



Wilkinson Basin area water masses: a revisit with EOFs

Alex Warn-Varnas^{a,*}, Avijit Gangopadhyay^b, J.A. Hawkins^c, Allan R. Robinson^d

^aNaval Research Laboratory, Code 7322, Stennis Space Center, MS 39529, USA

^bThe School for Marine Science and Technology, and the Department of Physics, University of Massachusetts Dartmouth,
285 Old Westport Road, North Dartmouth, MA 02747, USA

^cPlanning Systems Inc., Slidell, LA 70458, USA

^dDivision of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, MA 02138, USA

Received 13 May 2003; received in revised form 4 August 2004; accepted 3 September 2004

Available online 25 November 2004

Abstract

Hydrographic data sets from the Brooks (J. Geophys. Res. 90 (1985) 4687) survey are used for the initial study of water mass characteristics and distributions of the Wilkinson Basin area in the northwest corner of the Gulf of Maine. Parameters of Maine Surface, Intermediate, Bottom, and Slope waters are derived from the available conductivity-temperature-depth (CTD) casts and developed using cluster analysis with a distribution function for cluster point differences in temperature, salinity and depth. The cluster analysis yields two distinct surface water masses, one of which is warmer and less salty. This warmer water mass could be a result of the transport of fresher water into the western Wilkinson basin area by a branch of the Maine Coastal Current and subsequent solar heating.

A variance-based empirical orthogonal functions (EOF) analysis associated with water masses in the Wilkinson Basin area is simultaneously undertaken for the advancement of feature models. The cloud of points representing each water mass in terms of temperature, salinity, and depth are derived through cluster analysis and identified as matrix elements in the singular-value-decomposition (SVD) and covariance analysis. Covariance and correlation statistics of the water masses are analyzed. The vertical temperature and salinity is reconstructed through an EOF decomposition of each water mass cloud of points and a summation of all of them for the resultant vertical distribution. This constitutes the basis for a water mass-based feature model with latest-measurement update capability.

The percentage of the water volume occupied by each of the five water masses is derived and expressed as a function of depth and CTD cast location. The means, standard deviations, and ranges of the water masses are extracted in the temperature-salinity-depth space. The origin and formation of the water masses is considered together with the controlling physical processes. The prevalent dynamics is related to the water mass structure and characteristics. It is shown that slope water arrives into the Wilkinson Basin area from the Northeast Channel.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Wilkinson Basin; Water masses; Profile method; Cluster analysis; EOFs; Feature models

*Corresponding author. Tel.: +1-228-688-5223; fax: +1-228-688-4759.
E-mail address: varnas@nrlssc.navy.mil (A. Warn-Varnas).

1. Introduction

The water mass structure in the Gulf of Maine has been studied by Hopkins and Garfield (1979), Brown and Irish (1993) and others. Analyses of data sets indicate the presence of the following three interior water masses: Maine Surface Water (MSW), Maine Intermediate Water (MIW), and Maine Bottom Water (MBW). Three main boundary water masses are also present: Scotia Shelf Water (SSW), Georges Bank Water (GBW), and Slope Water (SLW, upper and lower).

The distribution of the water masses in temperature, salinity, and depth space show a degree of variability that inhibits the definition of absolute temperature and salinity ranges (Hopkins and Garfield, 1979). The variability can be due to a particular data set, the season (Brown and Irish, 1993) or inter-decadal variation (Colton, 1968). The water mass identification has to be based on the overall physical picture consisting of the temperature and salinity distribution together with the dynamical picture and structures that are at play.

Gangopadhyay et al. (2003) have summarized the synoptic circulation system in the Gulf of Maine and Georges Bank region using a feature-oriented approach. Prevalent oceanographic circulation structures are identified from previous observational studies, also known water mass characteristics and relevant dynamical processes responsible for formation and maintenance of the synoptic features are noted. Features include the buoyancy-driven Maine Coastal Current (MCC), the Georges Bank anticyclonic frontal circulation system, the basin-scale cyclonic gyres (Jordan, Georges, and Wilkinson), the deep inflow through the Northeast Channel, the shallow outflow via the Great South Channel, and the shelf-slope front (see Fig. 1).

The interior Gulf of Maine has three cyclonic circulation regions situated over three deep basins: Georges, Jordan and Wilkinson (see Fig. 1). These gyres are set up by the deep inflow of saline waters through the Northeast Channel, and they are forced by topography (Hannah et al., 1996; Lynch, 1999). In addition, it has been observed that the dominant temporal variability in or between the

gyres corresponds to the order of months (Xue et al., 2000).

Among the three basins, Wilkinson Basin area is the farthest away from the inflow region of the Northeast Channel and nearest to the outflow region through the Great South Channel. Furthermore, the underlying topography in this basin is fragmented, as compared to those in the Jordan and Georges Basins. Such physical constraints trigger vigorous water mass transformation processes in this particular basin throughout the year. In particular, the heavy vertical mixing that occurs in this basin (Brown and Irish, 1993) in the winter between the fresh MSW and the saltier MBW results in a water mass called 'MIW'. In the spring, coastal runoffs from rivers result in fresher surface water, and in the summer, solar heating stratifies the upper layers to produce a varied combination of water masses in the basin over time.

In addition to the gyre circulation transport, the regional currents in the Gulf of Maine also transport the water masses. A major circulation feature is the narrow MCC, with its bifurcation and trifurcation segments. In particular, the bifurcating southwestward branch of the MCC near Jeffery's Bank (68.5°W, 43.5°N) affects the water mass formation in the Wilkinson Basin area (Brown and Beardsley, 1978). Thus, the CTD stations to the outskirts of the 200 m isobath, which is sometimes called the boundary of the basin itself, (Gangopadhyay et al., 2003), are included in our analysis. We refer to this larger area (see Fig. 2 for CTD station locations) as the 'Wilkinson Basin area'. There is the circulation from Northeast Channel and Scotia shelf inflows and the outflow through the Great South Channel and the shelf-slope front. The tidal front around Georges Bank contributes to mixing and water mass formation. Buoyancy driving, bottom topography, tides, river inflow, atmospheric forcing and basin-wide pressure gradient set-up influence large-scale circulation. Together, the circulation, features, and phenomena control water mass formation and modification.

One objective of this work is to study the Wilkinson Basin area water mass characteristics through the application of two water mass models. One model is based on the water mass signature on

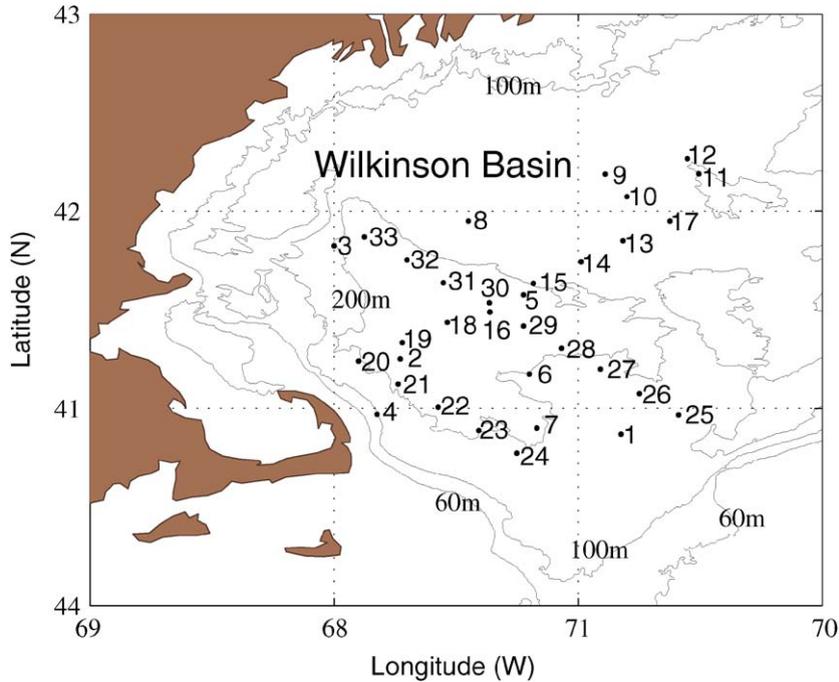


Fig. 2. CTD station location and numbering in the ‘Wilkinson Basin area’, which spans beyond the 200 m isobath, the boundary of the Wilkinson basin. The inner contour is the 200 m isobath. The next two are 100 and 60 m isobaths, respectively.

temperature and salinity profiles (Warn-Varnas et al., 1999). The other one is based on cluster analysis (Kim et al., 1991).

Another objective is to relate the water mass characteristics and structures to the dynamics of the Wilkinson Basin area and the Gulf of Maine. Understanding the formation and modifications of the water masses is also an objective. The last objective is to evolve an empirical orthogonal functions (EOF)-based characterization of the prevalent water masses in the Wilkinson Basin area and to relate the EOF-based characterization to feature models. Analyses of individual water masses are considered in detail with the aim to relate to various features and regions. In the future, this approach will be extended to feature-oriented nowcasting, forecasting, assimilation and process studies in other coastal regions of the world ocean.

A water mass is represented by a cluster of points consisting of temperature, salinity, and depth values. In CTD station measurements, the shapes of the vertical temperature and salinity

profiles indicate the water masses that are present. The line segments associated with the water masses in the measured profiles are observed throughout the survey region and suggest the existence of separate water masses. Such an approach for identifying water masses has been used by Nittis et al. (1993) and by Warn-Varnas et al. (1999).

Another approach that identifies water masses is with the application of cluster analysis to hydrographic data and previous knowledge. Individual points are clustered into groups from the temperature, salinity, depth, and geographic distribution, this can then be examined for their $T-S$ properties and spatial distribution. The application of cluster analysis with physical guidance to the Yellow and East China Sea data sets has yielded encouraging results. Kim et al. (1991) identified the Yellow Sea water masses. Also, Hur et al. (1999) described the monthly variations of water masses in the Yellow and East China Seas by this method.

In the Gulf of Maine, empirical functions have been derived from data sets to represent temperature and salinity profiles and the contained $T-S$

water mass characteristics in the cyclonic gyres (Bierweiler, 1999). These data-based empirical functions represented the hydrographic structures in the Jordan, Georges, and Wilkinson Basin area regions. The representation was a first step towards the construction of a synoptic feature model (Gangopadhyay et al., 2003).

Synoptic observational studies were used to extract feature-model water mass T – S structures for different gyres in the Gulf of Maine and Georges Bank region by Gangopadhyay et al. (2003). See profiles in Figs. 4, 5 and 8 of Gangopadhyay et al. (2003).

In this work, we relate the Wilkinson Basin area water masses to a variance-based EOF analysis associated with the cluster of points representing the temperature, salinity, and depth of each water mass. The derived EOF and eigenvalues enable the representation of vertical temperature and salinity profiles in the Wilkinson Basin area. This constitutes a feature model based on water masses and their variances, and it can be substituted for or supplemented with the existing “empirical” feature models.

In the following sections, we discuss the analysis of water masses (Section 2), EOF-based vertical feature models (Section 3), water mass distribution (Section 4), and our conclusions (Section 5).

2. Analysis of water masses

2.1. Data

We focus our study on revisiting the Wilkinson Basin area data sets available from Brooks (1985) surveys, archived in NODC, in order to identify the water masses and characterize them by clusters of points in temperature, salinity, and depth space and the active dynamic processes. Available NODC station data from the month of June were used for feature-model development study for the entire Gulf of Maine by Gangopadhyay et al. (2003). For the purpose of this study, we have separated a selected number of profiles located in Wilkinson Basin area. These station locations are shown in Fig. 2. All together there are 33 CTD stations over a depth range of 30–250 m. The

measurements were taken over a short time period in June 1982. These sections were primarily chosen to obtain a synoptic picture of the area. They encompass the Wilkinson Basin and surrounding regions.

2.2. Profile model of water masses

The water masses can be identified from their known signature (T – S properties) in the measured vertical temperature and salinity profiles. The shapes of the profiles indicate the water masses that are present. The profile parameterization method can be described in two steps. The first step is to identify the temperature, salinity and depth boundaries of all available water masses in every profile. The second step is to cluster these segmented water mass distribution from a regional perspective.

An illustration of the profile parameter method is provided in Fig. 3 corresponding to CTD station number 5 (latitude 42.5747°N, longitude–69.2247°W) in Fig. 2. The water masses are distinguished by their temperature and salinity line segment sets that define their associated profile span. Their vertical layer distributions and associations with temperature, salinity and depth ranges are first obtained. A similar methodology was used by Warn-Varnas et al. (1999). Parameters of depth, temperature, and salinity are introduced for locating the boundaries of the water masses (Fig. 3). For example, the MSW mass is defined by temperature and salinity line segments between the parameters T_S , S_S at the surface and parameters T_{SE} , S_{SE} at a depth D_S . The depth D_S is determined from the temperature and salinity slope inflection point. Proceeding downwards into the ocean, the MIW water follows. Similarly for MIW, the parameters at the temperature minimum location are also included, as D_{IC} , T_{IC} , S_{IC} . The parameters at the bottom of MIW water and the beginning of MBW water are D_{IE} , T_{IE} , W_{IE} and are determined from the temperature and salinity slope inflection point. The parameters at the bottom of the MBW water and the beginning of SLW water are: D_{BE} , T_{BE} , S_{BE} . They are determined from the temperature and salinity slope inflection point. The parameters at the

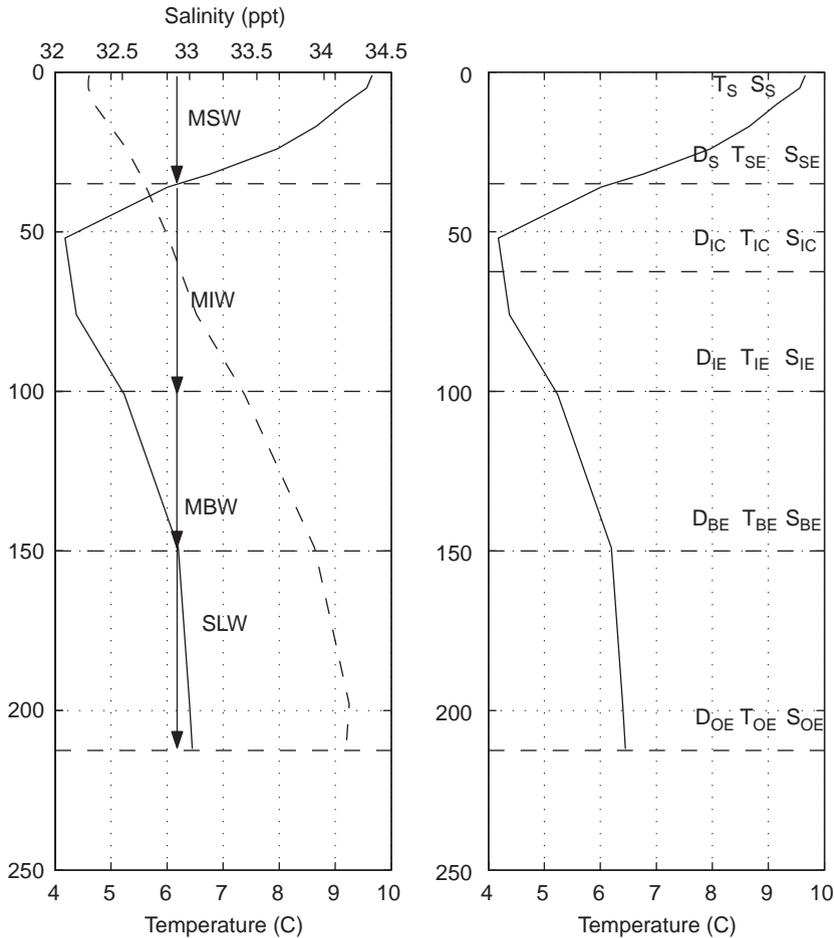


Fig. 3. Indication of water mass signatures on measured temperature (solid line) and salinity (dashed line) profiles. The four prevalent water masses are indicated in the left panel. The right panel contains the water mass parameters.

bottom of the SLW water are D_{OE} , T_{OE} , S_{OE} and are taken to be at the bottom of the temperature and salinity profile. Each water mass has its own specification in terms of a cluster of parameters in temperature, salinity and depth space. For example, the MSW water contains the set of parameters, $\{[T_S, S_S], [D_S, T_{SE}, S_{SE}]\}$ at the CTD stations. Note that individual CTD cast profiles may not contain all water masses.

From the sets of parameters the statistics of the water masses are calculated and are listed in Table 1. These include the means, standard deviation, range of variation and number of samples for each of the clusters of points representing a particular water mass. The depth

of the water masses increases from the surface down as one proceeds through SURF, MIW, MBW, and SLW.

We have calculated the variance ellipsoids for the Wilkinson Basin area water mass parameter distributions. The ellipsoids are constrained to be along the axes of the coordinates (Warn-Varnas et al., 1999). The centers of the ellipsoids represent the mean of each particular water mass. The axes are the standard deviations, and the volume defines the water mass domain within the standard deviation confinement along the axes, reflecting the variability about the mean.

Fig. 4 shows the results. There, the surface water mass is labeled SURF instead of MSW because

Table 1
Statistics of water masses from profile analysis

Water mass	No. of samples	Depth (m)	Mean D	σ_D	T (C)	Mean T	σ_T	Salinity	Mean S	σ_S
SURF	44	0–37	32.27	2.57	4.8–14.0	9.34	3.13	31.45–32.78	32.34	0.30
MIW	66	30–150	75.46	37.61	3.8–9.0	5.27	1.27	32.23–33.65	32.86	0.37
MBW	42	80–200	143.8	33.42	3.9–6.3	5.42	0.66	32.24–34.00	33.5	0.40
SLW	28	143–245	187.39	30.74	5.2–6.6	6.08	0.38	33.36–34.40	33.92	0.26

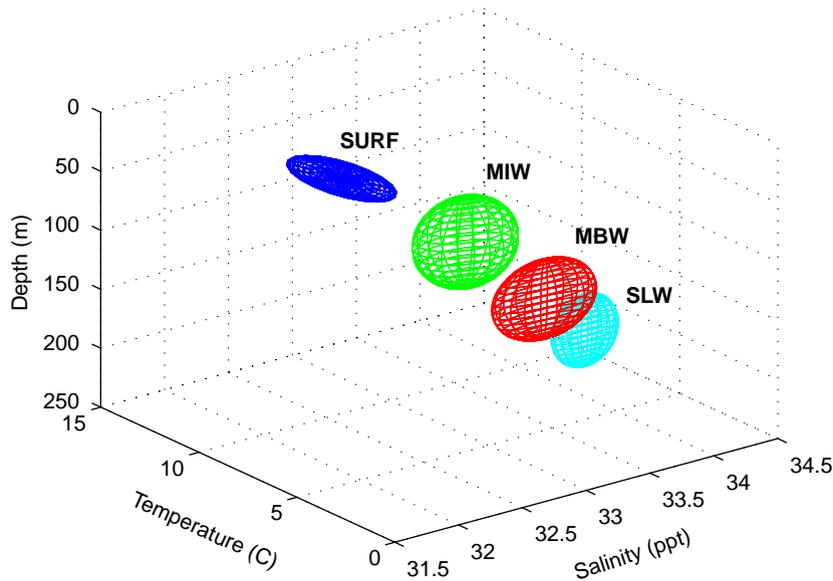


Fig. 4. Variance ellipsoids of water masses in (T, S, D) space from profiles analysis.

later in the paper, we will introduce the notation of MSW and Maine Hot Water (MHW) for two surface water masses. We encounter SURF, MIW, MBW, and SLW at successively greater depths (Table 1). The MIW occupies the largest volume and has the largest composite standard deviations. The layered structures of the water masses reflect the circulation dynamics and physical processes.

A more conventional way to approach the water mass clusters is to consider the variance ellipses of the temperature and salinity parameters with the data points in the background. Fig. 5a shows the results. We can now associate the water mass domains, within a specific standard deviation, with the data distribution as presented in a T – S diagram of the Wilkinson Basin area CTD casts.

The tracking of the data points by the ellipses is apparent in the background. The MIW mass encompasses a large domain in the T – S space and contains the temperature minimum of around 4°C . To the right of the temperature minimum are the MBW and SLW water masses, which have increasing salinities. To the left of the minimum is the SURF, with decreasing salinity. The SLW water mass exhibits only the lower branch, indicating that the upper branch of SLW is not present in the Wilkinson Basin area.

2.3. A cluster model of water masses

Another method for identifying water masses is cluster analysis, a statistical technique that

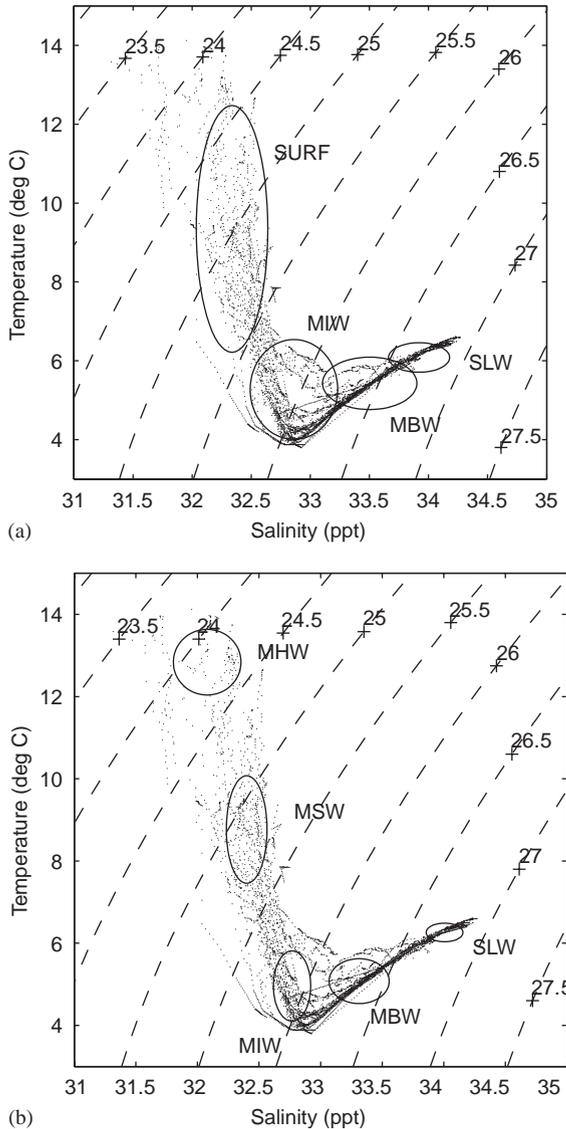


Fig. 5. Variance ellipses of water masses in (T, S) plane with data plotted in background. (a) Parameters from profiles analysis and (b) parameters from cluster analysis. Note that the cluster method re-classifies the surface waters in the profile method into two distinct water masses: MHW (a new water mass) and MSW.

classifies measurements of temperature and salinity into groups, or clusters, according to similarities in the respective values of temperature, salinity, and depth. The adjacent clusters are combined until the number of clusters corresponds

to the existing water masses. This approach has been used in the Yellow and East China Seas to delineate the monthly water mass variations (Kim et al., 1991; Hur et al., 1999). The latter study differentiated and identified the seasonal layers of water masses by this approach.

The approach is objective and uses a distance function for combining cluster points consisting of temperature, salinity, and depth. The distance function is defined as

$$F_D = T_n^2 + S_n^2 + d_n^2 \quad (1)$$

with the normalized temperatures, salinities, and depths are defined as

$$T_n = \Delta T_c / \sigma_T, S_n = \Delta S_c / \sigma_S, d_n = \Delta d_c / \sigma_d.$$

where ΔT_c , ΔS_c and Δd_c are the temperature, salinity, and depth distances between two clusters, and σ_T , σ_S and σ_d are the standard deviations.

The cluster analysis was applied to the Wilkinson Basin area CTD stations. The grid points were spaced 5 m apart in the upper 100 and 10 m apart from 100 to 250 m. On every grid point there is a temperature, salinity, and depth value.

The procedure consists of computing the distance function between all clusters of points on the vertical grids of CTDs and then forming a hierarchical tree of new clusters. The formation of new clusters is based on the average linkage method that tracks the average distance between clusters. The new clusters are grouped into a hierarchical tree according to an increasing distance function between them. The number of new clusters (or groups) is terminated at the expected number of water masses.

The iterations for combining clusters were stopped at 5 because this yielded the MIW, MBW and SLW water masses and two surface water masses. One of the surface water masses was unexpected and exhibited a warm trend. The results are summarized in Table 2. There, the number of samples and ranges of temperature, salinity, and depth are shown; with means and standard deviations also noted. The cluster analysis samples the profiles at the standard depths, every 5 m down to 100 m and then every 10 m down to 250 m. The profile analysis samples the slope inflection points and the maxima/minima. As

Table 2
Statistics of water masses from cluster analysis

Water mass	No. of samples	Depth (m)	Mean D	σ_D	$T(C)$	Mean T	σ_T	Salinity	Mean S	σ_S
MHW	51	0–20	5.39	5.18	10.55–14.12	12.84	0.8	31.32–32.53	32.08	0.27
MSW	196	0–45	19.59	10.90	5.84–12.12	8.77	1.31	31.99–32.74	32.4	0.16
MIW	351	30–100	62.64	17.19	3.81–7.23	4.96	0.86	32.31–33.13	32.77	0.15
MBW	248	80–180	121.55	23.97	3.85–6.20	5.07	0.55	32.32–33.94	33.31	0.24
SLW	79	160–220	185.82	16.30	5.72–6.61	6.26	0.23	33.68–34.27	34.01	0.15

a result, there are more samples in the cluster analysis (Table 2) than in the profile analysis (Table 1).

Ellipsoids are calculated for each water mass cluster of points using the respective means and standard deviations. The results are shown in Fig. 5b as a projection onto the T, S plane. The warm water mass MHW is found at the surface; MSW, MIW, MBW, and SLW inhabit progressively greater depths. The water masses span a range of temperature and salinity values. Also the structure of the data points outlines a path in the T, S , and D space where each point belongs to a particular water mass.

2.4. Interpretation of water mass models

One way to look at the water mass clusters is to consider the T, S variance ellipses and their superposition on the data points in the T, S plane. Fig. 5 illustrates the results for both water mass models. The cluster model delineates the Maine Surface Water into two surface water masses: MHW and MSW. MHW is appreciably warmer than MSW (by about 4°C) and is slightly less saline. Also, in the cluster model, the variance ellipses representing water masses do not overlap. Comparatively, in the profile model, they do. The axes of the variance ellipses span one standard deviation along the salinity and temperatures coordinates.

The statistics of the water mass model clusters indicate the temperature and salinity ranges associated with each water mass. The standard deviations of the cluster model are smaller than the standard deviations of the profile model. The differences in the mean temperatures and salinities

between the models are within the standard deviation variation of the models.

Absolute temperature and salinity ranges for each of the water masses can only be associated and established within a particular data set. Annual variations in the physical structure and properties of the water masses exist in the Gulf of Maine (Brown and Irish, 1993). Long-term trends in water type also exist in the Gulf of Maine. Hopkins and Garfield (1979) point out a warming trend of about 4°C relative to Bigelow (1927).

3. EOF-based vertical feature model

The previous cluster analysis enables the separation of CTD casts into clusters or clouds of temperature, salinity, and depth points representing the water masses. In the vertical, the water masses are separated according to their own cluster of points. The definition of a vertical water mass structure enables the coupling of water masses to a vertical EOF analysis.

3.1. Profile representation

3.1.1. Method

The approach of this analysis involves the construction of a matrix A from the measured temperatures at the CTD stations. The standard grid is defined by points spaced 5 m apart in the upper 100 m and 10 m apart from 100 m to 250 m. The matrix elements, a_{ij} , consist of rows (i index) and columns (j index) representing the variances of temperature relative to the horizontal average. They are defined as $a_{ij} = \bar{T} - T_{ij}$, where \bar{T} is the horizontal average and T_{ij} represents the

temperature on the i th row at the j th CTD station. Cluster analysis (Section 2.3) assigns a group number to each CTD cast measurement of temperature and salinity at the chosen standard depths. In this case there are 5 group numbers corresponding to 5 water masses. The matrix elements, a_{ij} , belonging to each water mass are sorted according to the group number.

3.1.2. Application to Wilkinson Basin area

For the Wilkinson Basin area, the original matrix A is decomposed according to water mass type as follows:

$$A = A_{\text{MSW}} + A_{\text{MHW}} + A_{\text{MIW}} + A_{\text{MBW}} + A_{\text{SLW}}. \quad (2)$$

Here, the subscripts denote the five different water mass constituents. When working with a particular water mass, $A = A_m$, where m is the water mass. In such a case the a_{ij} corresponding to water masses different from m are set to zero in matrix A . The matrix A_m is reconstructed for each water mass as: $A_m = \sum_{i=1}^k u_i \lambda_i v_i^T$, where u are the EOFs over column space, v represents the EOFs over row space with $A_m v_i = \lambda_i u_i$, λ_i are the eigenvalues, and m is the particular water mass in question. Also, $A_m A_m^T u_i = \lambda_i^2 u_i$. This decomposition is similar to that used by Fukumori and Wunsch (1991) for the North Atlantic water masses. We have used the EOF and SVD algorithms developed by Bjornsson and Venegas (1997).

A vertical EOF analysis is performed separately on each of these five constituent water masses. For the temperature structure of the MIW water mass, the first five vertical EOFs are shown in Fig. 6a. In the vertical, they are confined to the depth of the MIW water mass on the standard grid. However, some overshooting occurs at the boundaries of the grid. Their structure exhibits an increasing number of inflection points as the EOF number increases.

The average error of the reconstructed variance of the temperature field as a function of the number of EOFs is displayed in Fig. 6b. The average error is defined as $\bar{\epsilon} = \frac{1}{MN} \sum [A_m - A_m^r]$, where the summation is over M rows and N columns (CTD stations) and A_m^r denotes the reconstructed water mass variance summed over k EOFs.

The error, together with the standard deviation ($\sqrt{\bar{\epsilon}}$) envelopes, decreases with the EOF number. Thus, the gradient of the error, $\partial \bar{\epsilon} / \partial z$, decreases with the EOF number and indicates a change in trend at about 5 EOFs (Fig. 6b).

We considered the percent variance represented in the reconstruction of the temperature profiles for the five water masses. The variance error was expressed as a function of the number of EOFs or eigenvalues used.

The total variance of the reconstructed A is related to the eigenvalues as: $\|A\|^2 = \lambda_1^2 + \lambda_2^2 + \dots + \lambda_k^2$. Each eigenvalue, λ_i , is a measure of the variance represented by it.

Fig. 6c displays the results. We note that the five eigenvalues or EOFs account for about 95 percent or more of the variance in the reconstructed temperature field of each water mass. Table 3 lists the five eigenvalues for each water mass type. The prior availability of these eigenvalues enables the reconstruction of the temperature field pertaining to each water mass.

The variance of each water mass span a cluster of points along the depth of the water column, illustrated by the ellipsoids in Fig. 4. Each water mass has its own distribution of points along the depth dimension. As the location of the water masses changes in the ocean, a certain amount of overlap along the depth dimension results. In the overlap zone, each water mass has its own cluster of points (each point belongs to a particular water mass), as determined by the analysis. During the vertical EOF decomposition, the higher mode EOFs pick up the variances of the interaction zone and can be thought of as reflecting the interaction along the boundaries of the water masses. Some water masses such as the MIW can interact with the MSW and MBW. The surface water mass MSW, or the slope water mass SLW, can only interact with one other water mass MIW or MBW, respectively. The MHW water mass can interact with the MSW water mass. At some shallow CTD stations one has only the surface mass, MSW. At other stations no bottom water, MBW, or slope water, SLW, is present.

The improvement in explained variance (Fig. 6c), as the EOF number increases, tends to reflect the interaction that each water mass

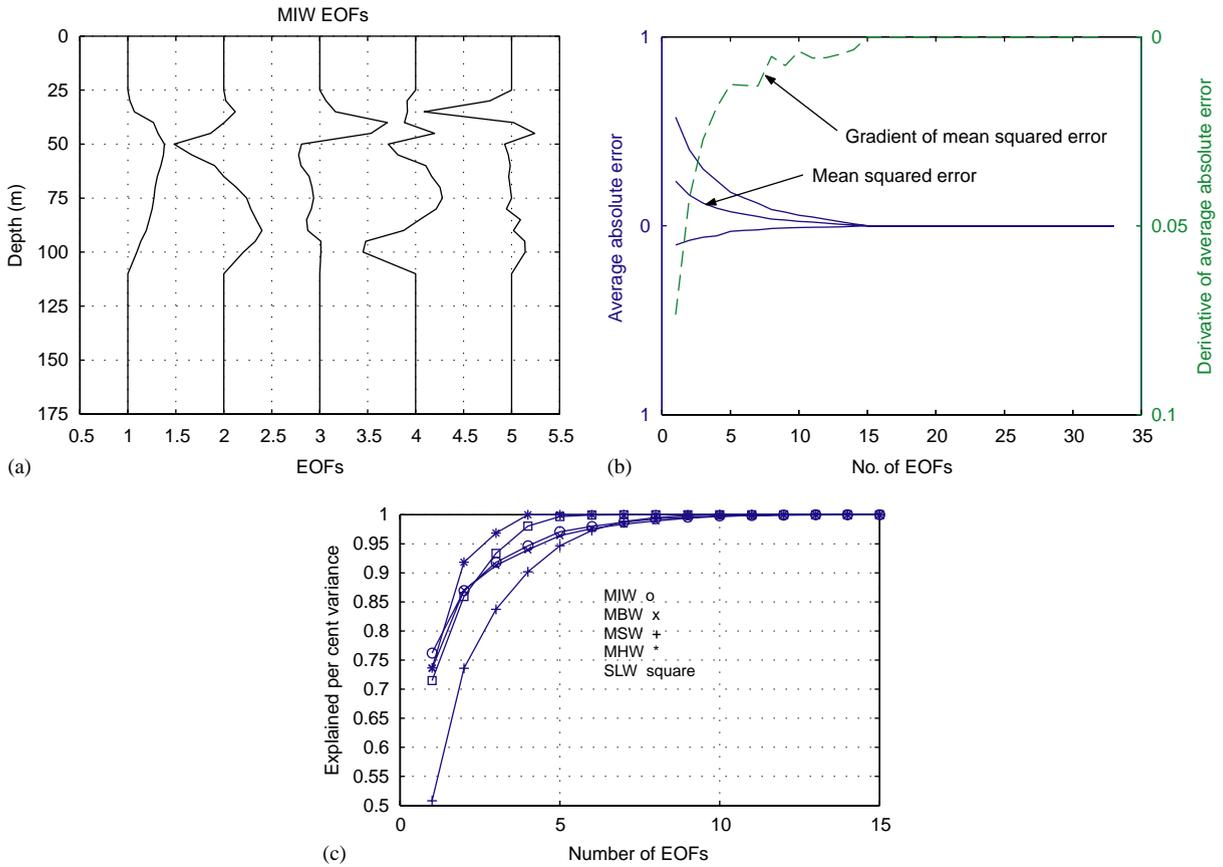


Fig. 6. (a) The first five vertical EOFs for MIW as a function of depth; (b) Average error of the reconstructed variance of the temperature field with standard deviation envelopes (solid lines) and the gradient of the mean squared error as a function of the number of EOFs (dashed line) and (c) Explained percent variance for each reconstructed water mass as a function of included EOFs.

Table 3
Eigenvalues for the first 5 EOFs

EOF mode #	MIW	MBW	MSW	MHW	SLW
1	12.0689	5.6407	10.1918	4.1686	1.3695
2	4.5181	2.4128	6.8238	2.0703	0.6161
3	3.0534	1.3593	4.5491	1.0897	0.4409
4	2.3334	1.0779	3.6359	0.8617	0.3510
5	2.1421	1.0264	3.0225	0.0000	0.2074

undergoes. The slope water, SLW, and hot surface water, MHW, require fewer EOFs to represent variance (Fig. 6c). These two water masses, MHW and SLW, also have lower occur-

rence statistics and boundaries with only one water mass.

The intermediate water MIW, which can interact with two water masses, requires five EOFs for reaching a 95% variance representation (Fig. 6c). The bottom water MBW, which interacts with one or two water masses, has a similar behavior (Fig. 6c). The surface water that can interact with one water mass requires about 5 EOFs for a 95% variance representation (Fig. 6c). The slower variance improvement for the surface water mass MSW, as the EOF number increases, suggests a higher complexity in its interaction with the intermediate water, MIW or MHW.

3.1.3. *Reconstruction of vertical profiles from the EOFs*

The temperature variances are reconstructed separately for each water mass. To obtain the total variance over depth, the reconstructed variances from each water mass are totaled. Matrix A was decomposed according to water masses as follows: $A = A_{MSW} + A_{MHW} + A_{MIW} + A_{MBW} + A_{SLW}$. Over the dimension of A , or all the variances, one has an orthonormal set that consists of

$$\begin{aligned} &(A_{MSW} + A_{MHW} + A_{MIW} + A_{MBW} + A_{SLW}) \\ &(A_{MSW} + A_{MHW} + A_{MIW} + A_{MBW} \\ &+ A_{SLW})^T u_i = \lambda_i^2 u_i \end{aligned} \quad (3)$$

The product $(A_{MSW} + A_{MHW} + A_{MIW} + A_{MBW} + A_{SLW})(A_{MSW} + A_{MHW} + A_{MIW} + A_{MBW} + A_{SLW})^T$ involves terms such as $(A_{MIW})(A_{MIW})^T + (A_{MIW})(A_{MSW} + A_{MHW} + A_{MBW} + A_{SLW})^T + \dots$. The cross terms represent products between one water mass and all the others. The water masses tend to be separated in T - S distribution and vertical locations (Figs. 4 and 5b). Since the cluster of points representing each water mass tends to be concentrated in a particular location, cross products such as $(A_{MIW})(A_{MSW} + A_{MHW} + A_{MBW} + A_{SLW})^T$ involve outlier cluster points, away from the main concentration of each water mass cluster point. The cross product over the main water mass cluster distributions is zero and AA^T reduces to approximately a summation over the individual water masses, enabling the decomposition of the matrix A into water mass components.

The question then arises: How accurate is reconstruction when each water mass is treated separately and then the reconstructed variances of each water mass are summed for the total field?

To test the accuracy, the Wilkinson Basin area temperature and salinity fields were reconstructed using 5 EOFs and compared against the CTD measurements. In the reconstruction, the two surface region water masses $A_{MSW} + A_{MHW}$, were combined into one in order to enable the EOF analysis to represent them together over the depth of the surface layer. An average of the reconstructed field at all the CTD stations is computed and compared against the averaged measurements.

The standard deviations are also considered. The results are shown in Figs. 7a and b, respectively.

Fig. 7a compares the horizontally averaged temperature obtained from data (solid line) and the reconstructed field (circles). The two curves

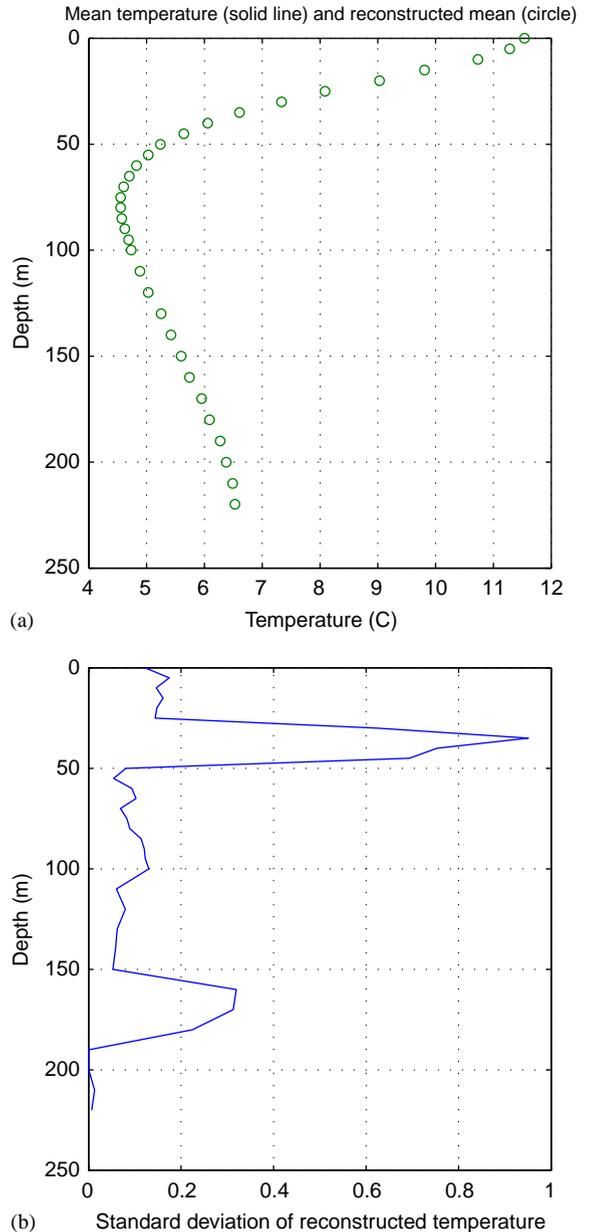


Fig. 7. (a) Measured mean temperature (solid line) and reconstructed mean temperature (circle); (b) standard deviation error of reconstructed mean temperature.

show good agreement. We computed the standard deviation error, arising from the reconstructed field with 5 EOFs, at each CTD cast location and averaged it over all the casts. The standard deviation error is $\sqrt{\bar{\epsilon}}$, where $\bar{\epsilon} = \frac{1}{MN} \sum [A_m - A_m^r]$ and the summation is over M rows and N columns (CTD stations) involving all water masses m and A_m^r represents the reconstructed water mass variance summed over 5 EOFs.

The results are shown in Fig. 7b. The largest error occurs at about 35 m, which is at the location where the surface water masses overlap the MIW water masses. We attribute this to the neglect of cross terms during the summation of reconstructed temperatures from each water mass. A smaller peak in the error field occurs at the overlap region between the MBW and SLW water masses.

We have compared the vertical profiles reconstructed from the water mass summation approach over 5 EOFs to vertical profiles reconstructed from the 5 EOFs of the entire water column, without decomposition into water masses. The analogous standard deviation error yields two peaks at 10 and 25 m of about 0.25 magnitudes on the scale of Fig. 7b. This error decreases versus depth to value of approximately 0.1 at 75 m and remains at about the same level. The two peaks near the surface reflect the higher variance of temperature there. The errors in the water mass overlap regions disappear, since there is no water mass overlap zone. The tracking of the individual water mass variances has versatility and enables one to update the clusters representing the water masses with new data.

The use of about 5 EOFs for representing each water mass is sufficient for the reconstruction of the temperature and salinity fields. When new measurements of temperature and salinity are taken, they can be assimilated into the existing data set and the computations of the EOFs repeated. If the new measurements are in the surface region, then only the surface region EOFs need be recomputed in view of the new data. This encompasses satellite, hydrographic, drifter buoy and any other platform measurement. Alternatively, if the measurements are in one particular water mass, only in that region need the coefficients be recomputed as is the case of autonomous

underwater vehicle measurements. This water mass-based approach constitutes a feature model for temperature and salinity in local areas of interest.

3.2. Covariance and correlation statistics

We have analyzed the covariance matrix, AA^T , of temperature data in the Wilkinson Basin area. The matrix is formed from the data using a 5 m vertical resolution. The elements are expressed as a variance by subtracting the horizontally averaged mean.

The covariance can be expressed in terms of the water mass covariance as $AA^T = (A_{MSW} + A_{MHW} + A_{MIW} + A_{MBW} + A_{SLW})(A_{MSW} + A_{MHW} + A_{MIW} + A_{MBW} + A_{SLW})^T$. We considered the separate covariance, correlations and statistics of each water mass, such as $A_{MIW}A_{MIW}^T$. This consideration neglects the cross terms between the water mass in question and the other water masses.

One can also form a normalized correlation matrix as:

$$\text{cor } B = \frac{AA^T}{\sqrt{\text{var var}}}, \quad (4)$$

where var is the variance of AA^T along the diagonal. Fig. 8a displays the temperature correlation matrix, of all the rows of A with themselves and one another, as a function of depth (the columns or stations are summed in the elements).

Along the sides of Fig. 8b, the correlation of the rows with themselves is exhibited. The correlation decreases, raises, goes through zero at 50 m, bottoms out in the negative range, and returns to zero at about 150 m. This suggests positive correlation scales of around 50 m and a total correlation of 150 m for the vertical length scales.

The correlations of the whole field and the separate water masses are shown in Fig. 8b. The clusters of the separate water masses contain fewer points and are distributed over shorter vertical depths. Their distribution in depth reflects their location as a function of depth. The MHW water mass correlation drops to zero at about 12 m, while the MSW water mass decreases appreciably at about 32 m. The deeper water masses have very

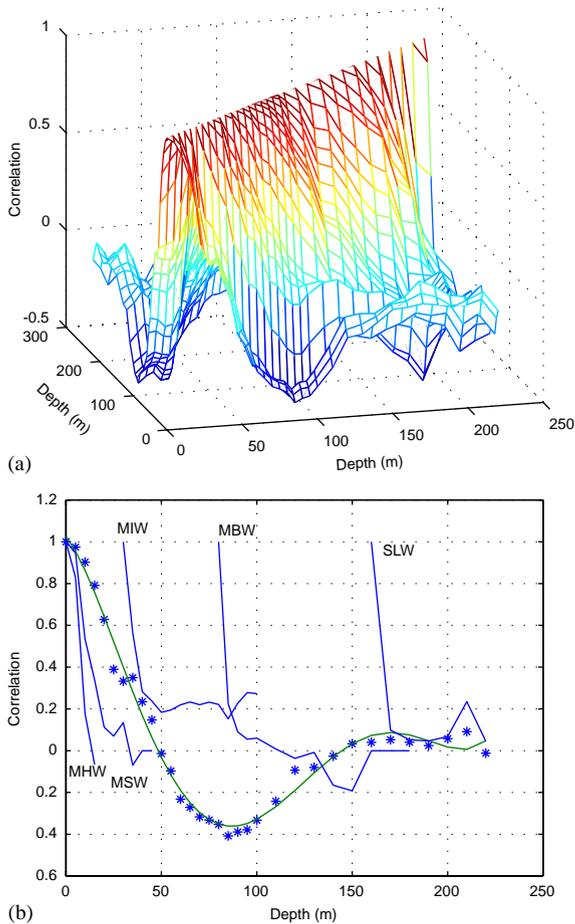


Fig. 8. (a) Temperature correlation matrix as a function of depth summed over CTD stations and (b) the correlations of the whole field and the separate water masses as a function of depth.

sharp initial drops that can be extrapolated to zero, about 10 m for MIW, 7 m for MBW, and 10 m for SLW. The MBW correlation turns negative and then returns to zero. The total correlation length for these water masses tends to extend over the depth span of the following water masses: 70 m for MIW, 80 m for MBW, and 55 m for SLW.

3.3. The role of EOF analysis for feature models

The core temperature or salinity distribution [$\psi_c(z)$] for the Wilkinson Basin area was given by Gangopadhyay et al. (2003) using the following

simple form:

$$\psi_c(z) = [\psi_o - \psi_b]\Phi(z) + \psi_b. \quad (5)$$

Here, ψ_o is the surface value, ψ_b is the bottom value, and $\Phi(z)$ is the non-dimensional vertical structure.

In principle, vertical EOFs could be used to determine the vertical structure function $\Phi(z)$ for any gyre/current system. In this study, we have applied an EOF analysis for the Wilkinson Basin area to derive $\Phi(z)$. This is a step beyond the simple analytical/empirical functional forms used in earlier studies by Gangopadhyay et al. (1997), Robinson et al. (2001), and Gangopadhyay et al. (2003). Furthermore, the advantage here is that new information obtained in real-time (on-board or via air-borne sensors) can be included in the EOF analysis to modify the constituents of the water mass eigenmodes.

This approach can also be extended to include horizontal variation along a current or across an eddy/gyre system to provide the x -dependence in $\Phi(x, z)$ in the generalized form of Eq. (5). (See Eq. (1)–(10) of Gangopadhyay et al. (2003) for variations of this function.) Vertical EOFs along stream might have variable eigenvalues along the axis of the current. The vertical depth dependence of the constituent water masses will also be available.

Recently, Gangopadhyay and Robinson (2002) have presented a generic methodology to apply the feature-oriented approach to any front, eddy or gyre system in any oceanic region. Our long-term goal is to be able to apply the presented EOF-based methodology to any region in the world's ocean. From a broader perspective, our study is a prelude to a more general study to develop EOF-based feature models for the whole of Gulf of Maine and Georges Bank circulation system.

In a general sense, for any new region, an EOF analysis of the water mass temperature, salinity, and percentage of volume occupied can be undertaken. The cloud of points representing each water mass by temperature, salinity, and depth are derived through cluster analysis and identified as matrix elements in the SVD and covariance analysis. Covariance and correlation statistics of the water masses are analyzed. The vertical

temperature and salinity are reconstructed through an EOF decomposition of each water mass' cloud of points and a summation over all of them for the resultant vertical distribution. The structure of the water mass variances can be projected onto an orthonormal set, through coordinate transformation or the Gram Schmidt procedure. This constitutes the basis for a water mass-based feature model with latest measurement update capability. The percent of volume occupied, or occurrence, provides guidance for the structure of the EOF-based feature model. The vertical eigenmodes are possibly linked with the prevalent water masses, and the corresponding eigenvalues might indicate their relative contribution to the composite water mass at a particular depth level. The quantification of such connectivity will be considered in a later study.

When new measurements of temperature and salinity are taken, they can be assimilated into the existing data set and the EOFs will be recomputed. If the new measurements are in the surface region, then only the surface region EOFs need to be recomputed in view of the new data. This encompasses satellite, hydrographic, drifter buoy and any other platform measurement. Conversely, if the measurements are in one particular water mass, only in that region do the coefficients need to be recomputed. This water mass-based EOF approach constitutes a feature model for temperature and salinity vertical distribution in local areas of interest.

Two-dimensional EOF analysis has been used for tracking satellite-sensed sea surface variability in the Gulf of Maine (Bisagni et al., 1996). The spatial and temporal variability was tracked for three years from 1993 to 1996. The first mode variance reflected the spatial structure in the seasonal heating cycle. The first mode spatial structure was associated with differences in mixed layer depth caused by tidal motion in shallow and deeper water. The third mode was related to regional differences in seasonal thermohaline and density characteristics. The second mode describes the inter-annual variability generated by cooling trends within the slope water region; this was caused by a decrease in warm-core ring activity during the study period. The surface SST and

features of currents and gyres derived from such an approach can be incorporated into a feature model of the region.

In the future, we will apply the EOF-based feature models along with other synoptic feature models, in regions of shallow water for phenomena such as internal solitary wave generation and propagation in the Yellow Sea, the Strait of Messina (Warn-Varnas et al., 2003), the Gulf of Gioia, the Primer shelf area, the Strait of Luzon and the ASIAEX area. In some cases, not enough data is available to guide the analysis. Even when there is data, the feature models will fill in the physical picture. EOF-aided feature models will thus help in cases of sparse data, among other situations.

In summary, we have presented an EOF-based approach for developing feature model vertical structures of a gyre system like that in the Wilkinson Basin area. We have also identified the EOFs and their relative contributions for constructing the water masses in a new environment.

4. Water mass distribution

4.1. Volumetric analysis

We considered the vertical extent or percent of the water column occupied by each water mass at the CTD stations in the Wilkinson basin area. The vertical extent or depth range occupied by each water mass was obtained from the cluster analysis. Then the percentage of the water column occupied by each water mass was computed relative to the column depth as

$$\alpha_m = \frac{d_m}{D}, \quad (6)$$

where d_m is the particular water mass depth range and D is the depth of the water column at the station.

Fig. 9a displays the percentage of the water column occupied by each of the water masses at the CTD stations together with the total fraction of water occupied by the particular water mass at all of the Wilkinson basin area CTD stations in

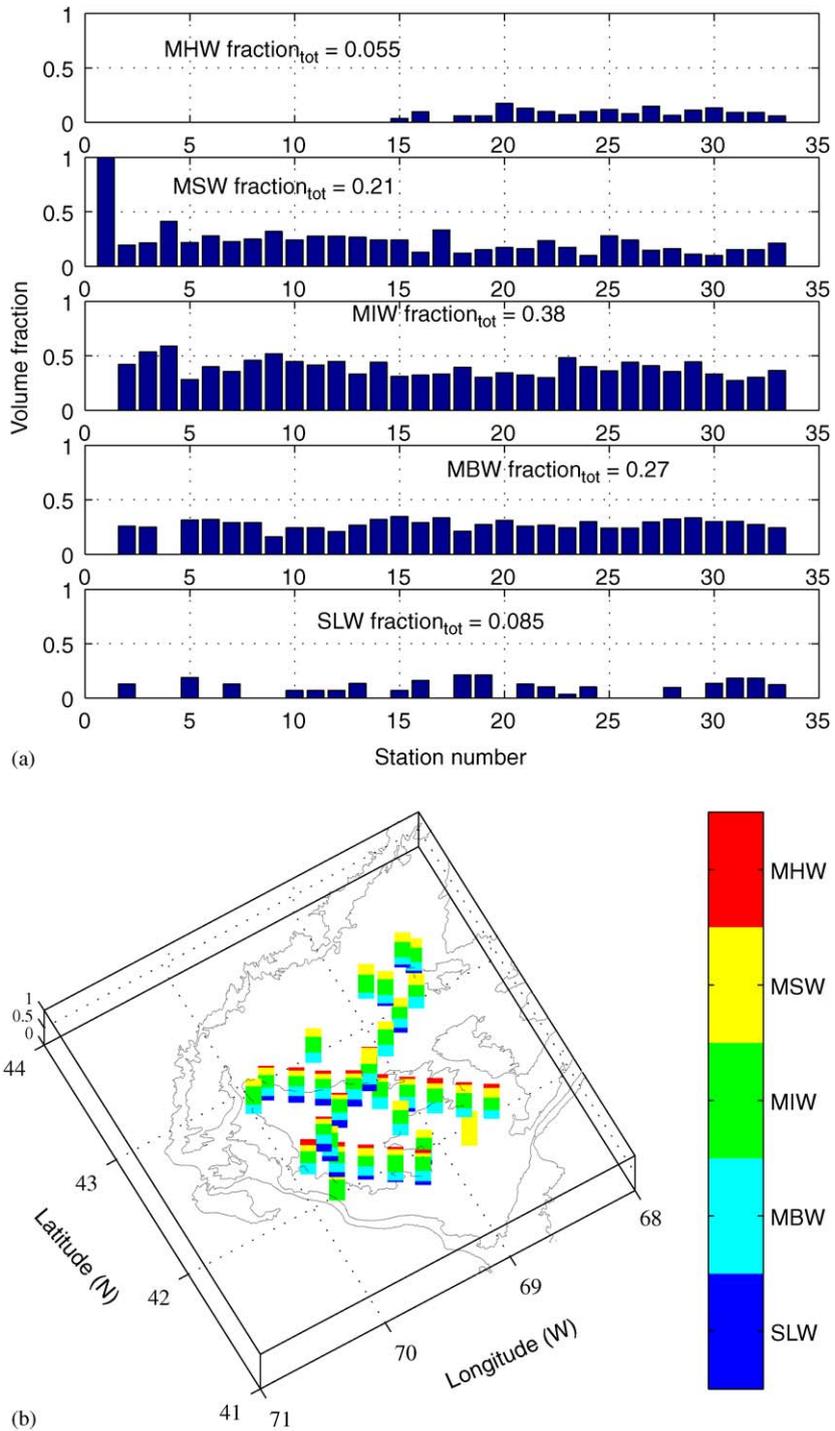


Fig. 9. (a) The percentage of the water column occupied by each of the water masses at the CTD stations. The total fraction of all water occupied by the particular water mass is shown on top. (b) Results of (a) shown at the geographic station location in terms of water mass percentage.

terms of bar diagrams. Some of the stations will not contain a particular water mass.

The MIW water mass occupies the largest percentage of water columns, with a total fraction of 0.38 or 38%. The MBW water mass follows with 27%. The MSW water mass accounts for 21% of the surface water. The MHW water mass occupies 5.5% of the water columns and is found only at some of the CTD stations. The SLW water, also, is not present at all of the CTD stations and occupies 8.5% of the water columns.

4.2. The dynamics of volumetric analysis

The origin of water masses is summarized in Table 4. The surface water mass originates with an inflow of Scotian Shelf Water and local fresh water runoff. The surface water characteristics, MSW water, are controlled by seasonal surface warming and cooling (Hopkins and Garfield, 1979). Local wind stress and heating events can affect the temperature and salinity structure.

The anomalous water mass called MHW was observed along both transects in the southern half of the domain. The mean salinity of this water mass is near 32.00 ppt (see Table 2). This suggests a Scotia Shelf Water origin (Flagg, personal communication). It could be that a branch of the MCC (Fig. 1) brings the fresher low salinity water into the western Wilkinson Basin area. The temperature and salinity range for the MCC in the Scotia Shelf region in June are about 3–8 °C

and 31.50–32.20 ppt, respectively (Gangopadhyay et al., 2003, Fig. 4(d), Table 3). The higher observed temperature range for MHW (10.55–14.12; Table 2) could be due to continuous solar-heating during its advection.

The MIW water is formed during winter cooling events and is exported from the formation region by advection. The MBW water is derived from SLW water and also by mixing with MIW. The SLW water originates from the Gulf Stream and is transported in through the Northeast Channel. Sometimes eddies impinge on the shelf from the Gulf Stream and contribute to the SLW characteristics.

The results of volumetric analysis (Fig. 9a) can be interpreted in a geographical perspective. Analysis of Fig. 9b indicates that three major water masses—MSW, MIW and MBW—were found during the survey period at most stations, i.e., throughout the Wilkinson Basin area. The occurrence is a characteristic of the Wilkinson Basin and surrounding area. MBW was not observed at station 4, which is at the Great South Channel. Most of the other stations in the deeper and shallower sections of the Wilkinson Basin area show the presence of MBW water.

Except for the MHW, the other four water masses are present throughout the extended basin area (Fig. 9a—top panel). The occurrence of SLW is restricted to the deeper troughs of the basin, as indicated by the blue bars superimposed on the bathymetry in Fig. 9b. Its presence is more sparse

Table 4
Origin of water masses

Water mass	Origin/formation	Controlling processes [Formation] {modification}
MSW	Inflow of Scotian Shelf water and local fresh water runoff	[{Wind stress, solar heating, latent and sensible heat flux}]
MHW	Atmospheric forcing acting on surface region Branch of Maine coastal current bringing fresher water into the western Wilkinson basin	{Advection, mixed layer dynamics} Solar heating, wind-driven forcing
MIW	Winter cooling	[Convection/atmospheric buoyancy exchange processes] {mixing, export through advection}
MBW	Derived from SLW and mixing with MIW	[Turbulent mixing in shear zone] {Isopycnal and diapycnal mixing}
SLW	Transport of SLW through the northeast channel	[Mesoscale dynamics of Gulf Stream] {eddies impinging on shelf}

than the MBW water. Slope Water did not reach the northeastern stations (9, 14 and 17) and was absent from stations 25, 26 and 27 for some unknown reason. The Slope Water could come in through the deeper bathymetry sections from the Georges Bank gyre region. We pursued this possibility by analyzing a prediction for the GOM region with a feature-oriented modeling approach (Gangopadhyay et al., 2003). The predicted salinity along a path from the Northeast Channel to the Wilkinson basin area is shown in Fig. 10 for March 2002. During the year, the monthly variation of salinity below 100 m is small and of the order of 0.25 ppt in the Wilkinson Basin area. Therefore, the same type of salinity structure is expected through the year at below 100 m (Gangopadhyay et al., 2003). Decadal variations in temperature and salinity values do occur (Brown and Irish, 1993; Hopkins and Garfield, 1979). These variations cause differences in specific

water mass temperature and salinity ranges. According to Table 2, the Slope Water salinities range from 33.68 to 34.27 ppt, with a mean of 34.01 ppt at a depth span of 160–220 m. In Fig. 10, the 34.00 ppt contour follows the topographic variations by raising and descending over topographic ridge-type features. This indicates an overflow of SLW mass over topographic ridges from the Northeast Channel into the Wilkinson Basin area. The situation is analogous to the creeping type of overflow of Levantine water over a ridge predicted in the Strait of Sicily (Warn-Varnas et al., 1999).

The whole water column at the location of Station 1 was comprised of a single water mass, the MSW. This station is a shallow (31 m) location at the northern flank of Georges Bank. The MSW and MHW water masses are located in the upper 45 m of the ocean (Table 2) and are found at all of the stations (Fig. 9b). The vertical distributions of volume between MSW and MIW in the Great

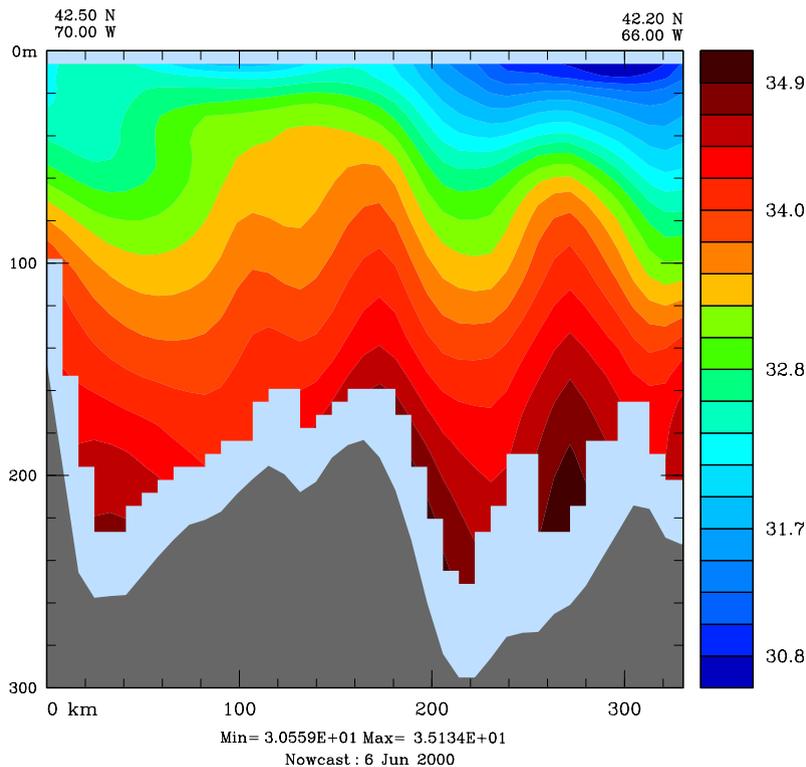


Fig. 10. Model predicted salinities with the Harvard Ocean Prediction System along transect from Northeast Channel to Wilkinson Basin area for March of 2002.

South Channel were about 40% and 60%, respectively. Such numbers can be used for designing an EOF-based feature model associated with the water masses.

5. Conclusion

The Wilkinson Basin area water mass characteristics were derived from the Brook's June 1982 survey (Brooks, 1985, 1990) by two different approaches. In one model, the water mass parameters were identified by their salinity and temperature profile structure (Warn-Varnas, 1999). In another model, cluster analysis was used to derive the water mass parameters through grouping (Kim et al., 1991; Hur et al., 1999). In both models, the means and standard deviations of the cluster of points that represent the water masses in temperature, salinity, and depth were derived. Away from the surface region, both approaches yielded comparable results within standard deviation errors. At the surface, the cluster analysis model yielded a MSW mass and a MHW water mass that was warmer by about 4 °C.

The percentage of volume occupied by each water mass was computed at the CTD cast and for the Wilkinson Basin area. A dynamic interpretation of the percentages of volume occupied by the water masses at the CTD casts was provided. It was found that the three major water masses are MSW, MIW, and MBW. The percentages of volume they occupied were 21%, 38%, and 27%, respectively. The origin of the SLW mass was considered, and it was shown to result from overflow over bathymetric ridges into the Wilkinson Basin area through the Northeast Channel, where the SLW mass flows into the region.

We have presented a variance-based EOF analysis associated with water masses in the Wilkinson Basin area and related it to feature models. This EOF-based approach will allow for consistent water mass dependent selection of the vertical structure in feature models.

Acknowledgements

The US Office of Naval Research through the US Naval Research Laboratory base program

supported this project. The US Naval Research Laboratory (at the Stennis Space Center) provided technical management. This research was partially funded by the National Aeronautics and Space Administration (NAG 13-48) and partially by the Office of Naval Research Grants N00014-03-1-0411 and N00014-03-1-0206 at the University of Massachusetts at Dartmouth. The team is grateful to all who contributed to its endeavors.

References

- Bierweiler, T., 1999. A data based model for the cyclonic gyre system in the Gulf of Maine, Graduate Research Report in Physics, University of Massachusetts Dartmouth, MA
- Bigelow, H.B., 1927. Physical Oceanography of the Gulf of Maine. Bulletin of Bureau for Fisheries US 40, 511–1027.
- Bisagni, J.J., Beardsley, R.C., Ruhsam, C.M., Manning, J.P., Williams, W., 1996. Historical and recent evidence concerning the presence of Scotian Shelf water on southern Georges Bank. *Deep-Sea Research* 43, 1439–1471.
- Bjornsson, H., Venegas, S.A., 1997. A manual for EOF and SVD analyses of climatic data, Technical Report from Department of Atmospheric and Oceanic Sciences and Center for Climate and Global Change Research, McGill University, Canada.
- Brooks, D.A., 1985. Vernal circulation in the Gulf of Maine. *Journal of Geophysical Research* 90, 4687–4705.
- Brooks, D.A., 1990. Currents at Linden Kohl Sill in the southern Gulf of Maine. *Journal of Geophysical Research* 95, 22173–22192.
- Brown, W.S., Beardsley, R.C., 1978. Winter circulation in the western Gulf of Maine, part I. Cooling and water mass formation. *Journal of Physical Oceanography* 8, 265–277.
- Brown, W.S., Irish, J.D., 1993. The annual variation of water mass structure in the Gulf of Maine: 1986–1987. *Journal of Marine Research* 51, 53–107.
- Colton, J.B., 1968. Recent trends in subsurface temperatures in the Gulf of Maine and Contiguous waters. *Journal of the Fisheries Research Board of Canada* 25, 2427–2437.
- Fukumori, I., Wunsch, C., 1991. Efficient representation of the North Atlantic hydrographic and chemical distributions. *Progress in Oceanography* 27, 111–195.
- Gangopadhyay, A., Robinson, A.R., Arango, H.G., 1997. Circulation and Dynamics of the Western North Atlantic, I: multiscale feature models. *Journal of Atmospheric and Oceanic Technology* 14 (6), 1314–1332.
- Gangopadhyay, A., Robinson, A.R., 2002. Feature oriented regional modeling of oceanic fronts. *Dynamics of Atmosphere and Oceans* 35 (4), 203–235.
- Gangopadhyay, A., Robinson, A.R., Haley, P.J., Leslie, W.G., Lozano, C.J., Bisagni, J.J., Yu, Z., 2003. Feature-oriented regional modeling and simulation (FORMS) in Gulf of

- Maine and Georges Bank. *Continental Shelf Research* 23 (3–4), 317–353.
- Hannah, C.G., Loder, J.W., Wright, D.G., 1996. Seasonal variation of the Baroclinic circulation in the Scotia Maine region. In: *Buoyancy Effects on Coastal and Estuarine Dynamics*, Coastal and Estuarine Studies, vol 53, American geophysical Union, pp. 7–29.
- Hopkins, T.S., Garfield, N., 1979. Gulf of Maine intermediate water. *Journal of Marine Research* 37, 103–139.
- Hur, H.B., Jacobs, G.A., Teague, W.J., 1999. Monthly Variations of Water Masses in the Yellow and East China Seas, November 6, 1998. *Journal of Oceanography* 55, 171–184.
- Kim, K., Kim, K.R., Rhee, T.S., Rho, H.K., 1991. Identification of water masses in the YS and East China Sea by cluster analysis. In: Takano, K. (Ed.), *Oceanography of Asian Marginal Seas*. Elsevier, Amsterdam, pp. 253–267.
- Lynch, D.R., 1999. A review of modeling in the Gulf of Maine. *Naval Research Review* 61(1).
- Nittis, K., Pinardi, N., Lascaratos, A., 1993. Characteristics of the Summer 1987 flow filed in the Ionian Sea. *Journal of Geophysical Research* 98 (C6), 10171–10184.
- Robinson, A.R., Rothschild, B.J., Leslie, W.G., Bisagni, J.J., Borges, M.F., Brown, W.S., Cai, D., Fortier, P., Gangopadhyay, A., Haley Jr., P.J., Kim, H.S., Lanerolle, L., Lermusiaux, P.F.J., Lozano, C.J., Miller, M.G., Strout, G., Sundermeyer, M.A., 2001. The development and demonstration of an Advanced Fisheries Management Information Systems (AFMIS). *Bulletin of the American Meteorological Society*, 186–190.
- Warn-Varnas, A., Sellschopp, J., Haley, P., Leslie, W., Lozano, C., 1999. Strait of Sicily water masses. *Dynamics of Atmospheres and Oceans* 29, 437–469.
- Warn-Varnas, A.C., Chin-Bing, S.A., King, D.B., Hallock, Z., Hawkins, J.A., 2003. Ocean–Acoustic soliton studies and predictions. *Surveys in Geophysics* 24, 39–79.
- Xue, H., Chai, F., Pettigrew, N.R., 2000. A model study of the seasonal circulation in the Gulf of Maine. *Journal of Physical Oceanography* 30, 1111–1135.