LIST OF FIGURES

LIST OF TABLES

Chapter 1 INTRODUCTION

The same stream of life that runs through my veins night and day, runs through the world and dances in rhythmic measures.

It is the same life that shoots in joy through the dust of the earth in numberless blades of grass, and breaks into tumultuous waves of leaves and flowers.

It is the same life that is rocked in the ocean-cradle of birth and of death, in ebb and in flow.

I feel my limbs are made glorious by the touch of this world of life.

And my pride is from the life throb of ages dancing in my blood this moment.

Tagore

A fundamental problem in the geophysical sciences is how to use data collected at a finite number of locations, and eventually at different times, to estimate values at any point of the space or of the space-time continuum. The ultimate aim of such estimate can be a simple visualization of the observed field, the diagnosis of physical processes of the use of the estimated field as input for numerical models, among other applications. A wide set of techniques have been developed for both diagnostic and prognostic analyses. Since the development of computers in the last decades made possible the automatic implementation of these techniques, they have been referred to as spatial objective analysis.

This work focuses mainly on the analysis of hydrodynamic variables in the shelf and slope areas of the Ebro Delta region, where a density shelf/slope front induces the presence of a geostrophic jet known as the Northern Current. The two- and three-dimensional distributions of these variables will be obtained through the application of spatial objective analysis techniques.

More precisely, the motivation for this work is to investigate the potential use of Empirical Orthogonal Functions (hereafter referred to as EOFs) altogether with a 2D spatial analysis technique (the Successive Corrections scheme, hereafter referred to as SC) to infer the three-dimensional distribution of observed variables. A related motivation is to investigate the estimation of the

dynamic height field in the presence of a shallow and sharp bathymetry. This is a crucial issue for diagnostic studies, since all the dynamical processes than can be inferred in the framework of the geostrophic and quasi-geostrophic theories rely on a proper estimation of the dynamic height field. The results obtained using the EOF-based formulation will be compared with those obtained using a more classical method proposed by Csanady for shallow shelf domains. Finally, another aspect dealt with in this thesis work is the potential of historic data (in front of data from a single oceanographic cruise) to give more statistical robustness to the EOF analyses.

While the above analyses have been successfully used in open ocean domains, a shallow shelf with an abrupt slope represents an interesting challenge, particularly for the geostrophic current estimates. In the Ebro Delta region, most of the shelf has depths shallower than 100 m, while the vertical integration of dynamic height must be performed from a reference level that cannot be shallower than 500 meters to be reasonably considered as a no-motion level. This implies that only 20%, or less, of the water column will be available. The remaining 80% or more will have to be inferred from deeper water data from the outer shelf, where the physical forcing mechanisms are often different than those governing the shelf dynamics.

Generally speaking, the most important forcing mechanism on the shelf are:

Atmospheric pressure and the resulting wind field.

The temperature and salinity distributions, which determine the density field. A distinctive feature in the study domain will be the fresh water discharge of the Ebro River.

Topographic and lateral boundary effects.

While the Ebro River runoff is a clear local forcing that has a strong effect on the temperature and salinity distributions, preferential heating and topography can also have a significant effect on the shelf variable distribution and dynamics.

With the above aspects in mind, the data analyzed in this work are essentially near-synoptic CTD profiles from four oceanographic campaigns:

three obtained in the framework of the FANS project, with a common distribution of CTD stations, and an additional one from the MEGO 94 project. [From now on the campaigns will be referred to as FANS I (carried out in November 1996), FANS II (in February 1997), FANS III (in July 1997) and MEGO (in March 1994).]

The general objectives of the European Community FANS (Fluxes Across Narrow Shelves) project were to improve the existing knowledge on the hydrodynamic, sedimentary and nutrient fluxes in the Ebro Delta shelf/slope region (Figure 4-1), as well as to study their interaction. The MEGO 94 campaign was part of the project "Dinámica de las Fluctuaciones de Mesoescala en la Plataforma Continental Catalano-Balear", whose goals were to study the physical mechanisms controlling the variability of the circulation on the Catalan Continental Shelf. The campaign focused mainly in the Gulf of Sant Jordi, with a dense CTD grid around the Ebro Delta region.

The inclusion of the MEGO campaign was actually aimed to the use of ADCP data to test a multivariate analysis (which allows to estimate both the geostrophic and ageostrophic components of the flow), but unfortunately the data turned out to be very noisy and not reliable enough for this purpose, as it had happened with the corresponding FANS II and FANS III ADCP data sets (the ADCP was not operative during the FANS I campaign). This data void be admitted as the main drawback of this work, since it will prevent to check the geostrophic circulation inferred by the different techniques against actual current data.

1.1 Objectives

The principal goal of this Thesis work is to apply an EOF-based spatial objective analysis techniques to the study of mesoscale oceanographic features in the Ebro Delta shelf/slope region. In particular, we focus on:

The use of Empirical Orthogonal Functions and the Successive Corrections algorithm to generate a 3D grid of observed variables.

The study of the potential use of historic data (in addition to single cruise data) to estimate the above fields.

The estimation of the geostrophic circulation in the Ebro Delta shelf/slope region from the observed 3D fields.

Compare the geostrophic circulation given by the EOF-based method with that given by the method proposed by Csanady.

The presentation has the following structure. Chapters two and three contain the basic aspects of the EOFs and the SC methodology. In chapter four we present the overall oceanographic aspects of the Ebro Delta shelf/slope region, as reported in state-of-the-art bibliography. Chapter five focuses on the data set used all along this work, from the set of CTD cast obtained during the campaigns to the data quality control procedures. It also includes complete information on the historical data from which the historical EOF modes are obtained.

The distribution of observed variables obtained from the analysis techniques are studied in the remaining chapters. In chapter 6, the largest, we first present $q - S$ diagrams (section 6.1) and the on-grid distributions of potential temperature, salinity, density and dynamic height (section 6.2). While the $q - S$ diagrams contain all observed data (with a vertical resolution of one meter), the 3D grid used to represent the fields has a vertical resolution of 10 meters (the horizontal resolution is 2.5 kilometres in both directions). In a third section, we compare the results obtained from the two EOF approaches: the vector modes derived from single cruise data (section 6.3), and those derived from historical data (section 6.4).

The geostrophic current results are presented in chapter 7. A brief introduction on the geostrophic theory (section 7.1) is followed by the presentation of Csanady's model (section 7.2) and the resulting dynamic thickness and circulation derived from it. In section 7.3 we present the dynamic thickness and currents that result from the EOF-based method.

The final sections are chapter 8, where the basic conclusions are outlined, and chapter 9, where some potential areas where this line of research can be continued are pointed out.

Chapter 2 EMPIRICAL ORTHOGONAL FUNCTIONS (EOF 'S)

2.1 Introduction

In its most general range of application, the main aim of EOF analysis is to compress the information contained in a certain data set into the smallest number of independent pieces of information. This procedure is equivalent to other data reduction methods such as principal component analysis (PCA) or factor analysis. The first reference to the application of EOF analysis to geophysical fluid dynamics seems to be a report by Edward Lorenz in 1956, in which he develops the technique for statistical weather prediction and actually coins the term "EOF".

A first comment, which is particularly relevant to the present application of EOF analysis, is to be made on the meaning of univariate/multivariate analysis. In geophysical fluid dynamics, "multivariate" refers to the analysis involving more than one physical variable and therefore considering the relationships between the variables as a crucial part of the analysis. In standard PCA, "multivariate" refers also to different variables, but here "variable" can refer for instance to the values of a unique physical variable at different levels. In this sense, PCA is always considered a "multivariate" analysis.

A second relevant comment is on the main result of EOF analysis: the partitioning of the data variance (this being temporal and/or spatial) into a set of "modes" (the EOFs themselves). When interpreting the meaning of these modes it should be clear that there is not necessarily a direct physical or mathematical relationship between them and any dynamical mode. Instead, EOFs are precisely referred to as "empirical" to emphasise that they are only determined by the covariance or correlation structure of the specific data set being analysed, without any explicit dynamical constraint. Hence, a single physical process may be spread over several EOFs, and also more than one physical process may be contributing to the variance of a single EOF. In fact, the method can be considered as an observation-based generalisation of a feature model approach.

Two major advantages of an EOF-based description of a data set are: a) it provides a very efficient method for data compression; b) the EOFs can be regarded as uncorrelated (i.e. orthogonal) modes of the data variability. In the context of objective analysis of hydrodynamical variables, two additional benefits are envisaged: c) EOFs can constitute a robust base for a consistent 3D interpolation; and d) EOF analysis can also become a statistically based extrapolation method (particularly useful for dynamic height analysis).

In this section we present the basic theory behind EOF analysis. It is developed in the framework of the particular application relevant to this work, i.e., the determination of the ocean vertical variability modes. Another usual application in oceanography is in the field of satellite imagery, namely to determine the horizontal (time-dependent) modes of surface variability. Most of the formulation follows the works by Haney et al. (1995) and Pedder and Gomis (1999). The first of these works focused on the feasibility of extrapolating the density field below the depth reached by CTD casts. The second focused mainly on the potential of EOFs as a consistent method for 3D analysis of observed variables to be used as input in diagnostic studies. In this context, the method was also used to infer dynamic height field in regions where data did not extend down to the reference level.

Our study is more on the line of the work by Pedder and Gomis (1999). Namely, we will use the method to carry out 3D interpolations of dynamically relevant fields (potential temperature, salinity, density and dynamic height). In this framework, the most challenging issue will be to test the capabilities of the method to estimate the geostrophic circulation in the presence of abrupt bathymetry.

The basic problem we face is how to estimate dynamic height on the continental shelf, where density profiles do not reach the depth considered as a reasonable reference level for geostrophic computations. [Hereafter profiles reaching the reference depth will be referred to as *deep* or *complete.* Conversely, the others will be referred to as *shallow* or *incomplete*.] A main underlying assumption to be tested will be that the dynamics of the deep region is representative of the whole spatial domain.

A secondary aim will be to compare the results obtained using vertical modes derived from historical data with those obtained using the modes as determined only from the data set being analysed.

2.2 Formulation

When the EOF analysis is applied to sets of quasi-synoptic CTD surveys, each vertical profile of any given variable j (such as potential temperature, salinity, specific volume anomaly or other) can be represented as

$$\underline{j}_j = <\underline{j}> + \sum_{m=1}^{L} a_{m,j} \, \pmb{p}_m \tag{2.1}$$

where $\pmb{j}\left(x_j, y_j, z_i\right)$ represents the j - th profile data vector (with L depth levels) located at $\left(x_j y_j\right)$, $<\underline{j}>$ is the mean profile calculated from the N deep soundings, $a_{m,j}$ is the expansion coefficient or amplitude for the j - th profile associated to the m - th eigenvector and $\pmb{p}_m$ is the m - th eigenvector or EOF. This is, the j - index is associated with horizontal space and the i - index with the vertical. In this way, each profile can be represented as the sum of m independent modes, each of them accounting for a known percentage of the variability of the signal.

The eigenvectors or EOFs (altogether with the corresponding eigenvalues) are obtained either from the covariance matrix (in which case the analysis will hereafter be referred to as ***non standardized analysis***) or from the correlation matrix (then being referred to as ***standardized analysis***). Both matrices can be obtained as follows.

Let $\pmb{j}\left(x_j, y_j, z_i\right)$ be any given variable measured at station or cast j, with horizontal coordinates $\left(x_j, y_j\right)$ and vertical coordinate z_i. Let $<\pmb{j}\left(z_i\right)>$ be the vertical average vector and $\pmb{s}\left(z_i\right)$ the variance at every level. The covariance matrix $C_{i,k}$ of increment profiles (i.e., the difference between actual profiles and the mean profile) is defined as:

$$C_{i,k} = \frac{1}{N}\left[\sum_{j=1}^{N}\left(\pmb{j}_{j,i} - <\pmb{j}_i>\right)\left(\pmb{j}_{j,k} - <\pmb{j}_k>\right)\right] \tag{2.2}$$

Similarly, the correlation matrix is given by:

$$C_{i,k} = \frac{1}{N}\left[\sum_{j=1}^{N}\left(j_{j,i} - <j_i>\right)\left(j_{j,k} - <j_k>\right)/\left(s_i s_k\right)\right] \tag{2.3}$$

where N is the number of deep casts available for a given data set. Thus, we could think of the covariance and correlation matrix elements as giving a measure of the statistical link between the values of a given variable at two different levels. The aim of the whole process is to take advantage of these links to reduce the dimensions of the problem.

An important difference between the covariance and correlation matrices is that, for the former, largest values usually correspond to covariance between levels close to the surface layers, where the variability of oceanographic fields is usually larger. For the correlation matrix, instead, the "weight" of the different levels with respect to the total variance does not depend on the level variability, since the matrix elements are normalized precisely by the standard deviations of the corresponding levels (e.g., all the elements of the diagonal are equal to unity for the correlation matrix), This difference will be emphasised later on, when looking to the differences between the two options.

Both matrices are real, symmetric and positive definite, so that their eigenvalues λ must be also real and positive and their eigenvectors π form an orthogonal basis of the L-dimension vector space:

$$\boldsymbol{p}_i \cdot \boldsymbol{p}_j = \left|\boldsymbol{p}_i\right|\left|\boldsymbol{p}_j\right|d_{ij} \tag{2.4}$$

where δ_{ij} is the Kronecker delta. Finding out the eigenvectors and eigenvalues is a matter of solving the system

$$\underline{C}\cdot\boldsymbol{p} = l\,\boldsymbol{p} \tag{2.5}$$

This has a non trivial solution only if

$$Det\left(\underline{C} - l\,\underline{I}\right) = 0 \tag{2.6}$$

which is an *N-th* degree polynomial in lambda whose roots are the eigenvalues. In principle, some of the L eigenvalues could be identical (multiple roots), in which case the system is referred to as "degenerate", and the eigenvectors would no longer be a base of the L -dimensional vector space. In practice, however, degenerate systems are very unlikely.

In our particular case, we use the numerical canned routine EVCSF from the IMSL (International Mathematics Scientific Library) package to solve the eigensystem (5). As usually, this routine reports eigenvalues in decreasing order of magnitude and each eigenvector is normalised to have euclidean length equal to unity.

It is not difficult to see that the diagonal elements of the covariance matrix are

$$C_{k,k} = \frac{1}{N}\left[\sum_{j=1}^{N}(j_{j,k} - <j_i>)(j_{j,k} - <j_k>)\right] = s_k^2 \qquad (2.7)$$

so that its trace is equal to the total variability of the field. Because the trace is an invariant to diagonalisation, it then turns out that

$$\sum_{i=1}^{N} l_i = Tr\underline{C} = \sum_{k=1}^{N} s_k^2 \qquad (2.8)$$

From this, it is not difficult to infer that the variance explained by the *i-th* mode is given by $l_i/Tr\underline{C}$. For the correlation matrix relationship (8) becomes

$$\sum_{i=1}^{N} l_i = Tr\underline{C} = N \qquad (2.9)$$

Once the *N* eigenvectors are found, any complete profile can be *exactly* expressed as a weighted sum of them, as given in (1). But moreover, usually the sum of a few (*M*) leading modes is enough to account for most of the variability of the field, so that eq. (1) becomes

$$\hat{\jmath}_j \approx <j> + \sum_{m=1}^{M} a_{m,j} p_m \qquad (2.10)$$

where $\hat{\jmath}_j$ is the estimated profile. The associated amplitudes $a_{m,j}$, can be obtained simply as the scalar product of the *m-th* EOF vector and the *j-th* increment profile

$$a_{m,j} = p_m \cdot j_j \qquad (2.11)$$

Profile Extrapolation

Obtaining the amplitudes for shallow casts (extending only down to *D* depth levels, where *D < L*) is not as straightforward as for complete profiles. Haney et al. (1995) proposed a sequential fitting method to estimate the

amplitudes in a progressive manner. This consists of obtaining the amplitude of the leading EOF by least-square fitting the D available data from the shallow cast to the first D components of the leading EOF. The second amplitude is obtained afterwards by fitting the residual of the first fit to the second EOF. Subsequent amplitudes are obtained in the same way. This is,

$$\hat{a}_{1,j} = \sum_{l=1}^{D_j} \boldsymbol{p}_{l,1} \boldsymbol{j}_{l,j} / \sum_{l=1}^{D_j} \boldsymbol{p}_{l,1}^2 \tag{2.12}$$

$$\hat{a}_{2,j} = \sum_{l=1}^{D_j} \boldsymbol{p}_{l,2} (\boldsymbol{j}_{l,j} - \boldsymbol{p}_{l,1} \hat{a}_{1,j}) / \sum_{l=1}^{D_j} \boldsymbol{p}_{l,2}^2 \tag{2.13}$$

According to Haney et al. (1995), using sequential fits ameliorates the fact that the EOFs are generally not orthogonal over the shallow depths D_j. Thus variability that can be explained equally well by several different modes due to their non-orthogonality is, in this way, always attributed to the lowest (leading) modes.

Once the amplitudes of the leading modes are computed, the shallow profiles can easily be extended down to the reference depth simply by multiplying the amplitudes by the full vertical extension of the EOFs.

Chapter 3 SPATIAL OBJECTIVE ANALYSIS

3.1 Introduction

Spatial objective analysis techniques were developed in operational meteorology by the second half of the twentieth century, as a fundamental step to initialize numerical weather forecasts. Prior to that, several attempts had been made to represent meteorological observations in a coherent fashion on charts or maps. These efforts had two motivations. First, it was hoped that the governing laws of atmospheric dynamics could be deduced from these charts. Second, it was felt that diagnosis charts of the past and present atmospheric states would help in the prognosis of futures states (Daley, 1991).

Despite spatial objective analysis techniques might seem to be little more than classical interpolation, they go well beyond that. Ideally, a spatial analysis algorithm should not only interpolate observations, but also filter out the noise inherent to the observations without appreciable alteration of the true field spectrum. It should also filter out the fraction of variance associated with scales which are too small to be properly resolved by the observing network and, if requested for initialization purposes, it should also eliminate fluctuations that do not satisfy the governing laws. Some or all these properties can be achieved by imposing the necessary constraints onto the interpolation. To put an example, the goal of statistical objective analysis is to minimise the ensemble average of the squared difference between the *"estimator"* and the "*true*" value of the observed field (or signal). The extent to which that difference (referred to as the analysis error variance) is minimized will determine the optimality of this kind of analysis schemes.

Before the use of computers, the estimation of observed fields was done "by hand" by specialized analysts. Because they were using their own subjective methods, the charts produced by two analysts could differ significantly. With the development of computers, automatic procedures to estimate the observed atmospheric variables on two or three dimensional grids from irregularly spaced observations were developed. These procedures,

referred to as ***objective analysis (OA)***, had to be robust enough to work without human intervention. In practice, the differences between the results of two different interpolation methods can sometimes be of the same order as the differences between subjective analysis produced by different forecasters (Daley, R., 1991). However, the advantage of the first over the latter is the underlying theoretical basis and the capability to reproduce exactly the same result when the same input data are used.

Most of the methods developed for the spatial objective analysis share a common formal representation. The reason is that most of the estimators are linear operators (matrices, in the discrete domain) acting on a set of input values (expressed as a vector ***X***) of a given variable and yielding a set of output values (vector ***Y***), usually of the same variable:

$$Y = H^T X$$

The elements of matrix ***H*** are usually referred to as "weights" (Thiebaux and Pedder, 1987). Worth noting features of this general representation are:

- The components of vector ***X*** could be observations from more than one field variable, the output values of **Y** then corresponding to one or all the input variables. In this case, the spatial objective analysis would be referred to as "multivariate", in front of the most common case in which all the values within **X** and **Y** correspond to the same variable ("univariate" analysis).

- Any number of analyses involving the same matrix ***H*** can be said to be realisations of the same analysis process. Therefore it is the elements of ***H*** which identify the specific procedure.

The first schemes proposed for computerized map analysis were anisotropic, inhomogeneous and quasi-statistical. Because of instability problems encountered when fitting polynomials over regions with sparse observations and also because of the computational constraints of the late 1950's, objective analysis procedures adopted for operational forecats had to seek for a compromise. The direction dependence and some location dependence of the early research analysis systems were sacrificed in order to produce a forecast in a reasonable time. (Thiebaux and Pedder, 1987).

Later on, when computational constraints were no longer a problem, more complex schemes were developed. All them involve the definition of some kind of model which serves as a basis for deriving a relationship between the observations and the estimated signal component of the analysed field variable. These models can be classified as belonging to one of the following types:

Empirical interpolation models. The signal estimate of the field variable at a given location is calculated as the weighted sum of observations located nearby. These weights depend fundamentally on the distance between the data location and the desired estimate location, and even though the data-weighting model is specified, there is no requirement regarding the differences between observed and estimated values in coincident locations. These models cannot, in principle, take into account the redundancy of information when two or more data points are adjacent to each other.

A statistical interpolation model. Even though the estimated value of a field variable is also a weighted sum of observations, statistical interpolation models explicitly use the ensemble spatial correlation structure of the whole field, relating values from data points to one another as well as to the analysed variable in the construction of a weighing scheme that minimises the analysis-error variance (Thiebaux & Pedder, 1987).

Following the pioneering work of Cressman in 1959, empirically formulated successive correction schemes have been widely used in meteorology for interpolating and smoothing spatial fields sampled by scattered observations. All them are based on the idea of correcting a prior estimate of the field at a particular location by adding to it a weighted mean of the differences between surrounding observations and their prior estimates, the weights applied to these increment data being calculated using a prescribed influence function of spatial separation.

A common approach to the design of empirical successive correction schemes is to consider the analysis system as a spatial filter acting on the observed field; analysis objectives are then expressed in terms of a theoretical wave-number response that is valid in the limit of continuous spatial sampling. Barnes (1964) was the first to apply this idea to the objective design of an empirical successive correction scheme. His work was later extended by

Maddox (1980), who developed a method of spatial filtering generally referred to as scale selection, based on a modified version of the Barnes analysis algorithm. Daley (1991) notes in his review that Barnes algorithm is one of the most successful SC algorithms and is still widely applied.

When a conventional SC approach is used as a basis of scale selection analysis, its net theoretical response is normally considered to predict the spectral relationship between the observed (input) field and the filtered analysis (output) field. However, in the case of inhomogeneous spatial sampling the actual (discrete) amplitude response of the analysis system cannot be considered as a spatially invariant property, and may differ substantially from its theoretical response. A similar problem arises due to the effect of domain boundaries, even when spatial sampling is relatively uniform (Pauley and Wu, 1990). Consequently the predicted theoretical response function is often an ambiguous and unreliable measure of scale selection properties. Pedder (1993) suggests an alternative approach to defining scale selection objectives, based on the idea of applying a continuous, spatially invariant filter to analysis rather than observed variables. The associated theoretical response is in this case also spatially invariant, and accurately describes the spectral relationship between the input field, which is a continuous approximation to the observed field, and an output analysis field. He also shows that it can be easily done when both the weights applied to the observed variables and the filter are derived from a simple Gaussian function of spatial separation.

The Successive Corrections algorithm used all along this present research is somehow a mixing of empirical and statistical methods. Despite it shares the simple formulation of empirical methods, the version developed by Bratseth (1986), with weights normalised in the "observation space", approaches Optimum Statistical Interpolation (the most widely used statistical method). Therefore, the scheme has the advantage of statistical methods regarding the consideration of the relative position of each data points relative to the others and not only to grid points. An additional normal-error filter convolution was also applied in the way proposed by Pedder (1993) in order to filter out the scales that could not be resolved by the sampling.

3.2 Formulation

The SC scheme is based on the following recursion formulas:

$$\mathbf{f}_g^k = \mathbf{f}_g^{k-1} + \mathbf{W}_g\left(\mathbf{f}_o - \mathbf{f}_o^{k-1}\right)$$

(3.1a)

$$\mathbf{f}_o^k = \mathbf{f}_o^{k-1} + \mathbf{W}_o\left(\mathbf{f}_o - \mathbf{f}_o^{k-1}\right)$$

(3.1b)

where $\mathbf{f}_g^k$ and $\mathbf{f}_o^k$ are the data vectors representing the analysis following *k* iterations evaluated on **g**rid (**g**-subscript) and **o**bservation (**o**-subscript) points, $\mathbf{f}_o$ is a vector of observations, $\mathbf{W}_g$ and $\mathbf{W}_o$ are matrices that contain weighting factors that interpolate increment variables onto the analysis and observation points respectively. The iteration cycle is initialised using

$$\mathbf{f}_g^0 = <\mathbf{f}_g>$$

(3.2)

$$\mathbf{f}_o^0 = <\mathbf{f}_o>$$

(3.3)

where $<\mathbf{f}_g>$ and $<\mathbf{f}_o>$ represent a "mean" field evaluated both on grid and observation points, which is estimated as a linear function of the observations. Instead of calculating a simple arithmetic mean, it is more appropriate to assume that it can be represented by a low order polynomial function of space, such as a linear trend surface estimated through ordinary least squares methods. This mean field analysis is assumed to account for variations associated with scales larger than the scale of the sampled domain.

The weighting factors $\mathbf{W}_g$ and $\mathbf{W}_o$ are not varied between iterations (they are in other SC methods, like in the one proposed by Maddox, 1980). They are generated assuming a gaussian correlation function model represented by

$$\Omega\left(x_{ij}\right) = \exp\left[-\frac{1}{2}\left(\frac{x_{ij}}{L}\right)^2\right]$$

(3.4)

where $x_{ij} = \left|x_i - x_j\right|$ is the distance between two points of the observed domain and *L* is a user-prescribed influence scale. *L* should be several times larger than the minimum resolved wavelength that can result from the data sampling.

Following the idea proposed by Bratseth in 1986, the weighting matrices $\mathbf{W}_g$ and $\mathbf{W}_o$ are normalized with respect to observation space, so that

$\mathbf{W}_g = \dot{\mathbf{U}}_g \mathbf{C}_o^{-1}$ and $\mathbf{W}_o = \dot{\mathbf{U}}_o \mathbf{C}_o^{-1}$, where **C** is an associated diagonal matrix of normalization constants

$$C_{kk} = \sum_{j=1}^{N} \Omega(x_{kj})$$

(3.5)

It can be demonstrated that in this way, the scheme (1) is equivalent to an analysis given by

$$\mathbf{f}_g^k = <\mathbf{f}_g> + \dot{\mathbf{U}}_g \mathbf{C}_o^{-1} \sum_{l=0}^{k-1} \left(\mathbf{I} - \dot{\mathbf{U}}_o \mathbf{C}_o^{-1}\right)^l \left(\mathbf{f}_o - <\mathbf{f}_o>\right)$$

(3.6)

Also the role of observational noise can be taken into account in a similar way to Optimum Interpolation by modifying the values of C_{kk} in a convenient way.

Another important practical consequence of using Bratseth's iteration scheme is that because the weighting factors in (1) are not varied between iterations, the analysis following k iterations can be represented by

$$\mathbf{f}_g^k = <\mathbf{f}_g> + \dot{\mathbf{U}}_g \boldsymbol{b}^k$$

(3.7)

where $\boldsymbol{b}^k$ is a parameter vector that depends only on observed variables, the number of iterations and the weighting matrix $\dot{\mathbf{U}}_o$, and is easily updated for each of the *k-th* iterations.

The theoretical one-dimensional amplitude response, considered as a function of spatial wavenumber, of a continuous linear filter that approaches the behaviour of the weighting process associated with the actual SC analysis system, is presented in Pedder (1993) (figure 1, not included in this text). He arrives to the following theoretical response function, which represents the spatial Fourier transform of the normal error influence function associated to the weights

$$R(\boldsymbol{m}) = \exp\left[\frac{-\boldsymbol{m}^2 L^2}{2}\right]$$

(3.8)

where $\boldsymbol{m} = 2\boldsymbol{p}/\boldsymbol{l}$, $\boldsymbol{l}$ being spatial wavelength. According to his results, as the iteration process proceeds, the response of SC is represented as

$$R^k(\boldsymbol{m}) = R(\boldsymbol{m}) \sum_{t=0}^{k-1} \left[1 - R(\boldsymbol{m})\right]^t$$

(3.9)

which in fact tends to an effective spatial scale separation process, damping structures smaller than the prescribed influence scale *L* (equation 3.4). A much deeper and complete explanation is found in his paper.

Pedder (1993) suggests an additional scale selection that does not involve repeating any iterations. Since the weighting factors are calculated from a continuous function of spatial lag, the analysis field can also be represented as a continuous function of location $\hat{f}(x_g)$.

Applying a continuous low-pass filter $g(x')$ to $\hat{f}(x_g)$, a new analysis field is defined by

$$\hat{f}^{1}(x_g) = \int_{-\infty}^{\infty} g(x')\hat{f}(x_g + x')dx' \tag{3.10}$$

If $g(x')$ is based on the normal error function similar to the SC analysis one, then it is defined by

$$g(x') = \frac{1}{(2\boldsymbol{p})^{1/2}L_g}\exp\left[-\frac{1}{2}\left(\frac{x'}{L_g}\right)^2\right] \tag{3.11}$$

where the parameter L_g is the filter scale parameter or cut-off wavelength. The scale selectivity of the filter is improved when applied in a recursive manner, but thorough a rather complex mathematical analysis, Pedder arrives to a filtered influence function solution which has the effect of a repeated filtering process.

Regarding the analysis parameters (characteristic analysis scale *L* that appears in equation 3.4, cut-off wavelength of the filtering on analysis results L_g), we have used the same values all along this work. The main reason is that they had already been tested in our study domain (Pinot et al., 1995). Moreover, checking the parameter values for the four alternatives of the EOF analysis (which will be detailed later on) and the four analysed variables of the four campaigns would increase the complexity of this research beyond affordable limits and might even obscure the interpretation of results.

For the above reasons, the parameters were fixed according to the following values. The prescribed influence scale for the analysis and for the

further spatial filtering, L and L_g, were set to 15 and 30 km according to correlation statistics (Pinot et al., 1995). All the data were treated as noise-free (noise to signal ratio equal to zero). Although this is obviously not the case, for CTD data the instrumental noise variance has been shown to be much lower than the variance of non-resolved scales, which are eliminated in a forthcoming step (Gomis et al, 2001). Different tests with values that ranged between 0.0 and 0.2 confirmed that nearly no differences were found on the resulting interpolated distributions.

In summary, each one of the analyses that were projected onto a regular 2.5 x 2.5 km were first treated with a first order detrending polinomial process (which eliminates any large scale signal), a constant length scale of 15 km for the weighting functions all along the 99 iterations. The further smoothing filter has a cut-off wavelength of 30 km. With this procedure we generate independent two dimensional grids. Their superposition with a vertical resolution of 10 m enables us to produce quasi-three dimensional grids of the hydrodynamic variables.

Chapter 4 GENERAL   OCEANOGRAPHIC   ASPECTS   OF   THE EBRO DELTA REGION

Antes de la peluca y la casaca
fueron los ríos, ríos arteriales:
fueron las cordilleras, en cuya onda raída
el cóndor o la nieve parecían inmóviles:
fue la humedad y la espesura, el trueno
sin nombre todavía, las pampas planetarias.

El hombre tierra fue, vasija, párpado
del barro trémulo, forma de la arcilla,
fue cántaro caribe, piedra chibcha,
copa imperial o sílice araucana.

    Tierno y sangriento fue, pero en la empuñadura
    de su arma de cristal humedecido,
    las iniciales de la tierra estaban
    escritas.

      Nadie pudo
      recordarlas después: el viento
      las olvidó, el idioma del agua
      fue enterrado, las claves se perdieron
      o se inundaron de silencio y sangre.

        No se perdió la vida, hermanos pastorales.
        Pero como una rosa salvaje
        cayó una gota roja en la espesura
        y se apagó una lámpara de tierra.

          Pablo Neruda
   (fragmento de "I - La Lámpara en la Tierra", del Canto General)

The Ebro Delta region is located in the western Mediterranean, within the Balearic Sea. In general terms there are four main basins within the western Mediterranean (Figure 4-1): from north to south we find the Liguro-Provenzal Basin, the Balearic Basin where we will focus our interest, the Algerian Basin and the Alboran Basin at the south-western extreme. The Liguro-Provenzal Basin comprises the Gulf of Lions, with a steep slope and a series of submarine canyons. The main circulation feature of this region is the Northern Current, which flows from the Gulf of Genova towards the southwest.-, It also receives the influence of the Rhone River discharge, which provides the main freshwater and particulate matter supply to the Gulf of Lions (the highest daily mean freshwater discharges reported by Durrieu de Madron et al. (1999) range between 3000 and 5500 m^3/s, while the lowest values are around 1000 m^3/s). The most remarkable aspect of the local oceanography is the deep and intermediate water formation by convection during winter, which gives origin to the Western Mediterranean Deep Water (WMDW) - which fills the deepest regions of the Basin- and to the Winter Intermediate Water (WIW).

South of the Gulf of Lions is the Balearic Basin. It covers a region limited to the west by the Spanish Peninsula, to the east by the Balearic Islands (Mallorca, Menorca, Ibiza and Formentera) and finishes to the south with the Gulf of Valencia at Cape La Nao. On the peninsular side the topography is characterized by a narrow shelf (less than 25 km south the Gulf of Lion) and submarine canyons. The shelf widens at Cape Salou to around 70 km and remains so down to Columbretes Islands. This region is known as the Ebro Delta Shelf and is our main study area. Between Cape La Nao and Ibiza is the Ibiza Sill ($\approx$ 800 m) between Ibiza and Mallorca is the Mallorca Sill ($\approx$ 600 m) and to the north is the Menorca Sill ($\approx$ 100 m). The island shelves are narrow and the slope is also abrupt. All over the Balearic Sea, the complex topography seems to play an important role in the regional circulation. (García Ladona et al., 1994; López-García et al., 1994; Millot, 1999; Pinot et al., 1995a).
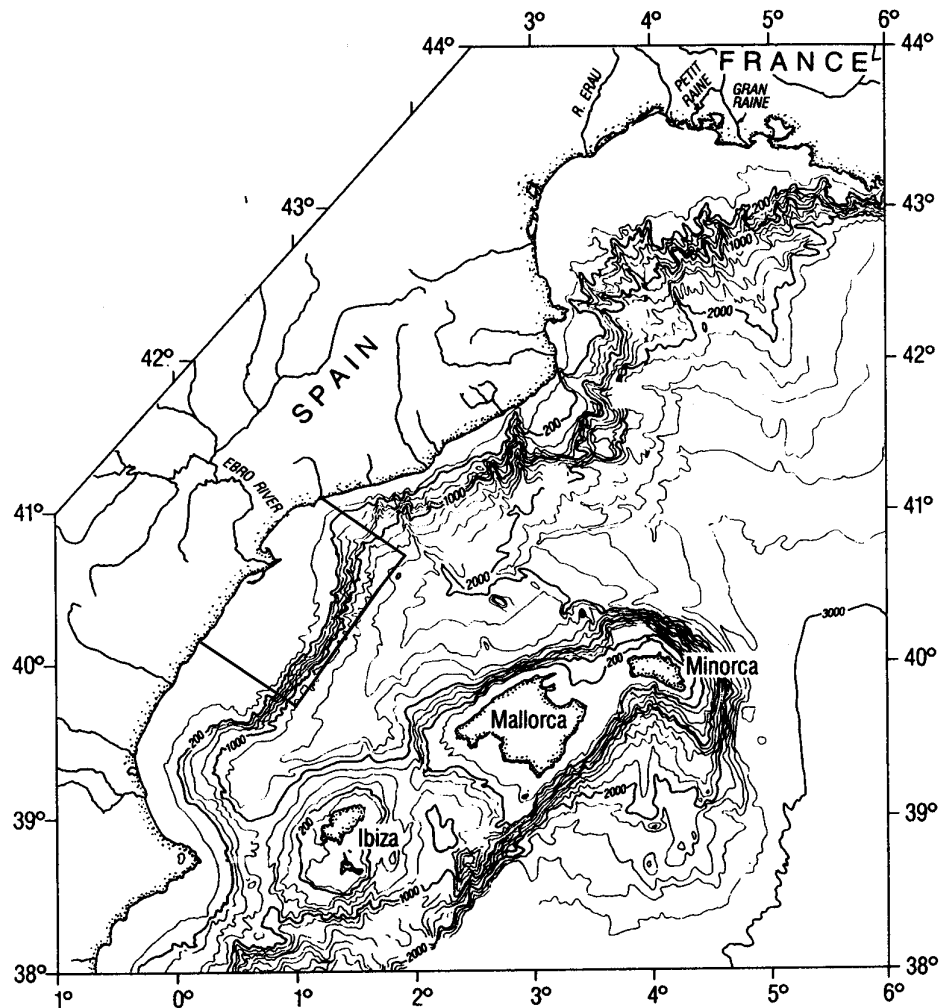
Figure 4-1 A general view of the Balearic Sea. The FANS and MEGO 94 domains are marked by the rectangle around the Ebro Delta.

Four major water masses are encountered within the western Mediterranean Sea (Millot, 1999). The term **Modified Atlantic Water** (**MAW**) is used to refer to the surface water all over the Mediterranean Sea. It originates from Atlantic waters that enter the Mediterranean through the Gibraltar Strait and become transformed through evaporation and mixing as they proceed through the Alboran and the Algerian basins. Except in some places, it forms a surface layer of around 100 - 200 m. The flow of MAW can be deflected towards the north to the Balearic Basin in the form of "recent" MAW, or continue eastward along the African slope as far as the coast of Sardinia, where part of the flow might deviate to the north and form the "old" MAW, and part continue to

the eastern Mediterranean. Its mean temperature below the mixed layer is approximately between 14-15 °C, and its salinity changes from 36.5 at Gibraltar to around 38-38.3 (Millot, 1999).

During winter, cold and continental winds, as well as air-sea temperature differences produce an important heat loss from the sea. This process, together with the mixing of the surface layer mostly through wind steering, lead to a relative homogenisation of the water column on the continental shelves of the Gulf of Lions and the Balearic Sea, so that MAW can be cooled without any intense mixing with the waters below. This leads to the formation of a water mass, named **Winter Intermediate Water (WIW)** (and also to the production of WMDW, see next paragraph) with temperatures ranging from 12.5 - 13.0 °C and salinity 38.1 - 38.3. The WIW is found at the base of the MAW layer and is characterized by a T minimum (Salat et al., (1978), Millot, (1999)). The coldest WIW is found in the north, from where it flows, as the MAW, along the Spanish continental slope and across the Algerian Basin. Its path can also be influenced by frontal eddies associated with the North Balearic Front. In any event, it appears that a relatively large amount of MAW can be transformed into WIW in cold winters. A similar process occurs in the Eastern Mediterranean, where saltier "old" MAW is transformed into **Levantine Intermediate Water (LIW)**, occupying the basin down to depths of around 800 m. This water mass is characterized by its relative temperature and salinity maxima. The LIW flows across the Sicilian Channel into the Western Mediterranean, where the LIW maximum temperatures range between 13.2ºC and 14 °C and the maximum salinities are of the order of 38.5 to 38.7. The LIW T and S max are generally found at 400 - 700 m.

Finally, the **Western Mediterranean Deep Water** (**WMDW**) fills the deepest regions of the Basin. Its typical temperature and salinity ranges are 12.75ºC - 12.80ºC and 38.44 - 38.46 respectively (Lopez García et al., 1994), (Millot, 1999). In a process similar to the formation of WIW in the Gulf of Lions, the vertical mixing process produces plumes of few hundred meters in diameter and average vertical velocities of around 10cm/s, that develop in a convection region with average sinking velocities of approximately 1 mm/s. If the process is

intense enough, WMDW is formed; otherwise convection stops at intermediate depths as WIW.

The transformation of MAW into WIW is an important stage of the transformation of Atlantic water into Mediterranean waters, as it produces a water denser than MAW which may eventually penetrate into the eastern Mediterranean Sea and can be more easily mixed with LIW in the western Mediterranean Millot, 1999). The various transformation and mixing processes concerning these intermediate waters occur on a wide range of space and time scales, so that adequate numerical models are needed to help quantifying the phenomena. It is important to mention that the hydrologic characteristics of the water masses are much more variable than thought in the past and are changing at decadal scales (Millot, 1999).

The thermohaline flows are the major forcing in the Balearic Basin, since winds have been shown to produce only transient perturbations (Font, 1990; Pinot et al., 1995a). There are two permanent thermohaline fronts in the Balearic Basin, one associated to the Northern Current flowing towards the south, whose signal can be traced as far as the Gulf of Valencia, and the other associated to the Balearic Current towards the northeast. These fronts are in quasi-geostrophic equilibrium in a density distribution characterized by lighter waters either from continental origin (Northern Current) or from recent MAW (Balearic Current), and the denser oceanic waters, consisting in its upper layer mostly by old Modified Atlantic Water (MAW).

The Northern Current in the Ebro Delta shelf/slope is a quasi-permanent feature that contours the continental slope with a transport of 1-1.5 Sv (Font, 1990, among others). Its signal on the surface usually lies above the 1000 m isobath and its nucleus normally stays in the upper 300 m. The flow velocities at the NC core are on average 25 - 30 cm/s from spring to autumn. In winter the jet velocity is enhanced as an effect of increased river discharge due to rainfall, as it increases the density gradients (mostly in the Rhone and the Ebro mouth areas). The thermal signature of the front can be masked in summer due to stratification, but the salinity gradients still reveal its presence. With the aid of satellite images, the presence of abundant and energetic mesoscale structures, such as eddies and filaments, has been observed (see e.g. La Violette et al.,

(1990); Arnau, (2000)), and a seasonality in their intensity has also been detected through current meter data spanning several years, with highest activity during autumn, a fast decline during winter and a continuous decrease until the end of summer, when the cycle begins again. (Font et al., 1995). The similarity with the overall dynamics in the Corsican Channel points out the possibility that the temporal variability of the Northern Current in the Balearic Basin reflects a general variability that affects the whole NW Mediterranean (Millot 1999). Mesoscale flow events lasting several days present a rather barotropic behaviour down to depths of at least 100 m. A possible explanation of its signal below the thermocline is energy transfer via internal waves at near inertial frequencies. In this part of the Mediterranean the inertial oscillations are as intense as the mean flow, inertial frequency signals spreading to near-inertial bands seem to evidence the interaction with a mean current, which causes the frequency shift through its vorticity field. In the lower layers, enhancing of inertial motions might be due to long-period waves generated at the coast and propagated offshore after a wind event (Font et al., 1995).

Durrieu de Madron et al. (1999) analysed one year of data on the Gulf of Lions continental margin and found that the mesoscale activity associated to the Northern Current presented a maximum in early spring. This discrepancy with the maxima values in autumn reported by Font et al., (1995) could well point out to the interannual variability of the whole system.

Another recurrent feature observed in IR images is the presence of warm water over the shelf of the Gulf of Valencia during the winter months, which seems to be recent MAW imported through the Ibiza Channel. It has been hypothesised (Lopez García et al., 1994) that the Algerian Current can become baroclinically unstable and develop large meanders and eddies which move eastward (and ultimately northward) for months at few km per day (Millot 1999). Numerical and physical models support the role that mesoscale eddies can have on the Current, deviating it and thus allowing recent MAW to reach directly the Balearic Channels Millot 1999). The eddies themselves may detach from the current and convey large amounts of recent MAW into the Balearic Basin.

Recent MAW seems to be fundamental in the dynamics of the Balearic Front because there is no major river (and therefore no input of light fresh

water) outflow from the Islands, and therefore the density gradient that sustains the Balearic front/current system must be given by the recent MAW entering the Balearic Channels from the south and the old MAW coming from the north. There is also evidence that at least part of the flow from the Northern Current does not cross the Ibiza Channel on its way to the south (Salat, 1995; Pinot et al., 1995b). A detached branch of continental influenced water deviating to the east and joining the Balearic Front has been often detected in the Gulf of Valencia. Some of this water might be trapped locally in an anticyclonic eddy, but there is not enough data coverage to assess this (Pinot, et al., 1995b). Mesoscale variability and seasonal fluctuations seem to be important in the Balearic Front, and there is evidence in its surface temperature signal of important changes in its position during winter. These fluctuations can not be tracked during the stratified season due to the low surface temperature gradients in the area. Like the front associated to the Northern Current, the Balearic Front has important mesoscale variability, and the intrusion of recent MAW from the south adds complexity to its dynamics (Lopez García et al., 1994) .

Bathymetry seems to play an important role in the dynamics of the region (Pinot et al., 1995a; Millot 1999). There seems to be a clear topographic steering of the surface layer currents in the western Mediterranean, being the continental slope the main guideline. The location and structure of the main front/current systems (such as the Algerian, Corsican and Ligurian currents, which follow slope bathymetric contours) seem to support this. Even though intrinsic instability of the jets must certainly play a role in meandering, local changes in the bathymetry seem to be one of the major mechanisms capable of generating significant frontal mesoscale recirculation and branching of the jets. An important feature of this kind has been detected off the northern Ebro margin, where an energetic, deep cyclonic eddy pulls slope waters offshore. This observation is reminiscent of large cool offshore plumes found recurrently at the same place by La Violette et al (1990) from an analysis of IR SST images. These plumes might be the surface signature of a systematic topography-induced branching of the Northern Current, with an estimated transport of about 0.1 Sv during the FE-91 mesoscale experiment in early

summer. Other mesoscale structures induced by topography seem to be important in the exchange of biogeochemical tracers, and also in the transport of fish larvae being forced offshore by currents induced by submarine canyons, where the flow pattern changes are due to vorticity adjustment (Pinot, et al., 1995b).

With the aid of a barotropic ocean model in which the topography simulated the Ebro Delta widening of the shelf, Gjevik, et al. (2002) studied the effects of friction and topography on the evolution of a jet (Northern Current) as it proceeds from a theoretical straight narrow shelf to a wide one. They find a tendency for topographic steering of the flow along the steeper part of the shelf/slope in the transition zone in all their simulations, developing over a time scale of 2 – 5 days. This shows that the adjustment to topographic steering takes place on a shorter time scale set by propagating wave modes. Through their numerical simulations, they also find that bottom friction reduces the flow level on the shelf and prevents the formation of eddies, which makes topographic steering a more prominent feature of the simulated current fields.

Winds in the area do not have a significant effect on the large scale circulation because there is not a wind pattern large and long enough to establish a current (Font, 1990; Font et al., 1995). Nonetheless wind events are an important forcing mechanisms on the upper layer circulation, both through the generation of inertial oscillations and interactions with local topography. Also, wind events are fundamental both because of their mixing power and their evaporative cooling, together with radiative cooling during winter, in the water formation processes in the Gulf of Lion.

The analysis of two years of wind data (April 1988 - March 1990) taken from three stations in the Ebro Delta region revealed the predominance of levante events, which are easterly and north-easterly winds blowing onshore and lasting more than 24 hours (Espino et al., 1998). Using a quasi-3D numerical model to solve the shallow water equations and applying it to the Gulf of Sant Jordi, Espino and colleagues (1998) concluded that the local circulation tends to be anticyclonic due to the topography and the forcing that the Northern Current exerts with its continuous southward flow, but the dominant wind events

(70% with an associated negative vorticity) tend to enhance this local circulation pattern.

Nonetheless, our study region comprises a shelf that widens from less than 25 km to around 70 km, and both the wind and frictional effects might well play a significant role on the local aspects of its circulation. A question addressed by the FANS project was to estimate the importance of the Northern Current as the dominant driving agent of the circulation, versus the influence of more local forcing mechanisms such as the wind, bottom friction and the buoyancy input through the Ebro River outflow. García et al. (personal communication) compared the low frequency variability of currents on the southern Ebro shelf/slope region with wind data measured in a nearby location. The data comprises two periods: from March to May, with weak stratification, and from July to October, with typical summer stratified conditions. They concluded that the mean shelf circulation south of the Ebro Delta is controlled both by the local wind and by the Northern Current, depending on the intensity of these forcing agents and on the mesoscale activity of the current. As mentioned before, the Northern Current jet is strongest during winter due to the increased density gradient that results from the largest fresh water input of the rainy season, and this is also the season when most of the mesoscale activity occurs. That is why they found that during winter-spring conditions the Northern Current plays a dominant role on the shelf circulation, with transient intensification and weakening periods associated with across shelf excursions of the current due to its mesoscale activity. They also find from the data analysis that a 15 cm/s velocity of the slope current is a threshold value for the Northern Current hydrodynamic control. On the other hand, during the summer conditions, the Northern Current signal is weaker and the mesoscale activity is also low, therefore the wind becomes the most important driving agent for the shelf circulation above the pycnocline.

Another possible effect of the wind on the shelf circulation was modelled by Gjevik et al. (2002) with the numerical model mentioned before. Events of northwesterly winds are not uncommon during winter time, and since the orography of the Ebro valley produces a channeling effect they performed a

numerical simulation to study the response on the current of a nortwesterly wind jet, with highest velocities on the transition zone where the shelf widens. A typical feature of the current response was the formation of a dipole eddy structure with its axis tilted at an angle of around 30 degrees relative to the wind jet. A strong current shear appeared between both gyres. The anticyclonic one, over the wider shelf, had a larger radius. The effect of the topography became evident with simulations in which the jet was well over the wide shelf, in this case both gyres had similar radii. They also mention the effect of bottom friction, which tends to decrease the velocities, particularly in the strong shear region between both gyres. Without bottom friction this shear zone eventually becomes unstable and a system with three eddies appears, two cyclonic and one anticyclonic. An almost symmetric eddy pair then propagates downstream on the wide shelf.

The Ebro fresh water discharge (positive buoyancy flux) interacts not only with the nearshore, but also with the shelf and slope regimes by means of a coastal plume spreading offshore. The dynamics of the plume regime, termed ROFI (**R**egion **O**f **F**resh water **I**nfluence) (Simpson et al., 1991) depends mainly on the balance between the buoyancy induced momentum flux spreading offshore (which itself depends on the amount of fresh water available) and the processes that lead to lower gradients and shears through mixing. While, as mentioned in previous paragraphs, the wind does not play a predominant role in the overall circulation of the Western Mediterranean, it may modify significantly the shelf/slope and nearshore circulation through wave breaking, momentum flux and upwelling or downwelling. The wind may also change the shape, extent and velocity of the plume (Maidana, et al., (2002)).

Chapter 5 THE DATA

5.1 Fans Campaigns

Within the frame of the FANS Project, three oceanographic cruises were carried out on board R/V García del Cid to obtain CTD data from the shelf and upper slope region, as far as the 1000 m isobath, over an area of around 7500 km^2. CTD casts were performed along transects perpendicular to the coast, with average spacing between stations of 3 to 6 km over the shelf, and even closer on the slope (Figure 5-1). The distance between transects was around 8 km. CTD casts were obtained with a GO MkIII C probe, equipped with a 12-bottle Rossette sampler. CTD data were recorded from 2 to 5 m below the surface down to around 5 m above the bottom under good weather conditions and when the bottom did not present abrupt changes. If the above conditions were not met, a much safer distance from the bottom was the lower limit of the measurements.

Water samples were taken at standard levels with Niskin bottles in alternate stations, and salinity measurements were carried out with a Guildline Portasal after the campaigns.

A hull-mounted 150 kHz NarrowBand ADCP was operational, in principle, during the FANS II and FANS III campaigns. Unfortunately, the ADCP data was not useful for the combined analyses we wanted to perform. The FANS III data was rather unreliable due to problems with the gyrocompass connections, and the raw FANS II data turned out to be extremely noisy, therefore useless for a precise estimate of ageostrophic currents through a multivariate CTD-ADCP data analysis. Since we were interested in the estimate of ageostrophic currents in the Ebro shelf/slope region, we tried to include in this work ADCP data from a previous campaign in the area, MEGO 94, which were reported to be reliable. But once again, the raw data turned out to be noisy, as with the FANS II data, and the problem was even more critical due to the fact that the recorded velocity values were low, which means that the imprecision added by the ship's

motion at stations could be very significant. As a result of that, no ADCP data analysis is included in this work.
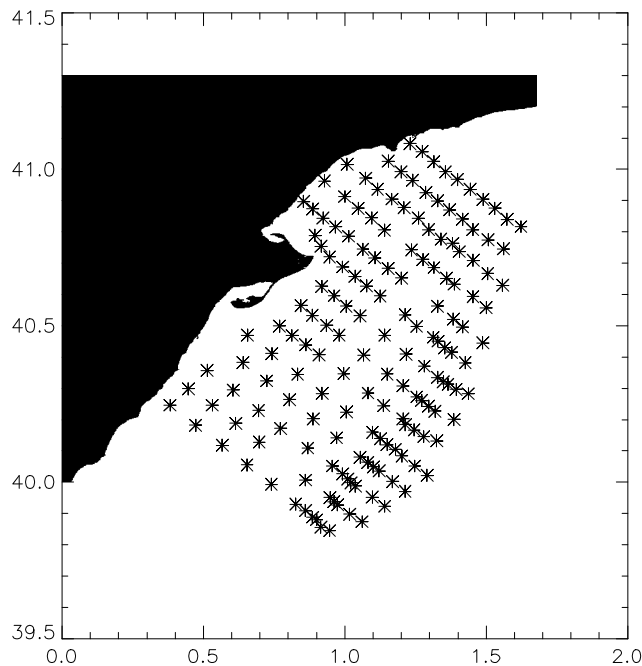
Figure 5-1 General stations distribution during the FANS project. Transects were designed perpendicular to the coast and stations are closer at the slope. This figure corresponds in particular to the FANS II position of CTD casts whose data was useful for the EOF analysis.

5.2 *Mego 94*

The MEGO 94 campaign was part of the project "*Dinámica de las Fluctuaciones de Mesoescala en la Plataforma Continental Catalano-Balear (AMB92-0251-C02-02)*", whose goals were to study the physical mechanisms that control the fluctuations in the circulation on the Catalan Continental Shelf. The MEGO campaign focused mainly in the Gulf of Sant Jordi, with a dense CTD grid all around the Ebro Delta region.

The CTD stations are much closer than during the FANS project around the Ebro Delta area, while the north and south transects have different orientation (Figure 5-2). In the figure we show the CTD stations that were actually used in the EOF analyses.

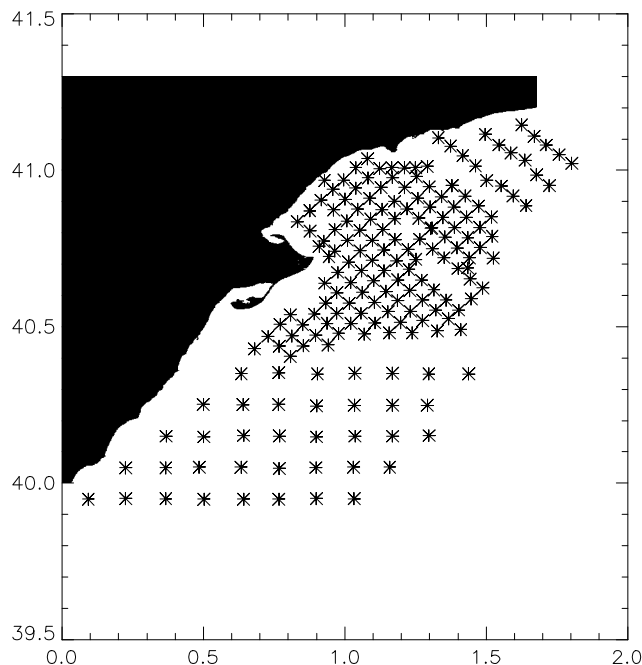**Figure 5-2 General stations distribution for the MEGO 94 project. This figure shows the location of CTD casts whose data quality made them useful for the EOF analysis.**

5.3 Ctd Data Quality Control

The CTD data quality control was performed in different steps: through CTD software, self developed software and further manual check and edition.

The CTD software, provided with the instrument, includes different processes: a) User-controlled spike detection in which a maximum difference between two successive observations is set; if that threshold is exceeded the corresponding data is flagged for the next step. b) User-controlled smoothing, in which the flagged data can be substituted by an averaged value based on neighbour observations. c) User-controlled vertical average is used to produce values at defined depth intervals. For all the campaigns we obtained an output at 1 m intervals.

Additional software routines developed by the author and manual checks were used for the following verifications:

a) Check for header positioning (stations header values should be according to campaign plan). During this process we took care of stations on land or outside the study area, correcting their positions.

b) Check for unreliable large property gradients.

c) Check for stability in the water column through density inversions.

d) CTD salinity values were compared with in-situ bottle measurements.

In the particular case of the FANS II campaign, some casts turned out with salinity profiles that seemed to "jump" in a step-wise manner towards higher values. Since no in-situ water salinity measurements were used for an specific calibration of these casts, the salinity values of the abnormal saltier casts were shifted to a line in θ - S space which was fitted to the casts where LIW and WMDW were observed. Some of the shallower casts which did not reach the required depth and whose salinity values were dubious were eliminated. Density was then recalculated with the corrected salinity values.

The final number of available CTD casts, after all the data quality control procedures, for each of the three FANS campaigns and for MEGO 94   is presented in Table 5-1.

Campaign	Date	No. CTD casts	No. CTD Deep casts
MEGO 94	4/Mar/94 -11/Mar/94	187	25
FANS I	1/Nov/96 – 11/Nov/96	133	25
FANS II	4/Feb/97 – 14/Feb/97	169	42
FANS III	7/Jul/97 - 16/Jul/97	148	28

Table 5-1 Date and number of casts with useful data for the EOF analyses. The profiles (eigenvectors) are estimated from the deep casts only.

5.4 Historic Data

The historic data used for this project was taken from the MEDATLAS data base. Information on the data quality control and further analyses can be found in the following web page:

http://www.ifremer.fr/sismer/program/medatlas/gb/gb_medat/htm

A key goal of the MEDATLAS project was to collect in a database format reliable hydrological data over the Mediterranean basin. Considering CTD data alone, there are more than 16,000 casts with good quality data over the whole Mediterranean Sea.

For this work, we used the basic statistics, eigenvalues and eigenvectors from 188 CTD casts for summer (April to October) historic data analyses and 85 casts for winter (November to March) ones. The basic statistics, together with the eigenvalues and eigenvectors, were kindly provided by Ananda Pascual and Dr. Damià Gòmis, from the Balearic Islands University in Palma de Mallorca, Spain.

In Table 5-2 we present the number of CTD casts included in the two seasons, while in Figure 5-3 and Figure 5-4 the actual geographic location is shown. It is important to notice that there are a significant number of casts outside our particular study domain, specially during winter, where an important number of deep casts is actually at the north, in a region where the bottom topography displays several submarine canyons and where winter water masses are believed to form and to sink.

Historic Data	Date	No. CTD casts	No. CTD Deep casts
Summer	April - October	658	188
Winter	November - March	308	85

Table 5-2 Seasonal historic data range, and number of CTD casts, for Summer and Winter from the MEDATLAS data base.
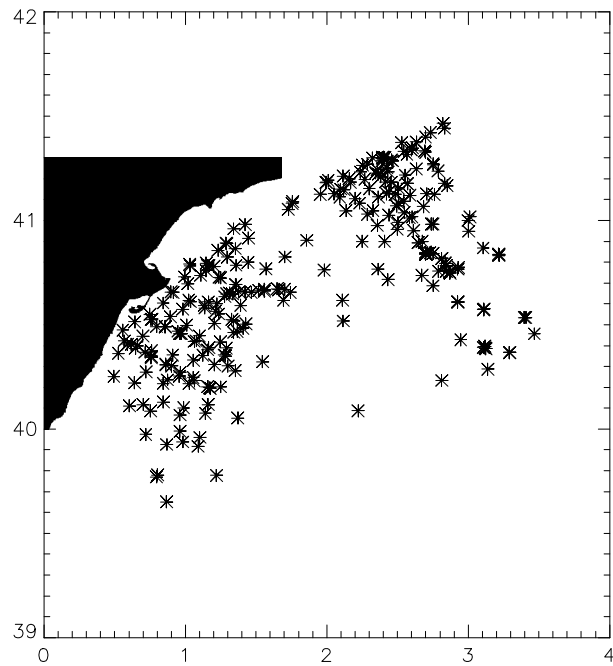
Figure 5-3 Location of all the historic summer stations. The EOF analyses with this historic data were performed with CTD deep casts.
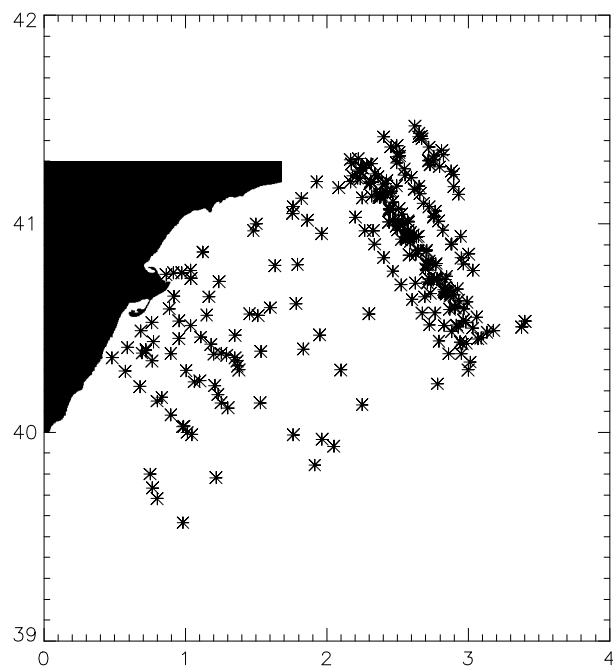

Figure 5-4 Location of all the historic winter stations. As in the previous case, EOF analyses were carried out with the deep casts only.