

# Ocean Interpolation by Four-Dimensional Weighted Least Squares—Application to the Waters around Australasia

K. R. RIDGWAY AND J. R. DUNN

*CSIRO Marine Research, Hobart, Tasmania, Australia*

J. L. WILKIN

*Institute of Marine and Coastal Studies, Rutgers University, New Brunswick, New Jersey*

(Manuscript received 15 June 2001, in final form 29 January 2002)

## ABSTRACT

A new four-dimensional ocean interpolation system based on locally weighted least squares fitting is presented. A loess filter is used to interpolate irregularly spaced data onto a uniform grid. This involves projecting the data onto quadratic functions of latitude and longitude while simultaneously fitting annual and semiannual harmonics by weighted least squares. The smoothness scale of the mapping method adapts to match the data density, thus producing gridded estimates with maximum resolution. The filter has a vertical dimension, such that the data on adjacent levels are included in the computation. This greatly reduces the effects of discontinuities in data distributions between adjacent levels, since the estimates at each level are no longer independent. The loess scheme has been further modified so that the weighting of data points is adjusted to allow for the influence of both bathymetry and land barriers. This allows the bathymetry to influence the mapped fields in a natural way, reduces leakage of structure between deep and shallow regions and produces far more realistic coastal gradients. The flexibility of the loess approach has allowed further adjustments to compensate for irregularities in spatial and temporal sampling. The mapping is shown to be statistically consistent with an objective measure of the a priori noise of the dataset. Departures of the mapped fields from independent surface temperature climatologies and mean vertical sections derived from withheld expendable bathythermograph (XBT) data are within error limits.

The method is applied to the major seas around Australia, New Zealand, Papua New Guinea, and Indonesia (50°S–10°N, 100°E–180°) to form a high-resolution seasonal climatology of temperature, salinity, oxygen, nitrate, phosphate, and silicate, referred to as the CSIRO (Commonwealth Scientific and Industrial Research Organisation) Atlas of Regional Seas (CARS). Stringent quality control procedures have been applied to a comprehensive dataset assembled from all known sources. The resulting maps successfully resolve both the large-scale structure and narrow coastal features and illustrate how the bathymetry influences the property distributions.

## 1. Introduction

The calculation of mean fields from large sets of historical data is a major task in oceanographic data analysis. The resulting regularly gridded maps are a valuable reference tool for characterizing ocean regions, providing a background field for climate studies, and for validating model results. A clear example of this is given by the widespread adoption of the *World Ocean Atlas* (Levitus 1982) as a standard reference. However, the irregular sampling density and accuracy of ocean observations, and lack of statistical stationarity, generally makes the production of maps a difficult exercise. Furthermore, where coastal geometry and bathymetry are complex, many of the commonly used interpolation

methods are not capable of obtaining realistic gridded fields (Brasseur et al. 1996; Dunn and Ridgway 2001, hereafter DR).

Many existing ocean climatologies are designed for resolving basinwide scales and hence are highly smoothed (Levitus 1982). Consequently, they are not capable of resolving boundary currents, frontal systems, eddy fields, and other permanent features with small spatial scales. New high-resolution observations from satellite platforms and output from general circulation models demonstrate that the spatial structure of the mean flow is influential down to the mesoscale thus, existing climatologies are clearly inadequate (Walker and Wilkin 1998; Roemmich and Sutton 1998; Webb 2000).

In this paper, we present an interpolation system that seeks to address many of these data shortcomings and regional complexities. When applied to an ocean region it provides mean fields that resolve both the large-scale structure and narrow coastal fronts and currents. Our

---

Corresponding author address: Dr. K. R. Ridgway, CSIRO Marine Research, GPO Box 1538, Hobart, TAS 7001, Australia.  
E-mail: ken.ridgway@marine.csiro.au

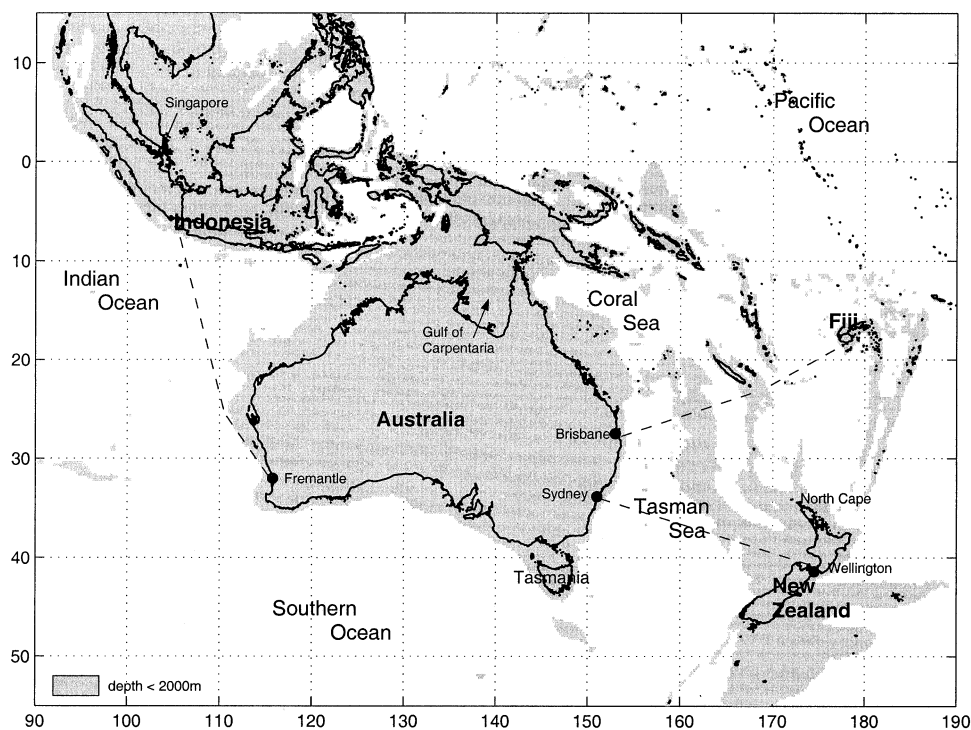


FIG. 1. The ocean region (50°S–10°N, 100°E–170°W) over which the new mapping methods have been applied. All geographical features, both on land and sea, are included in the figure. The three dashed lines show the location of regularly occupied XBT sections.

system is built around the weighted least squares quadratic or loess smoother of Cleveland and Devlin (1988). The computational demands of the method are less than other popular approaches such as Gauss–Markov estimation, yet the filtering characteristics are nearly as good (Chelton and Schlax 1994). The interpolation simultaneously fits seasonal terms along with the spatial components that greatly minimizes the temporal bias in the mean. The scheme also adjusts the weighting of data points to allow for the influence of both bathymetry and land barriers. This reduces leakage of structure between deep and shallow regions and produces far more realistic coastal gradients.

We demonstrate the components of the system within a case study covering the seas around the Australian continent. These waters contain many dynamically interesting and often unique features. In the Tasman Sea off eastern Australia, the East Australian Current (EAC) is a major western boundary current, with highly energetic mesoscale eddies associated with its poleward flow (Nilsson and Cresswell 1981; Mulhearn 1987). The Indonesian islands to the north act as a permeable barrier to flow from the Pacific to the Indian Oceans, thus playing a central role in the redistribution of mass and heat in the global system (Godfrey and Golding 1981). Furthermore, the very existence of the unique dynamics of the Leeuwin Current flowing poleward along the western Australian coast has only relatively recently been recognized (Cresswell and Golding 1980).

We describe procedures to assemble a complete in situ dataset for the region and the techniques that have been applied for estimating a gridded climatology. This climatology is entitled CSIRO (Commonwealth Scientific and Industrial Research Organisation) Atlas of Regional Seas (CARS). In section 2 we present the data used in the analysis, and detail the associated quality control methods. The loess methodology is described in section 3, including the topographic adjustments, and in section 4 we specify how the system is applied to the dataset. A description of sampling problems encountered and how they have been addressed is found in section 5. Finally the results are given in section 6. These include an analysis of the residuals, validation against independent data and an example of the mean fields.

## 2. Data

An archive of 65 000 vertical profiles of temperature, salinity, oxygen, phosphate, nitrate, and silicate has been assembled for the seas adjacent to Australia. Our mapping domain is (10°N–50°S, 100°E–180°, Fig. 1) although the mapping procedure requires data to be drawn from a slightly wider region (Fig. 2). These data were primarily obtained from the World Ocean Database (WOD98; Conkright et al. 1998), the Global Oceanographic Data Archaeology Rescue project (Levitus et al. 1994), CSIRO Marine Research archives, and the

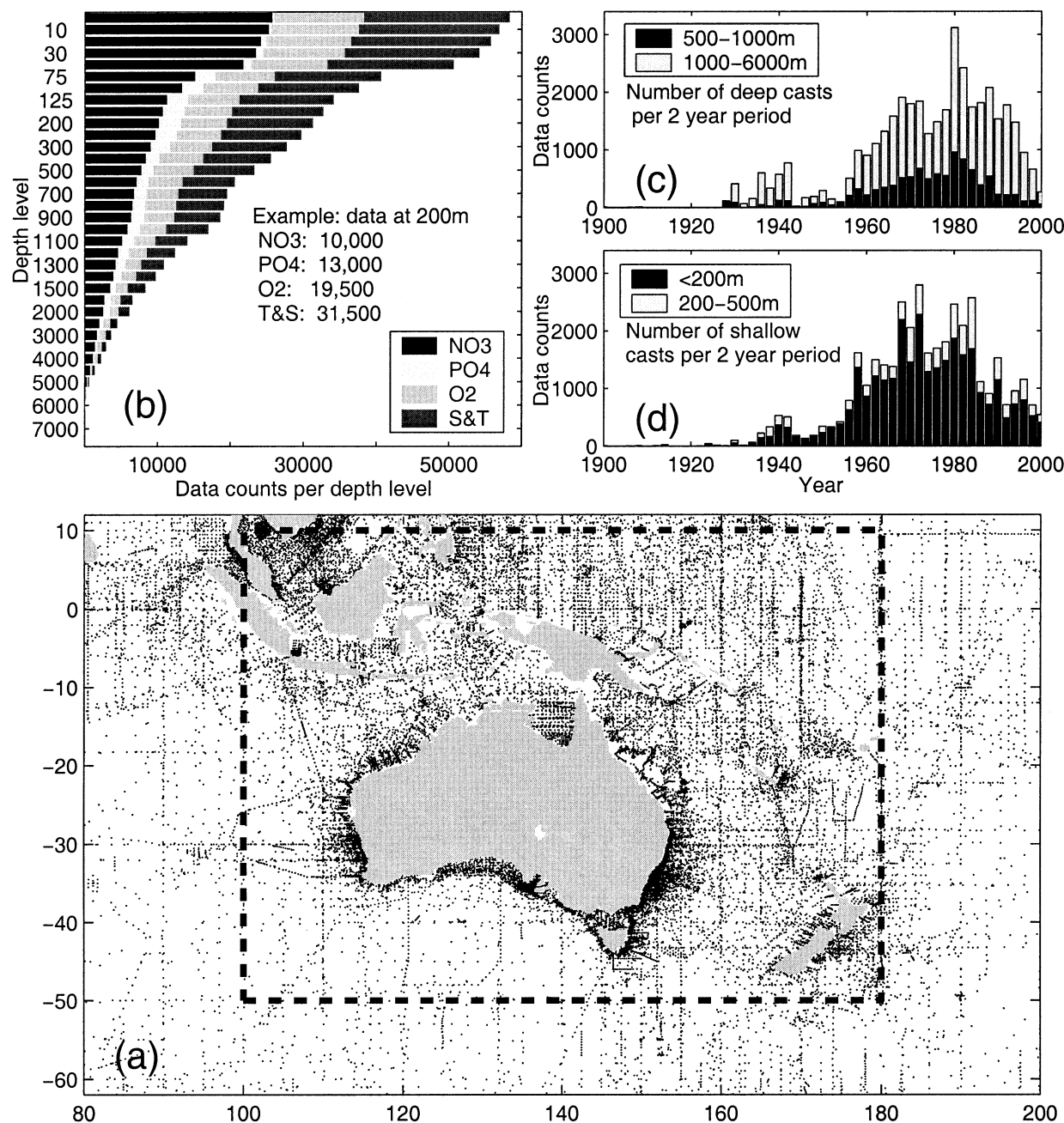


FIG. 2. The distribution of (a) the ( $T$ ,  $S$ ) casts as a function of position, (b) each property as a function of depth, (c) and (d) the ( $T$ ,  $S$ ) casts as a function of year at several depth ranges. The actual domain of the CARS climatology is shown by the dashed rectangle in (a).

New Zealand Institute of Water and Atmospheric Research. We have used both point-sampled stations (Nansen and Niskin bottle) and continuously sampled traces, obtained from conductivity–temperature–depth probes.

All observed level casts were interpolated onto a set of 56 “standard depth levels” (see Table 1) using an algorithm based on that of Reiniger and Ross (1968). This uses a weighted mean of values obtained from the linear interpolation of the two nearest data points, and the extrapolation of the pairs of data above and below.

Using these standard levels we are able to account for most of the structure in the surface mixed layer, the thermocline, and the deep water region.

Since data have been obtained from several data centers, many duplicates have been identified and removed. We also discovered numerous casts with erroneous positions and/or dates. In particular, casts with incorrect positions were often identified when their bottom depths were clearly greater than suggested by the local bathymetry. We used the ETOPO5 (Earth Topography—

TABLE 1. Standard depth levels used in the climatology.

Level	Depth (m)	Level	Depth (m)
1	0	29	700
2	10	30	750
3	20	31	800
4	30	32	850
5	40	33	900
6	50	34	950
7	60	35	1000
8	70	36	1100
9	75	37	1200
10	80	38	1300
11	90	39	1400
12	100	40	1500
13	110	41	1600
14	125	42	1750
15	150	43	2000
16	175	44	2250
17	200	45	2500
18	225	46	2750
19	250	47	3000
20	275	48	3250
21	300	49	3500
22	350	50	3750
23	400	51	4000
24	450	52	4250
25	500	53	4500
26	550	54	4750
27	600	55	5000
28	650	56	5500

5 minute) bathymetry (NOAA 1988) supplemented by a high-resolution dataset obtained from the Australian Geological Survey Organisation in Australian waters. The bathymetry data were smoothed and the rejection criteria were set such that casts were not rejected readily in regions of large depth gradient. If the indicated cast location was more than 30 km inland, it was deemed to be have been labeled incorrectly. Furthermore casts were also rejected if their documented locations implied that they were within the 200-m isobath, but their indicated depths were greater than 300 m.

Following the removal of these poorly documented casts, the data within each cast were systematically screened. In the case of temperature ( $T$ ) and salinity ( $S$ ) data, a prior screening process was performed. The casts were assessed in  $T$ - $S$  space using a method that shares features with those applied previously (Curry 1996; Gouretski and Jancke 1999). Essentially, outliers from a background  $T$ - $S$  field were identified and rejected (both  $T$  and  $S$  values). The climatology consists of fields of salinity evaluated at a series of  $0.5^\circ\text{C}$  temperature levels from  $-2.5^\circ$  to  $31^\circ\text{C}$ . The salinity is obtained on a  $1^\circ \times 1^\circ$  grid using the locally weighted methodology described in the following section. All casts lying outside a three standard deviation ( $\sigma$ ) range from the mean curve on each  $T$  level were ejected. In practice, the entire climatology and statistic fields were generated in an initial calculation, the variance profiles were vertically smoothed, and data with residuals outside  $3\sigma$  were

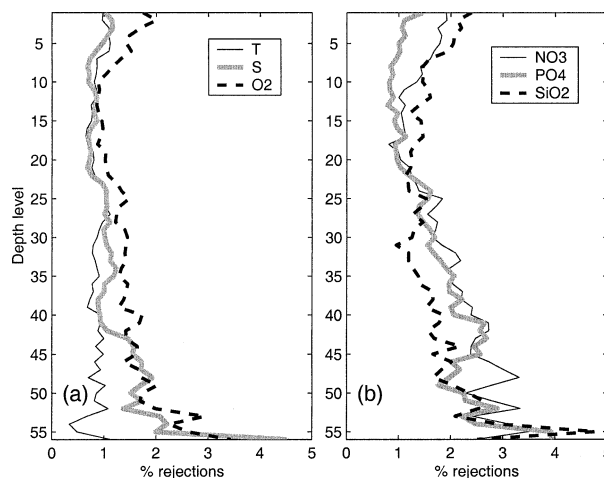


FIG. 3. The percentage of each property type rejected by the screening procedures plotted against depth level. (a) Temperature, salinity, and oxygen. (b) Nitrate, phosphate, and silicate.

flagged. Then the process was repeated, providing both an improved climatology and tighter statistics.

Next climatological fields of each property were obtained on a set of standard depth levels using the methods described in following sections. Residuals of each property were then obtained and data were again discarded when exceeding a specified multiple of  $\sigma$ . In this case, the multiple was set to 2.6, which for a normal distribution represents about 1% of the data. However, in a few regions perhaps containing whole cruises with data offsets or a localized freshwater runoff, the rejection rate was as high as 2.5%. For example, off western New Zealand, salinity values  $<20$  psu located in the deep-water fiords were rejected. In addition, casts with a high proportion of flagged data were rejected in their entirety. The overall percentage of data rejected is plotted against depth in Fig. 3. Note that higher proportions of the nutrients have been rejected and that the rejected rate tends to increase with depth.

The distributions of the data in space, time, depth, and property after quality control procedures have been applied are presented in Fig. 2.

### 3. Loess mapping

The diverse environment of the Australian region and the sampling shortcomings provide a challenge to any interpolation scheme and the loess filter is no exception. We have, therefore, developed simple extensions to the loess methodology, which substantially improve the outcomes if not completely addressing the above factors. The flexibility inherent in the loess method provides clear advantages over other interpolation schemes.

We interpolate our irregularly distributed dataset onto our chosen spatial grid by applying a space-time version of locally weighted least squares or "loess mapping" (Cleveland and Devlin 1988). The data are smoothed in



space by projecting onto spatial quadratic functions and simultaneously being fitted by annual and semiannual harmonic components. Fitting the spatial and temporal components in a single step minimizes the temporal bias in the mean. The temporal terms were only applied in the upper 1000 m.

We wish to determine the climatological estimate of a state variable  $\hat{\phi}(x_n, y_n, z_s, t_n)$  at some grid point  $(x_n, y_n)$  with standard depth  $z_s$  and time  $t_n$ . This is achieved by a weighted least squares fit to the  $K$  nearest data points to the grid point, of a four-dimensional surface defined by the following spatial quadratics, and temporal harmonics

$$\begin{aligned} \hat{\phi}(x_n, y_n, z_s, t) \\ = a_0 + a_1x + a_2y + a_3x^2 + a_4y^2 + a_5xy + a_6z \\ + a_7z^2 + \sum_{j=1}^4 a_{j+7}X_j(t), \end{aligned} \quad (1)$$

where  $x_k, y_k$ , and  $z_k$  are the longitude, latitude, and depth of a data point,  $x = x_k - x_n$ ,  $y = y_k - y_n$ ,  $z = z_k - z_s$ ,  $X_1 = \cos T$ ,  $X_2 = \sin T$ ,  $X_3 = \cos 2T$ ,  $X_4 = \sin 2T$ ,  $T = 2\pi t/365.25$ , and  $t$  is the day of year. Including further terms involving  $x, y$ , and  $t$  in (1) at best provides only marginal improvements and often leads to unrealistic “overfitting” of the data. The regression coefficients  $a_n$  in (1) are determined by minimizing the weighted sum of square errors

$$\Psi = \sum_{k=1}^K w_k^2 (\hat{\phi}_k - \phi_k)^2, \quad (2)$$

where  $\phi_k = \phi(x_k, y_k, z_s, t)$ . The weighting coefficients  $w_k$ , are defined by the tricubic function

$$w_k = \begin{cases} (1 - r^3)^3 & 0 < r < 1 \\ 0 & r > 1, \end{cases}$$

where  $r$  is the normalized distance metric, which defines the separation of observations and estimation (grid) point, and is defined as

$$\begin{aligned} r(x_k, y_k, z_k, d_k, d_n) = & \left[ \left( \frac{\beta(x_k - x_n, y_k - y_n)}{R} \right)^2 \right. \\ & \left. + \left( \frac{z_k - z_s}{R_z} \right)^2 + \left( \frac{r_b}{R_b} \right)^2 \right]^{1/2}, \end{aligned} \quad (3)$$

where  $R$  defines the horizontal radius of the data ellipse and  $\beta$  is a function that defines the adjustment of the pathways between observation and grid point due to land barriers. Furthermore  $R_z$  is the maximum vertical radius and  $r_b/R_b$  is a normalized bathymetry–distance function, which is dependent on the bottom depth at the data location ( $D_k$ ), and the bottom depth at the grid location ( $D_n$ ).

### a. 3D mapping

The terms in (1) containing the depth  $z$  and the second term in (3) have been included to minimize vertical discontinuities between fields due to major differences in data distribution between levels. In such cases we often obtain discontinuities in the temporal harmonics which can, for example, be manifested as unrealistic inversions in the temperature profiles. Solving for all depth levels simultaneously along with the spatial and temporal terms would minimize this behavior; however, this is beyond our computational resources. Therefore a more manageable scheme has been implemented, in which the data on adjacent levels are included after being downweighted appropriately. The vertical normalization radius  $R_z$  in (3) is chosen so that data points above and below the standard depth  $z_s$  have equal weight but are reduced by a factor of 0.5 times the weight of data actually on  $z_s$ . For example when  $z_s = 125$  m, then  $z_{s-1} = 110$  m, and  $z_{s+1} = 150$  m. Here,  $R_z$  is then an envelope 30 m above and 50 m below  $z_s$ .

### b. Topographic adjustment

The scheme to adjust the weighting of casts for the influence of varying bottom depth is called topographic adjusted relief (TAR). The final term in (3) is the bathymetry–distance function, which is used to minimize the contamination of deep oceanic regions by highly variable waters on the continental shelf. The normalized function has the form

$$\begin{aligned} \frac{r_b}{R_b} = 1 - & \left[ \max \left( 0, 1 - \mu \frac{D_k}{D_n} \right) \right. \\ & \left. - \max \left( 0, 1 - \lambda \frac{D_k}{D_n} \right) \right], \end{aligned} \quad (4)$$

where  $\lambda = 4$  controls the rate of inshore cutoff and  $\mu = 0.01$  controls the rate of offshore cutoff. These depths are further restricted to a threshold range of 25–2000 m. Figure 4 shows the form of this function at a range of bottom depths. Several examples of the improved results obtained using the TAR scheme are presented in DR. For example, in the waters adjacent to the South Island of New Zealand, the use of TAR has both reduced the spread of low salinity water into the deep ocean southwest of the South Island and also sharpened the coastal salinity gradient. Comparing the results with individual data points showed that these changes have produced far more realistic fields.

The function  $\beta$  adjusts data-point weighting to allow for the influence of land barriers and is a distinctive feature of the CARS methodology. It is designed particularly to reduce inappropriate leakage of structure across narrow headlands and islands. The method, which is termed barrier adjusted relief (BAR), represents the domain as a network of polygons containing subsets

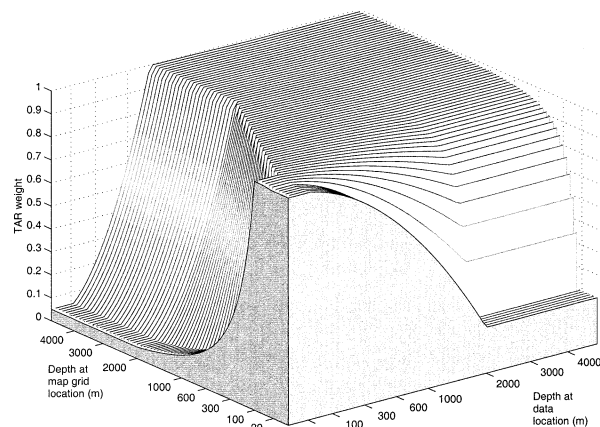


FIG. 4. The bathymetry–distance function showing its dependence on gridpoint depth and data depth. Note that when the gridpoint and data-point depths are similar the weights are close to unity and when they diverge the weights tend to a minimum value. For example, as the gridpoint depth increases, data points in shallow water are quickly downweighted to lessen the contamination of deep-water maps by highly variable coastal waters.

of the data, and relationships that are specified between pairs of these polygons. These relationships may significantly modify the distance functions between data and grid points in the vicinity of complex topography. Dunn and Ridgway. (2001) demonstrate that using the BAR system all but eliminates the unrealistic propagation of ocean properties directly across land barriers. In particular, erroneous features are removed from the interpolated fields for the waters around New Zealand and within the Indonesian archipelago. Figure 5 shows how the distribution of weights associated with a grid-point, change significantly when the TAR and BAR schemes are applied.

#### 4. Application to hydrology data

##### a. Resolution of the mean fields

The climatology is produced for the study region on a uniform  $0.5^\circ$  grid. This provides appropriate resolution for the deep ocean basins but is not sufficient to delineate the finescale structure at the coastal boundaries. Accordingly, we also generate supplementary higher-resolution maps ( $1/8^\circ$ ) adjacent to these boundary regions.

The actual resolution of the mean fields produced in our analysis is dependent on the choice of the data ellipse radius  $R$ . Following Cleveland and Devlin (1988) we allow  $R$  to vary with grid point by fixing the number of points  $N$  used for each estimate ( $N = 400$ ). This adjusts the effective bandwidth of the loess smoother to match data density. While enabling us to produce gridded estimates with maximum spatial resolution, this feature does tend to obscure the actual smoothing scale since the degree of smoothing of the loess filter is also defined by the magnitude of  $R$ . A lower limit of 200

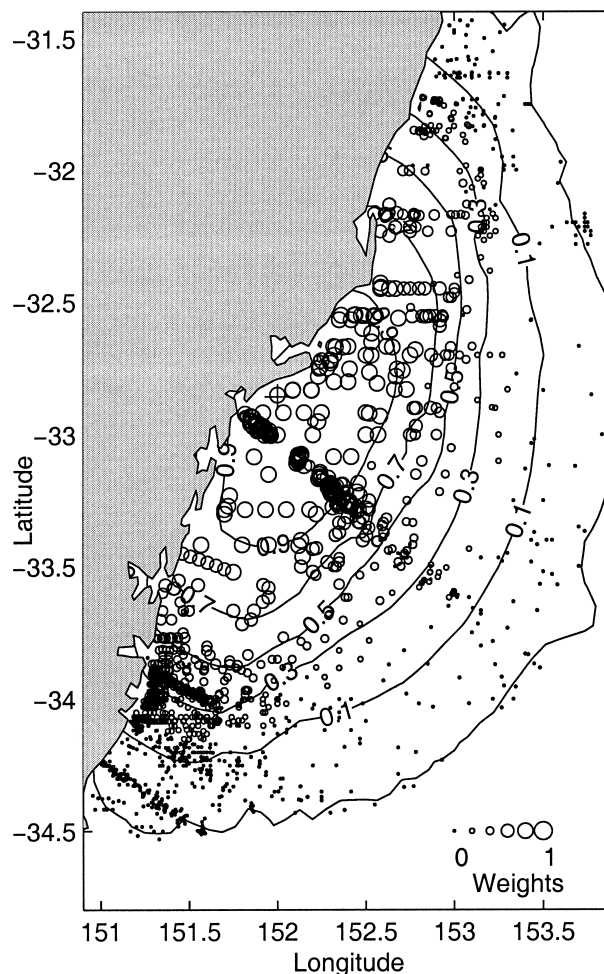


FIG. 5. The spatial distribution of weights allocated to the data casts used in the calculation of the mean salinity at the 20-m level for the indicated grid point (water depth 45 m). The weights for individual casts are represented by the size of the circles.

km is also imposed on  $R$  so that mesoscale eddy fluctuations are appropriately smoothed in regions of abundant data. This means that in such regions  $N$  is correspondingly larger. For example in Fig. 5 the grid point off southeast Australia includes some 2500 casts within the associated data ellipse. If  $N$  is reduced to say 300 casts, then the magnitude of the resolution decreases by about 10%; however, there is a noticeable increase in the amount of noise in the mean fields. In data sparse cases the 400-cast criterion is overridden and instead an upper limit is imposed on the source radius (1500 km), so that the implied length scale remains meaningful. In the rare cases where the number of data points falls below a minimum value, then no mapping is performed. A value at this grid point is then obtained by interpolation between adjacent grid points.

The resolution of the mean fields will vary over the domain since the shape and size of the data ellipse will tend to change slightly for each grid point. Hence, we



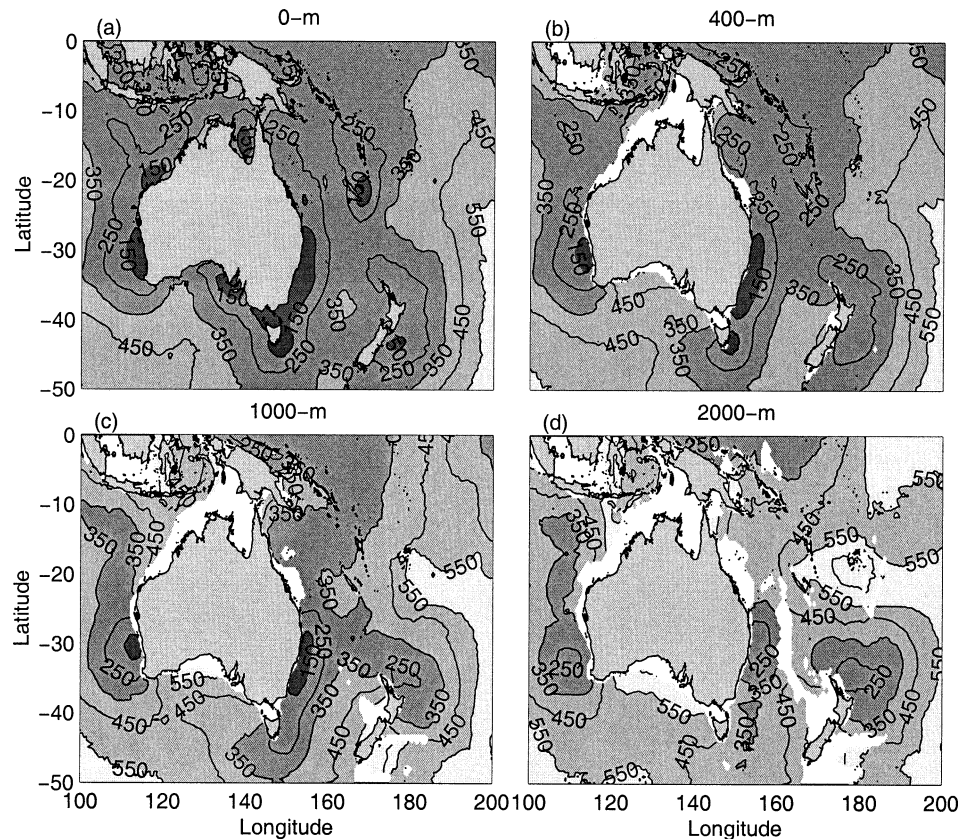


FIG. 6. The spatial resolution at (a) the surface, (b) 400-, (c) 1000-, and (d) 2000-m levels for  $T$  and  $S$  mapping. The quantity presented is the half-amplitude cutoff wavenumber ( $0.6R$ ), where  $R$  is the data ellipse radius, since this is approximately the same as that of a simple block average of the same width (Chelton et al. 1990).

have some difficulty in representing the effective resolution of any property over the whole domain. In Fig. 6, the variation in spatial resolution of the mean temperature is presented at several depths. This quantity is the half-amplitude cutoff wavenumber ( $0.6R$ ), where  $R$  is the data ellipse radius, since it approximates a simple block average of the same width (Chelton et al. 1990). At the surface the resolution ranges from 110 km at the Australian meridional boundaries to more than 550 km in the southwest corner of our domain, where data are sparse in the Southern Ocean. Furthermore,  $R$  is modulated significantly by the inclusion of the bathymetry and barrier adjustment. For grid points at a coastal boundary, the bathymetry dependency means that data ellipses will both be aligned along isobaths and stretched in the alongshore direction. Thus the interpolation will have its greatest resolution across the continental slope (DR). The effective resolution is thus a function of data density, local bathymetry, and geometry, and in general will be nonisotropic. The data ellipse in Fig. 5 associated with a grid point adjacent to the east Australian coast has an alongshore length double its cross-shelf length.

The loess formalism also allows us to determine a

locally weighted analog of the standard deviation of the observations from the CARS mean climatology

$$\sigma^2 = \frac{\sum_{k=1}^K w_k^2 (\phi_k - \bar{\phi}_k)^2}{\sum_{k=1}^K w_k}, \quad (5)$$

where  $\bar{\phi}_k = \phi(x_k, y_k)$  is the CARS mean value at  $(x_k, y_k)$ .

#### b. *A priori* noise

To determine whether the estimated mean fields fit the data at the appropriate scales, we estimate the *a priori* noise associated with each property. This is the small-scale variability arising from both unresolved ocean processes such as mesoscale eddies and internal waves as well as measurement errors. Following Holbrook and Bindoff (2000), we estimate this noise variance from the difference between properties measured at neighboring casts.

We have calculated the noise variance from the data

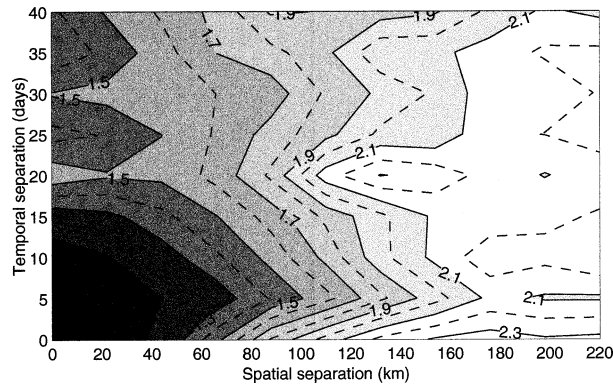


FIG. 7. Space-time distribution of the a priori noise variance of temperature for the Tasman Sea ( $40^{\circ}$ – $28^{\circ}$ S,  $152^{\circ}$ – $156^{\circ}$ E).

in a region adjacent to eastern Australia for a range of space-time scales to choose an appropriate spatial and temporal window (Fig. 7). We require the noise estimate that actually represents the combined effects of transient eddy processes and instrumental errors rather than the long-term, larger-scale signals that we wish to resolve. The results in Fig. 7 suggest that we choose casts that are separated by less than 100 km and 20 days. Further testing in other regions provides a more general definition that incorporates the variable resolution of the mapping procedure. Hence we define the a priori noise at any grid point as being casts that are separated by less than  $0.6R$  and 20 days. The value of  $0.6R$  is usually close to 100 km, but it does fluctuate slightly below or somewhat more above this figure depending on the data distribution.

The data does not allow seasonal and long-term trends in the a priori noise to be resolved. However, we are able to determine the spatial variation of the noise statistics over the study domain using the loess methodology. This approach improves on the binned-mean results obtained by Holbrook and Bindoff (2000). If  $\phi_r$  and  $\phi_s$  are observations of some property at cast  $r$ , and a neighboring cast  $s$ , then the a priori noise at depth  $j$  is

$$\sigma_j^2 = \frac{\sum_{k=1}^K w_k^2 (\phi_r - \phi_s)_k^2}{2 \sum_{k=1}^K w_k}, \quad (6)$$

where  $(\phi_r - \phi_s)_k$  is the  $k$ th difference between pairs of casts separated by some distance  $L$  and within  $N$  days.

## 5. Sampling issues

### a. Spatial-sampling problems

Initial results showed artifacts of the inhomogeneous spatial distribution of our dataset in the estimated fields. For example, the presence of close-packed clusters of data has several undesired outcomes. Such clusters tend to allow inappropriately short spatial scales in areas that

may have little data except the cluster. If a cluster is near the grid point, it will tend to dominate that mapping. More often the source radius will expand until it just touches the cluster and satisfies the data threshold. This means that all the data in the cluster, being at the maximum radius, will have a near-zero weight while the little remaining data will have an inappropriately large influence on the mapping.

We define a cluster to occur when more than 10 values are located within a radius of 250 m. To reduce the undesired effects above, clusters were removed before mapping, and replaced by monthly averages. For example, a cluster consisting of 50 data points in January, would be replaced by one point having the mean value, date, and position of all of those data. A cluster of 11 data points that occurred in different months would be replaced by 11 mean positions, values, and dates and so would be unchanged. This approach is somewhat akin to the “superobservation” procedure that is used in optimal interpolation schemes (e.g., Robinson and Leslie 1985).

The loess scheme tends to become unstable when there are gaps in the domain of any of the fitted functions. This often occurs when interpolating across large data voids or when extrapolating over much shorter distances adjacent to coastal boundaries. To compensate for such cases, prior to the interpolation we create an artificial (or “bogus”) data point at the grid point of the mapping. The bogus value is chosen to be a single data point of full weight, placed at the grid point, with all temporal coefficients set to zero, whose value is a distance-weighted mean of the nearest 10% of all data points in the ellipse. The end result is not unlike that found in optimal interpolation where the mapping scheme reverts to a background mean field where data are scarce. The technique is very efficient and it appears to be very successful in minimizing the instability. It was applied globally despite only being necessary and influential in data sparse regions. In Fig. 8 we test the scheme by first removing two blocks of data, one offshore and another adjacent to the coast, and then mapping the residue both with and without bogusing. Using the field obtained from the complete dataset as a comparison, we observe that the bogusing scheme has generally compensated for the removal of data and produced relatively smooth contours through the gaps (Fig. 8b). In contrast, in the case without bogusing the field in the offshore bin shows unrealistic small-scale structure while the coastal region contains an erroneous salinity gradient (Fig. 8c).

### b. Temporal-sampling problems

In most regions the use of a relatively high data threshold enables us to resolve the seasonal cycle. This is confirmed to some degree by the internal consistency of spatial patterns of amplitude and phase at each standard depth. However, in a small number of cases, gaps



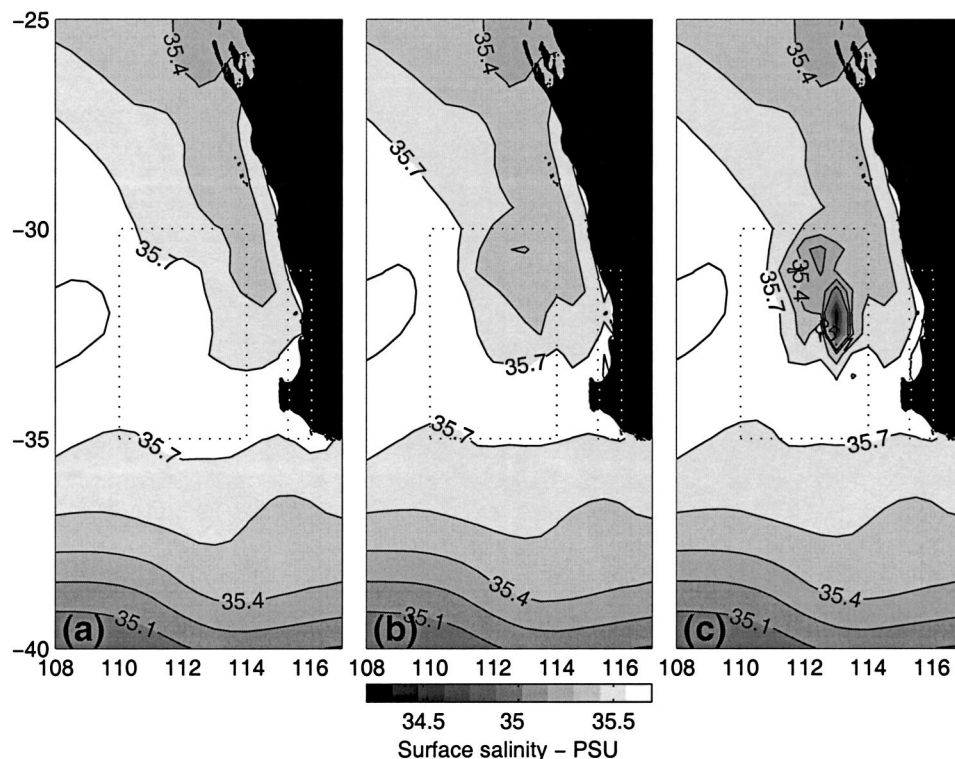


FIG. 8. (a) The surface salinity obtained using the complete dataset, (b) the field obtained when bogusing was used, and (c) the result obtained without bogusing. All the data in the rectangular region (shown by dotted lines) were removed in (b) and (c).

in the temporal distribution tend to produce unrealistic temporal harmonics. For example, at specific times of the year we have observed unrealistic temperature inversions in the vertical profiles. These occur in regions where the data distributions at adjacent levels diverge widely and induce vertical discontinuities in the seasonal cycles. Only a small phase difference between levels is required to produce a relatively large apparent temperature inversion. These features typically arise at higher latitudes in winter, where data gaps often occur (e.g., in the vicinity of the subtropical convergence between Tasmania and New Zealand).

At grid points containing such artificial inversions, a vertical temperature profile is generated for each month, and the inversions removed by iteratively smoothing toward the mean of the inverted sections of the profiles. Annual and semiannual harmonics are then fitted to these monthly values, but they tend to conform very strongly to all the nonadjusted values as they lie on the perfect harmonics from which they were originally derived. To overcome this an increased weight is applied to any values that have been adjusted. Note that these procedures are designed to only remove unrealistic temperature inversions. In regions such as the Antarctic Circumpolar Current where genuine inversions exist, the procedures have little or no influence on the temperature profiles. An alternative and more rigorous approach would be to use the static stability restoration algorithm

of Jackett and McDougall (1995) and this will be implemented in future versions of the atlas.

When there is a strong interannual signal, irregular temporal sampling may also introduce distortions in both the mean patterns and the seasonal harmonics. A localized example is found in the Gulf of Carpentaria where interannual variability in the regional precipitation induces corresponding large changes in the surface salinity. Applying a correction determined from a composite rainfall index successfully removed distortions in the spatial mean pattern caused by aliasing of the interannual signal.

In the southwest Pacific ( $0^{\circ}$ – $10^{\circ}$ S,  $165^{\circ}$ E– $180^{\circ}$ ), there is a very large El Niño–Southern Oscillation (ENSO) related signal that is manifested by a shoaling of the thermocline during an El Niño period. This is typically observed as a major decrease in the sea level over the whole region (Ridgway et al. 1993). A coupling of the episodic nature of this phenomenon with the irregular temporal sampling pattern produces major spatial aliasing of the mean fields and seasonal estimates. We address this problem to some extent by applying a correction derived from the Southern Oscillation index (SOI) directly to the individual data casts prior to the interpolation stage.

The slope of SOI versus property was obtained at every depth level on a  $1^{\circ}$  grid using locally weighted regression. The initial fields were somewhat noisy and

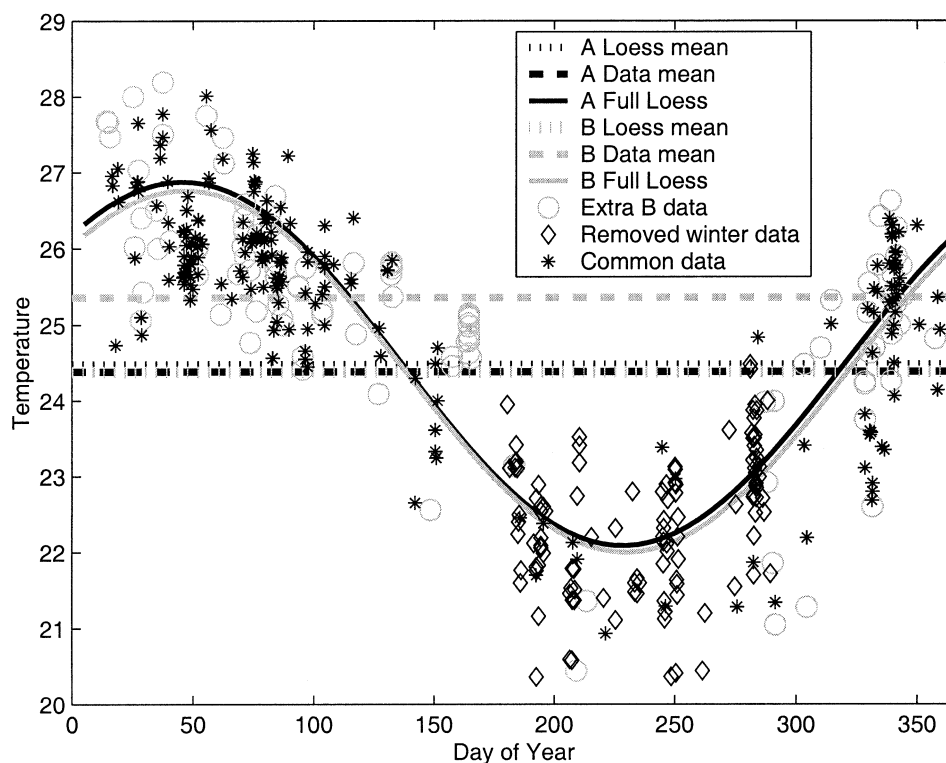


FIG. 9. Surface temperature plotted against day of year for the region ( $28.2^{\circ}$ – $21.85^{\circ}$ S,  $153.5^{\circ}$ – $160.5^{\circ}$ E). The figure legend shows the origin of each symbol and curve plotted in the figure.

a smoothing step was required. From the final form of these fields an SOI-property slope was estimated at each depth, latitude, and longitude. A slope value was then obtained at each cast position and multiplied by the SOI at the lagged cast times to provide a correction to each property value in the region. The resultant property values at each cast were then input to the loess interpolation stage. A more detailed presentation of these correction approaches is in preparation.

This approach is rather simple and has some obvious weaknesses. It is clear that the SOI only accounts for a proportion of the interannual variability in the region. The relationship between SOI and property values is determined from data that are very poorly distributed in space and time. The time lag of effects of ENSO events is likely to vary with depth and geographic location, whereas we have applied a constant value. In some regions it seems ENSO forcing results in modulation of the seasonal cycles, which may not be corrected by a simple linear adjustment. However, despite these problems, we are confident that following correction the maps are much improved and the effects of spatially aliasing has been greatly reduced. We also believe that we have achieved a correction of “interannual sampling bias” in some regions (where a disproportionate amount of the data has been collected in one pole of the ENSO cycle).

## 6. Results

The central outcome of any interpolation procedure is that the residuals from the original data should be both unbiased and also statistically consistent with the a priori estimate of the noise. Our mapping methods are not optimal in a strictly formal sense and hence there is no certainty that these criteria will be satisfied. Instead we need to verify them by analyzing a range of statistics arising from the interpolation.

### a. Fitting of temporal components

Within the upper layers of the water column, interpolation methods that only estimate the spatial mean of the data will contain biases in regions where data are not uniformly distributed by season. Unfortunately, oceanographic data tends to be skewed to periods having the most favorable weather conditions. The negative influence of such irregular data distributions has been greatly diminished within our approach by simultaneously fitting annual and semiannual components while estimating the mean field.

We illustrate this improvement by the example in Fig. 9. Consider the surface temperature within a region in the southern Coral Sea where the data are uniformly distributed in space and time. We first apply the full loess fit [Eq. (2)] for a central grid point ( $25^{\circ}$ S,  $157^{\circ}$ E)

and obtain the mean temperature and seasonal cycle indicated in Fig. 9. The curve closely follows the seasonal pattern of the individual data points. We note that the bin mean temperature is only  $0.1^{\circ}\text{C}$  less than the loess mean showing that the temporal distribution is very nearly uniform. We now remove 90% of the winter data and repeat each calculation. The loess scheme is forced to seek further casts to replace those eliminated resulting in a slightly longer length scale. Since the entire dataset has been winter decimated, most of the new points are in summer. However, the new seasonal curve is very similar to the original one and more importantly the mean is only some  $0.1^{\circ}\text{C}$  cooler than the “true” mean. In contrast, the bin mean is now about  $1.0^{\circ}\text{C}$  above the original value.

We also show how such localized biases may distort the final mapped field and actually reduce the spatial resolution. Figure 10a shows the mean SST for the New Zealand region obtained from the full procedures. The temperature contours reveal a smoothly varying spatial pattern of temperature that faithfully represents the seasonally corrected data. The plotted data were corrected using the harmonic components obtained during the actual computation. A mean field was then generated by only fitting loess spatial quadratics (Fig. 10b). The resulting pattern shows a large amount of small-scale structure that arises from spatial aliasing of the unresolved temporal variability (see the raw data included in the figure). Clearly, to reduce this “noise” in the mapped fields the length scale of the mapping would need to be increased thus reducing the overall spatial resolution of the procedure.

#### *b. A priori noise and analysis of residuals*

The vertical structure of a priori noise estimates, derived from our definition described in section 4, are presented in Fig. 11 for regions in both eastern and western waters. The curves display an interesting and varied vertical pattern. In all cases there are subsurface maxima with the Leeuwin Current region showing double maxima in both the  $T$  and  $S$  results. In the temperature case these maxima arise from eddy and internal wave-induced heaving of the thermocline waters (between 100 and 400 m). Hence the particularly large values within the EAC region due to the high mesoscale variability. We note also that the eastern regions roughly correspond to the subregions used by Holbrook and Bindoff (2000) in their analysis. Although the vertical distribution of the a priori curves in Fig. 11a agree with the previous study, our results are some 20%–30% larger. This is likely to be due to the rather broader spatial and temporal definition of a priori noise employed in our study, compared with that adopted by Holbrook and Bindoff.

The most fundamental test of any mapping procedure is to examine the residuals between the mapped fields and the original data. First, we examine the mean dif-

ference between CARS and the data for a Leeuwin Current and a Tasman Sea region (Figs. 11c,d). The curves show an excellent agreement, with near-zero biases in the data means throughout the water column, apart from small departures in the upper 50 m. The root-mean-square (rms) difference between climatology and data for four representative regions are included in Fig. 11. In general the climatological values are the full seasonal estimates including the mean and both annual and semi-annual components. In each case the rms residual curve lies very close to the corresponding a priori estimates. We have included one example of the annual mean-only CARS for the temperature in the Leeuwin Current region (Fig. 11b). A comparison with the seasonal CARS shows that including the seasonal components has reduced the residuals in the upper 250 m by up to  $0.8^{\circ}\text{C}$ . Similarly, in all other cases the addition of temporal components increased the variance accounted for by the mapped fields.

The overall result is a strong indicator that the residuals, and consequently the mapping, are statistically consistent with our objective measure of the a priori noise. The result is also quite robust, and shows little sensitivity to alternate definitions of the noise. Generally the results show that if anything the residuals tend to be smaller than the a priori noise (see Fig. 11a), which implies that the CARS seasonal mean may be very slightly overfitting the data. However, this is further qualified by examining the rms residuals between CARS and independent expendable bathythermograph (XBT) casts from the WOD98 archives (Fig. 11a). In the Tasman Sea the XBT residuals are larger than those obtained from the original data. In both cases they lie even closer to the a priori curves.

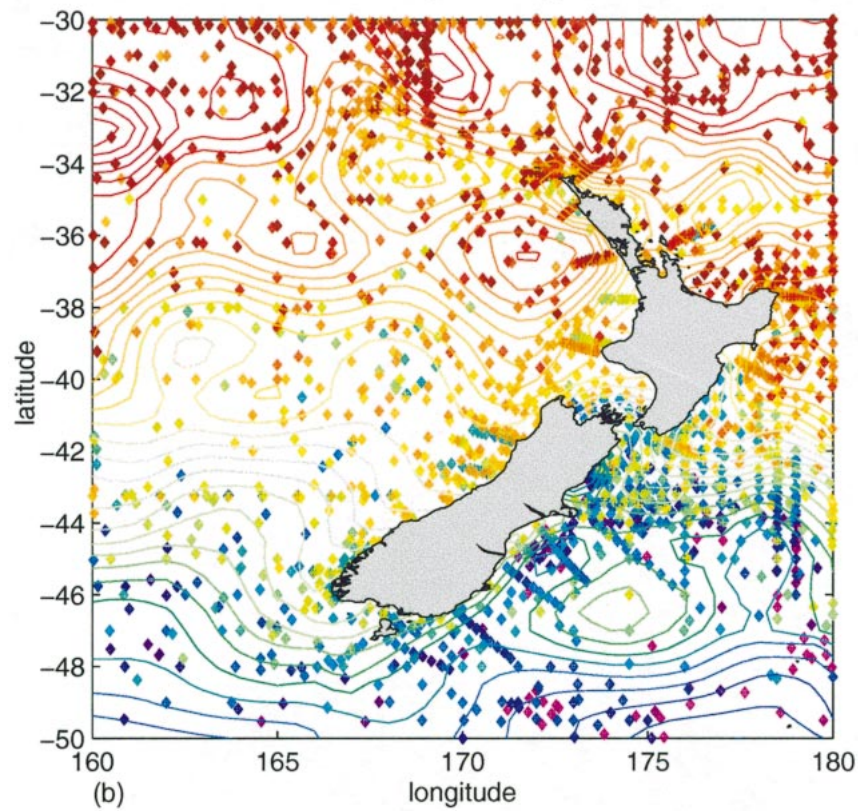
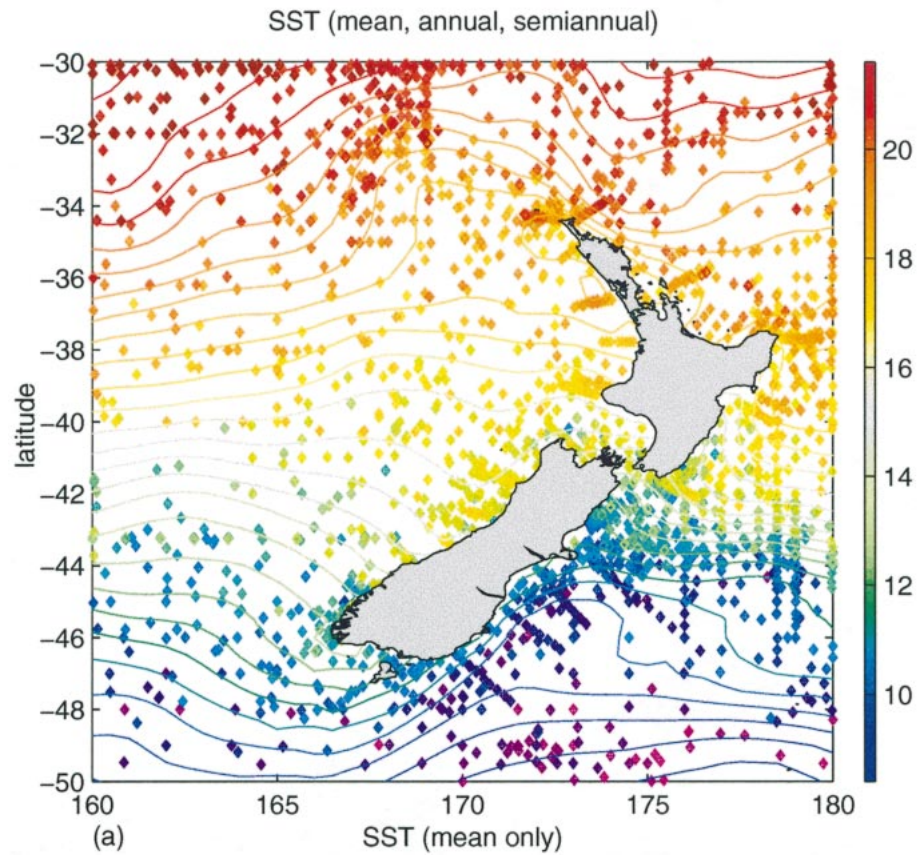
A wider perspective of the statistical consistency of the CARS mapping procedures is provided in Fig. 12. The normalized rms residual of temperature is presented at the surface and 300 m for the Tasman and Coral Seas. The spatial pattern of residuals is rather smoother than the equivalent distribution of noise, hence the patterns in Fig. 12 exhibit some spatial structure. However, the normalized residual lies between 0.5 and 1.3 confirming that the maps are statistically consistent within reasonable limits.

#### *c. Validation against independent data*

Unless climatological fields are supported by ancillary information, they represent nothing more than uncertain content of unknown accuracy (Roemmich and Sutton 1998). Hence we adopt a range of simple validation tests and case studies that verify specific elements of the results. Further validation examples are presented in DR.

The ocean surface is the one segment of the water column where complete, independent realizations of the mean field are available for validation. We use the Reynolds (1988) analysis, which primarily derives from sat-





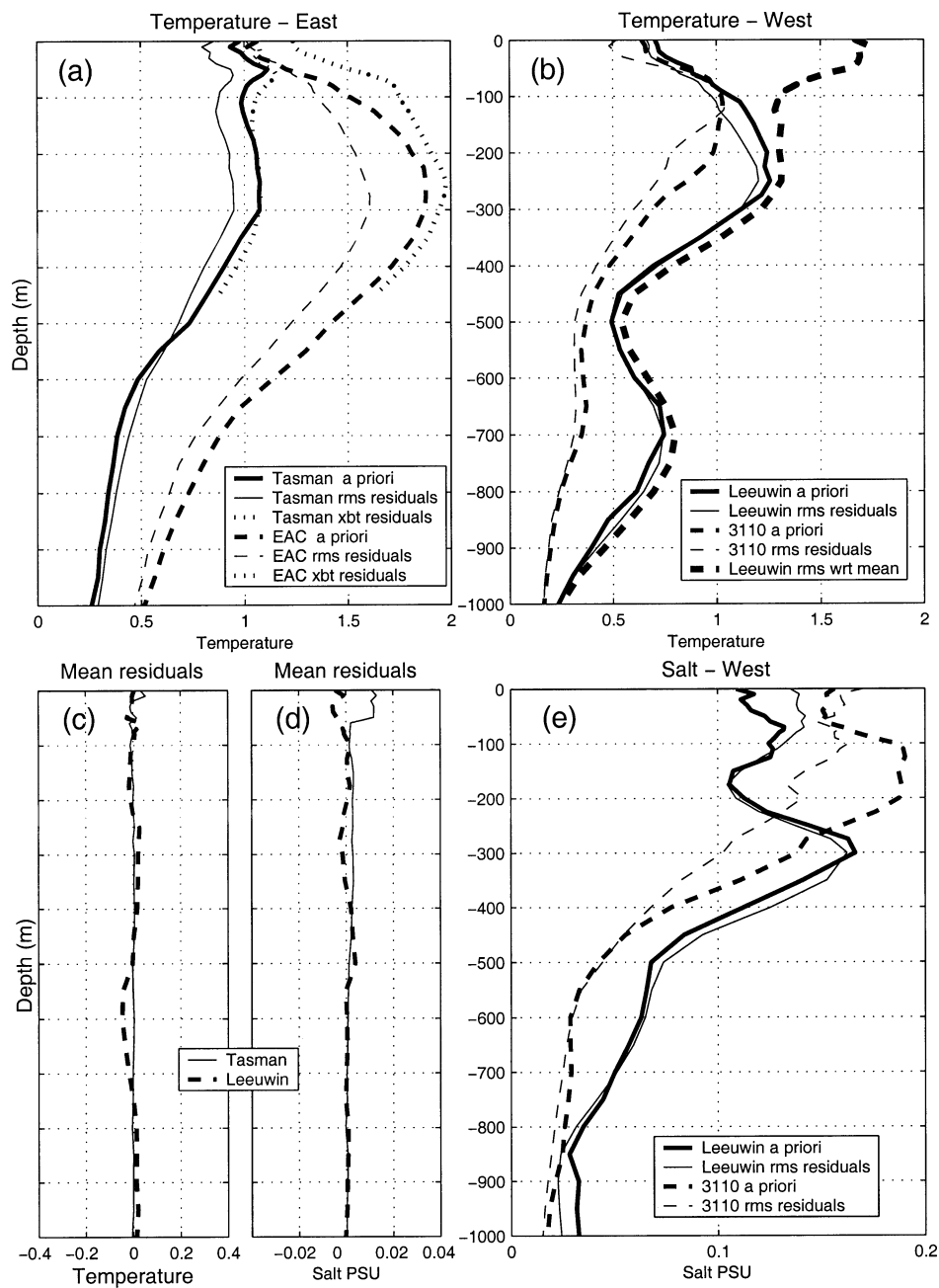


FIG. 11. (a) Rms residuals between the CARS climatology and the temperature data in regions enclosing the EAC and the central Tasman Sea (the two regions are shown in Fig. 12). Further curves show the residual for independent XBT data in each region. (b) As for (a) but for data in a Leeuwin Current region and the WMO square 3110. (c) Mean difference between observed  $T$  and  $S$  and CARS temporal climatology (full line), for the indicated regions. (d) As for (c) using salinity data. (e) As for (b) but using salinity data.

←

FIG. 10. (a) Mean surface temperature for the region ( $50^{\circ}$ – $30^{\circ}$ S,  $160^{\circ}$ E– $180^{\circ}$ ) obtained with simultaneous fitting of seasonal components. The colored diamonds represent the seasonally corrected individual data casts. (b) The mean temperature field obtained for a mean-only calculation. The data casts are now the unadjusted values.



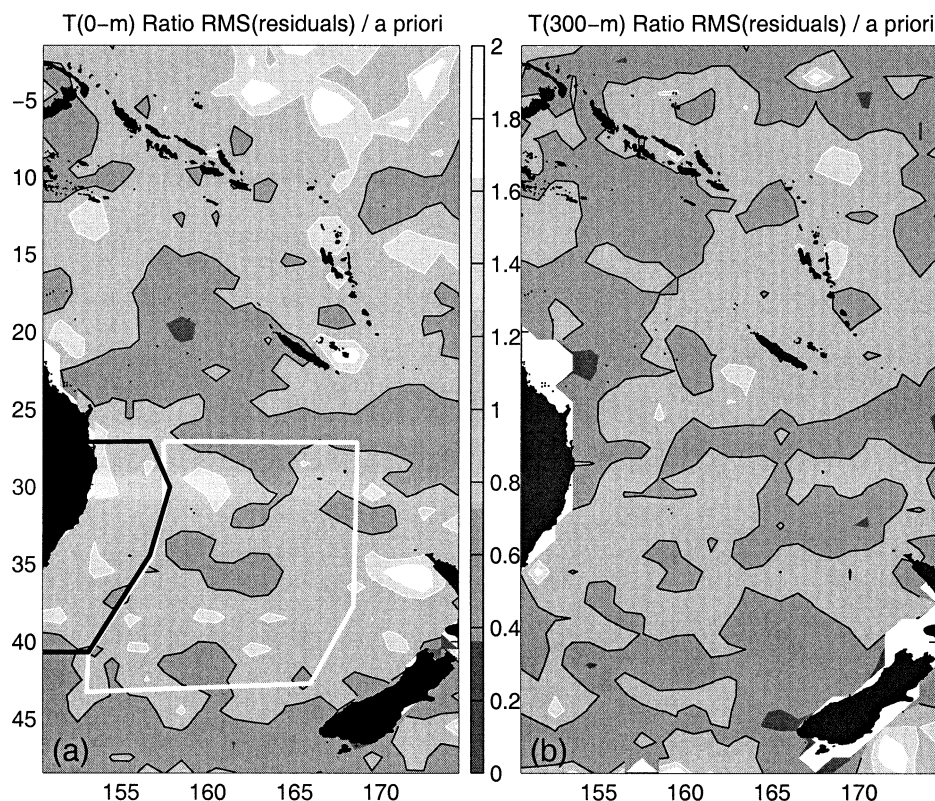


FIG. 12. Rms residuals (normalized by a priori noise estimates) between CARS and temperature data for the Tasman/Coral Sea region at (a) surface and (b) 300-m level. These residuals derive from locally weighted versions of the rms residual and the a priori noise [Eq. (6)].

ellite SST (1981–98) and the Comprehensive Ocean–Atmosphere Data Set (COADS, Woodruff et al. 1993) which is based on a variety of in situ observations (1945–89), but excluding all station casts. In Fig. 13 the Reynolds (1988) mean field is used as the control and the figure shows the departure of CARS and COADS from the Reynolds mean.

Figure 13a shows that CARS is in general agreement with the Reynolds mean, with the differences over the whole region mostly in the  $\pm 0.5^{\circ}\text{C}$  range. Within the difference pattern, there are both coherent large-scale features as well as more intense small-scale structure. The latter may partially be due to sampling shortcomings, however, at least a portion of this difference structure arises from the mismatch of the relatively coarse Reynolds mean with the high resolution of the CARS field. For example, the negative anomalies surrounding New Zealand result directly from the Reynolds fields smearing over the sharp frontal characteristics of the subtropical front, the East Auckland Current, and the Tasman front. Furthermore, both the degree of smoothing and the  $1^{\circ}$  gridding means that boundary currents such as the EAC and the Leeuwin Current are less well resolved in the Reynolds results.

Inspection of Fig. 13a suggests that there is an overall negative bias, which calculation confirms to be

$-0.10^{\circ}\text{C}$ . There is likely to be a contribution to this bias from the tendency for the surface station data to actually represent the temperature just below the surface. In addition the rms difference is only  $0.31^{\circ}\text{C}$ , which is only slightly greater than the value for the COADS–Reynolds difference ( $0.28^{\circ}\text{C}$ ). The most striking feature is in the southwest, where we observe a spatially coherent negative anomaly of up to  $1.0^{\circ}\text{C}$ . There are a few localized anomalies of either sign that may be caused by sampling shortcomings (e.g., the positive anomaly in the central Tasman). We have also produced (but have not shown) comparison plots of the annual amplitude from the respective climatologies. There is consistent agreement between all the products with CARS perhaps showing the most spatial structure. This is confirmed by the results from another high-resolution satellite product (Walker and Wilkin 1998).

A further comparison of CARS SST is now made with in situ observations. These consist of results from three long-term XBT sections (Fremantle–Singapore, Sydney–Wellington, and Brisbane–Fiji), which are shown in Fig. 14. In each of the sections, CARS agrees to within  $\pm 0.25^{\circ}\text{C}$  of the XBT mean, which is also well within the confidence limits. This is the about the same level of agreement as the other climatologies (Reynolds and COADS). That is apart from at the Sydney end of



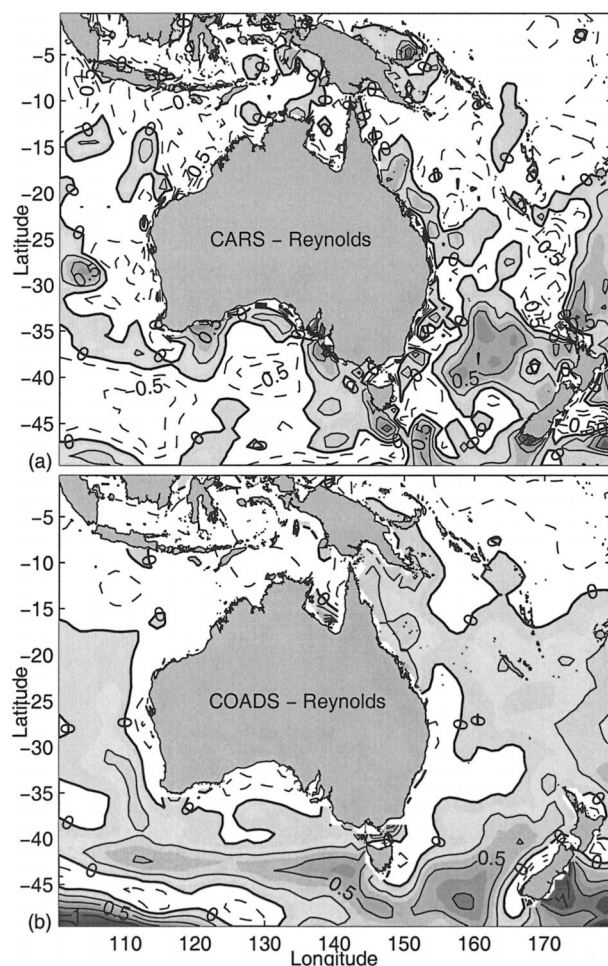


FIG. 13. Maps of the following SST climatology differences: (a) CARS–Reynolds, (b) COADS–Reynolds.

the PX34 section (Fig. 14 top). Here the three estimates diverge in concert by up to  $1^{\circ}\text{C}$  from the XBT value and lie well outside the 95% envelope. Since the XBT program covers the decade from 1991 and the data used in the other estimates are skewed toward a much earlier period, we suggest that this is a manifestation of a decadal or longer change in the regional circulation pattern. We also note that the PX34 transect crosses the large positive anomaly in the central Tasman Sea, seen in Fig. 13a. This comparison shows that in this region CARS agrees more closely with the XBT mean and that Reynolds is, if anything, slightly cool.

Finally we use the Fremantle–Singapore (IX1) section to examine the ability of CARS to represent the subsurface structure. The majority of the temperature difference between CARS and the XBT mean lies between  $\pm 0.25^{\circ}\text{C}$  although there are several isolated features that diverge by more than  $0.5^{\circ}\text{C}$  (Fig. 15a). For example, at  $26^{\circ}\text{S}$ , a large negative anomaly is caused by the XBT section resolving a small-scale structure associated with the local bathymetry, which is beyond the resolution of

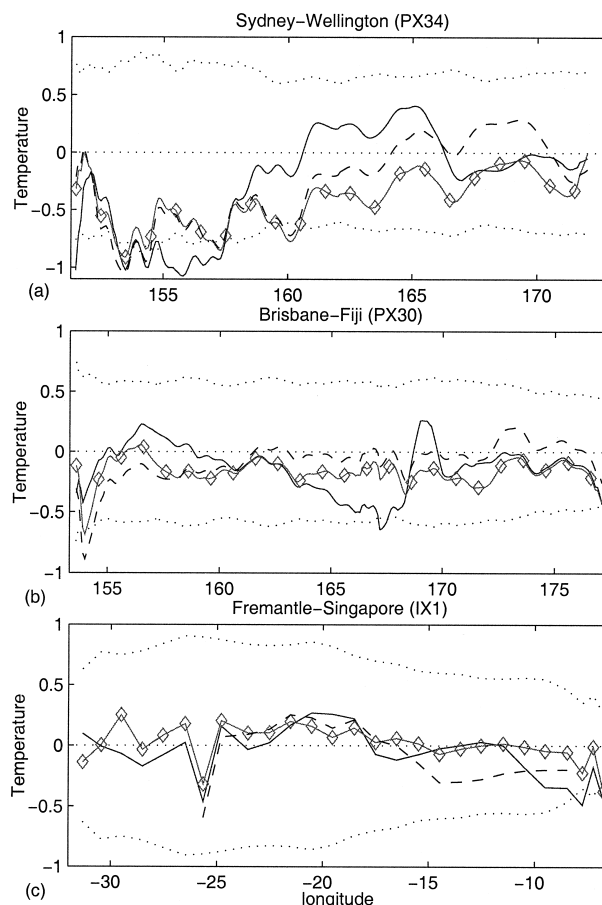


FIG. 14. Differences between the mean SST from long-term XBT observations and CARS (full line), Reynolds (diamonds), and COADS (dashed) along the three sections or transects (a) PX34, (b) PX30, and (c) IX1. The Tasman Sea transects PX34 and PX30 have been occupied four times per year since 1991 (Bailey et al. 1993). The IX1 transect has been occupied every 2–4 weeks since 1983 (Meyers et al. 1995). In each case the dotted lines represent the 95% confidence limits on the XBT mean.

CARS. The task of characterizing the annual cycle is far more demanding of the data than characterizing the mean. Here CARS appears to faithfully represent the general features of the annual amplitude and phase found in the XBT data (Figs. 15b,c). In particular, it captures the subsurface location of the maximum in the amplitude at  $12^{\circ}\text{S}$  and  $150\text{-m}$  depth. Above  $300\text{-m}$  depth, where the amplitude has a significant magnitude, the phase patterns show broad agreement.

#### d. An example of the CARS mean fields

We provide an example of the mean fields generated using the extended loess methods, restricting our attention to the eastern portion of the domain, the Tasman and Coral Seas ( $45^{\circ}\text{--}10^{\circ}\text{S}$ ,  $145^{\circ}\text{E--}180^{\circ}$ ). In Fig. 16 we present both the mean and rms of salinity at  $150\text{ m}$ . The mean field reproduces the large-scale pattern of water mass distribution and broad frontal structure observed

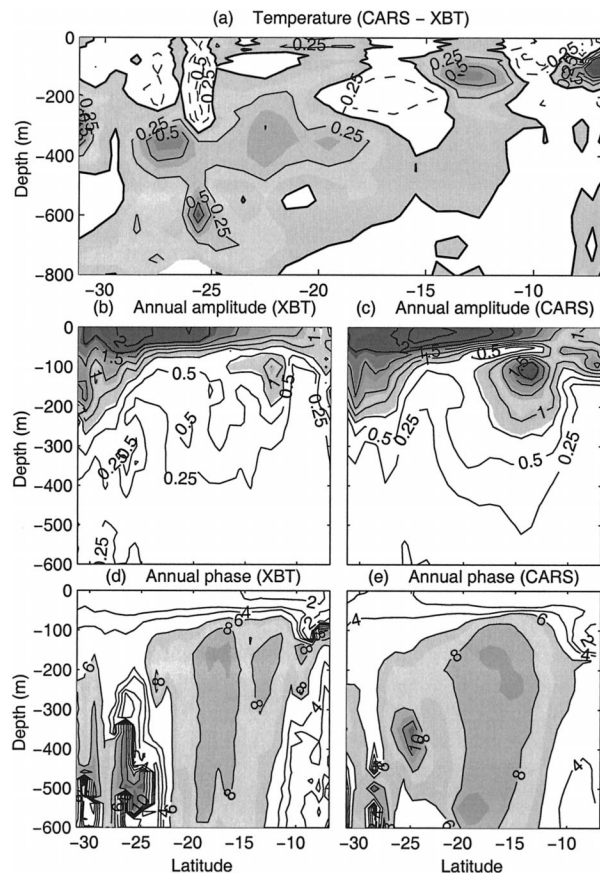


FIG. 15. Comparison of CARS temperature fields with results from the Fremantle–Singapore (IX1) high-density XBT section. (a) Mean temperature difference (CARS – XBT) along IX1 section, the amplitude of the annual cycle from (b) XBT data and (c) CARS, and the phase from (d) XBT data and (e) CARS.

in other climatologies (Levitus 1982; Sokolov and Rintoul 2000). However, the higher resolution of the CARS fields, obtained by the succession of measures described previously, has also been able to represent the interaction of these features with the bottom topography and coastal boundaries.

At this level the most prominent input of salinity occurs in the northeast, where a high-saline tongue spreads westward north of Fiji. This is associated with the Subtropical Lower Water, which is actually formed in the central Pacific region (Donguy 1994). Between 20° and 30°S the region is characterized by minimal salinity gradients and very low variability ( $<0.08$ ). However, adjacent to the eastern Australian coast there is a sharp alongshore gradient, giving evidence of both coastal freshwater input and a major offshore boundary flow. This is of course the EAC and the high-salinity filament (35.55–35.6) that spreads poleward along 155°E further reveals its influence. On the western flank of this feature we observe the separation of the main portion of the EAC from the coast (Godfrey et al. 1980) with a remnant proceeding southward. The high degree of variability is associated with the regular formation of mesoscale eddies in this region (Nilsson and Creswell 1981).

The separated EAC feeds into the Tasman front (Mulhearn 1987), which meanders its way across the central Tasman Sea (with increased variability). The path of this current is observed to be controlled at least in part by its passage across the major ridge systems bisecting the Tasman basin (Mulhearn 1987). Figure 16 shows that the CARS field has been able to resolve the details of the circulation around New Zealand. The Tasman Front reattaches to the land at North Cape and proceeds around the northern and eastern boundaries of New Zealand forming semipermanent eddy features along the way (Roemmich and Sutton 1998).

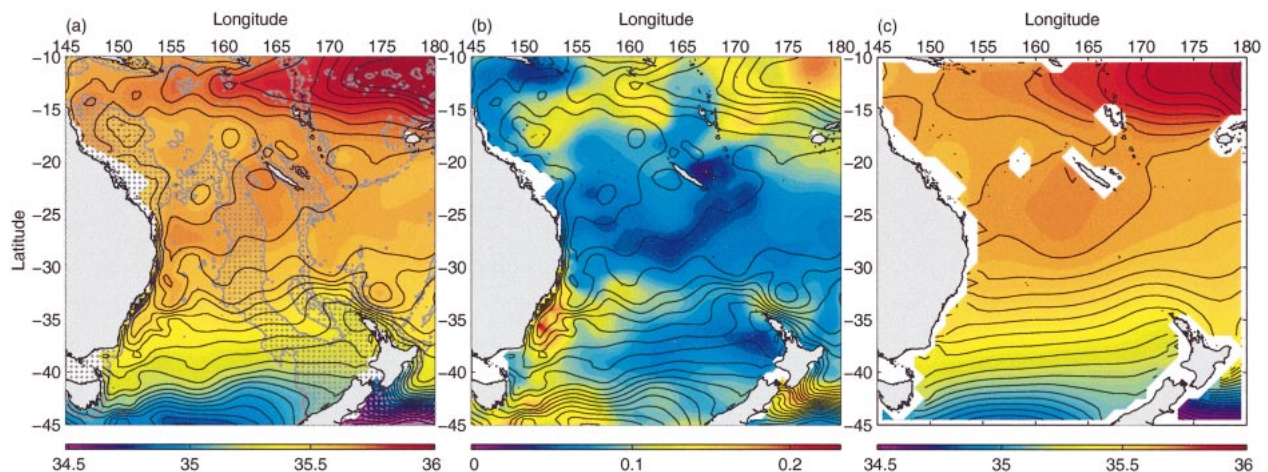


FIG. 16. Salinity at the 150-m surface for the eastern segment of the CARS domain (45°–10°S, 145°E–180°). (a) Mean from CARS, the hatched region indicates regions with depth  $<2000$  m, (b) rms variability (colour contours) with contours of the mean field [from (a), thin line contours] superimposed, and (c) mean from the WOD98 climatology (Conkright et al. 1998).



A compelling illustration of the improvements obtained using our interpolation system is obtained by comparing the CARS result in Fig. 16a with the same salinity field at 150 m obtained from a large-scale, global climatology (Fig. 16c, WOD98).

## 7. Discussion

### a. Choice of interpolation scheme

A central component of the mapping procedure is of course the choice of interpolation method. A variety of algorithms for estimating horizontally gridded values from irregular data have been described in the oceanographic literature. These include: binned averages (Wyrtki 1975), successive correction (Gomis et al. 1997), local least squares polynomial fitting (Chelton et al. 1990), natural or smoothing splines (McIntosh 1990; Lozier et al. 1995), variational inverse methods (Brasseur et al. 1996), and Gauss–Markov estimates (Bretherton et al. 1976). Further methods such as the averaging of vertical profile equations (Teague et al. 1990) and an empirical orthogonal function approach (Holbrook and Bindoff 2000) exploit the vertical correlation of oceanographic data. While the method of choice in many applications is Gauss–Markov (GM) or optimal interpolation (Bretherton et al. 1976), we have used a relatively unknown method of locally weighted least squares (loess).

The choice is not at all clear cut. The GM method is attractive because it provides an unbiased estimate of the gridded field along with estimates of the mapping error associated with the data. However, for large datasets it is often prohibitively expensive and hence is usually applied in a suboptimal form. It also requires a somewhat subjective choice of background mean and correlation function. These choices are not required in the loess scheme, which has at least an order of magnitude less arithmetic operations than does the GM algorithm (Chelton et al. 1990; Brasseur et al. 1996). This aspect has become less important as the actual CPU time used in the mapping component of the processing sequence represents only about 2.5% of the whole procedure. While the filtering characteristics of the loess smoother are almost as good as for GM (Chelton and Schlax 1994), a mapping error field is not readily extracted from the calculation.

The major reason for choosing the loess scheme came down to its greater flexibility in being able to incorporate features allowing for the spatial and temporal complexity of the data. For example, it is capable of fitting seasonal components in a single step with the mean, thus minimizing any temporal bias in the estimate. Furthermore it readily allows useful further extensions such as our TAR and BAR schemes and 3D mapping. This does not, of course, preclude the incorporation of these latter features into a GM scheme.

### b. Isobaric or isopycnal surfaces?

A further defining element of any mapping system concerns the surfaces on which properties are averaged. The traditional approach has been to average properties along isobaric levels (e.g., Levitus 1982). However, since mixing in the ocean occurs on density surfaces, recent studies indicate that such isobaric averaging may distort the mean fields (Lozier et al. 1994; Curry 1996; Gomis et al. 1997). We have chosen to use isobaric (actually depth levels) to perform the averaging for several technical reasons.

First, since we require values on depth levels, if the data are averaged along isopycnals, a subsequent vertical interpolation step is required. A more serious concern arises from the presence of dynamically unstable casts, which must be repaired prior to the averaging process (Gomis et al. 1997; Jackett and McDougall 1997). These difficulties are probably manageable; however, some more fundamental problems are encountered when we attempt to include temporal cycles in our curve fitting. This is because the depth of the neutral surfaces varies seasonally along with properties such as  $T$  and  $S$ , which are defined on these surfaces. Again this does not represent any difficulty as long as we remain in density space. However, as soon as we require results on depth surfaces we are faced with a myriad of technical dilemmas with no clear and unambiguous solutions. A further disadvantage of operating in an isopycnal framework is that we cannot easily incorporate the influence of the bathymetry into the mapping scheme.

In summary, we are again faced with a choice between two worthy approaches. While isopycnal averaging is clearly more consistent with the underlying physics, it is far more demanding, requiring us to solve some quite complex technical problems. It also lacks the flexibility inherent in using the loess approach with depth levels. Having said that, we are continuing development of methods that will produce an isopycnally averaged version of the interpolation system.

We note that there does not appear to be a consensus on this issue within the oceanographic community. While Lozier et al. (1995) have presented a new mean field of the North Atlantic on isopycnal surfaces, many recent climatologies have used an isobaric approach (Boyer and Levitus 1994; Brasseur et al. 1996; Reynaud et al. 1998).

### c. General application of the method

The methods presented here provide an improved approach to the interpolation of ocean properties. The results from the Australasian case study demonstrate that the range of procedures included in our system produce mapped fields that faithfully represent the data and effectively delineate boundary currents and frontal regions. The separate components of the system combine



to produce an accurate representation of this structure although the relative contributions of each component remain difficult to define. In addition, the relative importance of each component is likely to vary depending on the local characteristics of the region.

This leads us to question how the mapping system may be applied more generally. Would climatologies of the open ocean benefit from these procedures? In what areas of the global ocean would the techniques improve existing climatologies? Are there regions in which this system would actually produce inferior results? Clearly the greatest benefit of the TAR and BAR schemes is obtained in regions with complex geomorphology such as the Indonesian archipelago and New Zealand. Away from land masses, in deep water, the advantage gained may not be so obvious. However, open oceans all have boundaries of some form that require more detailed coverage. It is at these boundaries where the large-scale climatologies fail to resolve the narrow but very important current systems. Furthermore, with increasing depth the geometry becomes complicated even in open basins with implications for the trajectories of deep and bottom water masses.

We have shown that the inclusion of the seasonal terms in the interpolation both reduces the seasonal bias and increases the spatial resolution within the surface layers. This feature has perhaps the most universal applicability. The only qualification is likely to be where data densities are low causing instabilities in the higher-order temporal terms. In fact, we did not include semi-annual components in the nutrient calculations for this reason.

We note that despite the innovations presented here, the importance of obtaining a high-quality dataset remains of central importance. We therefore continue to refine our data assessment techniques as data are added to our archive. Future developments of our system include extending the region of coverage, mapping on isopycnal levels, and determining appropriate mapping error schemes.

Finally, the mean fields of CARS are available online at [http://www.marine.csiro.au/~dunn/eez\\_data/atlas.html](http://www.marine.csiro.au/~dunn/eez_data/atlas.html). The fields include temperature, salinity, oxygen, nitrate, phosphate, and silicate at 56 standard depths. Full documentation is also included along with several simple routines to access the fields.

**Acknowledgments.** We wish to thank Nathan Bindoff, David Griffin, Scott Condie, and Richard Coleman for their helpful comments, which have considerably improved the paper. We especially acknowledge the numerous researchers over many years, who have been involved in collecting data often under extremely difficult conditions. This project forms part of the Regional Oceans analysis system within the MUMEEZ program of the CSIRO Marine Research.

## REFERENCES

- Bailey, R., D. Roemmich, G. Meyers, W. Young, and B. Cornuelle, 1993: Transports of the major current systems of the Tasman Sea. Preprints, *Fourth Int. Conf. on Southern Hemisphere Meteorology and Oceanography*, Hobart, Australia, Amer. Meteor. Soc., 50–51.
- Boyer, T. P., and S. Levitus, 1994: Quality control and processing of historical oceanographic temperature, salinity and oxygen data. NOAA Tech. Rep. NESDIS 81, 64 pp.
- Brasseur, P., J. M. Beckers, J. M. Brankart, and R. Schoenauen, 1996: Seasonal temperature and salinity fields in the Mediterranean Sea: Climatological analyses of a historical data set. *Deep-Sea Res.*, **43**, 159–192.
- Bretherton, F. P., R. E. Davis, and C. Fandry, 1976: A technique for objective analysis and design of oceanographic instruments applied to MODE-73. *Deep-Sea Res.*, **23**, 559–582.
- Chelton, D. B., and M. G. Schlax, 1994: The resolution capability of an irregularly sampled dataset: With application to Geosat altimeter data. *J. Atmos. Oceanic Technol.*, **11**, 534–550.
- , —, D. L. Witter, and J. G. Richman, 1990: Geosat altimeter observations of the surface circulation of the Southern Ocean. *J. Geophys. Res.*, **95**, 17 877–17 903.
- Cleveland, W. S., and S. J. Devlin, 1988: Locally weighted regression: An approach to regression analysis by local fitting. *J. Amer. Stat. Assoc.*, **83**, 596–610.
- Conkright, M. E., and Coauthors, 1998: *World Ocean Database 1998: Documentation and Quality Control, Version 2.0*. CD-ROM Dataset Documentation, National Oceanographic Data Center Internal Report 14.
- Cresswell, G. R., and T. J. Golding, 1980: Observations of a south-flowing current in the southeastern Indian Ocean. *Deep-Sea Res.*, **27A**, 449–466.
- Curry, R. G., 1996: HydroBase: A database of hydrographic stations and tools for climatological analysis. Woods Hole Oceanographic Institute Tech. Rep. WHOI-96-01, 44 pp.
- Donguy, J. R., 1994: Surface and subsurface salinity in the tropical Pacific Ocean: Relations with climate. *Progress in Oceanography*, Vol. 34, Pergamon Press, 45–78.
- Dunn, J. R., and K. R. Ridgway, 2001: Mapping ocean properties in regions of complex topography. *Deep-Sea Res.*, **24**, 591–604.
- Godfrey, J. S., and T. J. Golding, 1981: The Sverdrup relation in the Indian Ocean, and the effect of the Pacific–Indian Ocean throughflow on Indian Ocean circulation and on the East Australian Current. *J. Phys. Oceanogr.*, **11**, 771–779.
- , G. R. Cresswell, T. J. Golding, A. F. Pearce, and R. Boyd, 1980: The separation of the East Australian Current. *J. Phys. Oceanogr.*, **10**, 430–440.
- Gomis, D., M. A. Pedder, and A. Viudez, 1997: Recovering spatial features in the ocean: Performance of isopycnal vs isobaric surfaces. *J. Mar. Syst.*, **13**, 205–224.
- Gouretski, V. V., and K. Jancke, 1999: A description and quality assessment of the historical hydrographic data for the South Pacific Ocean. *J. Atmos. Oceanic Technol.*, **16**, 1791–1815.
- Holbrook, N. J., and N. L. Bindoff, 2000: A statistically efficient mapping technique for four-dimensional ocean temperature data. *J. Atmos. Oceanic Technol.*, **17**, 831–846.
- Jackett, D. R., and T. J. McDougall, 1995: Minimal adjustment of hydrographic profiles to achieve static stability. *J. Atmos. Oceanic Technol.*, **12**, 381–389.
- , and —, 1997: A neutral density variable for the world's oceans. *J. Phys. Oceanogr.*, **27**, 237–263.
- Levitus S., 1982: *Climatological Atlas of the World Ocean*. NOAA Prof. Paper 13, 173 pp. and 17 microfiche.
- , R. Gelfeld, T. Boyer, and D. Johnson, 1994: Results of the NODC and IOC Oceanographic Data and Archaeology Rescue Project: Report 1: Key to Oceanographic Records Documentation, No. 19, NODC, 73 pp.
- Lozier, M. S., M. S. McCartney, and W. B. Owens, 1994: Anomalous

- anomalies in averaged hydrographic data. *J. Phys. Oceanogr.*, **24**, 2624–2638.
- , W. Brechner-Owens, and R. Curry, 1995: The climatology of the North Atlantic. *Progress in Oceanography*, Vol. 36, Pergamon Press, 1–44.
- McIntosh, P. C., 1990: Oceanographic data interpolation: Objective analysis and splines. *J. Geophys. Res.*, **95**, 13 529–13 541.
- Meyers, G., R. J. Bailey, and A. P. Worby, 1995: Geostrophic transport of Indonesian Throughflow. *Deep-Sea Res.*, **42**, 1163–1174.
- Mulhearn, P., 1987: The Tasman Front: A study using satellite infrared imagery. *J. Phys. Oceanogr.*, **17**, 1148–1155.
- Nilsson, C. S., and G. R. Cresswell, 1981: The formation and evolution of East Australian Current warm core eddies. *Progress in Oceanography*, Vol. 9, Pergamon Press, 133–183.
- NOAA, 1988: Digital relief of the surface of the earth. Data Announcement 88-MGG-02, NOAA National Geophysical Data Center, Boulder, CO. [Available online at <http://www.ngdc.noaa.gov/mgg/global/et0p05.html>.]
- Reiniger, R. F., and C. F. Ross, 1968: A method of interpolation with application to oceanographic data. *Deep-Sea Res.*, **9**, 185–193.
- Reynaud, T., P. LeGrand, H. Mercier, and B. Barnier, 1998: A new analysis of hydrographic data in the Atlantic and its application to an inverse modelling study. *International WOCE Newsletter*, No. 32, WOCE International Project Office, Southampton, United Kingdom, 29–31.
- Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.
- Ridgway, K. R., J. S. Godfrey, G. Meyers, and R. Bailey, 1993: Sea level response to the 1986–87 ENSO event in the western Pacific in the vicinity of Papua New Guinea. *J. Geophys. Res.*, **98**, 16 387–16 395.
- Robinson, A. R., and W. G. Leslie, 1985: Estimation and prediction of oceanic fields. *Progress in Oceanography*, Vol. 14, Pergamon Press, 485–510.
- Roemmich, D., and P. Sutton, 1998: The mean and variability of ocean circulation past northern New Zealand: Determining the representativeness of hydrographic climatologies. *J. Geophys. Res.*, **103**, 13 041–13 054.
- Sokolov, S., and S. R. Rintoul, 2000: Circulation and water masses of the southwest Pacific: WOCE section P11, Papua New Guinea to Tasmania. *J. Mar. Res.*, **58**, 223–268.
- Teague, W. J., M. J. Carron, and P. J. Hogan, 1990: A comparison between the Generalized Digital Environmental Model and Levitus climatologies. *J. Geophys. Res.*, **95**, 7167–7183.
- Walker, A. E., and J. L. Wilkin, 1998: Optimal averaging of NOAA/NASA Pathfinder Satellite sea surface temperature data. *J. Geophys. Res.*, **103**, 12 869–12 883.
- Webb, D., 2000: Evidence for shallow zonal jets in the South Equatorial Current region of the southwest Pacific. *J. Phys. Oceanogr.*, **30**, 706–720.
- Woodruff, S. D., S. J. Lubker, K. Wolter, S. J. Worley, and J. D. Elms, 1993: Comprehensive Ocean–Atmosphere Data Set (COADS) Release 1a: 1980–92. *Earth Syst. Monitor*, **4**, 1–8.
- Wyrski, K., 1975: Fluctuations of the dynamic topography of the Pacific Ocean. *J. Phys. Oceanogr.*, **5**, 450–458.

