

## Use of the Temperature–Salinity Relation in a Data Assimilation Context

ALBERTO TROCCOLI AND KEITH HAINES

*Department of Meteorology, University of Edinburgh, Edinburgh, United Kingdom*

(Manuscript received 4 September 1998, in final form 12 February 1999)

### ABSTRACT

A data analysis using conductivity–temperature–depth (CTD) measurements in the western tropical Pacific is carried out to get an estimate of the timescale over which temperature–salinity ( $T$ – $S$ ) relationships are preserved. Results show that the  $T$ – $S$  preservation holds for periods on the order of a few weeks.

A new method for assimilating upper-ocean temperature profiles with salinity adjustments into numerical ocean models is then proposed. The approach would use a  $T$ – $S$  relation that is more local in space and time than is the climatological  $T$ – $S$  relation used in previous studies. The assimilation method avoids convective instability as the temperature data are introduced.

The CTD data and instantaneous fields from an ocean model are used to test the assimilation method by combining one profile with another. These tests recover the salinity profiles and the 0–500-m dynamic height very well (differences are smaller than 1 dyn cm). By contrast, analyses that used a climatological  $T$ – $S$  relation did not provide a good salinity profile or dynamic height (errors were greater than 3 dyn cm).

If used for data assimilation, the method would allow the recovery of a good salinity and density field when only temperature data were available, at intervals of, say, two to four weeks. There is evidence that the same conclusions could be drawn for many other oceanic areas.

### 1. Introduction

Ocean data assimilation has been investigated continuously over the last decade [reviews are provided by Ghil and Malanotte-Rizzoli (1991), Anderson et al. (1996), and Malanotte-Rizzoli and Tziperman (1996)]. Most research has focused on either satellite altimeter data or near-surface temperatures, as measured by either expendable bathythermographs (XBT) or moored buoys. Recently it has been realized that salinity also plays an important role in determining density and circulation not only at high latitudes [e.g., Reverdin et al. (1997)] but also in the Tropics [e.g., Donguy (1994)]. However, little attention has been paid to updating salinity  $S$  in ocean circulation models used for data assimilation. This lack of attention is due mainly to the paucity of salinity data. As a result, the easiest solution usually is adopted, namely, salinity is left unmodified by data assimilation [i.e., the a priori  $S(z)$  profile is preserved, where  $z$  is depth]. This solution, however, may lead to increased errors in the velocity field (Cooper 1988). It is the main objective of this paper to present a simple approach to including salinity field adjustment,

in a data assimilation context, when only temperature measurements are available.

Density  $\rho$  at any given depth is a function of the independent variables  $S$  and temperature  $T$ . However, if, for instance,  $S$  could be expressed as a single-valued function of  $T$ , then the variables for  $\rho$  would be reduced to only one,  $\rho = \rho[T, S(T)]$ , and  $S$  would not need to be known directly. This formulation would be an enormous simplification of the problem, for it would imply that, at a fixed depth,  $\rho$  can be derived from knowing  $T$  only. In particular, this solution is possible in circumstances for which the  $T$ – $S$  relation does not vary much or varies only slowly in time. Of course, this case is not always true, but there is observational evidence that it applies to many parts of the ocean (Emery and Dewar 1982). This fact usually is referred to as  $T$ – $S$  preservation (or conservation) and stems from the fact that the  $T$ – $S$  water mass properties often are conserved for a long time. The  $T$ – $S$  preservation hypothesis is the core assumption adopted in the present work.

It is clear that, for the method to work, a good  $T$ – $S$  relation is required, from which  $S$  can be derived, given  $T$ . Since the Lagrangian  $T$ – $S$  properties undoubtedly are well preserved because of the weak mixing in most parts of the ocean, it would at first sight be plausible to use a long-term mean  $T$ – $S$  relation based on some climatological dataset, and this solution is the most widely adopted one. Hitherto, its principal application has been for evaluation of reliable dynamic heights (Stommel

---

*Corresponding author address:* Dr. Alberto Troccoli, Dept. of Meteorology, University of Edinburgh, The King's Bldgs., Mayfield Rd., Edinburgh EH9 3JZ United Kingdom.  
E-mail: alberto@met.ed.ac.uk

1947; Emery 1975; Emery and Dewar 1982; Kessler and Taft 1987). Dynamic height is quite important in oceanography, for it allows an estimate of surface currents, relative to some reference level. Emery (1975), for instance, computed the dynamic height in three North Pacific areas and found that, for two of them, satisfactory results could be obtained by using a mean  $T$ - $S$  relationship (averaged over a 21-yr period), along with the measured temperature profiles. At the third location, however, a thermal inversion led to large salinity fluctuations in the  $T$ - $S$  curves, thus compromising the dynamic height calculation. In fact, temperature inversions do not allow a unique definition of salinity as a function of  $T$ . A solution to this problem is presented in this work. Furthermore, abundant altimeter datasets are now available for validation of dynamic height estimates.

Although dynamic height is a useful quantity, it represents only the vertical integral of the density. Thus, a precise dynamic height estimation does not ensure that the derived density, using  $T$ - $S$  preservation, is very close to the true density, because compensations in the vertical can be present (e.g., see Fig. 5 in Kessler and Taft 1987). This complication may not matter if the main focus is on surface geostrophic currents; however, it becomes an issue when more complete descriptions of the internal dynamics are required.

When initializing an ocean model, for example, a realistic vertical density structure is more vital than is a good initial velocity field (Gill 1982; Philander et al. 1987; Moore et al. 1987) because the potential energy is greater than the kinetic energy. Since the density is dominated, especially in low latitudes, by  $T$  as opposed to  $S$ , the latter was not included as a prognostic variable in tropical ocean models until the late 1980s (Cooper 1988). However, Cooper (1988) showed that, even using a tropical ocean model, it is essential to include  $S$  along with  $T$  if the goal is to reduce rms errors in the velocity field, and that assimilating only  $T$  can be deleterious for the dynamics described by ocean models (Cooper 1988; Woodgate 1997). However, from a practical viewpoint, salinity data simply are not available in most data assimilation situations so a means of deriving salinity changes in terms of temperature is required.

Using a climatological  $T$ - $S$  relation is the most obvious choice but will restrict considerably the variability in  $S(T)$ . Consider, for example, the tropical Pacific, which exhibits large interannual variations in water properties because of El Niño. Recent work by Vossepoel et al. (1999) uses a CTD dataset for the tropical Pacific to tackle this issue. They found that a climatological  $T$ - $S$  relation represents a good background (or first guess) field [they used Levitus et al. (1994) and Levitus and Boyer (1994), together referred to hereafter as “Levitus”] but that corrections to the salinity field are necessary to get a reliable density field. Here we consider using a different background  $T$ - $S$  relation rather than climatological data, and thereby we hope to

dispense with the need for additional corrections. The approach differs from previous work because the salinity analysis is purely local, with the background  $T$ - $S$  properties being taken either from those most recently available from CTD data or from a model prediction of  $T$ - $S$  in the context of sequential data assimilation.

By avoiding use of long-term averages (e.g., climatological data), larger density variations for any given  $T$  profile are permitted. However, in a modeling context, unlike the case in which a climatological  $T$ - $S$  relation is used, this approach could permit large and possibly detrimental salinity drifts. A solution could be to prevent the  $T$ - $S$  relation from drifting too far from the climatological one by using some weighted  $S(T)$  between the model prediction and the climatological data. An advantage of using the model  $S(T)$  alone, however, is that, given its local implementation, there is a considerable reduction of the computational costs. In this sense, this method is comparable to the “suboptimal” technique suggested by Cooper and Haines (1996) for altimeter data.

The work is organized as follows. In section 2, the CTD dataset is introduced. In section 3,  $T$  and  $S$  variances and covariations in the tropical Pacific are analyzed to obtain an estimate of the timescale over which the  $T$ - $S$  relation is preserved. The presentation of the assimilation technique is in section 4. Results obtained by the assimilation of  $T$  profiles, taken both from CTDs and from model realizations of the Hamburg Ocean Primitive Equation (HOPE) model, are shown in section 5. The comparison with “persistence” [i.e., preserving  $S = S(z)$ ] and “climatology” (using long-term  $T$ - $S$  relationships) based estimates of  $S$  and the density will be discussed here. In section 6, the method is discussed and comments on its possible applications are made.

## 2. Data

The CTD data used in this study were obtained from the database maintained at the U.S. National Oceanographic Data Center (NODC) and collected as part of major campaigns. The data are located in a limited area of the western tropical Pacific between 4.5°N–4.5°S and 154°–159°E and they refer to periods between November 1992 and February 1993. The total number of CTD casts is 1252. All the casts reach a depth greater than 300 m, and 76% of them reach more than 500 m. The vertical resolution is more than 1 datum per 10 m for each profile. A map with the geographical distribution of the CTDs is shown in Fig. 1, from which it can be seen that the data are concentrated into three subregions: one north of the equator at 3.6°–4.4°N, 155.5°–157°E (named N) and the other two south of the equator at 2.4°–1.6°S, 154.2°–156.2°E (named SW) and 2.7°–1.9°S, 157.5°–159°E (named SE). At a single location denoted S1, 104 CTD casts have been collected within a 10-day period in November 1992. The four locations

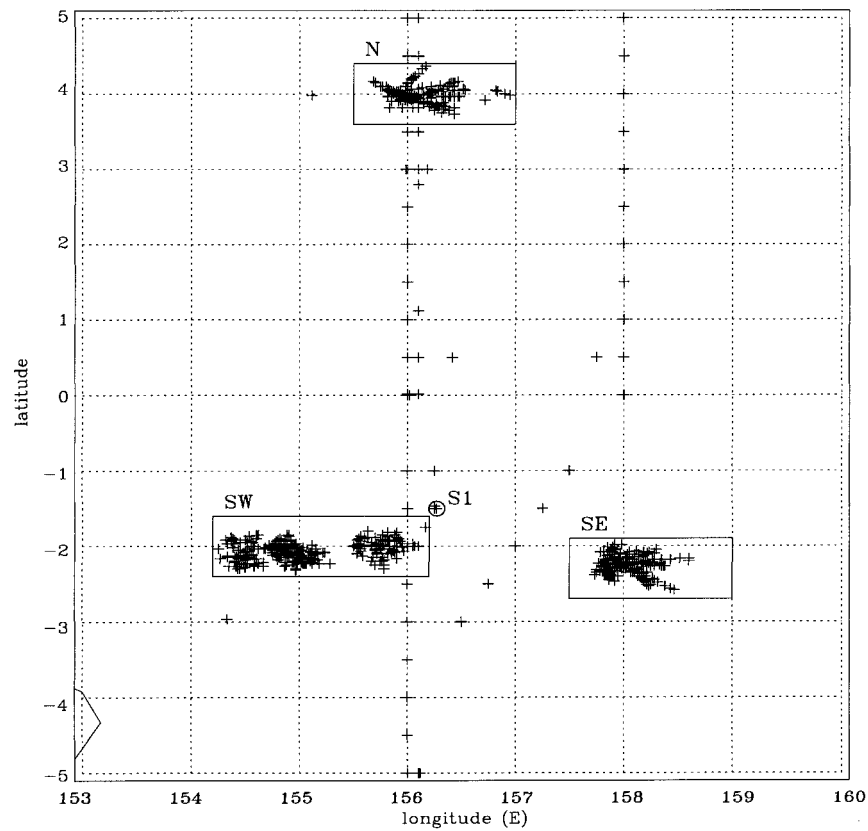


FIG. 1. Spatial coverage of the CTD data. The three boxes (N, SW, and SE) and the single location (S1) have been highlighted.

(N, SE, SW, and S1) are hereinafter treated as individual grid locations.

There are several reasons for choosing this region for study of  $T$ - $S$  variability. The most important are the following: (a) in this period an extensive observational campaign was commissioned for which high quality data were gathered (e.g. Anderson 1995); (b) the western tropical Pacific plays a very important role in the onset of El Niño (Philander 1990); (c) due to high precipitation, particularly south of the equator, important upper water column salinity features, such as the “barrier layer” (Lukas and Lindstrom 1991; Roemmich et al. 1994; Sprintall and McPhaden 1994; Ando and McPhaden 1997; Vialard and Delecluse 1998a,b) often appear (when the mixed layer does not coincide with the isothermal layer); (d) several previous studies have analyzed the western tropical Pacific, so that these can be used for reference and comparison (Emery and Dewar 1982; Delcroix et al. 1987; Donguy 1994; Vossepoel et al. 1999).

### 3. Evaluation of $T$ - $S$ functions for deriving salinity profiles

Two approaches to deriving salinity do not require use of a  $T$ - $S$  relation. First, climatological salinity  $S_{cl}(z)$

might be assumed constant [e.g., salinity from Levitus et al. (1994)], which would remove any active role for salinity in current variations. Alternatively,  $S(z)$  from a numerical model being run with temperature profile assimilation could be left unaltered whenever new  $T$  data are available. This approach commonly is done because it is simple but leads to the problems already outlined.

If a  $T$ - $S$  relation is used, the  $S_{cl}(T)$  always has been used in previous works. This approach probably is acceptable when nothing better can be obtained because it is the best estimate of the long-term average [although care should be taken because most climatological datasets are not constructed with the aim of accurately representing a mean  $T$ - $S$  relation, as pointed out by Lozier et al. (1995)]. However, in a data assimilation context with  $T$ -profile data being assimilated every week or so, it may be better to use a  $T$ - $S$  preservation assumption, that is, to use the local  $T$ - $S$  relation predicted by the numerical model. By keeping  $S(T)$  rather than  $S(z)$  fixed at assimilation time, we are assuming implicitly that a model is able to predict  $S(T)$  better than  $S(z)$ . It will be demonstrated that this assumption is reasonable by studying the observed variability of salinity profiles on differing timescales, in both data and a numerical model.

The approach involves statistically analyzing the data

in the three selected boxes (N, SW, and SE) plus those in the single location S1 (1.5°S, 156.25°E), at various frequencies from daily (for S1) to weekly and seasonally (for N, SW, and SE). For the daily and weekly periods, the analysis follows a similar procedure. The variability of the single profiles [i.e.,  $T(z)$ ,  $S(z)$ , and  $S(T)$ ] within the period considered is studied, focusing attention on the  $S$  variability and its correlations with  $T$ . The rms variations for  $S(z)$  and  $S(T)$  are compared numerically on short timescales, as discussed below. On the seasonal timescale, the discussion essentially is based on the graphical comparison between  $S = S(T)$  and  $S = S_{cl}(T)$  profiles, with  $T = T(z)$  and  $S = S(z)$  also shown.

*a. High-frequency analysis of 10-day period:  $S(z)$  versus  $S(T)$*

For this analysis, 104 CTD casts over a 10-day period in November 1992 at the same location (S1) are examined. Figures 2a–c show the variability of  $T(z)$ ,  $S(z)$ , and the  $T$ – $S$  relation.

In  $z$  space there is a pronounced  $T$  and  $S$  variability (Figs. 2a, b), which, if interpreted as vertical displacements, imply displacements as large as 40 m in the top 300 m, presumably caused by internal waves. Visually there appears to be less variability in the  $T$ – $S$  relation in Fig. 2c, but it is difficult to make numerical comparisons by eye. Therefore, Fig. 2d provides a more consistent way of directly comparing the  $S(z)$  and the  $S(T)$  variability. Rms salinity variance is shown as a function of  $z$  space (solid line). The dashed line shows rms  $S(T)$  variance calculated from the mean  $S(T)$  relation but with the variances transformed back into  $z$  space using the mean  $T(z)$  profile for the period. The tight  $S(T)$  curves (Fig. 2c) do express the fact that on the daily timescale there is little mixing of water masses and little influence of horizontal advection, which is consistent with much of the variability being due to internal waves. Figure 2d shows rms variances in  $S(T)$  that are several times smaller throughout much of the water column.

This high  $T$ – $S$  correlation reflects the potential for obtaining improved  $S$  analyses from  $T$ . For instance, if one were interested in determining salinity at 100 m, the maximum uncertainty in  $S$  would be 0.35 psu using the  $S = S(z)$  profiles (Fig. 2b), whereas, if  $T$  were known, uncertainty would be almost halved to 0.2 psu, if  $T$ – $S$  curves are used. There are, however, depths, or temperatures, for which inference of  $S$  from  $T$ – $S$  properties still leads to large uncertainties. These events happen, for example, in places where advective intrusions come about; the spike at about 17°C and 35 psu in Fig. 2c and the rms variance peak at about 200 m in Fig. 2d are examples. This feature is caused by inclusion of profiles from the last three days of the period. Also, use of the  $T$ – $S$  relation does not work well where a halocline divides an isothermal layer (e.g., the 20°C water in Fig. 2c and the absolute rms variance  $T$ – $S$  maximum in Fig.

2d). In spite of these problems, the  $S(z)$  rms variance is almost everywhere larger than the  $S(T)$  rms variance (Fig. 2d), and therefore  $S(T)$  generally would be preferable over  $S(z)$  for these data.

A feature in Fig. 2 worth noting is the significant  $S$  variability in the upper water column (e.g., 0–80 m, Fig. 2b), which seems to appear only in the rms variance in  $z$  space. In fact, the  $S(z)$  rms variance is greater than the  $S(T)$ -based rms variance by a factor of 7–8 at about 80 m (Fig. 2d). This feature is the barrier layer, a rather common feature in the western tropical Pacific [e.g., Vialard and Delecluse (1998a)] that lies between 30 and 70 m (roughly 40-m thick). The highly stratified water column probably encourages internal wave propagation, which then preserves the  $T$ – $S$  relation, leading to high  $S$  variability but excellent results from correlating  $S$  with  $T$  in this region of the water column.

The conclusion is that it is very useful to use an  $S(T)$  relation that is derived locally on these short timescales. We would expect this result to hold almost anywhere in the global oceans. This result is not particularly surprising, considering that the 10-day period is rather short for distinct water mass variations to occur. Nonetheless, changes like the intrusion are likely to happen, even over a few days. Unfortunately, this temporal density of  $T$  data is hardly ever collected, and the scope of the analysis needs to be enlarged to longer time periods, which will permit other mechanisms to affect the  $S$  variability.

*b. Weekly analysis:  $S(z)$  versus  $S(T)$*

The focus here is on the intraseasonal variability seen when comparing weekly averaged profiles. The choice of a week corresponds to the typical time between assimilation cycles for current seasonal forecasting applications and for previous ocean assimilation studies (Derber and Rosati 1989; Rosati et al. 1995; Rosati et al. 1996; Carton and Hackert 1990; Kleeman et al. 1995). The representative box N is discussed (Fig. 3), although the results from the other boxes are similar. Nine profiles, one for each week (49–57), are shown along with the “mean” profile in bold. Table 1 shows the number of individual CTD profiles available for each week.

The weekly mean  $T$  and  $S$  profiles (both in  $z$  and  $T$  spaces) were first calculated by omitting data from the first level of most casts (top 15 m), which showed inconsistencies and had to be ignored [see e.g., Ando and McPhaden (1997)].

It is evident that there are considerable differences between  $S(z)$  and  $S(T)$  variabilities. The latter presents a fairly compact series of profiles, suggesting that mixing and nonlocal dynamics play minor roles. This observation also can be read from Fig. 3d, which shows how the  $S(T)$  rms variance almost everywhere is smaller than the  $S(z)$  rms variance. At depths where this relation is not true (about 160 m and below 320 m), the  $S(T)$

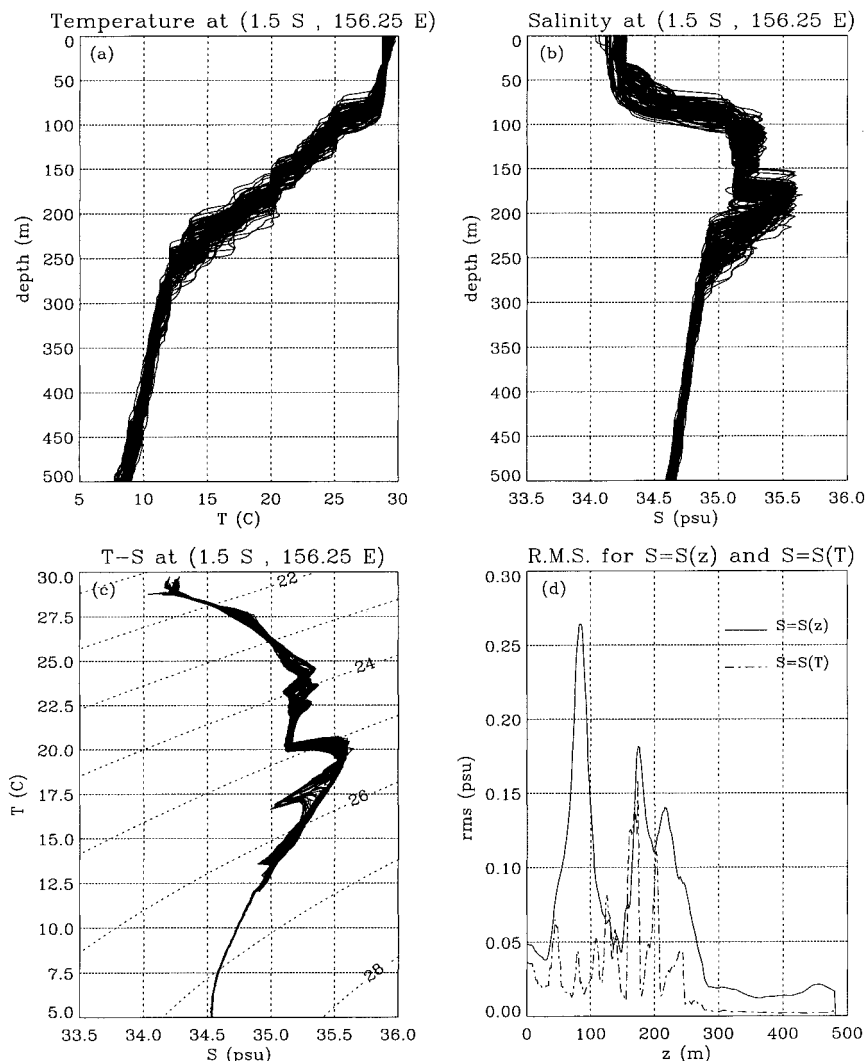


FIG. 2. Data analysis. From the top left: (a) temperature profile, (b) salinity profile, (c)  $T$ - $S$  diagram, and (d)  $S(z)$  to  $S(T)$  rms variance comparison for the 104 CTD casts at S1 location for a 10-day period in Nov 1992 (see Fig. 1). Note that to transform the  $S(T)$  rms from  $T$  to  $z$  space [in (d)] the average temperature profile [from (a)] has been used.

rms variance values are decidedly smaller than 0.05 psu and not very different to the  $S(z)$  rms variance (although the former variance contains some noise caused by the data processing). The  $S(z)$  rms minimum at about 160 m is peculiar. By looking at the single  $S(z)$  profiles and noting the time of each profile, it is found that in the last three weeks of January 1993 there is a considerable and sudden increase in the upper-layer salinity (top 150 m in Fig. 3b). The cause appears to be a first-mode Kelvin wave, as the respective temperature profiles also indicate, and the changes are consistent with an upward displacement of about 50 m (corresponding to an average vertical velocity of about  $8.3 \times 10^{-5} \text{ m s}^{-1}$  over one week). The presence of a salinity maximum near this depth then gives the anomalously small  $S$  variability

because the salinity at 160-m depth just happens to remain the same after the displacement.

Similar results hold for the other two boxes SW and SE (not shown), which again suggests that  $S(T)$  relationships are persistent on the weekly timescale. One could compare the validity of a climatological  $S_{cl}(T)$  on this timescale, but this comparison is discussed in section 3c.

### c. Seasonal analysis: $S(T)$ versus $S_{cl}(T)$

The results to this point have demonstrated that salinity and temperature are strongly correlated, at least on timescales of weeks in the tropical Pacific, and that this correlation supports the use of  $S(T)$  functions for

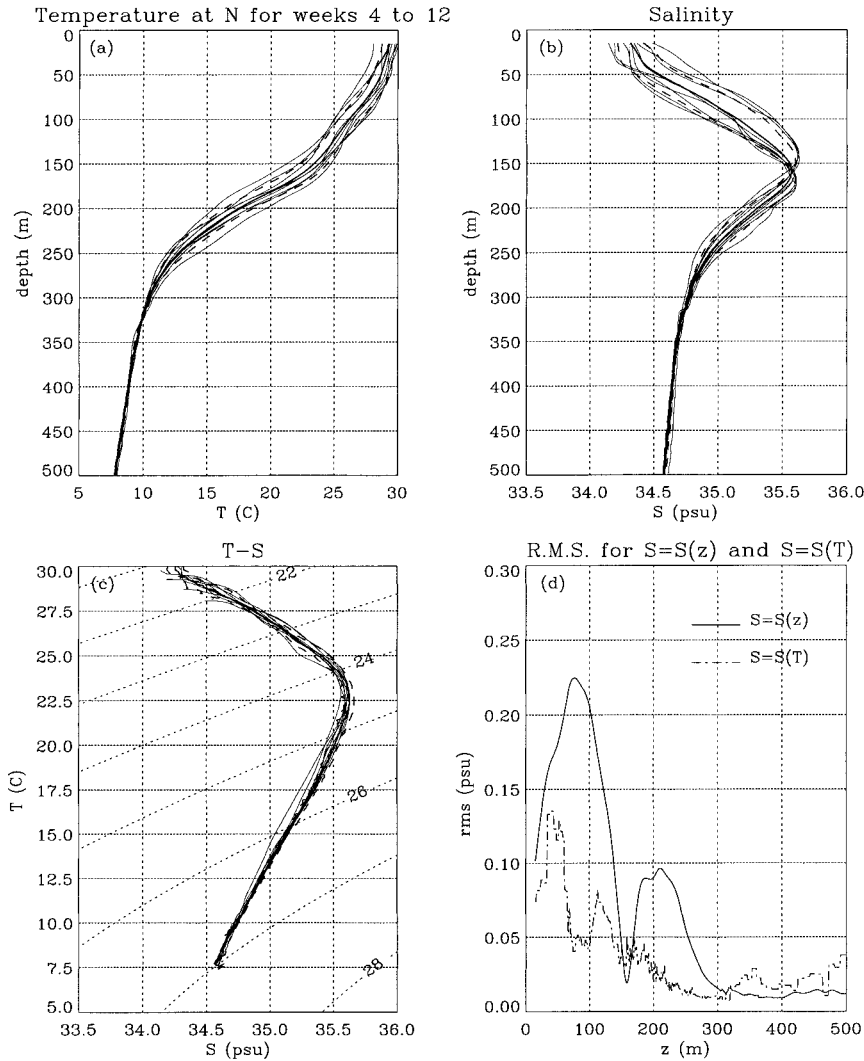


FIG. 3. As in Fig. 2 but for the N box. The solid lines in (a)–(c) are the mean profiles and the dashed lines their respective rms variances. Data refer to the weeks displayed in Table 1.

analyzing for salinity when temperature data are available. However, at some timescale, variability of the  $S(T)$  relationship is likely to become important. This section compares, mostly qualitatively, the seasonally averaged  $S(T)$  function from the NODC data in this study with  $S_{cl}(T)$  from a climatological dataset (e.g., Levitus). The former is obtained by averaging the individual weekly profiles discussed in section 3b (cf. also Table 1). Also, since the Levitus seasonal definition and spatial grid do not coincide well with those of this study’s data, linearly interpolated averages of seasonal Levitus data are used to get as close as possible to an appropriate climatological value. In Fig. 4, these averages are compared with the mean profiles from each of the boxes N and SW along with associated rms variability, based on the weekly variations, for the study dataset (the behavior of the SE box, not shown, is similar to the SW box). Although the two datasets are not entirely independent

(the climatological dataset used all data from 1900 up to 1993, including those from this study), significant salinity differences appear in both  $z$  and  $T$  spaces. There are differences in all boxes, especially N. The  $S(T)$  properties for the south differ mainly in the upper water column ( $T \geq 22.5^\circ\text{C}$ , corresponding to a depth of about 160 m), whereas for the N box there is disagreement throughout the main thermocline (from about 100 to 310 m). Given the large variability in the evaporation–precipitation forcing in this region [e.g., Ando and McPhaden (1997)], it is not unexpected to find mismatches in the upper water column, but the differences in the N-box  $S(T)$  properties are surprising. A plausible explanation is that this is a region of high meridional salinity gradient; therefore, a long-term average can smooth out much of the variability. Last, by comparing the rms variability for the  $S(T)$  curves, it is observed that even a long time period, such

TABLE 1. Temporal distribution of the CTD data for the selected subregions. The second line is the number of weeks (from Jan 1992, for convenience). The first column refers to the selected boxes (see Fig. 1). Bracketed weeks are ignored in the analysis.

	Weeks in 1992								Weeks in 1993							
	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
N	(0)	9	(0)	(0)	54	46	40	42	34	12	40	36	(2)	(0)	(0)	(1)
SW	20	24	25	61	51	(2)	22	29	30	14	(1)	11	29	23	26	13
SE	(1)	70	71	17	30	16	30	52	68	42	48	4	(0)	(0)	(0)	(0)

as a specific season, is not described adequately by the climatological dataset, which falls well outside the observed seasonal variability over a considerable temperature range. As a consequence, any temporally closer source of  $S(T)$  information would represent a better option for inferring salinity. This result is why we have chosen to seek different background  $S(T)$  fields for data assimilation purposes. In the following sections some results from a numerical model are examined, and how  $S(T)$  relationships vary in a modeling context is considered.

4. A salinity analysis algorithm

This section considers how the results derived above can be applied in the context of data assimilation with  $T$  profile data. The problem is how to combine a priori (or forecast) temperature and salinity fields with the measured in situ temperature (e.g., XBT) profile. The emphasis is on trying to get the correct salinity, but the dynamic heights obtained with various  $S$  reconstructions also are compared, since these also can be obtained from altimeter data. Maintaining the a priori  $S(T)$  relation is

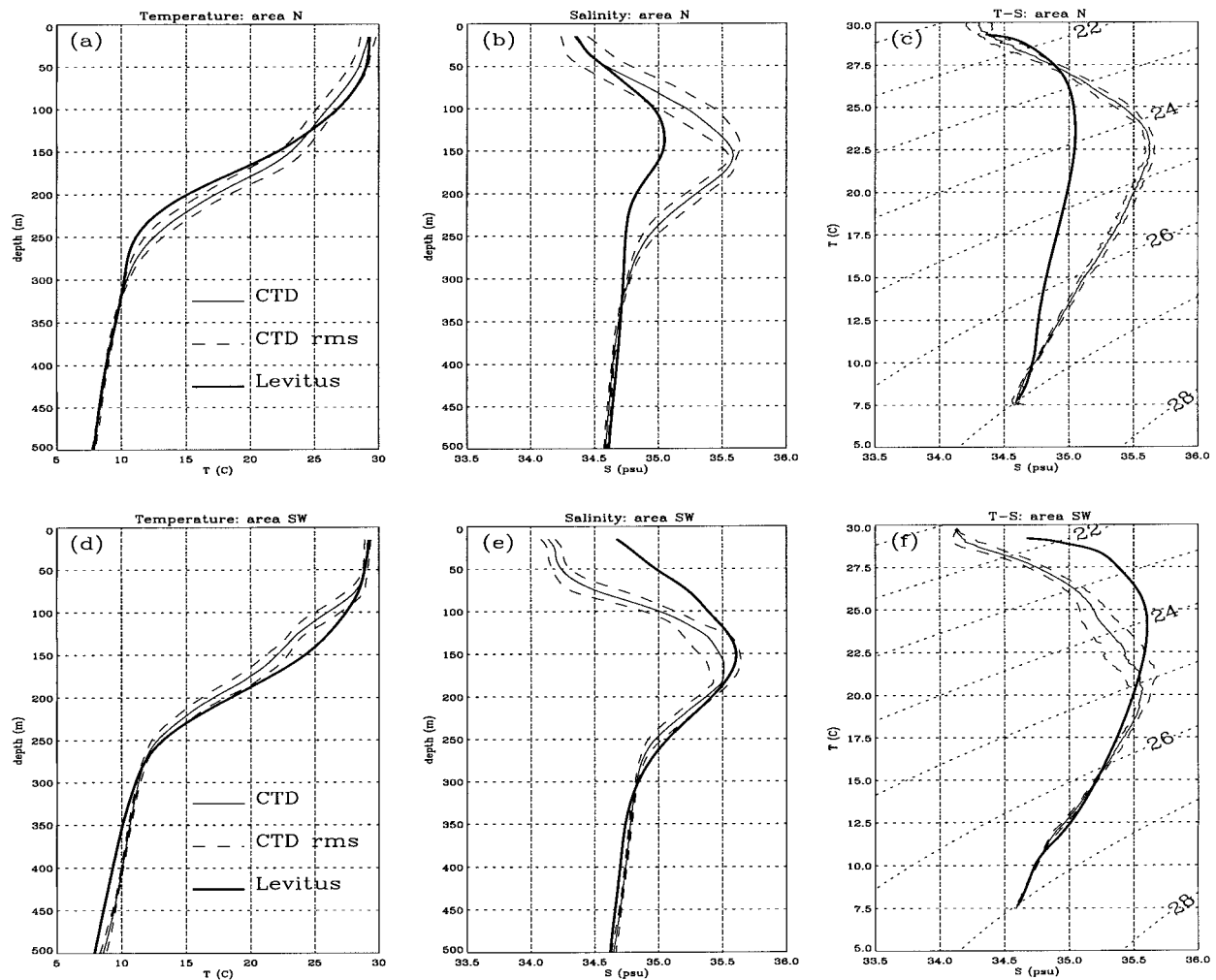


FIG. 4. Comparison between climatological data (thick line) and seasonally averaged CTD data (lighter line) for the N (top) and SW (bottom) boxes. From left to right: (a) and (d) temperature profile, (b) and (e) salinity profile, and (c) and (f)  $T-S$  diagram.

the preferred method and is compatible with the principle that water mass properties should be changed only when experimental evidence is available, as suggested in Haines (1994).

Although the data analysis in the previous sections refers only to limited areas in the western Pacific Ocean, the assimilation method with model data presented in this section is not so restricted and could be applied over wide regions of the oceans, including difficult areas such as those where temperature inversions are present (Emery 1975; Emery and O'Brien 1978). To simplify the initial discussion, we assume that the temperature profiles are monotonic. This mostly is true in regions within the band  $\pm 40^\circ\text{N}$  [e.g., Emery and Dewar (1982)]. At higher latitudes or in regions of particularly high salinity variability (e.g., the eastern Atlantic, just off the Gibraltar Strait, or the Mediterranean) the monotonicity assumption is no longer met. The approach for this more general case is discussed later in the section.

Even when monotonic temperature profiles are considered, the limited vertical extent of XBTs (typically 500 m) prevents one from determining uniquely the full temperature profile. Usual XBT data assimilation methods leave the water column unmodified below the deepest temperature available ( $T_{\text{ref}}$ ), trusting that the internal dynamics of the model will provide a realistic response. Here a different approach, similar to that used by Cooper and Haines (1996) for altimeter assimilation, is introduced. It requires that  $T_{\text{ref}}$  will match the model water of the same temperature by vertically shifting the model water column until this match is the case. This method preserves that part of the  $T$  and  $S$  hydrography below the deepest temperature level, although  $T$  and  $S$  can change as a function of depth. A big advantage of this method is that it avoids initiating convection, which could generate new water masses. This vertical displacement is the first stage of the assimilation procedure. To get the salinity analysis within the depth range of the XBT, the second stage of assimilation subsequently assumes that the background  $T$ - $S$  properties are preserved and are taken from the a priori model state. Figure 5a schematically shows how the water column is modified by the analysis, both by vertical displacement and when  $S(T)$  is preserved; this model state also is compared to the persistence analysis (Fig. 5b), which leads to permanent water mass modifications and easily can lead to convective instability.

In the nonmonotonic case, for which  $S = S(T)$  is not uniquely defined (i.e., it is not a function), we adopt the solution of considering  $S = S(T, z)$  [or, alternatively,  $S = S(T, \sigma_t)$ , where  $\sigma_t$  is the density anomaly]. The idea stems from the fact that vertical water column displacements do occur, because of internal waves or mesoscale features (Emery and Dewar 1982). Normally, these processes do not modify to any great extent the water mass properties, and therefore one can distinguish between two (or more) isothermic parcels according to their depth. Salinity is then recovered from the nearest  $T$  point

in  $Z$  in the background field. Also, to mimic the physical picture, an upper limit to the vertical displacement can be assigned. This limit will depend on the time lag between assimilation steps and on the location, particularly the latitude.

## 5. Assimilation results

The data assimilation method presented in the previous section is tested by combining CTD data profiles, one with another, and by combining different model realizations. The procedure for the CTD data is as follows. The a priori hydrography is taken as the complete CTD  $T$  and  $S$  profile data at time  $t_1$ . Then only the upper water column (500 m) of the  $T$  field from a second CTD at a different time  $t_2$  (e.g., a week later) is used to simulate an observed XBT at the analysis time  $t_2$ . The corresponding analysis  $S$  field is derived by vertical displacement followed by  $T$ - $S$  property preservation based on the a priori  $T$  and  $S$  profiles. The true  $S$  field at time  $t_2$  is available as a control field to check the accuracy of the method. The reference depth  $z_{\text{ref}} = 500$  m is chosen as this depth is the normal maximum reached by XBT probes. An analogous approach is adopted later for combining numerical model fields.

There is no unique way of assessing techniques like this one. Three measures of success in reproducing fields are considered. The first is the dynamic height integral, 0–500 m, which often has been used (Emery 1975; Emery and Dewar 1982; Kessler and Taft 1987; Delcroix et al. 1987; Vossepoel et al. 1999). However, a measure of the details of the vertical density structure also is defined by employing a 1-norm quantity defined by

$$\Delta\sigma = \sum_i [|\sigma_A(i) - \sigma_O(i)|\Delta z(i)]/z_{\text{ref}}, \quad (1)$$

where the subscripts  $A$  and  $O$  stand for assimilation (or analysis) and observation/truth, respectively;  $\Delta z$  is the layer thickness; and  $i$  is the level index. Thus Eq. (1) supplies an estimate of the error in the density at all depths, preventing compensations that can occur in the dynamic height measure. Further, to check whether the assimilation improves the vertical  $S$  profiles, a relative assessment (in percent), using a 1 norm, is calculated:

$$\Delta S = \left( 1 - \frac{\sum_i |S_A(i) - S_O(i)|\Delta z(i)}{\sum_i |S_B(i) - S_O(i)|\Delta z(i)} \right) 100, \quad (2)$$

where the subscript  $B$  stands for background or a priori field. According to Eq. (2),  $\Delta S$  ranges from a maximum of 100% in the ideal case to  $-\infty$  when  $\sum_i |S_B(i) - S_O(i)|\Delta z(i)$  tends to zero. A negative  $\Delta S$  thus will signify that the assimilation has worsened the background  $S$  profile. The assessment indices (1 and 2) are evaluated for the part of the water column that extends from 15 to 500 m. Retaining the structure of section 3, that is, decreasing frequency from daily to seasonally, some of the assimilation method results are presented.

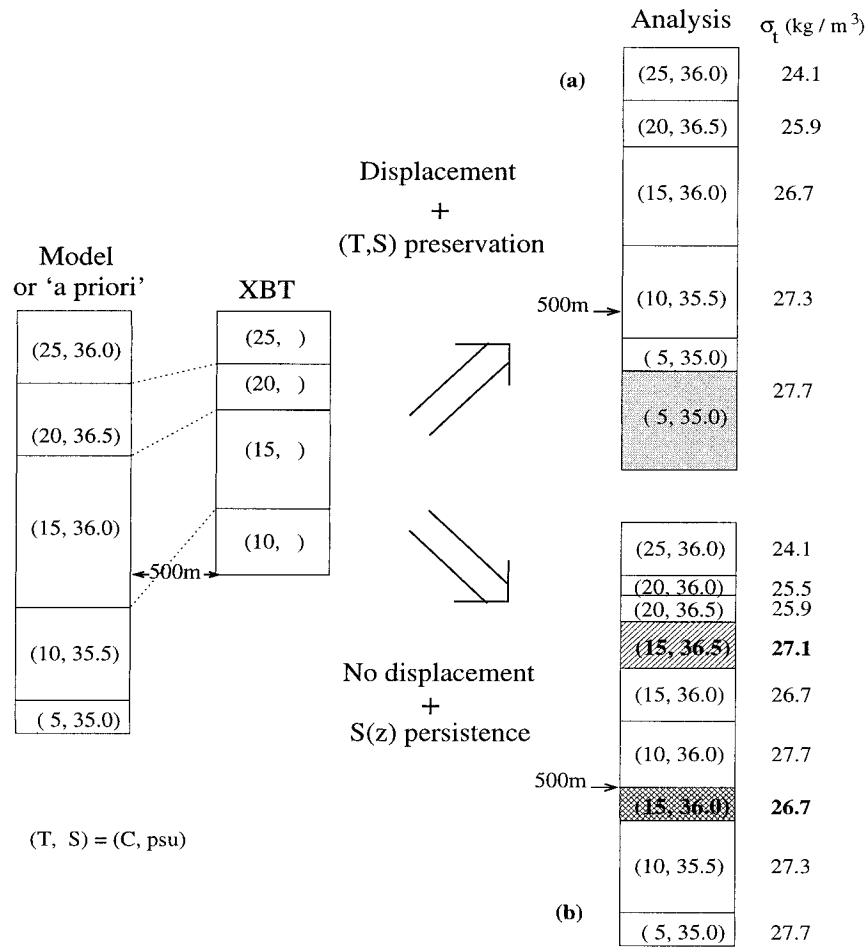


FIG. 5. Schematic representation of the two stage analyses (right) obtained by combining model and XBT water columns (left), for the monotonic case. On the right, the density anomaly ( $\sigma_t$ ) is shown. (a) This paper's method: the model is displaced to match the bottom of the XBT (stage 1) and the  $T-S$  preservation is applied (stage 2). The shaded layer ( $5^\circ\text{C}$ ,  $35$  psu) highlights added water, caused by the vertical displacement. (b) Direct insertion of the XBT into the top 500 m of the model: the absence of a displacement causes an unstable layer ( $15^\circ\text{C}$ ,  $36$  psu, crosshatched). Also, the unaltered  $S$  profile introduces another unstable layer ( $15^\circ\text{C}$ ,  $36.5$  psu, forward hatched).

a. Daily assimilation

The first test has been carried out by selecting two CTD profiles at the S1 location (Fig. 1), eight days apart, that is,  $t_1 \equiv$  day 0 and  $t_2 \equiv$  day 8 (Fig. 6).

Despite a water mass intrusion that has occurred at  $t_2$ , for  $\sigma \geq 26 \text{ kg m}^{-3}$ , the  $T-S$  curves for the two profiles very closely resemble each other (Fig. 6c). This resemblance suggests that a good salinity analysis is attainable by preserving the  $T-S$  relation. By contrast, the salinity in  $z$  space varies appreciably, therefore discouraging a persistence-type approach.

The actual salinity analyses are shown in Fig. 6b. Here  $S_o$  is the observed true profile,  $S_{sz}$  is the a priori profile that would be preserved if salinity were unaltered, and  $S_{st}$  is the analysis preserving the  $T-S$  relation. It is evident that the  $S_{st}$  analysis represents the control profile very well, and this result is confirmed by the

very high  $\Delta S$  index (above 81%). One of the biggest  $|S_o - S_{st}|$  discrepancies is near the surface, within a shallow layer (roughly 25 m thick). The water intrusion at about 200 m also is not picked up by this paper's method, because the background field does not contain that water mass type. This failure shows up in the difference between  $S_o$  and  $S_{st}$  at about 215 m. Last, the positive impact of the assimilation is highlighted further by the small figures of both the  $\Delta\sigma$  index and the dynamic height difference ( $D_o - D_{st}$ ) (Fig. 6d).

b. Weekly assimilation

The assimilation also was applied to the weekly averaged profiles (see section 3b), assimilating between weekly averaged profiles two weeks apart. The example here is from the SE box. In addition, a parallel analysis,

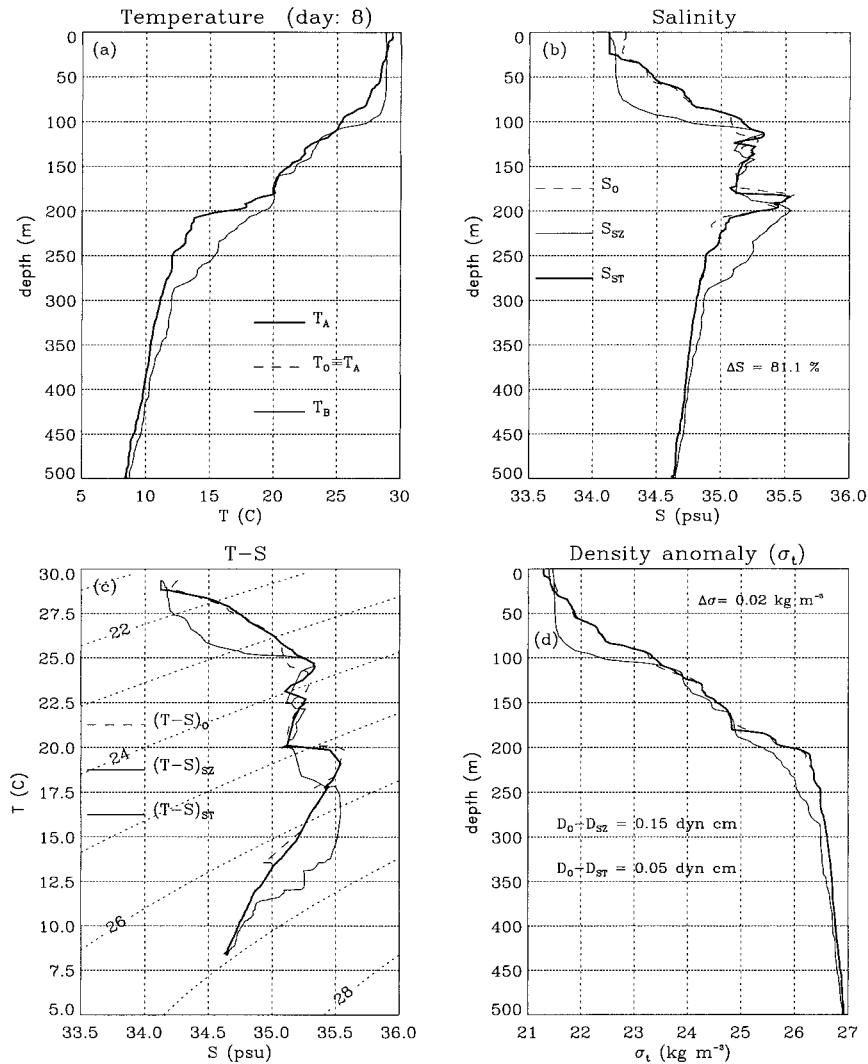


FIG. 6. Data assimilation. From the top left: (a) temperature profile, (b) salinity profile ( $S_{ST}$  is this paper's analysis), (c)  $T$ - $S$  diagram, and (d) density anomaly for two profiles eight days apart at S1 location. Note that the assimilated and control  $T$  profiles are coincident, having assumed no errors in the  $T$  measurement. The  $\Delta S$  figure, calculated by Eq. (2), is shown in (b) whereas the  $\Delta\sigma$  [Eq. (1)] and the dynamic height differences are in (d). For consistency,  $(D_o - D_{sz})$  and  $\Delta\sigma$  have been evaluated by using the  $T_o$  profile, instead of  $T_B$ .

which uses the Levitus climatological  $T$ - $S$  for this region, also was carried out.

This study considered  $t_1 \equiv$  week 53 and  $t_2 \equiv$  week 55 (Fig. 7). As implied by the  $\Delta S_{ST}$  index, the  $T$ - $S$  background yields a 59% better analysis than does leaving  $S(z)$  unmodified. A further confirmation is given by the dynamic height estimation [ $-1.2$  dyn cm for  $S(z)$  compared with  $0.2$  dyn cm for  $S(T)$ ]. For the climatological dataset comparison, the two  $T$ - $S$  dynamic heights differ substantially (i.e.,  $|D_o - D_{ST}| \ll |D_o - D_{STcl}|$ ), as do the  $\Delta\sigma$  indices ( $\Delta\sigma_{ST} \ll \Delta\sigma_{STcl}$ ) (Fig. 7d), mainly due to differences in the top 100 m. This result is because the climatological dataset  $T$ - $S$  curve differs considerably from the measured one (Fig. 7c).

To check the impact of the vertical displacement (stage 1, Fig. 5), the 500–1000-m dynamic height for the  $S_{SZ}$  and  $S_{ST}$  analyses also was computed. Although the displacement is not very large (16 m), it does slightly improve the description of the water column below  $z_{ref}$ .

An interesting feature in the  $S_{ST}$  salinity profile (Fig. 7b) is seen above 50-m depth where an isohaline layer is present. When observed temperatures are higher than any temperature in the a priori field (always near the surface), we assume that the water is just heated, and therefore the salinity is not changed and the whole of the warmer layer takes the salinity of the warmest waters in the background field (cf. Figs. 7a, b).

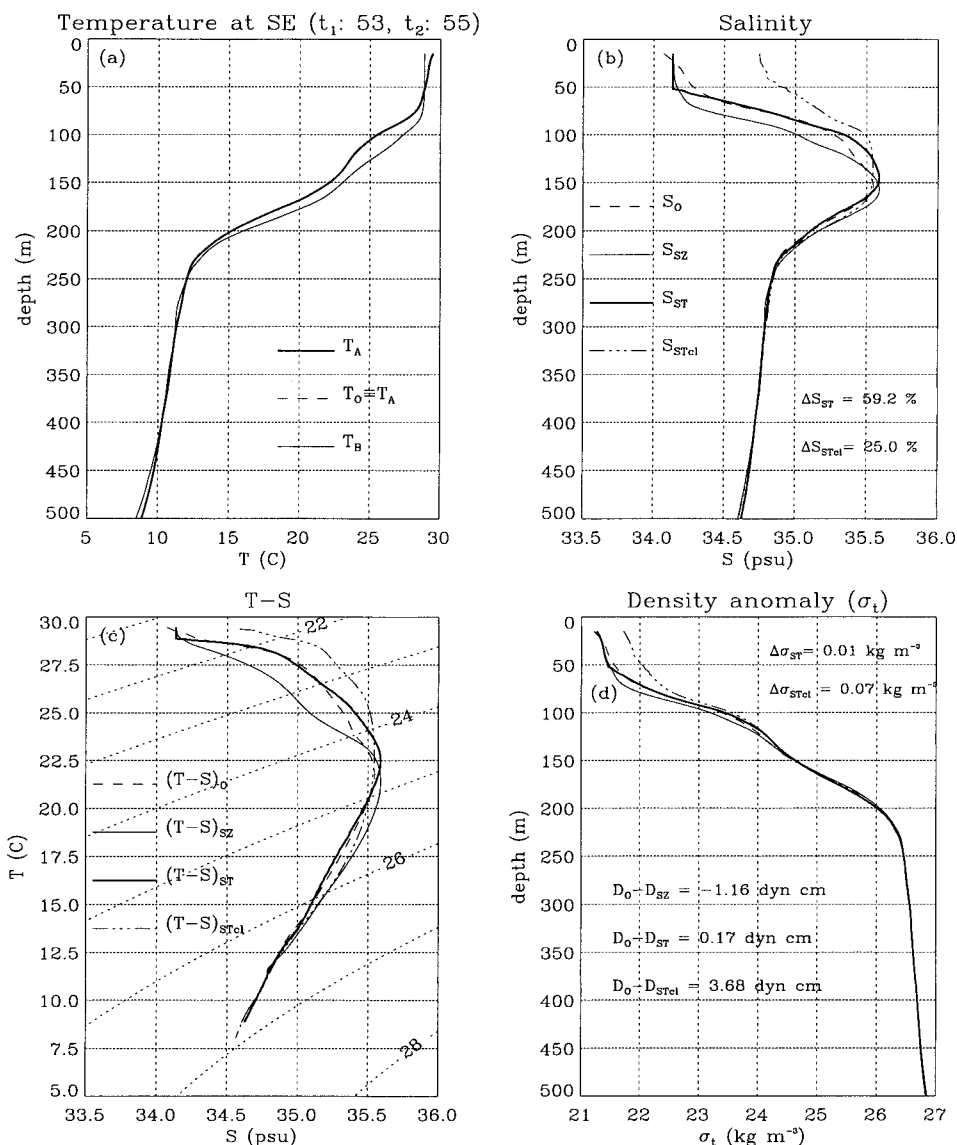


FIG. 7. As in Fig. 6 but for the two weekly averaged profiles two weeks apart, for the SE box. Here also the assimilation performed by using climatological  $T$ - $S$  relations is shown (subscript “cl”). Moreover,  $(D_o - D_{stcl})$  and  $\Delta\sigma_{stcl}$  have been evaluated by using the  $T_o$  profile.

*c. Seasonal assimilation*

Two weekly averaged profiles 10 weeks apart for the SW (Fig. 8) box are considered. In this case  $t_1 \equiv$  week 46 and  $t_2 \equiv$  week 56.

In this SW box, the background and observed  $S(z)$  profiles differ substantially (Fig. 8b) and the corresponding  $T$ - $S$  curves also show considerable differences (Fig. 8c). In particular, a water intrusion with  $\sigma_t$  between approximately 23.7 and 24.4  $\text{kg m}^{-3}$  is present at  $t_2$ . Although this feature cannot be captured by the assimilation method (cf.  $S_o$  and  $S_{st}$  in Fig. 8b, between 80- and 120-m depth), the  $\Delta S_{st}$  is still over 70%, indicating a very positive impact of the assimilation scheme. This impact also is seen in the  $|D_o - D_{st}|$  index, which is

equal to about 0.3 dyn cm (Fig. 8d). The climatological dataset does not perform nearly as well, with  $|D_o - D_{stcl}| \approx 3.9$  dyn cm, even bigger than  $|D_o - D_{sz}|$ .

As in the weekly assimilation, the vertical displacement (of the same magnitude, 16 m) again provides an improved 500–1000-m dynamic height.

*d. Model field assimilation*

We have taken the first steps toward implementing this scheme in a model. In collaboration with the European Centre for Medium-Range Weather Forecasts (ECMWF), a further test has been performed by using two instantaneous HOPE model outputs. This ocean

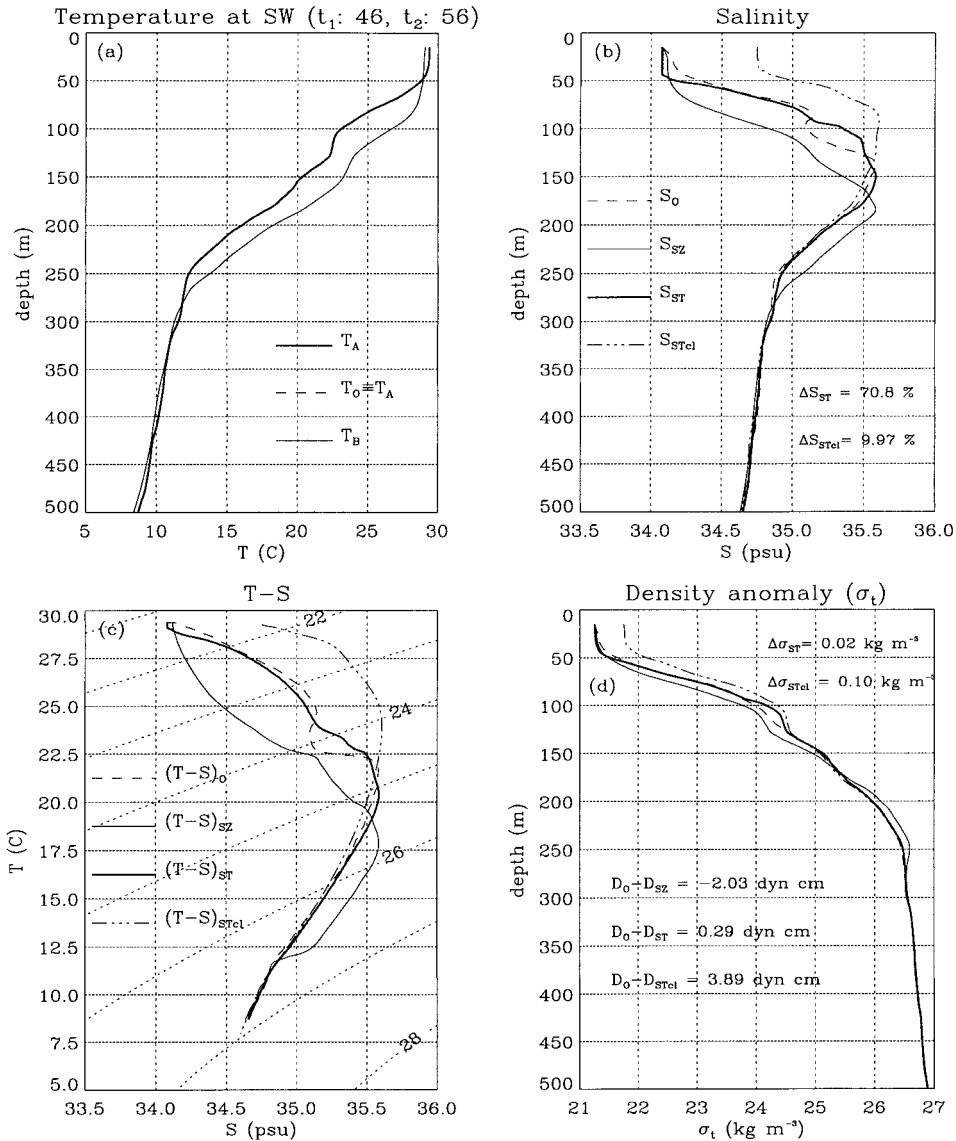


FIG. 8. As in Fig. 7 but for two profiles separated by 10 weeks (seasonal assimilation), and for the SW box.

model is the one used as part of a coupled seasonal prediction project at ECMWF. The model is primitive-equation-based, three-dimensional, and uses  $z$  coordinates with variable surface height. Forcing is provided by surface solar heat flux, wind stress, and precipitation [see Wolff et al. (1997) for a more detailed description]. Only the equatorial Pacific cross section is shown here. Figure 9a shows the thermal field for the upper water column showing the strong stratification and an upward thermocline slope to the east. This configuration implies that vertical displacements or thickness variations of the water columns would induce large changes to the density distribution.

The two realizations discussed are initialized with the same fields but then run for three months each, using

different wind stress forcing. The surface heat fluxes also will differ to the extent that they depend on the SST and wind fields. This time integration is long enough to generate large differences in the hydrography, as shown by the temperature difference section (Fig. 9b). The main temperature differences (as much as  $-1.8^\circ\text{C}$ ) are found east of the international date line for depths up to 200 m, where the thermocline is more pronounced in Fig. 9a. Large differences (about  $1.8^\circ\text{C}$ ) are present also in the upper water column between  $260^\circ$  and  $280^\circ\text{E}$ . Salinity differences also are quite distinct as shown by the difference cross section (Fig. 9c), and are spread over most of the top 200 m across the basin. In particular, large values are found between  $200^\circ$  and  $240^\circ\text{E}$ , ranging from 0.45 psu in the upper water column

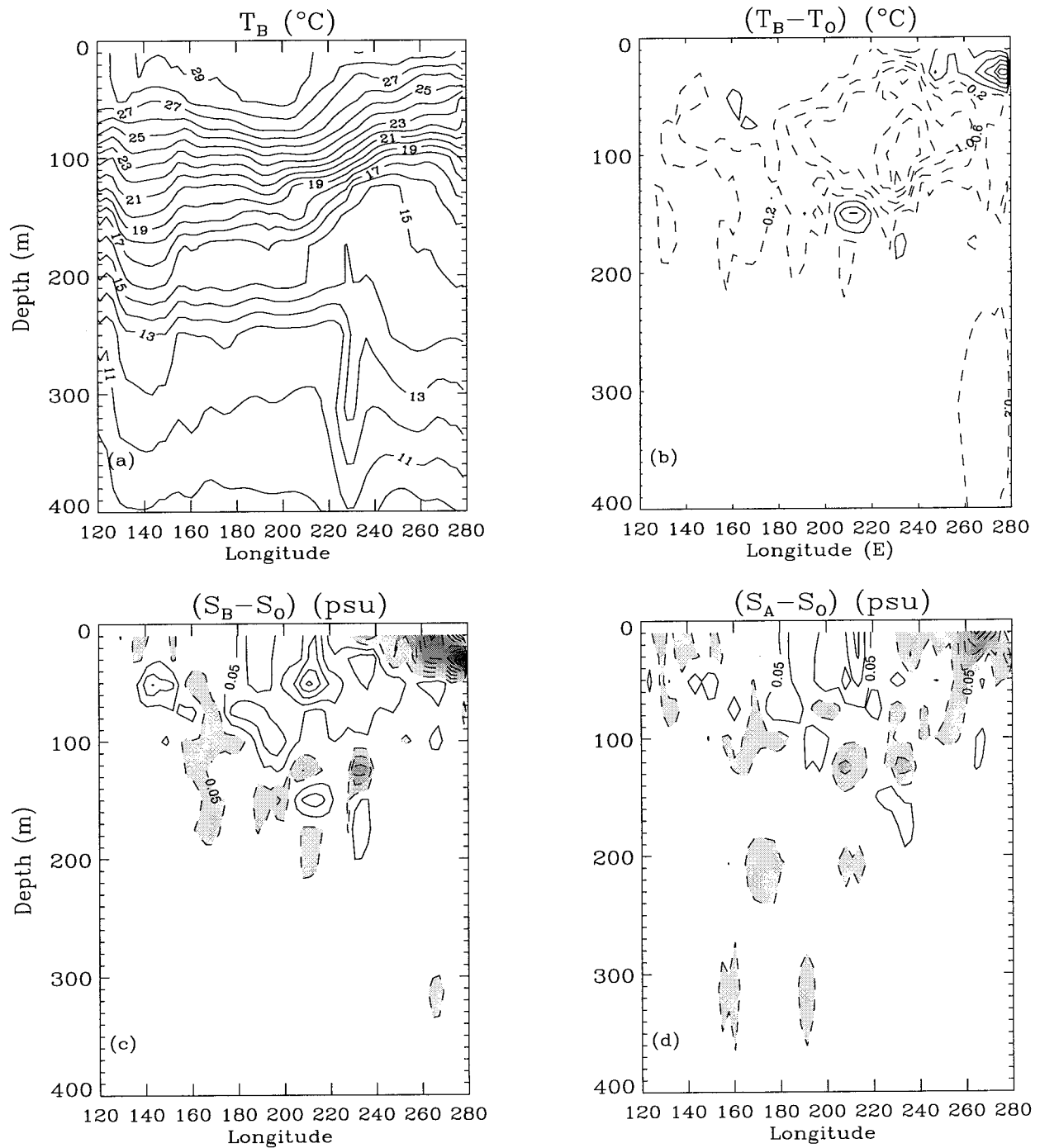


FIG. 9. Data assimilation using HOPE model outputs along the equatorial tropical Pacific. The two 3-month runs were initialized in the same way and thereafter only the wind stress forcing was different. (a) Background temperature section after 3 months; (b) temperature difference between background and control run (contour interval is  $0.4^{\circ}\text{C}$ ); this increment is the one used by the assimilation scheme. Salinity difference (contour interval is  $0.1$  psu): (c) before and (d) after assimilation.

to  $-0.25$  psu lower down. A considerable minimum also is seen in the eastern part of the section (about  $-0.85$  psu). The sign of the salinity differences generally is opposite to that of the temperature differences such that there is not a density compensation.

Assimilation is applied to run B (background, shown in Fig. 9a), using the temperature increments shown in Fig. 9b but maintaining the  $S(T)$  relationship from run B. The salinity difference results are presented in Fig. 9d and show a clear improvement over most of the

section. All of the extreme values present before assimilation seen in Fig. 9c have been reduced considerably, especially between 200° and 220°E and east of 260°E. A minor negative effect has been induced below the main thermocline between 150° and 200°E, presumably by discrepancies in the  $T$ - $S$  relationships that develop over the 3-month period. In spite of this effect, the assimilation has led to a satisfactory  $\Delta S$  index of about 23% improvement.

## 6. Discussion and further work

An alternative approach to including a salinity analysis as part of a temperature-profile XBT data assimilation framework has been presented. It embraces the hypothesis of  $T$ - $S$  relationships as its principal cornerstone. The novelty comes from using  $T$  and  $S$  fields from recent data, rather than from any climatological dataset. The main motivation is that the usual approach of using a long-term average  $T$ - $S$  relationship from climatological data generally is found to be inadequate in reconstructing the density field. We suggest therefore that a temporally closer  $T$ - $S$  field, such as from a model prediction or any recent CTD data, will provide a better solution.

A CTD dataset for a region of the western tropical Pacific was analyzed and it was shown that one can use  $T$ - $S$  relationships to link them very accurately for time periods on the order of a few weeks, except sometimes in the uppermost layer (about 50 m). The assimilation scheme is useful both for CTDs and model realizations. In the former case, values for the 15–500-m dynamic height differences (i.e.,  $|D_o - D_{stl}|$ ) rarely bigger than 2 dyn cm were obtained. Comparing this accuracy to altimetry shows that 2–3 dyn cm is a lower limit given by the satellite observational accuracy when low-frequency barotropic changes can be neglected (Katz et al. 1995). It also has been shown that with assimilation performed using the climatological  $T$ - $S$  relation, only mediocre results are obtained, with  $|D_o - D_{stcl}|$  normally bigger than 3.5 dyn cm. This result indicates the inadequacy of the climatological dataset in representing the water mass properties at any given time, possibly partly because of the high degree of smoothing in  $z$  space that was used. In particular, note that keeping  $S(z)$  fixed sometimes leads to better dynamic height analyses than does using climatological water properties (e.g., Fig. 8d; although in this case compensations in the vertically averaged density field contribute to producing this result).

While the results shown are promising, it should be emphasized that the analyzed CTD data only refer to a few limited areas, and therefore it has not been shown that this assimilation scheme will perform well globally. However, even when temperature inversions or water intrusions occur, the scheme still can yield positive outcomes at other levels in the water column. One of the implicit requirements of the scheme is that a sufficient

stratification, that is, a wide enough  $T$  range, be present in the a priori water column. In the tropical regions this condition is met easily but problems may arise as one moves poleward, as described by Emery and Dewar (1982). Their suggested borderline at around 40°N calls for further work to examine the XBT assimilation problem in midlatitudes.

Other measurements also could be employed to improve the  $S$  analysis, especially near the surface. Vossepel et al. (1999), for instance, propose a way of inferring salinity corrections by incorporating a further constraint to the  $T$ - $S$  preservation given by independent observations of sea surface height (SSH) (however, they also assumed that the sea surface salinity was known). However, good statistics are required to deduce which part of the salinity profile is likely to require altering. More important, the SSH accuracy, together with the uncertainty in the level of dynamic height reference, means that dynamic height may not provide sufficient additional information to be useful if the a priori  $T$ - $S$  relation is sufficiently well known. Last, even when an accurate enough dynamic height is available to provide independent data, some salinity errors may compensate each other in the density profile and thus still be undetectable.

Despite these pitfalls, we believe the synergistic use of  $T$  (correcting  $S$  using  $T$ - $S$  functions) and SSH datasets will be the most useful way forward in reconstructing the density field for ocean and coupled modeling experiments.

*Acknowledgments.* The authors are grateful to David Anderson and Oscar Alves for providing the HOPE model fields and for several constructive discussions, and to Donald Collins for providing the CTD data from the NODC data bank. We also thank Alan Fox for fruitful comments on an earlier version of the manuscript. Author AT would like to thank William Emery for valuable epistolary suggestions. Author AT was supported by the EC Grant ENV-CT95-0113 AGORA project and a grant from SOC.

## REFERENCES

- Anderson, D. L. T., 1995: The Tropical Ocean Global Atmosphere programme. *Contemp. Phys.*, **36**, 245–265.
- , J. Scheinbaum, and K. Haines, 1996: Data assimilation in ocean models. *Rep. Prog. Phys.*, **10**, 1209–1266.
- Ando, K., and M. J. McPhaden, 1997: Variability of surface layer hydrography in the tropical Pacific Ocean. *J. Geophys. Res.*, **102**, 23 063–23 078.
- Carton, J. A., and E. C. Hackert, 1990: Data assimilation applied to the temperature and circulation in the tropical Atlantic, 1983–84. *J. Phys. Oceanogr.*, **20**, 1150–1165.
- Cooper, M., and K. Haines, 1996: Altimetric assimilation with water property conservation. *J. Geophys. Res.*, **24**, 1059–1077.
- Cooper, N., 1988: The effect of salinity on tropical ocean models. *J. Phys. Oceanogr.*, **18**, 697–707.
- Delcroix, T., G. Eldin, and B. A. Hénin, 1987: Upper ocean water masses and transports in the western tropical Pacific (165°E). *J. Phys. Oceanogr.*, **17**, 2248–2262.

- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.
- Donguy, J. R., 1994: Surface and subsurface salinity in the tropical Pacific Ocean. Relations to climate. *Progress in Oceanography*, Vol. 34, Pergamon Press, 45–78.
- Emery, W. J., 1975: Dynamic height from temperature profiles. *J. Phys. Oceanogr.*, **5**, 369–375.
- , and A. O'Brien, 1978: Inferring salinity from temperature or depth for dynamic height computations in the North Pacific. *Atmos.–Ocean*, **16**, 348–366.
- , and L. J. Dewar, 1982: Mean temperature–salinity, salinity–depth, and temperature–depth curves in the North Atlantic and North Pacific. *Progress in Oceanography*, Vol. 16, Pergamon Press, 219–305.
- Ghil, M., and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Advances in Geophysics*, Vol. 33, Academic Press, 141–266.
- Gill, A. E., 1982: *Atmosphere–Ocean Dynamics*. Academic Press, 662 pp.
- Haines, K., 1994: *Dynamics and Data Assimilation in Oceanography*. Vol. 19, *Data Assimilation: Tools for Modelling the Ocean in a Global Change Perspective*, P. Brasseur and J. Nihoul, Eds., NATO ASI, 1–32.
- Katz, E. J., A. Busalacchi, F. Bushnell, F. Gonzalez, L. Gourdeau, M. McPhaden, and J. Picaut, 1995: A comparison of the ocean surface height by satellite altimeter, mooring, and inverted echo sounder. *J. Geophys. Res.*, **100**, 25 101–25 108.
- Kessler, W. S., and B. A. Taft, 1987: Dynamic heights and zonal geostrophic transports in the central tropical Pacific during 1979–84. *J. Phys. Oceanogr.*, **17**, 97–122.
- Kleeman, R., A. M. Moore, and N. R. Smith, 1995: Assimilation of subsurface thermal data into a simple ocean model for the initialization of an intermediate tropical coupled ocean–atmosphere forecast model. *Mon. Wea. Rev.*, **123**, 3103–3113.
- Levitus, S., and T. P. Boyer, 1994: *Temperature*. Vol. 4, *World Ocean Atlas 1994*, NOAA Atlas NESDIS, 129 pp.
- , R. Burgett, and T. P. Boyer, 1994: *Salinity*. Vol. 3, *World Ocean Atlas 1994*, NOAA Atlas NESDIS, 111 pp.
- Lozier, M. S., W. B. Owens, and R. G. Curry, 1995: The climatology of the North Atlantic. *Progress in Oceanography*, Vol. 36, Pergamon, 1–44.
- Lukas, R., and E. Lindstrom, 1991: The mixed layer of the western equatorial Pacific Ocean. *J. Geophys. Res.*, **96** (Suppl.), 3343–3357.
- Malanotte-Rizzoli, P., and E. Tziperman, 1996: *The Oceanographic Data Assimilation Problem: Overview, Motivation and Purposes*. Vol. 17, *Modern Approaches to Data Assimilation in Ocean Modelling*, P. Malanotte-Rizzoli, Ed., Elsevier, 455 pp.
- Moore, A. M., N. S. Cooper, and D. L. T. Anderson, 1987: Initialization and data assimilation in models of the Indian Ocean. *J. Phys. Oceanogr.*, **17**, 1965–1977.
- Philander, S. G. H., 1990: *El Niño, La Niña and the Southern Oscillation*. Academic Press, 289 pp.
- , W. J. Hurlin, and R. C. Pacanowski, 1987: Initial conditions for a general circulation model of tropical oceans. *J. Phys. Oceanogr.*, **17**, 147–157.
- Reverdin, G., D. Cayan, and Y. Kushnir, 1997: Decadal variability of hydrography in the upper northern North Atlantic in 1948–1990. *J. Geophys. Res.*, **102**, 8505–8531.
- Roemmich, D., M. Morris, W. R. Young, and J. R. Donguy, 1994: Fresh equatorial jets. *J. Phys. Oceanogr.*, **24**, 540–558.
- Rosati, A., R. Gudgel, and K. Miyakoda, 1995: Decadal analysis produced from an ocean data assimilation system. *Mon. Wea. Rev.*, **123**, 2206–2228.
- , —, and —, 1996: *Global Ocean Data Assimilation System*. Vol. 17, *Modern Approaches to Data Assimilation in Ocean Modelling*, P. Malanotte-Rizzoli, Ed., Elsevier.
- Sprintall, J., and M. J. McPhaden, 1994: Surface layer variations observed in multiyear time series measurements from the western equatorial Pacific. *J. Geophys. Res.*, **99**, 963–979.
- Stommel, H. S., 1947: Note on the use of the  $T$ – $S$  correlation for the dynamic height anomaly computations. *J. Mar. Res.*, **2**, 85–92.
- Vialard, J., and P. Delecluse, 1998a: An OGCM study for the TOGA decade. Part I: Role of salinity in the physics of the western Pacific fresh pool. *J. Phys. Oceanogr.*, **28**, 1071–1088.
- , and —, 1998b: An OGCM study for the TOGA decade. Part II: Barrier layer formation and variability. *J. Phys. Oceanogr.*, **28**, 1089–1106.
- Vossepoele, F. C., R. W. Reynolds, and L. Miller, 1999: Use of sea level observations to estimate salinity variability in the tropical Pacific. *J. Atmos. Oceanic Technol.*, **16**, 1401–1415.
- Wolff, J., E. Maier-Reimer, and S. Legutke, 1997: The Hamburg Ocean Primitive Equation Model. Tech. Rep. 13, Deutsches KlimaRechenZentrum, Hamburg, Germany.
- Woodgate, R. A., 1997: Can we assimilate temperature data alone into a full equation of state model? *Ocean Modelling*, **114**, 4–5.