose, the estimated disturbance profile is modeled as a weighted sum of empirical vertical modes (EOFs). The EOFs are computed from the surface to 2000 m, using 126 largely independent CTD stations off Point Sur, California. Separate computations are made for the summer half-year (mid-April to mid-October) and the winter half-year (mid-October to mid-April). For each observed density profile, the EOF weights that determine the estimated profile are obtained by performing a successive least-squares fit of the disturbance density profile above D to the first N EOFs. In this study, N is taken to be 7, which is the number of EOFs that account for the "signal" in the profiles as determined by the methods of Preisendorfer et al. and Smith et al. The estimated profiles are then verified against the observed profiles to 2000 m, and the results are presented as a function of the depth D.

In general, the vertical extension method is moderately successful at estimating density fluctuations at and below 500 m from data entirely above 500 m. Observed density profiles to depths shallower than 500 m can be extended to 500 m, with a correlation that depends on the time of year as well as on the depth of the observed profile. For example, a minimum of 200 m of data is needed to perform a useful extension to 500 m, and in all cases extensions are more successful in winter than in summer. As might be expected, correlations between the estimated profiles and a seven-mode reconstruction of the observed profiles, representing the "signal" part of the observed profiles, are somewhat higher. The dynamic height of the sea surface relative to 500 m, an important integral quantity, can be estimated quite well with only 300 m of data. A practical result of this study is that data down to only 200 or 300 m, as might be acquired by a SeaSoar CTD survey, can be extended to 500 m or more using the EOF-based method with a known and useful level of skill. Tests with a small sample of independent data confirm the above results. The success of the method is attributed to the fact that in this part of the ocean the dominant EOFs represent variability in the upper ocean that is also reflected at deeper depths.

1. Introduction

A classic problem in synoptic ocean analysis is that of estimating the deep ocean density structure from observations in the upper ocean alone. The problem is of general interest simply because there are far more observations in the upper ocean than in the deep ocean. In addition, new ocean instruments and measuring devices such as SeaSoar now make it possible to carry out hydrographic surveys of the oceanic mesoscale that are nearly synoptic in time. This is possible, however, only if the measurements are restricted to rather shallow depth ranges, for example, the upper 200 m. The extent to which such an upper ocean survey can adequately describe the dynamical features in and below the pycnocline at a given time is therefore an important question. Subpycnocline density fluctuations on synoptic scales are important in their own right and they also influence the dynamics of the pycnocline and upper ocean itself. For example, if hydrographic data from the upper ocean are utilized in a numerical model, the extension of the data to the deeper ocean influences the upper ocean currents through the hydrostatic and geostrophic constraints and strongly influences the subsequent model predictions (Hurlburt et al. 1990).

Several recent studies have been directed at various aspects of this problem. For example, Smith et al. (1985) evaluated the skill with which the amplitude of

Corresponding author address: Robert L. Haney, Department of Meteorology, U.S. Naval Postgraduate School, 589 Dyer Road, Bldg. 235, Rm. 254, Monterey, CA 93943-5114.

quasigeostrophic dynamical modes (which have important signals at midocean depths) could be estimated from mixed CTD–XBT surveys in the California Current. For shallow CTD and XBT casts (<750 m), the best results were obtained using a method based on the covariances between the quasigeostrophic modal amplitudes and the amplitudes of the first few empirical vertical modes. Another approach for specifying density disturbances in the deeper ocean from upper ocean observations involves the use of feature models, which are relationships between sea surface temperature patterns and subsurface thermal features developed for the Gulf Stream and its warm and cold core rings (Robinson et al. 1988). Another approach is the use of statistical relations between sea level and subsurface temperature or pressure anomalies determined from numerical model simulations (Hurlburt et al. 1990; Mellor and Ezer 1991). These studies have shown that fluctuations in model sea level are highly correlated with fluctuations in model density at depths of 1–2 km.

In the present study, we test an empirical method that in some ways represents an observations-based generalization of the feature model approach. The method consists of fitting a disturbance density profile from the upper ocean to empirical vertical modes (EOFs) determined from historical CTD (conductivity–temperature–depth) data taken in the region of interest. The vertical EOFs, which extend from the surface to the deep ocean, thereby represent a set of "feature models" in this method. The method is potentially useful for real-time ocean analysis and for initializing numerical ocean prediction models.

2. The vertical extension method

The problem being considered is that of constructing the best estimate for the density profile below a certain depth D given the observed density profile above that depth. As described in detail below, the estimated disturbance density profile is modeled as a weighted sum of the first N empirical vertical modes. To begin the estimation procedure, the potential density Θ is expressed in terms of a mean part $\bar{\Theta}(z)$ and a disturbance part $\Theta'(z)$,

$$\Theta(z) = \bar{\Theta}(z) + \Theta'(z). \tag{2.1}$$

In practice $\bar{\Theta}(z)$ would have to be specified from historical data. In this study $\bar{\Theta}(z)$ is computed as the mean of the profiles used in the analysis (see Fig. 2). The disturbance profile Θ' is modeled as a weighted sum of the first N vertical EOFs. Denoting it $\hat{\Theta}(z)$, it is expressed as

$$\hat{\Theta}(z) = \sum_{i=1}^{N} A_i \Theta_i(z), \tag{2.2}$$

where $\Theta_i(z)$, $i = 1$–N, are the EOFs and A_i are their amplitudes (weights). The A_i are obtained by performing a successive least-squares fit of Θ' to the first N empirical vertical modes above the depth D. Thus, performing a least-squares fit of Θ' to the first mode determines A_1,

$$A_1 = \frac{\langle \Theta' \Theta_1 \rangle}{\langle \Theta_1 \Theta_1 \rangle}, \tag{2.3}$$

where the angle brackets represent a vertical integral (numerical sum) from the surface to D. Performing a least-squares fit of the residual profile $\Theta' - A_1\Theta_1$ to the second mode determines A_2,

$$A_2 = \frac{\langle (\Theta' - A_1\Theta_1)\Theta_2 \rangle}{\langle \Theta_2 \Theta_2 \rangle}, \tag{2.4}$$

and so forth up to A_N. Using sequential fits ameliorates the fact that the EOFs are generally not orthogonal over the shallow depths D. Thus, variability that can be explained equally well by several different modes owing to their nonorthogonality is, in this method, always attributed to the lowest mode. This assumption is justified simply on statistical grounds. In the following section the method is tested using a large number of deep CTD stations off Point Sur, California.

3. The POST CTD data

The CTD data used in this study were collected as part of the Point Sur Transect (POST) program (Tisch et al. 1992). The observation site extends from 20 km to about 250 km offshore of Point Sur, California, near 36°N (Fig. 1). All of the CTD stations were taken in water depths greater than 2000 m, and most of them were taken seaward of the region of strong coastal upwelling next to the coast. Table 1 gives a summary of

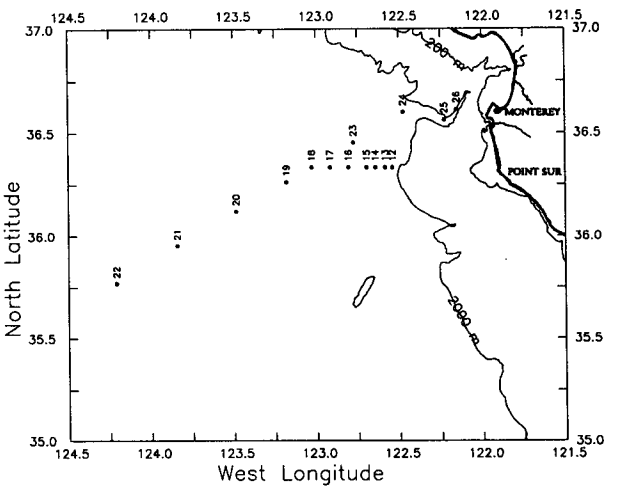

FIG. 1. Location of the POST CTD stations used in this study.

TABLE 1. Dates of the POST cruises from which the CTD data for this study were taken. Of the 250 total CTDs, only about half of them (126) were selected for analysis, while the other half (124) were eliminated (culled). The selection procedure consisted of culling out neighboring stations from the same cruise that had a correlation of 0.5 or greater with a selected station. In this way, the average station spacing is about 25 km, and the profiles so selected have a greater degree of independence than would otherwise be the case.

	Number of CTD stations		
Month	Selected	Culled	Total
April 1988	7	5	12
August 1988	5	10	15
September 1988	7	5	12
November 1988	11	2	13
February 1989	9	3	12
March 1989	7	6	13
May 1989	7	4	11
July 1989	5	6	11
September 1989	5	9	14
November 1989	8	7	15
January 1990	5	9	14
March 1990	6	7	13
May 1990	7	8	15
June 1990	5	10	15
August 1990	5	8	13
October 1990	5	8	13
December 1990	10	4	14
February 1991	6	8	14
April 1991	6	5	11
Total	126	124	250

all of the data that were taken approximately every two months during the three year period from April 1988 through April 1991. Details of the quality control and processing of the raw CTD data to compute the density perturbations used in the following analysis can be found in Tisch et al. (1992).

Since the goal of this study is to present regional EOFs and to use them to test a vertical extension method, we have tried to estimate the statistical significance of our computations in several ways. First, we utilize an a priori knowledge of the annual variability in the region by considering two halves of the year separately. One half-year runs from mid-April to mid-October, hereafter referred to as the upwelling or summer season, and the other half-year runs from mid-October through mid-April, hereafter referred to as the winter season. With this choice, all of the POST cruises assigned to the upwelling season occurred after the well-known spring transition that occurs each year near the end of March in this region (Strub et al. 1987; Lentz 1987). Because the stations are spaced close together, it is clear that not all of the CTD casts from a single cruise can be considered independent. In order to retain as large a sample of CTDs as possible, and at the same time to remove some of the redundancy caused by the station spacing, we culled out neighboring CTD stations that had an overall correlation greater than 0.5 with a station previously selected for analysis. With this screening procedure, approximately one-half the stations were culled out (see Table 1). As a result, the number of profiles selected for analysis was ND = 64 in summer and ND = 62 in winter, and the average spacing between selected stations was about 25 km. We thereby feel justified in applying standard statistical techniques to the data. As a partial check on the robustness of our results, the vertical extension method was also evaluated using a small sample of independent CTDs (ones not used to compute the EOFs). However, unless stated otherwise, all the analysis shown below is based on the selected POST data.

Figure 2 shows the mean potential density $\bar{\Theta}(z)$ and the buoyancy frequency from the surface to 2000 m for both the summer and winter halves of the year. The stratification associated with the mean pycnocline occurs in the upper 300–500 m, with the maximum stratification near 60-m depth. The annual variation is rather small and confined to above 100 m where the water is less dense and more stratified in summer than in winter. As noted earlier, the study area is located seaward of the narrow upwelling zone immediately next to the coast and therefore it experiences a typical, but weak, annual cycle.

The density EOFs are shown in Fig. 3. The statistical significance of the EOF shapes is important when attempting to ascribe a particular physical meaning to a given individual mode. This significance was estimated using an eigenvalue method (North et al. 1982) and a bootstrap method (Smith 1984; Smith et al. 1985). In the eigenvalue method, one compares the sampling error $\delta\lambda$ of a particular eigenvalue λ, $\delta\lambda \sim \lambda(2/\mathrm{ND})^{1/2}$, to the spacing $\Delta\lambda$ between λ and a neighboring eigenvalue (North et al. 1982). The shape of the EOF corresponding to λ is statistically significant only if $\delta\lambda < \Delta\lambda$ holds for both neighboring eigenvalues of λ. The interpretation is that if a group of true eigenvalues lies within one or two $\delta\lambda$ of each other (i.e., if $\Delta\lambda \leqslant \delta\lambda$), then they form an "effectively degenerate multiplet," and the sample eigenvectors (EOFs) are simply a random (nonunique) mixture of the true eigenvectors (North et al. 1982). Table 2 shows the results of the eigenvalue test. The shapes of no more than the first five EOFs in summer and the first four EOFs in winter appear to be significant, with overlapping degenerate pairs occurring among higher-order EOFs. However, fewer EOFs may in fact be significant since North et al.'s criteria uses a relatively low level of significance and our data are not entirely independent. In the bootstrap method, a random number generator is used to draw a resample of ND new density profiles, with replacement, from the original ND profiles (this resample may have more than one copy of some profiles and none of others), and an EOF analysis of this resample is performed. This procedure is repeated 100 times, and the distribution of the results is used to

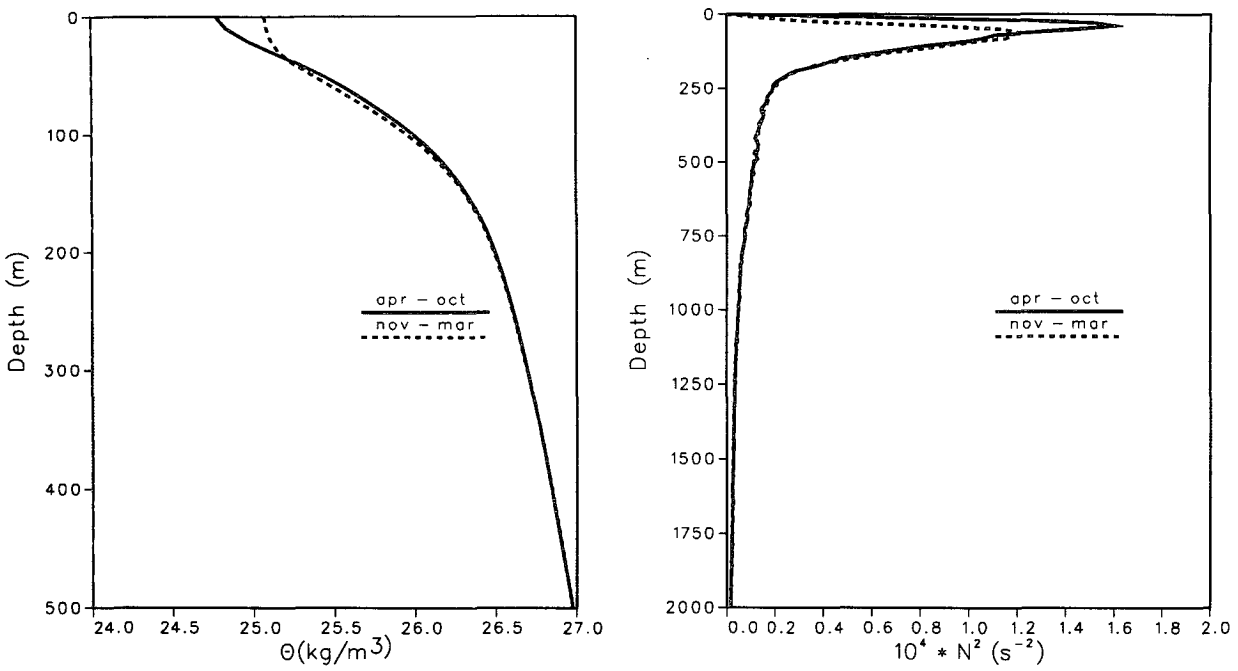

FIG. 2. Mean profiles of density (Θ − 1000) to (a) 500 m and buoyancy frequency to (b) 2000 m. For April–October the sample size is 64 and for November–March it is 62 (see Table 1).

specify the confidence limits shown in Fig. 3. The 95% confidence limits on the shapes of the EOFs shown in Fig. 3 also indicates that the first five EOFs in summer and the first four EOFs in winter are significant, that is, have tight confidence limits away from zero.

Although the shapes of the first four or five EOFs may be robust in the sense discussed above, we will only ascribe a physical meaning to the first two modes. In general, the density EOFs in Fig. 3 are similar to those presented in previous studies of this region (Chelton 1980; Smith et al. 1985; Rienecker et al. 1987). Comparing Figs. 3a and 3b, it can be seen that there is very little seasonal difference apart from a slight vertical displacement of the modes in summer compared to winter. The first mode represents variability associated with the equatorward vertical shear of the California Current in the main seasonal pycnocline that slopes upward toward the coast in the upper 500 m (Chelton 1980). Note that in the present analysis, in which the CTD stations are from different locations as well as at different times, variability represented by the EOFs can be either spatial or temporal or both. The second mode, which is similar to the first mode except for a change in sign above about 70 m, resembles the second, or "seasonal mode," identified off Northern California by Bray and Greengrove (1993). It is interpreted here as the offshore response of the mixed layer and main pycnocline to local variations in the alongshore wind stress. Seaward of the shelf and slope, upwelling favorable winds, for example, produce offshore transport in the mixed layer and downward motion in the pycnocline below. This leads to density increases in the mixed layer (due to horizontal advection) and density decreases in the pycnocline (due to vertical advection), with a zero crossing near the base of the mixed layer as seen in Fig. 3. Note that mode 1, and also mode 2 in summer, has a significant amplitude only in the upper 500 m. These relatively shallow modes of variability obviously cannot be very helpful in estimating density fluctuations below 500 m from observations above that depth. On the other hand, the third and fourth modes, and the second mode in winter, which have significant signals below 500 m, can be useful in this regard if they can be detected in shallow CTD casts.

The above results concerning the statistical significance of the EOF shapes is consistent with the study of North et al. (1982). They showed that for many cases of geophysical interest the sampling errors are often unacceptably large for samples of even a few hundred independent realizations. It is important to note, however, that in this study *we are not especially concerned with the shapes of the EOFs per se.* In the present application, we wish only to have a convenient basis set for representing disturbance density profiles. The sample EOFs computed from our finite dataset form such a complete basis set, and a subsum of these EOFs is expected to account for about as much variance as the corresponding subsum of the true EOFs (North et al. 1982).

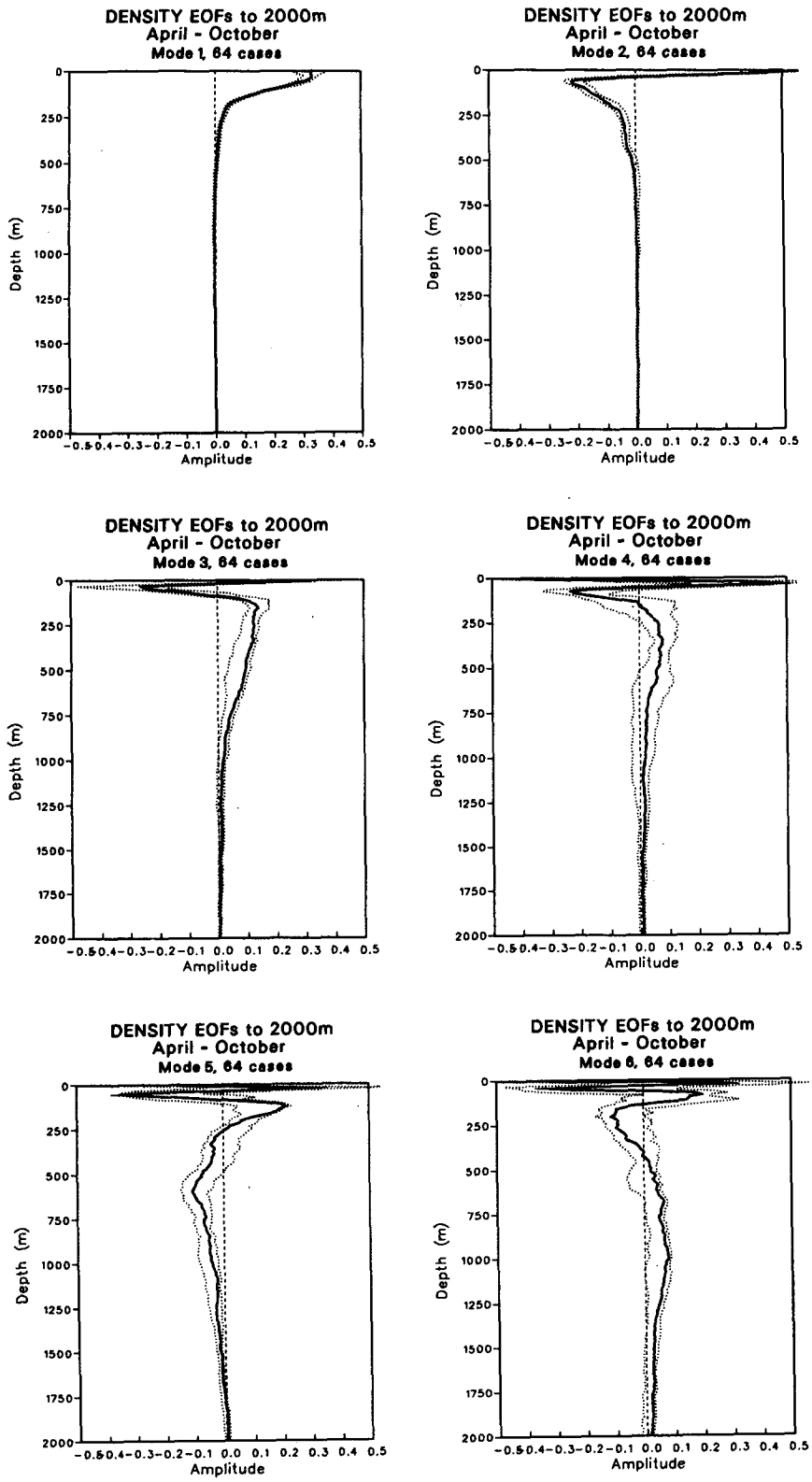

FIG. 3. Density EOFs (kg m^{-3}). (a) April–October, 64 stations and (b) November–March, 62 stations. The 95% confidence limits are estimated using the bootstrap method described in the text.

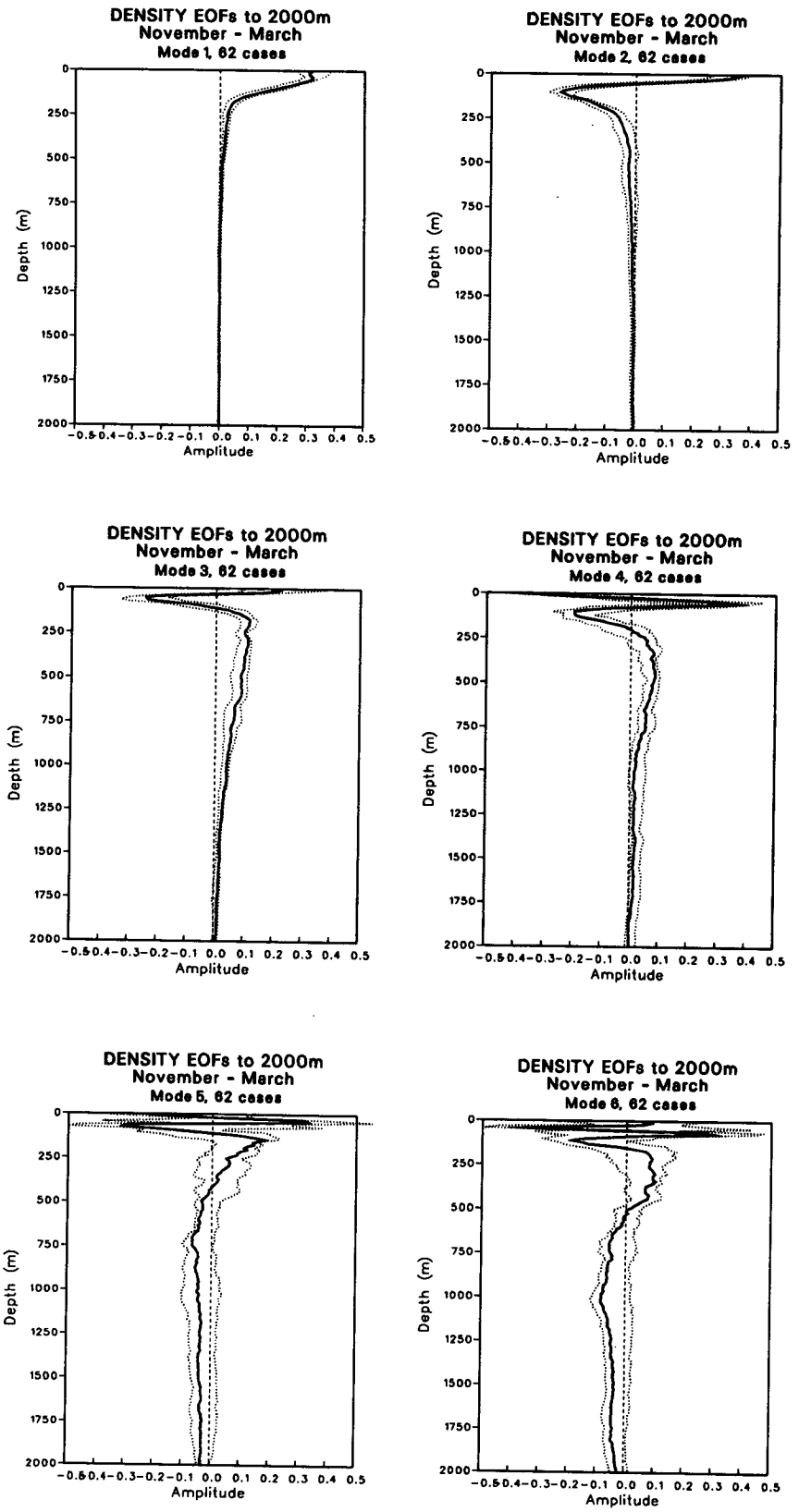

FIG. 3. (*Continued*)

TABLE 2. First 12 eigenvalues and their estimated standard error using the method of North et al. (1982) described in the text. The standard error $\delta\lambda \sim \lambda\epsilon$, where $\epsilon = (2/N)^{1/2}$ and N is the number of profiles used (64 in summer and 62 in winter).

	Summer half-year			Winter half-year		
Mode	$\lambda - \delta\lambda$	λ	$\lambda + \delta\lambda$	$\lambda - \delta\lambda$	λ	$\lambda + \delta\lambda$
1	1.0672	1.2963	1.5255	0.6131	0.7473	0.8815
2	0.1734	0.2106	0.2479	0.1482	0.1806	0.2131
3	0.0476	0.0578	0.0680	0.0494	0.0602	0.0710
4	0.0259	0.0315	0.0370	0.0142	0.0173	0.0204
5	0.0142	0.0172	0.0202	0.0077	0.0094	0.0111
6	0.0079	0.0096	0.0113	0.0064	0.0078	0.0092
7	0.0065	0.0079	0.0093	0.0035	0.0042	0.0050
8	0.0036	0.0044	0.0051	0.0031	0.0038	0.0045
9	0.0021	0.0025	0.0030	0.0018	0.0022	0.0025
10	0.0016	0.0020	0.0023	0.0015	0.0018	0.0021
11	0.0013	0.0016	0.0019	0.0013	0.0015	0.0018
12	0.0011	0.0013	0.0016	0.0010	0.0012	0.0015

Note that the above tests on the EOF shapes provide no guidance in determining the number of EOFs that can represent the profiles efficiently. For this purpose, additional assumptions about the nature of the "noise" in the data must be made. In this study we follow the ideas of Preisendorfer et al. (1981) and Smith et al. (1985). To distinguish signal from noise, Preisendorfer et al. (1981) compare the computed eigenvalues or (equivalently) the percent of explained variance with the eigenvalues (or the percent of explained variance) generated by simulated density data having the same variance profiles as the original data but having no correlation with depth (Fig. 4). Although there is some subjectivity in applying this method, the results in Fig. 4 suggest that only the first two EOFs can be attributed to a vertically coherent signal. As pointed out by Smith et al. (1985), this test does not allow each successive eigenvalue to be evaluated independently of the lower modes. To allow for this, and to compare each successive eigenvalue equally, Smith et al. (1985) compare the percent of *residual* variance explained by a given mode to the corresponding value generated by the sim-

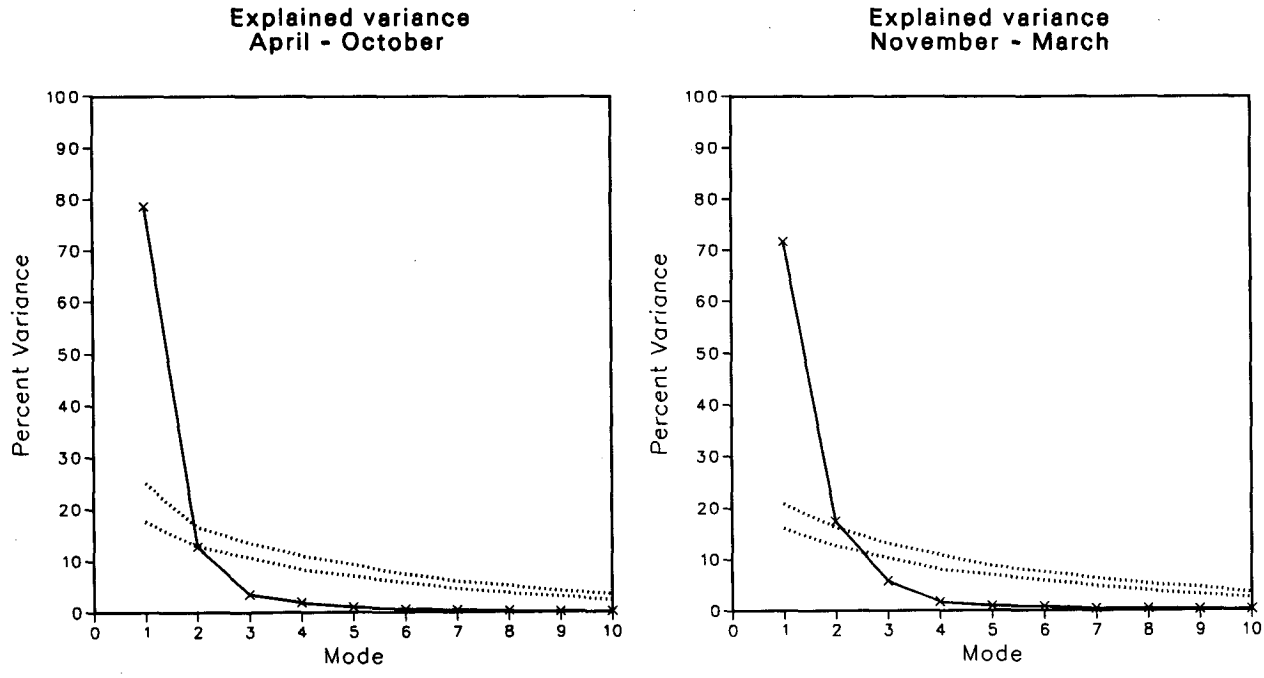

FIG. 4. Percent of variance (scaled eigenvalue) accounted for by each EOF, up to mode number 10, for (a) April–October and (b) November–March. The 95% confidence limits on the simulated data (dotted lines) are estimated using the bootstrap method described in the text.

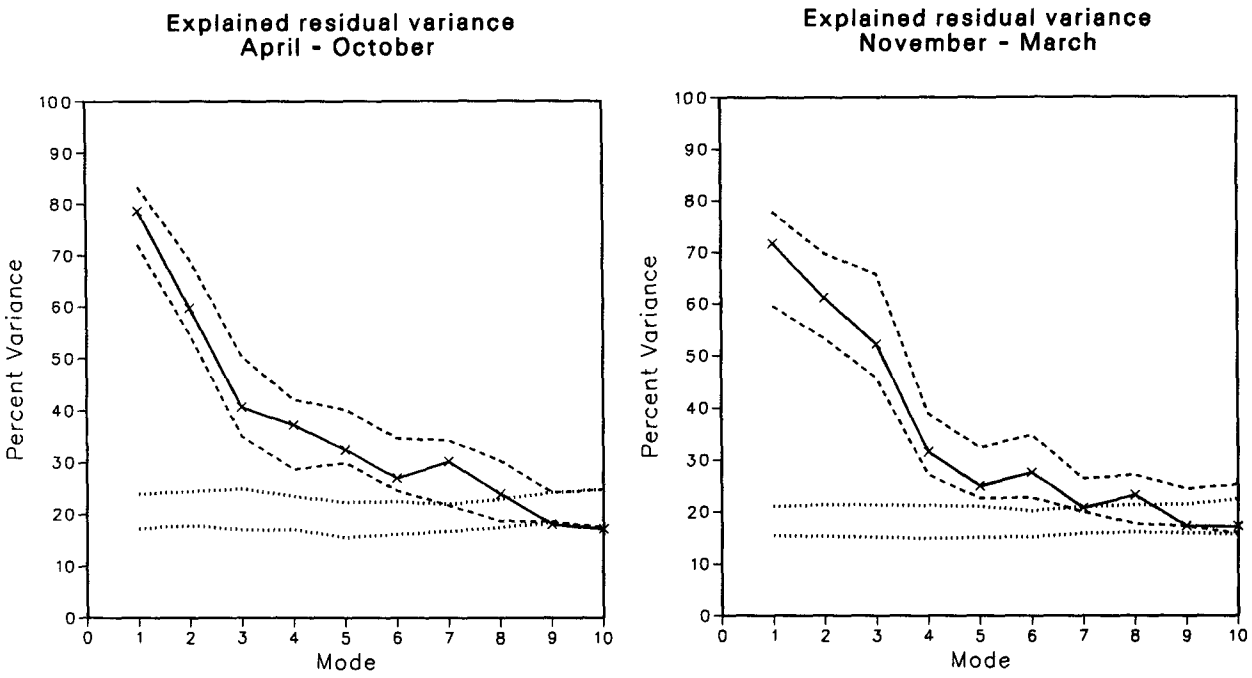

FIG. 5. Same as Fig. 4, except the percent of residual variance accounted for by each EOF is shown. The 95% confidence limits on the residual variance is also shown (dashed lines).

ulated data. In this way, the variance explained by the first EOF does not affect the evaluation of the second EOF, and the variance explained by the second EOF does not affect the evaluation of the third EOF, etc. The result of this calculation (Fig. 5), taking into account the uncertainty of the eigenvalues (bootstrapped confidence limits), indicates that all of the vertically coherent signals can be represented by the first seven EOFs. The result of this test is therefore quite different from that of the Preisendorfer et al. (1981) test (Fig. 4). A subjective comparison between the observed profiles of disturbance density and the corresponding profiles reconstructed from the first seven EOFs (see Fig. 11, first frames) supports the choice of seven modes for the cutoff between signal and noise. At the same time, a similar comparison between the observed profiles and those reconstructed from only two modes (not shown) reveals large differences that most physical oceanographers would not consider attributable simply to "noise." On the basis of this analysis, subjective as it admittedly seems to be, we consider all of the profile variability that is contained in the first seven modes to be "signal" and the remaining variability to be "noise." In total, the first seven modes account for over 99% of the variance in both winter and summer (Table 3). Having identified the first seven EOFs as a convenient basis set for describing the observed "signal," we now examine the extent to which these modes can be identified from shallow casts and used to estimate the profiles at depth.

TABLE 3. Percent variance (PV) and cumulative percent variance (CPV) accounted for by the first seven EOF modes. Data are taken from Fig. 4.

	Summer half-year		Winter half-year	
Mode	PV	CPV	PV	CPV
1	78	78	71	71
2	13	91	18	89
3	4	95	6	95
4	2	97	2	97
5	1	98	1	98
6	1	99	1	99
7	<1	99	<1	99

4. Results of vertical extension

In this section, density profiles below the depth D are estimated by fitting the first seven EOFs to the selected profiles above D. The accuracy with which an estimated profile (E) matches an observed profile (O) is measured by their correlation coefficient r,

$$r = \frac{\langle E'O' \rangle}{[\langle (E')^2 \rangle \langle (O')^2 \rangle]^{1/2}}. \qquad (4.1)$$

In (4.1), angle brackets denote an average, at a fixed depth, over all of the selected profiles for the half-year in question, and the prime denotes the departure from this average.

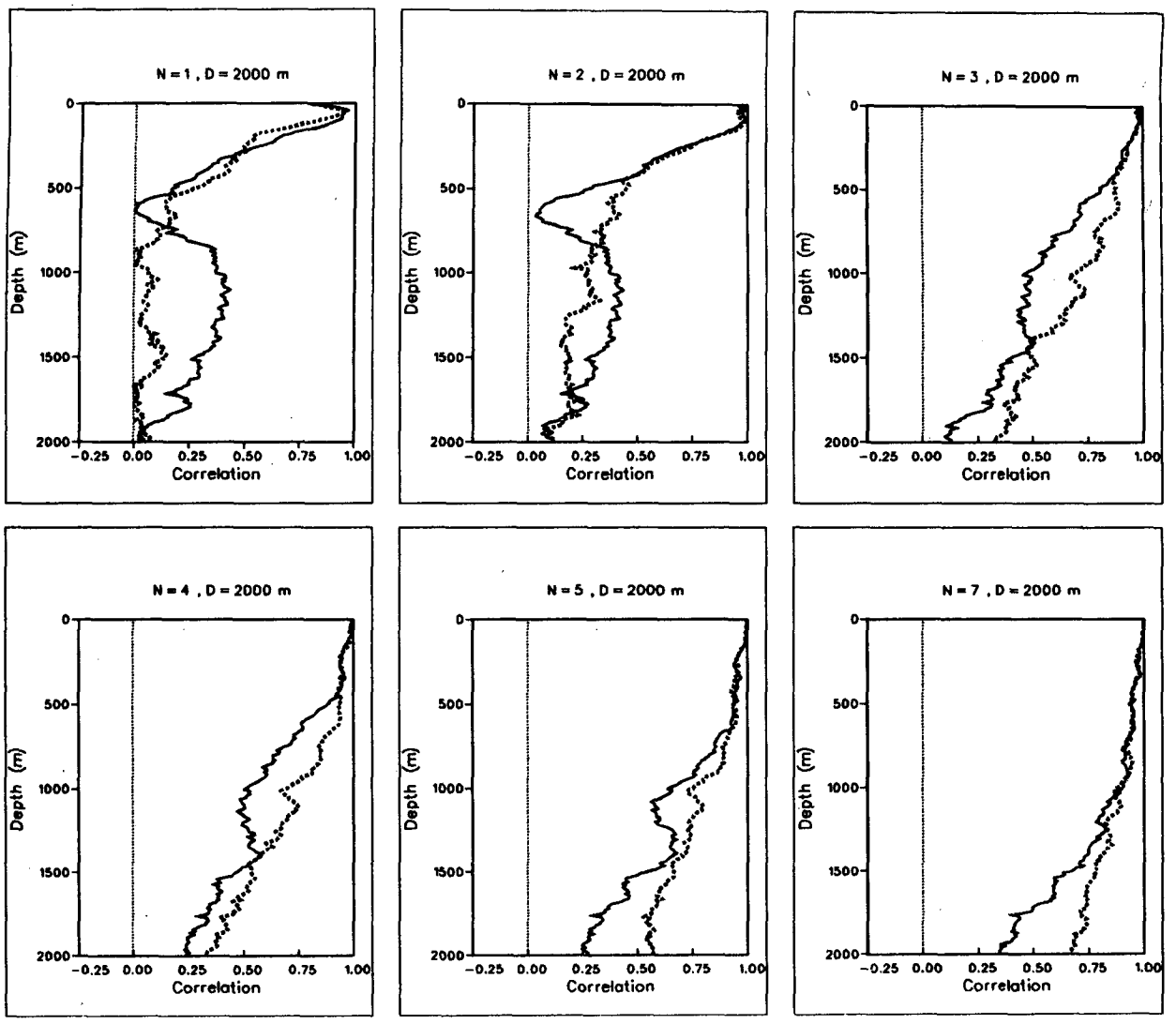

FIG. 6. Correlation r between the estimated and the observed profiles when the fitting is performed to 2000 m and N EOFs are used. Profiles of r are shown for N = 1, 2, 3, 4, 5, and 7 for both the summer half-year (April–October, ND = 64, solid) and the winter half-year (November–March, ND = 62, dashed).

Figure 6 shows how the estimated profiles approach the observed profiles as more and more modes are used, and the fitting is performed over the full water column (D = 2000 m). In both halves of the year the correlation is seen to increase as the number of EOF modes that are used to estimate the profiles is increased. The correlation tends to be greater in the upper part of the water column and to decrease systematically with depth. This is because the successive EOFs are constructed so as to explain the maximum amount of vertically coherent residual variance, and this is much larger in the upper ocean than at depth. From the earlier discussion about profile "noise," it is clear that the relatively poor correlations below about 1500 m when seven EOFs are used in the fit (last frame) are entirely due to noise, representing no more than 1% of the total variance (Table 3). Nevertheless, the small correlations at depth indicate that only a small part of the variance at these depths is actually accounted for by the first seven EOFs. It is interesting that the way in which the individual modes contribute to the complete profile is different in summer than in winter. For example, the second and third EOFs make important contributions to the correlations at all depths in winter. However, in summer, the second EOF seems to be important only in the mixed layer, while the third EOF is very significant in the depth range of 300–1000 m.

The central result of this study, the skill of estimating density profiles from data in the upper ocean alone, is presented in Fig. 7. For clarity, we show the results in

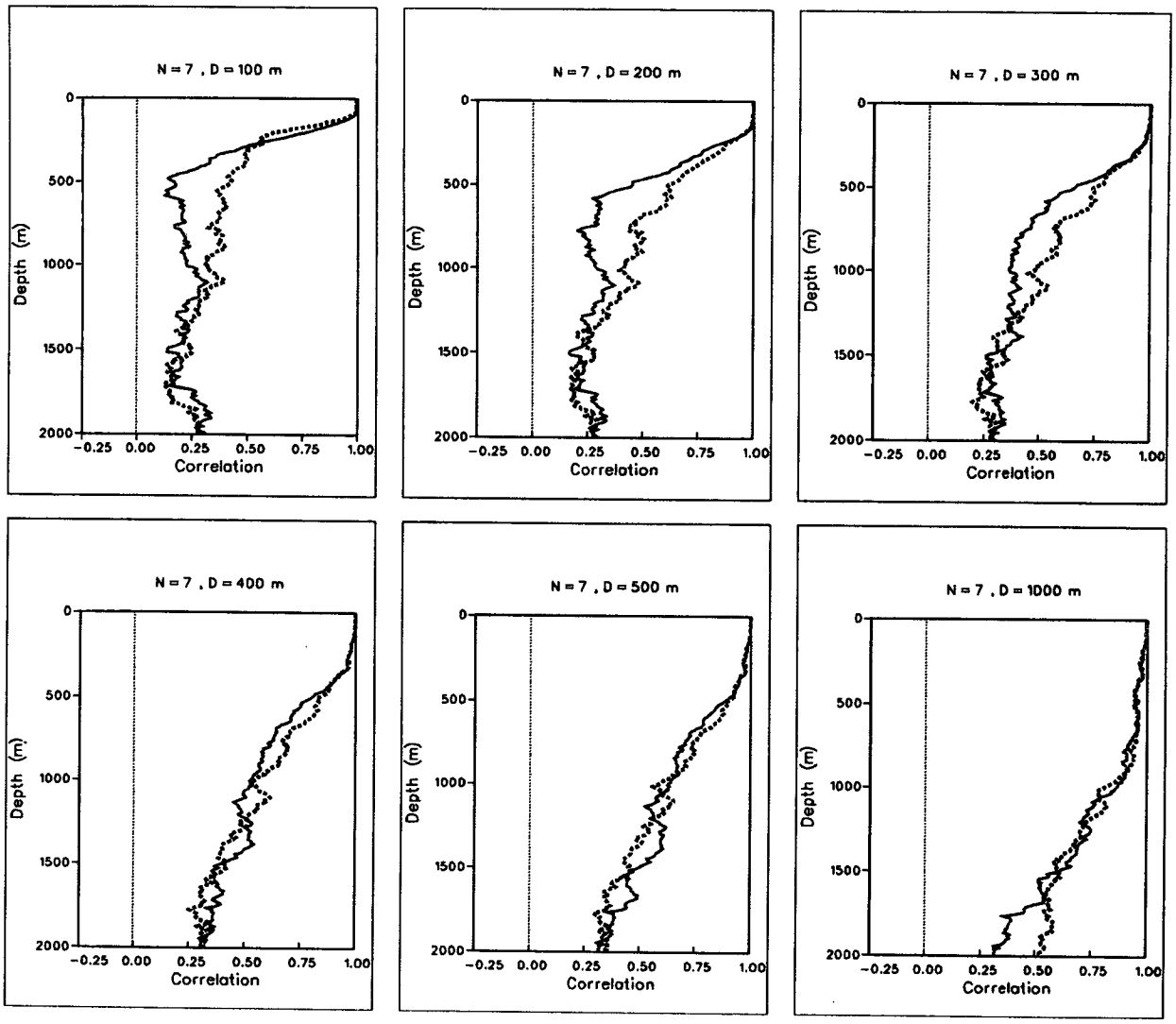

FIG. 7. Correlation r between the estimated and the observed profiles using the first seven EOFs and data above D for D = 100, 200, 300, 400, 500, and 1000 m. Profiles of r are shown for both the summer half-year (April–October, ND = 64, solid) and the winter half-year (November–March, ND = 62, dashed).

six separate frames, each frame corresponding to a different value of the depth D used to define the "upper ocean." The successive frames show how well the estimated profiles match the observed profiles, as measured by the correlation r, as the data-EOF fitting is performed over an increasing depth of the upper ocean. In general, the correlation decreases rather rapidly below D. Furthermore, the decrease with depth is more rapid in summer than in winter. In addition, the correlations seem to depend more strongly on D in summer than in winter. The difference between the correlation curves for D = 200 and D = 300 is especially notable in summer. These results are operationally significant because shallow casts to 200 or 300 m are typical of the SeaSoar CTD system. The dropoff in estimation skill below D (for $D \leq 500$ m say) is due in part to the relatively shallow nature of the major baroclinic features represented by the first seven EOFs in this region of the California Current.

We now examine the effect of observational "noise" in the CTD profiles. As a result of the analysis presented above, density fluctuations accounted for by EOF modes higher than seven are considered indistinguishable from vertically incoherent "noise." Since no vertical extension method can be expected to explain such noise, it is of interest to know how well the present EOF method estimates the "noise-free" part of the density profile, that is, the "signal." To address this issue, we present in Figs. 8 and 9 the same information as in Figs. 6 and 7, except that here the correlation is

computed between the estimated profiles and noise-free profiles obtained by fitting the observed profiles (to 2000 m) to the first seven EOFs. Figure 8 shows how the fit (to 2000 m) to the seven EOF profiles, as measured by r, improves as more modes are used. The fit is nearly exact when six modes are used (last frame). The ways in which the various modes contribute to the filtered profiles are different in summer than in winter, as seen before in Fig. 6.

The success with which the EOF method can estimate the seven EOF profiles using data entirely above D is shown in Fig. 9. As with estimating the observed profiles, the correlations also decrease below D, but the decrease is not nearly so rapid for $D \geq 400$ m. In fact, the results show that observations down to 500 m can be used to estimate the seven EOF profiles below 500 m quite accurately. For shallower casts, that is, $D \leq 300$ m, there is also a noticeable improvement compared to estimating the observed profiles, but this occurs mostly below 1000 m. The improved correlations in the deeper ocean are entirely due to the removal of what we consider to be noise.

To evaluate the potential effect of the vertical extension on the computation of surface geostrophic currents, we show in Tables 4 and 5 a comparison of the surface dynamic height d computed from the observed and estimated profiles. Although the correlation between the observed and estimated dynamic heights relative to 500 m (Table 4) is quite high for all values of D, the root-mean-square differences are a significant

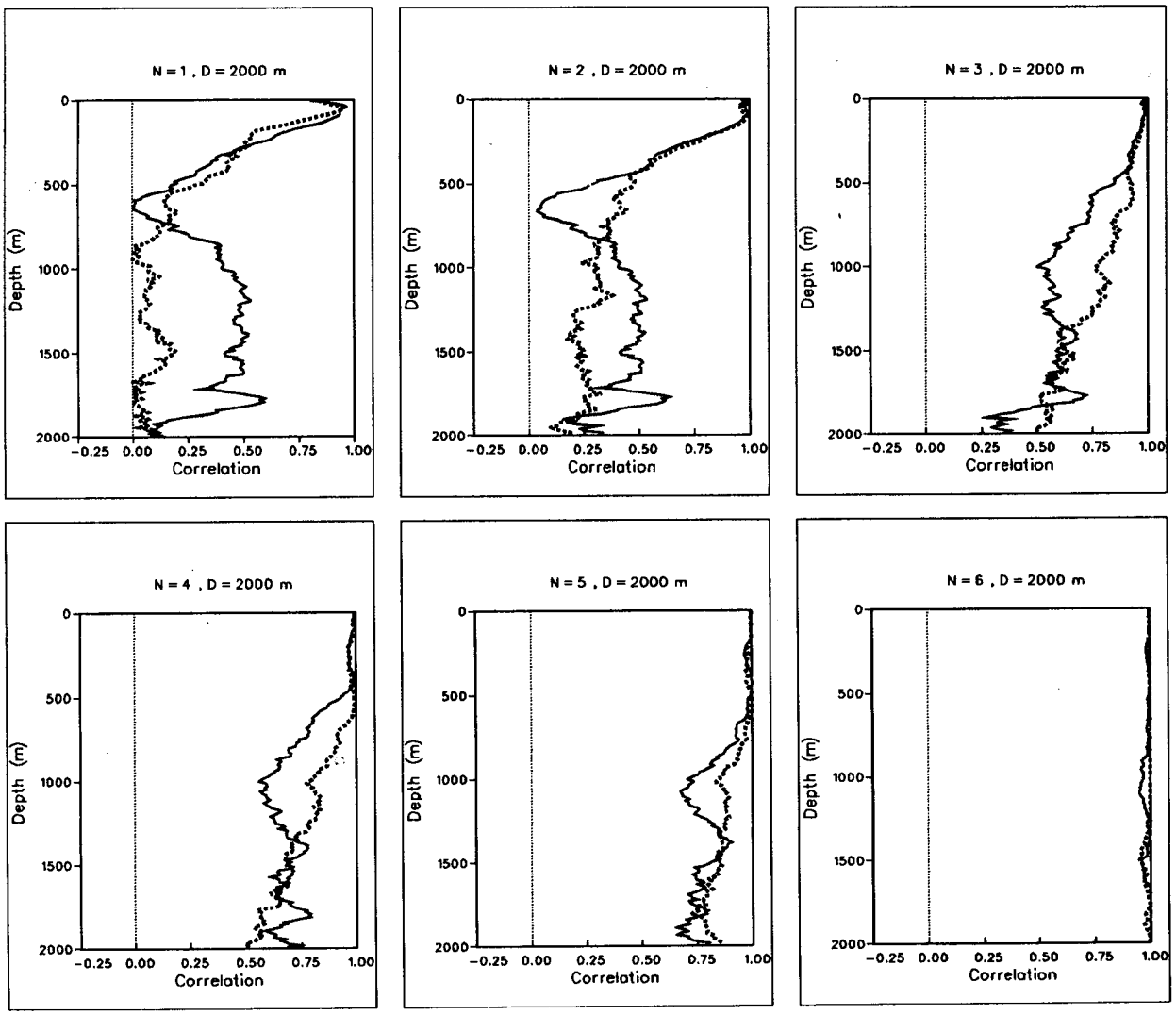

FIG. 8. Correlation r between the estimated and noise-free (seven EOF) profiles when the fitting is performed to 2000 m and N EOFs are used. Profiles of r are shown for N = 1, 2, 3, 4, 5, and 6 for both the summer half-year (solid, ND = 64) and the winter half-year (dashed, ND = 62).

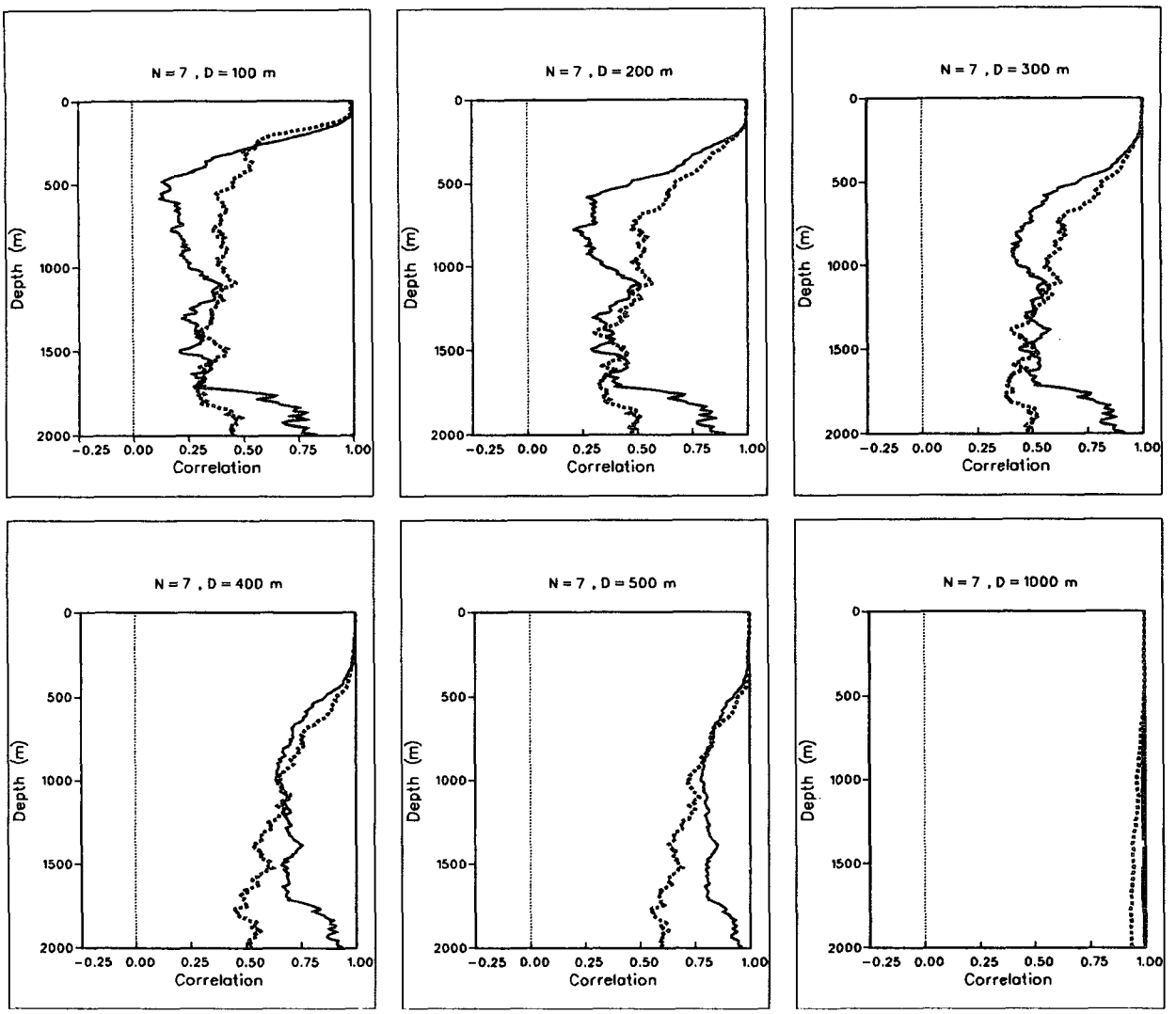

FIG. 9. Correlation r between the seven EOF profiles and the estimated profiles using the first seven EOFs and data above D for D = 100, 200, 300, 400, 500, and 1000 m. Profiles of r are shown for both the summer half-year (solid, ND = 64) and the winter half-year (dashed, ND = 62).

fraction of the observed standard deviation unless $D \geqslant 300$ m. This is true for both seasons of the year and it indicates that 300 m of data are required for the vertical extensions to be useful for this purpose. The results for the dynamic height of the sea surface relative to 1000 m (Table 5), a more reliable indicator of the surface geostrophic currents in this region, are not as good because of the decreased skill of estimating the density structure below 500 m.

To partly test the robustness of the vertical extension procedure, we applied the method to a small but independent set of hydrographic profiles taken in the area of the Gulf of the Farallones, which is located approximately 100 km north of the POST area (Jessen et al. 1992). Many of the CTDs during these Gulf of the Farallones cruises were taken in shallow water, but 16 profiles from the winter half-year and 34 profiles from the summer half-year reached to 2000-m depth and also passed the same kind of screening for mutual independence that was used with the POST data. In this test, the independent profiles were extended below the depth D using the mean and first seven EOFs that were computed from the dependent (POST) data (Figs. 2 and 3). The results are shown in Fig. 10. In general, the vertical extensions to depths less than 1000 m are about as successful as they were with the dependent data (Fig. 7). However, a significant degradation in the profile correlations (relative to the results from the dependent data) occurs below 1000 m, where the correlations are negative during the summer half-year.

TABLE 4. Statistical comparison of the dynamic height of the sea surface relative to 500 m computed from the observed profiles d_o and from the estimated profiles d_e. Shown are the means ($\bar{}$), the standard deviations σ, the root-mean-square difference (rmsd), and the correlation r as a function of D for both the winter and summer.

D (m)	$\bar{d}_e$ (cm)	σ_{d_e} (cm)	rmsd (cm)	$r(d_o, d_e)$
Winter half-year (n = 62, $\bar{d}_o$ = −0.11 cm, σ_{d_o} = 79 cm)				
100	−0.04	75	28	0.93
200	−0.06	79	12	0.99
300	−0.08	79	8	1.00
400	−0.09	79	5	1.00
500	−0.10	79	3	1.00
Summer half-year (n = 64, $\bar{d}_o$ = − 0.10 cm, σ_{d_o} = 100 cm)				
100	−0.03	97	28	0.96
200	−0.06	99	15	0.99
300	−0.08	100	9	1.00
400	−0.09	100	5	1.00
500	−0.09	100	4	1.00

Examination of the individual Gulf of the Farallones profiles as in Fig. 11 (not shown) reveals distinct density fluctuations with vertical scales of several hundred meters at depth that are clearly *not* associated with any fluctuations in the upper ocean. No vertical extension method can be expected to estimate such fluctuations from upper ocean data alone. The results of the surface dynamic height computations from the independent data (not shown) are quite similar to those from the dependent data when the reference level is 500 m, and they are only marginally less successful when the reference level is 1000 m.

To complement the above statistical information and to show some performance characteristics of the estimation method, we present several direct comparisons between observed and estimated profiles in Fig. 11. Many more such examples are shown in Haney et al. (1994). The top left frame in each figure shows the observed and seven-EOF-based profile. The other three frames show the seven-EOF-based profile and estimated profile for D = 200, 300, and 500 m, respectively. The character of the noise in the observed profiles is seen by comparing the observed and seven-EOF profiles in the top left frames. As defined here, "noise" is simply density fluctuations not accounted for by the first seven EOFs. One of the profiles is seen to have very little noise (Fig. 11d); some have noise in especially short vertical scales (Figs. 11b,e); and other profiles have noise in somewhat larger vertical scales (Figs. 11a,c). The way in which the estimation improves as D increases from 200 to 500 m is seen by comparing the profiles in the other three frames. Occasionally, an entire profile can be estimated quite accurately from only 200 m of data (Fig. 11d), while in most cases, especially in summer, 200 m of data is clearly inadequate.

Another point worth noting from the results in Fig. 11 is the way in which the amplitude of the different modes that make up the estimated profile changes as the depth D of data used in the fitting procedure is changed. In Fig. 11a, for example, the amplitude of the first mode is almost the same in all four frames, indicating that it is generally well established by data down to only 200 m. This characteristic of the first mode also holds at the other locations shown in Fig. 11. It also holds to a lesser extent for the second mode but not for the third and higher modes. This result indicates that the higher modes, unlike the lowest two modes, are not well established by shallow data alone. The lowest modes can be rather well established by fitting over a shallow depth (200 m) because they have their strongest characteristic signal in the upper 200–300 m. On the other hand, the higher modes ($n \geq 3$) have a greater part of their signal below that depth (Fig. 3). When these higher modes are active in a given profile, deeper data are needed to estimate the given profile accurately.

In most cases the estimated profile using 500 m of data matches the seven-EOF profile quite well. This is reflected in the rather high correlation between the seven-EOF profiles and the estimated profiles with D = 500 m shown in the bottom middle frame in Fig. 9. Occasionally, however, even with 500 m of data, the estimation appears to miss some potentially important features (Fig. 11a). Of course, the extent to which an estimate of the seven-EOF profile can be considered satisfactory depends on the use to which the estimated profile will be put. The above results indicate that the method can estimate noise-free (seven EOF) profiles using data from the upper 500 m that would be satis-

TABLE 5. Dynamic height computations as in Table 4 except the reference level is 1000 m as described in the text.

D (m)	$\bar{d}_e$ (cm)	σ_{d_e} (cm)	rmsd (cm)	$r(d_o, d_e)$
Winter half-year (n = 62, $\bar{d}_o$ = −0.21 cm, σ_{d_o} = 88 cm)				
100	−0.04	78	42	0.88
200	−0.08	86	27	0.95
300	−0.11	87	20	0.97
400	−0.13	87	15	0.99
500	−0.15	88	12	0.99
1000	−0.19	88	3	1.00
Summer half-year (n = 64, $\bar{d}_o$ = −0.20 cm, σ_{d_o} = 102 cm)				
100	−0.03	97	43	0.91
200	−0.07	101	33	0.95
300	−0.10	101	23	0.97
400	−0.12	101	16	0.99
500	−0.14	101	12	0.99
1000	−0.19	102	3	1.00

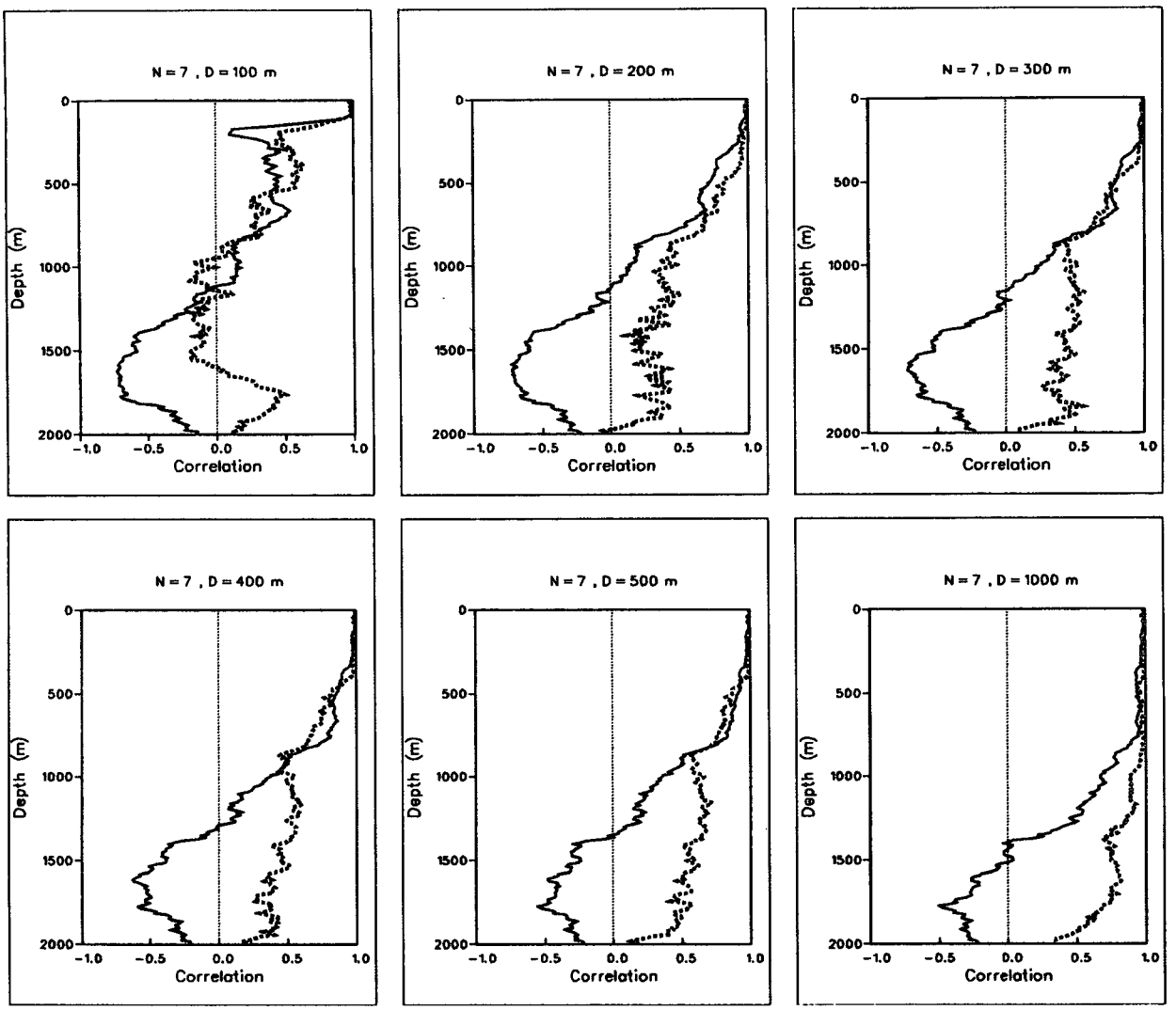

FIG. 10. Same as in Fig. 7, except using the independent data as described in the text. Note that the abscissa is different from that in Fig. 7, and the sample size is 16 in summer and 34 in winter.

factory for initializing operational numerical models designed to nowcast and forecast the oceanic synoptic scale. The method might also be useful for extending historical CTD profiles (e.g., the CALCOFI data) to greater depth.

5. Summary and conclusions

A method for extending upper ocean density profiles to the deeper ocean has been developed and tested using a large number of deep CTD stations off Point Sur, California. The method involves fitting a shallow density profile to the first N full column EOFs determined from historical data. Here N is the number of EOFs that are needed to represent the vertically coherent "signal" part of the observed profiles efficiently. Our analysis, following the ideas of Preisendorfer et al. (1981) and Smith et al. (1985), resulted in $N = 7$ for our data. The EOFs found in this study are quite similar to those described by Chelton (1980), Rienecker et al. (1987), and Bray and Greengrove (1993) for the nearby California Current region. The first mode is due to fluctuations in the baroclinic structure associated with the equatorward vertical shear of the California Current in the main pycnocline, while the second mode appears to represent a response in the mixed layer and the pycnocline to variations in alongshore wind stress. The third and fourth modes have deeper structures. All of these modes make important contributions to the success of the vertical extension method.

When the vertical extension method is tested against the observed profiles, the results are moderately suc-

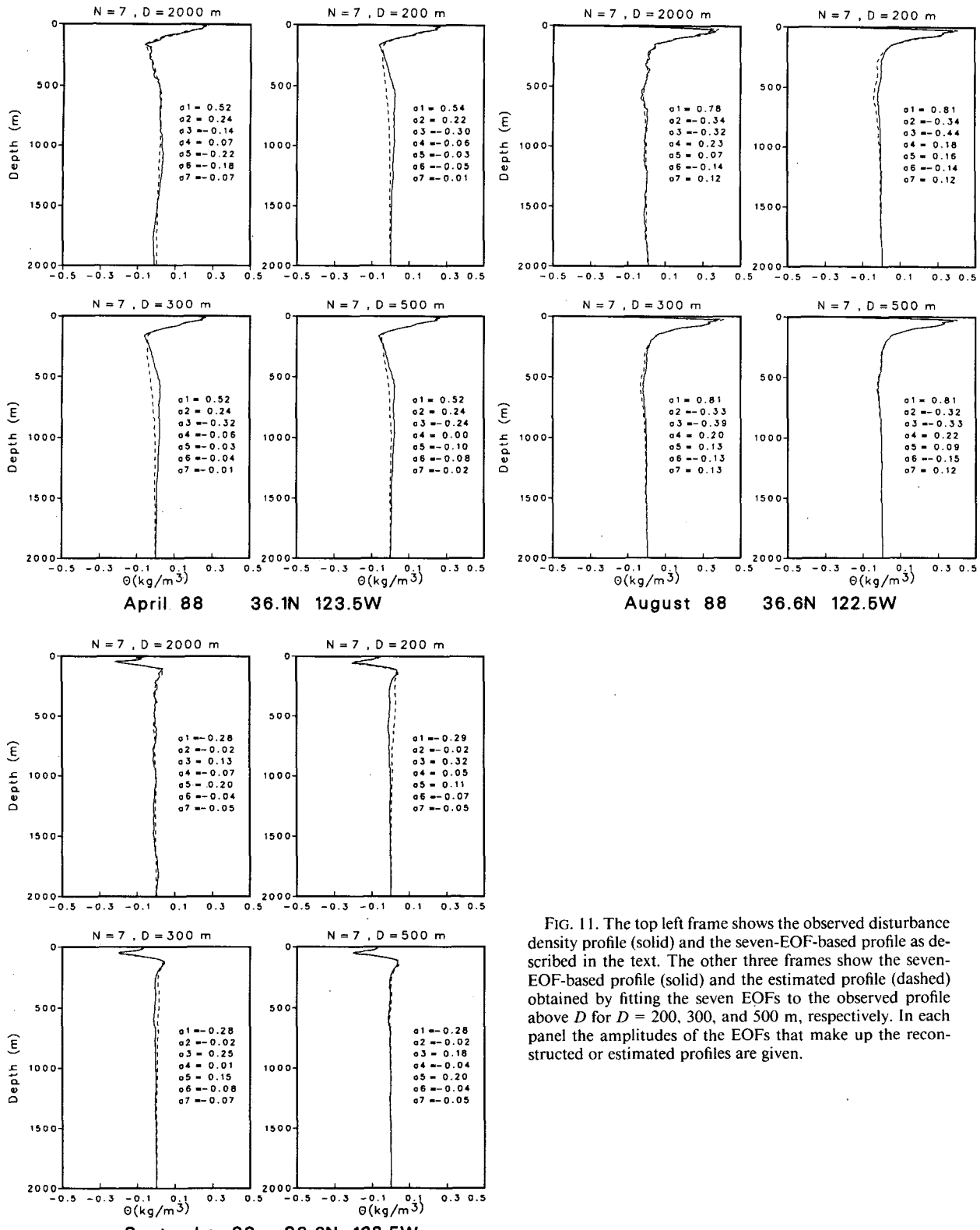

FIG. 11. The top left frame shows the observed disturbance density profile (solid) and the seven-EOF-based profile as described in the text. The other three frames show the seven-EOF-based profile (solid) and the estimated profile (dashed) obtained by fitting the seven EOFs to the observed profile above D for D = 200, 300, and 500 m, respectively. In each panel the amplitudes of the EOFs that make up the reconstructed or estimated profiles are given.

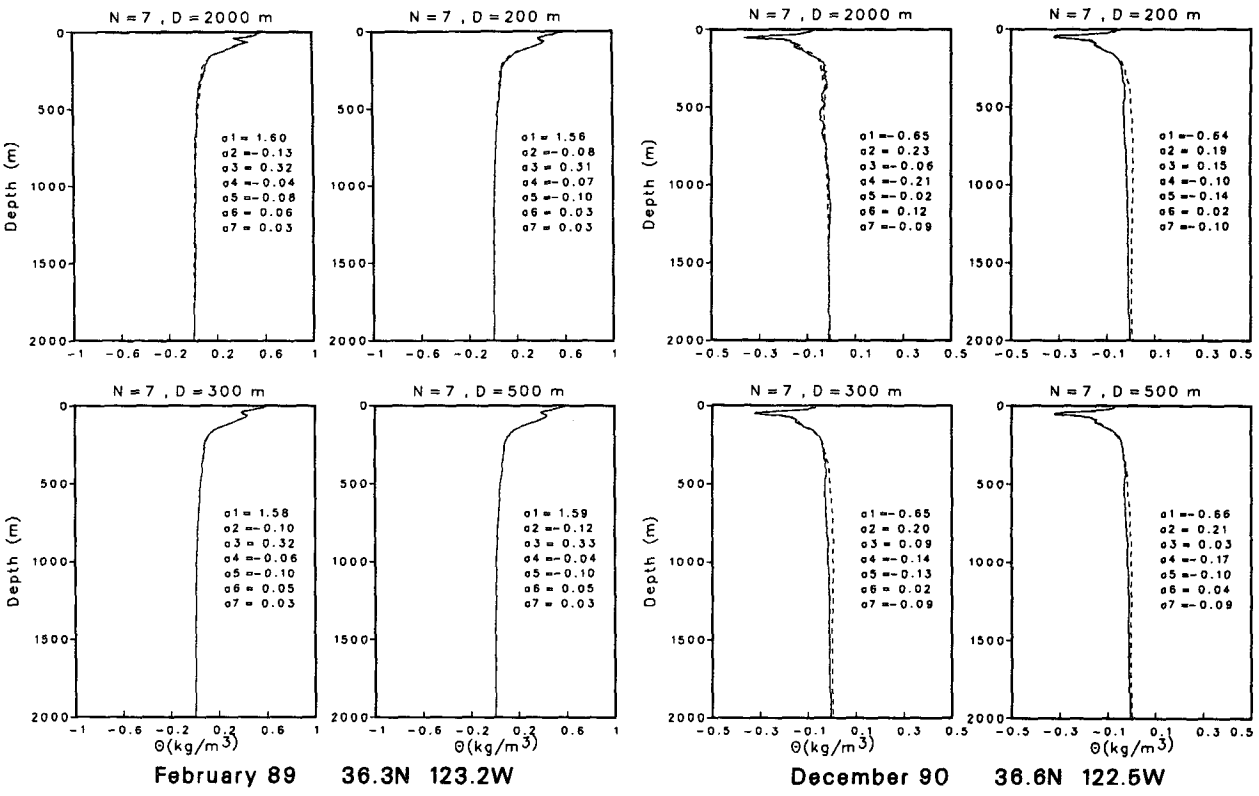

FIG. 11. (*Continued*)

cessful (Fig. 7). Profiles to depths shallower than 500 m can be extended to 500 m with an over all correlation and skill that depends on the time of year as well as the depth of the profiles. In all cases the estimated profiles correlate better with noise-free (seven-EOF) profiles, because the EOF-based method cannot estimate variability associated with vertically incoherent noise. The improved correlations due to removing such noise is significant at all depths below D in winter but only below about 1000 m in summer. The method was also shown to be useful for estimating the dynamic height of the sea surface relative to 500 m from observations extending down to only 300 m, the kind of observations that are common with SeaSoar. The dynamic height relative to 1000 m can also be estimated quite well but it requires observations down to at least 500 m. These results were largely confirmed with a set of independent profiles from the Gulf of the Farallones but with less success below 1000 m in summer.

The above results raise important questions about the generality of the vertical extension method and whether it would be successful in other geographical regions. The success of the method rests on its ability to detect and differentiate between the EOF modes in the region of interest from only a limited amount of upper ocean data. As such, its success depends on the distinctive shapes of the different modes, on the extent to which they are uncorrelated (orthogonal) over the shallower depths, and most importantly on whether the modes have significant signals at depth. Its success also depends on the validity of the assumption, inherent in the sequential fitting procedure, that upper ocean density fluctuations that are equally attributable to several different modes (due to the nonorthogonality of the EOFs over shallow depths) are actually due to the lowest of those modes. On the basis of these considerations, we expect the method would be equally successful in all the dynamically active parts of the World Ocean where the internal density fluctuations are associated with moderate to deep baroclinic structures similar to those that exist in the California Current off Point Sur.

Acknowledgments. We are grateful to Dudley Chelton, Michele Rienecker, Jerry Smith, and Leonard Walstad, as well as our NPS colleagues J.-M. Chen and Pat Harr, for especially valuable discussions and input on the topic of this paper. This work was sponsored by the Applied Oceanography Division (Code 124) of the Office of Naval Research and the Naval Postgraduate School, whose support is gratefully acknowledged. Ms. Penny Jones typed and revised this paper.

REFERENCES

Bray, N. A., and C. L. Greengrove, 1993: Circulation over the shelf and slope off Northern California. *J. Geophys. Res.*, **98,** 18 119–18 145.

Chelton, D. B., Jr., 1980: Low frequency sea level variability along the west coast of North America. Ph.D. dissertation, Scripps Institution of Oceanography, La Jolla, CA, 212 pp.

Haney, R. L., R. A. Hale, and C. A. Collins, 1994: Estimating subpycnocline density fluctuations in the California coastal region from upper ocean observations. Tech. Rep. NPS-MR-94-001, Naval Postgraduate School, Monterey, CA, 71 pp.

Hurlburt, H. E., D. N. Fox, and E. J. Metzger, 1990: Statistical inference of weakly-correlated subthermocline fields from satellite altimeter data. *J. Geophys. Res.*, **95**(C7), 11 375–11 409.

Jessen, P. F., S. R. Ramp, C. A. Collins, N. Garfield, L. K. Rosenfeld, and F. B. Schwing, 1992: Hydrographic and acoustic Doppler current profiler (ADCP) data from the Farallones shelf and slope study, 13–18 Feb 1991; 16–21 May 1991; 12–18 Aug 1991; 29 Oct–3 Nov 1991; 7–17 Feb 1992. Tech. Rept. Nos. NPS/OC92003-NPS/OC92007, Naval Postgraduate School, Monterey, CA, 846 pp.

Lentz, S. J., 1987: A description of the 1981 and 1982 spring transitions over the Northern California shelf. *J. Geophys. Res.*, **92**(C2), 1545–1567.

Mellor, G. L., and T. Ezer, 1991: A Gulf Stream model and an altimetry assimilation scheme. *J. Geophys. Res.*, **96**(C5), 8779–8795.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110,** 699–706.

Preisendorfer, R. W., F. W. Zweirs, and T. P. Barnett, 1981: Foundations of principal component selection rules. SIO Ref. Series 81-4, Scripps Institution of Oceanography, 192 pp.

Rienecker, M. M., C. N. K. Mooers, and A. R. Robinson, 1987: Dynamical interpolation and forecast of the evolution of mesoscale features off Northern California. *J. Phys. Oceanogr.*, **17,** 1189–1213.

Robinson, A. R., M. A. Spall, and N. Pinardi, 1988: Gulf Stream simulations and the dynamics of ring and meander processes. *J. Phys. Oceanogr.*, **18,** 1811–1853.

Smith, J. A., 1984: Empirical and dynamic modes in the CCS. NPS Tech. Rep. NPS 68-84-003, Naval Postgraduate School, Monterey, CA, 42 pp.

——, C. N. K. Mooers, and A. R. Robinson, 1985: Estimation of quasi-geostrophic modal amplitudes from XBT/CTD survey data. *J. Atmos. Oceanic Technol.*, **2,** 491–507.

Strub, P. T., J. S. Allen, A. Huyer, and R. L. Smith, 1987: Large scale structure of the spring transition in the coastal ocean off western North America. *J. Geophys. Res.*, **92**(C2), 1527–1544.

Tisch, T. D., S. R. Ramp, and C. A. Collins, 1992: Observations of the geostrophic current and water mass characteristics off Point Sur, California, from May 1988 through November 1989. *J. Geophys. Res.*, **97**(C8), 12 535–12 555.