Forecast verification of a 3D model of the Mediterranean Sea. Analysis of model results and observations using wavelets and Empirical Orthogonal Functions.

Aida Alvera Azcárate

Departement of Astrophysics, Geophysics and Oceanography

University of Liège

A thesis submitted for the degree of

Docteur en Sciences

Committee members

Prof. Jean-Paul Donnay, president (University of Liège, Belgium)
Prof. Jean-Marie Beckers (University of Liège, Belgium)
Prof. Robert H. Weisberg (University of South Florida)
Dr. Pierre-Yves Le Traon (CLS Space Oceanography Division, France)
Dr. Damià Gomis (IMEDEA, CSIC, Spain)
Dr. Marilaure Grégoire (University of Liège, Belgium)
Dr. Michel Rixen (NATO/SACLANT Undersea Research Centre, Italy)

Acknowledgements

There are many persons I would like to thank. Without them, this work would not have been possible.

First, I would like to specially thank the supervisor of my thesis, Professor Jean-Marie Beckers. His deep knowledge on ocean modelling made all my questions an easy task to be solved. His advise has been always of undeniable value for me, and without his help and support this work would not have been possible.

Special thanks also to Dr. Michel Rixen, who encourage me all along these three years of work. His enthusiastic comments and suggestions have always been of great help.

I would like to express my gratitude to the other members of my commitee, Prof. Jean-Paul Donnay, Prof. Robert H. Weisberg, Dr. Pierre-Yves Le Traon, Dr. Marilaure Grégoire and Dr. Damià Gomis for their disponibility to review my thesis and for their efforts to find a convenient date for the defence!

I wish to thank also my colleagues, Zied Ben Bouallègue, Luc Vandenbulcke, Caroline Raick and Abderrahman Elkafazi. They have contributed to create a friendly atmosphere, whith fruitful discussions that have enriched my work.

Many thanks also to those who have helped me with the correction of the manuscript, Stéphenie Lange, Aline Gral, Dr. Miltos Markatos and Bénédicte Delsemme.

I do not know how to thank Alexander Barth. He has made enormous efforts to help me every time I needed and his encouraging words have always come in the right moment. This work would not have been possible without him.

I keep my warmest words to my family, for their loving support and encouragement. They have been always willing to help me, and they have encouraged me since the very first time I decided to study oceanography, even if that meant hundreds of kilometers between us.

The National Fund of Scientific Research (FNRS, Belgium) is acknowledged for the financing of a supercomputer. The work was realised in framework of the Satellite-based Ocean ForecasTing project (EVK3-CT-2000-00028) of the European Union.

Abstract

The quality assessment of the three-dimensional GHER (GeoHydrodynamics and Environmental Research) model of the Mediterranean Sea is presented in this work. The verification of the model results is done in a spatio-temporal approach. Traditional error measures (*i.e.* correlation, mean error, etc) are very useful to assess the quality of a model, but they do not take into account the high complexity of three-dimensional models. The verification process is thus done in three main parts: first, the model is compared to observations and climatology in a qualitative approach, in order to make a preliminar study about the model behaviour. Then, the error assessment is done, using traditional statistic measures. In order to take into account the complexity of the model and observations, the last step in the verification process consists in a spatio-temporal analysis using wavelets and empirical orthogonal functions. This last analysis will allow us to have an insight about the model quality in a more detailed way.

This verification process has been applied to the GHER model. This model is implemented in a two-way nesting approach in the Mediterranean Sea, Liguro-Provenąl basin and Ligurian Sea, where the highest resolution is reached. Assimilation of sea surface temperature and sea level anomaly is made during a nine-week experiment. Another test is carried out, to assess the quality of sea surface temperature from the SOFT predictor of the Ligurian Sea. The predicted sea surface temperature is assimilated in the model and the quality of the forecast is compared to the first assimilation experiment. The assimilation of the SOFT statistical predictors can be very useful to force models in a real forecast experiment, where no observations are available.

Contents

Chapter 1

Introduction

1.1 The verification process

Forecast verification is an important of the modelling process, because the verification of a model allows us to better understand its behaviour, to assess its weaknesses and strengths, and to establish the causes or sources of the error. However, model forecast verification is often an underestimated aspect of the modelling process. For example, when assessing the error of a hydrodynamic model, it is very tempting to draw an error curve of a given variable, as the root mean square error of the surface temperature, and establish the quality of a model based on this curve. However, this curve usually summarises two complex fields, three-dimensional model fields of different variables and the corresponding observations fields into one single value. This two fields evolve in time and space, at different spatial and temporal scales. The complexity of these two fields cannot be summarised into a single error measure, because they can be different in many aspects(*e.g.* different variability, missing processes, intensity of a variable, etc), and because there is no single error measure capable of capturing all this information (Murphy, 1996). Moreover, unobserved variables affect the behaviour of the model and the quality of the model in representing the observed variables. The causes of the model error can be multiple. A good choice of the verification tools is thus necessary to make a complete study of the model. In addition, these verification measures must be interpreted carefully to extract as much information as possible from the model. Usually, a series of different scores measuring different aspects of the model are combined to assess its quality (Wilks, 1995; Murphy, 1996; Jolliffe and Stephenson, 2003). Murphy and Winkler (1987) claimed for the need of a general framework for forecast verification. The use of an homogeneous verification process would in fact make the comparison between different models easier. In this work we will develop a verification process that takes the complexity of the model and observations into account. The first step is to identify the main sources of error that can arise in our model.

The verification of a three-dimensional, two-way nested model with assimilation of sea surface temperature and sea level anomaly will be made in this work. There are thus

at least five possible causes of error:

1. The boundary and initial conditions, the atmospheric forcings
2. The transition between nested domains
3. The assimilated observations or the assimilation scheme
4. The bathymetry
5. The model parameterisation and numerical schemes

Of course, these reasons have been enumerated in a broad way. The verification process must determine the relative contribution of each of these main reasons to the model error, and further study them to concretise it as much as possible.

The preparation of the appropriate verification strategy must be done carefully. The definition of verification, although it seems to be clear (*i.e.* to assess a model's quality), is not that trivial, because the forecast verification process should vary depending on a wide variety of factors. Among others:

- The type of forecast we are dealing with (probabilistic, binary, continuous field forecast...)
- What we want to verify (*e.g.*, the overall quality of a model, the improvement compared to a previous version of the model, a specific process located in time and/or space, etc).
- Which data (observations) can be used for the verification (are they available for the same period than the forecast? quality of the data, what scales are resolved by the data, etc.)
- Which relationship exists between the data and the model (are they independent from the model forecast, or have they been used to force the model?)
- What the final use of the verification results is (is there a final user, is it an operational product, who and how will use the model forecast...)

The design of a forecast verification programme depends thus on what we want to verify and at which spatio-temporal scales. It depends also on how we define the goodness of a model. Murphy (1993) defined three types of goodness, mainly based on (i) the consistency, or how the model matches the modeller expectations, (ii) the quality, or how the model matches the observations, and (iii) the value, or what is the benefit that is obtained from the model. In this work we concentrate in the second point, the quality of

the model, as the first and third ones are more subjective and can change depending on the situation (Roebber and Bosart, 1996). The verification process must be designed to find the answers to the questions that may arise when dealing with forecast results. In this work we focus on the verification of a three-dimensional model of the Mediterranean Sea. The verification of such models is often called field forecast verification, gridded, or spatial verification (Hoffman *et al.*, 1995; Wilks, 1995; Ebert and McBride, 2000; Jolliffe and Stephenson, 2003). In field forecast verification, some of these questions are:

- Does the model accurately represent the observed magnitudes?
- Is the model better than other prediction systems? (climatology, persistence...)
- Is the model variability similar to what is observed?
- Are the model variables systematically over or underestimated?
- Are there specific missing processes in the model? (such as a recurrent gyre or a front)
- If there is assimilation of observations into the model: is the skill of the model improved? how does the assimilation of one variable affect the other variables?

Answering some of these questions cannot be done with a unique error measure. To answer questions related to the spatial distribution of a variable, the temporal average of the model can be computed. However, in field forecast verification special attention must be paid to the fact that the spatial correlation between grid points makes the study of skill distribution in space very difficult (Livezey and Chen, 1983). It is very unlikely that two adjacent points in a model grid are completely independent, thus the interpretation of a spatial skill becomes ambiguous (Wilks, 1995; Briggs and Levine, 1997; Jolliffe and Stephenson, 2003).

Time evolution of the error is thus preferred to avoid the correlation problem. An average over all points in a grid at a given time is thus made for the verification. The temporal evolution of the error is very useful, as it allows to obtain a general idea about the quality of the model. It is often the only error measure applied. However, to answer some of the questions specified above, one needs to keep the spatial distribution of the studied variable. Spatio-temporal techniques can help us to study the evolution of a variable in time, keeping the information of the spatial distribution, and avoiding the problem of correlation between adjacent points. Multi-scale techniques allow us to study the behaviour of a model at different spatial scales, or even to focus on a specific scale of interest. Nested models, as the one we are working with, can also be considered as a spatio-temporal approach. Each nested level is a refinement of the parent model, so the verification at those two model grids can also give us an idea about the model's behaviour at different spatial and/or temporal scales.

There are many ways to study the spatio-temporal characteristics of a model, due to the wide variety of processes at different spatial and temporal scales. For example, error decomposition methods. Hoffman *et al.* (1995) studied the error on a 500-hPa geopotential height field. The error was decomposed into three main components: displacement error, amplitude error and residual error. In this way, the contribution of each of these error types to the total error was determined. The displacement error accounts for the error that can be explained by the misplacement of the model results compared to observations. The amplitude error accounts for the error explained by a change in amplitude on the displaced model. The same approach was used by Ebert and McBride (2000) for rain events, although in an object-oriented approach. Ebert and McBride (2000) defined "contiguous rain areas" as the area of contiguous observed and/or forecast rainfall enclosed within a specified rain threshold, and applied the decomposition of the error to these entities. They decomposed the error on location, volume and pattern. Of course, this kind of object-oriented approach is applied to easily identifiable objects, such as rain areas. In oceanography this approach should be adapted to field distributions as in Hoffman *et al.* (1995) or applied to well defined structures, as a permanent eddy or a front.

A different approach to study the spatio-temporal characteristics is the decomposition of the model into different resolution scales, and study how these scales evolve in time. There are several ways to decompose the model into spatial scales. For example, Zepeda-Arce *et al.* (2000) used spatial averages to reach different resolution levels, and compared the model to observations at each level. They applied this technique to a rainfall model, although it can be applied to other variables.

Wavelet Transforms are a powerful tool for scale-decomposition analysis (Daubechies, 1992; Torrence and Compo, 1998). They are an extended form of the Fourier Transform, with a window that is translated and dilated over the studied domain, allowing us to separate the signal into orthogonal components (Mallat, 1998). Two-dimensional wavelets have been recently applied in the frame of forecast verification (Briggs and Levine, 1997; Casati *et al.*, 2004). In these works, a model is decomposed into several spatial scales using wavelets, and the quality of the model at each of these scales is assessed. The main interest in using this approach is that the scales that have a bigger contribution to the total error can be identified. As these scales have a physical meaning, we can identify the error with a physical process characteristic of this scale (Briggs and Levine, 1997).

The study of a model in a multi-scale approach can be applied to all the variables of a model. Briggs and Levine (1997) and Casati *et al.* (2004) worked with 500-hPa geopotential height fields and rain fields respectively. Alvera-Azcárate *et al.* (2004a) work with sea surface temperature fields. Liang and Robinson (2004a,b) established a multi-scale Energy and Vorticity Analysis (MS-EVA), that uses wavelets for the multi-scale decomposition. Liang and Robinson (2004c) used MS-EVA to study the energy and vorticity balances at different spatial and temporal scales of the Iceland-Faroe Front and considered the transfer and distribution of energy and vorticity between the

large scale, meso-scale and sub-mesoscale. A similar approach was used by Fournier (2002) and Fournier (2003a) to atmospheric fields.

Empirical Orthogonal Functions (EOFs) are also widely used in spatio-temporal analyses. EOFs allow to decompose a matrix into orthogonal modes representing the major patterns of variability found in the data. They are classified in order of importance, and usually the two leading modes account for more than 80% of the total variability (*e.g.* Yiou *et al.*, 2000). For example, Lermusiaux (2001) studied the three-dimensional variability of the Massachusetts Bay using EOFs. Lermusiaux (2002) made an analysis of the sensitivity of a mapping method to the size of the considered subspace, the studied scales and the system dynamics. Molcard *et al.* (2002) realised an EOF analysis of the wind stress over the Mediterranean Sea, to study the relationship between wind stress and the ocean response. EOFs have been also used for the reconstruction of missing data sets (Beckers and Rixen, 2003; Alvera-Azcárate *et al.*, 2004b), as it will be also shown in this work.

Some works have studied the feasibility of merging the two mentioned techniques, wavelets and EOFs. For example, Yiou *et al.* (2000) used the advantages of wavelets to make a multi-scale singular spectrum analysis (MS-SSA). Their approach allowed to analyse non-stationary time series with data-adaptative wavelets, using EOFs in the role of a wavelet. Wang and Chern (2000) combined also wavelets and EOFs to analyse internal tides. The combination of these two powerful techniques is still a field under developement, and it is giving a new insight into the understanding of geophysic processes.

Wavelets and EOFs present some similarities and differences:

Variability: In the interpretation of decomposition results, often the most dominant EOFs are retained, and the others are usually considered as noise. Only a part of the total variability is considered, but the less important EOFs add little information and it is sometimes desirable to eliminate them from the analysis. Wavelets conserve the initial variability of the decomposed field, if all scales are taken into account. Wavelets decompose the data in a limited number of scales depending on the initial size of the data set. The limited number of scales makes it easier to manage and interpret the decomposition.

Scales: Wavelets divide the domain into successive spatial scales. Each scale has a clearly defined resolution, and it is orthogonal to the other scales. EOFs divide the domain in an order of importance of the features detected, but not in scales. It is therefore difficult to relate a mode with a physical process, because they are not identified with a specific scale.

1.1.1 Forecast Verification in oceanography

All oceanographic models must be verified when released. This verification can be made by comparison with available observations, or by inter-comparison with other models (*e.g.* MEDMEX project (Beckers *et al.*, 2002), MERSEA intercomparison of operational models (Crosnier *et al.*, 2004), the Ocean Model Intercomparison Project (OMIP `http://www.clivar.org/organization/wgomd/pomip.htm`), etc).

Operational oceanography systems (*e.g.* MFSTEP (Pinardi *et al.*, 2002), MERSEA, EuroGOOS association, HYCOM (Bleck, 2002), MERCATOR (Bahurel *et al.*, 2001, 2004), etc.) need a systematic data quality control of their products. This quality control assessment must provide easily understandable statistics that may vary depending on the final user. Operational projects need to identify possible problems at an early stage, so a verification programme must be carefuly designed to be applied in an operational frame. Even in a non-operational frame, all model products should pass quality assessment tests.

This work has been realised in the frame of Satellite-based Ocean ForecasTing (SOFT) project (Alvarez, 2003; Alvarez *et al.*, 2004). SOFT project aims to improve operational systems with the assimilation of statistical predictors. Statistical predictors are better than climatology in a short-range frame, so its assimilation on an operational model can lead to better short-range forecasts. Statistical predictors are built from recent observations, thus they may contain specific events that are not contained in a climatological field.

1.2 Aim of this work

In this work, we present the verification of a three dimensional model of the Mediterranean Sea. There are two embedded models in this one, centred in the Ligurio-Provençal Basin and in the Ligurian Sea respectively. The highest resolution is reached in the Ligurian Sea model, so a big part of our work is concentrated in this model. An assimilation experiment has been realised. Sea surface temperature and sea level anomaly are assimilated into the model once a week, during nine weeks. The aim of this work is then to assess the impact of the assimilation on the model skill. An additional test has been carried out with the Ligurian Sea model: sea surface temperature SOFT statistical predictors have been assimilated in the same period as the mentioned assimilation test. SOFT assimilation results are then compared to observed sea surface temperature assimilation results, in order to establish the quality of SOFT systems.

In an attempt to consider all the aspects that have been enumerated in this introduction for a good verification process, the study of the model has been made in three major parts. First, the error is assessed in a qualitative way, that is, the model is compared "visually" to observations or bibliographic data. This first quality assessment allows us

to have an overall idea about how the model represents the observed values. We can also detect in an early step any major problem present in the model results. The second part of the forecast verification consists in a classical error assessment. Traditional error measures, such as the root mean square error or the correlation between two fields, are used to quantify in an objective way the error of the model. Spatial averages are used here to study the evolution of the error with time. The spatial information is thus lost. In the third part we will consider the spatio-temporal characteristics of the model, and we will analyse it using wavelets and EOFs. This spatio-temporal analysis will finish the quality assessment of the model. The three parts into which the verification process is divided follow in a logical scheme. This work aims to give a complete idea about the model behaviour, the effect of the nesting and the impact of the assimilation experiment.

1.3 Organization of the thesis

This thesis is organised as follows:

- In a first part, an overview about the data and model used is presented. The first chapter describes the Mediterranean Sea main circulation and characteristics of the water masses. Then the model used in this work is described, with the assimilation experiment and the SOFT predictors. Finally, the data used for the forecast verification are presented.

- The second part of this work contains the theory and methodology on which this work is based. There is a description of the forecast verification error measures, the theory of wavelets and a method for the reconstruction of missing data. This method will be used to reconstruct missing data in satellite fields. These reconstructed fields will be used in the spatio-temporal analysis of the model.

- The last part shows the results of the verification. The results are shown in the same order as the verification procedure is applied: first the qualitative error assessment, then the quantitative error assessment and finally the spatio-temporal analysis. There is an additional chapter where the results obtained by the assimilation of the SOFT predictors are analysed. This part follows the same structure qualitative, quantitative and spatio-temporal analysis.

Part I

Study site, Model and Data

Chapter 2

The Mediterranean Sea

The aim of this chapter is to give detailed information about the Mediterranean main circulation paths. The dynamic characteristics of the Mediterranean Sea are described for each basin, following the water course since it enters the Mediterranean to its exit, at the Strait of Gibraltar. People who are not familiar with the Mediterranean Sea dynamics may find here the basic information to understand some parts of this work.

2.1 General description

The Mediterranean is a semi enclosed sea with an extension of approximately 3000 km in the zonal direction and 1500 km in the meridional direction. It is connected to the Atlantic Ocean by the Strait of Gibraltar, a narrow canal of 20 km width and 300 m depth.

The Mediterranean Sea (see Figure 2.2) is an evaporation basin, because the freshwater input due to precipitation and river discharges do not balance the evaporation. This evaporation/precipitation balance mainly controls the circulation in the Mediterranean Sea. In the Strait of Gibraltar, the cool and light (S<36.5) Atlantic Water (AW) enters the Mediterranean above the saltier (S>38.4) and warmer Mediterranean Water (Viudez *et al.*, 1996). Due to evaporation and mixing, the Atlantic Water gradually changes its physical and chemical properties, forming the Modified Atlantic Water (MAW). The MAW forms a surface layer of about 100-200 m depth all over the Mediterranean Sea (Millot, 1999). Circulation in the Mediterranean Sea follows a cyclonic path all over its basins and sub-basins. The MAW contributes to the formation of an intermediate water mass in the Eastern Mediterranean due to warming and evaporation processes. The circulation pattern is completed by its return to the west of the basin, where it finally exits through the Strait of Gibraltar below the entering flow of AW. Other water masses formed in several parts of the Mediterranean Sea, contribute to diversify and to complete the circulation. They usually describe complex patterns, which in some cases are not completely understood.

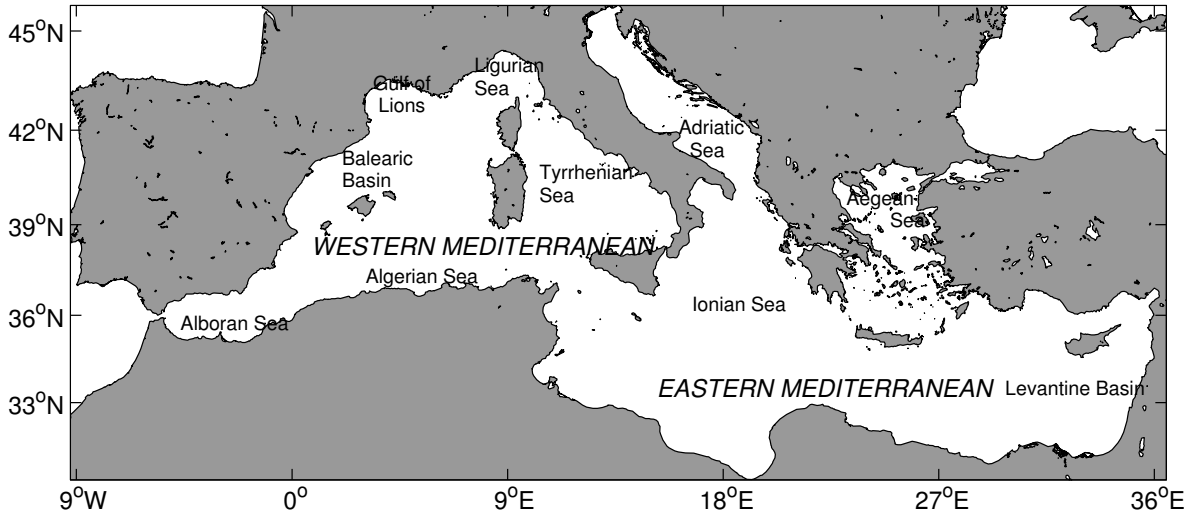

Figure 2.1: The Mediterranean Sea and its main basins and sub-basins

The Mediterranean Sea has two major basins, the Western basin and the Eastern basin, separated by the Strait of Sicily. Both basins are divided into several sub-basins. The Western basin is formed by the Balearic subbasin, limited to the north by the Gulf of Lions and to the south by the Algerian subbasin and the Alboran Sea; the Liguro-Provençal subbasin, at the north of the Corsica Strait; and the Tyrrhenian sub-basin, between the Italian peninsula and the islands of Sardinia and Corsica. The Adriatic, the Ionian, the Aegean and the Levantine sub-basins form the Eastern basin (see Figure 2.2.

2.2 Winds

Winds in the Mediterranean basin are mainly westerlies. There are two important wind regimes: the Mistral and the Etesian winds. The Mistral, a north-westerly wind, is a strong jet that dominates in winter, giving an important meridional component to the wind pattern. It increases the variability of the basin average wind speed amplitudes and directions. In early summer the Mistral is greatly reduced and its location is shifted to the north (Pierini and Simioli, 1998). The Etesian winds dominate in summer, giving a strong northerly component to the winds pattern, mainly in the Eastern Mediterranean (Castellari *et al.*, 1998).

At a subregional scale, the interaction of the westerlies with the local orography in winter, and with land-sea temperature contrasts in summer, characterise the surface atmospheric flow field (Pinardi and Masetti, 2000). Its persistence in winter leads to deep-water formation processes in the Western Mediterranean, because of the water

surface layers cooling.

The Mistral stress curl induces a cyclonic vorticity on the eastern flank of the wind jet in the north Mediterranean basin, and an anticyclonic vorticity on the western side of the jet current in the south Mediterranean basin.

2.3 Water masses

The names given to water masses are conceived to describe the place where they are formed, and the depth they are usually found because of their salinity and temperature characteristics. In general, all authors use a unique nomenclature. This allows to easily identifying all water masses. In some cases, there are several names used by different authors that refer to the same water mass. In this work we use the most common nomenclature for all water masses.

Some examples of water masses in the Mediterranean Sea are the Modified Atlantic Water (MAW) in the surface layer, the Levantine Intermediate Water (LIW) and the Winter Intermediate Water (WIW) in the intermediate layer, and finally the Western Mediterranean Deep Water (WMDW) and the Tyrrhenian Deep Water (TDW) in the deeper layers. The most relevant ones for this work are described here. A summary of this water masses characteristics is made in Table 2.1

- The MAW is the Atlantic Water (AW) that enters through the Strait of Gibraltar and travels eastward occupying a surface layer of about 100-200 m depth all over the Mediterranean basin. This water is characterised by a low temperature (about 12-13 °C) and a low salinity (less than 36.5). Both parameters increase progressively due to evaporation processes in the MAW eastward journey (Sparnocchia *et al.*, 1995).

- The LIW is generally formed in the Rhodes Gyre, in the Levantine basin, because of strong cooling processes during winter, i.e. cold winds that affect the whole region. The stratification in the centre of the gyre is strongly weakened, giving rise to convective processes. The thermocline, situated around 200-300 m in autumn, erodes down to a depth of 500-600 m during these events (Buongiorno Nardelli and Salusti, 2000). The LIW is characterised by high temperatures, with a maximum of 13.2-14 °C, and salinities of about 38.6, and is found at depths ranging from 200 m to 1000 m (Sparnocchia *et al.*, 1995; Millot, 1999). After its formation, the LIW circulates and disperses to the West. Some LIW re-circulates within the eastern basin, and some exits to the Western Mediterranean, flowing below the MAW surface layer through the Strait of Sicily. The LIW is the predominant source of

the salty Mediterranean water mass in the Northern Atlantic (Marullo *et al.*, 1999; Send *et al.*, 1999).

- The Winter Intermediate Water (WIW) is originated in the western basin, mainly in the Gulf of Lions, the Balearic Basin and the Ligurian Sea (Bouzinac *et al.*, 1999). It is formed as a result of the moderate winter cooling of the MAW surface layer due to the north-westerly winds. These moderate events are not strong enough to form a deep water mass. The WIW is not salty enough (salinities of 38.1-38.3) to be mixed with the LIW or deep water masses, so it can be recognized as a relatively low temperature (<13°C) thin layer situated between the MAW and the LIW (Albérola *et al.*, 1995; Millot, 1999; Send *et al.*, 1999).

- The Adriatic/Aegean Intermediate Water (AIW) is a water mass that has been detected along the Tunisian slope and at its bottom. It enters the Western Mediterranean through the Strait of Sicily, below the LIW. The AIW is a dense water mass, denser than the Western Mediterranean Deep Water (WMDW). The AIW flows along the Sicily Strait bottom and enters the Tyrrhenian Sea, where it may contribute to the formation of the Tyrrhenian Deep Water TDW (Send *et al.*, 1999). Other authors refer to this water mass as the transitional Eastern Mediterranean Deep water (tEMDW), Astraldi *et al.* (*e.g.* 2002), or as the Cretan Intermediate Water (CIW), Klein *et al.* (*e.g.* 1999).

- The Western Mediterranean Deep Water (WMDW) is formed because of the persistent cooling processes in the central zone of the western Mediterranean, mainly in the Gulf of Lions. It is the most characteristic water mass formed in the western Mediterranean, and fills the deeper levels of the western basin (Send *et al.*, 1999). The WMDW is characterized by a potential temperature of 12.7-12.9°C, a salinity of 38.4 and a density of about 29.12 σ_θ units. It is situated below the LIW, from 600 m to the bottom of the basin (Albérola *et al.*, 1995; Sparnocchia *et al.*, 1995).

- The Tyrrhenian Dense water is the result of an active mixing of LIW, Eastern Mediterranean Deep Water (EMDW, explained later) and WMDW when they enter the Tyrrhenian Sea. The result is a water mass denser than the LIW, but not dense enough to be mixed with the bottom layers water (Millot, 1999).

- The Tyrrhenian Deep Water (TDeepW) is warmer and more saline than the WMDW. The Tyrrhenian Sea is nearly totally enclosed below a depth of 400 m, with the Tyrrhenian Trough (1900 m) as the unique deep connection, in the Sardinia Channel. The TDeepW is formed by inflowing WMDW that mixes with

Table 2.1: Summary of some of the mediterranean water masses characteristics

Name	Depth	Temperature	Salinity
MAW	0-200 m	12-13°C	<36.5
LIW	200-600 m	13.2-14°C	38.6
WIW	200 m	<13°C	38.1-38.3
WMDW	600-bottom	12.7-12.9°C	38.4

the overlying warmer and more saline LIW and the dense AIW. The outflow of the TDeepW from the Tyrrhenian Sea is estimated in 0.4 Sv, and its signal can be found all around the Western Mediterranean, from 600 to 1900 m depth (Send *et al.*, 1999). Some authors (*e.g.* Millot, 1999) do not make a distinction between TDeepW and WMDW water masses, identifying the WMDW as the unique water mass in the bottom of the Tyrrhenian Sea.

- The wintertime Northern Adriatic Dense Water (NADW) is formed in the north part of the Adriatic Sea. As it flows southwards, it sinks due to an open ocean deep convection process, reaching a depth of about 1200 meters. This water mass flows at depth out of the Otranto Strait and is transformed in Eastern Mediterranean Deep Water (EMDW). It flows along the deep western boundary of the Ionian basin following the isobaths and drives the entire Eastern Mediterranean deep thermohaline circulation (Bergamasco *et al.*, 1999; Marullo *et al.*, 1999). Another water mass is identified in the Adriatic Sea, the Adriatic Deep Water (ADW) formed in the south of the basin, and that contributes also to the formation of the EMDW (Marullo *et al.*, 1999).

2.4 Surface Circulation

The circulation in the Mediterranean Sea occurs at mainly three space scales: basin scale, subbasin scale and mesoscale. It can also be thought in terms of time scale, as interannual scale (i.e. deep water mass formation and variations in volume transport) and seasonal scale (i.e. thermocline variations, surface gyre variations) and smaller scales (Brankart and Brasseur, 1998; Pinardi and Masetti, 2000).

At a basin scale, the most important pattern is the zonal circulation of MAW. It enters the Strait of Gibraltar and travels eastward to the Levantine Basin, where it takes part in the LIW formation processes. This newly formed LIW returns to the west and exits the Mediterranean through the Strait of Gibraltar below the MAW. The subbasin

scale is characterized by gyres (subbasin scale gyres) and permanent or quasi permanent cyclonic and anticyclonic structures interconnected by intense jets and meandering currents (Malanote-Rizzoli *et al.*, 1999).

The path of the main currents in the Mediterranean Sea can be seen in Figure **??**.

2.4.1 Western Mediterranean

Alboran Sea

The AW enters the Mediterranean Sea through the Strait of Gibraltar, forming a jet of about 30 km width. The total inflow transport is of about 0.92 Sv, and the outflow transport varies between 0.84 Sv and 1.26 Sv (Bryden and Kinder, 1991; Bethoux and Gentili, 1999; Hopkins, 1999; Send *et al.*, 1999). The balance must be positive (inflow greater than outflow) to compensate the excess of evaporation that takes place in the basin. As the water passes the strait, the jet forms a meandering structure that results in the formation of a quasi-permanent anticyclonic gyre in the West Alboran Sea, called the Alboran gyre. The Alboran gyre has a strong seasonal variability, with a maximum in summer (Larnicol *et al.*, 1995, 2002). After this, the circulation pattern becomes more variable and presents most of the time a second gyre, also anticyclonic (Viudez *et al.*, 1996; Beckers *et al.*, 1997). The MAW continues to flow eastwards from Spain to Algeria, forming the 'Almeria-Oran jet' and reaching the Algerian Sea (Millot, 1999).

Algerian Sea

The surface current in the Algerian Sea is relatively narrow (about 30 km) and deep (200-400 m at the coast) and with a transport of about 1.7 Sv (Benzohra and Millot, 1995; Viudez *et al.*, 1996). The current is unstable, and presents several branches and meandering structures of 50-150 km width. These structures lead to the formation of coastal cyclonic and short-lived eddies, and two permanent coastal anticyclonic eddies that propagate a few kilometers per day and last for months (Send *et al.*, 1999). These structures, with a diameter of 50-100 km, continue to propagate to the East with the main current, along the Algerian slope. An upwelling is associated with the seaward part of the second anticyclonic eddy. The whole Algerian Current is deviated seawards, due to the interaction with open sea eddies, *i.e.* anticylonic eddies observed in the middle of the basin with a diameter of about 200 km, and that could be generated by coastal eddies which separate from the coast near the Channel of Sardinia (Benzohra and Millot, 1995; Millot *et al.*, 1997; Millot, 1999).

At this point, the Atlantic jet splits into two branches in its journey to the East. The first one appears as a result of the eddy detachment from the Algerian Current. This branch flows northward from the Algerian Sea at a longitude of 1°E. It passes then

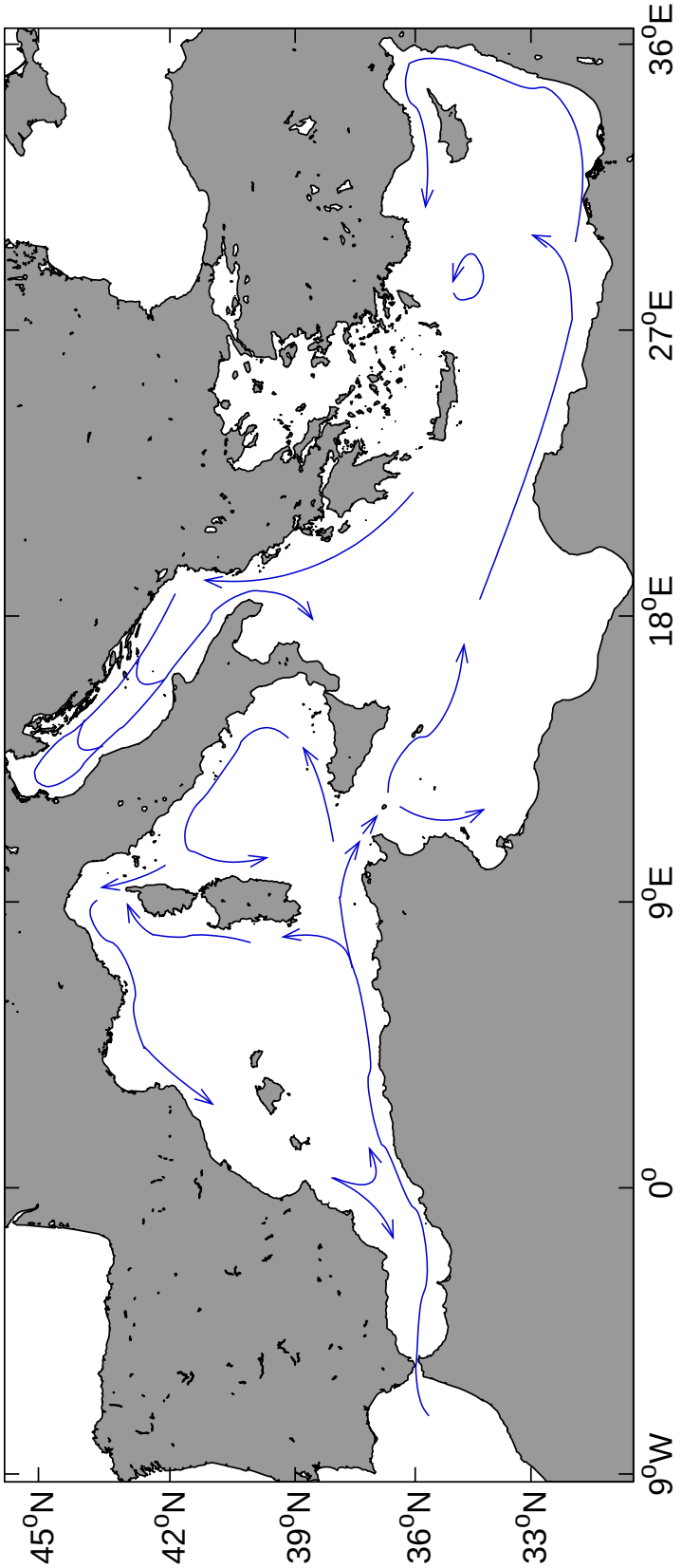

Figure 2.2: The Mediterranean Sea and its main basins and sub-basins

near the west coast of Corsica where it forms the Western Corsican Current, and joins the Northern Current in the Liguro-Provençal basin (Albérola *et al.*, 1995). Along the African coast, the eastward flow continues and passes the Channel of Sardinia. After this, before the Strait of Sicily is reached, the current is divided again, with the Atlantic jet flowing to the east towards the Ionian basin, and a northward branch that passes the west coast of Sicily and enters the Tyrrhenian Sea (Astraldi and Gasparini, 1992; Pierini and Simioli, 1998; Astraldi *et al.*, 2002).

Tyrrhenian Sea

The circulation in the Tyrrhenian Sea presents a seasonal variation, mainly due to changes in the wind patterns. In winter, a basin-wide cyclonic gyre is present. The branch of MAW that turns northward after the Sardinia Channel, flows along the west coast of Sicily and joins here this cyclonic circulation. The MAW can follow the whole cyclonic gyre and then exit the Tyrrhenian Sea following the south coast of Sardinia, where it joins the MAW to form the West Corsican Current (Astraldi *et al.*, 2002). Some of the MAW in the Tyrrhenian Sea flows northward and exits the Tyrrhenian Sea through the Corsica Strait. It joins there the Northern Current as the East Corsican Current. Some authors (*e.g.* Vignudelli *et al.*, 1999) have observed a reversal of this current in summer. The cyclonic gyre is smaller and weaker in summer, usually restricted to the north of the basin, due to the reduced wind intensity. The circulation can also switch from cyclonic to anticyclonic (Astraldi and Gasparini, 1992; Ayoub *et al.*, 1998; Pierini and Simioli, 1998).

Liguro-Provençal basin

The circulation in the Ligurian Sea describes a cyclonic gyre, which is more intense in winter and is mainly due to wind curl stress (Larnicol *et al.*, 1995). The two northward currents that surround the coast of Corsica, the West Corsican Current and the East Corsican Current, feed the circulation pattern of this area. Both branches join in the Ligurian Sea, north of Corsica, forming the Northern Current, that flows south-westward following the French and the Spanish coasts along the continental slope. The signal of the Northern Current extends from the north of Corsica to as far as the Catalan Sea (Albérola *et al.*, 1995; Sammari *et al.*, 1995; Astraldi *et al.*, 1999; Millot, 1999).

This region is exposed to intense atmospheric processes in winter leading to deep-water formation. This deep-water formation causes an annual deficit of heat and water in the basin. This loss of heat and water is mainly compensated by the Eastern Corsican Current, as it will be seen in the following section (Astraldi *et al.*, 1999).

- **Seasonal variability:** the Northern Current presents its highest speed values in late January, with a mean of 15-30 cm/s. During the winter season, the current is

narrow (less than 20 km wide) and deep and it is placed close to the coast, mainly from late February to mid-march (Albérola and Millot, 2003). The transport is also maximal in this period, with about 1.6 Sv in December. From April to the end of the year, the Northern Current is wider (more than 30 km wide) and shallower, with lower current speeds than in January, and a transport of about 1 Sv (Albérola *et al.*, 1995). This seasonal variability in the Northern Current is mainly due to seasonal changes in the East and West Corsican Currents:

→ The West Corsican Current varies seasonally, reaching its maximum transport values at the beginning of spring. Then it decrease progressively to reach a minimum in autumn. These seasonal changes are mainly due to the dense water formation processes occurring in the deep winter in the Liguro-Provençal basin (Sammari *et al.*, 1995). Its volume transport, considering a current of about 30 km width and 600 m depth, is of 1.15 Sv, and it presents no annual variability. Although more important in volume than the East Corsican Current, this last one influences to a major extent the Northern Current characteristics (Astraldi and Gasparini, 1992).

→ The East Corsican Current is driven by thermohaline conditions, and its signal is maximum in early winter (Sammari *et al.*, 1995). High transport values persist all the cold season, weakening in spring to reach its minimum signal in summer and autumn. This seasonal oscillation can be observed in the surface MAW as well as in the LIW. The flow through the Corsica Channel is composed by about 80% of MAW and a minor part of LIW (Astraldi *et al.*, 1999). The volume of surface and intermediate flows crossing the Corsica Channel has been established in about 0.5 Sv for the surface layer and about 0.15 Sv for the LIW. Nearly 80% of this transport is concentrate in the winter and early spring seasons. It is thought that the East Corsican Current is responsible for the heat and water exchanges with the Northern Current. These heat exchanges, the highest of the Western Mediterranean Sea, induce a density gradient between the Tyrrhenian and the Ligurian Seas that forces the transport from the first to the second one. The heat coming from the Tyrrhenian basin restores the winter losses, the major factor conditioning the basin surface layer, and responsible for the deep water formation in this part of the Mediterranean basin (Astraldi and Gasparini, 1992; Astraldi *et al.*, 1999).

- **Mesoscale variability:** a seasonal variation of the Northern Current due to the seasonal behaviour of the Eastern and Western Corsican Currents has been described (Albérola *et al.*, 1995). However, an intense mesoscale variability is also present, that appears as meandering structures. The meanders can vary in length (from some 10 to 100 km) and in velocity (from 10 to 20 km d^{-1}). Thus,

these meandering structures can be as wide as the Northern Current and they can advance at the same velocity (Sammari *et al.*, 1995). The mesoscale features are very energetic in winter (Albérola and Millot, 2003), and have longer life periods, from 10 to 20 days. These structures are visible all the year, but the weaken from spring to autumn. Their life periods shorten to about 10 days in this period. The mesoscale features propagate seawards (Albérola *et al.*, 1995; Ayoub *et al.*, 1998).

As already mentioned, the Northern Current signal extends as far as the Catalan Sea, and thus, the seasonal and mesoscale variations described above can be also seen all along its south-westward flow (Font *et al.*, 1995).

Balearic Basin

The Balearic Basin is limited to the north by the Liguro-Provençal basin and to the south by the Algerian basin. The Balearic Islands form an arc that perturbs the flow that passes through them. The circulation in the Balearic Basin is characterised by a weakened Northern Current flowing southward through the Channel of Ibiza, and that results in the formation of mesoscale eddies that are present most of the time. These eddies, called "weddies" as they are formed by WIW, cause a recirculation of the Northern Current in the Arc of Balearic Islands causing the recirculation of the southward transport through the channels (Pinot *et al.*, 1995).

After passing this barrier, the water continues to flow southward, and the LIW exits the Mediterranean through the Strait of Gibraltar while the MAW joins again the fresher MAW entering the Strait in its travel to the East (Astraldi *et al.*, 1999).

2.4.2 Eastern Mediterranean

Ionian Basin

The MAW enters the Eastern Mediterranean through the Strait of Sicily, often meandering due to the complex local topography that divides the strait into two channels. The atmospheric forcings also contribute to the instability of the surface current. The current is now named the Atlantic-Ionian Stream (AIS), and brings the water to the Levantine Basin across the central Ionian basin (Ayoub *et al.*, 1998).

After passing the Strait of Sicily, at 37°N, $16 - 17$°E, the IAS bifurcates into two branches. One turns to the south part of the Ionian basin to the Tunisian coast, describing an anticyclonic pattern. The other continues to the East, crosses the basin and turns south to arrive to the Levantine basin through the Cretan passage as the so-called Mid-Mediterranean Jet, MMJ (Malanote-Rizzoli *et al.*, 1997; Robinson *et al.*,

1999; Lermusiaux and Robinson, 2001). Once the jet enters the Levantine basin, it travels along the African coast. At about 27° E, the MMJ turns north-east and its signal weakens; it has arrived to the end of its eastward travel, and now it will contribute, by evaporation and mixing, to the deep-water formation processes that happen in this part of the Mediterranean basin (Ayoub *et al.*, 1998).

Levantine Basin and Aegean Sea

The circulation in the Levantine basin is very unstable. There are several stationary mesoscale patterns, such as the Rhodes cyclonic gyre, considered as the primary source of LIW, situated at 34°N, 28°E. The seasonal cycle is very strong in this part of the Mediterranean, with energetic boundary currents in winter, and weaker in summer. This variability is mainly caused by the wind curl (Molcard *et al.*, 2002).

The Aegean basin presents many factors that make the circulation very variable, as the irregular bottom topography, the numerous channels and island chains, and the inflow of the colder and less salty Black Sea waters (Zervakis *et al.*, 2002). In general, the surface circulation presents a cyclonic gyre in winter and a southerly movement in summer, caused mainly by the wind regime, the Etesian (Poulos *et al.*, 1997).

Adriatic Sea

The Adriatic Sea is a subbasin of the Mediterranean Sea (see Figure 2.3) of about 800 by 200 km, connected to the Ionian Sea by the Strait of Otranto. The northern and central parts of the basin are very shallow, with maximum depths of 270 m, and a mean depth of 35 m (Artegiani *et al.*, 1997a). The southern part is deeper, reaching 1200 m, but in the Strait of Otranto, the depth decreases again to 780 m (*e.g.* Cushman-Roisin *et al.*, 2001). The circulation in the basin is cyclonic, with water coming from the Ionian basin and entering the Strait of Otranto to the east of the Adriatic Sea (Eastern Adriatic Current, EAC). This current is wide and weak and brings at depth warm and salty modified Levantine Intermediate Water to the northern basin. A western current flowing southward closes the cyclonic circulation. This current is called WAC (Western Adriatic Current) and is thinner than the EAC. In winter, it brings at surface cold and fresh water to the southern basin. Salty and cold water coming with the WAC fills the depth at this southern part. Embedded in this main circulation path are three sub-basin cyclonic gyres, in the northern part of the Adriatic, between the Istrian Peninsula and the Jakuba Pit, and in the southern Adriatic Sea, respectively (*e.g.* Orlić *et al.*, 1992; Poulain, 2001). The basin-wide cyclonic surface circulation is mainly produced by winds. There are two dominant winds: the steady south-easterly wind, called Sirocco, which enters the basin by the Strait of Otranto, and the north-easterly wind called Bora, which enters the Adriatic Sea from the Northeast (Bergamasco *et al.*, 1999; Poulain, 1999). The fresh water input in the Adriatic Sea is mainly due to rivers in the northern and eastern

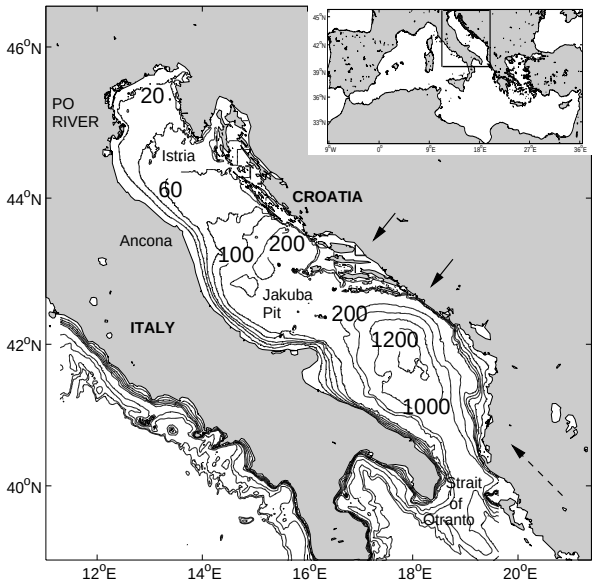

Figure 2.3: The Adriatic SeaSea and its bathymetry (depth in meters). The arrows show the typical winds found in this area. In solid line, the Bora, and in dashed line, the Sirocco.

coasts. The Po River accounts for 28% of this runoff (Cushman-Roisin *et al.*, 2001). The mean surface temperature in winter is 14°C. The southern part is warmer (24°C) than the northern part (22°C) in summer (values obtained from the MEDAR/Medatlas climatology, (MEDAR-Group, 2002)). In the zonal direction, the water is warmer in the open sea than in the eastern coastal waters. A coastal front can be found throughout the year, but its position varies between seasons. For a detailed description of the Adriatic's characteristics see *e.g.*, (Orlić *et al.*, 1992), (Artegiani *et al.*, 1997a), (Artegiani *et al.*, 1997b), (Cushman-Roisin *et al.*, 2001).

2.5 Thermohaline Circulation

2.5.1 Eastern Mediterranean

In the Levantine Basin, mainly in March, severe cooling events lead to the formation of LIW. This new water spreads at a depth of about 500 m, and it is situated below the MAW. Some LIW exits the Eastern Mediterranean through the Strait of Sicily and some re-circulates in the basin (Marullo *et al.*, 1999).

Some changes have been discovered concerning the Eastern Mediterranean circulation in the last decade. Studies in the 80s showed that deep-water formation processes happened mainly in the Southern Adriatic. In this basin ADW exits from the Otranto

Straits, transforms into EMDW and spreads in the Eastern Mediterranean bottom layer. In the decade of the 90's, the circulation path has changed, affecting mainly the deep and intermediate water masses. This change has been named Eastern Mediterranean Transient (EMT) (*e.g.* Theocharis *et al.*, 2002).

Some studies indicate that the LIW is confined in the Levantine Basin, limited in its westward spread by an anticyclonic structure situated in the south of the basin. To replace this LIW, water from the Aegean Sea (Cretan Intermediate Water, CIW) exits by the Cretan Arc Straits to the Ionian Basin. Another water mass of Cretan origin has been also recognized. This water spreads to the Ionian and Levantine basins. This water mass is saltier (greater than 38.8) and warmer (more than 13.6° C) than the CIW and also than the Cretan Deep Water (CDW). This water mass has been named Cretan Sea Overflow Water (CSOW) by Klein *et al.* (1999). The Ionian and Levantine Seas have suffered a massive salt increase in the bottom layer, below 1200 m (Lascaratos *et al.*, 1999; Malanote-Rizzoli *et al.*, 1999), and a moderate increase of the salinity content in the upper 500-1400 m intermediate and surface layers. The intermediate layers in the Eastern Mediterranean are weaker due to the observed EMT. As a consequence, the flux of intermediate waters, mainly LIW and CIW, to the Adriatic Sea is less important. This results in the decrease of salt content on the Adriatic Deep Water (Klein *et al.*, 1999). The main deep-water formation basin is now the Aegean Sea, instead of the Adriatic Sea.

This change in the origin of dense water from the Adriatic to the Aegean Sea has several consequences. First of all, the rate of dense water formation in the Aegean is nearly 3 times greater than in the Adriatic, being now of about 1 Sv (Klein *et al.*, 1999; Malanote-Rizzoli *et al.*, 1999). In addition, its higher salinity has caused the isopycnals to rise about 500 m. This new water has replaced 20 % of the old deep water. Recent studies, however, indicate that the Adriatic Sea is still a source of deep water. In spite of the EMT, very cold winters and the recirculation of the CIW to the Adriatic Sea may guarantee the continuation of deep water formation in the southern Adriatic (Manca *et al.*, 2002; Theocharis *et al.*, 2002).

The reasons for the increasing rate of dense water formation in the Aegean Sea could be the extreme climate conditions suffered in the Eastern Mediterranean in the period 1988-1993: long dry periods and exceptionally cold winters. The meteorological conditions may also affect the circulation patterns of the intermediate water masses, LIW and CIW (Lascaratos *et al.*, 1999).

The signal of this recent change has spread from the Eastern to the Western Mediterranean. It takes about two or three years to reach the Sicily Strait (Astraldi *et al.*, 2002), and about one year and a half to reach the south Tyrrhenian Sea. The Tyrrhenian Sea is considered as the transition basin between the Eastern and the Western basins, and the effects of this change will be transferred to the Western Mediterranean via the deep circulation pattern, through the deep Tyrrhenian Trough. It is not clear

yet if this reversal in the climatic conditions is a permanent change, and a new reversal to the original conditions has been proposed to happen in 1998 by Astraldi *et al.* (2002).

The water present in the Strait of Sicily spreads to the West in the return path of the basin-scale circulation. It is composed by LIW formed in the winters before the circulation change, and by CIW (Malanote-Rizzoli *et al.*, 1999).

2.5.2 Western Mediterranean

The water flowing from the Eastern Mediterranean to the Western Mediterranean is mainly composed by LIW in the upper part (from 200 to 800 m) and EMDW at the bottom layers. When this flow passes the strait of Sicily, it describes a mainly cyclonic path around the eastern and northern boundaries of the basin, so its signal is weak in the interior and south of the basin. After passing the Sicilian coast it turns thus northward, following this cyclonic path. As it enters the Tyrrhenian Sea, the dense flow (EMDW) cascades along the bottom, to reach a maximum depth of 1850 m, above the WMDW (or TDeepW). This cascading process involves strong mixing between water masses, which lead to the formation of TDW. The LIW continues to flow at its characteristic depth, but it also suffers partial mixing with this TDW (Millot, 1999; Sparnocchia *et al.*, 1999). The Tyrrhenian Sea can be considered as the place where the East Mediterranean Water and the Western Mediterranean Water mix up. Some of these water masses disappear to generate the new ones (Astraldi *et al.*, 2002).

The intermediate flow of LIW follows the path described by the MAW. Part of it exits the Tyrrhenian Sea by the Corsica Channel and arrives to the Ligurian Sea. The water that do not pass through the narrow and shallow Corsica Channel, continues the cyclonic circulation to exit through the Sardinia Channel.

In the southern branch along the coast of Sardinia, the LIW vein (now mixed with the TDW) became wider, shallower (from 50 km to 120 km wide and from 800 to 750 m depth) and cooler. The reason for this abrupt change is the interaction with the Algerian Current, namely with the Algerian eddies. Another hypothesis would be the instability of the LIW vein, which can lead to the formation of anticyclonic structures. Once the Sardinia Channel is passed, the cyclonic circulation continues, also following the path of MAW along the west coasts of Sardinia and Corsica. Once in the north of Corsica, the LIW joins the outflow of the Corsica Channel. There, it contributes to the formation of deep-water masses. The circulation continues south-westward along the French and Spanish continental slopes, although the interaction with several mesoscale structures may also spread the LIW seawards. After the Spanish coast, it continues to the Alboran Sea and exits to the North Atlantic Ocean through the Strait of Gibraltar. Not all the LIW exits the Mediterranean Sea, as the interaction with the Almeria-Oran jet deviates some of the water eastward to the Algerian coast. This deviated water flows

eastward and joins the more recent LIW west of Sardinia (Millot, 1999; Fuda *et al.*, 2000).

The WIW is formed in the Gulf of Lions, the Ligurian Sea and the Balearic Sea, because of moderate cooling processes in winter. It follows the cyclonic circulation path of the MAW, along the French and Spanish slopes. However, the WIW circulation is more affected by the topography than the MAW. It can also be transported to the interior of the basin due to the interaction with mesoscale eddies. The formation of some WIW lenses in the North-western Mediterranean has been observed, (Gasparini *et al.*, 1999). They may be the result of instabilities in WIW formation processes (Benzohra and Millot, 1995; Fuda *et al.*, 2000).

Deep-water formation processes in the Gulf of Lions, at the north of the Western Mediterranean, lead to the formation of WMDW. The deep circulation also follows a cyclonic path, following the French and Spanish continental slopes. At the Algerian Sea the topography forces the WMDW to enter the Tyrrhenian Sea through the Tyrrhenian Trough, and not to continue to the Eastern Mediterranean because of the shallow depths of the Sardinia Channel. After flowing in the Tyrrhenian Sea (always in a cyclonic path) it exits again by the Tyrrhenian Trough, as the Corsican Channel is also too shallow (450 m) to allow the deep water circulation (Millot, 1999). The authors that see the WMDW and the TDeepW as different water masses (*e.g.* Send *et al.*, 1999) consider that the inflow of WMDW to the Tyrrhenian Sea may induce a compensatory outflow of TDeepW to the Algero-Provençal basin (always through the Tyrrhenian Trough). As a result, the deepest layers of the whole Western Mediterranean would be filled by this deep-water mass, with the WMDW lying below it.

Chapter 3

The Model

3.1 Description of the GHER model

The GHER model, a three dimensional model (Beckers, 1991) has been applied to the Mediterranean Sea. The model is governed by the primitive equations under the Boussinesq and *beta*-plane approximations, it has free surface and is implemented in an Arakawa C-grid: scalar variables (temperature, salinity, elevation and turbulent kinetic energy) are defined at the center of each grid box and the normal velocity component is defined at the interface of them. In the vertical the model uses a double-sigma coordinate, with 31 levels. The numerical scheme conserves mass, heat and salt. In particular, the advection scheme is a monotone TVD scheme preserving fronts and gradients. Other details of the GHER model can be found in Beckers (1991).

A 1' bathymetry has been used (Smith and Sandwell, 1997) in the model. The initial salinity and temperature are computed from the MODB database (Brasseur *et al.*, 1996). Two river freshwater inputs are also included: the Rhone (in France) and the Arno (in Italy) rivers, obtained from Tusseau and Mouchel (Tusseau and Mouchel, 1994) and from Rinaldi (personal communication), respectively. Heat and momentum fluxes are calculated from bulk formulaes. Temperature of the air at two meters, cloud coverage, wind speed at 10 meters, dew point temperature and pressure fields are obtained from the European Center for Medium range Weather Forecasting (ECMWF), and used for the computation of fluxes. The fluxes used in the model are interactive, *i.e.*, they depend of the surface temperature provided by the GHER model.

This model has been applied in a two-way nesting aproach to the Ligurian Sea (Barth *et al.*, 2003). The model consists in three nested sub-domains: the coarsest grid represents the whole Mediterranean, with 15' of resolution (22×28 km or $1/4°$), the intermediate grid covers the Liguro-Provençal basin (3' resolution, 4×6 km, $1/20°$), and the finest grid is centered in the Ligurian Sea, with 1' resolution (1.5×1.9 km, $1/60°$). This three domains can be seen in Figures 3.1 to 3.3, where the bathymetry is also represented.

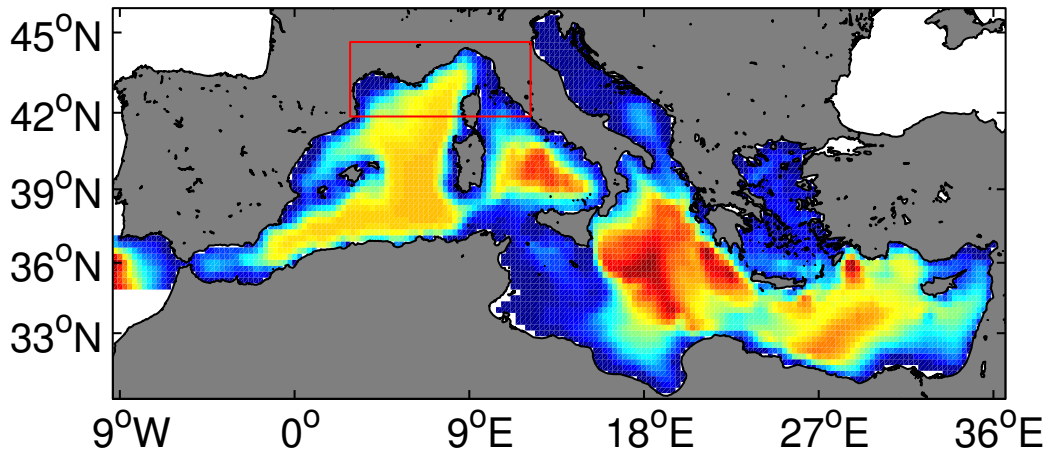

Figure 3.1: The Mediterranean Sea grid (22 × 28 km resolution) and its bathymetry. The red square represents the next nested domain, the Provençal Basin.

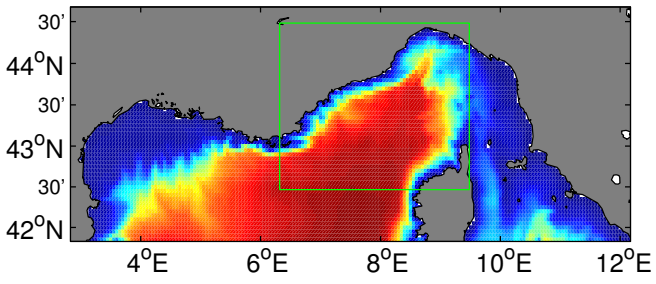

Figure 3.2: The Provençal Basin (4 × 6 km resolution) and its bathymetry. The green square represents the next nested domain, the Ligurian Sea.

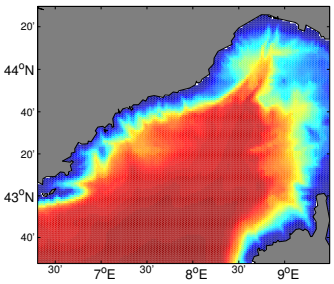

Figure 3.3: The Ligurian Sea grid (1.5 × 1.9 km resolution) and its bathymetry.

The two-way nesting approach allows the three grids to change information between them. Nesting is used to attain a higher resolution in a specific zone, where the small scale features can be resolved, as a highly variable current or mesoscale eddies. In the two-way aproach the information (velocity, temperature, salinity and turbulent kinetic energy) is passed from the coarse grid model to the fine grid model by interpolation of the coarse grid boundary conditions, which are imposed to the fine grid model. The coarse and fine grid models are integrated one step ahead, and then the fine grid model gives a feedback to the coarse grid model. The fine grid values are averaged and imposed to the coarse grid model. The time step of the three models is the same. As they run in a parallel computer no computational time would be gained if a higher time step was chosen for the coarse models. The refinement is done in the horizontal, but not in the vertical, so the sigma-layers are the same for the three models. The state vector contains all three grids together. Nesting procedures are used when it is necessary to resolve the physical processes of a particular zone in more detail. The increased resolution allows to study in detail a given zone, without the increased computational effort that it would result if the resolution is increased over the whole domain. The two-way approach is very beneficial when the physical features of the fine grid domain extend over the coarser domains. The exchange of information between all levels of resolution will improve the representation of such features in the fine and coarse grids.

3.2 The Assimilation Experiment

Data assimilation is a common tool used to control the error of a model. Model dynamics are combined with available data in an optimal way, to constrain the model to the observed dynamics. The uncertainty of the model is thus reduced. The combination of observed data with model dynamics can be done in several ways, and research in this field has given rise to a wide variety of algorithms for data assimilation.

A spin up of the GHER model from the 1st January 1998 to the 5th July 2000 is used to compute initial conditions. From the 6th July to the 1st September 2000 an assimilation experiment has been carried out. Each week Sea Surface Temperature (SST) and Sea Level Anomaly (SLA) satellite data are assimilated into the model via a Singular Evolutive Extended Kalman filter (SEEK) (Pham *et al.*, 1998; Brasseur *et al.*, 1999; Brankart *et al.*, 2003), which is a reduced order Kalman filter. For the computation of the model error covariance, an ensemble of 200 members has been generated. Initial conditions and atmospheric forcings are perturbed and then the ensemble of 200 members runs for 2 weeks, beginning on 5 July. The final ensemble is used to calculate the error covariance of the model.

The state vector for the assimilation contains the elevation, temperature and salinity for all three nested grids. The correlation of the variables between nested domains can

be taken into account by this approach. A unique state vector can avoid, for example, artificial gradients at the boundaries after an assimilation cycle. See Barth *et al.* (2004) for a detailed description about the assimilation experiment analysed in this work.

3.3 The SOFT System

Satellite-based Ocean Forecasting (SOFT) systems have been also assimilated in the model. The aim of this part of the work is to assess the differences in skill of predicted data when assimilating SST and when assimilating statistical predictors of the ocean state. The use of statistical predictors can be of great help in a real forecasting frame, when no observations are available to force the model. The SOFT predictors are constructed from an EOF decomposition of a time series set of data of a determined zone. They have been tested in a variety of locations (Álvarez *et al.* (2000) in the Alboran Sea, Alvarez (2003) in the Adriatic Sea, Alvarez *et al.* (2003) in the Tuscan Arphipelago, and Alvarez *et al.* (2004) in the Ligurian Sea). In this work the SOFT predictors of the Ligurian Sea are used (Alvarez *et al.*, 2004). These predictors are based on an EOF decomposition obtained from a series of 393 averaged satellite images of the Ligurian Sea ranging from 1 March 1993 to 11 October 1999. The images have been processed by the German Aerospace Research Center (DLR), and its processing will be described in section 4.2.

The empirical prediction of satellite observations is done in three major steps. First, the space-time variability of the data is obtained with an EOF decomposition:

$$F(x,y,t) = A_T^1(t)P_T^1(x,y) + ... + A_T^N(t)P_T^N(x,y) + T(x,y) \tag{3.1}$$

where $A_T^n(t), n = 1...N$, are one-variable time series and $P_T^n(x,y)$ are spatial patterns. T refers to temporal variance decomposition and $T(x,y)$ is the time mean of the satellite images substracted in the decomposition.

Once the EOF decomposition of these images has been done, the most dominant EOFs N_r are retained in order to eliminate noise from the time series. The last step is the time series prediction. This is done by means of a dynamical model that is obtained for each of the time series $\tilde{A}_T^n(t), n = 1...N_r$:

$$\hat{A}_T^n(t) = \gamma_n(\tilde{A}_T^n(t-\tau), \tilde{A}_T^n(t-2\tau), ..., \tilde{A}_T^n(t-m\tau)) \tag{3.2}$$

where $n = 1, ..., N_r$, m is an embedding dimension and τ is a time lag unit.

A genetic algorithm called DARWIN, (Alvarez and Tintoré, 2001; Alvarez, 2003) for time series prediction is used for the approximation of γ_n. This predictor randomly generates a set of equations for each n-amplitude equation $A_n(t)$. The equations have the form of Eq. 3.2 and are formed by random sequences of the variable A at passed times,

combined with real-number constants and the basic arithmetic functions (summation, substraction, multiplication and division). Small parts of the equations are mutated randomly. Two additional operators are included to avoid sub-optimal solutions: an iterative operator, that runs DARWIN several times with different initial conditions, and a random immigration of some members, that are replaced with new ones. A fit to the data is done to measure how well an equation behaves. The prediction of SST and SLA in the Ligurian Sea is done with this system. The prediction has been done at weekly time scales, and these predictions will be assimilated to the GHER model, in the same way as observed SST. We will made a comparison of forecast skill for these two assimilation procedures will be made. The aim of this experiment is to see the feasibility of the predictors assimilation. In a real forecast experiment, when no future observations are available, it can be very useful to work with these predictors.

Chapter 4

The Data

4.1 Satellite Data

Weekly Sea Surface Temperature data

The model is forced by data assimilation of satellite data. These data are choosed to cover the whole period from 6 July to 1 September. Advanced Very High Resolution Radiometer (AVHRR) SST data from the German Aerospace Research Center (DLR) (`http://eoweb.dlr.de:8080/servlets/template/welcome/entryPage.vm`) is used. These data are mean weekly composite temperature fields. The mean is computed from the daily maximum temperature images, for every pixel. Clouded pixels are excluded from the data, so that the composites are only calculated from clean pixels. The use of maximum temperature for the computation of the weekly composite may add some bias to the observations, because skin temperature in summer can attain high values and it is not eliminated from these images. To minimise the bias between model and observations, the model output is done at 14:00 pm, so that it contains also daily reheating. The resolution of these images is of 1 km. in the zonal and in the meridional directions.

Sea Level Anomaly data

Sea Level Anomaly (SLA) data is obtained from the Naval Research Laboratory, from `http://www7300.nrlssc.navy.mil/altimetry/data.html`. The data used in this work is a combination of several satellites (TOPEX/POSEIDON+ERS-2+GFO) allowing a higher coverage of the Mediterranean that with a single satellite. The track of those satellites are grouped together in one-week sets, and then they are assimilated into the model. The coverage at each week is shown in Figure 4.1. We can see that there is in general an homogeneous distribution of the satellite tracks.

A series of geophysical corrections have been made to the SLA data (dry troposphere path delay, wet troposphere path delay, ionosphere path delay, electromagnetic bias and static inverse barometric correction). In addition, these data have been quality-controlled

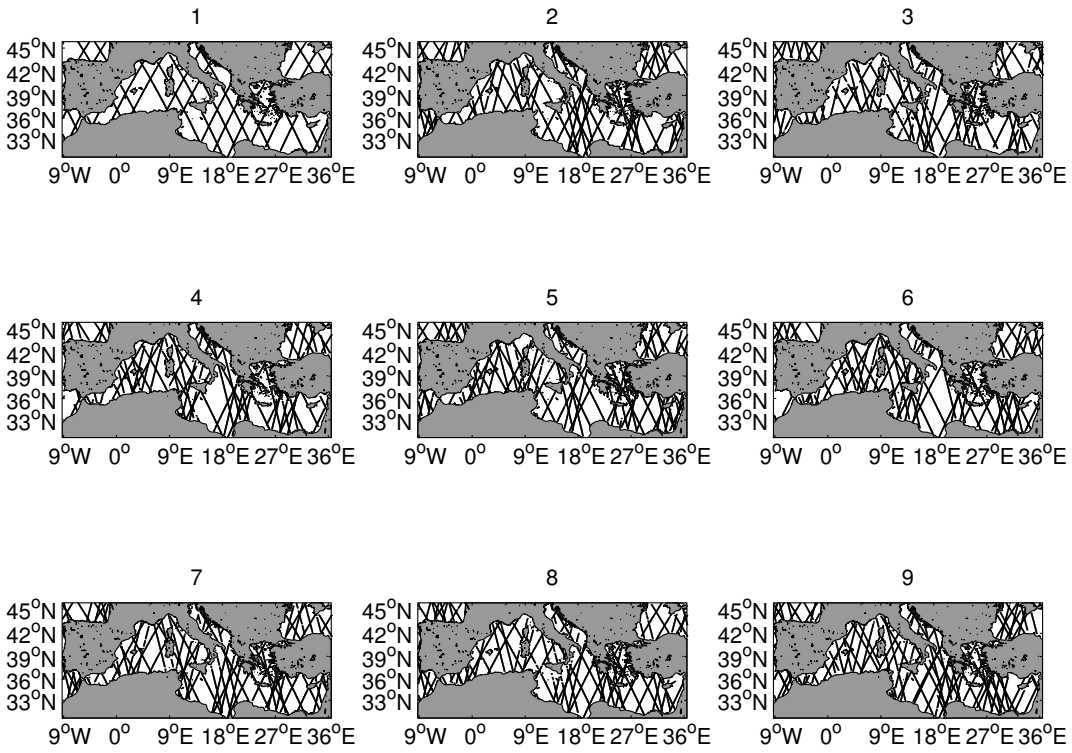

Figure 4.1: Coverage attained from the combination of TOPEX/POSEIDON+Ers-2+GFO satellites at each of the experiment weeks.

by the detcection of outliers and by verifying the variability error. After these tests, the tracks are interpolated along the ground tracks to fixed reference points. Then the tides are removed from the altimeter data, although near the coasts no correction is made due to the coarse resolution of tidal models. Two lasts correctaions stand for orbit error removal (which consists in a long-wave component removal) and the referencing of the data to a consistent mean (in this case, the mean was calculated over 5 years of TOPEX/POSEIDON data).

SST and SLA are both assimilated into the model. Due to the lack of available high resolution independent data, they are also used to establish the skill of the model. However, to minimise the artificial skill generated by this fact, the comparison between model and observations is made before they are assimilated into the model.

Daily Sea Surface Temperature data

AVHRR Pathfinder v5SST data are also used in the verification of the model. They consist on daily averaged global SST maps. A subset ranging from 3 July to 3 September 2000 and covering the Mediterranean Sea is used. Only descending (nighttime) passes are selected, to minimise skin temperature effects. They are derived from the 5-channel AVHRR on board the NOAA-7, 9, 11, 14, 16 and 17 polar orbiting satellites. Together with the SST data, quality flags files can also be obtained. These files classify the SST data into seven quality levels, ranging from 0 (the worst) to 7 (the best). Clouds are identified from these quality flags, so that each user can decide the mask that will be applied to the data. In this work, quality flags under 5 were classified as clouds or missing data. Figures 4.2 and 4.3 show the SST for the 8 July and the quality flags for this image, respectively. The process of giving each pixel a quality flag passes by eight tests. These tests and how they are related to each quality flag are summarised in Appendix A. The average accuracy of the Pathfinder v5 nighttime data has been found to be 0.1±0.33°C (Kearns *et al.*, 2000). For further information see Kilpatrick *et al.* (2001). Pathfinder data are available via anonymous ftp in the Jet Propulsion Laboratory web site (`ftp://podaac.jpl.nasa.gov`), or through the POET interface at `http://podaac.jpl.nasa.gov/poet`. The high resolution of these data and their high quality make them very useful for verification purposes.

4.2 *In situ* data

In situ data is also used, altough they are not assimilated into the model. These data are then used as independent data, which is necessary to establish the skill of the model without any ambiguity.

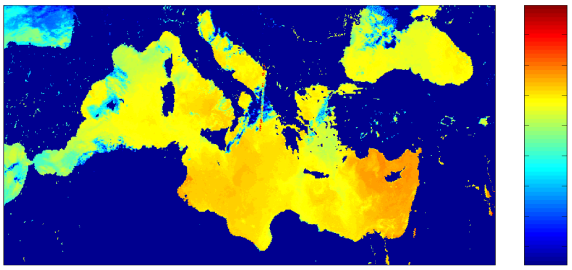

Figure 4.2: Rough Pathfinder SST for 8 July

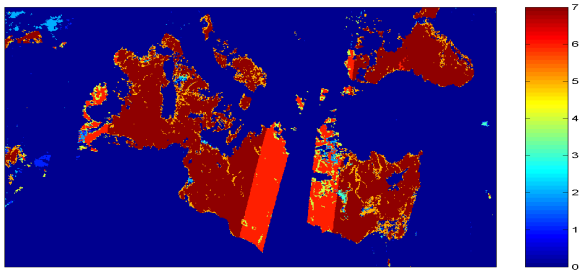

Figure 4.3: Quality flags provided to 8 July.

Salinity and temperature are obtained from 51 CTD (Conductivity, Temperature and Depth) profiles from the Sirena 2000 cruise of the SACLANT Centre (La Spezia, Italy). They cover the period from 22 to 30 August 2000, and their distribution can be seen in Figure 4.4. Temperature and salinity of the Liguran Sea model will be compared with Sirena profiles. In addition, data from the MEDAR/Medatlas database (MEDAR-Group, 2002) are used to validate the model. Data from the whole Mediterranean Sea is used, but only specific zones have a good coverage (*e.g.* ship tracks, permanent stations, etc). Their temporal coverage goes from 17 July to 28 September 2000, but they are irregularly spaced in time. The distribution of these data is shown in Figure 4.5. Most of the MEDAR/Medatlas data obtained come from XBT's, so only temperature can be studied.

4.3 Bias between the observations

The data used in this work come from different sources (different measurements, as Sirena and MEDAR/Medatlas data, or the same measurements with different post-treatment, as DLR and Pathfinder v5 SST data). There are, logically, some biases between those data. This has been measured and included here, as any posterior error analysis of the model with these data will give different results in function of the data used.

We compare all data with DLR, as they are assimilated into the model, thus they affect the model results. The difference between DLR SST and Sirena data at 1 m depth is about -0.6°. Sirena is warmer than DLR SST, which is an unexpected result, as the DLR is a daily maximum temperature composite of skin temperature and Sirena data do no resolve the skin layer temperature, and is taken at various times of the day. This bias can be seen in Figure 4.6.

The difference between DLR SST and MEDAR/Medatlas data is smaller, due to the fact that this difference is taken over the whole Mediterranean, and Mediterranean DLR

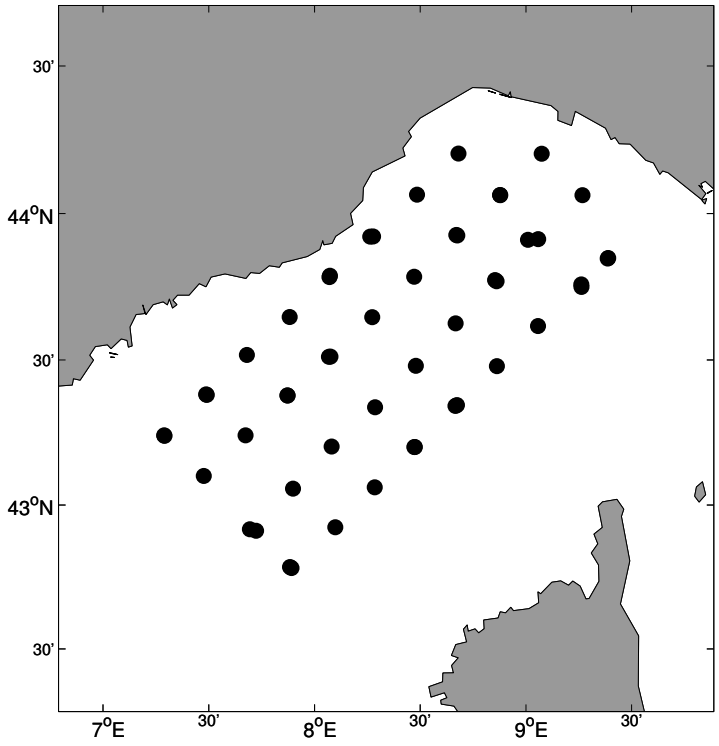

Figure 4.4: Distribution of the Sirena 2000 stations.

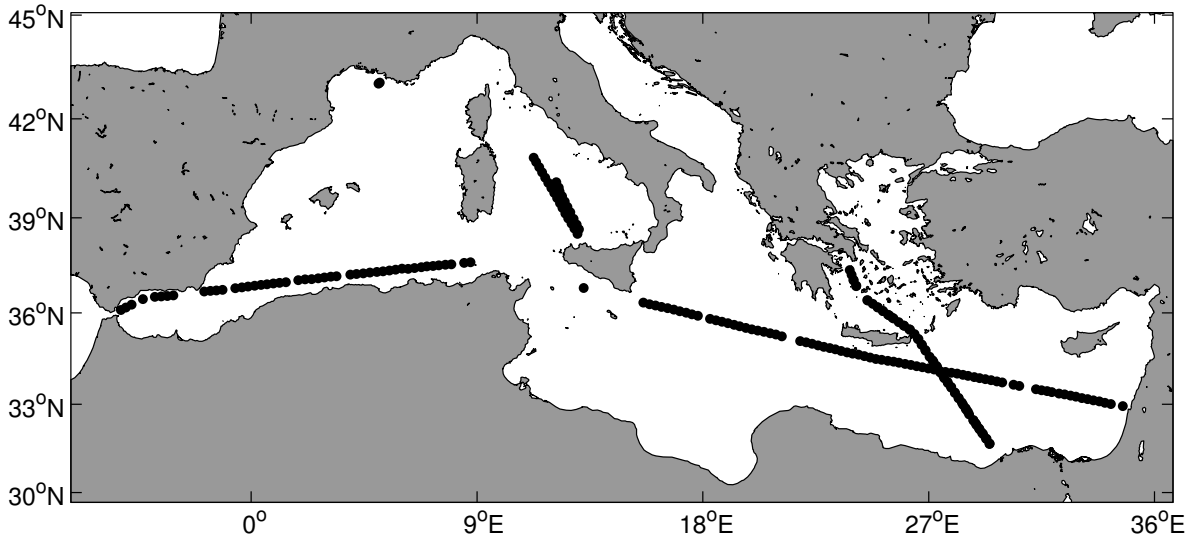

Figure 4.5: Distribution of the MEDAR/Medatlas stations used in this work.

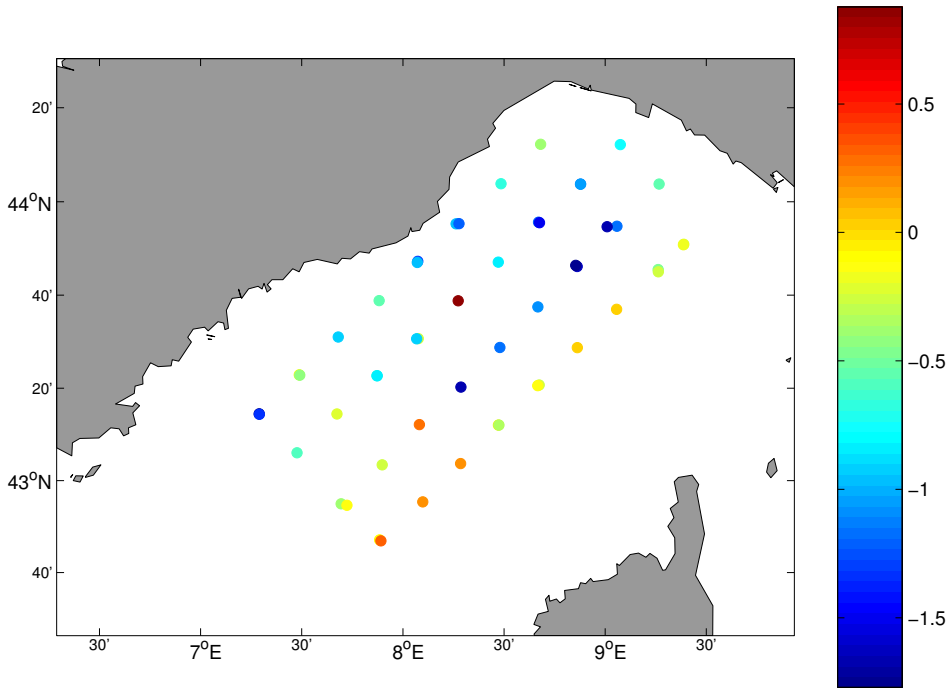

Figure 4.6: Difference between DLR SST and Sirena surface values.

data have a coarser resolution. The mean bias between DLR and MEDAR/Medatlas is 0.13°C, so again MEDAR/Medatlas is warmer than DLR SST. This results are shown in Figure 4.7.

Finally, DLR SST has been compared to Pathfinder v5 SST data. The difference is −0.5°C, being Pathfinder warmer than DLR. This is again an unexpected result, as we use nighttime Pathfinder data, and they are compared to the maximum daily temperature of DLR. However, the treatment of each of those data make that the results are different. The diference between these two data sets is averaged over the space, and the time evolution of the Pathfinder v5 and DLR temperatures is shown in Figure 4.8.

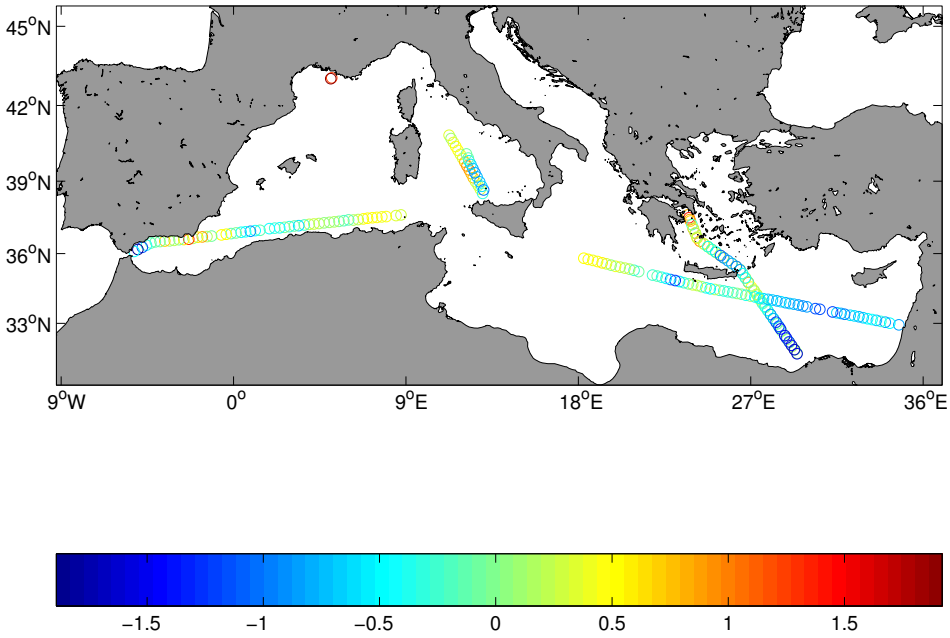

Figure 4.7: Difference between DLR SST and MEDAR/Medatlas surface values.

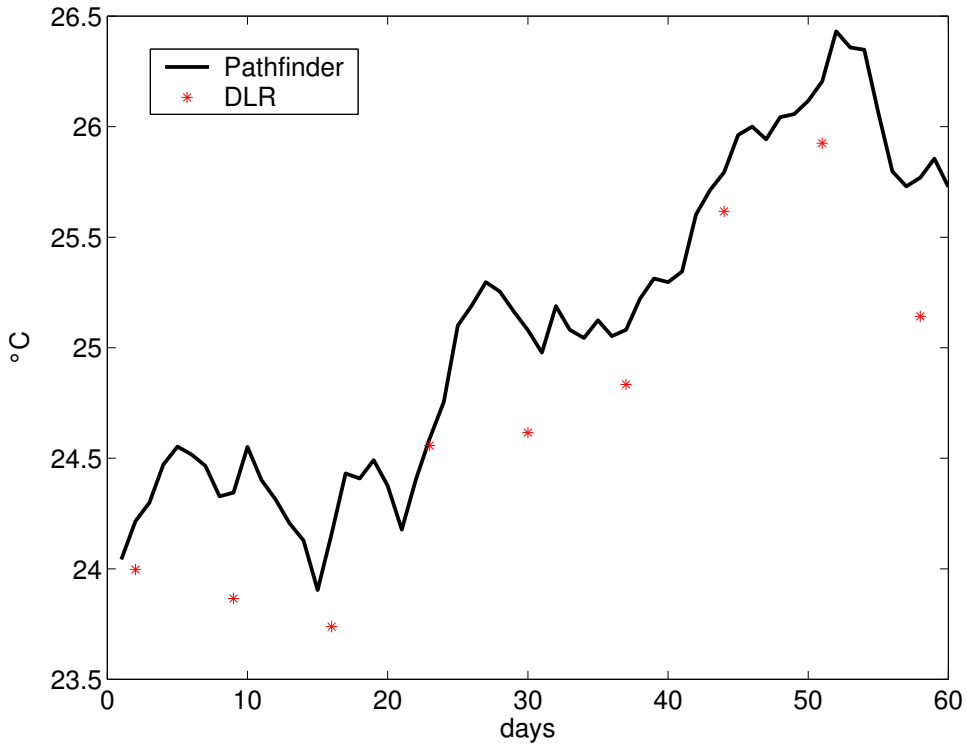

Figure 4.8: Difference between DLR SST and MEDAR/Medatlas surface values.

Part II

Methodology

Chapter 5

Error Measures in Forecast Verification

5.1 Introduction

Definition of Forecast Verification
Forecast Verification was defined by Murphy and Winkler (1987) as "the process and practice of determining the quality of forecast". The definition by Jolliffe and Stephenson (2003) is more explicit: "forecast verification is the exploration and assessment of the quality of a forecasting system based on a sample or samples of previous forecasts and corresponding observations". These concepts must be further defined, in order to explain the meaning of "quality". The quality of a forecast is however hard to establish, as it depends on a variety of factors (the initial configuration of the forecast system, the end user expectations, the interpretation of the results...).

Following Murphy and Winkler (1987), there are two major approaches to the verification of a forecast: the *measures-oriented* approach, where the whole system of forecasts and observations is reduced to a single measure (or limited set of measures); and the *distribution-oriented* approach, that uses the joint distribution of forecasts and observations. This second approach can be very useful when there are several forecasts to analyse, but the dimensionality of the problem can be very high. With the first approach, the analysis is reduced to a single number, or a small set of numbers, but some information is lost (Brooks and Doswell, 1996). Although it should be desirable to analyse a forecast from the second point of view, the dimensionality of the problem can limit the choice, and usually the first option is chosen.

Why doing Forecast Verification
Forecast verification is an important part of the forecasting process. It helps, in an early step, to detect the weaknesses of the forecasting system and to make the necessary improvements (Wilks, 1995; Buizza, 2001), a process also called "diagnostic verification" by Murphy and Epstein (1989) and Murphy and Winkler (1992). Once the best

forecast that can be produced by the model is released, the verification gives important information to the end user, that will decide how to use the forecast results based on this quality assessment. Moreover, a good quality assessment leads to better decision making. By means of the verification procedure we aim to identify missing physical processes, over and under estimation of a given variable, biased behaviours, punctual sources of error, etc. In a more general framework, a good quality assessment is also necessary in the ongoing monitoring of operational forecast, where in addition the rapidity in the detection of problems is a key factor.

How to choose the verification procedure
There are several ways to realise the verification of a forecast, but all are thought in terms of accuracy and skill of the forecast when compared with the corresponding observations. Indeed, forecast accuracy is based on the comparison between the forecast and the corresponding observations, and forecast skill is based on the comparison between the forecast and a reference system (Wilks, 1995; Jolliffe and Stephenson, 2003). These two terms will be explained later in more detail. The process of choosing the verification forecast procedure is very important, as the information obtained from the model results will depend on the measures used. The verification process should vary depending on the kind of forecast, the variables to be verified, the reference forecast, the existing observations, etc. Verification can be defined as the basis of a program of forecast improvement (Brooks and Doswell, 1996).

An important number of verification techniques, including different combinations of skill values, graphics, etc... have been proposed (*e.g.* Taylor, 2001). The verification method applied to a given forecast must be carefully chosen, and the possibilities of combining several methods, or the creation of new diagrams depends on each case. In this section, the basic tools needed for forecast verification are included, those who allow the study of the main characteristics of the forecast-observation pair. For a further study, these basic tools may be applied to understand and improve the forecast weaknesses. First, a brief definition of the different existing types of forecast is presented, followed by the description of the basic concepts of accuracy and skill. Finally, the error measures used in the verification of quantitative forecasts is presented, as it is the type of forecast we deal with in this work.

5.2 Types of forecasts

- **Binary forecast:** a forecast in which the occurrence or non occurrence of an event is predicted. Oonly one event can happen, e.g., the occurrence of an eddy at a particular place in a given moment. It is also called categorical forecast. In Meteorology, this kind of forecast is the most used when forecasting rain events. This kind of forecast contains no expression of uncertainty. The categorical forecast

can be applied to discrete or continuous variables.

- **Probabilistic forecast:** a forecast given in terms of the probability that a considered event would or would not happen. Used, for example, in ensemble forecast verifications, where a probability distribution function (pdf) can be computed.

- **Quantitative forecast:** a given number of the variable predicted (e.g. temperature, 15 °C), without specifying a measure of probability or uncertainty (von Storch and Zwiers, 1999). Some authors (*e.g. Wilks, 1995)*, define this kind of forecast as a categorical forecast, because they do not include any probability or expression of uncertainty as part of the forecast format.

5.3 Verification tools

The verification system becomes difficult because of the high complexity of the forecasts. It is necessary to develop some measures to establish the forecast performance in an easy way, and usually this is achieved with some scalar verification measures. They allow to obtain useful information about the forecast quality, as the bias or the accuracy, although some information about the joint behaviour of forecast and observations is lost (Brooks and Doswell, 1996). To calculate the forecast skill, one or several of this scalar measures are used. The most common one is the accuracy, compared with the accuracy of a reference forecast (Wilks, 1995; Jolliffe and Stephenson, 2003).

5.3.1 Scalar Statistics

Bias: it indicates the degree of correspondence between the average forecast and the average observations (Wilks, 1995). The bias indicates, on average, if the forecast are under-predicted (bias>0, the forecast value is too low) or over-predicted (bias<0), the forecast value is too high, (Cook, 1999). The bias does not depend on the individual correspondence between the forecast and the observations, so it is not a measure of accuracy. We need an error measure that allows to say if the forecast matches exactly the observations, and if this is not the case, how far is the forecast from the observations. This is achieved by an accuracy measure.

Accuracy: the accuracy can be defined as the average degree of correspondence between forecasts and observations, (Wilks, 1995). Scalar measures of accuracy summarize the overall quality of a forecast. The Mean Averaged Error (MAE), Mean Square Error (MSE) or the Root Mean Square Error (RMSE), which will be introduced later in this chapter, are accuracy measures.

Correlation: is a measure of the relationship between the forecast and the observations. It measures if two sets of data change together. Values close to 1 are the best, and no correlation occurs when values in both sets are unrelated. Positive correlation indicates that large forecast values are associated with large observed values. However, high correlation does not necessarily denote high accuracy.

Standard deviation: it measures the dispersion from the mean of a given parameter of the forecast. This measure does not consider the observations. A large SD indicates large dispersion from the mean. In forecast verification, the standard deviation is used to measure the extent that the forecast error differs from location to location from the mean, (Cook, 1999).

Reliability (or conditional bias): it represents the degree to which the correspondence between the mean forecast and the observations, given the forecast, varies with the forecast. In a probabilistic sense, reliability refers to the correspondence between the forecast probability of an event and the relative frequency of the event, conditional upon the forecast probability (Mason, 1982).

Resolution: this measure refers to the degree to which the forecast sort the observed events into groups that are different from each other. Both reliability and resolution are concerned with the properties of the conditional distributions of the observations given the forecast. Resolution refers to the differences between the conditional averages of the observations for different values of the forecast, whereas reliability compares the conditional averages of the observations with the forecast values themselves, (Wilks, 1995).

Discrimination: the opposite of resolution. It summarizes the conditional distributions of the forecast given the observations. The discrimination attribute reflects the ability of the forecasting system to produce different forecast for those occasions having different realized outcomes of the predicted variable, (Wilks, 1995).

Sharpness: is an attribute of the forecasts alone, without regard to their corresponding observations. Measures of sharpness characterise the unconditional distribution (relative frequencies of use) of the forecast (Wilks, 1995).

In table 5.1 these scalar measures are summarized. For those that are most used in this work, some examples are given.

5.3.2 Skill Scores

The skill is a relative quality measure of the forecast of interest compared to that of a reference forecast. The reference forecast can be a simple strategy as recurrence, persistence, climatology or random guessing. The output of a forecast model or an ancient version of the studied forecast can be also used as a reference. In the latter case, the use of an

Table 5.1: Forecast attributes

Attribute	Definition	Related measures
Bias	Correspondence between mean forecast and mean observation	bias
Association	Strength of linear relationship between pairs of forecasts and observations	covariance, correlation
Accuracy	Average correspondence between individual pairs of of observations and forecasts	MAE, MSE, RMSE
Skill	Accuracy of forecasts relative to accuracy of forecasts produced by a standard method	MSEE or RV
Correlation		ACC, Correlation
Standard Deviation		SD
Reliability	Correspondence of conditional mean observation and conditioning forecast, averaged over all forecasts	-
Resolution	Difference between conditional mean observation and unconditional mean observation, averaged over all forecasts	-
Sharpness	Variability of forecasts as described by distribution of forecasts	-
Discrimination	Difference between conditional mean forecasts and unconditional mean forecasts, averaged over all observations	-
Uncertainty	Variability of observations as described by the distribution of observations	-

ancient version of the forecast allows to know if the new version ameliorates the old one, and in what extent (Wilks, 1995; Brooks and Doswell, 1996; von Storch and Zwiers, 1999).

Skill Score: The general expression of a skill score is:

$$skill\ score = \frac{score_{forecast} - score_{reference}}{score_{perfect} - score_{reference}} \cdot 100$$

The terms $score_{forecast}$ and $score_{reference}$ refer to a given measure of accuracy (*e.g.* MSE) of the forecast and the reference system, respectively. $Score_{perfect}$ is the accuracy measure that would be achieved by perfect forecasts.

A skill score measures the percentage of improvement in accuracy over a reference forecast. A skill score has a value of 0 for a strategy that issues forecasts with accuracy equal to that of the reference forecasts. Positive scores indicate that the forecast accuracy is an improvement over that of the reference forecasts. Depending on the type of studied forecast, different verification procedures may be applied.

5.4 Error Measures for Quantitative Forecasts:

The difference between categorical and quantitative forecasts is that in the second one there is no threshold value that identifies the occurrence or non-occurrence of an event. In this work, we deal with quantitative field forecasts (also called 'categorical filed forecasts' by Wilks (1995)), as the forecast output is presented on a gridded field. Because of the multiciplity problem (Wilks, 1995; Jolliffe and Stephenson, 2003) already mentioned, spatial averages of the gridded forecast are used. With this approach, the evolution of the error over all spatial points with time is studied. For the study of the more complex spatial features, other methods (wavelets, Empirical Orthogonal Functions) will be used. The verification methods are here applied to the N points of the grid, and the forecast is perfect if it is equal to observations at each point of the grid, $x_m^f = x_m^o$.

The verification of field forecasts is a problem of very high dimensionality, even for small grids, because two gridded fields (observed and forecasted) can differ in many ways. For this, although the joint distribution of forecasts and observations is preferred to analyse the forecasts, usually the correspondence between forecasts and observations is studied using scalar measures.

The following list contains the most common measures used to analyse quantitative field forecasts in the spatial domain, both scalar measures and skill scores. The nomenclature used in all this work is the same: f is for forecast, o for observations and N

is the total number of values used, number of points in a grid or number of time steps, depending if the error measure is applied in time or in space. Of course, all the measures defined here can be applied with a time averaged forecast (and to obtain the distribution of the error in space), just doing the summation and averages over time, rather than over space.

5.4.1 Bias

This concept has been already explained (5.3.1). Two types of bias can be identified: unconditional or additive bias, and multiplicative bias. The expression for the unconditional bias is:

$$\text{Bias}(x) = \frac{1}{N}\sum_{n=1}^{N}\left(x_n^f - x_n^o\right) \tag{5.1}$$

Values near zero are the best. The expression for the multiplicative bias:

$$\text{MB}(x) = \frac{\sum_{n=1}^{N} x^f}{\sum_{n=1}^{N} x^o} \tag{5.2}$$

where the best value is the unity. The unconditional bias is widely used, as it is the best way to detect drifts or permanent errors in the model forecast. However, the only measure of bias does not give a complete idea of the forecast error. A forecaster making a temperature forecast 10 °C too warm and a forecast 10 °C too cold will get the same bias as a forecaster making two forecasts that match the observations exactly. To correct this problem the errors need to be nonnegative, and this will give an idea of the accuracy of the forecast. For this, additional measures should be used to assess the skill of the model.

5.4.2 Accuracy measures

Several measures of accuracy are used in forecast verification. These measures do not present the same problem of cancelling negative errors with positive, as bias does, so they can add valuable information about the forecast skill.

Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) expression is:

$$\text{MAE}(x) = \frac{1}{N}\sum_{n=1}^{N}\left|x_n^f - x_n^o\right| \tag{5.3}$$

The MAE is less expensive to compute than other accuracy measures as Mean Squared Error (MSE), because of the absence of multiplication. However, the MSE is often preferred in practice.

Mean Square Error (MSE)

The expression for MSE is:

$$\mathrm{MSE}(x) = \frac{1}{N}\sum_{n=1}^{N}\left(x_n^f - x_n^o\right)^2 \tag{5.4}$$

The MSE is more sensitive to outlier errors than is MAE (Jolliffe and Stephenson, 2003) due to the square. The MSE is very sensitive to systematic errors. It favours forecasting schemes that avoid extremes and tend not to deviate greatly from the observations, because the penalty grows as the square of the error (Wilks, 1995).

The MSE can be decomposed into a bias term and a centred MSE term. If we add and substract the forecast and observed spatial means at each grid point, Eq. 5.4 can be expressed this way:

$$\mathrm{MSE} = \frac{1}{N}\sum_{n=1}^{N}\left[(x_n^f - \bar{x^f}) - (x_n^o - \bar{x^o}) + (\bar{x^f} - \bar{x^o})\right] \tag{5.5}$$

By arranging the terms of Eq. 5.5, we obtain:

$$\mathrm{MSE} = \left[\frac{1}{N}\sum_{n=1}^{N}\left(x_n^f - x_n^o\right)\right]^2 + \frac{1}{N}\sum_{n=1}^{N} x_n^f - x_n^o + (\bar{x^o} - \bar{x^f}) \tag{5.6}$$

The first term of the right-hand side can be identified as the square of the bias, and the second term can be called centred MSE (MSE'), so the MSE is composed by:

$$\mathrm{MSE} = \mathrm{bias}^2 + \mathrm{MSE'} \tag{5.7}$$

The advantage of decomposing the MSE this way is that it allows to assess the error without the effect of bias. In long-term model runs, where important systematic bias can develop, it is interesting to separate this two terms. The elimination of bias from the MSE signal can ameliorate the model skill if the bias drift between the model and forecast is high.

In addition, the decomposition of MSE has been used in the interpretation of the error measures presented in this section, as it allows to relate some of them. It is also possible to construct a graphic based on this decomposition that helps to improve the knowledge of the model. This two points will be further developed in this chapter.

Root Mean Square Error

The square root of the MSE, Root Mean Square Error (RMSE) is easier to interpret as it has the same unities than the forecast. The expression for RMSE is:

$$\mathrm{RMSE}(x) = \left[\frac{1}{N}\sum_{n=1}^{N}\left(x_n^f - x_n^o\right)^2\right]^{1/2} \tag{5.8}$$

Lower values of accuracy show in general best performance of the forecast (e.g., for MAE, MSE and RMSE, a value of zero denotes a perfect forecast). The RMSE is the widest used error measure, because of the conservation of unities. However, it should be kept in mind that the use of RMSE does not give a complete idea of the causes of the error, so it should be used with other complementary error measures.

5.4.3 Mean Square Error Skill Score (MSESS)

From the accuracy MSE measure, a skill score can be computed. It averages the individual squared differences between the observed and forecast fields at each grid point. The expression for the Mean Square Error Skill Score (MSESS) is based on the MSE (Eq. 5.4):

$$\mathrm{MSESS} = \frac{MSE_{forecast} - MSE_{reference}}{MSE_{perfect} - MSE_{reference}} \tag{5.9}$$

Considering that $MSE_{perfect}$ is zero, the expression for MSESS can be finally expressed as:

$$\mathrm{MSESS} = 1 - \frac{MSE_{forecast}}{MSE_{reference}} \tag{5.10}$$

A MSESS value of 1 indicates a perfect forecast. A value of zero denotes no improvement over the reference system. A negative value implies that the model forecast is worse than the reference system, although it does not indicate that the model has no skill at all. A skill score is always defined with regard to a reference system. The MSESS can be also thought in terms of a percentage score, by multiplying Eq. 5.10 by 100. We can thus talk about the improvement percentage of the model.

We can now go back to the decomposition of the MSE. The MSE can be further decomposed to obtain an expresion that relates it to the correlation. Murphy (1988) decomposed the MSE for the time averaged case. Murphy and Epstein (1989) made the decoposition for the space averaged MSE. We are refering to this late one now. Developing the square on Equation 5.6 and re-arranging the terms we have:

$$\text{MSE} = \text{bias}^2 + S_f^2 - S_o^2 + 2rS_fS_o \tag{5.11}$$

where:

$$S_f^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n^f - \bar{x^f})^2 \qquad ; \qquad S_o^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n^o - \bar{x^o})^2 \tag{5.12}$$

are the forecast anomaly field sample variance and the observed anomaly field sample variance respectively. The last term of Eq. 5.11 is twice the covariance between observed and forecast anomaly fields multiplied by a coefficient of correlation between the anomalies in the forecast and observed fields, r. If the anomaly is calculated over the past observations, this correlation is called Anomaly Correlation Coefficient, and it will be explained in the next section.

Murphy and Epstein (1989) applied the decomposition of the MSE to the MSESS, using as reference system a climatology. Substituting Eq. 5.11 on Eq. 5.10 we arrive to the following expression:

$$\text{MSESS} = \frac{\{r^2 - [r - (S_f/S_o)]^2 - [\text{bias}/S_o]^2 + (o'/S_o)^2\}}{[1 + (o'/S_o)^2]} \tag{5.13}$$

We can identify the following terms in this equation:

1. r^2: the square of the correlation.

2. $[r - (S_f/S_o)]^2$: a nondimensional measure of the conditional bias in the forecast anomalies. This term can be also interpreted as an indicator of the amount of forecast variability ($S_f^2 = S_o^2 r^2$) necessary to eliminate this term (*i.e.* to eliminate the conditional bias).

3. $[bias/S_o]^2$: a non dimensional measure of the unconditional (or overall) bias in the forecast anomalies.

4. $(o'/S_o)^2$: is the only term independent from the forecast. This term represents the square of the coefficient of variation of the anomalies in the observations, and should be very small compared to unity, unless the analysis is done in a small area.

The division of the MSESS in these four terms allows to interpret the skill score jointly with other error measures, as the correlation or the bias. For example, the correlation is the skill if both the conditional and unconditional bias (terms (b) and (c) in the previous list) could be eliminated, with the slight correction made by term (d). For further explanations about the decomposition of the MSESS, see Murphy and

Epstein (1989) and von Storch and Zwiers (1999). Note that the choice of the climatology or mean field substracted to compute the anomalies will influence the obtained results.

There are many skill assessment studies based on the MSE decomposition, as the Taylor diagram (Taylor, 2001, explained later in this chapter), the LEPS score, (Potts *et al.*, 1996), etc. Murphy (1996) made an additional decomposition of MSE-based skill scores, where the verification is made by conditioning on the forecast or in the observations. This approach allows to obtain different attributes of the forecast skill starting with a MSESS.

5.4.4 Anomaly Correlation Coefficient (ACC)

The ACC is largely used to evaluate large-scale / medium-range forecasts skill. It provides a reliable indication of the overall model skill. The ACC measures the relationship between the observed and forecast deviations from the climatological field. To compute the ACC, the observed and forecast values are first transformed into anomalies. The climatological average value of the observed field (at each grid point) is thus subtracted. The expression for the ACC is:

$$\mathrm{ACC} = \frac{\sum_{m=1}^{M}\left[(x_m^f - C_m)(x_m^o - C_m)\right]}{\left[\sum_{m=1}^{M}(x_m^f - C_m)^2 \sum_{m=1}^{M}(x_m^o - C_m)^2\right]^{1/2}} \tag{5.14}$$

where C_m is the climatological value of the observed variable x_m^o at the point n, given by the average over m prior observations of the field

$$C_m = \frac{1}{n}\sum_{k=1}^{n} x_m^o(k) \tag{5.15}$$

The ACC is designed to detect similarities in the patterns of deviations from the climatological field. Those values of forecast and observations that are simultaneously above or below the climatological value will sum to the ACC. The ACC for a perfect forecast has a value of 1, being -1 the minimum possible value. Another characteristic of the ACC is that it is not sensitive to bias in the forecast. The ACC corresponds to the correlation between the forecast and observed anomalies only if the averages over the M grid points of the two anomalies is zero, that is, only if $\sum_m(x_m^f - C_m) = 0$ and $\sum_m(x_m^o - C_m) = 0$, (Miller *et al.*, 1995; Wilks, 1995).

5.4.5 Correlation Skill Score

Correlation between the forecast x^f and observations x^o. It measures if two data sets change together. The correlation skill score is insensitive to some types of systematic error: it is not affected if the forecast has a constant bias, or if the amplitude of two differ by a constant factor.

The expression for the Correlation Skill Score is:

$$\rho_{FO} = \frac{Cov(\mathbf{x^f}, \mathbf{x^f})}{\sqrt{Var(\mathbf{x^f})(\mathbf{x^o})}} \tag{5.16}$$

Values close to unity indicate best skill, but high correlation does not necessarily denote high accuracy. Values of zero or less than zero denote random forecasts, (von Storch and Zwiers, 1999).

5.4.6 Proportion of Explained Variance

Proportion of the variable variance that is explained by the forecast (von Storch and Zwiers, 1999). Its expression is:

$$R^2_{FO} = \frac{Var(x^o) - Var(x^f - x^o)}{Var(x^o)} = 1 - \frac{Var(x^f - x^o)}{Var(x^o)} \tag{5.17}$$

5.4.7 Standard deviation

The standard deviation (SD) can be used to see if the variability of the forecast is similar to the observed one. The SD is calculated separately for the forecast and the observations, and the results are compared. The SD is easier to interpret than the variability, because the units of the SD match those of the forecast. The expression for the SD is:

$$SD = \left[\frac{1}{N-1} \sum_{n=1}^{N} (x_n - \bar{x})^2 \right]^{1/2} \tag{5.18}$$

x_n and $\bar{x}$ are the sample value and mean value of the forecast. A large SD indicates large dispersion from the mean. In terms of error measure, the standard deviation is used to measure the extent that the forecast error differs from location to location from

the mean, (Cook, 1999).

5.5 Taylor diagram

The verification of a model does not simply consists in a collection of error measures telling us about its perfomances. These measures should be combined and interpreted together with the aim of obtaining the maximum information possible to understand the origin of errors. When working with a limited number of model results this combination can be done easily. However, in the frame of a multi-scale model decomposition, as it is done in this work, the model is decomposed into several spatial scales, and at each scale we can apply the error measure described above. The joint interpretation of these results can be very tricky. We can think on two solutions: the computation of a single measure giving us the more convenable information, or the representation of some error measures in a single diagram. The first option is not desirable, insofar as there is no single measure that brings together all the information contained in a complex model.

The second option is more convenable, but the representation of several measures on a single graphic must be useful and easy to interpret (otherwise there is no advantage on using it). In this section, we will describe the Taylor diagram (Taylor, 2001), as it has been very useful in our work. There are many works that have used this diagram to analyse complex systems with many variables, as Denis *et al.* (2003) for a climate model over the north Atlantic Ocean and Raick *et al.* (2004) for an biochemical model on the Ligurian Sea.

Taylor diagram is based on the MSE decomposition presented in section 5.4.2. If we express Eq. 5.7 as a RMS decomposition rather than MSE decomposition:

$$\mathrm{RMS}^2 = \mathrm{bias}^2 + \mathrm{RMS'}^2 \tag{5.19}$$

then the centred RMS (RMS') can be expressed as:

$$\mathrm{RMS'}^2 = S_f^2 + S_o^2 - 2rS_fS_o \tag{5.20}$$

Note the relation between Eq. 5.20 and the law of the cosines:

$$c^2 = a^2 + b^2 - 2ab\cos\phi \tag{5.21}$$

where a, b and c are the sides of a triangle and ϕ is the angle opposite to c. With this relation, we can easily plot into one diagram several statistics of a field. We are interested in the comparison between the model and the observations, more exactly how the model resembles the observed data. We can thus plot into the same graphic their standard deviations, the correlation of the model and the RMS'. The graphic is presented in Figure 5.1, and the mentioned statistics are ploted as follows:

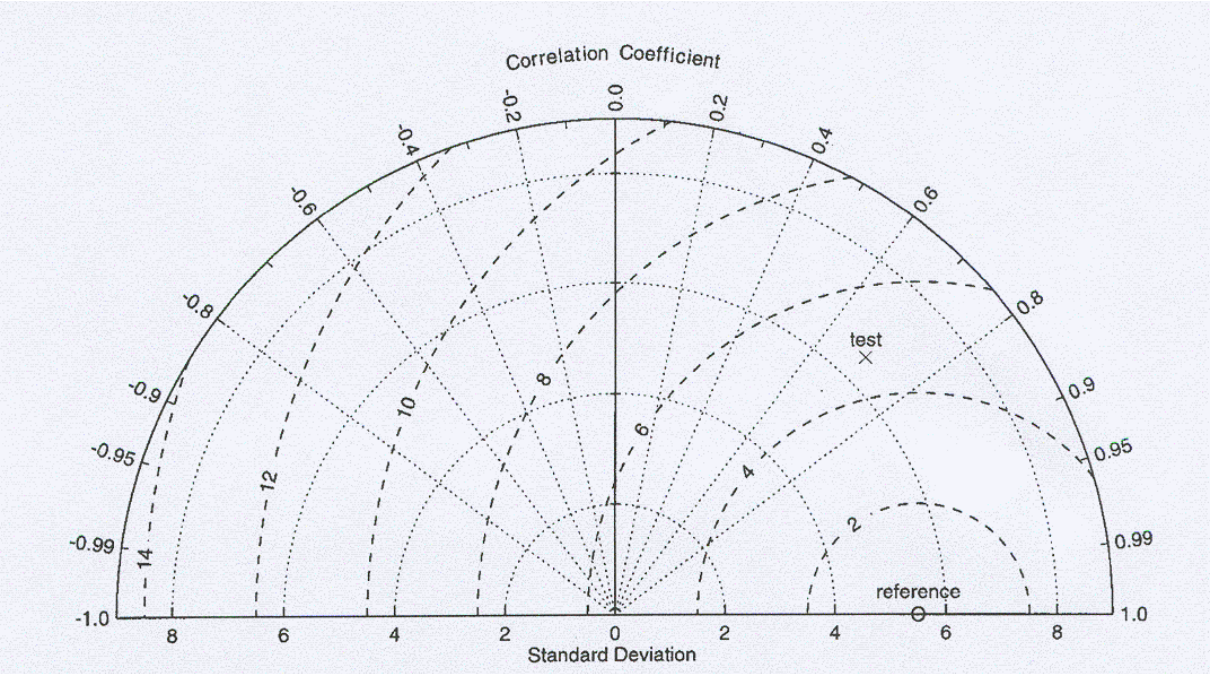

Figure 5.1: Taylor diagram, obtained from Taylor (2001). The standard deviation is represented by the distance to the origin of the graphic. The correlation is given by the angle, and the RMS' by the distance between the test and reference points on the graphic.

- The standard deviation is plotted as the radial distance to the graphic origin.
- The correlation between both fields is given by the azimutal position.
- The RMS' is the linear distance between the observations and the model points (represented by a dashed line in Figure 5.1.

The standard deviation and the RMS' to be represented into the Taylor diagram can be normalised by the standard deviation of the observations ($\hat{RMS}' = RMS'/S_o$; $\hat{S}_f = S_f/S_o$ and $\hat{S}_o = 1$), so that the observations lay in the abscissa axis, with a normalised standard deviation of 1. The other variables on the plot are thus referred to the observations. The correlation is not affected by this normalisation.

A point represented this way into the graphic will thus have the same standard deviation as the observations if it is situated at a distance of one from the origin. The standard deviation will be higher if this distance is higher, and the opposite if the distance is smaller than one. The bigger the angle separating the model and the observations, the worse its correlation. With this representation we can have a condensed information about the model behaviour compared to the observations.

Chapter 6

Mixed Layer Depth

The mixed layer is a zone with a high variability. The ocean surface is affected by the atmospheric conditions, and this information is transmitted through the water column to a depth that varies from season to season. The position and width of the mixed layer depth (MLD) varies thus depending on the season, and even day/night variations affect it. The position of the mixed layer is a key factor for the surface ocean biology. It is thus important to have a good representation of the MLD in model results. A good knowledge of the MLD is necessary for many studies in the ocean, for example, for a good estimation of the sea level by the thermal structure of the water column (*e.g.* Bouzinac *et al.*, 2003). Many works assess the problem of a good methodology for the detcetion of the mixed layer position (*e.g.* Soetaert *et al.*, 2001; Thomson and Fine, 2003).

The comparison of the model output with the Sirena 2000 Cruise and the MEDAR/Medatlas database has been made in two ways. First, station by station, merging the profiles to obtain a single number for comparison between model and observations. Then an analysis with depth is made. The structure of the water column is compared to the observed profiles averaging over the stations of a same geographical zone. The Sirena profiles, as they are concentrated in the Ligurian Sea, can be averaged together. However, an station-by-station study is also done in order to depict the differences in vertical structure of the water column that may exist in the Ligurian Sea. The distribution of the Sirena and MEDAR/Medatlas database can be seen in Figures 4.4 and 4.5. Figure 6.1 shows the maximum depth attained by Sirena profiles.

When comparing the first time the model to the Sirena cruise, the RMS error appear to be smaller than 2°C when comparing the first 150 m. However, when looking to the vertical structure of the profile, it appeared that at the thermocline the error was very big (bigger than 4°C). The forecast was even doing worse than the free model run. The assimilation of SST and SLA perturbs the field at depth, so that the structure of the water column does not match with what is observed.

In an attempt to see what happens to the water column structure, a study of the

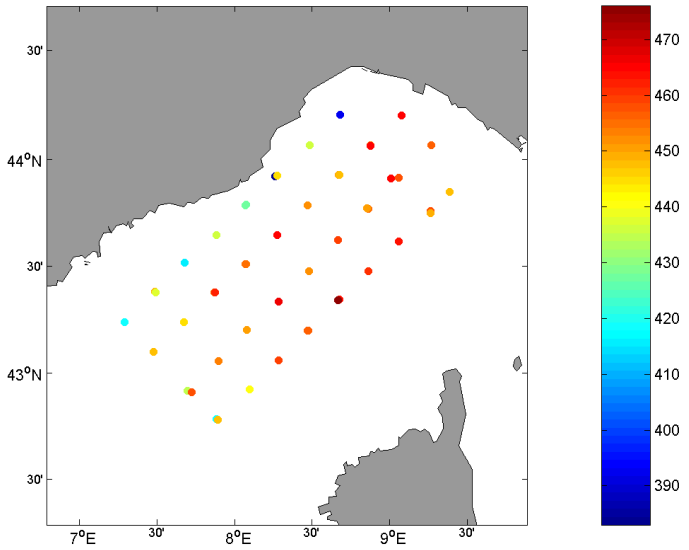

Figure 6.1: Distribution of the Sirena 2000 stations, with the maximum depth attained by each of the profiles.

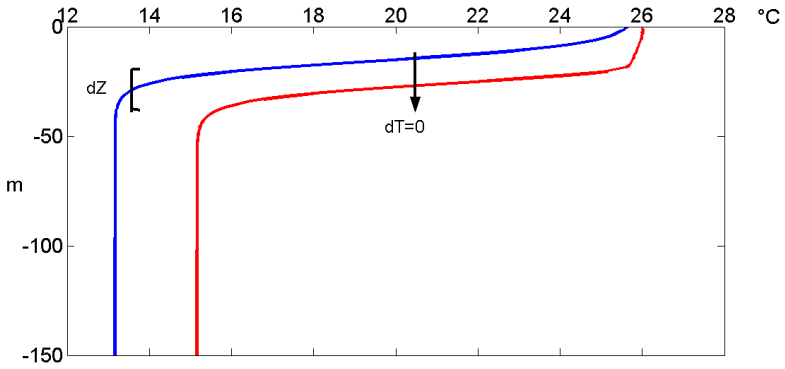

Figure 6.2: The notion of double penalty when assessing the error of a profile is shown. Both curves show a similar temperature at the thermocline, but the depth is different. The error in temperature is thus very high when comparing both profiles.

Mixed Layer Depth (MLD) is done. The idea is simple, and it is related to the notion of double penalty (Ebert and McBride, 2000): if the thermocline is too deep or too shallow, when assessing the error between model and observations a double error is taken into account. First, an error in the position of the mixed layer, and secondly an error in temperature, due to a shift in depth. This is shown in Figure 6.2 If the depth error is considered independently from the error in temperature, the cause of the error can be detected more easily.

To calculate the MLD, several methods have been tried, with unreliable results. For example, the maximum gradient of the profile was very hard to calculate if the profile itself was not smooth. For the model profiles, a smooth curve is often obtained, but *in situ* measurements are very noisy. This makes very difficult the detection of the mixed layer depth. Some works have been done to calculate the depth of the mixed layer (*e.g. Soetaert et al., 2001; Thomson and Fine, 2003)*

The technique that gave the best results, and that is used in this work, is the adjustment of each individual profile to a sigmoid function:

$$s = T_u + \frac{T_b - T_u}{1 + e^{(0.5 * \frac{x-D}{W})}} \tag{6.1}$$

where:

- T_u: Temperature at the upper thermocline
- T_b: Temperature at the bottom of the thermocline
- D: depth of the thermocline
- W: width of the thermocline

These parameters can be seen in Figure 6.3. As a result of the fit, the mentioned parameters are easily obtained. The parameter that is used in this work is the depth of the thermocline, D.

This technique stands as the most useful for the purposes of the study. A Matlab routine that solves non-linear curve-fitting problems is used. Problems of this type occur in a large number of practical applications especially when fitting model functions to data, i.e., nonlinear parameter estimation. The coefficients that best fit the function F are calculated:

$$\min_x \in R^n f(x) = \frac{1}{2} \parallel F \parallel_2^2 = \frac{1}{2} \sum_i (F(x)_i)^2 \tag{6.2}$$

They are also prevalent in control where one wants the output, $y(x,t)$, to follow some continuous model trajectory, $s(t)$ (the sigmoid), for vector x and scalar t. This problem can be expressed as:

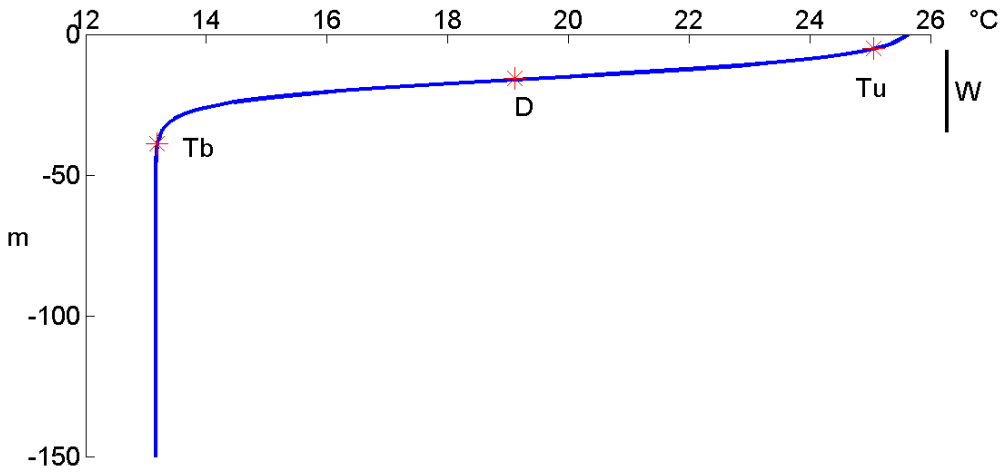

Figure 6.3: Profile fit to a sigmoid function. 'Tu' is the temperature of the upper thermocline, 'Tb' is the temperature at the bottom of the thermocline, 'D' is the Depth of the thermocline and 'W' is the width of the thermocline

$$min_x \in R^n f(x) = \sum i = 1m(\bar{y}(x, t_i) - \bar{s}(t_i))^2 \tag{6.3}$$

The residuals of the adjustment are also obtained, which give an idea of the error made. An example can be seen in Figure 6.4. The averaged initial forecast profile and its fit are shown in Figure 6.5. The fit matches very good the initial profile, with the exception of the surface temperature were the forecast is warmer.

The study of the MLD is twofold: first, it allows us to know if the forecast is able to represent with enough accuracy the position of the MLD. In addition, comparing the MLD of the forecast to what is obtained by the free run can be useful when studying the impact of the assimilation of SST and SLA at depth.

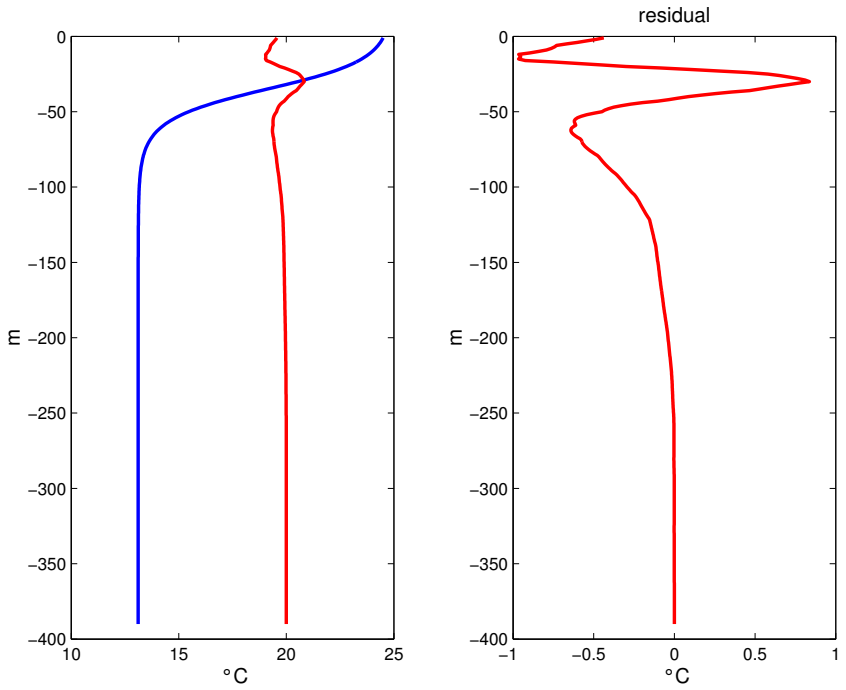

Figure 6.4: Residuals of the averaged fit respect to the initial profile. In the figure of the left, the residual superimposed to the profile. At right, the residuals alone.

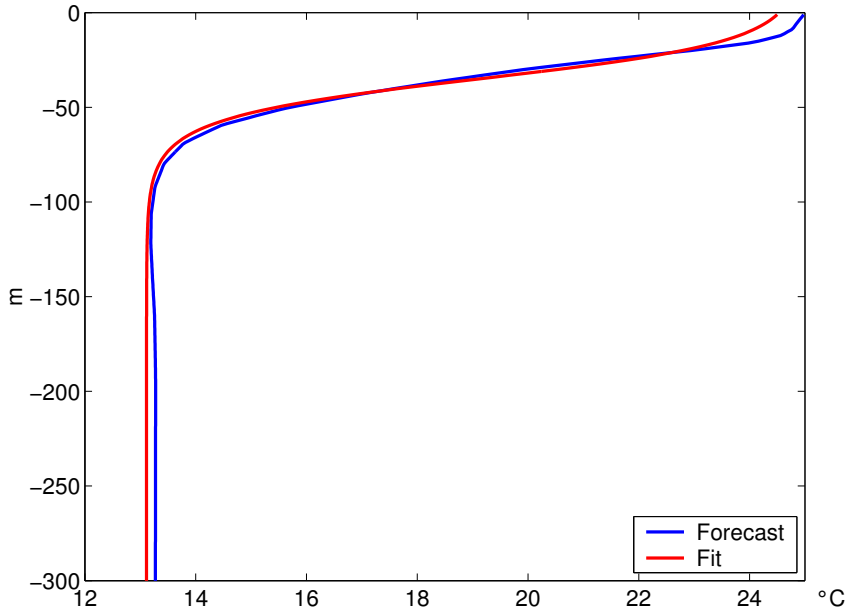

Figure 6.5: Profile and fit. In blue, the forecast profile, and in red, the fitted profile where the information about the MLD is obtained

Chapter 7

Wavelet Transforms

7.1 Introduction

The ocean is a highly variable system, both in time and in space. A wide variety of processes are present simultaneously at different spatial and temporal scales in the ocean, and they interact between them. A 3D hydrodynamic model that aims to resolve these processes is thus also a complex system, which aims also to represent those features at different spatial and temporal scales. In the forecast verification process, this complexity must be taken into account, using verification methods that do not over-simplify the comparison between the ocean and the model.

Traditional methods used in forecast verification (*i.e.*, the methods presented in chapter 5) reduce the comparison between the observations and the model into a set of numbers that explain, for example, the evolution of the error in the entire domain with time. This is certainly very useful, because it allows the forecaster to have a global idea about how the model behaves in relation with observations. But when looking in more detail for the causes of the error, in order to correct it, or when examining the capability of the model to reconstruct small scale features present in the observations, all this space and time variability information has been lost.

Multiscale methods (first developed by Mallat (1989)) allow separating a time series or a field into statistically independent components. These components are the result of the decomposition into several frequencies or scales of the initial data. In two-dimensional fields (*e.g.* SST), the decomposition is made in terms of spatial scales, and thus the model and the observations can be compared at those resolution levels, using, at each scale, the traditional error measures presented in chapter 5. This approach makes easier the analysis of the model results.

There are several introductory works on wavelets. Daubechies (1992) makes a complete description of the wavelet theory with mathematical rigor. This book was made after a CBMS conference in wavelets, where Daubechies gave ten key lectures about the

theory of wavelets. This work is addressed to a wide community of scientists, although a background on Fourier analysis is quite helpful. The work by Mallat (1998) presents also a strong mathematical basis, with a large component on the multiresolution theory. This book is easier to read than the first one, and it presents more examples and applications. Both works can be used as a good reference when working with wavelets. Jawerth and Sweldens (1994) made a review of wavelet analysis, but also with a particular focus on multiresolution analysis. More application-oriented is the work by Lau and Weng (1995), where the analogy between wavelets and music is exploited to explain the wavelet theory (very talkative to persons with a musical background), but it is limited to time series analysis. Vidakovic and Muller (1994) have explained wavelets in a very readable way (although its title, "Wavelets for Kids" is just a little far too optimistic). The approach by Sweldens (1997), with the computation of 'Second Generation Wavelets' (wavelets that are not necessarily translates and dilates of one fixed function) by a lifting scheme is a different and interesting approach with interesting applications, as wavelet computation on irregular grids. The lifting scheme is a powerful and simple way to compute the Wavelet Transform, which does not rely on a Fourier Transform. For more information about the lifting scheme, see Sweldens (1997) and Daubechies and Sweldens (1998). Finally, Kumar and Foufoula-Georgiou (1997) make a state of the art of the application of wavelets in geophysics, with a good introduction to the theory and some examples.

Wavelets have become widely used in geophysics since the 1980s. Many works have been done in time series analysis, and more specifically about the study of oscillatory behaviours, where wavelets stand as a popular analysis tool. Liu and Miller (1996) and Liu (2000), for example, analyse a current velocity time series and a wave time series on the Great Lakes respectively. In Liu and Miller (1996), the current dynamics are studied. Inertial periodic oscillations are detected, as well as other high-frequency oscillations due to the action of the wind. In Liu (2000), the continuous Morlet wavelet is used to study the grouping characteristics of wind waves. Wavelet Transforms allowed to identify energy density parcels in the time-frequency domain. Also working with waves, Meyers *et al.* (1993) have studied the dispersion of Yanai waves (mixed Rossby-gravity waves) from a time series of sea surface height obtained by a reduced gravity equatorial model. The authors study the propagation of the different wavelengths, and the results obtained agree with theory. Meyers and O'Brien (1994) have studied the temperature variations in the equatorial Indian Ocean. The aim of their work is to study the time evolution (from 1987 to 1988) of the wavelet transform of a SST time series in the Indian Ocean. Periods from 2 to 70 days are studied, showing for example that the wavelet coefficients of 10-30 days period decrease their amplitude (up to a half) from 1987 to 1988. Gambis (1992) studies the El Niño/Southern Oscillation Index (SOI) variation at intraseasonal and interannual time-scales, and its relationship with the Length Of the Day (LOD) index. The author examines the correspondence between the El Niño events and low and high-frequency oscillation patterns in the SOI and LOD. Wang and Chern (2000) combined two powerful techniques (wavelets and EOF) to analyse internal tides. The work by Wang and Chern (2000) tried to extend the classical EOF concept including the

modal decompositions of the original data to the analysis. Yiou *et al.* (2000) uses also the advantages of wavelets (namely, their finite moving window) to construct a multi-scale singular spectrum analysis (MS-SSA). Their approach allowed to analyse nonstationary time series with data-adaptative wavelets, using EOFs in the role of a wavelet. The combination of two powerful techniques as EOF and wavelet analysis is still a field under developement, and is giving interesting new insight into the geophysic processes analysis.

Wavelets are also widely applied in meteorology. For example, Weng and Lau (1994) have used a Haar wavelet to study the scale separation and period doubling of a synthetic data set, and a Morlet wavelet to a real case in the tropical Pacific IR radiance data. In this late case they identify two regimes in the variation of the IR radiance data, caused by wet and dry periods. In addition, intraseasonal and high-frequency signals are identified and analysed. Gamage and Blumen (1993) have done an interesting work in which the signature of a cold front is analysed with Fourier analysis, EOFs and Wavelets, in order to intercompare those methods. The conclusion of their work is that wavelets are superior when analysing sharp gradients. As it will be mentioned later, the strength of wavelets lies in its location property, which makes them suitable to identify and analyse abrupt changes such as gradients in a time series.

Two-dimensional wavelets have also many applications in geophysics. Kumar and Foufoula-Georgiou (1993a) and Kumar and Foufoula-Georgiou (1993b) present an interesting work where wavelets are used to analyse rainfall data. Rainfall fields are highly variable in time and in space, as well as in intensity. The segregation of large scales from the small ones using orthogonal wavelets appears to be very useful for this kind of data. Ostrovskii (1995) has studied the SST variability in the Japan Sea, in order to understand the annual mixing patterns of temperature in this zone. Teti and Kritikos (1992) have applied the wavelet transform to SAR fields filtering. Farge and her studies on two and three-dimensional turbulence (*e.g.* Farge *et al.* (1999) and Farge *et al.* (2001)) is also an interesting example on wavelet applications.

Fournier (2000) has applied Orthogonal Wavelet Transforms to observed geopotential height fields in order to demonstrate the efficiency of wavelets when analysing spatial structures in comparison to the classical Fourier Transform analysis. Recent works (Fournier (2002), Fournier (2003a) and Fournier (2003b) in the atmosphere and Liang and Robinson (2004a), Liang and Robinson (2004b) and Liang and Robinson (2004c) in the ocean) have develop a new methodology that consists on a multi-scale decomposition of the energy and vorticity terms (MS-EVA, Multi-Scale Energy and Vorticity Analysis) using wavelets, in order to quantify the budget of these terms at different spatial and temporal scales, as well as to understand the transfer and distribution of energy between those scales. In Liang and Robinson (2004c) this technology is applied to the study of the Iceland-Faroe Front variability and more particularly a cold meandering intrusion observed by *in situ* data. They worked with three scale levels, the large-scale, the meso-scale and the sub-mesoscale levels. The aim was to study how the energy is

distributed among these scales, and how the energy transfers between them.

The use of two-dimensional wavelets in field forecast verification was introduced by Briggs and Levine (1997) for meteorology. They applied a Daubechies 8 wavelet to decompose a forecast of the 500-mb height and a reference system into different resolution levels, allowing them to calculate the error for each resolution scale. Casati *et al.* (2004) applied also two-dimensional wavelets in meteorology. They applied a Haar wavelet to simulated rain fields and observations, in order to assess the skill of the model at each different spatial scale. This allows them to separate the error as a function of the scale and the intensity of the rain event.

Wavelets are used worldwide and to a high variety of data. Their potential as an analysing tool make them very appropriate to forecast verification purposes. In this chapter, the basic theory of wavelets is presented, focusing on two-dimensional Discrete Wavelet Transforms, which are used for multiscale decomposition. The aim is to decompose the model forecast and the observations into different resolution levels, representing each one a physical scale. Then, the comparison between model and observations can be done separately in each of these scales. This allows to identify the scales that are mainly contributing to the global error, and thus to have a closer look into the behaviour of these fields.

This chapter is organised as follows: first Fourier Transforms are briefly introduced to see differences and analogies with. Then, the multiresolution approach is explained, with a little example about how it works, using a time series of Sea Surface Temperature (SST). The theory of wavelets is then explained, from one-dimensional to two-dimensional wavelets, which are used in the present work. We will finish this chapter with some specific error measures that will be applied to the model results in the frame of wavelet decomposition.

7.2 From Fourier Transforms to Wavelet Transforms

There are several methods of data analysis. One of the most known is the Fourier Transform, and the Windowed Fourier Transform, a generalization of the first one. Since wavelet Transforms are based on a Fourier Transform, a brief explanation of each of these concepts is given here.

- Fourier Transforms.

  Fourier Transforms (FT) of a function $f(t)$,

$$(\mathcal{F}f)(w) = \frac{1}{\sqrt{2\pi}} \int e^{-iwt} f(t) dt \tag{7.1}$$

use sine and cosine base functions that have infinite span and are globally uniform in time. FT do not contain any time dependence of the signal, so they cannot provide any local information about the time evolution of its spectral characteristics (Lau and Weng, 1995). FT are very well suited to analyse stationary signals, but when looking for a local event in time, it can be very complex to identify it with FT (Daubechies, 1992). For example, a given signal and its inverse in time will lead to the same FT spectrum.

- Windowed Fourier Transforms.

  In an attempt to correct the problem of time localization with FT, Windowed Fourier Transforms (WFT) were developed. They include a sliding window, constant in time and frequency, which is used to analyze small parts ($f(s)$) of the function $f(t)$:

$$(WFT\ f)(w,t) = \int f(s)\ g(s-t) e^{-iws} ds \tag{7.2}$$

  Because of the constant time window, a large number of high-frequency cycles are included, and only a few of low frequency cycles. This results in a non-uniform representation of different frequency components (Kantha and Clayson, 2000).

- Wavelet Transforms.

  Wavelet Transforms (WT) use a windowing technique with variable-sized regions. This allows using long time intervals where low frequency information is needed, and short time intervals to retain high frequency information. WT can be thought as a generalized form of FT (Kantha and Clayson, 2000).

  WT use wavelets (ψ) that can stretch and translate over the signal, with a flexible resolution in both time and frequency (Lau and Weng, 1995). The expression for a Wavelet Transform is:

$$Wf(m,n) = \lambda^{-m/2} \int f(x) \psi(\lambda^{-m} x - n x_0) dx \tag{7.3}$$

The terms on this expression will be explained later. Now we look to its form. WT and WTF take both the inner products of f with a family of functions characterized by two parameters: $g_{w,t}(s) = e^{iws} g(s-t)$ for Equation 7.2 and $\psi_{m,n}(x) = 2^{-m/2} \psi(2^{-m} x - n)$ for Equation 7.3. But the shape of these two functions, as already said, is not similar. The

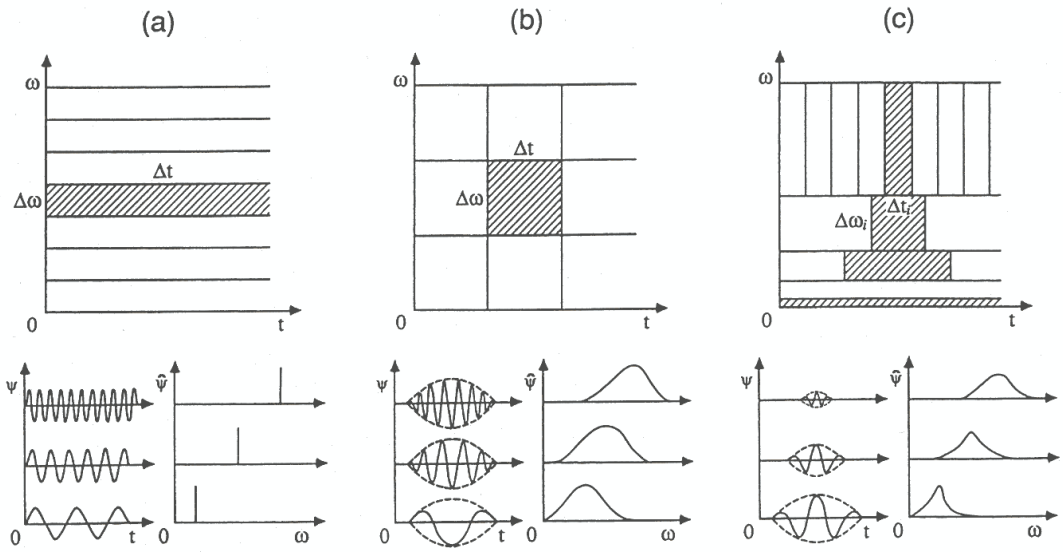

Figure 7.1: Time-frequency domains for Fourier Transforms (a), Windowed Fourier Transforms (b) and Wavelet Transforms (c). (From Lau and Weng 1995).

WFT functions $g_{w,t}$ consist all in the same envelope function g that is translated to the proper time location and has the same form, regardless of the frequency. In the opposite, WT functions $\psi_{m,n}$ have time-widths adapted to their frequency: high-frequency $\psi_{m,n}$ are very narrow and low-frequency $\psi_{m,n}$ are wider (Daubechies, 1992). This characteristic allows the wavelets to isolate high and low frequencies, and so to achieve the optimal resolution with a minimum number of basis functions (Lau and Weng, 1995). In Figure 7.1 the different domains for FT, WFT and WT are shown.

7.3 Multiresolution analysis

The ocean is a complex system with different processes acting at different spatial and temporal scales, and interacting between them. A multiresolution analysis allows to study these processes at each scale, and to understand the relationship between scales. Wavelets are useful for multiresolution analysis (Mallat, 1989). The advantage of using wavelets to analyse a set of data is because of their capability to decompose a function into different resolution scales (or frequencies). Each scale is statistically independent, which is an attractive characteristic for verification purposes. The resolution of the data becomes coarser and coarser with the increasing scale. At a given resolution scale, the data are decomposed into two sets. These two sets are called the Approximation and the Details. The Approximation retains those features larger than a given resolution m and the Details retain the small features characteristic of this scale. Details can be also saw as the difference between the Approximation at scale $m-1$ and the Approximation at

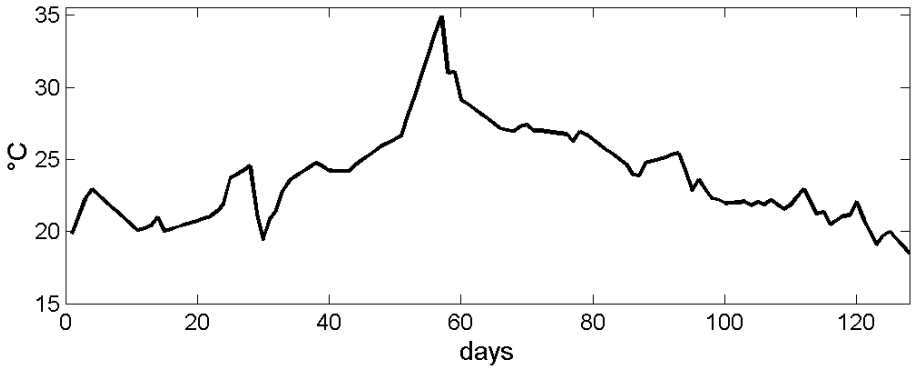

Figure 7.2: Sea Surface Temperature at the mouth of the Po River, in the northern Adriatic. Day 1 is for 28 May.

scale m. Given a function $f(x)$, the approximation of this function at the scale $m+1$ plus the details at this scale, will recover the signal at a finer scale (f_m).

$$f_m(x) = f_{m+1}(x) + f'_{m+1} \tag{7.4}$$

where f'_{m+1} are the details at scale m+1 and f_{m+1} are the approximation data at scale m+1. Let's illustrate the multiresolution analysis with an example. In Figure 7.2 we can see the sea surface temperature obtained from AVHRR satellite imagery at Po River mouth, in the Northern Adriatic Sea. The data set goes from 28 May to 2 October 1995 and it has a length of 128 days. In wavelet analysis, it is desirable to work with data sets of dyadic length, because the decomposition into scales or frequencies is made in powers of two. However, it is always possible to obtain a set of dyadic size simply by extending the initial set with 'dummy' data. There are several ways to fill our data: with zeros, linear extrapolation, periodically... This choice must be done carefully, as the Wavelet Transform will be always affected, and it is desirable that the impact is the minimum possible.

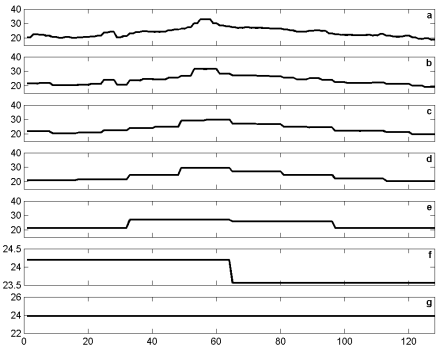

Figure 7.3: Approximation from scale 1 (a) to scale 7 (g) of the Po time series. Day 1 is 28 May. Note that (f) and (g) does not present the same y axis as (a) to (e). X-axis shows days from 28 May and y-axis show the temperature in °C.

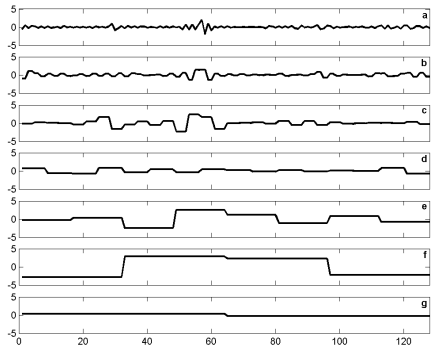

Figure 7.4: Details from scale 1 (a) to scale 7 (g) of the Po time series. X-axis shows days from 28 May and y-axis show the temperature in °C.

In Figures 7.3 and 7.4 the decomposition in Approximation and Details of the Po SST time series are presented respectively. The decomposition is done in dyadic data sets, from two times the initial time scale (2^1) to 128 times the initial time scale (2^7) days. The time scale of the original data set is one day, so the decomposition is done into $2^1, 2^2, ..., 2^7$ days, and they will be called scales 1, 2, ..., 7. The wavelet used for the decomposition is called 'Haar wavelet' and it will be explained later in this chapter.

Figure 7.3a is thus the approximation at scale 1. It retains the information that in the original data set has more than 1 day duration. Details, or features with 1 day duration, are retained in the details at scale 1, in Figure 7.4a. Note that the approximation at the coarser scale (2^7) represents the mean of the original data set. Between the finest and the coarsest scales, we find all the features present in the original data set, divided into seven resolution levels. For example, in Figure 7.4f, details at scale (2^6) retains those features with a period of 64 days, that is, we can see the seasonal cycle, with summer months of July and August being detected in this figure. The approximations are obtained by the application of a Scaling Function, that smoothes the data. Details are extracted by means of a Mother Wavelet. The concepts of Mother Wavelet and Scaling Function will be explained later. Finally, it should be noticed that the sum of the approximation at the coarsest scale (the mean value of the data set) with the details of all scales, gives the original signal, that is:

$$\text{Original Signal} = \text{Approximation}_7 + \sum_{m=1}^{7} \text{Details}_m \tag{7.5}$$

The Scaling Function can be also thought as a low-pass filter and the Mother Wavelet as a high-pass filter. In fact this is the concept exploited by Sweldens (1997) in the computation of second-generation wavelets.

7.4 Wavelets

7.4.1 General Description

A wavelet (ψ) is a real function with a localized waveform and an average value of zero. Wavelets are characterised by a location and a scale, which make them suitable for multiresolution analysis. These are some of their general characteristics (Kumar and Foufoula-Georgiou, 1997; Torrence and Compo, 1998):

- Compact Support, or fast decay. Wavelets are not sustaining waves as they are in Fourier analysis. This means that the practical non-zero zone is localised, which allows location in space.
- Zero mean. This property ensures that the wavelet has a waveform.
- Unit energy ($\int |\psi|^2 = 1$)

It should also be pointed out that the time-frequency localization property of wavelets is limited by the Heisenberg uncertainty principle. One cannot measure with arbitrarily high resolution in both time and frequency (Kumar and Foufoula-Georgiou, 1997; Mallat, 1998).

The mathematical expression for a wavelet is:

$$\psi_{m,n}(x) = \lambda^{-m/2}\psi(\lambda^{-m}x - nx_0) \tag{7.6}$$

The wavelet ψ is dilated and translated by means of parameters m and n, so that $\psi_{m,n}$ is the wavelet at a scale m and at the position n. The complete set of all $\psi_{m,n}$ forms a basis function, $\{\psi_{m,n}\ \forall m, n \in \mathbb{R}\}$. Consider the function $f(x) \in L^2(\mathbb{R})$, where the orthogonal basis $L^2(\mathbb{R})$ denotes the vector space of measurable, square-integrable one dimensional functions $f(x)$. $f(x)$ can be represented as a linear combination of these basis functions. The Wavelet Transform of $f(x)$ is obtained from the convolution between the function $f(x)$ and the wavelet $\psi_{m,n}(x)$, and its expression was shown on Equation 7.3.

There are many types of wavelets, which can be classified into two main groups: Continuous Wavelets and Discrete Wavelets. The choice of the type of wavelet being used depends on the problem one wants to solve. For multiresolution purposes, Discrete

Wavelets must be used. Another additional characteristic that the wavelet chosen must fulfil is orthogonality. This characteristic allows making the decomposition into statistically independent components, and thus the analysis of these components is much easier.

7.4.2 Discrete Wavelet Transforms

The orthogonality of a wavelet depends on the choice of parameters λ and x_0 on Equation 7.6. If the wavelet basis formed by all $\psi_{m,n}$ $\forall m, n \in \mathbb{R}$ is an orthogonal basis, the decomposition of a function $f(x)$ will be made into orthogonal components (Daubechies, 1992). The particular choice of $\lambda = 2$ and $x_0 = 1$, gives an orthogonal wavelet basis $\{\psi_{m,n}(t)\}$:

$$\psi_{m,n}(x) = 2^{-m/2}\psi(2^{-m}x - n) = \frac{1}{\sqrt{2^m}}\psi\left(\frac{x - n2^m}{2^m}\right) \tag{7.7}$$

The Wavelet Transform is discrete if parameters m and n are discrete. Note that the wavelet is continue in itself, but the location and scale parameters are discrete. The wavelet is stretched by a factor of 2^m and translated by a factor of $n2^m$. The choice of m and n depends on the size of the function to be decomposed. The dilated and translated family $\{\psi_{m,n}(t)\}$ forms an orthornomal basis of $L^2(\mathbb{R})$. Several wavelets work better when the length of the function to analyze is dyadic, *i.e.* it has a size 2^M, $M \in \mathbb{Z}$. The values for m will then be $m = 1, 2, ..., M$.

Wavelets are dilated as scale m increases. The decomposition of a function with a wavelet stretched by a factor m is called the decomposition at *scale* m. A wavelet of scale m+1 have a support double than a wavelet at scale m. The process of wavelet decomposition is as follows: for a function of size 2^M, first a stretching factor of 2^1 is applied to the wavelet, the data of the function are taken 2 by 2, and the inner product between ψ and f(x) is calculated. This is the decomposition at scale 1. Then the stretching factor is incremented to 2^2, so data are taken 4 by 4. This gives the decomposition at scale 2. The procedure is repeated until arriving at a stretching factor of 2^M, that is, the whole data are taken at once. This is decomposition at scale M. At each of these scale decompositions, the wavelet is translated over the whole function length.

As mentioned at the beginning of this chapter a Mother Wavelet and a Scaling Function are necessary to realise the multiscale analysis. The expression for the Mother Wavelet was given in Equation 7.7. The Scaling Function has the following expression:

$$\phi_{m,n}(x) = 2^{-m/2}\phi(2^{-m}x - n) = \frac{1}{\sqrt{2^m}}\phi\left(\frac{x - n2^m}{2^m}\right) \tag{7.8}$$

Going back to Equation 7.4, now the terms of this equation can be determined

$$f_m(x) = \sum_{n=-\infty}^{\infty} C_{m,n}\phi_{m,n}(x) \tag{7.9}$$

$$f'_m(x) = \sum_{n=-\infty}^{\infty} D_{m,n}\psi_{m,n}(x) \tag{7.10}$$

$C_{m,n}$ and $D_{m,n}$ are the Scaling Function and the Mother Wavelet coefficients respectively, and are given by:

$$C_{m,n} = \int f(x)\phi_{m,n}dx \tag{7.11}$$

$$D_{m,n} = \int f(x)\psi_{m,n}dx \tag{7.12}$$

The Scaling Function ϕ acts as a smoothing function, filtering out the details, so that the Approximation is obtained. The Mother Wavelet retains these details. The Approximation is further decomposed into the following scales and at each level the Details are taken out of the signal. The decomposition into an approximation field and a detail field, however, does not increase the number of points. If the initial field has a size of $2^M \times 2^M$ pixels, each of the decomposed fields at a coarser scale has $2^{M-1} \times 2^{M-1}$ (Mallat, 1989). This is due to the orthogonality characteristic and prevents the increase of data.

The Scaling Function and the Mother Wavelet must derive from the same type of wavelet. In Figures 7.5 to 7.10 you can see a few examples of Mother Wavelet and Scaling Functions.

7.4.3 The choice of the wavelet: The Haar wavelet

The wavelet form chosen to decompose a data set will undoubtedly affect its decomposition. This is why a careful choice of the Mother Wavelet is very important. Torrence and Compo (1998) names several factors that should be taken into account when choosing the Mother Wavelet: 1) Orthogonal or non-orthogonal wavelets (which has already been discussed); 2) Complex or real wavelets (complex wavelets are also largely used in geophysics, *e.g.* Spedding *et al.* (1993), Chapa *et al.* (1998) and Rao and Murthy (2001)); 3) the width of the wavelet (or compact support); and 4) the shape of the wavelet; in fact, the shape of the wavelet should be chosen accordingly to the characteristics of the data to analyse, because the results change when using different wavelet forms. There are other methods to choose the more adequate waveform. In Briggs and Levine (1997), the wavelet that minimises the entropy of the Wavelet Transform is chosen. Casati *et al.* (2004) worked with binary fields of rain, and the square-shaped Haar wavelet appeared

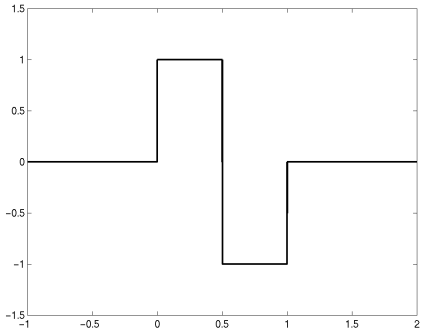

Figure 7.5: Haar Mother Wavelet

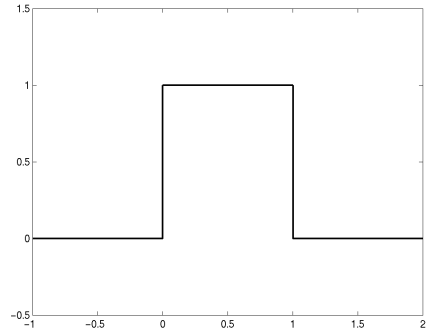

Figure 7.6: Haar Scaling Function

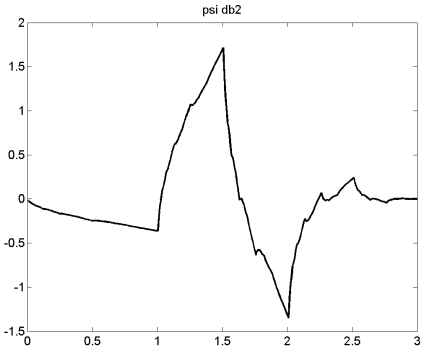

Figure 7.7: Daubechies 2 Mother Wavelet

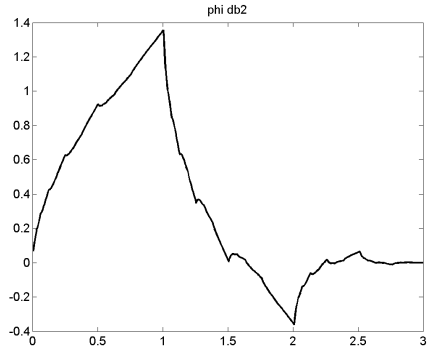

Figure 7.8: Daubechies 2 Scaling Function

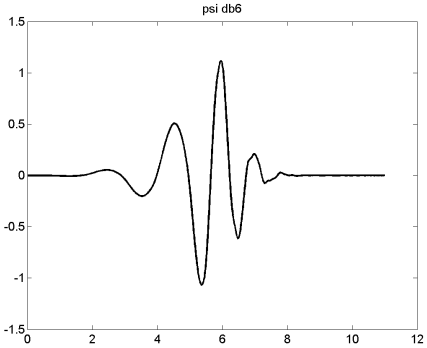

Figure 7.9: Daubechies 6 Mother Wavelet

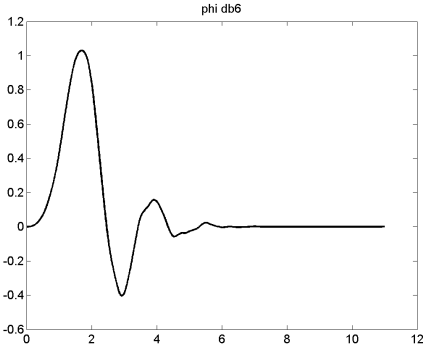

Figure 7.10: Daubechies 6 Scaling Function

to fit the best to the data.

In Oceanography, there is an additional degree of difficulty when using wavelets. Often the oceanographic domains contain irregular boundaries at the limit between land and sea, which is not the case in the atmosphere, for example. As the wavelet travels along the domain, it encounters these boundaries. The abrupt change from sea to land points results in highly perturbed wavelet coefficients. For this reason, the use of a small wavelet will help to minimise this effect. Wavelets have a determined compact support, or effectively non-zero zone, and the wider it is, the further the perturbation will be felt. For example, Daubechies Wavelet family has a compact support of 2N-1, with N the order of the wavelet. The Haar Wavelet has a support width of two, the smaller support width of all the wavelet families. In Figures 7.5 to 7.10 we can see the mentioned wavelet families.

The Haar Wavelet has been chosen for this work, to minimise the boundary effects. Other factors have been taken into account to minimise these effects, and will be explained latter in this chapter.

The Haar wavelet is the simplest one from the whole family of wavelets. The Mother Haar Wavelet has the following values:

$$\psi_{m,n}(x) = \begin{cases} 1 & 0 \leq x \leq 1/2 \\ -1 & 1/2 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.13}$$

and the Scaling Function:

$$\phi_{m,n}(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.14}$$

Both functions can be seen in Figures 7.5 and 7.6.

7.4.4 Two-Dimensional Discrete Wavelets

In this work, two-dimensional Haar Wavelets are used. Two-dimensional wavelets are obtained from the orthogonal product of one-dimensional wavelets. Two dimensional wavelets basis have the same characteristics as those presented for the one-dimensional case. The signal to be analyzed now is $f(x, y) \in L^2(\mathbb{R}^2)$ ($L^2(\mathbb{R}^2)$ denotes the vector space of measurable, square-integrable two dimensional functions $f(x, y)$). As a first definition, one can think at the two-dimensional Haar Wavelet as a spatial averaging filter over $2^m \times 2^m$ data, with m the scale of the decomposition. The Scaling Function computes the average value of the data at a given scale m and the Mother Wavelet retains the differences between this averaged data and the finer scale $m - 1$. A little diagram about how it works can be found in Figure 7.11. In this figure, one can see

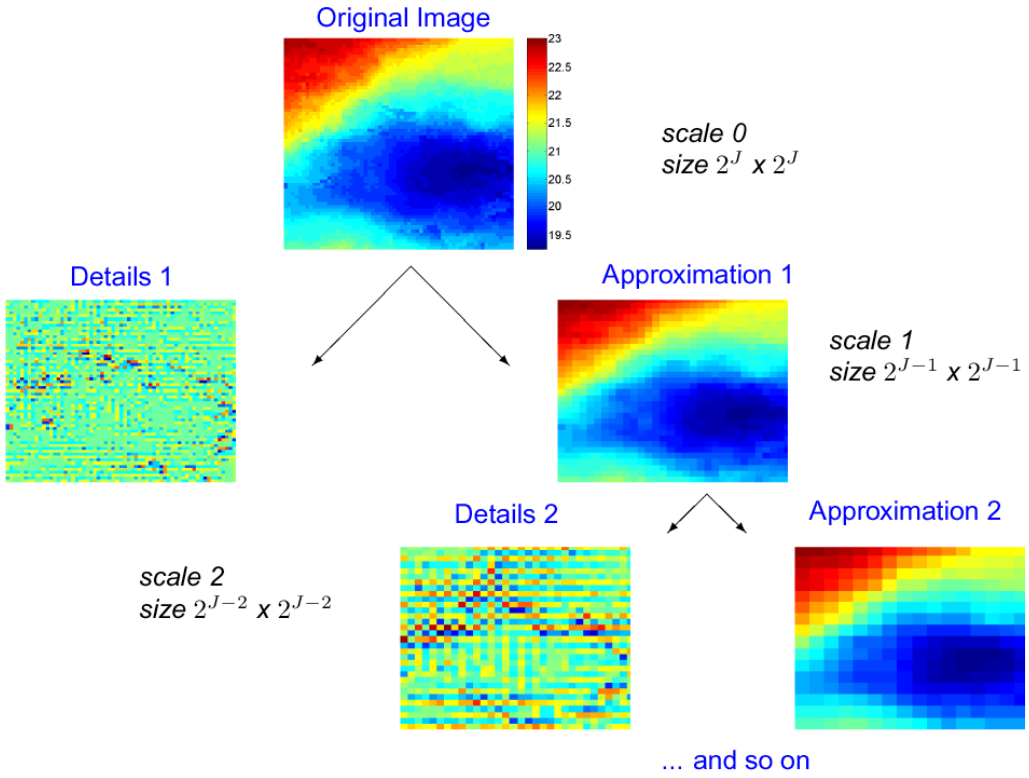

Figure 7.11: Three first steps in the two-dimensional decomposition of the Sea Surface Temperature of the Ligurian Sea. At the top of the figure we see the original image, and successive decompositions follow below it.

the three first steps of the decomposition of a two-dimensional image of the sea surface temperature on the Ligurian Sea. Starting from the original data (which can be referred as 'scale 0'), the mean over 2×2 pixels is calculated, obtaining the approximation at scale 1. The difference between the approximation at scale 1 and the original image (scale 0) gives the details at scale 1. Then the approximation at scale 1 is again averaged, in groups of 4×4 pixels, and the approximation at scale 2 is obtained. The difference between the approximation at scale 2 and the approximation at scale 1 gives the details at scale 2. This procedure is repeated until the whole data set is averaged at once, so that the approximation at scale M represents the mean value of the original image. The details at scale M are again the difference between the approximation at scale M and the approximation at scale M-1.

In a more strict way, the two-dimensional Mother Wavelet can be expressed as (Mallat, 1989; Kumar and Foufoula-Georgiou, 1993a; Mallat, 1998):

$$\psi_{m,n,k}(x,y) = 2^{-m/2}\psi(2^{-m}x - n, 2^{-m}y - k) \tag{7.15}$$

and the Scaling Function has the respective form:

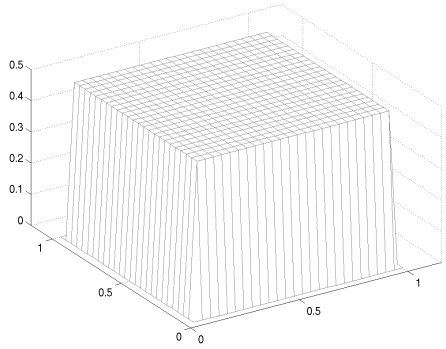

Figure 7.12: Approximation two-dimensional Haar Wavelet

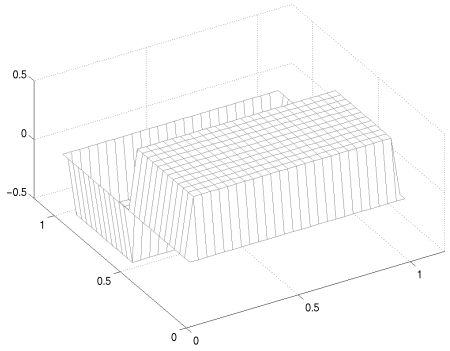

Figure 7.13: Vertical Details, $\phi(x)\psi(y)$

$$\phi_{m,n,k}(x,y) = 2^{-m/2}\phi(2^{-m}x - n, 2^{-m}y - k) \tag{7.16}$$

The approximation of function $f(x,y)$ at a resolution m is obtained through the inner products:

$$A_m f = \left\{(f, \Phi_{mnk})_{(n,k)\in Z^2}\right\} = \left\{(f, \phi_{mn}\phi_{mk})_{(n,k)\in Z^2}\right\} \tag{7.17}$$

For the details, the Mother Wavelet and Scaling Function are combined as follows:

$$\Psi^1(x,y) = \phi(x)\psi(y);\ \Psi^2(x,y) = \psi(x)\phi(y);\ \Psi^3(x,y) = \psi(x)\psi(y); \tag{7.18}$$

so that the details are obtained by the inner product of $f(x,y)$ with each of these vectors:

$$D_m^1 f = \left\{(f, \Psi_{mnk}^1)_{(n,k)\in Z^2}\right\} \tag{7.19}$$

$$D_m^2 f = \left\{(f, \Psi_{mnk}^2)_{(n,k)\in Z^2}\right\} \tag{7.20}$$

$$D_m^3 f = \left\{(f, \Psi_{mnk}^3)_{(n,k)\in Z^2}\right\} \tag{7.21}$$

The decomposition of $f(x,y)$ is thus made into the sum of three components, that account to different spatially oriented frequency channels (Mallat, 1989; Kumar and Foufoula-Georgiou, 1993a): vertical high frequencies (high horizontal correlation of the data) retained by $D_m^1 f$; horizontal high frequencies (high vertical correlation), retained by $D_m^2 f$; and high frequencies in both directions (high vertical and horizontal correlation, features like corners, for example), retained by $D_m^3 f$. This property can be very useful when working with fields with directional behaviour (as, *e.g.* rainfall fields, in Kumar and Foufoula-Georgiou (1993a) and Kumar and Foufoula-Georgiou (1993b)).

In Figures 7.12 to 7.15 you can see the aspect of a two dimensional Haar Mother Wavelet and the three components of Details.

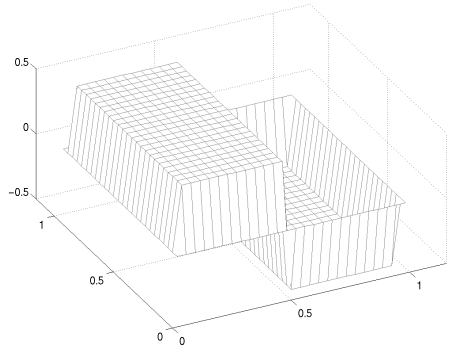

Figure 7.14: Horizontal Details, $\psi(x)\phi(y)$

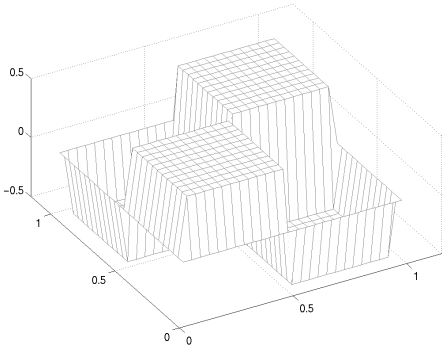

Figure 7.15: Diagonal Details, $\psi(x)\psi(y)$

7.4.5 Inverse Wavelet Transform

The general expression for the Inverse Wavelet Transform in the one-dimensional continuous case is as follows (Daubechies, 1992):

$$f(x) = \frac{1}{C_\psi} \int_{\infty}^{-\infty} \int_{\infty}^{-\infty} m^{-2} W f(m,n) \psi_{m,n}(x) dm dn \tag{7.22}$$

where C_ψ is a constant that depends on the choice of the wavelet (Daubechies, 1992; Kumar and Foufoula-Georgiou, 1997). The original signal $f(x)$ can be thus reconstructed, with no loss of data, from the wavelet coefficients with the expression 7.22. In the multiscale framework one can think of a signal as a superposition of wavelet coefficients at different resolution levels (as in Equation 7.5).

7.5 Error Measures applied to Wavelet Decomposition

Once the model and the observations (for example SST images, as we need two-dimensional fields to carry on the 2D WT) are decomposed into a series of resolution scales, the verification process can start. The data used in this work for multiscale decomposition is the sea surface temperature obtained by the model in the Ligurian Sea and the DLR SST in the same zone. The aim is to compare the model to observations at each of the resolution scales. In Table 7.1 the resolution of each scale for the data used is specified, in the wavelet space and in the data space. One can choose to work directly in the wavelet space (the decomposed fields), or in the data space (Inverse Wavelet Transform fields). Each scale is isolated before computing the inverse wavelet, as explained in Section 7.4.5 so one recovers each scale separately in the data space, but with the resolution of the initial data. In the wavelet space, one can measure how each scale h of wavelet coefficients resembles the structure of the Mother Wavelet. In the data

Table 7.1: Scales and resolution attained with the wavelet decomposition in the data used in this work. Scale 0 is the original data. The number of data of the different scales once transformed back to the data space remains constant.

scale	resolution (km)	number of data	
		wavelet space	data space
initial data	1		256 × 256
1	1	128 × 128	256 × 256
2	2	64 × 64	256 × 256
3	4	32 × 32	256 × 256
4	8	16 × 16	256 × 256
5	16	8 × 8	256 × 256
6	32	4 × 4	256 × 256
7	164	2 × 2	256 × 256
8	128	1 × 1	256 × 256

space, the different scales provide an interpretation of the data at each level of resolution. In this work, this last option is chosen, because of the already mentioned boundary problems. If we want to assess the error at each scale, but the number of pixels covered by land decreases at each scale, the relationship between scales is somewhat confusing. When working on the data space, all scales have the same number of pixels, and the effect of boundaries is less important, or at least, it affects in the same way all the scales.

First, the variability of each scale is calculated, to compare the behaviour of the model to observations about the distribution of energy between scales. The variability of the observations for each scale h is expressed as:

$$V_h = \sum_{i=1}^{M}\sum_{j=1}^{N}(x^o_{i,j} - \bar{x^o})^2 \tag{7.23}$$

The standard deviation is preferred in order to obtain interpretable results (with the same units as the data):

$$SD_h = \left[\frac{1}{N-1}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(x^o_{i,j} - \bar{x^o}\right)^2\right]^{1/2} \tag{7.24}$$

The size of the data set is $M \times N$ with $M = N$. In general, the error measures explained in chapter 5 are applied at each scale h on forecast and observations (ACC_h, $RMSE_h$...).

Because of the linearity property of the Wavelet Transform, one can calculate an error score both in the wavelet and in the data space with no change in the result. However, in the case of the RMSE this is not true, because the number of data is contained in the score. The total number of points in the data space is $N \times N$, and in the wavelet space it is $n_h \times n_h$ where $\sum n_h = N$ (as shown in Table 7.1). The equivalence between the RMSE in the data space and in the wavelet space is:

$$\frac{N^{1/2}}{n_h^{1/2}} RMSE_h(data) = RMSE_h(wavelet) \tag{7.25}$$

Additionally, the energy ratio (following Briggs and Levine (1997)) is calculated. Each scale has a different variability, and the energy ratio can help to know whether the forecast and the observations have the same variability at each spatial scale:

$$ER_h = |1 - \xi| \tag{7.26}$$

with

$$\xi = min\left(\frac{\| x_h^o \|_2^2}{\| x_h^f \|_2^2}, \frac{\| x_h^f \|_2^2}{\| x_h^o \|_2^2}\right) \tag{7.27}$$

so that ER_h is bounded between 0 and 1. If the forecast is near to 0, its variability is similar to the observed one. When the variability of the forecast differs from the observations, it tends to 1. If the forecast and observed matrices are expressed as vectors, the ER_h can be also interpreted as the ratio of the lengths of these vectors (Briggs and Levine, 1997). Moreover, the correlation ACC_h can be viewed as the cosine of the angle between both vectors, and the $RMSE_h$ as a function of a distance between forecast and observations, as explained in chapter 5 (the relation between these two variables can be viewed in Figure 7.16). These three measures, ER_h, ACC_h and $RMSE_h$ will be thus interpreted together.

The percentage each scale contributes to the total error can be also calculated for ACC and MSE (MSE is preferred to RMSE to calculate a percentage score because of the presence of a radical in the RMSE). The expression for the $\%ACC$ derives from the expression for ACC given in chapter 5:

$$\%ACC_h = \frac{\left[(x_h^f - C_h)(x_h^o - x_h^o)\right]}{[(x^f - C)(x^o - C)]} \tag{7.28}$$

And for the $\%MSE$:

$$\%MSE = \frac{\left(x_h^f - x_h^o\right)^2}{\left(x^f - x^o\right)^2} \tag{7.29}$$

All the given error measures can be calculated in the wavelet space or in the data space. When doing on the data space, all scales different from h must be set to zero before doing the Inverse Transform.

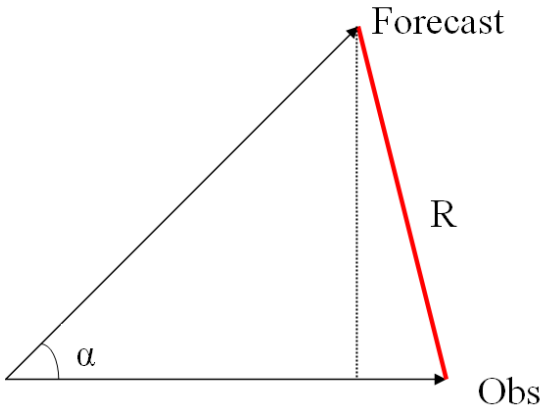

Figure 7.16: Geometrical interpretation of correlation and RMSE. The correlation between Forecast and Observations can be thought as the cosine of angle α. The RMSE is thus the distance between these two vectors.

7.6 The boundary problem

As already mentioned, the irregular boundaries present in most oceanic models can be a problem when working with wavelets. The perturbation caused to the wavelet coefficients when they analyse points near the coast may affect the decomposition, and consequently in the error measures realised with them. The problem can be viewed as an extension of the one-dimensional case: when the analysed time series do not have a dyadic length, it is normally padded with extra values up to the nearest dyadic length. These added points can be constructed in many different ways, (*e.g.* with zeros, or a periodic extension of the original data, a gently decrease to zero, the mean value of the variable, etc). The same can be done in two dimensions to reach to the dyadic length trying to minimise the perturbation of the wavelet coefficients. Boundary problems have been treated by *e.g.* Sweldens (1997), Oh *et al.* (2001) and Naveau and Oh (2004). Second generation wavelets (Sweldens, 1997) can be very helpful when working on irregular grids, and can also deal with irregular boundaries. The approach by Oh *et al.* (2001) and Naveau and Oh (2004) based on a polynomial wavelet regression, is specially designed to deal with irregular boundaries.

In this work a gently decrease to zero is used. The values nearest to the coast are extended towards land points, reaching a constant value at a given point on land. The effect of this procedure can be seen in Figure 7.17.

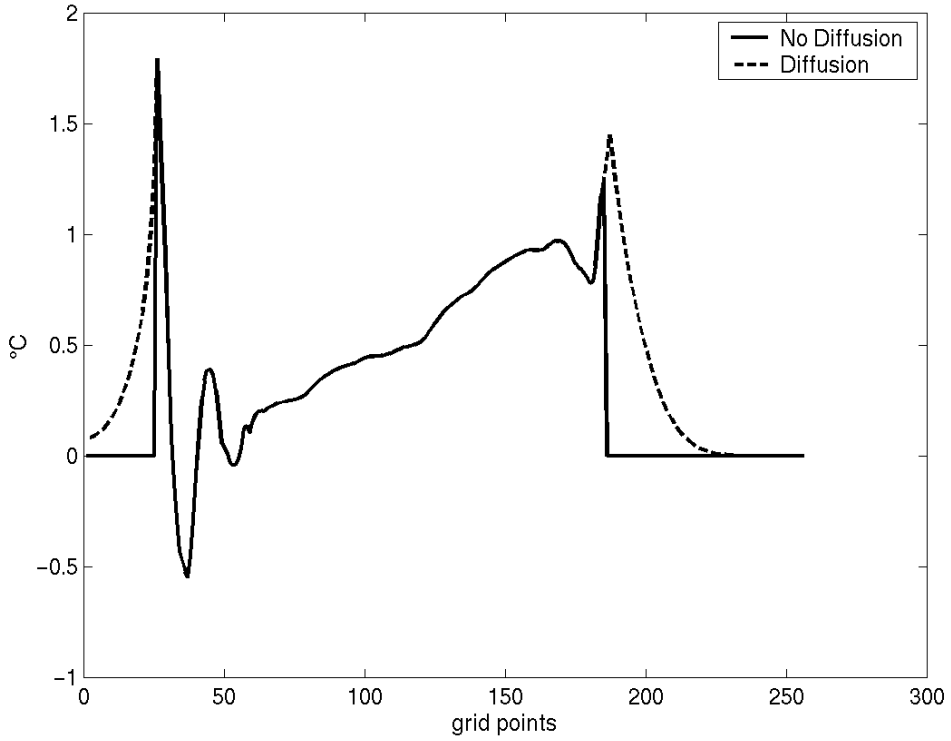

Figure 7.17: Effect of smoothing the transit between sea and land in a meridional transect on the Ligurian Sea. The solid line represents the initial field and the dashed line represents the smoothed field.

Chapter 8

Reconstruction of satellite images

8.1 Introduction

Satellite images are commonly used in oceanography. They allow realising general studies of the sea surface characteristics (*e.g.* Gacić *et al.*, 1997; Borzelli *et al.*, 1999) as well as studies at depth that present a signal at the surface, such as internal waves (*e.g.* da Silva *et al.*, 1998; Liu *et al.*, 1998). Satellite images are widely used due to their extensive coverage, in time and space. No other data acquisition method gives the coverage and precision of satellite images as quickly. Obtaining observational data in this way is therefore very useful in operational oceanography and in near real time studies.

There are several kinds of data measured by satellites, depending on the type of sensors used: Sea Surface Temperature (SST), chlorophyll or Sea Surface Height (SSH). The receptors working in the visible and infrared ranges present however one disadvantage: the clouds present in the atmosphere can totally or partially cover the studied area. The loss of data due to clouds can reach a high percentage in some periods. There can also be noise, or malfunctions in the satellite that reduce its coverage. A complete data set is crucial for many applications using satellite images, as in Empirical Orthogonal Function (EOF) analysis, when tracking features in the ocean or in the study of zones with high spatial and temporal variability. In addition, complete fields of sea surface temperature are used to force meteorological models. For other studies where it might not be crucial to have complete satellite coverage, it is always preferable to have a maximum amount of information.

Several methods have been used in the past when dealing with recovery of missing data. Spline interpolation (Emery and Thomson, 1998) has been used by, *e.g.*, Everson *et al.* (1997). Inverse methods, such as optimal interpolation (OI) (Bennett, 2002), have also been widely used in the reconstruction of SST data sets (*e.g.* Chu *et al.*, 1997; Fieguth *et al.*, 1998; Houseago-Stokes, 2000; Gomis *et al.*, 2001; He *et al.*, 2003), as well as in the reconstruction of Sea Level Anomaly (SLA) data sets (*e.g.* Fieguth *et al.*, 1998; Le Traon *et al.*, 1998; Le Traon and Didarboure, 1999; Le Traon *et al.*, 2001). Many

studies have reconstructed historical SST and *in situ* data sets (see *e.g.* Reynolds and Smith, 1994; Smith *et al.*, 1996; Kaplan *et al.*, 1997, 1998) using OI. Reynolds and Smith (1994) used OI to realise a global SST analysis. In Smith *et al.* (1996), the OI analyses are used to compute the most dominant EOFs. These EOFs are then fitted to *in situ* data to compute the reconstructed SST set. The work by Kaplan *et al.* (1997) first calculates the EOFs from the covariance matrix, and then a least-squares fit is done to estimate the best reconstruction of a 136-year SST set on the Atlantic Ocean. In Kaplan *et al.* (1998) the same technique is applied, and three different fitting methods are compared. One disadvantage of the optimal interpolation as a method for reconstructing missing data is the necessity for *a priori* information about the error statistics of the data, generally poorly known (Bennett, 2002). The computational cost of these techniques may be prohibitive when using large matrices (Kaplan *et al.*, 1997).

The alternative methodology presented by Beckers and Rixen (2003) is a self-consistent, parameter-free technique for the reconstruction of gappy data that presents the advantage of not needing this kind of *a priori* information. The method allows calculating the missing data from an optimal number of EOFs determined by a cross-validation technique (*e.g.* Wilks, 1995; Brankart and Brasseur, 1996; von Storch and Zwiers, 1999). This cross-validation also gives an error estimate of the filled data. The method is based on the fact that an EOF analysis aims to extract a small number of significant degrees of freedom, present in the physical system, from a large data set. These reduced variables should represent a large fraction of the original variability of the data set (*e.g.* Wilks, 1995). The combination of the dominant EOFs and their amplitudes can therefore help recover missing data values.

The aim of this work is the application of the reconstruction method presented by Beckers and Rixen (2003), hereafter called DINEOF (Data INterpolating Empirical Orthogonal Functions) to a realistic case: a data set covering a whole subbasin at high resolution. The data chosen for this purpose is a series of Advanced Very High Resolution Radiometer (AVHRR) images covering the entire Adriatic Sea for a 6-month period (Alvera-Azcárate *et al.*, 2004b).

The chapter is organised as follows: in section 2.4.2, a description of the main characteristics of the Adriatic Sea is presented, followed by a review of applications of satellite images in this zone. The data set used in this work is described in section 8.3.1. In section 8.2 the method used for the reconstruction of missing data is briefly presented, as well as some estimates of the performance of the code and computational time. Reconstruction results of the Adriatic images are shown in section 8.3.3, with additional validation studies, apart from the cross-validation itself. A comparison with a classical OI reconstruction is made in section 8.3.5. To show the robustness of the results obtained, the EOF decomposition of the reconstructed data is presented in section 8.3.7. Finally, we present our conclusions.

8.2 The Method

Beckers and Rixen (2003) have presented a self-consistent method for the reconstruction of missing data in oceanographic data sets, DINEOF. Consider that $\mathbf{X}$ is the initial matrix of dimensions $m \times n, m > n$ (with m the spatial dimension and n the temporal dimension), containing the observations. It may also contain some unknown values corresponding to the missing data. For the reconstruction of these data, a Singular Value Decomposition (SVD) technique is used to compute the EOFs of the matrix, in which a first guess has been introduced for the missing data. The equation:

$$\mathbf{X} = \mathbf{USV}^T \tag{8.1}$$

allows calculating the spatial EOFs, $\mathbf{U}$, with dimension $m \times r$, the temporal EOFs, $\mathbf{V}$, with dimension $n \times r$, and their singular values $\mathbf{S}$, with dimension $r \times r$. The value r is the rank of the matrix, with $r \leq \min(m, n)$. Only the most significant spatial and temporal EOFs are necessary for the reconstruction method. The k largest singular values and singular vectors can also be calculated by eigenvector decomposition:

$$\mathbf{XX}^T\mathbf{u}_i = \rho_i^2\mathbf{u}_i \tag{8.2}$$

with $\mathbf{u}_i$ the *ith* column of $\mathbf{U}$ and ρ_i the corresponding singular value, $i = 1...k$. To avoid using Eq. 8.2, which implies the $m \times m$ matrix $\mathbf{XX}^T$, we can rather use:

$$\mathbf{Av}_i = \rho_i^2\mathbf{v}_i \tag{8.3}$$

$$\mathbf{u}_i = \frac{\mathbf{Xv}_i}{\rho_i} \tag{8.4}$$

where $\mathbf{A} = \mathbf{X}^T\mathbf{X}$ is a real symmetric $n \times n$ matrix.

DINEOF can be explained as follows:

- The average value of the matrix is subtracted once for the entire procedure and the missing data points are initialised to zero in order to have an unbiased first guess. The missing data, however, are 'flagged' to differentiate them from those existing points on the mean. This demeaned matrix is used throughout the whole procedure.
- Two steps are then repeated for a given k until convergence:
  - → An EOF decomposition of the matrix is realised, with only the first k EOFs, to obtain a first estimate of the singular values and singular vectors.
  - → The elements $\mathbf{X}_{i,j}$ corresponding to the flagged missing data are now replaced by the value obtained with the EOF series:

$$\mathbf{X}_{i,j} = \sum_{p=1}^{k} \rho_p(\mathbf{u}_p)_i(\mathbf{v}_p^T)_j \tag{8.5}$$

An improved guess has thus been introduced for the missing data, so we recompute the EOFs and obtain a new value for the missing data.

- Once the convergence is reached, the number of computed EOFs is increased, from $k = 1...k_{max}$, so at the end we have k_{max} estimates for the missing data reconstructed with $1, 2...k_{max}$ EOFs. But which estimate is the best? The answer is obtained by cross-validation.

- We calculate the optimal number of EOFs from the series of k_{max} EOFs. To do so, a random set of data is initially set aside from the valid data to apply a cross-validation technique, as described in, *e.g.*, Wilks (1995, chap. 6), Brankart and Brasseur (1996), von Storch and Zwiers (1999, chap. 18) and Beckers and Rixen (2003). This data set has a size of $\min(0.01 * m * n + 40, 0.03 * m * n)$, and for this particular case, 99532 points are retained for the cross-validation. The optimal number of EOFs is the one that minimises the error between the data set aside and the values obtained at these points with the reconstruction method.

- Once the optimal number N of EOFs is known, the whole procedure is repeated, now including the data set aside for cross-validation, but only with the N first EOFs considered as optimal. Final values for the missing data are then computed.

This is a general description of how the method works. For a more detailed description, see Beckers and Rixen (2003).

In the present work, we apply the reconstruction method for missing data DINEOF to a large matrix. We use a Lanczos method (see, *e.g.* Chatelin, 1993; Toumazou and Cretaux, 2001) for the EOF decomposition phase to make the application of DINEOF effective when working with large matrices. The desired characteristic of the EOF decomposition algorithm is the possibility to compute only the k largest EOFs at a small computational cost, since it must be used several times during the DINEOF iterations.

Toumazou and Cretaux (2001) have compared three different EOF decomposition methods, one based on the SVD algorithm, and two that express Eq. 8.1 as an eigenvalue problem: the QR strategy and the Lanczos method. They have shown that a Krylov-type method, called the Lanczos method, is a good choice when using large matrices. The EOF analysis performed by the mentioned methods gives similar results, although the Lanczos method requires half the storage memory than the others, and up to 22 times less computational time for large matrices. Another attractive characteristic of the Lanczos method is that it does not need to compute all the singular values, only the k largest ones. For these reasons, the Lanczos solver provided by Toumazou and Cretaux (2001), which uses the ARPACK freeware (Lehoucq *et al.*, 1997), has been implemented in DINEOF.

The main characteristic that makes the Lanczos method suitable when dealing with a large matrix is that, instead of working with the $n \times n$ matrix $\mathbf{A}$ of Eq. 8.3, a $p \times p$ ($p << n$) tridiagonal matrix is used, obtained by the projection of $\mathbf{A}$ onto the sub-space Krylov $K_p(\mathbf{A}, q)$ (Chatelin, 1993). The eigenvalues are calculated in this reduced matrix, in an iterative way until a convergence criterion is satisfied. This stopping criterion is based on the backward error (Bennani and Braconnier, 1994). For a more detailed explanation of the Lanczos method, see, *e.g.*, Chatelin (1993) and Toumazou and Cretaux (2001).

8.2.1 Code performance

A comparison between several codes that calculate EOF decomposition has been already made by Toumazou and Cretaux (2001) and is not the aim of this work. However, several tests were carried out to test the EOF reconstruction code performance. A Silicon Graphics MIPS R12000, 400MHz. is used to run DINEOF. In Table 8.1 we show the computation time of the Lanczos method for EOF decomposition (*i.e.* Lanczos method called once), calculated with several subsets obtained from the initial data. For a matrix of dimensions $m \times n$, the cost of the Lanczos method is proportional to $m^{1.23}n^{1.5}$. The total computational time for DINEOF is also presented in Table 8.1, being the cost of the total reconstruction method $m^{1.25}n^{1.17}$, with, always, $m > n$. As can be seen, for the larger matrix studied (94755×135) a total time of 119 minutes is necessary, and for the smallest case (8281×20), the whole procedure is finished in less than a minute. In an attempt to further improve computational time, interesting for very large matrices reconstruction, we have introduced different convergence criteria in the calculation of the singular values, as well as different criteria for the iterations made when searching the convergence of the singular values. We tested different values for missing data initialisation, stopping criteria to find the optimal number of EOFs, and the control of the iterations made by the Lanczos method. Only slight changes in computational time were attained, accounting for up to 10% of saved time.

8.3 Reconstruction of the Adriatic data set

8.3.1 Data Set

A set of 135 Advanced Very High Resolution Radiometer (AVHRR) images of the Adriatic Sea obtained from `http://radlab.soest.hawaii.edu/avhrr/adria/cdrom/html/` (Dousset *et al.*, 1998) is used to test the DINEOF method. The images range from 9 May 1995 to 22 October 1995. Their size is 248×709 pixels, with a resolution of 1.25 km. Land points are not used in this method, so the final spatial size is 94755 (out of 175832). In order to minimise skin temperature effects due to diurnal heating, only nighttime images are used. The mean cloud coverage is 61%. Some of these images present extreme cloud coverage (more than 95%) and have been eliminated in order to obtain reliable

Table 8.1: Time (in s) for EOF decomposition with the Lanczos method, and for the whole reconstruction process. The size of the matrix is $m \times n$, with m the spatial dimension and n the temporal dimension.

m	n	time (SVD decomposition)	time (reconstruction)
94755	135	27.96	7148.72
94755	105	20.31	5389.77
94755	75	15.19	3992.92
94755	50	13.46	3589.61
94755	20	1.72	822.3
39616	135	15.94	3846.17
39616	105	10.53	2182.37
39616	75	8.11	1941.81
39616	50	3.5	1018.5
39616	20	0.62	307.83
23205	135	6.85	1597.34
23205	105	5.0	1026.96
23205	75	2.68	668.25
23205	50	1.91	385.32
23205	20	0.34	138.2
8281	135	1.39	371.7
8281	105	1.17	283.51
8281	75	0.84	209.42
8281	50	0.47	133.69
8281	20	0.12	49.12

results. Images containing less than 5% of data do not provide useful information, and might affect the final result. A subset of 105 images (hereafter called the 'Complete Set' for clarity) with less than 95% of cloud coverage is kept for the analysis. The mean cloud coverage of this data set is 52%. In Figure 8.1 we can see that the distribution of the cloud coverage in the Complete Set is very homogeneous, with a slightly higher cloud concentration in the northern Adriatic and in the southwestern part of the basin. The coastline has extreme cloud coverage, probably due to slight errors that occurred when treating the images at these zones. This point highlights the importance of a correct image processing to avoid high data loss. Moreover, clouds in the satellite images must be precisely identified to obtain correct results in the reconstruction. Otherwise, artificial features may appear in the results. This is why a good cloud detection algorithm must be used. For the Adriatic data set, three cloud detection tests were conducted with nighttime images (see `http://radlab.soest.hawaii.edu/avhrr/adria/cdrom/html/index.htm` for detailed information on the cloud detection algorithm). These tests aimed to 1) detect temperature variations due to convective clouds (clear water pixels present very uniform brightness temperature, in contrast to clouded pixels); 2) compare to MODB climatology (Brasseur *et al.*, 1996; Rixen *et al.*, 2001); the maxima and minima of this climatology were increased and lowered by 1°C, and pixels falling outside this range were eliminated; and 3) detect stratiform clouds (based on differences between the infrared channels). These tests followed the work by Saunders and Kriebel (1988). However, as we will show, the reconstruction results in a filtering of the original data, which can correct some of the artificial features that were not detected by these tests (see for example, Figure 8.5).

8.3.2 Satellite observations in the Adriatic Sea

Remote sensed data has been used to study the Adriatic Sea in different works (*e.g.* Gacić *et al.*, 1997; Borzelli *et al.*, 1999; Cushman-Roisin *et al.*, 2001; Mauri and Poulain, 2001; Alvarez, 2003). A large number of features can be identified from the SST field. For example, the signature of the Western Coastal Layer (WCL), formed to the south of the Po delta and related to the WAC. Filaments advecting cold waters from the eastern coast (detaching at the Istrian Peninsula and at southern positions along the Croatian coast, between 43° and 44°N) to open seawaters have also been detected from satellite images (Borzelli *et al.*, 1999). These filaments can extend up to a hundred of kilometres, with a width of about 10-20 km, and have been observed in summer. Mauri and Poulain (2001) have also observed cold waters (21-23°C) in the nearby the Croatian coast in September and October. The signal of the Po River can be identified from satellite images, as it can spread over the entire northern basin in the summer season. The Po River water is slightly warmer than seawater in summer, so other variables, such as chlorophyll, act as better tracers of the Po plume than temperature. In winter the plume is weaker, the water flows mainly southward and the Po River water is several degrees colder than seawater. The work by Gacić *et al.* (1997) presents a complete analysis of the seasonal and interannual variability of the surface temperature of the Adriatic Sea.

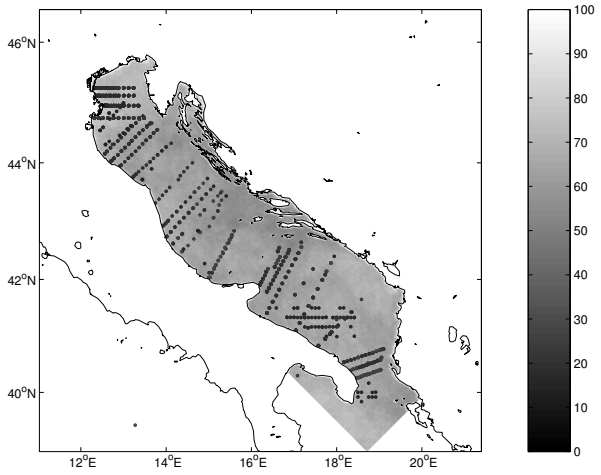

Figure 8.1: Mean percentage of cloud coverage for the Complete Set of images. The points show the distribution of the *in situ* data obtained from the MEDAR/Medatlas database used for validation of the reconstruction. A total of 452 stations were extracted from this database.

Studies of this kind could benefit from any method recovering missing information due to clouds.

8.3.3 Reconstruction results

The first test we conducted is the reconstruction of the Complete Set. DINEOF keeps 10 EOFs as optimal for reconstruction (number of EOFs that minimises the error in cross-validation), as can be seen in Figure 8.2, which minimise the expected error to 0.6°C. These 10 EOFs explain 98.97% of the initial variance, which has been calculated as the variance of those points not covered by clouds.

In Figures 8.3 to 8.5 we can see three examples of the quality of DINEOF results. They show three original images, with blanks where there are no data, and their reconstruction. Figure 8.3 is of 23 July, and we can see the warm plume of the Po River, which reaches 32°C and spreads over a large surface in the northern Adriatic and flows southward following the west coast, forming the WCL. We can also see a weak cold filament going from the Istrian Peninsula (in the east coast) to open sea. Finally, at the southeast, on the Albanian coast, there is a strong cold signal. This kind of cooling event has been reported by *e.g.* Bergamasco and Gacić (1996) and Gacić *et al.* (1997), and the cause may be the cooling and mixing due to the Bora wind that blows parallel to this coast. This image will be further commented later in this work. Figure 8.4 corresponds to 3 September. In this image, a strong cold filament can be observed at the same location as in Figure 8.3. This filament has a temperature of about 20°C, which corresponds to the situation described by Mauri and Poulain (2001) for September-October. In

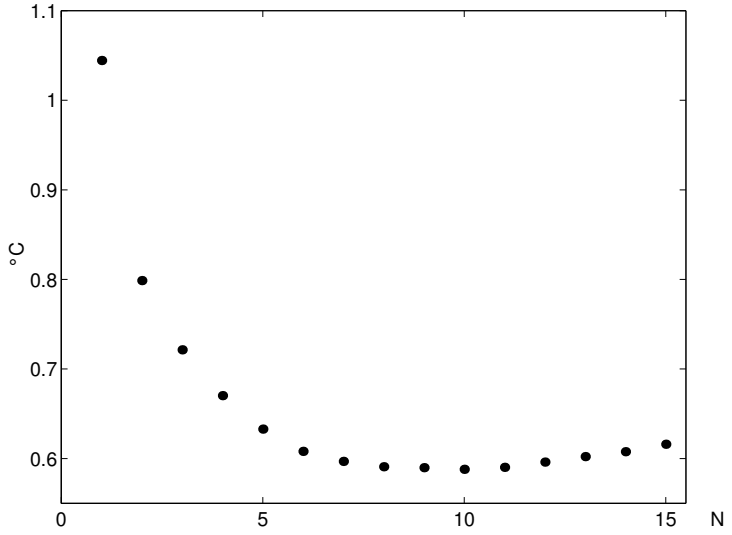

Figure 8.2: Error obtained with cross-validation for reconstruction of the Complete Set. N is the number of EOFs.

Figure 8.5, of 5 October, a strong cooling event occurs in the northern Adriatic, where temperature decreases to 18°C. This situation is typical of autumn, when temperatures begin to decrease in this zone (Gacić *et al.*, 1997). In this image, we clearly see the effect of filtering of the reconstruction method, due to the rejection of some EOFs by cross-validation: in the cloudy image, some cold pixels are present at the limits of clouds in the centre of the Adriatic. The gradient with surrounding waters suggests that it is not a real feature, and that some cloud pixels may have been interpreted in the initial data set as sea pixels. In the reconstruction these features have been eliminated, and the result is a more realistic representation of surface temperature. Also real features, such as currents or meanders, are smoothed in the reconstructed data. Preventing the smoothing of real features by DINEOF is currently under study. The whole series of 105 images reconstructed by DINEOF, as well as the other results presented later in this work can be seen at `http://modb.oce.ulg.ac.be/alvera`.

In Figure 8.6 a time series of a random point (point located at a latitude of 42N and a longitude of 17.3E, situated in the centre Adriatic, south of the Croatian Islands) in the Adriatic Sea is shown, with the reconstructed and the original values, including the gaps. As we can see, the reconstruction process smoothes the original data. The mean temperature of the initial matrix, calculated over existing points, has been also plotted. The reconstructed series shows some spikes that are also present in the initial data, as for example, around the 10 September. This may indicate that there are some days that show smaller temperature in the whole basin, and this effect is also reflected in the reconstruction.

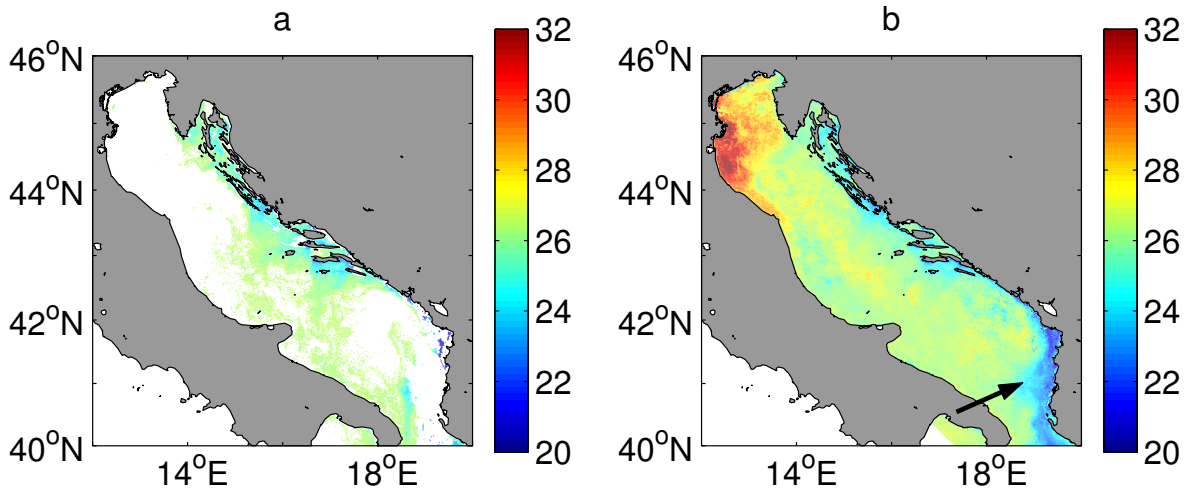

Figure 8.3: Original cloudy image (a) from the Complete Set for 23 July and its reconstruction (b). A warm Po plume is clearly seen, as well as cold waters along the Albanian coast, marked with an arrow.

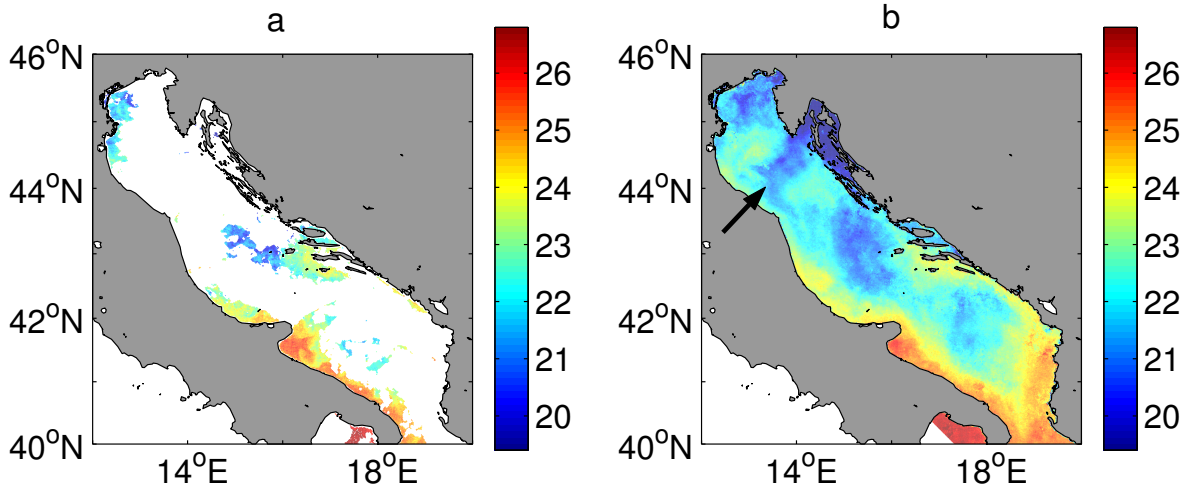

Figure 8.4: Original cloudy image (a) from the Complete Set for 3 September and its reconstruction (b). The arrow shows a strong cold filament detaching from the Istrian Peninsula.

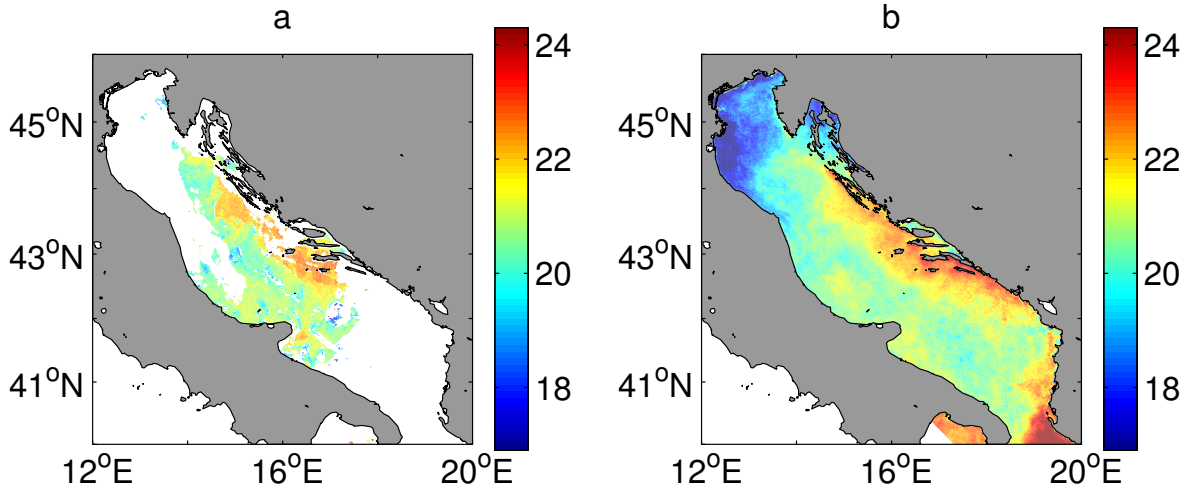

Figure 8.5: Original cloudy image (a) from the Complete Set for 5 October and its reconstruction (b).

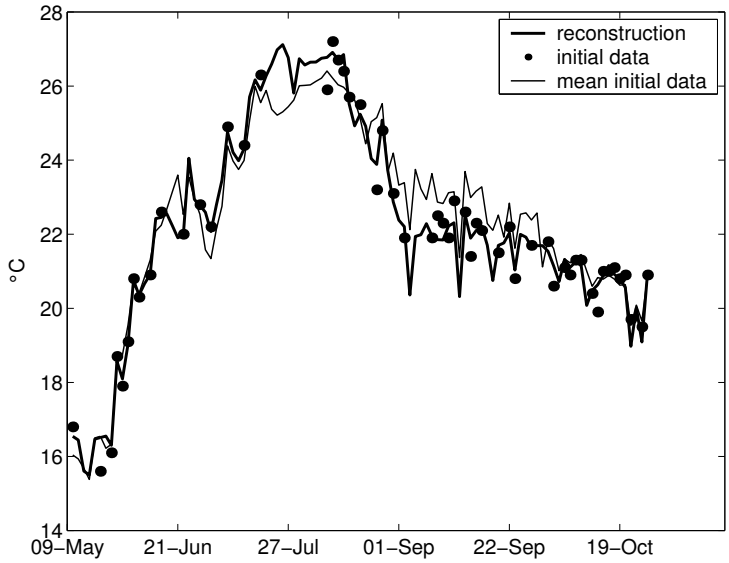

Figure 8.6: Time Series of a random point (point at [42N,17.3E]) in the Adriatic Sea. Solid coarse line represents the reconstructed field, • represents the original data, with gaps where there are no data, and thin solid line represents the mean initial temperature

8.3.4 Validation

To establish the efficiency of DINEOF, a third data set extracted from the previous one is used. In this data set, only the cleanest images are taken into account. The result is a sequence of 15 images with mean cloud coverage of 18%. Extra cloud coverage is then added to this data set, so that we can then compare the reconstruction to the original images. Clouds are extracted from other images from the complete data set, so that the coverage presents a realistic aspect. The extra clouds added increase mean cloud coverage up to 40, 60 and 80%, in order to construct three different data sets, each with different amounts of cloud coverage.

Continuity in time of the images is not necessary to effectuate the reconstruction. Good results can be obtained with irregularly time-distributed data. However, if the set of images is too sparse in some periods (*e.g.*, only one or two images in summer and the rest in autumn), the reconstruction will be deteriorated by irrelevant EOFs describing the variability of a different period of the year, since the information for the former period is insufficient. To avoid this problem, the set of 15 images contains only data from September and October.

An example of reconstruction of these three sets can be found in Figures 8.7 to 8.9 which correspond to 16 October 1995. The reconstructed image from the 40% of added clouds (Figure 8.7), shows a good agreement with the original field, free of clouds, as well as the image that was covered with 60% of clouds (Figure 8.8). The main physical features are realistically represented, such as cold temperatures south of the Po River and the warm current entering by the Strait of Otranto and following the east coast. Temperature distribution is maintained. Figure 8.9 is a little bit noisier, due to the extreme cloud coverage of this data set, 80%. In this last figure, the maximal temperatures are slightly weaker in the reconstruction than in the original field, although the temperature distribution is well represented (again, the warm current entering the Adriatic Sea by the Strait of Otranto is maintained).

In Figure 8.10 the difference between the reconstruction of the 40% added clouds 8.7c and the initial clean image 8.7a is shown. The difference between them is very small, with the higher values, as it would be expected, in the zone where clouds were added, although there are differences of the same magnitude in the centre Adriatic.

In these tests, only 2 EOFs were retained for the reconstruction (minimum error obtained by cross-validation, see Figure 8.11). The data loss due to added clouds results in a decrease of variance on the reconstructed images. The variance of the three data sets (with 40%, 60% and 80% of cloud coverage) is 94%, 90%, and 80% of the original data variance respectively. Note that the variance has been calculated in relation to the images where extra clouds were added.

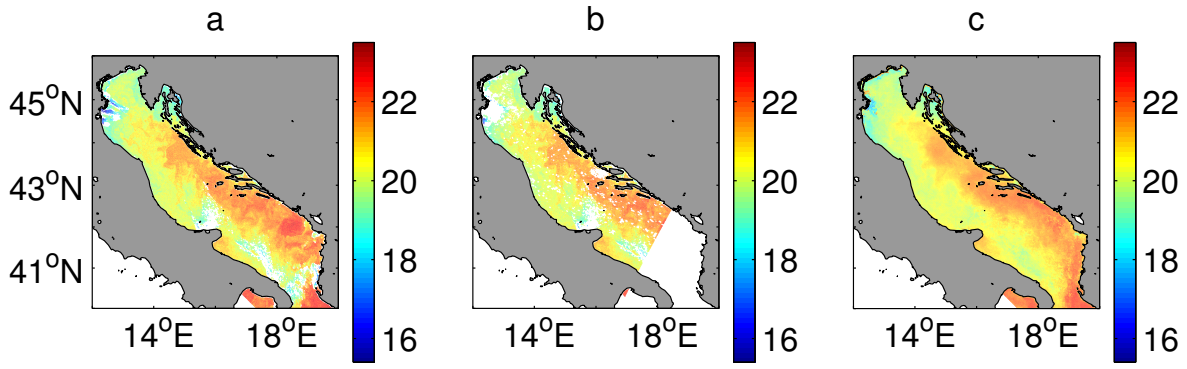

Figure 8.7: Image of 16 October. (a) is the original image, with low cloud coverage; (b) is the same image with 40% extra cloud coverage added; (c) is the reconstruction of the image (b).

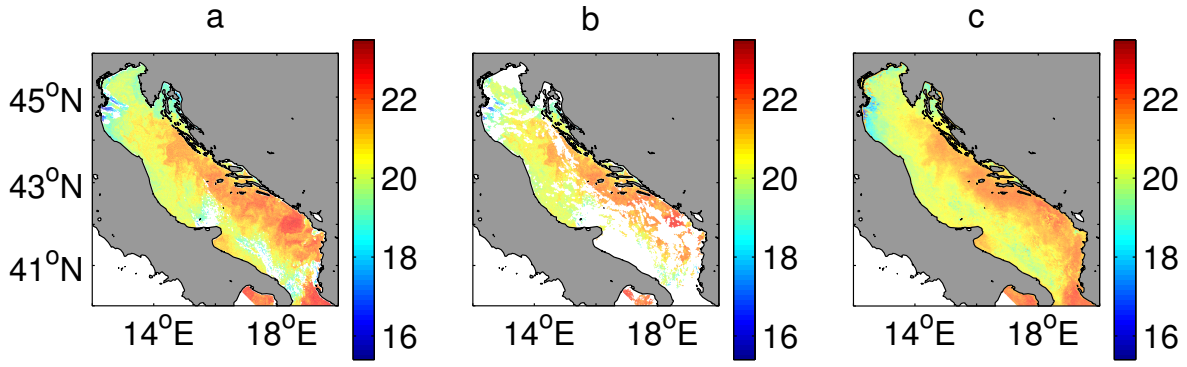

Figure 8.8: Image of 16 October. (a) is the original image, with low cloud coverage; (b) is the same image with 60% extra cloud coverage added; (c) is the reconstruction of the image (b).

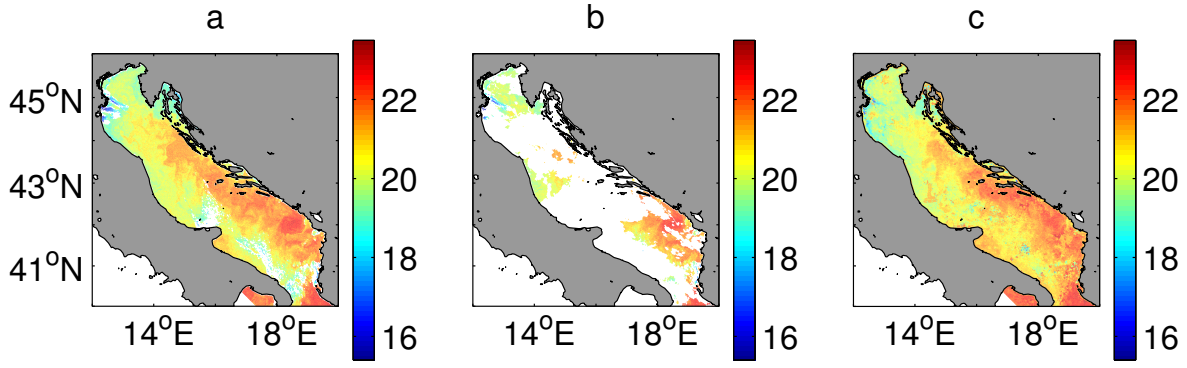

Figure 8.9: Image of 16 October. (a) is the original image, with low cloud coverage; (b) is the same image with 80% extra cloud coverage added. (c) is the reconstruction of the image (b).

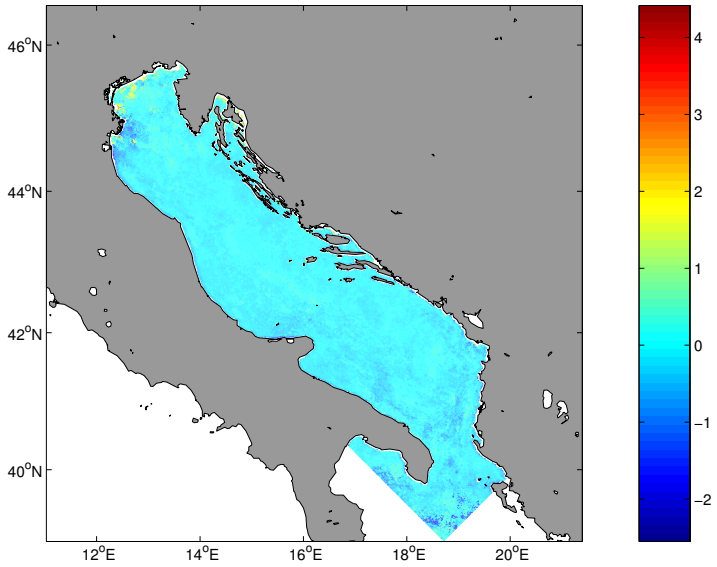

Figure 8.10: Difference between the 40% extra clouds reconstruction and the initial clean image.

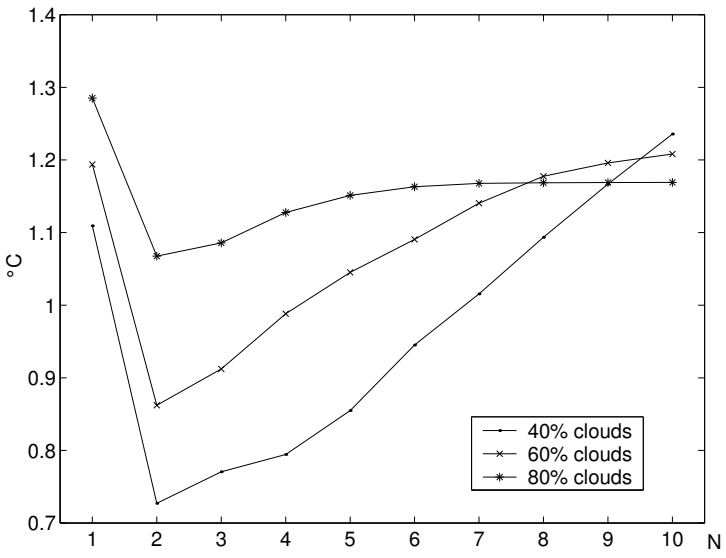

Figure 8.11: Error obtained with cross-validation for the validation test. N is the number of EOFs.

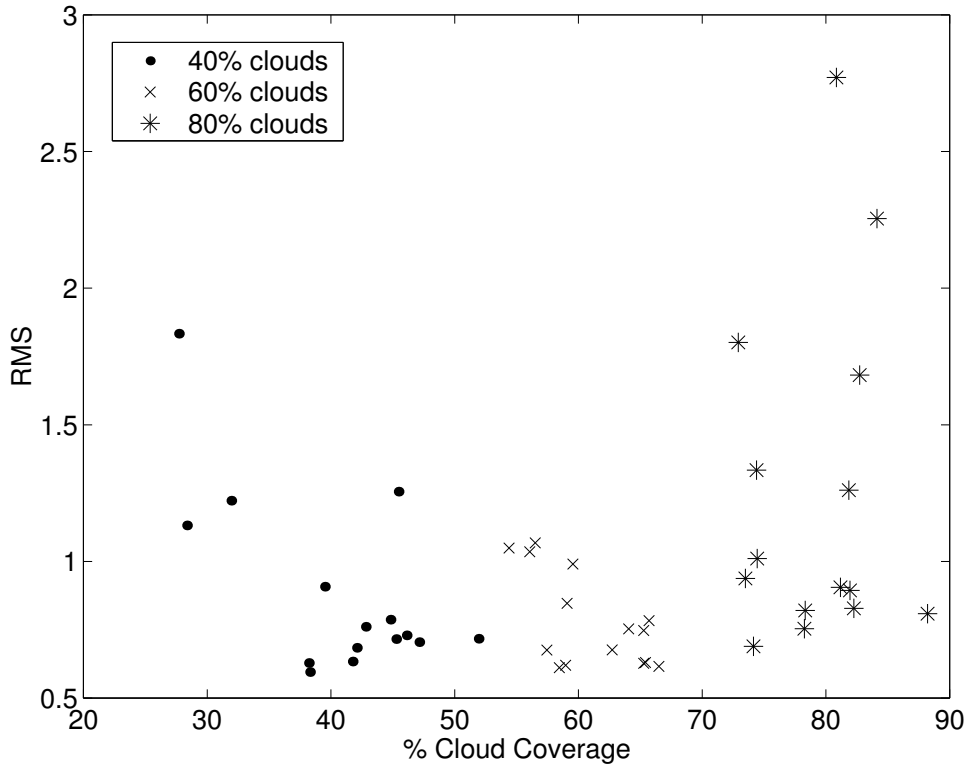

Figure 8.12: RMS error related to cloud coverage.

Figure 8.12 shows the Root Mean Square (RMS) error between the three reconstructed subsets of 15 images and the original one, related to cloud coverage. Error increases with cloud coverage, but not very much. As can be seen, there are points of near 90% coverage that show only errors of 0.7°C, while other points show higher errors and smaller cloud coverage. The RMS error for each data set is about 0.89°C, 0.78°C and 1.25°C for the 40, 60, and 80% of added cloud coverage respectively, which is comparable to the error estimate obtained with cross-validation (Fig. 8.11).

It may appear surprising that the set of 40% added cloud coverage presents a slightly higher error than the set of 60% added clouds. This is simply due to the fact that the added clouds are 'real' clouds, *i.e.*, they are taken from other images of the data set, so they are not homogeneously distributed, as would be obtained with an artificial random coverage. As a result, the set with 40% extra cloud coverage has zones where cloud coverage is very high, while clouds in the 60% added clouds set are more homogeneously distributed. As this situation is certainly possible in reality, we decided not to modify the cloud distribution, although it does have to be considered when looking at the obtained RMS error.

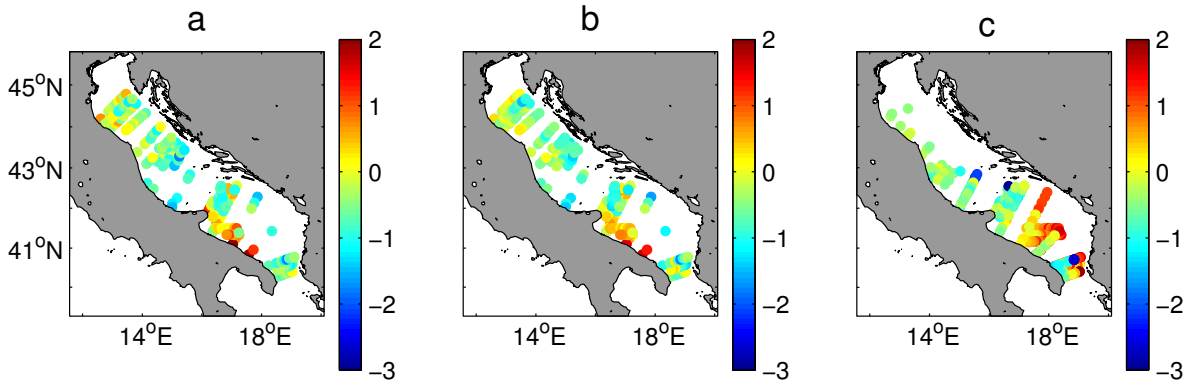

Figure 8.13: Difference between the satellite images and *in situ* data from MEDAR/Medatlas database. a) shows this difference for initial non clouded points, before reconstruction. b) shows this difference for initial non clouded points, after reconstruction. c) shows this difference for initialy clouded points, reconstructed by DINEOF

Independent data

In situ data from the MEDAR/Medatlas database (MEDAR-Group, 2002), covering the period of the Complete Set, have been used as a source of independent data for the validation. In Figure 8.1 we can see the distribution of the stations extracted from MEDAR/Medatlas. A total of 452 observations taken at these stations are considered. Surface points from MEDAR/Medatlas database are warmer than the satellite images used here, mostly in summer, due to daily reheating. As the SST images used in this work are night images, the data from MEDAR/Medatlas have been taken at 5-m depth in order to avoid skin temperature effects. The error between *in situ* and reconstructed data is minimised at this depth. The difference between MEDAR/Medatlas data and the reconstruction of the Complete Set is presented in Figure 8.13. In this figure we can see the difference between MEDAR/Medatlas data and: a) the original points (*i.e.* not covered by clouds) on the SST data set before applying DINEOF; b) these same points after the reconstruction; and c) those points that were missing in the initial data and whose value have been obtained using DINEOF. The total RMS error for the original satellite data (before DINEOF) is 0.71°C, which we can consider as the error of the satellite measures in relation with *in situ* measures. Other works have found similar errors in satellite data (*e.g.* Wick *et al.*, 1992). After the reconstruction, the error of these points is 0.67 °C. This reduction in error is due to the fact that the truncate series of EOFs used for the reconstruction does not contain noise that may exist in the initial data. For the clouded points that have been reconstructed, the RMS error is 0.95°C, *i.e.*, only 0.28°C higher than the real initial points. The difference between the satellite data and *in situ* data is higher in southern stations, with the satellite data generally warmer than *in situ* data.

Table 8.2: Main parameters of the OI reconstruction

Number of influential points	20
Radius of influence	25 km
Influential time window	6 days
Observation error	0.1°C
Correlation, zonal zero crossing	60 km
Correlation, meridional zero crossing	60 km
Zonal decorrelation (decay) scale	40 km
Meridional decorrelation (decay) scale	40 km
Temporal decorrelation (decay) scale	6 days

8.3.5 Comparison with an Optimal Interpolation method

The Harvard Ocean Descriptive Predictive System (HOPS) OI package (Davis, 1985; Robinson and Leslie, 1985; Carter and Robinson, 1987) has been used to reconstruct the Complete Set in order to compare this reconstruction to DINEOF. The HOPS system is based on a minimisation of a pre-selected error norm, chosen to be the mean square difference between the estimate and the true value of the field. A multivariate space-time objective analysis scheme is used to produce regularly gridded fields in time and space (analyses) from the initial data (*i.e.*, irregularly distributed observational data). In this case, the clouded images are used as initial data. The main parameters used in this reconstruction are summarised in Table 8.2.

The results of reconstruction by the OI method have been compared to MEDAR/Medatlas *in situ* data. The OI approach has a high computational cost, and to obtain results similar to DINEOF, a total computational time of 40 hours 30 minutes is necessary. It must be said that for this OI reconstruction, only one point in 100 was used, because the reconstruction of the total Complete Set with this method would have taken too long. The RMS error between initially non-clouded points and *in situ* data is 1.78 °C. For clouded points, the RMS error is 2.4°C. The OI reconstruction thus presents a higher error than DINEOF (errors given in Section 8.3.4). Of course, a higher number of points could have been used for the reconstruction, thus obtaining a smaller error. As the test realised here was at the limit of the computer resources, we were unable to take a higher number of points.

The OI results can be improved by a fit to the initial data. The procedure explained by Smith *et al.* (1996) is followed. First, the 15 most dominant EOFs are calculated from the OI reconstruction and then the amplitudes are fitted, in the least-squares sense, to the initial existent data. The RMS error is now 0.89°C for initially clouded points,

and 0.65°C for non-clouded points. This improvement in the OI reconstruction is easy to obtain and entails only small additional computational cost. The errors of the fitted reconstruction are thus similar to those obtained by the DINEOF reconstruction, with a difference in time required to calculate them.

An additional test was carried out, to determine what the behaviour of the OI procedure would be if the same computational time as used for DINEOF were allotted, *i.e.*, one hour and a half. To stay within this time, only one point out of 1600 could be taken, otherwise the allowed computational time would have been exceeded. The reconstruction obtained in this case with the OI procedure is very poor, and the mean value of the initial existing points is usually recovered as the reconstruction, due to the lack of influential points.

8.3.6 Moving Patterns

Extended EOFs (ExEOFs), first introduced by Weare and Nasstrom (1982) are an extension of the classical EOFs. The reconstruction of a complex geophysical field evolving in space and time can be improved if its time correlation is considered. ExEOFs extract space-time (moving) patterns of data by taking both the spatial and temporal correlations into account (Kantha and Clayson, 2000; Jolliffe, 2002; Kim and Wu, 2000). In the satellite images used in this work, we can see a variety of small-scale features, such as eddies or filaments, that change their shape and position in space with the evolution in time. The ExEOF can ameliorate the representation of such features by using a lagged version of the analysed matrix $\mathbf{X}$ (Weare and Nasstrom, 1982):

$$\mathbf{X}_{i,j} = \left(\begin{array}{cccc} X_{1,1} & X_{1,2} & ... & X_{1,N} \\ X_{2,1} & X_{2,2} & ... & X_{2,N} \\ X_{3,1} & X_{3,2} & ... & X_{3,N} \\ ... & & & \\ X_{M,1} & X_{M,2} & ... & X_{M,N} \\ \hline \\ X_{M+1,1} & X_{M+1,2} & ... & X_{M+1,N} \\ X_{M+2,1} & X_{M+2,2} & ... & X_{M+2,N} \\ ... & & & \\ X_{2M,1} & X_{2M,2} & ... & X_{2M,N} \end{array} \right) \tag{8.6}$$

$\mathbf{X}$ has now a size of $2m \times n$. By taking into account both the spatial and temporal correlation of the data, the ExEOFs resolve moving patterns in a more accurate way (von Storch and Zwiers, 1999; Ghil *et al.*, 2002; Jolliffe, 2002; Kim and Wu, 2000). A second positive aspect for using ExEOFs is that a missing data point can be reinforced using

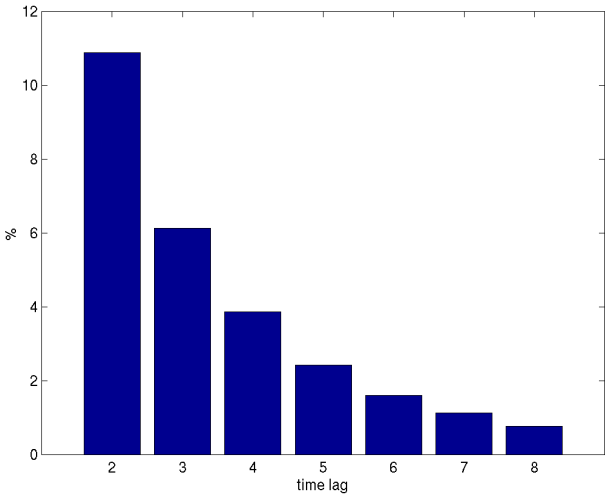

Figure 8.14: Percentage of points that, covered one day, are still covered during a given time lag (in the x-axis).

additional information of a future or past non-missing point. The obtained results show that the application of ExEOFs to the Adriatic data set gives more accurate results.

A small test has been carried out, to see the impact of ExEOFs in the reconstruction of the Adriatic data set. The time lag should be chosen in a way that it optimises the reconstruction. A compromise must be found; if the time lag is too short, there is a high probability that the lagged point is cloudy, and that will add no additional information for the reconstruction; on the opposite, if the time lag is too large, the time correlation between both parts of the matrix will be very small. Figure 8.14 shows the probability that a point covered one day is still covered p days after. As can be seen, the probability that a covered point is still covered the previous day (time lag of 1) is near 12%. Of course, this probability decreases, and the probability that a covered point is also covered 5 days later is less than 2%.

We have made eight different reconstruction tests of the Adriatic data set with ExEOFs, using a time lag that varies from one to eight days. Each of these reconstructions have been compared to MEDAR/Medatlas *in situ* data, as it has been done in section 8.3.4, with classical EOF reconstructed data. In Figure 8.15 we can see the RMS error for each of these eight tests. Initially cloudy points have been examined separately from initialy clean points. A black line in each of these plots indicates the error obtained using classical EOFs that we want to improve with ExEOFs. For a time lag of seven days, a decrase of the RMS error is observed for both cloudy and clean points. Other time lags improve only one of the two subsets, but increases the error of the other one.

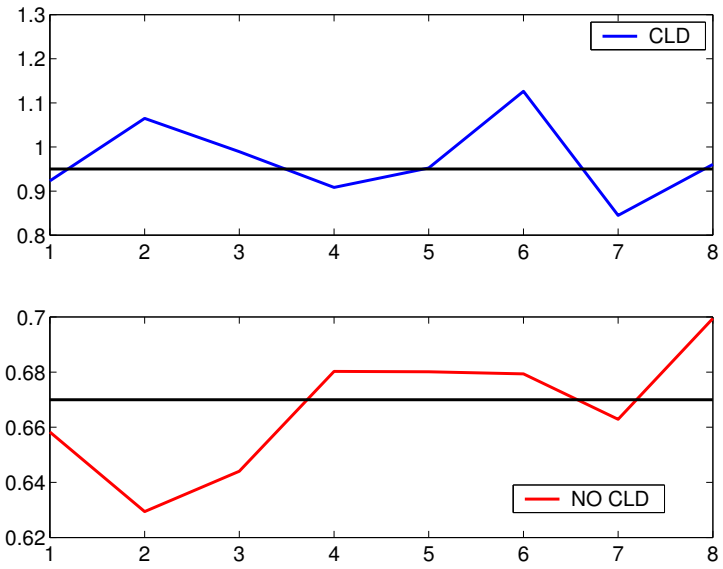

Figure 8.15: RMS error obtained when using ExEOFs with a time lag of one to eight days. Upper pannel shows the RMS for cloudy points, and black line represents the error obtained using classical EOFs. In the lower panel, the RMS error of initially non cloudy points, and a black line marks the RMS obtained with classical EOFs. Going below this line indicates improvement of the ExEOFs over the classical technique.

Finally, a time-lag of seven days was chosen. The decrease in the RMS error is not very high. However, there are some improvements in the reconstruction that show the beneficial impact of the ExEOFs approach. One example is shown in Figure 8.16. In this figure we can see a sequence of the initial cloudy images (in the left of the figure) and the reconstruction with classical EOFs and with ExEOFs (in the middle and right parts of the figure, respectively). The three images correspond to three consecutives images of the data set, more precisely, days 28 September, 5 October and 6 October (they are consecutive in the matrix, since there is no data from 28 September to 5 October). As we can see in this figure, the classic EOF reconstruction presents a cold water zone around the Po River the 5 October, although the days before and after the reconstruction does not show this cold water zone. ExEOFs present a smoother transition from 28 September to 6 October, and the cold water zone is no longer present. This last reconstruction is physically more coherent, since those rapid temperature changes are not observed. We can conclude that the use of ExEOFs slightly improves the reconstruction of the matrix, although further work should be done in this line.

8.3.7 An application: EOF analysis

In this section we will use the reconstructed matrix to realise an EOF decomposition. We aim to show the quality of the reconstructed images with this example. Any kind

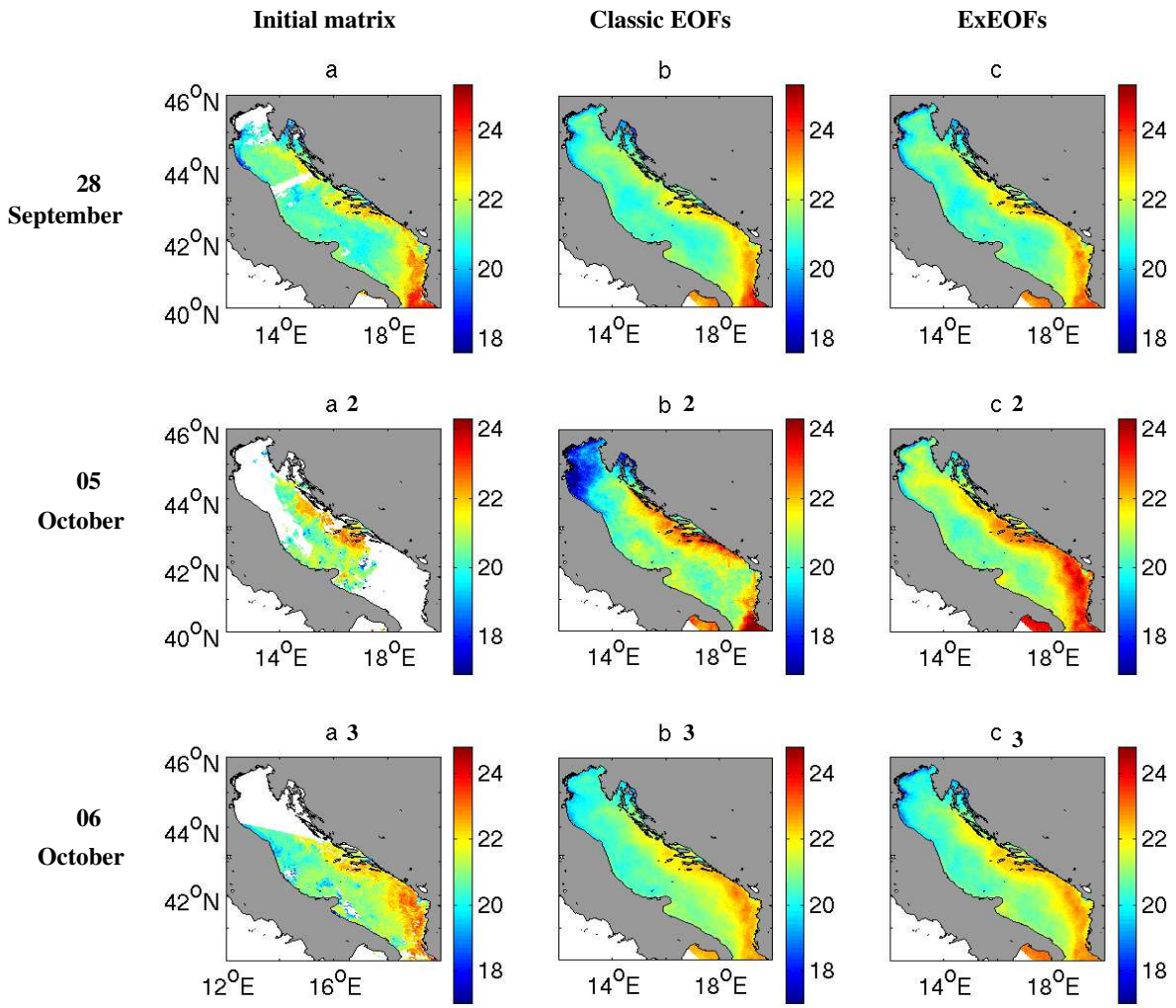

Figure 8.16: Initial SST matrix of the 28 September (a), 5 October (a2) and 6 October (a3). The reconstruction made with classic EOFs is shown in (b), (b2) and (b3) for those three days respectively, and the reconstruction with ExEOFs is shown in (c), (c2) and (c3)

of study can be done with the reconstructed data set, as there is no longer a coverage limitation. An EOF study will show in addition some interesting aspects of the SST distribution in the Adriatic Sea. The first three EOFs are presented, and we pay special attention to a cold event that occurred on the Albanian coast in July (Figure 8.3).

Figure 8.17 shows the first three spatial modes. These three modes account for 94% of the variability (84.6%, 7.2% and 2.2% respectively). The first mode (Figure 8.17a) is positive almost everywhere in the basin, indicating a general warming or cooling of the Adriatic Sea. The temperature distribution corresponds to that described as the typical situation in the Adriatic Sea. Indeed the east coast presents smaller values than the west coast. The zone surrounding the Po River has the strongest values, indicating that the influence of the Po River is very high. When looking at the first temporal mode (Figure 8.18a) we see the seasonal cycle, with a warming trend when approaching summer and a cooling when approaching autumn. The first mode thus represents the seasonal cycle ranging from May to October.

In the first and second modes, there is a clear signal of the EAC entering the Ionian Sea. The different time evolution of both modes contributes to the modulation of this current. The first mode reaches a maximum when the second one has its smallest values, *i.e.* at the end of July. Then the second mode presents its highest values while the first one decreases towards zero.

The second spatial mode (Figure 8.17b) also presents the WAC exiting the Adriatic Sea. We can see the signal of the filament detaching from the Istrian Peninsula. In the second temporal mode (Figure 8.18b), we can see that the fluctuations reach their highest values from 1 September to the end of the record. The EAC is more intense in autumn (Artegiani *et al.*, 1997b; Cushman-Roisin *et al.*, 2001; Poulain, 2001), and when looking at the Complete Set, we can see that both currents, EAC and WAC, increase their intensity in this period. We can also observe a north-south division of the basin temperature in the second spatial mode, with the largest negative values on the Croatian coast, south of the Istrian peninsula. In summer, the temperature in the northern basin is higher than the temperature in the southern basin. The opposite occurs in winter (Cushman-Roisin *et al.*, 2001). As it can be seen when looking at the second temporal mode, the values from May to August are near zero or negative, which indicates that the northern basin is warmer than the southern one in the summer season, as the second spatial mode presents negative values in the north. In autumn, the second temporal mode reaches its highest values, indicating that the northern basin is then colder than the southern basin.

In the third spatial mode (Figure 8.17c), we can clearly see the signal of the Po plume, extending over a large zone. This situation is mainly found in summer, when winds decrease and the Po River water can spread over the entire northern basin. In winter, the plume is reduced, and winds force it to follow the west coast of the Adriatic Sea

(Cushman-Roisin *et al.*, 2001; Mauri and Poulain, 2001). This situation can be observed in the third temporal mode (Figure 8.18a), where we see a gentle decreasing trend all along the time series as autumn approaches. The water coming from the Po River is warmer than the surrounding waters from May to August. In September-October, when the temporal mode becomes negative, the Po River water is colder than that in open sea. This situation has been described by Gacić *et al.* (1997) and Cushman-Roisin *et al.* (2001). Again, the largest negative values are found on the Croatian coast. These waters are cold in summer, as we the temporal mode presents positive values. When the EAC inflow increases in September, the hot waters also affect this zone, which becomes warmer than in summer (negative values of the temporal mode). South of the Istrian peninsula, there is the signal of a filament with negative values.

One feature that attracted our attention is a peak that is present in the second and third temporal modes, occurring on 23 July. This abrupt change is negative in the second mode, and positive in the third one. This peak occurs on 23 July; this image has already been shown (Figure 8.3). In this figure, we can see a cold tongue of water along the Albanian coast, in the southeast of the Adriatic, which is also partially visible in the initial cloudy data set (Figure 8.4). This feature can be observed from 22 to 26 July. Bergamasco and Gacić (1996) and Gacić *et al.* (1997) describe the occurrence of cold waters in this zone, but its cause is not well documented. It may be an upwelling due to the action of the northeasterly Bora wind, when it blows parallel to the coast, or simply the vertical mixing induced by the wind. The study of the wind series for this period is thus important to understand this event. We used the ECMWF Re-Analysis (ERA) wind series from May to October. This data set covers the Adriatic Sea with a half-degree resolution in both latitude and longitude. As a result, a total of 14x13 points over the Adriatic are used. In Figure 8.19 the projection of this time series in the NE direction is shown in order to identify Bora wind. An arrow marks a peak occurring on 20 July, when the wind reaches high values in Bora direction. Looking at the wind distribution these days (Figure 8.20), we observe that for a period of six days, from 17 to 22 July, the wind blows constantly parallel to the Albanian Coast, with the maximum speed reached on 20 July. The speed ranges from 3 to 5 m/s, which is not considered as a strong Bora event. However, this speed is only exceeded on 9 October, so the event on 20 July is one of the strongest of the period May-October 1995. The figures shown suggest that the cold tongue observed at the Albanian coast is caused by the action of the Bora. The reconstructed data set very accurately reflects this event.

8.4 Reconstruction of the Mediterranean Pathfinder data set

The second experiment was carried out with a set of 63 SST Pathfinder v5 images of the whole Mediterranean Sea. The images can be obtained via ftp at

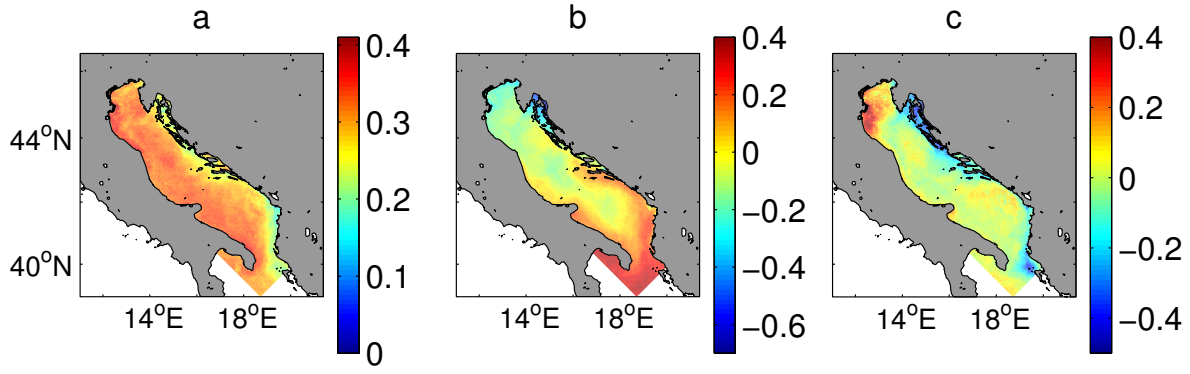

Figure 8.17: Spatial EOFs, modes 1 to 3 (a-c). Units are °C

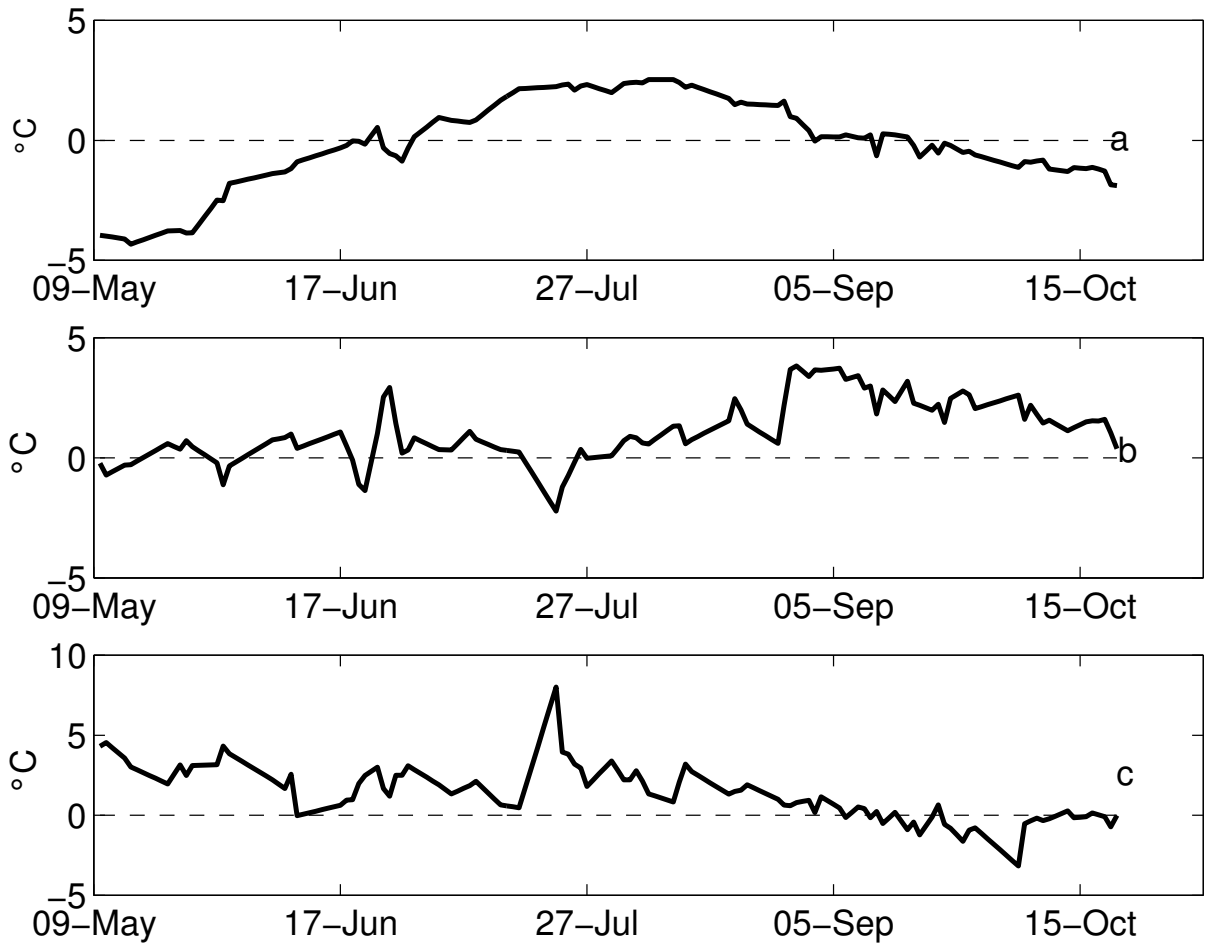

Figure 8.18: Temporal EOFs, modes 1 to 3 (a-c). Units are °C

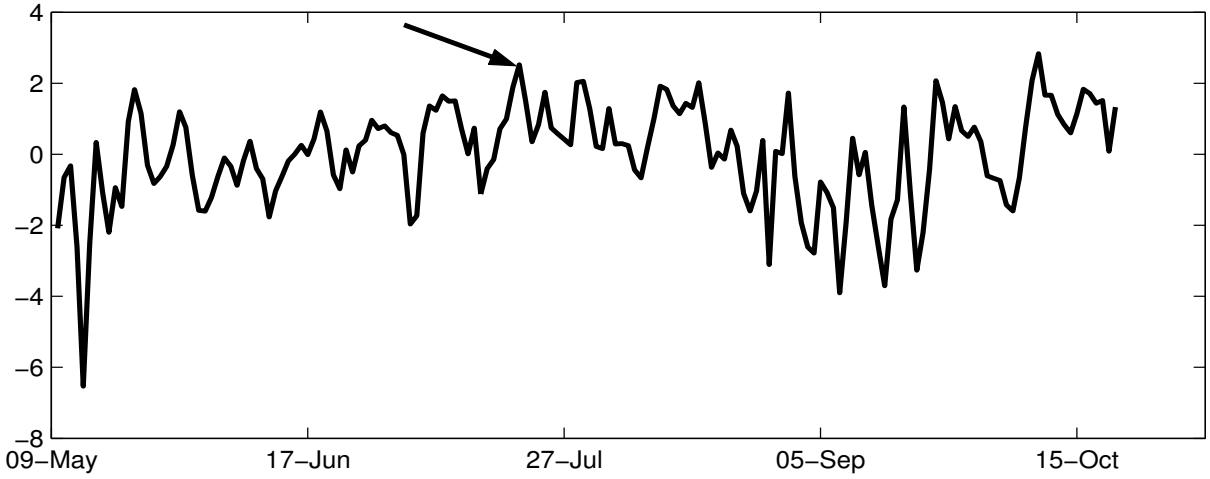

Figure 8.19: ECMWF wind series projected on the NE direction. The arrow shows the peak of wind speed on 20 July 1995.

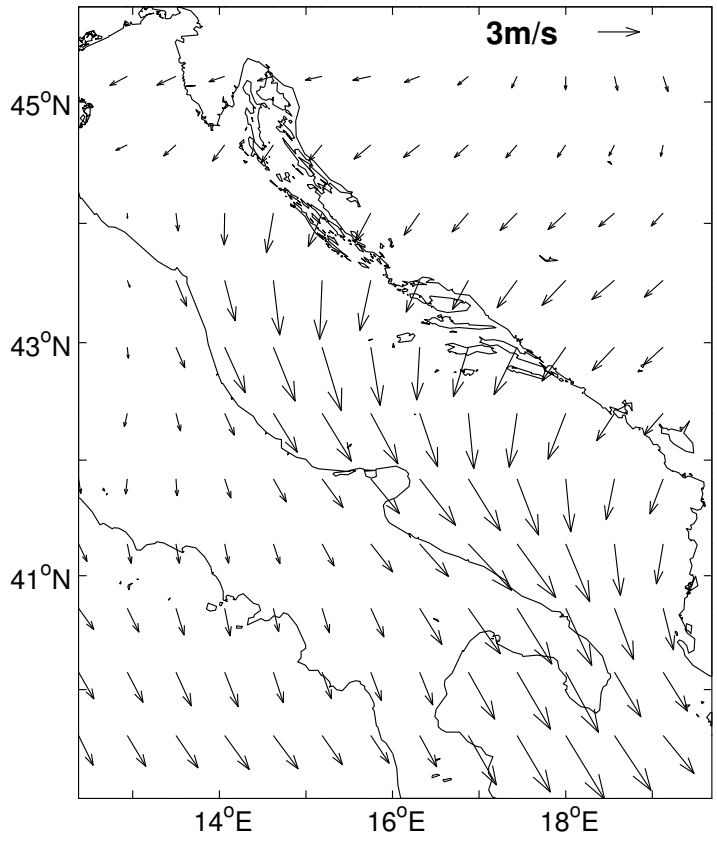

Figure 8.20: Wind mean distribution for the period from 17 July to 22 July 1995.

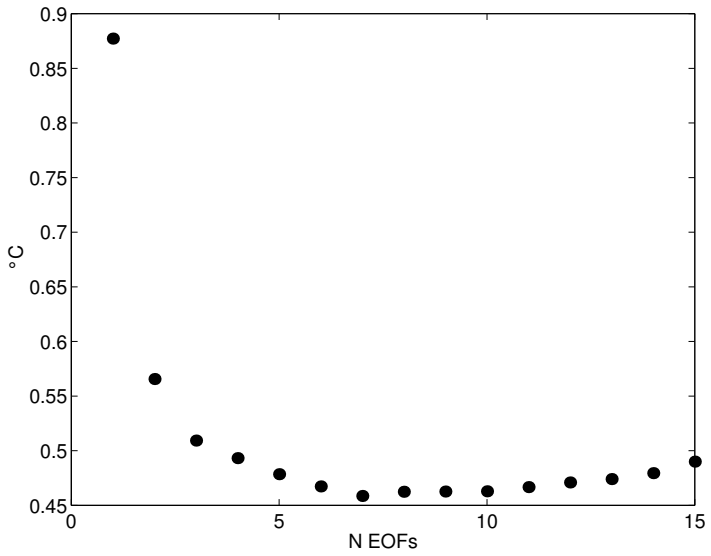

Figure 8.21: Error obtained with cross-validation for reconstruction of the Mediterranean data set. N is the number of EOFs.

`ftp://podaac.jpl.nasa.gov/pub/sea_surface_temperature/avhrr/mcsst/` and they have a resolution of 4 km. Only nighttime images were used, and they range from 3 July to 3 September, thus covering the period when the assimilation experiment described in chapter 3 takes place. Pathfinder SST v5 data set was already described in section 4.

Pathfinder v5 data has a pixel-by-pixel quality flag classification. Quality flags values go from 0 (the worst) to 7 (the best). For this work, all pixels corresponding to quality flag smaller than 5 were considered as missing data. With this choice there is a 38% of missing data in average.

DINEOF retained 7 EOFs for the reconstruction of the Mediterranean data set, which reduced the cross-validation error to 0.46 °C (see Figure 8.21). The computational time for the reconstruction was 120 min, and for a single Lanczos iteration, 18.82 s.

Figures 8.22 to 8.24 show some of the results obtained from the reconstruction of the Pathfinder v5 data set with classic EOFs. Figure 8.22 shows the 20 July data. In the initial cloudy image (Figure 8.22 upper pannel) we can see clouds covering partially the Meditarranean Sea, and also a missing satellite track in the Western Mediterranean. Recall that the classification between missing and non-missing data is made upon a threshold. If this threshold is decreased, some features will be classified as non-missing data. The reconstruction show some features, as the cold jet flowing eastwards south of Sicily and some meso-scale activity at the Eastern Mediterranean that was not clearly visible in the initial data. Figure 8.23 presents the initial data and its reconstrucion

for 5 August. This image is extensively covered by clouds, mainly in the Western Mediterranean and the Eastern Mediterranean. Only the central Mediterranean has a good coverage. The reconstruction of this image (bottom pannel of Figure 8.23) shows a good reconstruction of the main physical features present in the Mediterranean Sea, as cold waters in the Gulf of Lions, cold water entering the Mediterranean by the Strait of Gibraltar, cold water entering by the Bosphore Strait, or meso-scale eddies in the Levantine Basin. As the reconstruction of the whole series of images is done simultaneously, the information for the reconstruction of images such as the one presented in Figure 8.23 beneficiates from the information provided by previous and next images with smaller cloud coverage.

Finally, Figure 8.24 presents the reconstruction for day 27 August. In the initial cloudy image we can see a cold water band in the Western Mediterranean, from the French coast to the south-east, traversing the Tyrrhenian Sea. This band is possibily the result of clouds that have been considered as sea points. The reconstruction of this image shows smoother values at this zone, and the cold water band is no longer visible. Also the missing data in the Gulf of Lions have a cold edge that is smoothed in the reconstruction. This image shows again the benefit of the truncate EOF series used for the reconstruction. The whole reconstructed SST series can be seen at `http://modb.oce.ulg.ac.be/alvera`.

Pathfinder v5 reconstruction has also been validated with MEDAR/Medatlas *in situ* data. A total of 256 stations was used. 208 stations correspond to non-cloudy points, and only 48 to initially cloudy points. Their distribution can be seen in Figures 8.25 and 8.26, as well as the difference between MEDAR/Medatlas and Pathfinder data. In Figure 8.25, the data after the reconstruction is presented. Difference between MEDAR/Medatlas and Pathfinder v5 data before the reconstruction is not shown, as its distribution and magnitude are very similar to the reconstructed data results. A depth of 2 meters is taken for the MEDAR/Medatlas data, as this depth minimises the difference between both data sets. The RMS error of the initial Pathfinder v5 data, before reconstruction, is 0.92°C. The RMS error for those points after DINEOF reconstruction is 0.93°C, only 0.01°C higher than the initial points. For reconstructed data, the error is 0.81°C.

When looking at the spatial distribution of the difference between MEDAR/Medatlas and Pathfinder v5 data (Figures 8.25 and 8.26), we can see that it is very homogeneous, *i.e.* there are no zones that present higher errors than others.

8.5 Summary and Conclusions

A reconstruction method, called DINEOF, has been successfully applied to two realistic cases. Two high resolution SST matrices of the Adriatic and the Mediterranean Sea have been reconstructed. The method is robust, simple to use and does not need any *a*

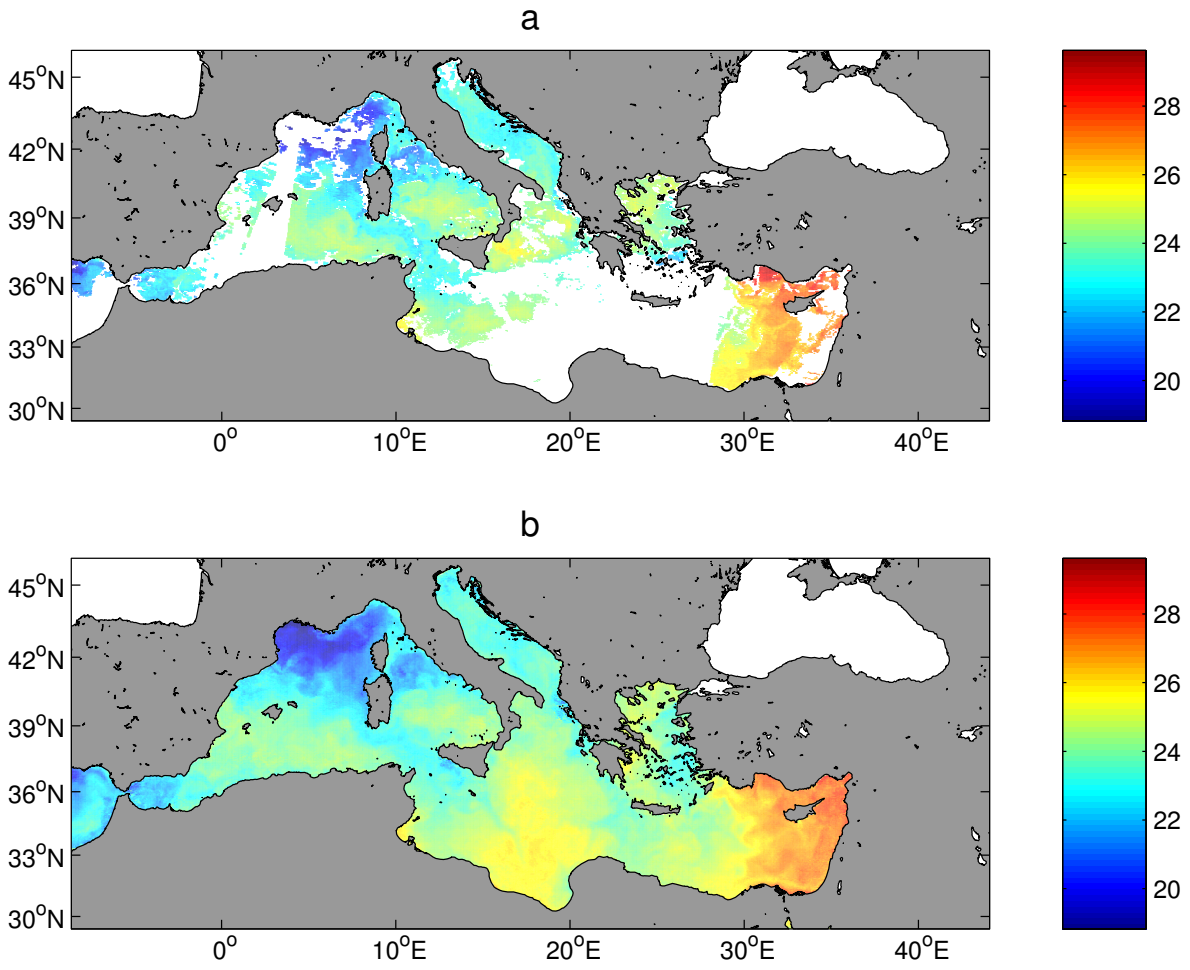

Figure 8.22: Pathfinder v5 SST on 20 July. Upper pannel shows the initial data, with clouds. Bottom pannel shows the reconstruction made with DINEOF

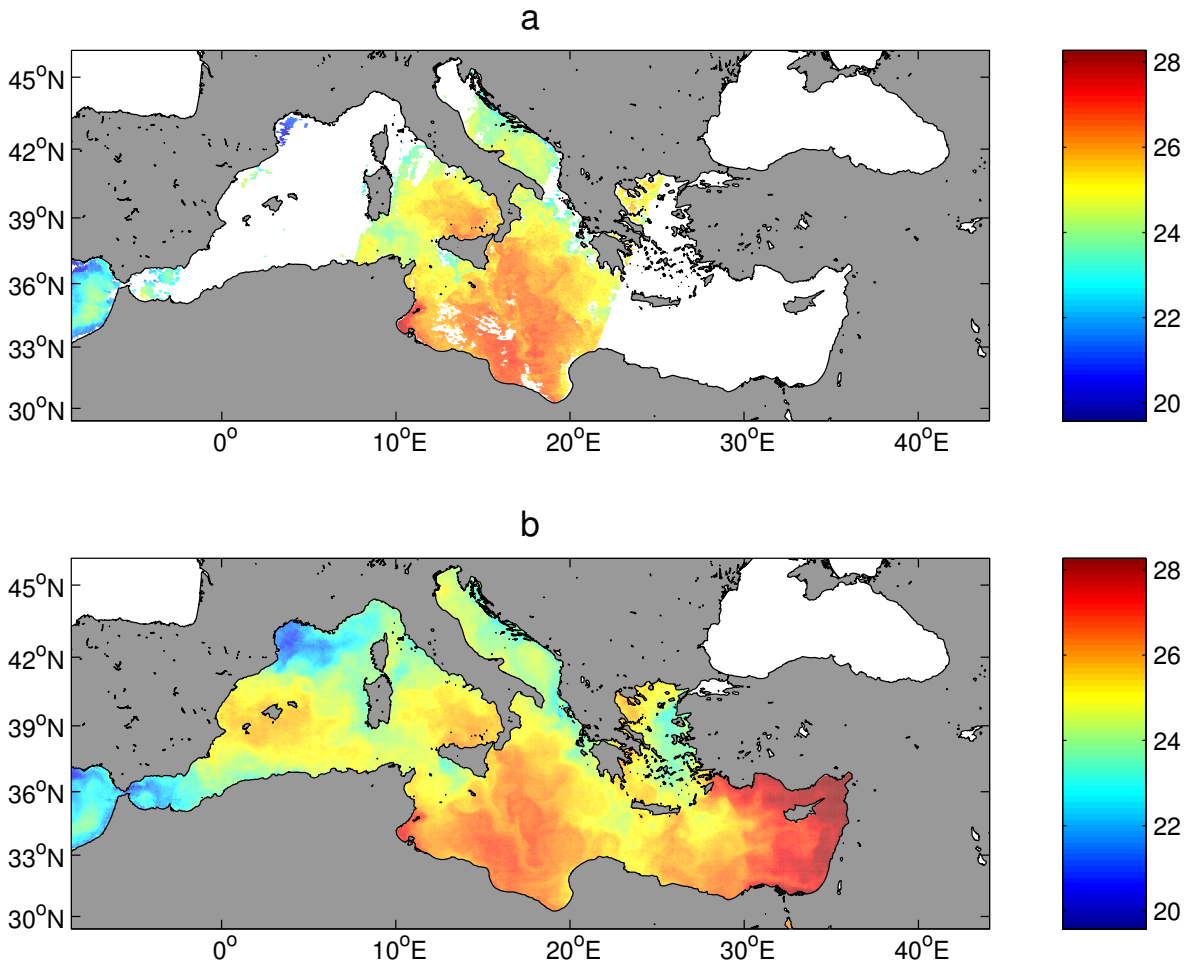

Figure 8.23: Pathfinder v5 SST on 5 August. Upper pannel shows the initial data, with clouds. Bottom pannel shows the reconstruction made with DINEOF

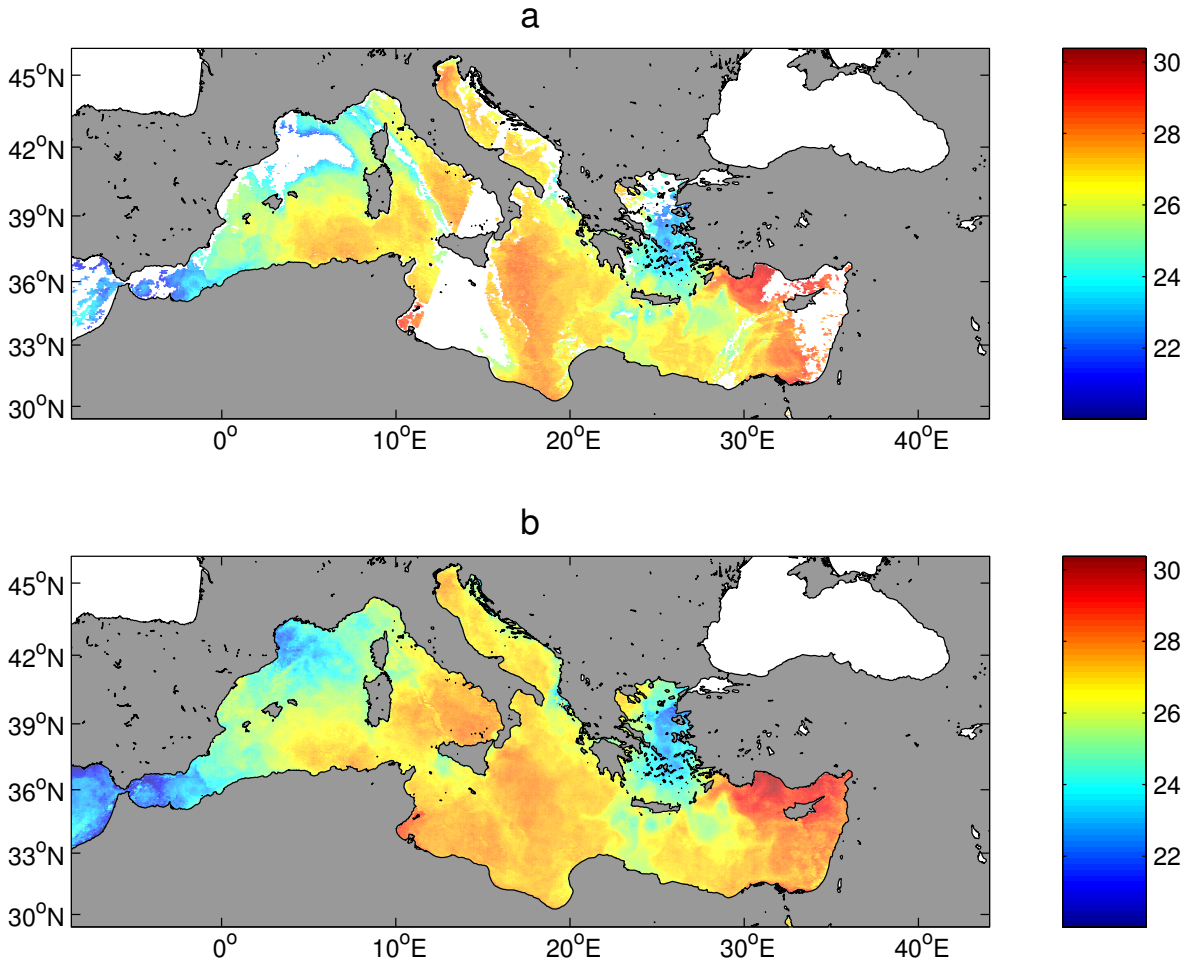

Figure 8.24: Pathfinder v5 SST on 27 August. Upper pannel shows the initial data, with clouds. Bottom pannel shows the reconstruction made with DINEOF

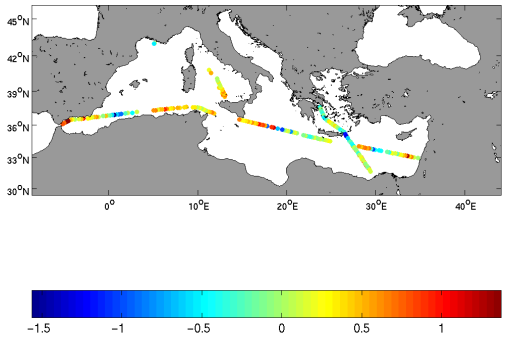

Figure 8.25: Difference between MEDAR/Medatlas stations and initially non covered data.

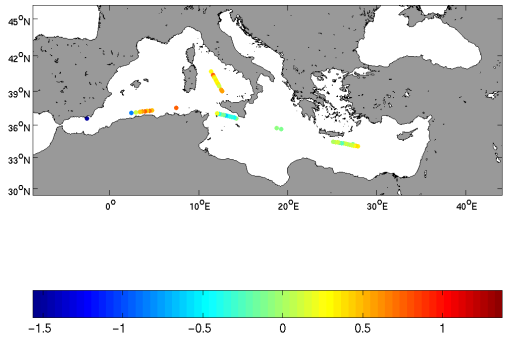

Figure 8.26: Difference between MEDAR/Medatlas stations and covered data that have been reconstructed with DINEOF.

priori information about the error statistics of the data. The results obtained have been analysed, giving an example of their reliability and usefulness.

The aim of this chapter was twofold. First, the application of the method to a realistic case, a SST data set covering the whole Adriatic Sea for a 6-month period. With this example we wanted to establish the capacity of DINEOF to reconstruct large matrices in a convenable time. Second, the application of DINEOF to a SST data set of the Mediterranean Sea, that covers in space and time the assimilation experiment realised with the GHER model. The reconstructed data are used in the skill assessment of the model.

A Lanczos solver was combined with DINEOF, in order to calculate the EOFs in an optimised way. Computational times are acceptable for the examples given in this paper. However, for the application of the methodology to very large matrices some additional optimisation work could be done, such as stop criteria for the calculation of the optimal number of EOFs, which would indicate when the minimum error calculated by cross-validation has been attained. EOFs do not need a regular time distribution for the analysis, although the accuracy of the DINEOF reconstruction could beneficiate from a data weighting related to the time distribution (North *et al.*, 1982; Shen *et al.*, 1994).

The results of the Adriatic Sea SST reconstruction are accurate and smoothly included in the final result, as seen in the visual examples given and in the validation studies made with a small set of the matrix. The validation was carried out in a data set with increasing amounts of missing data, where 40%, 60% and 80% of data loss were added. The comparison of the reconstructed fields with the original one shows

that the error is small, about 0.89°C, 0.78°C and 1.25°C for the 40%, 60% and 80% of missing data respectively. The comparison with *in situ* data from the MEDAR/Medatlas database also reveals an RMS error of the same magnitude, of about 0.95 °C for the points that are covered by clouds in the Complete Set. When visually checking the reconstruction results, we can see that the main physical features are recovered in the final result, such as the Po River Plume, cold filaments generated in the eastern coast, or the warm water current entering the Adriatic Sea by the Strait of Otranto in autumn.

Some artificial features in the initial data are also filtered out of the final result using DINEOF. This is highly interesting when dealing with data sets that have not been properly treated and present abrupt temperature changes in the vicinity of clouds.

In order to compare DINEOF reconstruction method to other commonly used methods, the Adriatic data set was reconstructed using an Optimal Interpolation (OI) method. The clouded data are reconstructed with OI, and then the EOFs are calculated and their amplitudes are fitted to the initial data. The results show that both methods, fitted OI and DINEOF, perform very similarly. The main difference is the computational time, which is nearly 30 times greater for the OI reconstruction. In many situations, such as an operational frame, this time difference is very important.

An additional test using Extended EOFs (ExEOFs) has been realised in the Adriatic data set. ExEOFs are very useful when the decomposed matrix has spatio-temporal moving patterns. In the case of missing data, this technique can help the reconstruction of a clouded point with additional information. The results showed that the ExEOFs approach gives slightly better results, with more coherent transition between the matrix images. Other techniques based on this extension of the initial matrix, such as multivariate EOFs (de Mey and Robinson, 1987; Fukumori and Wunsch, 1991; Sparnocchia *et al.*, 2003), where the matrix is composed by different variables, can also reinforce the reconstruction of the initial matrix. For example, wind field or air temperature are highly correlated to the SST, so the combination of these variables can lead to more consistent reconstructions.

An EOF analysis of the reconstructed field shows the utility of the reconstructed images. The first mode shows the seasonal cycle, from spring to autumn. The second mode presents the modulation of two general currents, the EAC and the WAC. Finally, the third mode shows the modulation of the Po plume. A cold event on the Albanian coast has been studied, and compared to a time series of wind obtained from ECMWF. This cold event is shown to be related to Bora winds that blow over the Albanian coast in the studied period.

The reconstruction of the Mediterranean Sea SST data set gives also good results. The comparison with MEDAR/Medatlas *in situ* data shows an error of 0.78°C for the reconstructed data and of 0.93°C for the initially clean data. This matrix will be further used in this work, namely in the frame of the spatio-temporal study.

In this study, DINEOF has been applied to temperature fields. Further work may show its utility when reconstructing other variables, such as salinity or chlorophyll. When dealing with chlorophyll fields, the patchiness that usually characterises their distribution (Martin, 2003; Strass, 1992) makes the reconstruction of these data a difficult task. Chlorophyll fields are very decorrelated, and techniques like optimal interpolation may have difficulty reconstructing this data. The characteristics of DINEOF make it suitable for reconstruction of such data. Any other field variable obtained from satellites (from passive receptors that cover a large area and thus provide a good data coverage) or from an observational network is suitable to be reconstructed with DINEOF.

Part III

Results

In this part of the work the skill assessment of the experiment where SST and SLA are assimilated into the model is presented. This experiment extends over nine weeks, and at each week the available data are assimilated into the model. The following data sets are used:

- The model:

  → Model forecast, just before the assimilation of SST and SLA each week. At this step, the forecast is independent from the observations that will be assimilated into the model. We will refer to this run as `Forecast` in this work.

  → Model analysis: the result of the assimilation of observations. This step is thus non-independent from the observations. It is considered in this work to study the impact of the assimilation step into the forecast process. It will be called `Analysis` in this work.

  → Free run of the model, without assimilation of observations. This set is used to study the contribution of the assimilation to the skill of the model. Apart from the data assimilation, the free run does not have any other difference with the forecast (model parameterisation, forcings...). It is called `Free run` in this work.

- Assimilated observations:

  → DLR SST data, the data assimilated on the model. They are used to establish the skill of the model in a weekly time step.

  → Sea Level Anomaly (SLA) tracks from the Naval Research Laboratory, also assimilated on the model.

- Non-assimilated observations:

  → Pathfinder v5 SST daily images at 4km resolution, to study the evolution of the model skill between two assimilation cycles. The data source is the same for both DLR and Pathfinder v5 SST, although they have been treated differently.

  → *in situ* data from Sirena cruise, in the Ligurian Sea, from 22 to 30 August 2000. This period covers the last week of the experiment, from the eighth to the ninth assimilation cycle.

  → *in situ* data from MEDAR/Medatlas database (MEDAR-Group, 2002). They cover the period 17 July to 1 September 2000. The measurements are taken all over the Mediterranean Sea.

  For further information on these data sets, see section 4.2.

- MEDAR/Medatlas climatology

This part of the work follows the already introduced scheme 'Qualitative verification', 'Quantitative verification' and 'Spatio-temporal analysis'. In chapter 9 we analyse the mean behaviour of the model compared to observations, climatologic data, or bibliography in a qualitative way. We will have a first idea about the model performance. Then, in chapter 10 we establish the error of the model results by means of the classical measures introduced in chapter 5. This will allow us to quantify the model error. Finally, in chapter 11 the model results and the observations are analysed in a spatio-temporal frame, which allows us to better understand the error sources and to have a complete idea about the model skill.

Chapter 9

Qualitative error assessment

9.1 Introduction

The first step in the verification process is the qualitative assessment of the model skill. Before applying any quantitative error measure, we must confirm that the model is able to represent the mean observed patterns. Any major problem concerning the temperature or salinity distribution, currents etc, should be detected here. This kind of quality assessment also allows obtaining a first, subjective idea about the model skill (it is subjective in terms that no statistic measure score is used). Mean fields of temperature, salinity and velocity will be analysed and compared to observations or climatology. Water masses can be studied with TS diagrams. The transports in some specific areas of interest will also be calculated and compared to the existing bibliography.

This kind of qualitative verification is often used in model intercomparison (*e.g.* Crosnier *et al.* (2004) for the MERSEA project, or Beckers *et al.* (2002) for MedMEX intercomparison project). When different models are compared, it is very difficult to apply and interpret an error measure, because of their different parameterisations. Qualitative error assessment allows to examine the model goodness and avoid problems on the scores interpretation.

A major advantage of the qualitative error assessment is that it allows to rapidly detect major problems. On the other hand, as there is no rating of the model, its quality is established in a subjective way. Ideally, this error assessment must be combined with a quantitative error assessment, where the model is rated in an objective way.

This chapter is organised as follows: first, the Mediterranean Sea mean temperature, salinity and velocity spatial distribution is studied, at the surface, 100 m depth and in some chosen vertical sections. Then, the water masses characteristics are established, with the help of TS diagrams. Finally, the transports of the more representative currents of the Ligurian Sea are calculated. After this study we will made a brief summary of this chapter.

9.2 Mean Fields

The mean horizontal field of the Mediterranean SST for both the `Free run` and the `Forecast` is compared to MEDAR/Medatlas climatology (Rixen *et al.*, 2001) and Pathfinder v5 SST (Kilpatrick *et al.*, 2001). The model has a surface relaxation to MEDAR/Medatlas climatology, so it cannot be considered as independent. However, as long as no quantitative error measure is applied here, this dependency is not a problem. This section aims to establish the overall similarity of the model with the main observed features in the Mediterranean.

Figures 9.1 to 9.4 show the averaged SST over a period of 60 days from 5 July to 2 September, for `Free run`, `Forecast`, climatology and Pathfinder v5 respectively. Climatology average has been calculated over the months of July and August. Mediterranean MEDAR/Medatlas climatologic fields have a resolution of 1/5°, with 25 vertical levels.

Both the `Free run` and the `Forecast` present overall a good distribution of the SST and surface salinity (Figures 9.1 and 9.2). We can see cold Atlantic water entering the Strait of Gibraltar and getting warmer as it spreads eastwards. Cold water is also entering the Bosphore Strait, coming from the Black Sea. In fact, input water coming from the Black Sea is not included in the model configuration, so the signal observed in the model is the result of the relaxation to the climatology. The `Free run` is colder than climatology or observed SST at the Gulf of Lions, spreading over a higher surface. At the eastern Mediterranean, the temperature of the `Free run` is slightly warmer than climatology and observed SST. The `Forecast` improves these mismatches, and the mean SST distribution is very close to Pathfinder v5 and climatology. In both the climatology and Pathfinder v5 SST we can see a two-lobbed water mass exiting the Aegean Sea at the East and West sides of Crete. This structure is also visible in the `Forecast`, but not in the `Free run`. Fine structures, as the Mid-Mediterranean Jet spreading southeastwards from Sicily are clearly seen in the `Forecast`. The `Free run` SST is generally smoother than the `Forecast`. This later run is thus improved by the assimilation of the observations.

When looking at the mean surface salinity in figures 9.5 to 9.7 for the `Free run`, `Forecast` and MEDAR/Medatlas climatology, we can see the input of fresh Atlantic Water in the Strait of Gibraltar, and how it becomes saltier as it spreads eastwards, transforming into Modified Atlantic Water. This Atlantic Water is slightly spreading to the North in the model results, and this is not observed in the climatology. Both `Free run` and `Forecast` behave in a similar way. Fresh water in the Gulf of Lions, northern Adriatic and Aegean Sea, coming respectively from the Rhone River, the Po River and the Black Sea discharges into the Mediterranean Sea are too weak in the model (`Free run` and `Forecast`) when compared to climatology. The fresh water coming from the Arno River (northeast Italian coast) is also visible in the model results, although its signal is very weak, as it has a smaller discharge volume than the Po River or the Rhone River. The `Free run` salinity in the East basin is

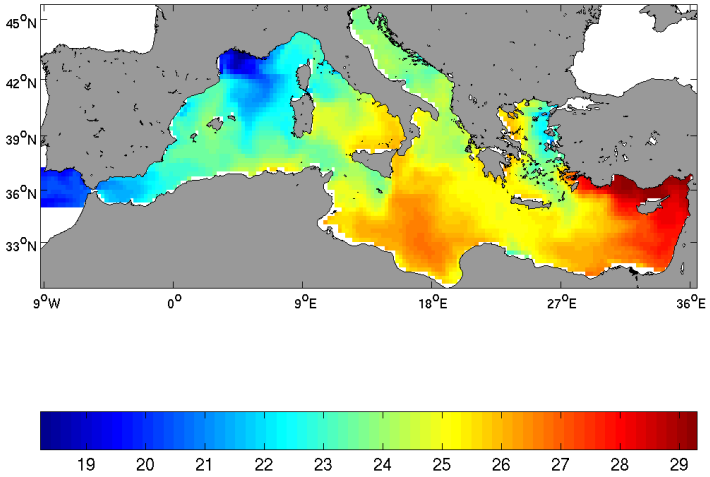

Figure 9.1: `Free run` mean SST.

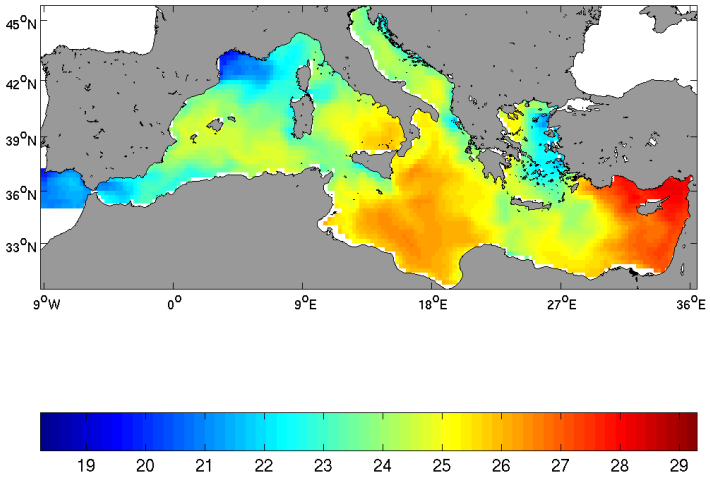

Figure 9.2: `Forecast` mean SST.

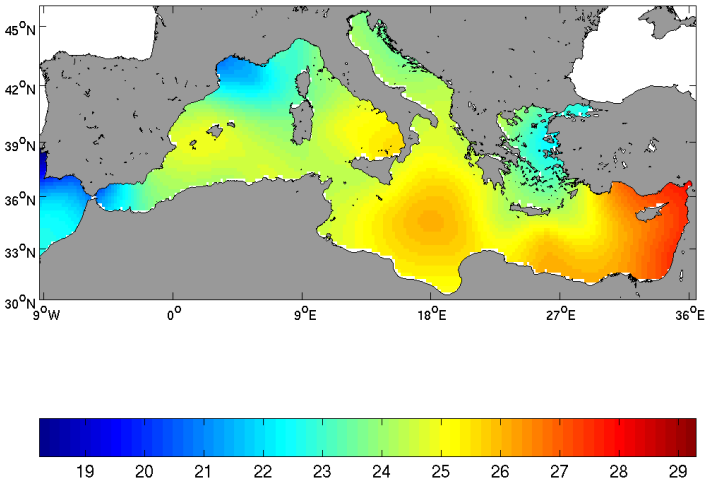

Figure 9.3: MEDAR/Medatlas mean SST.

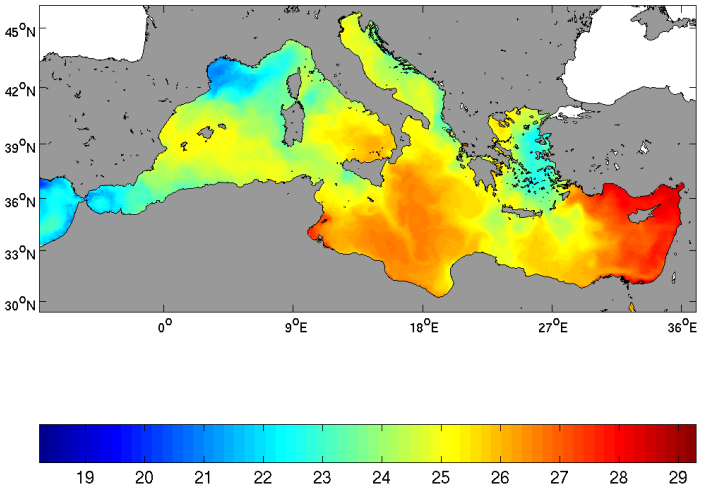

Figure 9.4: Pathfinder v5 mean SST.

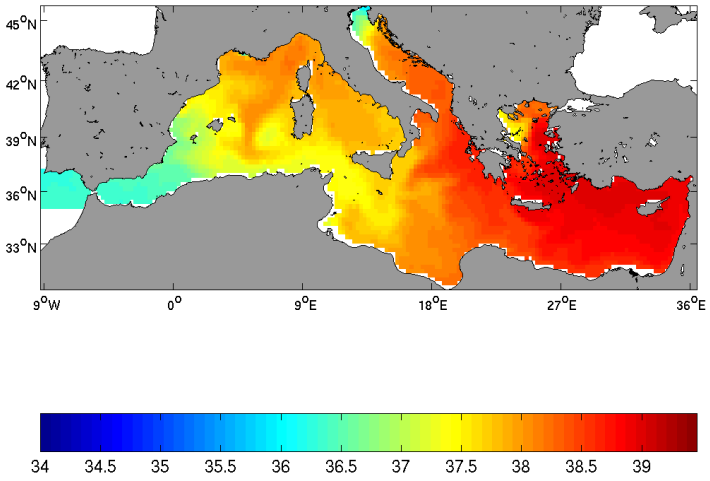

Figure 9.5: `Free run` mean surface salinity.

higher than the observed. The `Forecast` salinity in this zone is closer to climatology, which is an interesting result because no salinity is assimilated into the model. The mean distribution of surface salinity over the Mediterranean is in general well represented.

The mean horizontal surface fields of velocity obtained by the model reflect the main currents found in the Mediterranean. A depth of 10 meters has been chosen to avoid the effect of wind at the surface, which would be reflected in the velocity fields. Figures 9.8 and 9.9 present the mean surface velocity fields for the `Free run` and `Forecast` respectively. The main surface currents observed in the Mediterranean Sea are present. MAW enters by the Strait of Gibraltar and describes a cyclonic path over the whole Mediterranean. The model velocity detaches from the Algerian Current to the north, joining again the observed circulation at Sicily. This deviation to the north has been also observed in the temperature `Forecast` fields. At the Strait of Sicily, the current flows mainly along the southern Sicily coast, with a weaker counter-current along the Tunisian coast, feature that has been seen in other models (Roussenov *et al.*, 1995; Beckers *et al.*, 2002). The model is relaxed to MODB climatology, and this feature is also found in the mean geostrophic velocity field (Brankart and Brasseur, 1998). Also smaller features can be seen, as the cyclonic gyre in the Ligurian Sea, or gyres in the Alboran Sea.

The current from the Tyrrhenian Sea to the Ligurian Sea is flowing mainly southwards in the model results. Although this current flows usually northwards, periods of inversion of this current have been observed in summer (Vignudelli *et al.*, 1999). However, this reversal is only observed in the surface layers, as the total transport across this channel

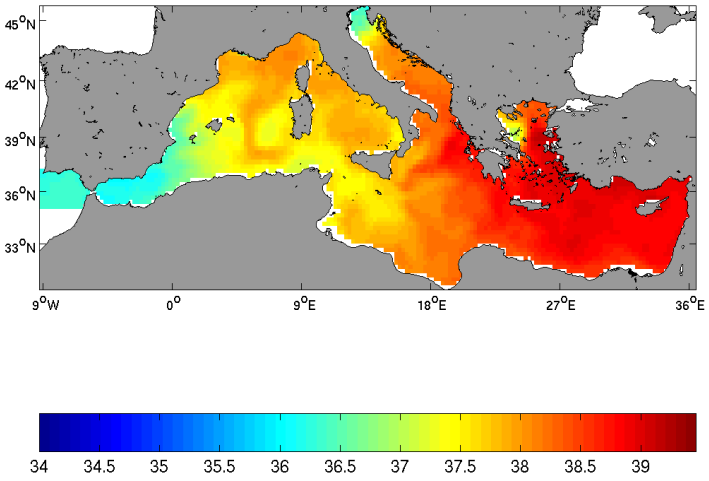

Figure 9.6: `Forecast` mean surface salinity.

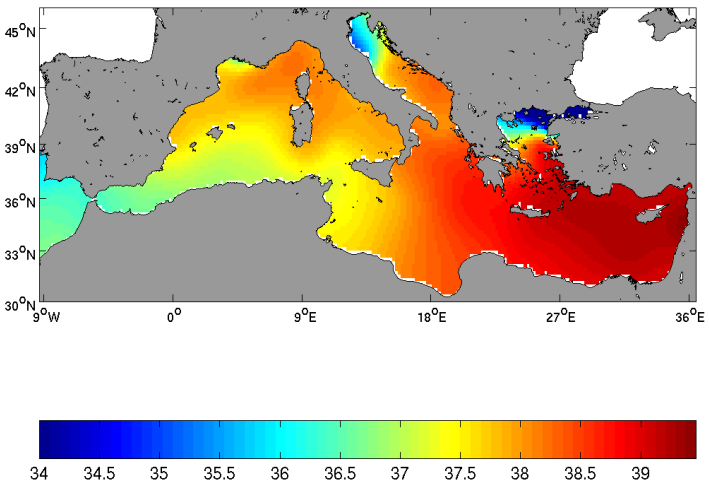

Figure 9.7: MEDAR/Medatlas mean surface salinity.

goes northwards (this will be shown later).

The mean horizontal distribution of temperature and salinity has also been studied at a hundred meters depth. Figures 9.10 to 9.12 present the temperature distribution at 100 m for the `Free run`, `Forecast` and climatology. The temperature distribution is overall well represented by the model, although eastern waters are too cold compared to climatology.

When looking to the horizontal salinity distribution at 100 m depth (Figures 9.13 to 9.15) we can see that in the climatologic field the fresh water input from the Atlantic is still present, whereas the `Free run` and `Forecast` Atlantic Water input is smaller and confined to the African coast. In the Western basin, the Tyrrhenian Sea shows the highest values of salinity. The model eastern basin salinity is similar to climatology, although it was seen that the largest differences in temperature were located in this basin.

Vertical sections

Vertical sections allow us to study the vertical distribution of the column water at some specific areas of interest in which we want to focus our study. Two vertical sections have been done in the Western basin. They will be compared to climatology by making an average over the 2 months period mentioned at the beginning of this chapter. In addition, a vertical section covering the whole south Mediterranean from East to West is made and compared to profiles obtained from the MEDAR/Medatlas database. These profiles have been taken in a few days ship trip, so no average is made. The model results are shown for the same days as the profiles.

The first vertical section is situated in the Northern Western Mediterranean, from the Gulf of Lions to the Italian coast. It covers the first 100 meters of the water column. In Figure 9.16 a map shows the location of this section. In the same figure, the climatology temperature section is shown. Climatology salinity distribution is shown in Figure 9.17. Figures 9.18 and 9.19 show the same transect for the `Forecast` temperature and salinity respectively.

The climatology has a thinner mixed layer in the Gulf of Lions than the `Forecast`. The mixed layer depth for climatology goes from about 10 meters depth in the Gulf of Lions to 25 meters between Corsica and the Italian coast. The model presents a higher mixing of surface waters at the Gulf of Lions, so the surface layer reaches a higher depth, of about 30 meters. Surface waters in the Gulf of Lions are colder in the model than in the climatology. Water is also fresher in the climatology, because of the Rhone River fresh water input. It was already noticed in the mean salinity surface field that the impact of the Rhone River into the Gulf of Lions surface waters is smaller in the model results. At the eastern side of the section, the model surface layer is very similar to

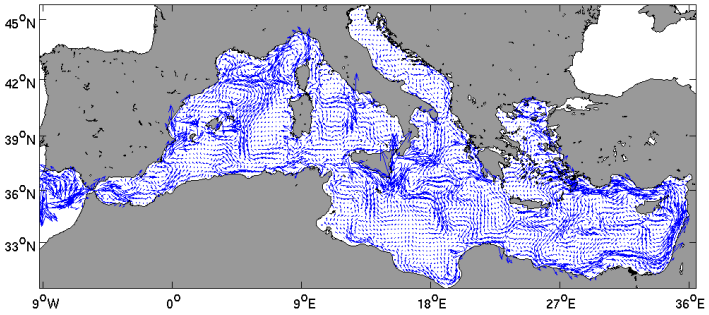

Figure 9.8: `Free run` mean surface velocity field.

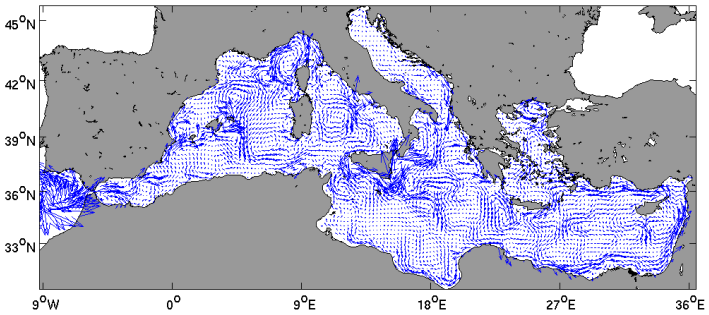

Figure 9.9: `Forecast` mean surface velocity field.

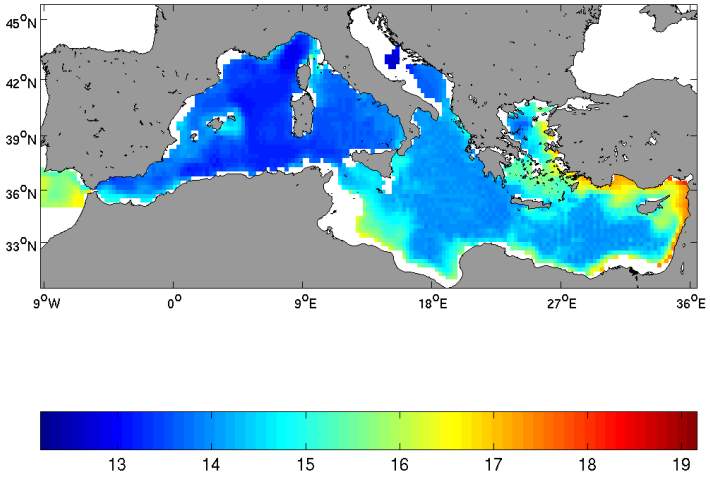

Figure 9.10: `Free run` mean temperature at 100 m depth.

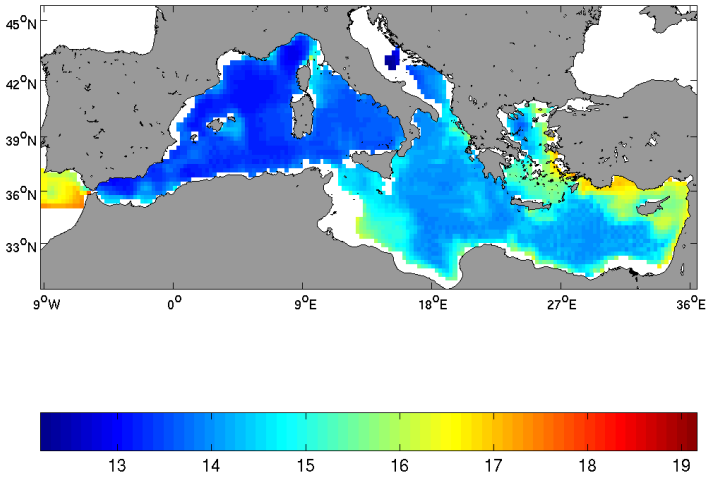

Figure 9.11: `Forecast` mean temperature at 100 m depth.

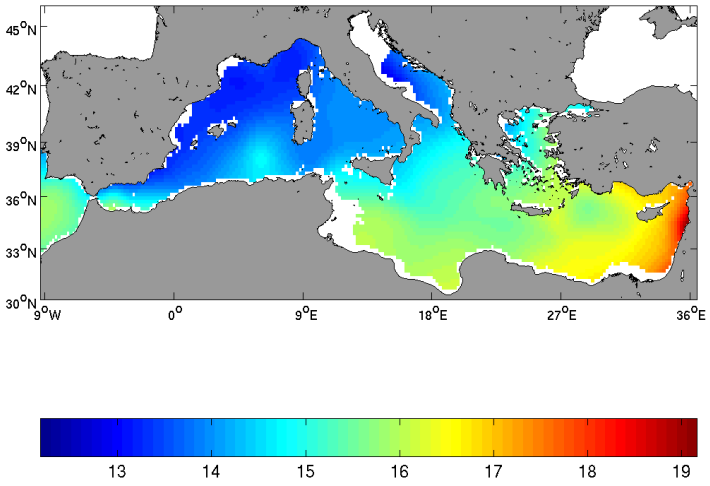

Figure 9.12: MEDAR/Medatlas mean temperature at 100 m depth.

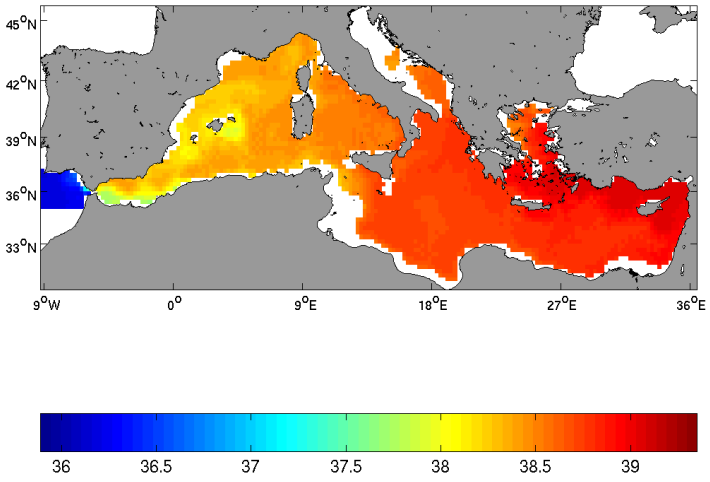

Figure 9.13: `Free run` mean salinity at 100 m depth.

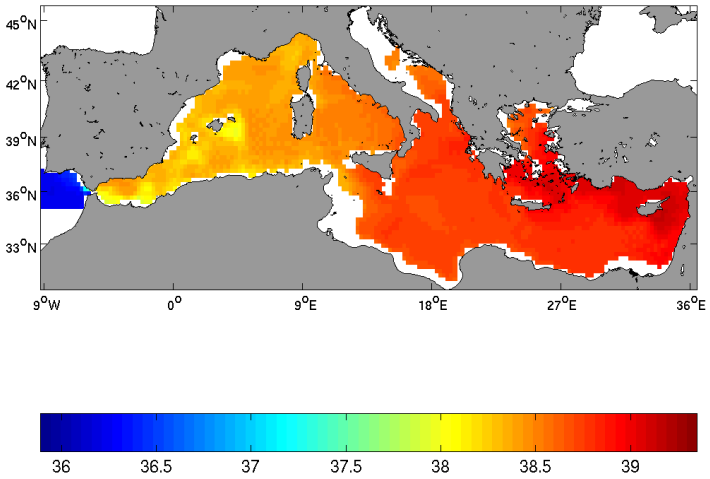

Figure 9.14: `Forecast` mean salinity at 100 m depth.

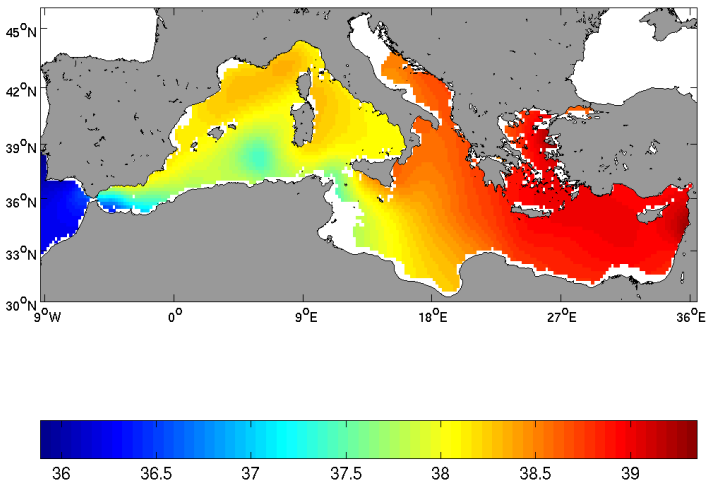

Figure 9.15: MEDAR/Medatlas mean salinity at 100 m depth.

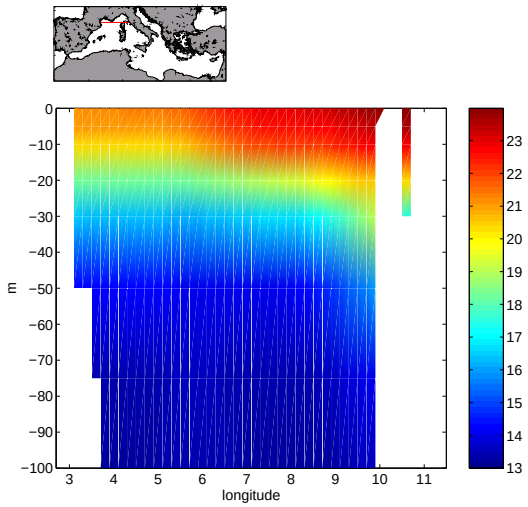

Figure 9.16: Vertical section in the Gulf of Lions for climatology temperature. The red line inthe upper map shows the localisation of the section.

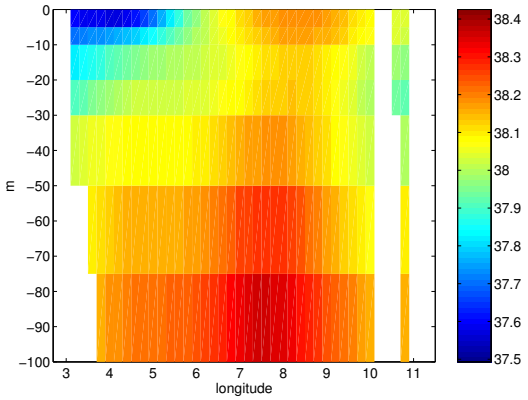

Figure 9.17: Vertical section in the Gulf of Lions for climatology salinity.

climatology. A surface temperature maximum is found at this place, mainly between Corsica and the Italian coast. In the model, the surface MAW entering the Ligurian Sea is less saline than climatology. At this place, warm waters coming from the Tyrrhenian Sea enter the Ligurian Sea to form the Northern Current. The water column in this place presents the structure that can be seen in the model results, with MAW at the surface layer and LIW at depth.

The second section made in the Western Mediterranean covers the basin from North to South. In Figure 9.20 we can see the map with the position of the section and the climatology temperature. The climatologic salinity is shown in Figure 9.21. Figures 9.22 and 9.23 show the same section for `Forecast` temperature and salinity respectively. In these figures, we can see the Northern Current signal, with a shallower mixed layer than in the centre of the Mediterranean. The signal of the Northern Current can also be seen in the doming of the salinity at this place. At the Algerian coast, the climatology presents high temperatures at the surface (Figure 9.20), corresponding to the MAW entering the Mediterranean by the Strait of Gibraltar. In the salinity distribution (Figure 9.21), this current is even more visible, with fresher water covering the first 70 meters of the water column. This water is confined to the African coast. MAW current is also present in the model (Figures 9.22 and 9.23), although its depth is higher, about 30 meters. The salinity of the model presents two separate fresh water branches, one thinner confined to the African coast that can be identified with the Algerian current and one thicker seaward, in the middle of the Western basin. This fresh water structure is also visible in the surface horizontal salinity figure for the `Forecast` (Figure 9.6).

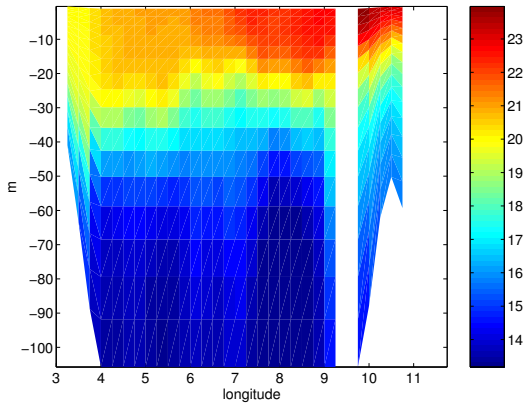

Figure 9.18: Vertical section in the Gulf of Lions for the `Forecast` temperature.

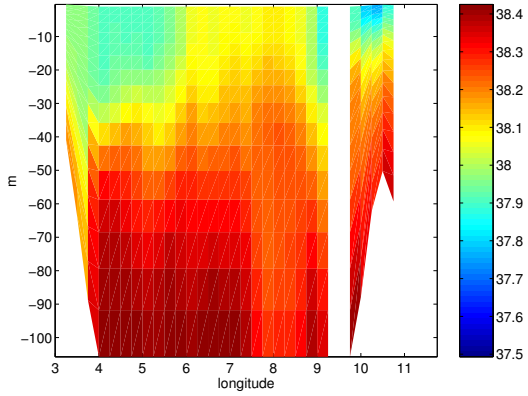

Figure 9.19: Vertical section in the Gulf of Lions for the `Forecast` salinity.

Fresh MAW circulates northwards at the west coast of Corsica to form the WCC. In this case, it seems that fresh water has been trapped in a cyclonic gyre (see surface currents at Figure 9.9, at 38°N, 7°E) and transported northwards by the Algerian Current. The MAW extends over a depth of about 200 meters, and below this warmed surface layer, we can find the typical characteristics of the Atlantic water, already mixed with the warmer and saltier Mediterranean water.

The third vertical transect covers the south Mediterranean, from East to West. The position of the MEDAR/Medatlas profiles is shown in Figure **??**. The temperature of these stations is shown in Figure 9.24. The `Forecast` temperature at the same locations is shown in Figure 9.25.

The main feature that must be commented in these figures is the mixing at the Strait of Gibraltar. As it can be seen, the model presents a uniform temperature over the water column at this location, which is not visible in the observed profiles. These profiles are situated at the interior of the Mediterranean Sea, but still near to the Strait of Gibraltar, thus affected by its dynamics. Fluxes at the Strait of Gibraltar are very strong, thus causing a high mixing in the model. When looking to the profiles at a higher depth, we can see that this feature is related to a high gradient in the bathymetry (see Figure 9.26). The steep bathymetry found at this point can be also the cause for the instability.

9.3 TS diagrams

Water masses can be identified by their temperature and salinity characteristics. TS diagrams are widely used for this purpose. The main water masses of the Mediterranean Sea were described in chapter 2, and a summary can be found in table 2.1. In Figures 9.27 to 9.32 the TS diagrams for the Western and Eastern Mediterranean basins for the `Free run`, the `Forecast` and the climatology are respectively shown. The TS diagrams

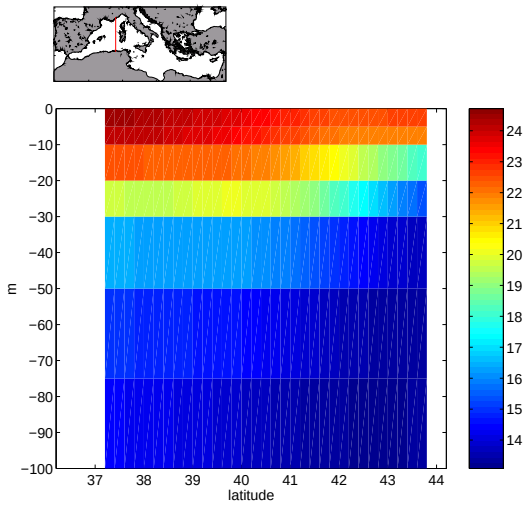

Figure 9.20: Vertical section in the Western Mediterranean for climatology temperature. The red line in the upper map shows the localisation of the section.

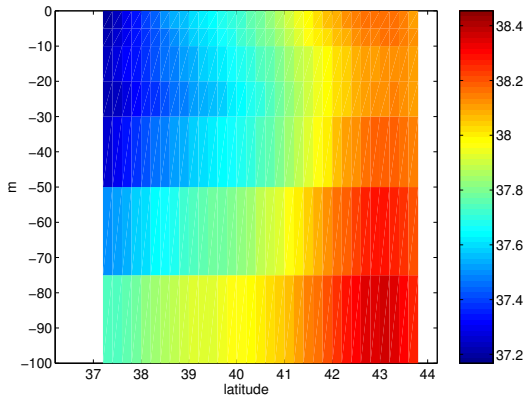

Figure 9.21: Vertical section in the Western Mediterranean for climatology salinity.

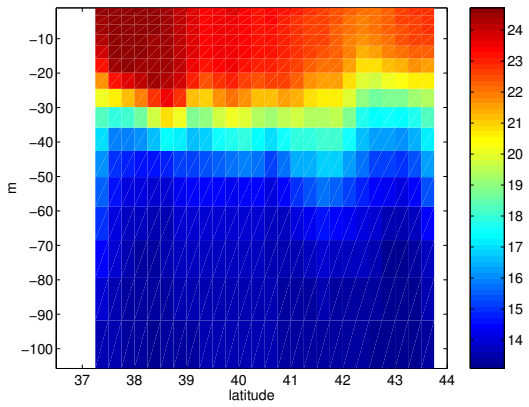

Figure 9.22: Vertical section in the Western Mediterranean for the `Forecast` temperature.

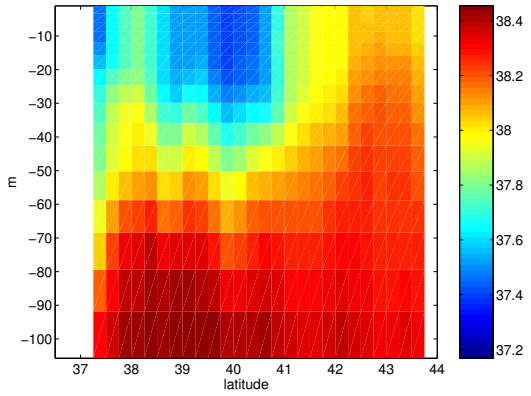

Figure 9.23: Vertical section in the Western Mediterranean for the `Forecast` salinity.

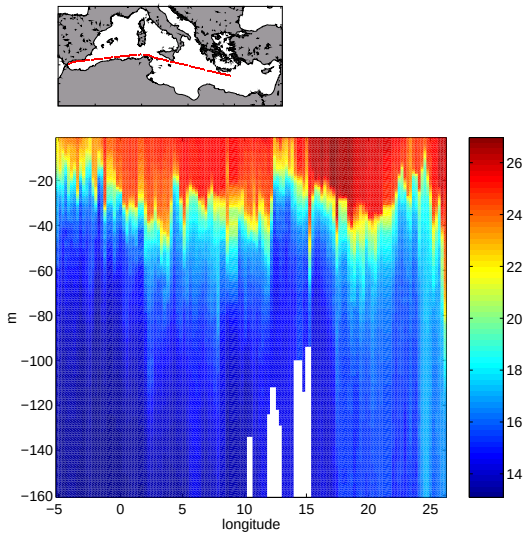

Figure 9.24: MEDAR/Medatlas profiles temperature.

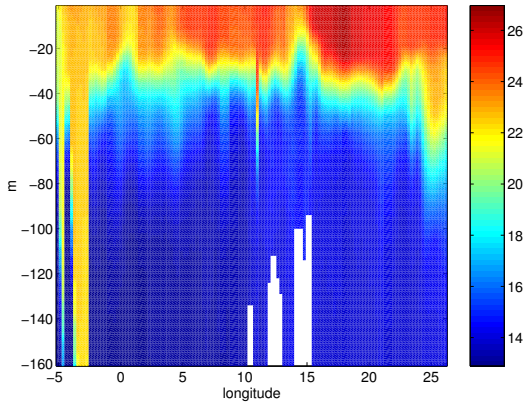

Figure 9.25: `Forecast` run temperature at the position of MEDAR/Medatlas profiles (Fig. 9.24) for the first 100 meters

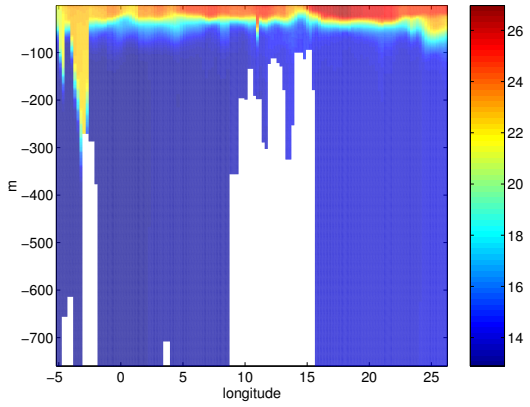

Figure 9.26: `Forecast` temperature at the position of MEDAR/Medatlas profiles (Fig. 9.24) for the first 700 meters.

for the Western basin and the Eastern basin are different. Western basin waters are colder and fresher than the Eastern basin waters. Both the `Free run` and the `Forecast` present a similar shape, very close to climatology one. In the western basin, we can identify a branch of fresh cold water situated at depths from 500 to 200 meters. This water mass is Atlantic Water, situated outside the Strait of Gibraltar. Once in the interior of the Mediterranean basin, the Atlantic Water is mixed with the saltier and warmer Mediterranean water.

Looking in detail to the deep-water masses of the TS diagrams in the Western Mediterranean (Figures 9.33, 9.33 and 9.33 for climatology, `Free run` and `Forecast`, respectively) we can distinguish between the signal of LIW, warm and saline (about 14°C and 38.7 psu) of about 500 meters depth, and the WMDW, colder and less saline (about 13.4°C and 38.5 psu) filling the depth of the Western basin. Model and climatology present similar characteristics of these water masses.

In the Eastern basin, the climatology (Figure 9.32) and the model (Figure 9.30) TS diagrams are slightly different. For example, the climatology surface is less saline than the `Forecast`. Salinity and temperature are higher in the Eastern basin (note how the Eastern and Western TS diagrams have a complementary form). When looking to the detail at depth, (Figures 9.34, 9.34 and 9.34 for climatology, `Free run` and `Forecast`), we can see that climatology deepest waters are less saline and colder than those on the model. Again, we can distinguish between an intermediate water mass present at the Eastern basin, the LIW and a deep water mass, the EMDW.

9.4 Transports

The transports corresponding to the most relevant currents of the Ligurian Sea model have been calculated: the ECC between Corsica and Italy, the WCC at the west of Corsica, and the NC at the north of the Ligurian basin. The places where these transports have been estimated are indicated in Figure 9.39.

Information about the transports across the main channels and straits of the Mediterranean basin is not very extensive in the bibliography. Most of the time, annual means are available. Summer transports are often the weakest of the year, so the comparison of these values with our model results can be difficult. The comparison between bibliography values and the model is made with this constraint in mind. Astraldi *et al.* (1999) made a good review of transports in the Mediterranean straits and channels, and we use their data and references in their work here for comparison. Other sources are consulted, and they will be cited at appropriate time.

Mean transports over the mentioned transects are summarised in Table 9.1. The

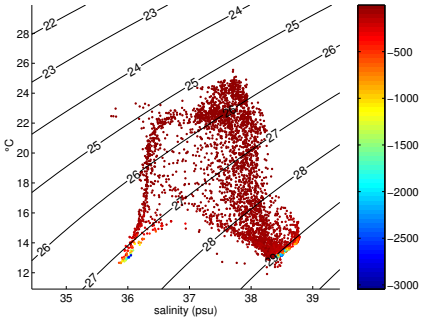

Figure 9.27: TS diagram of the Western Mediterranean for the `Free run`

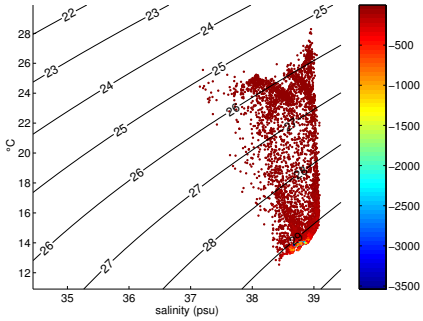

Figure 9.28: TS diagram of the Eastern Mediterranean for the `Free run`

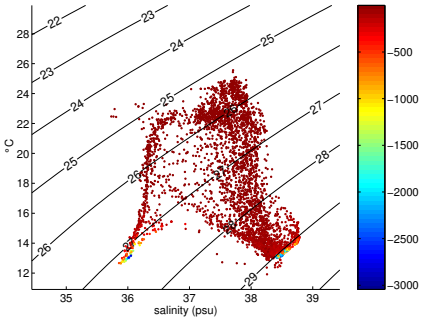

Figure 9.29: TS diagram of the Western Mediterranean for the `Forecast`

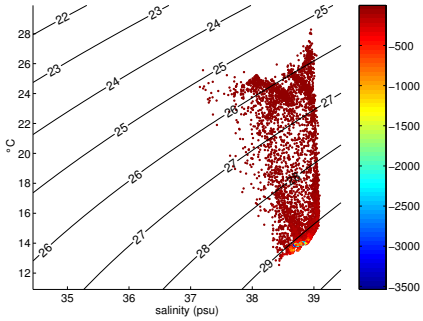

Figure 9.30: TS diagram of the Eastern Mediterranean for the `Forecast`

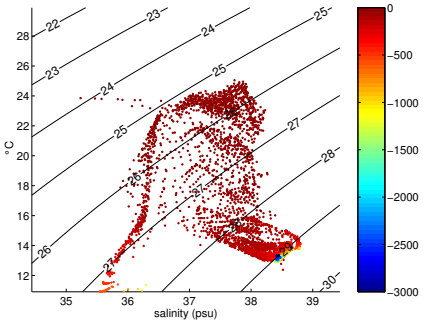

Figure 9.31: TS diagram of the Western Mediterranean for the climatology

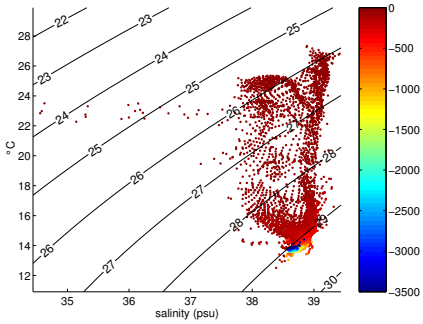

Figure 9.32: TS diagram of the Eastern Mediterranean for the climatology

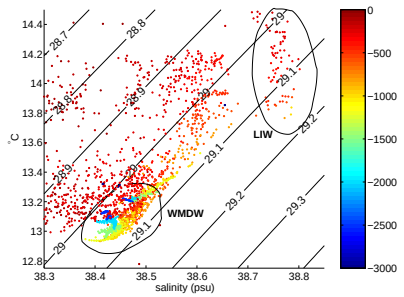

Figure 9.33: Detail at depth, TS diagram of the Western Mediterranean for the climatology

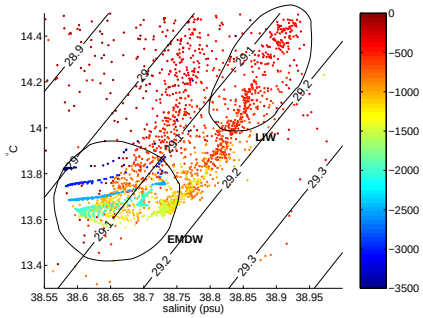

Figure 9.34: Detail at depth, TS diagram of the Eastern Mediterranean for the climatology

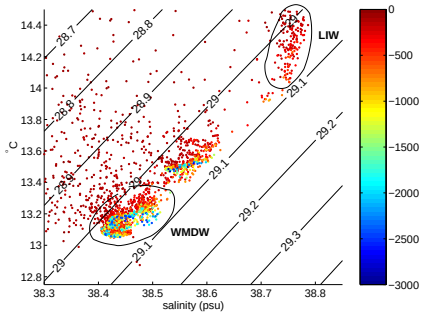

Figure 9.35: Detail at depth, TS diagram of the Western Mediterranean for the `Free run`

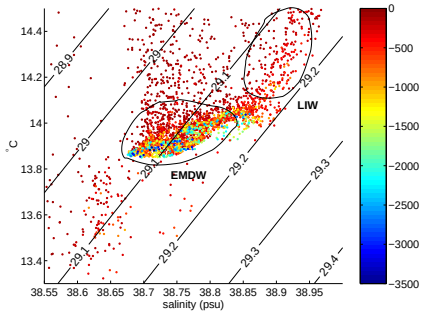

Figure 9.36: Detail at depth, TS diagram of the Eastern Mediterranean for the `Free run`

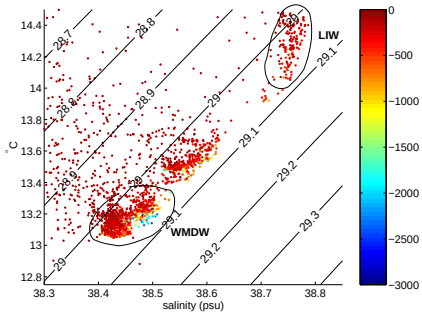

Figure 9.37: Detail at depth, TS diagram of the Western Mediterranean for the `Forecast`

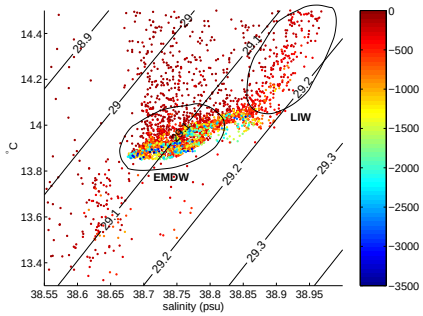

Figure 9.38: Detail at depth, TS diagram of the Eastern Mediterranean for the `Forecast`

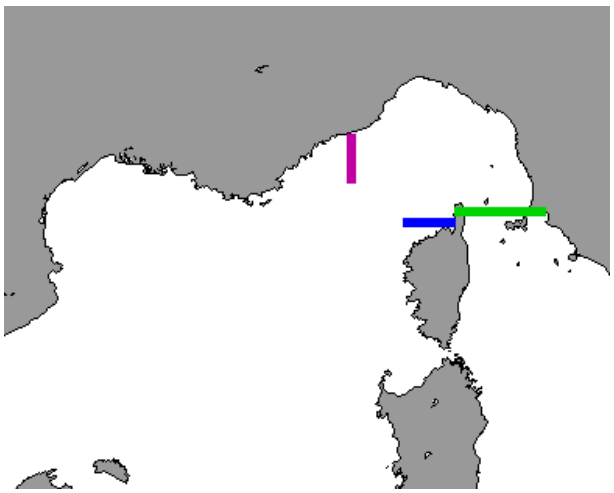

Figure 9.39: Places where the transports have been calculated. In green the ECC, in blue the WCC and in magenta the NC.

transport of the ECC at the surface layer is underestimated for both `Free run` and `Forecast`, but for the deep layer (from 200 to 600 meters) the model results are good. The surface current reversal that we have observed (Figure 9.9) is the cause of the weak surface transport, as well as the difference between summer and winter values. The `Forecast` WCC transport is closer to observations than the `Free run`, although the difference is very small. The WCC observed value is an annual average, and it has been observed that in summer there is a decrease in the transport. Although the WCC is more important than the ECC in terms of volume, this last one has a greater impact in the heat content of the Ligurian Sea, as it brings a warmer water mass to this basin. The NC is underestimated in the `Free run`, and in the `Forecast` the transport increases up to observed values.

9.5 Summary

In this chapter we have realised a first, qualitative assessment of the model results. This kind of studies are necessary in a verification process. They allow to detect major problems in an early step of the verification process. This study will guide the future quantitative assessment of the model performances.

Model fields, averaged over the two months length of the assimilation experiment, have been compared to MEDAR/Medatlas climatology, and when available, to direct observations (Pathfinder v5 SST and bibliographic data). On average, model surface temperature presents the main observed distribution, for both the `Free run` and the `Forecast`. However, `Forecast` temperature is somewhat closer to observations. Salinity fields are less realistic than temperature fields. This difference can be due to the fact

Table 9.1: Mean transports on the Western Mediterranean Sea. Observed values are taken from bibliography.

Current	Observed	`Free run`	`Forecast`
ECC surface	0.5 Sv[a]	0.18 Sv.	0.2 Sv.
ECC depth	0.15 Sv[b]	0.16 Sv.	0.16 Sv.
WCC	1.15 Sv[c]	0.86 Sv.	1.02 Sv.
NC	0.8-1.6 Sv[d]	0.57 Sv.	0.85 Sv.

[a]from Astraldi and Gasparini (1992); Astraldi *et al.* (1999). Annual mean value. 80% of the transport takes place in winter.

[b]from Astraldi and Gasparini (1992); Astraldi *et al.* (1999). Annual mean value. 80% of the transport takes place in winter.

[c]from Astraldi and Gasparini (1992). Annual mean. Maximum observed rates in spring, minimum in Autumn.

[d]from Albérola *et al.* (1995).

that surface temperature is assimilated into the model. However, `Free run` presents also better temperature than salinity, and no SST assimilation is done in this model.

The main currents in the Mediterranean Sea are represented in the model, with some mismatches, as the deviation to the north of the Algerian current, the reversal of the current along the Tunisian coast, or even the reversal of the ECC at surface, between the Tyrrhenian and Ligurian Seas. Some of these reversals have been observed (*e.g.* ECC, (Vignudelli *et al.*, 1999)).

The model seems to be closer to climatology in the Western Mediterranean. The grid refinement is applied in this part of the Mediterranean Sea, so they are more accurate. The benefits of the two-way nesting approach have already been studied by Barth *et al.* (2003). In that work, the impact of the two-way nesting was assessed between the Ligurian Sea model and the Liguro-Provençal basin, leading to the conclusion that both models benefit from the two-way nesting approach.

The qualitative error assessment is thus a first step to have a general idea about the model performance. We have already seen that the assimilation and the two-way nesting improve the model results, although the main improvement is done in the assimilated variables. The non-assimilated variables, as the salinity `Forecast`, show smaller improvement compared to what is obtained in the `Free run`.

Now it is necessary to establish the model skill in a quantitative way. We cannot simply say that the model is "good", as it will not be very helpful if we want to improve the model skill, and it will be of no aid to any potential user of the model results. We

must find where the bigger errors are, and which factors are causing them. This will help us to understand the causes of the error and, as far as possible, to correct these error sources in the model.

Chapter 10

Quantitative error assessment

10.1 Introduction

After the qualitative assessment of the model performances, we can establish its skill in a quantitative way. We have seen that the model represents the major physical features of the Mediterranean Sea, consistent with observations, although we have not made a rating of the model with regard to observations and to the reference system. Now we can measure the model skill objectively, and study the effect of the assimilation into the model. Classical error measures, presented in chapter 5, such as RMS or the called skill scores, are useful to measure the perfomances of a model. If they are carefuly chosen and the obtained results are interpreted jointly, they can also help us to find the causes or sources of the error. This is the aim of this chapter. The extraction of a maximum of information about the model skill will help us to better understand it, and to make the pertinent corrections that will lead to better forecasts.

The verification of spatial scales is easily performed if spatial averages are made. We can study the time evolution of the errors. However, a spatial distribution of the error is more difficult to interpret, because of the correlation that can exist between adjacent model points (Livezey and Chen, 1983). In this chapter we will use spatial averages to study the error time evolution? For the spatial distribution characteristics of the model we will apply spatio-temporal techniques in the next chapter.

This chapter is organised as follows: first, the skill of the model is studied in comparison with the weekly-assimilated data, with SST in the Ligurian Sea and with SST and SLA in the whole Mediterranean Sea. DLR SST and TOPEX/POSEIDON-GFO/ERS-2 SLA tracks are used in this part. Then, in section 10.3 a comparison to independent data is made, namely with Pathfinder v5 SST data, Sirena cruise in the Ligurian Sea and *in situ* data from MEDAR/Medatlas database. The use of *in situ* data will allow us to study the water column structure, rather than only its surface. To conclude this chapter section 10.4 presents the conclusions of this chapter, with some guidelines to the improvement of the verification process.

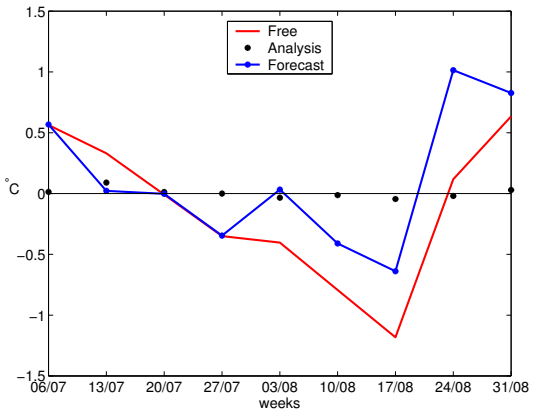

Figure 10.1: Bias evolution of `Forecast`, `Analysis` and `Free run` SST related to DLR SST data.

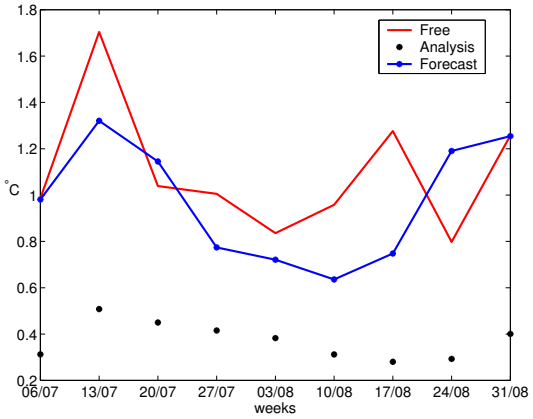

Figure 10.2: Root Mean Square Error of `Forecast`, `Analysis` and `Free run` SST related to DLR SST data.

10.2 Weekly error measures

The assimilation of observed data is done weekly. The first test we will make to the model consists in studying at each assimilation cycle the impact of the assimilation, to identify weaknesses or problems that may appear because of the assimilation itself. We make the comparison with SST data in the Ligurian Sea model first, whose resolution is the highest due to the nesting. After the skill assessment in the Ligurian Sea, the model SST is compared to the DLR SST in the whole Mediterranean. The comparison with altimetry data is also done in the whole Mediterranean.

10.2.1 SST error evolution in the Ligurian Sea

The time evolution of `Forecast`, `Analysis`, and `Free run` are compared to the weekly DLR SST data. For the `Forecast`, the model results are compared to observations before they are assimilated, so they are still independent. The `Free run` results are taken at the same dates than the `Forecast`, and the `Analysis` is the result of the assimilation of observations, one step after the `Forecast`. The time evolution of the error is a common way to study the model accuracy progression, because it allows in a rapid way to quantify the skill and to detect drifts in model results. Two frequently used measures in forecast verification are the bias and the RMS error, and they are presented in Figures 10.1 and 10.2.

Bias

Bias allows us to know if a variable is over or underestimated, and to see if the assimilation process or the model configuration introduce a systematic bias into the model. In this case, the bias is smaller for the `Forecast` than for the `Free run` for the first 7 weeks. Both runs present almost the whole period a negative bias. `Forecast` and `Free run` do not follow the SST summer warming as observations do, and temperature is thus underestimated. From 24 to 31 August, the `Free run` has a smaller bias than the `Forecast`. In fact, both the `Forecast` and `Free run` change the bias sign from 17 to 24 August, making a total increase in temperature of 1.5 °C compared to observations. The behaviour is very similar for both runs: first, from 3 to 10 August, the bias of both the `Forecast` and `Free run` becomes more and more negative. Both runs are warmer than observations on 24 August. This episode will be further studied with the following error measures. The `Forecast` presents almost no bias on 13 July, 20 July and 3 August. At the first instant of the graphic, on 6 July, the `Free run` and the `Forecast` are strictly the same, because no data assimilation has been done into the model, so both runs have the same bias. The `Analysis` is always close to zero, and it helps to reduce the `Forecast` bias to a great extent, so no systematic bias is introduced with SST assimilation.

Root Mean Square Error

The RMS error measures the accuracy of the model rather than its bias. The bias does not depend on the individual correspondence between the model results and the observations, so it cannot be considered as an accuracy measure. An accuracy measure, as the MSE or the RMS error, will give us an idea about the average degree of correspondence between the model and the observations. We use the RMS error rather than the MSE because the units of the former are those of the studied variable, so it is more easily interpretable. The model's RMS error evolution (Figure 10.2) is somewhat similar to the bias behaviour: the `Forecast` is better than the `Free run`, except for 24 August, where the `Forecast` error increases noticeably. The `Free run` has also a slightly smaller RMS error than the `Forecast` on 20 July. Surprisingly, the RMS error of both `Free run` and `Forecast` reaches its maximum value on 13 July, where there is almost no bias. The same result is obtained on 20 July. This example shows the utility of using more than one measure to establish the skill of a model. The mean temperature of the `Forecast` is very close to the mean temperature of the observations on 13 July, but the `Forecast` does not accurately represent the SST of the Ligurian Sea. We can see this in Figure 10.5, where the observations, `Free run`, `Forecast` and `Analysis` SST on 13 July are given. The `Forecast` temperature is colder than observations at the southwest of the basin, but warmer temperatures in the northeast balance this difference and make the bias to be zero. However, the distribution of SST does not match the observed one, thus giving a high RMS error. The `Free run` southwest-northeast temperature gradient is even sharper than for the `Forecast`. The assimilation of the observations into the model (called "Assim" in the figure) redistributes the SST in the Ligurian Sea in agreement

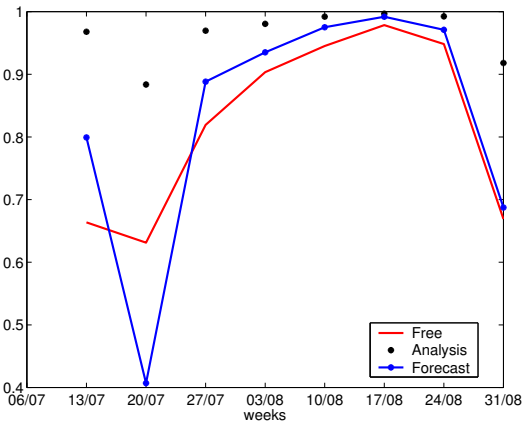

Figure 10.3: Anomaly Correlation Coefficient for the `Forecast`, `Free run` and `Analysis` SST related to DLR SST data.

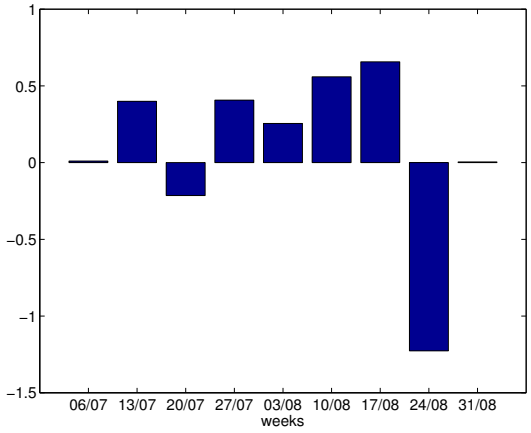

Figure 10.4: Mean Square Error Skill Score for the `Forecast` SST. The `Free run` is used as reference system.

with observations.

During the whole experiment, the `Analysis` decreases the error of the `Forecast` of about 0.5°C at each assimilation cycle, but it seems the `Forecast` drifts away from observations during the week, between two assimilation cycles. The daily behaviour of the `Forecast` will be studied later in this chapter.

Anomaly Correlation Coefficient

In addition to accuracy, measured by RMS, we can measure the correlation between the model and the observations. Figure 10.3 presents the Anomaly Correlation Coefficient (ACC). As for all correlation measures, 1 represents the best, and -1 the worst. The ACC will indicate us if the `Forecast` and the observations anomalies follow the same patterns of evolution. This measure does not take into account the bias, or the accuracy of the model, so we will extract different information from it, namely if the `Forecast` follows the observed pattern. In this graphic, we can see that on 20 July the `Forecast` presents a bad ACC (ACC=0.4), even worse than the `Free run`. An ACC of 0.6 is usually taken as the lower limit to consider a forecast as useful[1]. However, the rest of

[1]It was shown by Murphy and Epstein (1989) that, taking Eq. 5.13 under certain conditions (mean anomaly in the observed field close to zero, negligible unconditional bias and a variability of the forecast close to observations) the MSESS can be expressed as:

$$\mathrm{MSESS} \sim 2\mathrm{ACC}_{f'o'} - 1 \tag{10.1}$$

This way, the skill remains positive if the ACC is greater than 0.5. An ACC of 0.6 will give a skill of 0.2, so the forecast can be considered as useful (Murphy and Epstein, 1989; von Storch and Zwiers, 1999).

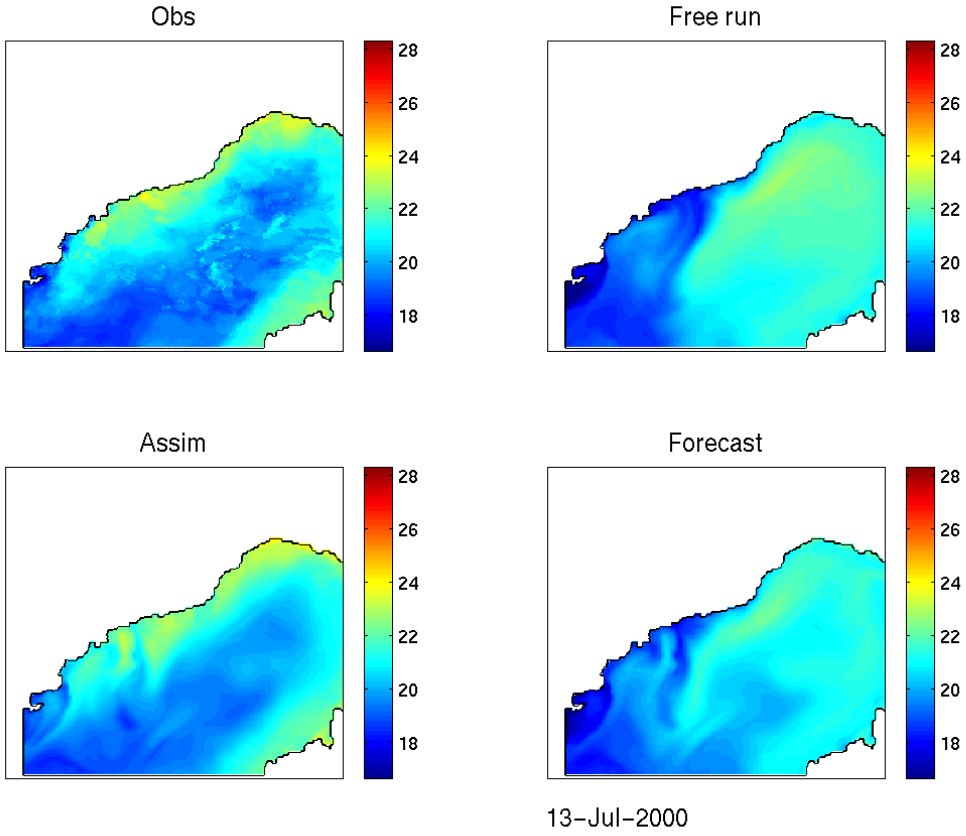

Figure 10.5: SST distribution in the Ligurian Sea on 13 July for the observations, `Free run`, `Forecast` and `Analysis` ("Assim"). `Forecast` is thus still independent from the shown observations.

the experiment the `Forecast` has a slightly better ACC than the `Free run`. In terms of ACC, the 24 August `Forecast` seems not to be as problematic as with the other error measures. Thus, the error at this week is due to the small accuracy and high bias of the `Forecast`, but the `Forecast` follows the observed patterns.

Mean Square Error Skill Score

Until now, the performance of the `Free run` and the `Forecast` has been explained together, but none of the used measures compare them directly. This comparative study can be done with the help of skill scores. In this work, the Mean Square Skill Score (MSESS) is used. In Figure 10.4 the MSESS is presented. With this kind of error measures we can see the improvement (or deterioration) of the model skill with regard to a reference system. In this case, the `Free run` is used as reference system. The MSESS is based on the MSE, so the behaviour is related to the RMS graphic on Figure 10.2. The `Forecast` improves the `Free run`, except on 20 July, 24 and 31 August. The skill score on 24 August is very bad, being the improvement of the `Free run` over the `Forecast` of more than a hundred percent (recall that the skill scores can be expressed as percentages).

Error summary

To summarise the error measures applied to the Ligurian Sea weekly SST, we can say that there are three instants that show problems, due to different reasons. For example, on 13 July, there is a problem of accuracy in the `Forecast`, whereas there is no bias. The following week (20 July) there is still no bias, but the ACC shows very bad results. The 17 August, the `Forecast` is cold compared to observations (high negative bias), but the other error measures give rather good results. This may indicate that the high errors on the following two weeks (24 and 31 August) start as a bias problem. The averaged errors for `Forecast` and `Free run` for the Ligurian Sea SST appear in table 10.1. The `Forecast` improves the `Free run`, but due to the bad `Forecast` results on 24 and 31 August, the improvement is very low on average. The `Forecast` SST is in average overestimated (bias of 0.12°C), with an RMS error of 0.97°C and a good ACC. All these scores would be improved if we did not take into account the two last weeks of the experiment (but that's cheating).

Discussion

Let us take a closer look at what happens on 24 August. As we have already mentioned, the error starts as a high negative bias on 17 August. Looking at the mean SST of the DLR data, `Forecast` and `Free run` (Figure 10.6), it seems that there exists a delay between the surface temperature of the model (with and without assimilation) and the observations. The maximum observed temperature is reached the 17 August, while the

Table 10.1: Mean errors for the `Forecast` and the `Free run` SST in the Ligurian Sea.

Measure	`Free run`	`Forecast`
Bias	-0.12 °C	0.12°C
MAE	$0.92°C^2$	$0.84°C$
RMSE	1.1 °C	0.97°C
MSE	$1.27\ °C^2$	$1.01°C^2$
ACC	0.81	0.83
MSESS		0.095

model reaches its maximum the 24 August. We can see in this figure that the maximum reached by the `Free run` the 24 August is in addition very close to the temperature of the observations at this date, while the `Forecast` peak is sharper, about 1 °C warmer than the `Free run`. This would explain the fact that the `Free run` has a smaller error than the `Forecast` the 24 August, as it presents (it could be added "by hazard") a mean temperature that is closer the observations. The fact that the `Free run` presents also this delay in the mean temperature maximum, implies that it is not a result of the assimilation process. Data assimilation has a beneficial impact into the model, because the `Free run` appears to be colder than the observations the whole period, and the `Forecast` is closer to observations, with the exception of 24 August where the assimilation warms up too much the model. From 3 to 17 August the `Free run` is systematically colder than the observations, and the assimilation makes the `Forecast` to be closer to the real temperature.

To study this event in more detail, the atmospheric forcings have been studied. In Figure 10.7 the air temperature in the Ligurian Sea is presented. The observed temperature is also included in this figure, although one must bear in mind that temperature scales for air and sea are not the same. This figure is included to study any mismatch between the evolution of both temperatures. The maximum air temperature is reached about the 28 August, one week before the DLR SST. This difference in the maximum temperature date can be one cause of the model delay, and thus to the error present in the model. The 17 August, both `Forecast` and `Free run` are colder than observations, and the assimilation of these observations already at their maximum makes the forecast to continue warming. The air temperature continues also to warm up, so on 24 August, when observed SST have already begun to decrease, the difference between `Forecast` and observations is the higest for all the period, thus giving a bigger error for this week.

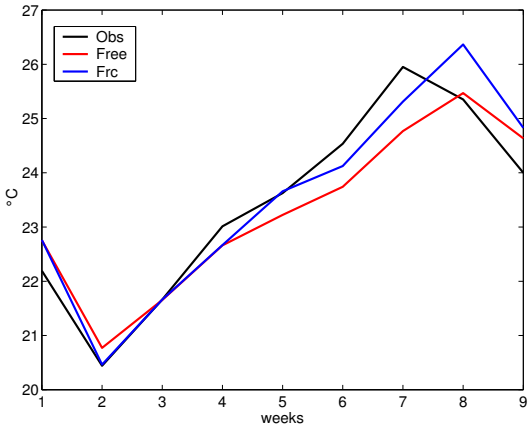

Figure 10.6: Mean temperature evolution as a function of time for DLR observations, `Free run` and `Forecast` in the Ligurian Sea.

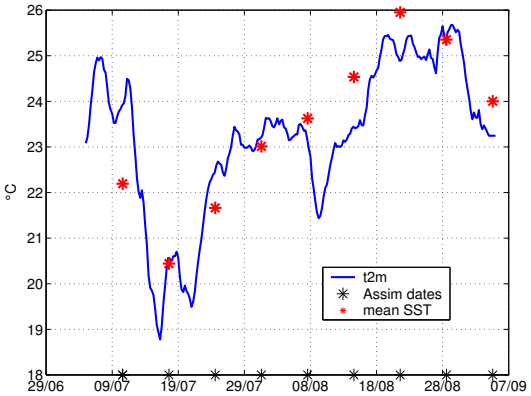

Figure 10.7: Air temperature evolution in the Ligurian Sea. Red asterisks show the temperature of the assimilated SST and black asterisks (at the time axis) show the moments where the assimilation is done.

10.2.2 SST error evolution in the Mediterranean Sea

In this section, we make a comparison between the SST results from the coarse resolution model, which covers the whole Mediterranean Sea, and DLR weekly images with the same coverage. The aim is to study the differences between the Ligurian and the Mediterranean model skill, as well as to study the model performance in the whole Mediterranean Sea. First, we want to assess the impact of the two-way nesting, which should improve all nested levels. Second, the high resolution reached at the Ligurian Sea model should improve the representation of the SST, although the increased resolution can also involve a decrease in skill. This decrease in skill of high resolution models is due to the double penalty effect (Hoffman *et al.*, 1995; Ebert and McBride, 2000). This means that the high resolution model, because of the increased resolution, represents more realistically the small-scale features than the coarse grid model. However, it is easy that such features are misplaced respect to the observed ones. When measuring the error between model and observations these small features, the model is penalised twice: once for missing the actual feature at the good place, and another for forecasting it in a bad place. The coarse resolution model can thus present a better skill than the fine resolution model.

The assimilation of DLR images is also done at the coarse level. Their resolution is degraded to 10 km for this assimilation. The same error measures applied to the fine resolution grid of the Ligurian Sea are calculated here, and again the `Free run` is used as reference system to assess the impact of the assimilation. The SST spatial average is computed to study the time evolution of the error.

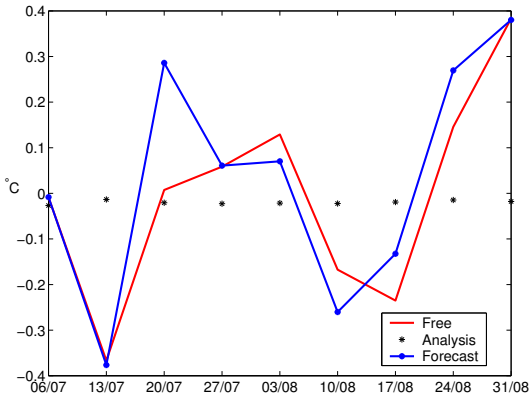

Figure 10.8: Bias evolution for the SST of the `Forecast`, `Free run` and `Analysis` in the Mediterranean Sea model.

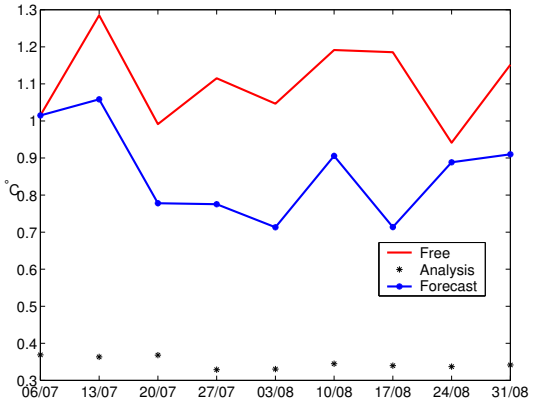

Figure 10.9: RMS error evolution for the SST of the `Forecast`, `Free run` and `Analysis` in the Mediterranean Sea model.

Bias

Figure 10.8 shows the bias of `Forecast` and the `Free run`. Both the `Free run` and the `Forecast` biases show a similar time evolution. There are two important sign changes in the bias, as noticed also in the Ligurian Sea. From 13 July to 20 July, the `Forecast` bias changes from -0.4°C to 0.3°C. The `Free run` does not seem to follow this warming. From 17 August to 24 August, however, both the `Free run` and the `Forecast` present the behaviour already observed in the Ligurian Sea model, with an increase of the temperature bias of 0.5°C. On average, (see Table 10.2) the bias is smaller for the Mediterranean model than for the Ligurian Sea model.

Root Mean Square error

As shown before, the bias alone is not a complete measure for model skill. Figure 10.9 shows the RMS error time evolution. The `Forecast` presents a smaller error than the `Free run` the whole period. On 24 August, the RMS error increases for the `Forecast`, and decreases for the `Free run`, as observed in the Ligurian Sea model, but not enough to make the `Free run` better than the `Forecast`. The `Analysis` step decreases the model error up to 0.6°. The assimilation of SST has a more important effect in the coarse grid model than the experimented in the finest grid model. This may be due to the already mentioned double penalty effect.

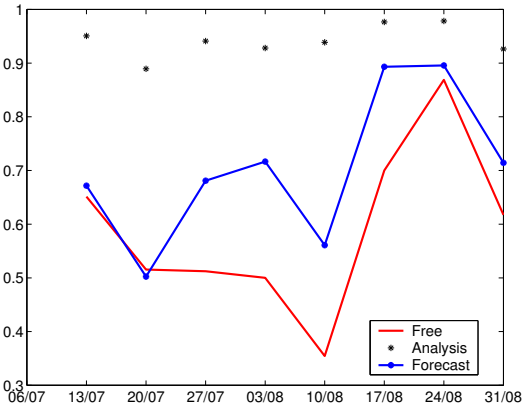

Figure 10.10: ACC error evolution for SST of the `Forecast`, `Free run` and `Analysis` in the Mediterranean Sea model.

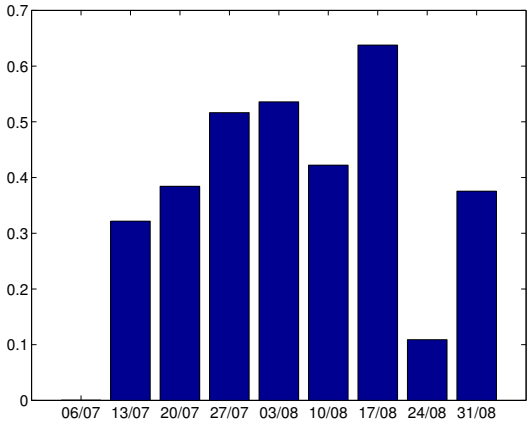

Figure 10.11: MSE Skill Score evolution for the SST of the `Forecast`, `Free run` and `Analysis` in the Mediterranean Sea model.

Anomaly Correlation Coefficient

The coarser resolution of the Mediterranean Sea model can affect the ACC, because this model is missing some structures, so the model pattern compared to the observed one is not good. The ACC can be seen in Figure 10.10. Indeed, the average ACC is worse than the produced by the Ligurian Sea (see Table 10.2). The `Forecast` is overall better correlated to observations than the `Free run`, except on 20 July where the `Free run` is slightly better. Again, the ACC is the best of the whole period on 24 August, when the model shows a high bias. This indicates that the error at this week is due to the accuracy of the model, although the pattern of the observed field is well represented.

Mean Square Error Skill Score

Finally, to compare the skill of the `Forecast` with the `Free run`, the MSESS is calculated. Figure 10.11 shows the MSESS of the `Forecast` when compared to the `Free run`. We can see that the `Forecast` is more skillful compared to the `Free run` during the whole period. Only the 24 August the skill goes under 0.2. The first day of the time series, the 6 July, the `Forecast` shows no skill compared to the `Free run` because at the first assimilation cycle both models are the same (no assimilation has been still done). This week is not included for the computation of the mean MSESS in Table 10.2. On average the MSESS is 0.37, which means an improvement of about 40% of the `Forecast` over the `Free run`.

Table 10.2: Mean errors for the `Forecast` and the `Free run` SST in the Mediterranean model.

Measure	`Free run`	`Forecast`
Bias	-0.01 °C	-0.03°C
MAE	$0.88°C^2$	$0.67°C^2$
RMSE	1.1 °C	0.86°C
MSE	$1.23°C^2$	$0.76°C^2$
MSESS		0.3670
ACC	0.59	0.7

Error summary

In Table 10.2 the mean errors for the 9-week period experiment in the Mediterranean Sea model are shown. As explained, the mean error in the Ligurian Sea, supposed to be smaller due to the nesting, is very close to the obtained at the whole Mediterranean Sea. In addition to the double penalty, we can find an explanation for these results in the warming event at the end of the experiment that was highly felt in the Ligurian Sea model. This event increases the error in the Ligurian Sea model, but it is weaker in the Mediterranean Sea model.

10.2.3 SLA error evolution in the Mediterranean Sea

Until now, we have been looking to the evolution of the surface temperature. Sea Level Anomaly (SLA) tracks are also assimilated on the model, in the three model grids (Ligurian, Provençal and Mediterranean). The TOPEX/Poseidon+ERS-2+GFO tracks (described in section 4) have been binned into weekly data sets, and then assimilated on the model every week. As the coverage of the SLA tracks is lower than the one given by the DLR SST data, the verification of this variable will be done directly in the Mediterranean Sea model. The `Free run` is again used as the reference system. The TOPEX/Poseidon+ERS-2+GFO tracks have initialy a higher resolution than the Mediterranean Sea model, so an alongtrack filtering has been made to eliminate from the observed data those features that cannot be resolved by the model before they are used for the assimilation. This filtered data is used here for the error assessment. SLA data are ususally filtered to eliminate artificial features (*e.g.* Larnicol *et al.*, 2002). At each week, the spatial average of the binned tracks and the spatial average of the model SLA at the same positions than the tracks are computed. Figure 10.12 shows the evolution in time of the SLA bias. We can see that the `Free run` presents over the entire period a negative bias, so the SLA of the model is smaller than the observed one. The `Forecast`

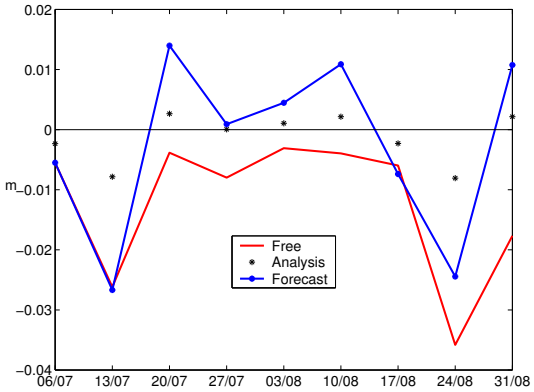

Figure 10.12: Bias SLA evolution in the Mediterranean Sea model.

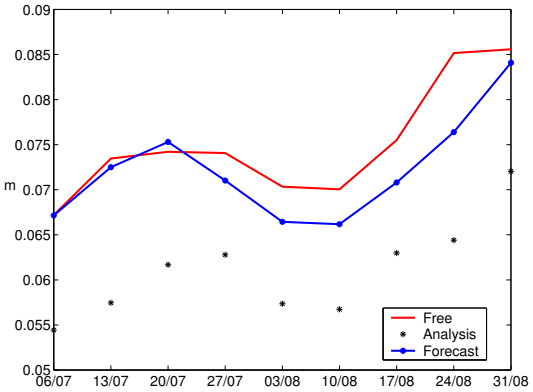

Figure 10.13: RMS SLA error evolution in the Mediterranean Sea model.

Table 10.3: Mean errors for the `Forecast` and the `Free run` SLA in the Mediterranean model.

Measure	`Free run`	`Forecast`
Bias	-0.01 m	-0.002 m
RMSE	0.075 m	0.072 m
MSE	0.006 m^2	0.005 m^2
MSESS		0.07

has always higher values than the `Free run`. On 13 July and 24 August, the `Forecast` presents its worst values, a behaviour that is also observed in the `Free run`. If we look to the point-by-point SLA difference between the `Forecast` and the observations (Figure 10.14), we can see that the distribution of this anomaly data is very homogeneous. However, points situated near the coast have the highest errors, due to the difficulty of correcting the tidal signal on the observed tracks. Third week (20 July), present some points with high positive anomalies (more than 0.25 m), which are contributing to the high positive bias on this week.

Figure 10.13 shows the SLA RMS error time evolution. The `Forecast` is better than the `Free run` during all the experiment. The effect of the `Analysis` on the model performances decreases the `Forecast` RMS in ∼1 cm. The highest RMS errors are again found at the end of the experiment, on 24 and 31 August, as it hapenned with the SST error. The improvement of the `Forecast` over the `Free run` is not very high, as it can be seen in the MSE skill score (Figure 10.15). The mean MSE skill score over the nine weeks of the experiment is only 0.07, or 7% of improvement of `Forecast` over `Free run`. The mean errors for the SLA data can be found in Table 10.3.

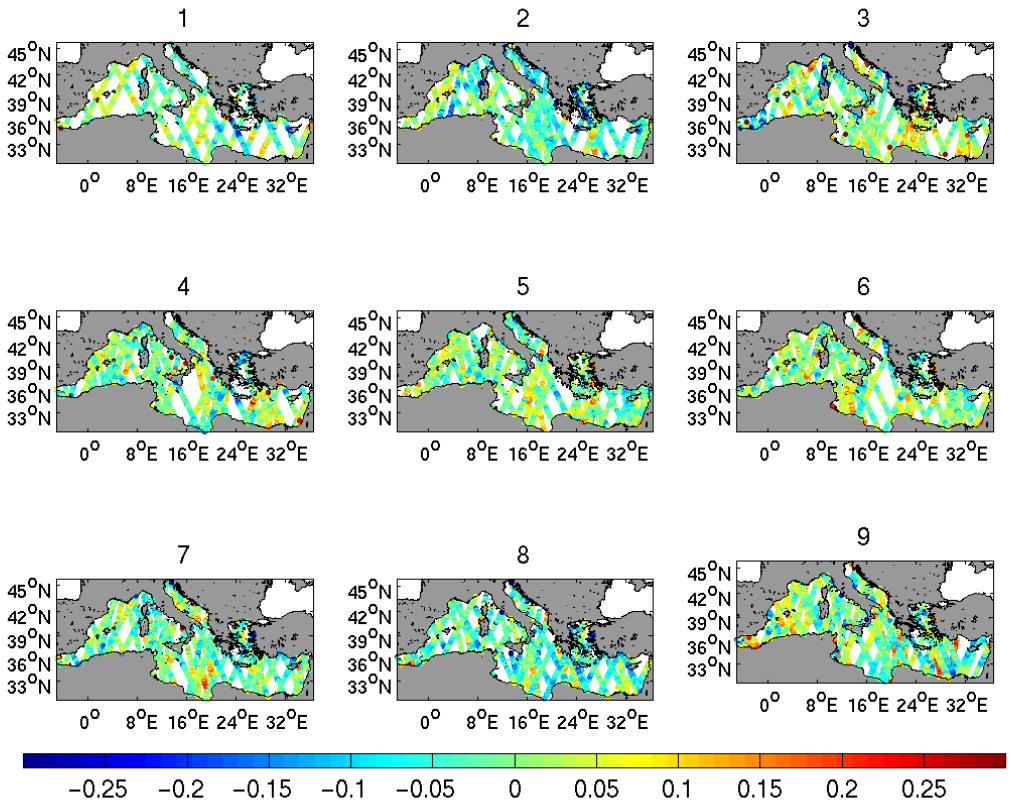

Figure 10.14: Point-by-point difference between the `Forecast` and the observed tracks for the nine weeks experiment. Weeks are marked over each figure.

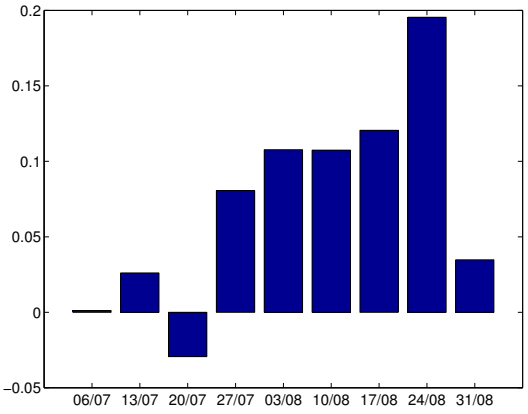

Figure 10.15: MSESS evolution for the SLA of the Mediterranean Sea model.

10.3 Comparison with independent data

In this section, the model results will be compared to three data sets that have not been assimilated into the model, and thus they are independent. The comparison with independent data is necessary because of the artificial skill that can be added to the model, being the information provided by the observations somewhat included in the model. The comparison with the SST data sets made in sections above is however necessary. No other data set as the satellite data gives a nearly complete coverage of the studied zone, in both time and space. The comparison with other types of data, as *in situ* data, is localised in time and space. We can thus obtain only a limited vision about the model skill. However, the advantage is that other zones (rather than the sea surface) and other variables (*e.g.* salinity) can also be studied.

10.3.1 Daily error assessment with SST Pathfinder v5 high resolution data

The next step is to see the evolution of the model skill between assimilation cycles, at a daily time step. Again, the `Free run` is taken as a reference system, to study the impact of the assimilation on the model. The Pathfinder v5 SST data set is built from the NOAA AVHRR satellite, as the DLR data assimilated into the model. Thus, the Pathfinder v5 data and the model with assimilation are again non-independent, although both data have been treated in a different way. This fact must be taken into account when studying the model skill. The daily output of the model is made as an average of one day values, and not at 14:00 hours as happened with the weekly results used in the section before. Before any error measures is made, the model is interpolated into the Pathfinder v5 grid.

Figures 10.16 and 10.17 show the bias and the RMS error respectively, for the model with `Forecast` and the `Free run`. As already shown in chapter 4, Pathfinder v5 data are warmer than the DLR data. A detailed study about the discrepancies between DLR and Pathfinder v5 data is out of the scope of this work. This difference in temperature is also felt in the comparison between the model and Pathfinder v5. The model is 0.24°C colder than Pathfinder v5 on average (see table 10.4). We will make the interpretation of the error between Pathfinder v5 and the model bearing in mind this bias. The RMS error (Figure 10.17) of the `Forecast` shows overall better results than the `Free run`. At the dates of the DLR data assimilation (marked with an asterisk on the x-axis), we can see that the error of the `Forecast` is improved. This is not the case on 24 August (8th cycle of assimilation, the 51st day on the figure), when the RMS error shows an increasing trend. The MSE nad MSESS are also shown (Figures 10.18 and 10.19), and we can see that the `Forecast` improves the `Free run` more than a 50% during the whole period.

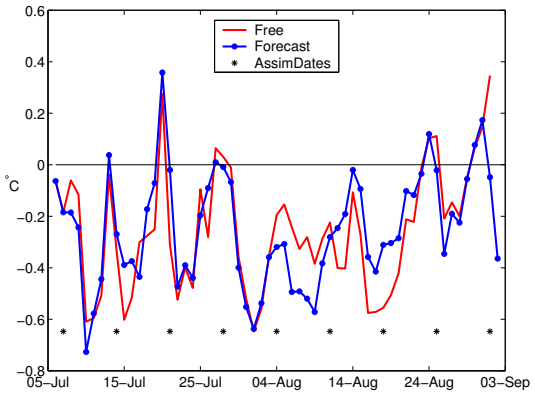

Figure 10.16: Bias evolution for the SST of the Mediterranean Sea model compared to Pathfinder v5 data.

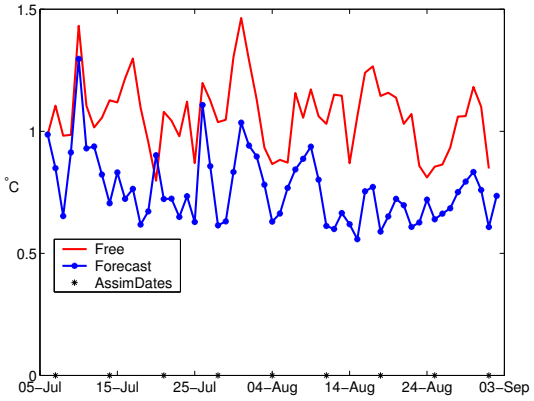

Figure 10.17: RMS error evolution for the SST of the Mediterranean Sea model compared to Pathfinder v5 data.

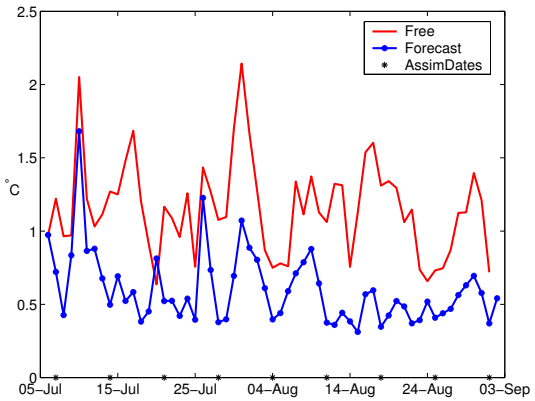

Figure 10.18: MSE evolution for the SST of the Mediterranean Sea model compared to Pathfinder v5 data.

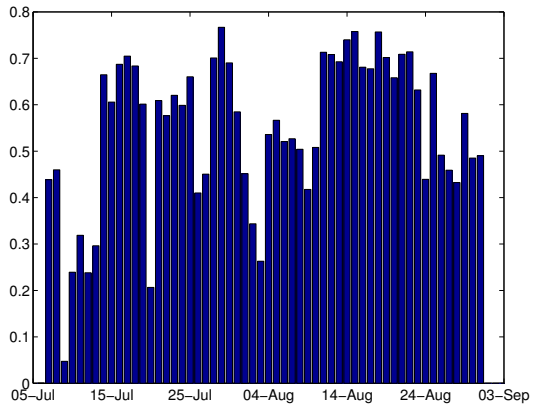

Figure 10.19: MSESS evolution for the SST of the Mediterranean Sea model compared to Pathfinder v5 data.

Table 10.4: Mean errors for the `Forecast` and the `Free run` SST for the Mediterranean model, compared to Pathfinder v5 data.

Measure	`Free run`	`Forecast`
Bias	-0.26°C	-0.24°C
MAE	$0.81°C^2$	$0.54°C^2$
RMSE	1.05 °C	0.7°C
MSE	$2.01°C^2$	$0.91°C^2$
MSESS		0.54
ACC	0.43	0.57

10.3.2 *In situ* data from Sirena Cruise

The Sirena Cruise consists of 51 CTD profiles in the Ligurian Sea. The comparison with *in situ* profiles allows us to study the impact of the assimilation of SST and SLA at depth. These measurements coincide with the 8th assimilation cycle, so the warming event that takes place at this moment that affects the skill of the model can be further studied. However, the comparison of the model with *in situ* data just in a period when the model presents low skill may give a biased perception about the model performance. Sirena cruise was already described in Section 4.2.

Temperature

First, the average of all profiles is computed, in order to study the distribution of error with depth. Figure 10.20 presents the temperature bias between the model and the observations. Both the `Free run` and the `Forecast` are about 0.5°C colder than Sirena profiles at surface. At depth the model is warmer than the observations. At the mixed layer depth (MLD), the `Forecast` presents a bias of more than 4°C, whereas the `Free run` has a bias of about 2.3°C. We can see here that the warming caused by the assimilation and atmospheric forcings on 24 August of the assimilation experiment is felt in the model, even at depth. However, the `Free run` bias at the MLD is smaller, indicating that the assimilation modifies the shape of the profiles at depth. Recall from section 4.2 that on average, Sirena data is warmer than DLR data at surface . The model bias distribution with depth indicates that the model surface temperature is colder than Sirena, but warmer at depth. This situation takes place because the model profile is misplaced respect to observations (see Figure 10.25, where the mean `Forecast`, `Free run` and observed profiles are shown for the first 100 meters of the water column), giving rise to a big error in the mixed layer. After the maximum at the MLD, the bias decreases to zero. From 75 m depth to the end of the profile, there is almost no bias.

Obviously, the RMS temperature error reflects also a high error at the MLD. It behaves in a similar way as the bias does. At surface, an error of about 1°C is found, for both the `Free run` and the `Forecast`. At depth, mainly at the MLD, the RMS error is again higher for the `Forecast` than for the `Free run`. The `Free run` has an error of about 4°C, and the `Forecast` has an error of 5°C. We will further study in this section the erroneous position of the MLD.

Salinity

The behaviour is different for the salinity. Recall that no salinity is assimilated into the forecast, so it is very interesting to study the differences between the `Free run` and the `Forecast` for this variable. Figure 10.22 presents the salinity bias for the `Free run` and the `Forecast`. The bias is negative at surface for both the `Free run` and the

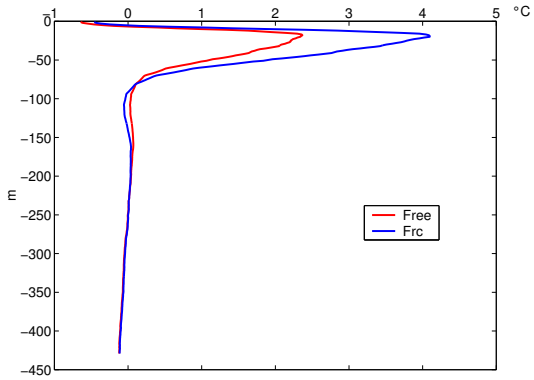

Figure 10.20: Temperature bias distribution with depth for the `Free run` and `Forecast`.

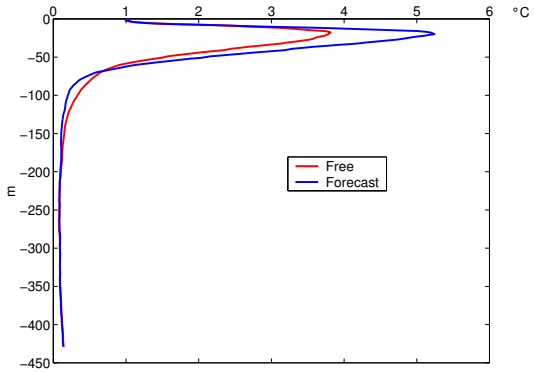

Figure 10.21: RMS temperature error distribution with depth for the `Free run` and `Forecast`.

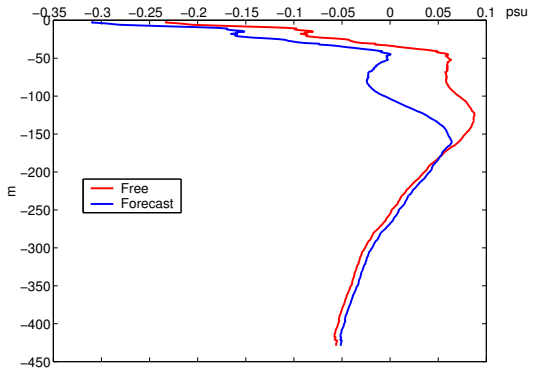

Figure 10.22: Salinity bias distribution with depth for the `Free run` and `Forecast`.

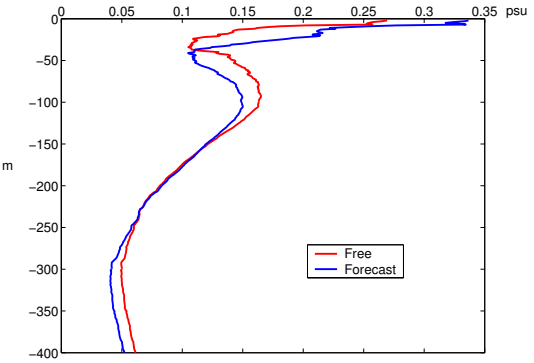

Figure 10.23: RMS salinity error distribution with depth for the `Free run` and `Forecast`.

`Forecast`, and it decreases progressively to zero in the first fifty meters of the water column. The `Forecast` is fresher at surface is bigger than the bias of the `Free run`, of about 0.07 psu. The SST and SLA assimilation deteriorates the salinity at the surface, making surface waters fresher than without assimilation. This effect is corrected with depth. If we make a zoom over the first 100 meters of the water column, we can see (Figure 10.24) that surface model salinity is too low compared to observations in the first 30 meters. At a depth of about 30-40 meters, both model and observations present finally a similar salinity, with `Forecast` closer to observations than the `Free run`. This is the cause of the minimum in the error at this depth. Below 50 meters, the bias of the `Forecast` is smaller than the bias of the `Free run`. From 200 meters to the end of the profile, the bias of both runs evolve in a similar way (Figure 10.22). However, the bias at depth is slightly negative, about -0.05 psu. It is not zero as it happens with temperature bias. There are two causes that can explain this bias at depth: the fact that no salinity is assimilated into the model or an initial condition biased with respect to Sirena salinity profiles. Figure 10.23 show the RMS error for the salinity of the `Free run` and the `Forecast`. Its behaviour is also different from what is observed in the temperature error profile. For instance, the high error observed at the MLD for the temperature is observed here as a minimum in the RMS error of the salinity. The maximum error is found at surface, where the `Forecast` has a bigger error than the `Free run`. The RMS error decreases up to 50 m depth, a little shallower for the `Free run`.

Calculation of the mixed layer depth

The high temperature error observed at the MLD may be caused by a bad representation of the water column structure. The MLD is now studied to see if its depth is well represented by the model and if the temperature of the MLD is accurate. The method used for the detection of the MLD was explained in section 6. First, we will look at the temperature profiles. In Figure 10.25, a zoom over the first 100 meters of the observed and model profiles and the bias between them is shown. We can see that, indeed, the model profiles are too deep compared to observations, as it was said before when looking to the distribution of the bias with depth. Both the `Forecast` and `Free run` are colder than observations at surface, and then at depth the profiles are "pushed down", appearing too deep. `Forecast` is even deeper than `Free run`, which can be an effect of the assimilation of a warm SST temperature.

If we decompose the error between depth and temperature, we can see the amount of the error explained by each part. In Tables 10.5 and 10.6 the bias and RMS errors for the `Free run` and the `Forecast` are presented, decomposed in depth and temperature errors.

The model error observed at the MLD is mainly explained as an error in depth, because the model mixed layer is situated too deep compared to the observed MLD

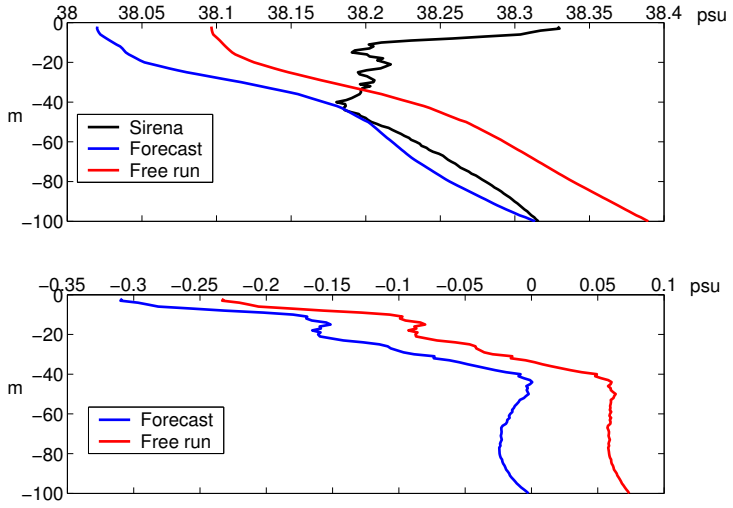

Figure 10.24: Zoom over the first 100 meters of the salinity profiles. Upper panel shows the salinity of observations, `Forecast` and `Free run`. Lower panel shows the salinity bias between observations and model runs.

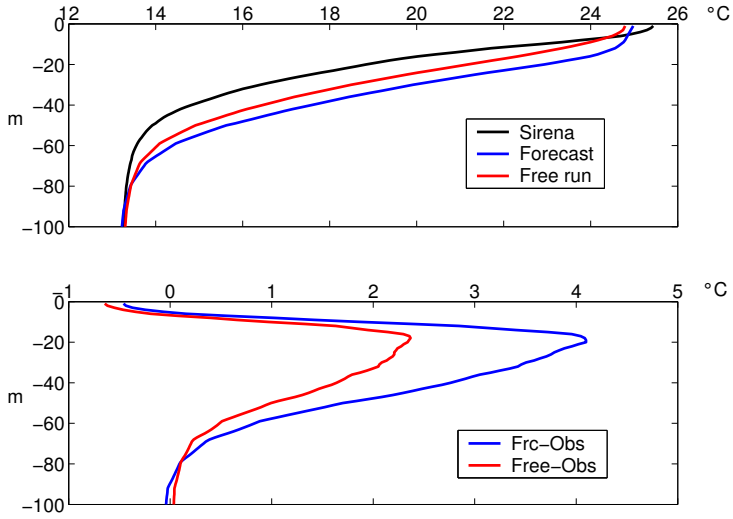

Figure 10.25: Zoom over the first 100 meters of the averaged temperature profiles. Upper panel shows the temperature of observations, `Forecast` and `Free run`. Lower panel shows the temperature bias between observations and model runs.

Table 10.5: RMS error and bias in the depth representation of the Mixed Layer.

	`Free run`	`Forecast`
RMS	12 m	16.2 m
bias	-8.6 m	-14.28 m

Table 10.6: RMS error and bias in the temperature representation of the Mixed Layer.

	`Free run`	`Forecast`
RMS	0.78°C	0.76°C
bias	-0.38°C	-0.25°C

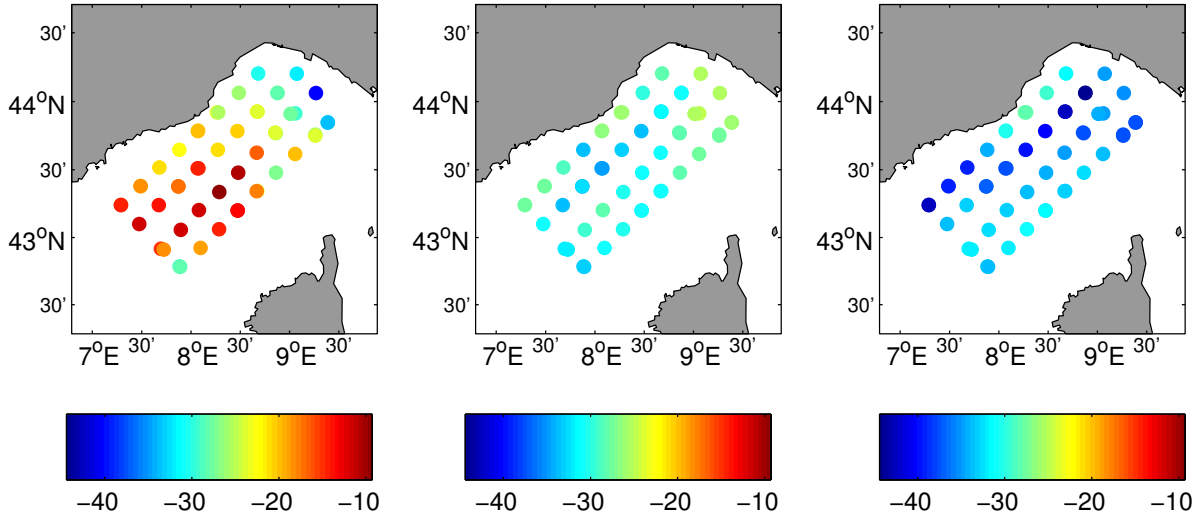

Figure 10.26: Depth of the mixed layer for Sirena profiles (left panel), `Free run` (central panel) and `Forecast`(right panel).

(negative bias in the depth error). The `Forecast` presents a deeper mixed layer than the `Free run`, so it can be said that the assimilation "pushes down" the water column. The depth of the mixed layer for Sirena profiles, the `Free run` and the `Forecast` is shown in Figure 10.26. Northern stations for the model (`Free run` and `Forecast`) agree with the depth of the mixed layer, but southern stations are too deep in the model. The `Forecast` presents overall a deeper mixed layer than the `Free run`. The temperature of the MLD (in table 10.6) is well represented, in both the `Free run` and the `Forecast`. The high errors observed when looking to the vertical distribution of temperature bias and RMS (Figures 10.20 and 10.21) are then due to a bad positioning of the MLD. The assimilation process has the effect of deepening of the mixed layer.

Depth averaged errors

The next step is to study the behaviour of the model at each of the profile locations averaged over depth. As seen in the previous figures, the error in the first 100 m is much higher than in the rest of the water column, so an integration over the whole depth will

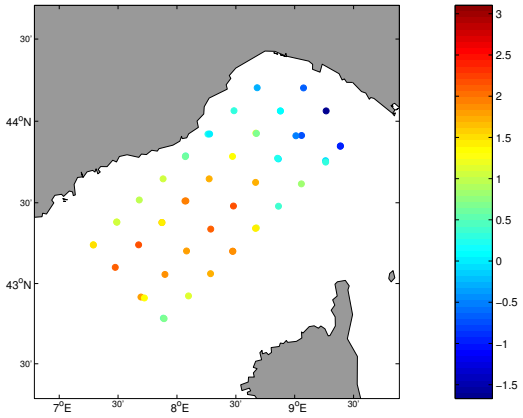

Figure 10.27: Temperature bias of the `Free run` calculated over the 100 first meters of the water column

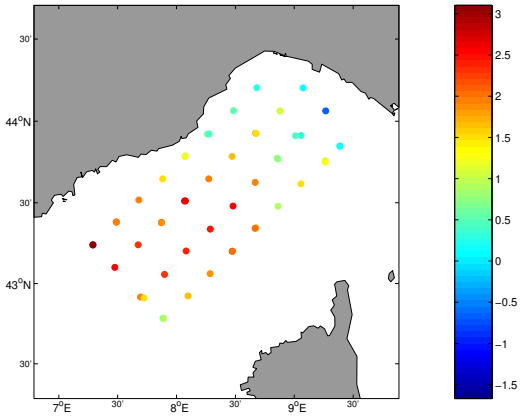

Figure 10.28: Temperature bias of the `Forecast` calculated over the 100 first meters of the water column

distort the error between the observations and the model. Only the first 100 m are used in Figures 10.27 to 10.34 to assess the temperature and salinity error at each profile location.

Temperature bias of the `Free run` and the `Forecast` is presented in Figures 10.27 and 10.28, respectively. The `Free run` presents a smaller bias than the `Forecast`, as depicted from the vertical distribution of the error. Both runs have the same bias spatial distribution. Northern stations, near to the Italian coast, present a negative bias for both the `Free run` and the `Forecast`. The model is thus colder than the observations at these locations, although the assimilation decreases this difference. The bias of the `Free run` is thus higher at these locations. The opposite is observed at the southern stations, where the `Forecast` and `Free run` are both warmer than the observations. The assimilation has here the effect of increasing the `Forecast` bias compared to the `Free run` bias. The assimilation warms overall the Ligurian basin and this is again the consequence of the maximum of temperature reached at the eighth assimilation cycle, shown in Figure 10.6 (recall that the Sirena cruise takes place just before this assimilation cycle).

When looking at the RMS distribution (Figures 10.29 and 10.30), northern stations present again smaller errors than southern ones. The RMS error is smaller also following the French coast, so that finally it can be said that open sea stations present higher errors.

In Figures 10.31 and 10.32, the bias of the salinity integrated over the first 100 meters at each station is presented. Northern stations have positive bias, and southern stations have a negative bias. Model results at northern stations (both `Free run` and `Forecast`) are thus saltier and colder than the observations. This part of the basin is influenced by waters coming from the Tyrrhenian Sea (the Eastern Corsican Current, ECC), which is warm and salty. The characteristics of this current are not accurately enough represented, although the difference is not very big. The north/south behaviour

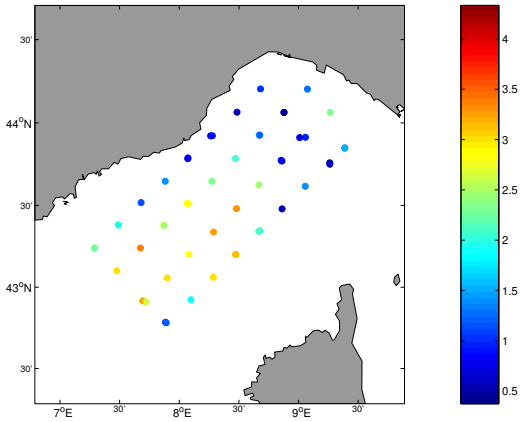

Figure 10.29: RMS temperature error of the `Free run` calculated over the 100 first meters of the water column

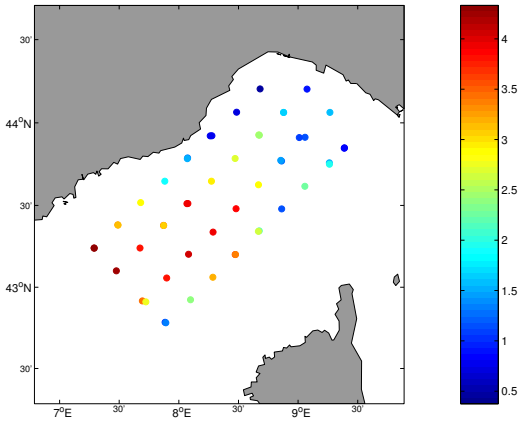

Figure 10.30: RMS temperature error of the `Forecast` calculated over the 100 first meters of the water column

is observed less clearly in the distribution of the RMS error of the salinity (Figures 10.33 and 10.34). Northern stations have higher errors than southern stations, but open sea stations also do (more exactly the third line of stations parallel to the French coast). The distribution of the error differs from what is found in the temperature RMS.

Profile-by-profile study

The spatial distribution of the error in temperature indicates that there are at least two types of profile in the model and/or the observations. A profile-to-profile study will allow us to better understand the water column structure. The relation between Sirena profiles, model profiles and the SST DLR data assimilated is studied at each of the points where a Sirena profile has been done. A detailed look at each of the Sirena profiles combined with the model results gives four different types of profile types. Figures 10.35 to 10.38 show these four different types of profiles, explained below:

- Type 1: Sirena and model profiles shape show a good agreement. Sirena at surface is warmer than DLR.
- Type 2: `Forecast` is too warm at depth; `Free run` is cold, and generally with a bad shape. Sirena at surface is warmer than DLR.
- Type 3: Sirena profile and `Forecast` are similar to type 2. `Free run` shape is similar to `Forecast`. DLR warmer than Sirena at surface.
- Type 4: Sirena and model profiles show a good agreement, as in type 1. DLR is warmer than Sirena values at surface.

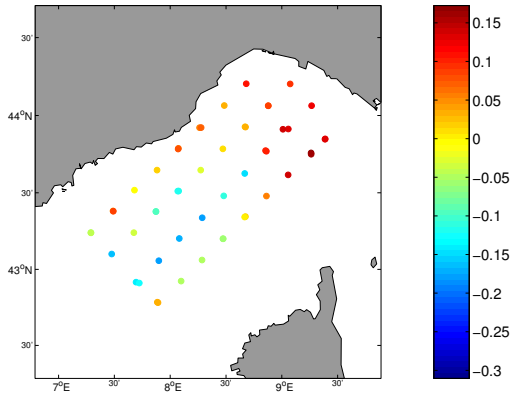

Figure 10.31: Salinity bias of the `Free run` calculated over the 100 first meters of the water column

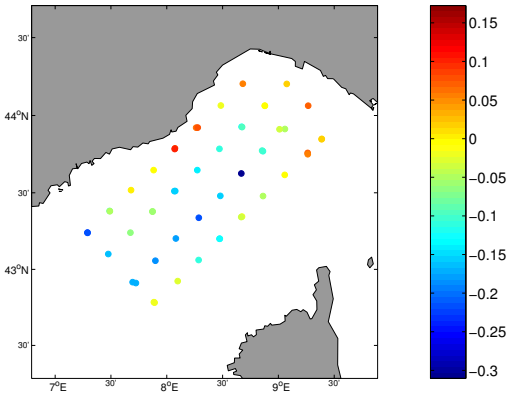

Figure 10.32: Salinity bias of the `Forecast` calculated over the 100 first meters of the water column

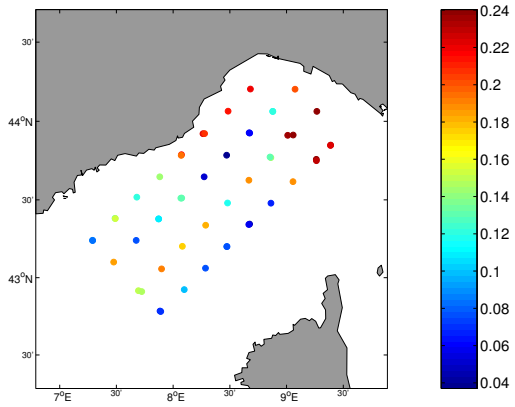

Figure 10.33: RMS salinity error of the `Free run` calculated over the 100 first meters of the water column

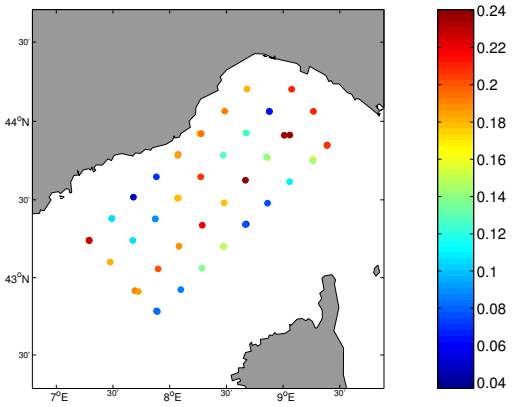

Figure 10.34: Rms salinity error of the `Forecast` calculated over the 100 first meters of the water column

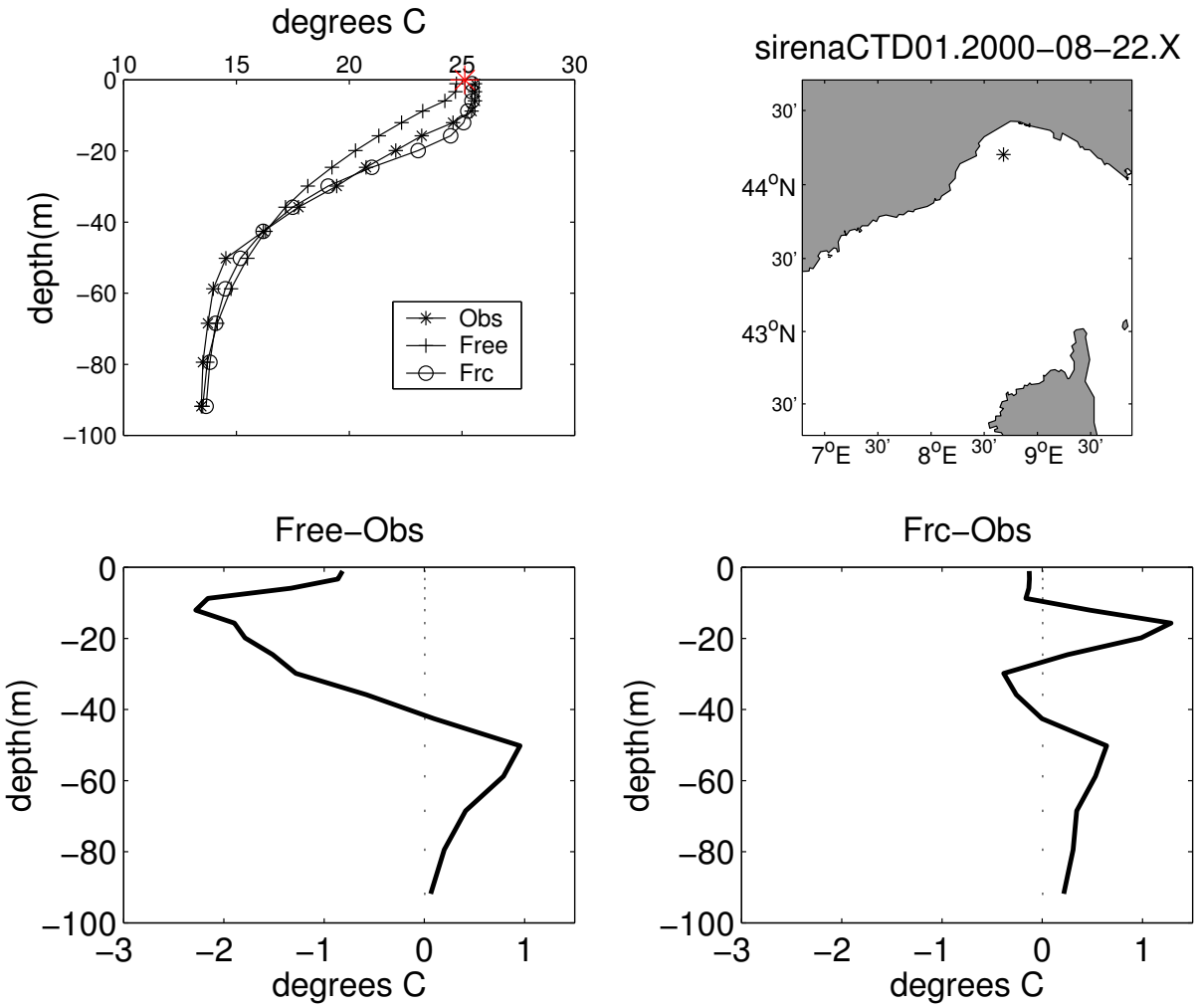

Figure 10.35: Type 1 profile. Sirena profile and comparison with the model and free run. Red asterisk in the upper left figure is the SST from DLR at this location.

These four profile types present also a specific spatial distribution: profile type 1 is found in the northern part of the basin, following the path of the Eastern Corsican Current and the Northern Current along the Italian coast. Type 2 is found mainly along the French coast. Type 3 is in open water, in the middle of the Ligurian Sea, and type 4 corresponds to the southeastern stations, situated also in open water, but nearer to Corsica. The profile locations shown in Figures 10.35 to 10.38 give an idea of the general location of each type.

In Figure 10.35 we can see type 1 profile. The `Free run` is colder than the assimilated SST, which is represented by a red asterisk at the surface, and also colder and shallower than the Sirena profile. The assimilation of SST data improves the model results, because the surface temperature of the `Forecast` is closer to observations. Type 1 profiles show

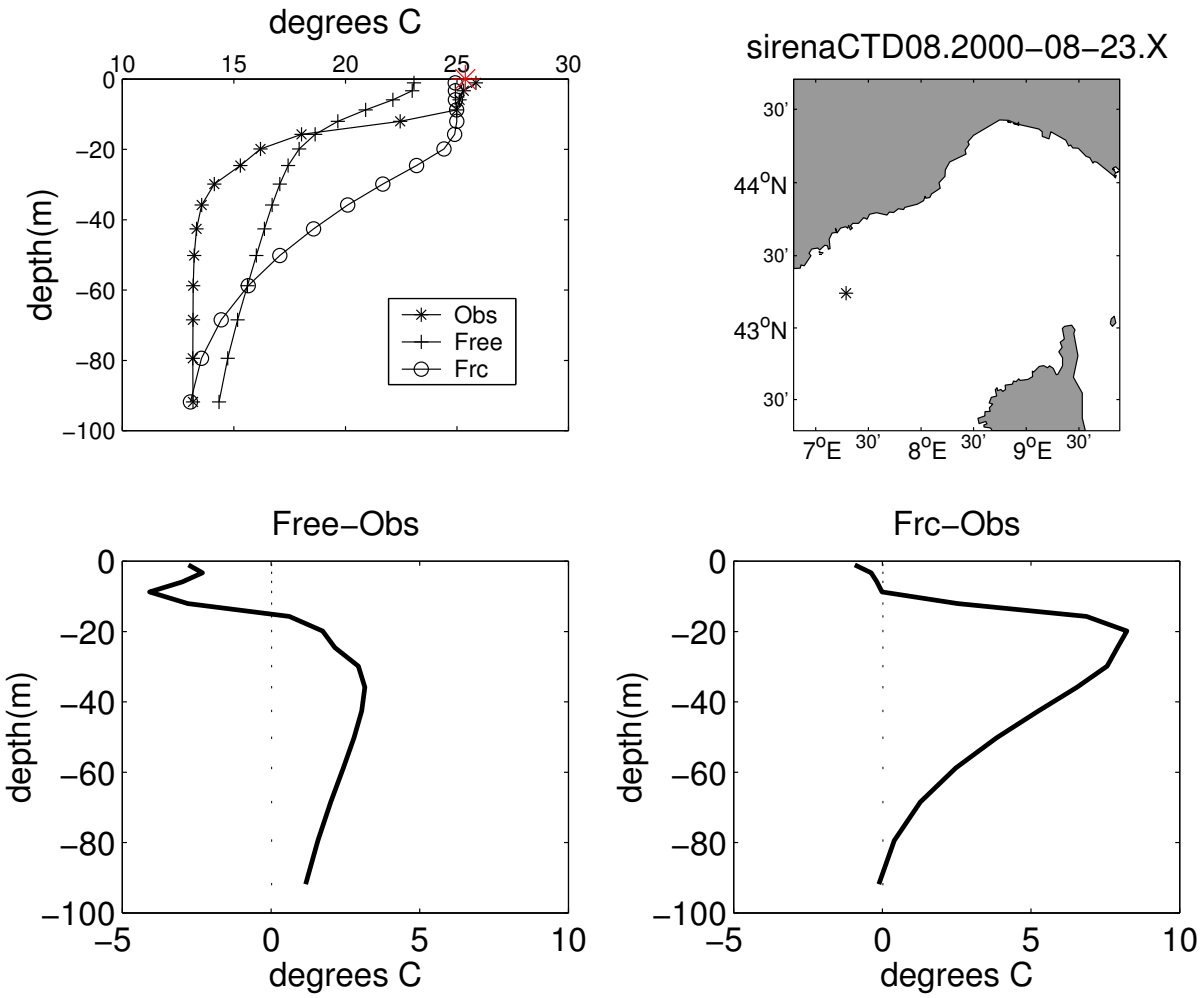

Figure 10.36: Type 2 profile. Sirena profile and comparison with the model and free run. Red asterisk in the upper left figure is the SST from DLR at this location.

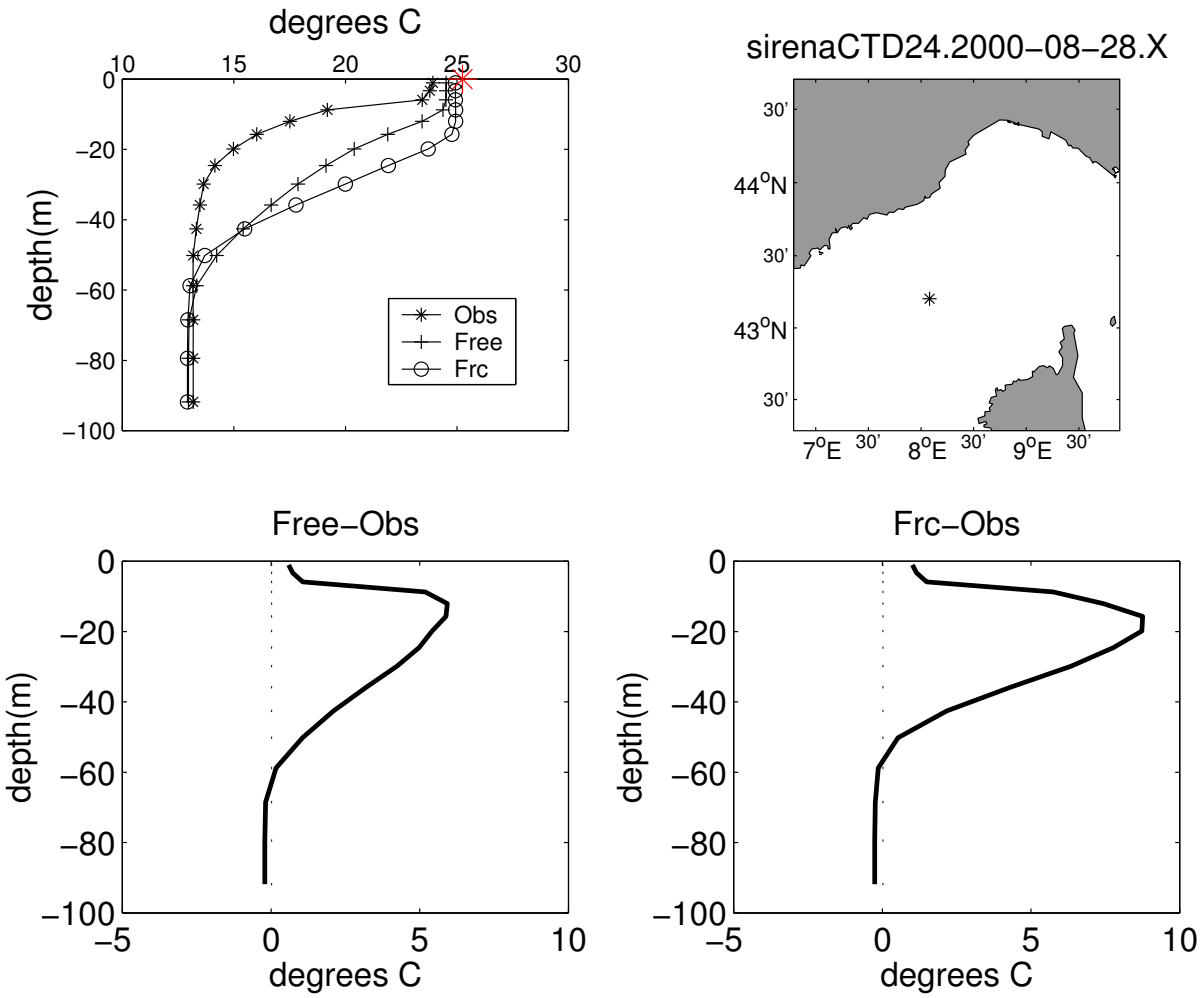

Figure 10.37: Type 3 profile. Sirena profile and comparison with the model and free run. Red asterisk in the upper left figure is the SST from DLR at this location.

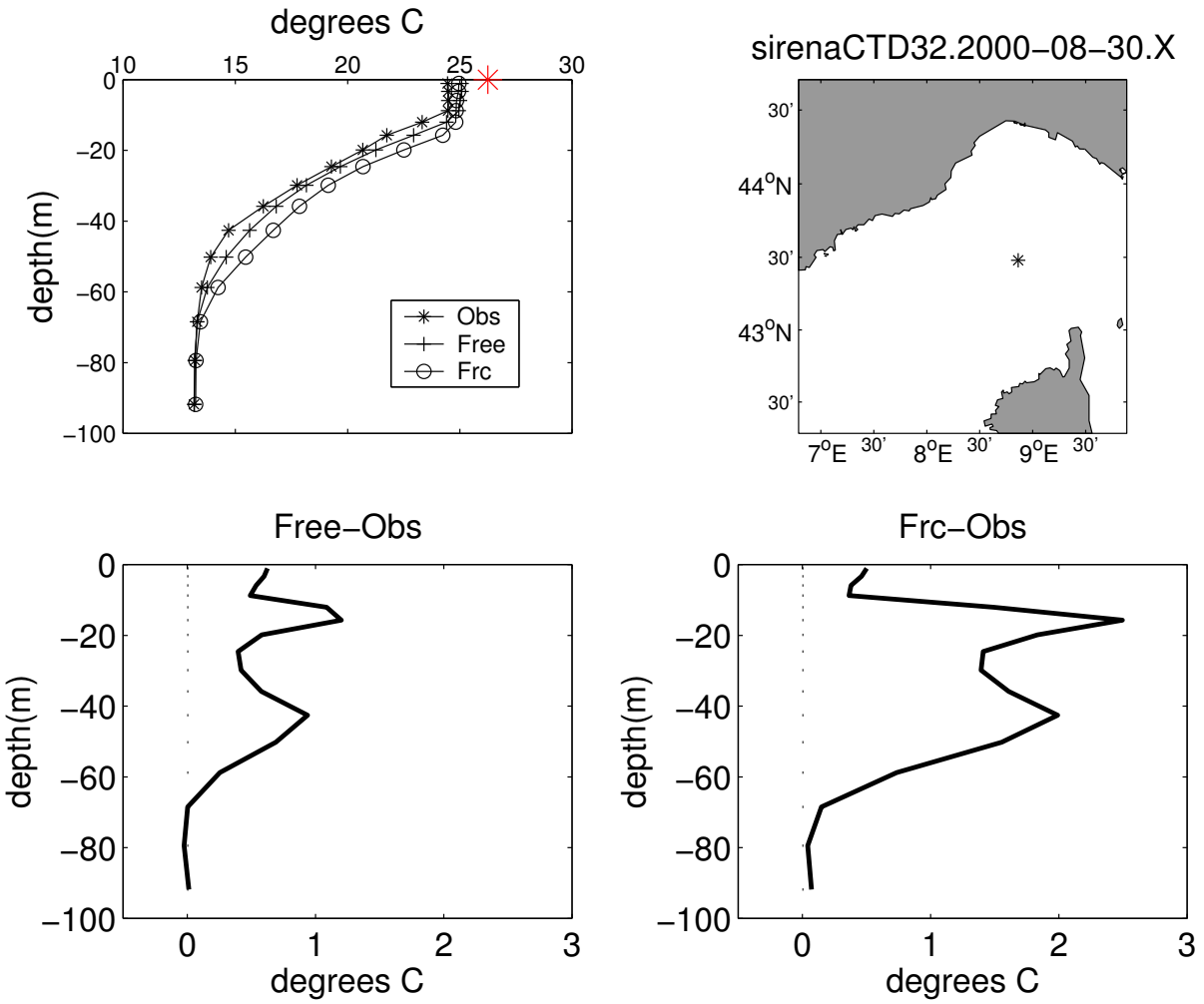

Figure 10.38: Type 4 profile. Sirena profile and comparison with the model and free run. Red asterisk in the upper left figure is the SST from DLR at this location.

a positioning of the mixed layer similar to the Sirena profiles (Figure 10.26) and a small error in temperature the first 100 m of the profile (Figure 10.30).

Type 2 profile is presented in Figure 10.36, and it shows a very different structure. Sirena MLD is shallower than in type 1, as it can be also seen in Figure 10.26. The `Forecast` profile is deeper than in type 1, so finally there is a big disagreement between observed and modelled profiles. `Free run` profiles are also very different from Sirena. They are too cold at surface and too warm under the mixed layer. Here the assimilation has a strong effect on the shape of the profile, as it warms at surface about 2°C and it deepens the whole water column.

In Figure 10.37 we can see the type 3 profile. It is similar to type 2 but now the `Free run` is too deep compared to the observed profile. The observed profile is even shallower than in type 2 profile. The assimilation of SST, warmer than the Sirena profile at surface, seems not to influence very much the model: both `Free run` and `Forecast` profiles have a similar shape, with a deep mixed layer.

Finally, Figure 10.38 shows type 4 profile, which is similar to type 1. The main difference with profile type 1 is that the DLR SST here is warmer than the Sirena profile. The `Forecast` profile is deeper than the `Free run` profile, and the stratified surface layer of the observations is more pronounced than in the other types.

10.3.3 *In situ* data from MEDAR/Medatlas database

The model results were already compared with MEDAR/Medatlas profiles in section 9, only in a visual comparison frame. We saw then that the `Forecast` agree with observations except for the temperature representation near the Strait of Gibraltar. Now we look at the positioning of the mixed layer. In tables 10.7 and 10.8 we can see the error in position and in temperature respectively. As with Sirena profiles, the most part of the error is concentrated in the position of the mixed layer, which is too shallow for the model. The `Forecast` shows a smaller bias, although the RMS error is higher than for the `Free run` profile.

The temperature error at the mixed layer is very small compared to the error at depth. The bias is positive, of about 0.3°C, and there is 1°C of RMS for both runs. The error in depth is higher for the MEDAR/Medatlas profiles that for Sirena profiles. This could be because MEDAR/Medatlas profiles are spread over the whole Mediterranean Sea, and Sirena profiles are concentrated in the Ligurian Sea, where the higher resolution model is situated.

Table 10.7: RMS error and bias in the depth representation of the Mixed Layer when comparing to MEDAR/Medatlas data.

	Free run	Forecast
RMS	16.67 m	17.2 m
bias	6.95 m	4.42 m

Table 10.8: RMS error and bias in the temperature representation of the Mixed Layer when comparing to MEDAR/Medatlas data.

	Free run	Forecast
RMS	1.15°C	1.17°C
bias	0.28°C	0.26°C

10.4 Summary and Conclusions

In this chapter, we have studied the quantitative error assessment of the model. A good choice of the applied error measures must be done in advance, in order to extract the maximum of useful information from the model. We have seen that the bias alone is not a good indicator of model skill, and the RMS error, which can be considered as a good measure, needs to be complemented as well by other measures. The addition to these two measures of the ACC and the MSESS helped us to understand the nature of some errors.

The quantitative error assessment has been done first in comparison with the assimilated variables. This has allowed us to establish the impact of the assimilation on the model. We have seen that the `Forecast` SST behaves better than the `Free run` SST. This improvement is very small in the Ligurian Sea model, because of a warming effect that takes place the two last weeks of the experiment. This warming effect is due to a shift between the observed temperature and the air temperature that is used as atmospheric forcing. Both the `Free run` and the `Forecast` present this bigger error at the end of the experiment, although in the `Forecast` this error is bigger. The assimilation of SST deteriorates the model skill at the end of the experiment, and this is at the expenses of the overall skill of the `Forecast`, which is on average only slightly better (about 10% better) than the `Free run`. When we assess the error of the model at the whole Mediterranean Sea model, this effect is smaller, so the improvement of the `Forecast` over the `Free run` is higher, of 23%.

We have also seen the effect of the grid refinement on the model skill, which is usually called “double penalty”. The finest grid model represents more accurately the observed features, but the model can be penalised twice if these small-scale features are shifted respect to the observed position. Because of this double penalty, and because of the small influence of the warming event at the end of the experiment, the Mediterranean model presents better skill than the Ligurian Sea model.

SLA data have also been compared to the model results. The `Forecast` is overall

better than the `Free run` in terms of RMS error. When looking at the point-by-point difference between `Forecast` and observations we can see that it is homogeneously distributed, with only some points near the coast that present very high errors. However, the skill of the `Forecast` compared to the `Free run` is not very high, thus the improvement reached by SLA assimilation is not very high.

After the comparison with assimilated data, it is necessary to establish the skill with independent, non-assimilated data. Independent data have allowed us to establish the skill of the model without ambiguity. First, Pathfinder v5 daily SST is used. This data set is not exactly independent, because it comes from the same source as the AVHRR SST data DLR assimilated data. We use them because they have a very good coverage in time and space (the whole Mediterranean Sea at a daily time step), which make them very valuable for forecast verification purposes. The comparison of the model with this data has shown us the influence of the weekly data assimilation, and the slight drift that the model suffers between two assimilation cycles. The improvement of the `Forecast` over the `Free run` is very high (57 %).

In order to assess the model quality at other zones than the surface, *in situ* profiles from Sirena cruise and from the MEDAR/Medatlas database are used. The use of *in situ* profiles helps us to see the impact at depth of the assimilation of surface data. Sirena cruise data have been taken between the two assimilation cycles with a high `Forecast` error. Sirena data were thus used to better understand this event, but the obtained skill was not representative of the whole experiment.

When studying how the error is distributed at depth, we have observed that a significant contribution of the error is concentrated at the Mixed Layer Depth (MLD). The split of the error at the MLD in a positioning error and a temperature error has shown us that the MLD is too deep compared to observations, but the temperature of the MLD is well represented by the model. The assimilation deepens the water column, thus making the `Forecast` skill poorer than the `Free run` skill.

Overall, the quality of the `Forecast` is better than the `Free run`. Because of punctual big problems, such as the warming effect at the end of the experiment, the `Forecast` and the `Free run` skill are very similar, with RMS error of about1°C in surface temperature. The `Free run` is even better than the `Forecast` at some weeks in terms of RMS, namely on 24 and 31 August. However, other error measures, as the ACC give a good score for these weeks. This highlights the importance to use different error measures to assess the skill of a model, because we can have a limited view (too good or too bad) that will influence the future modifications on the model. A feedback between the modeller and the "verificator" (if they are not the same person) is indeed very important to improve the performances of a model.

However, the spatial distribution of the error has not yet been studied. To do so,

spatio-temporal techniques should be applied to the model results. The study of the model results behaviour in a spatio-temporal frame will complete our study about the model skill.

Chapter 11

Spatio-temporal analysis

In this section, the analysis of the model results is done in a spatio-temporal approach. First, the model and observations are decomposed into eight different resolution levels, using Wavelet Transforms. This analysis is applied to the Ligurian Sea model SST. The study of the model at different spatial scales can reveal some valuable information about its behaviour that cannot be studied with a traditional error assessment. At each resolution scale, the error measures usually applied to assess a model error can be used, so we obtain the information of how each of these scales is contributing to the total error. The scales can be easily identified to physical processes at small, meso and large scale. We can know this way if there are specific processes that are the cause of the error in the model.

After the wavelet decomposition, a wide used spatio-temporal technique is used. EOF decomposition of the SST will allow us to study how the model represents the variability compared to observations. EOFs capture the main patterns of variability in a data set and will help us to compare the behaviour of the model to the observations.

This chapter is thus divided into two main sections: first, the wavelet analysis of the Ligurian Sea SST is done, followed by the study of the main EOFs of the model.

11.1 Wavelet Analysis

The wavelet analysis results will be presented in this section. We have decomposed model and observations SST into eight resolution levels with a Haar wavelet. The initial resolution of the observations is 1 km, and the model has also been interpolated to this resolution. Thus, the first wavelet decomposition of the model and observations has 1 km of resolution (the initial resolution of the model times 2^0, scale 1). Subsequent decompositions are 2 (2^1, scale 2), 4 (2^2 scale 3, etc.), 8, 16, 32, 64 and 128 km. These resolution levels were described in table 7.1, in chapter 7. Once the decomposition is done, the traditional error measures introduced in chapter 5 are applied to each of these resolution

levels. Note that only Details are used and not the Approximations. This is because Details contain the information of one specific resolution level, and the Approximation contains the cumulative information of all scales coarser than the one we analyse. The use of Details allows us to analyse only features of a given scale independently from the others.

As it has been mentioned in chapter 7, the analysis is made in the data space rather than in the wavelet space. This is achieved by isolating the Details of a given scale and then by applying the Inverse Wavelet Transform to them. The analysis can be performed in both spaces, the wavelet and the data space (Briggs and Levine, 1997). We choose to work in the data space because the number of land points remains constant at all scales, making easier the interpretation of the results.

We concentrate the study on the higher resolution grid to assess the impact of the nesting and the assimilation simultaneously. In addition, land points mainly cover the half-upper left part of the domain (the French coast). Land ponts are thus concentrated in one part of the domain, so they have smaller impact on wavelets. All data sets are transformed into a dyadic field by padding with zeros. In the meridional direction, the padding is done in the upper zone of the matrix, so that they extend the already existing land points. In the zonal direction, the padding is done at the western side. The final matrix has a size of 256×256 pixels, for each time instant. An example of these padded matrices can be seen in Figure 11.1. There are nine SST fields, corresponding to the nine assimilation cycles.

The effect of boundary treatment explained in section 7.6 is directly seen in the wavelet coefficients. In Figure 11.2 we can see the wavelet coefficients (in data space) for the three first spatial scales, with and without boundary treatment. The major problems are found in the artificial boundaries at the west of the domain that result from the padding made to obtain a dyadic image size. Wavelet coefficients are also perturbed at the natural boundaries of the domain.

However, the impact of the boundary treatment on the error measures is very small. In Figures 11.3 and 11.4, we can see the SST RMS error at each of the spatial scales for the case where no boundary treatment has been made and for the case with boundary treatment, respectively. The error on smaller scales is reduced by the boundary treatment, although the error decrease is very small. At larger scales, the decrease in the error is even smaller. The distribution of the RMS error through the different scales is similar for both cases (with and without boundary treatment), indicating that the robustness of the boundary treatment. Further work should be done to minimise the boundary effect. Some work has been done in the frame of wavelets on irregular domains, as the work by Sweldens (1997) and Daubechies *et al.* (1999) on second generation wavelets, or the work by (Oh *et al.*, 2001; Naveau and Oh, 2004) on polynomial wavelet regression. This last approach is currently applied to one-dimensional data sets, but for two-dimensional data sets work is still ongoing (Hee-Seok, *personal communication*).

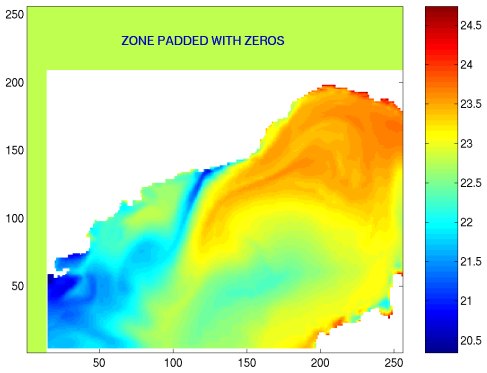

Figure 11.1: Matrix padded with zeros to obtain a dyadic size.

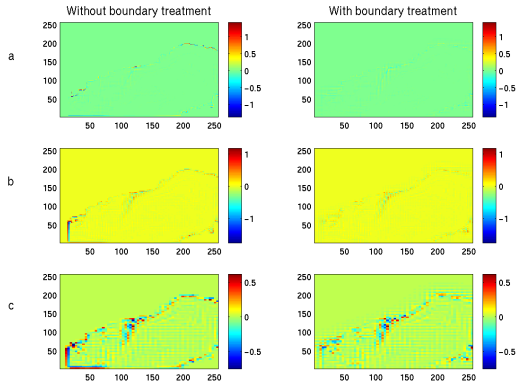

Figure 11.2: Three first scales wavelet coefficients without boundary treatment (left side) and with boundary treatment (right side). (a) is the finest scale, at 1 km resolution, (b) is the second scale with 2 km resolution and (c) is the third spatial scale, with 4 km resolution.

Standard deviation

We begin our analysis by looking at the SST standard deviation (SD) of each spatial scale, to see how it is distributed. Larger scales have higher SD, as we can see in Figures 11.5 and 11.6. Larger scales account for more energetic processes than small scales, so their inherent variability is higher. By variability we mean the statistics of the differences between the model and the observations. Both model and observations evolve in time, and have been decomposed into eight spatial scales. This definition is similar to the one given by Lermusiaux (2001). For example, in recent years it has been a change on the places where deep water formation occurs in the Mediterranean Sea. This change includes processes at large scales (basin and sub-basin scale) so its impact is very high in terms of heat exchange between sub-basins. On the other side if a small-scale feature, as an eddy, changes its position the effect will not have the high impact of the previous example.

Figure 11.5 shows the SD for each of the scales at each of the nine weeks of the experiment. Figure 11.6 shows the average over the nine weeks. The SD distribution for each of the nine weeks (Figure 11.5) follows a clear trend (larger scales have higher SD) with all the model curves (`Free run`, `Forecast`, and `Analysis`) following the structure of the observations. However, SD distribution at weeks 6, 7 and 8 is different from the behaviour observed the rest of the experiment. The SD of the observations does not present the structure of the previous weeks: it is weaker and the difference between small

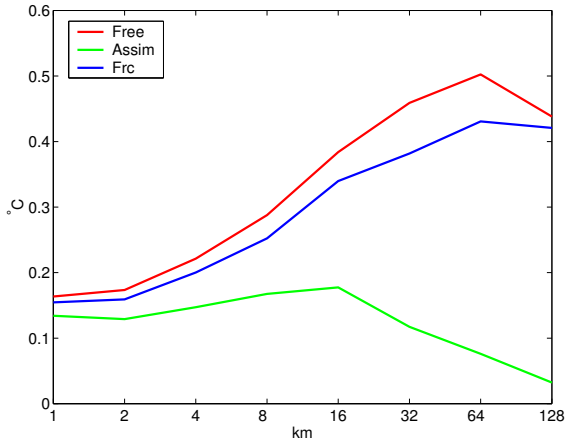

Figure 11.3: SST RMS error at each scale without boundary treatment.

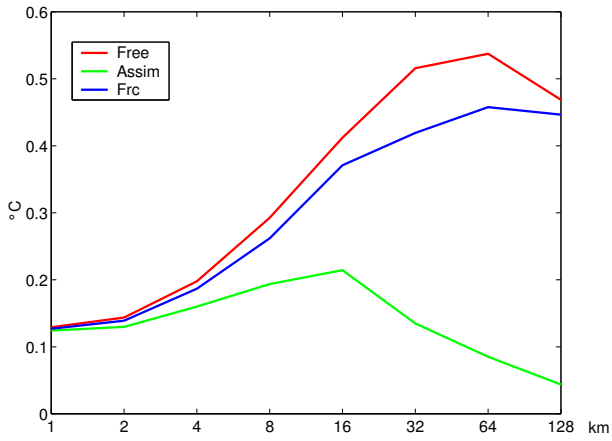

Figure 11.4: SST RMS error at each scale with boundary treatment.

and large scales is no longer as high as previous weeks. In fact, we can see a decreasing trend in the SD of larger scales from weeks 2 to 8. Recall that week 8 (24 August) presents a high RMS error (10). The distribution of SD among scales is different than the rest of the experiment.

When looking at the time averaged SD (Figure 11.6), we can see that the SD is underestimated at all scales for both the `Forecast` and the `Free run`, although this later is nearer to observations. The biggest difference between observations and model is at scales one to three (resolution of 1 to 4 km). The analysis step causes a decrease of the model variability, instead of bringing it closer to observations.

The fact that scale 8 presents a small SD for all data sets is again due to the presence of land points. Scale 8 is represented by 4 pixels covering the whole Ligurian basin, yet nearly half of the points of the initial matrix correspond to land. The small amount of sea points representing scale 8 is the cause of its small SD.

Root Mean Square Error

The RMS follows the same trend that has been observed with the SD, with higher RMS error at large scales (Figure 11.4). We can look now to the time evolution RMS for each of the spatial scales, in Figure 11.7. The RMS distribution at each of the assimilation cycles is very similar, and the differences at weeks 6, 7 and 8 seen in the SD are less evident here. The second week has the highest RMS error, as is has been already seen in the quatitative error assessment of chapter 10. One thing that should drawn our attention to is the fact that the `Analysis` error does not follow the trend "larger scales, higher error". The assimilation has a bigger effect on larger scales. Physical features at

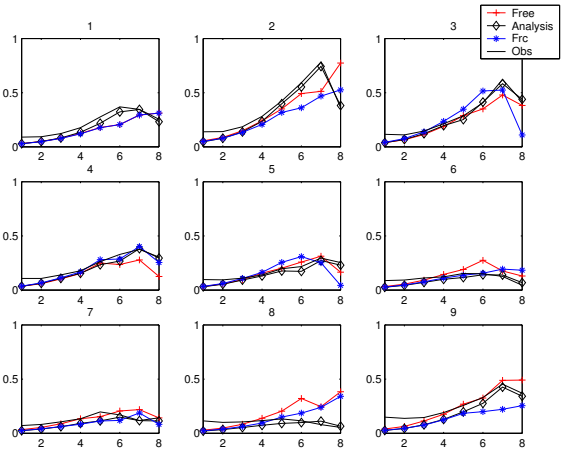

Figure 11.5: SST standard deviation at each of the spatial scales, for `Free run`, `Analysis`, `Forecast` and observations, for each of the nine weeks of the experiment (weeks are specified at the top of each figure).

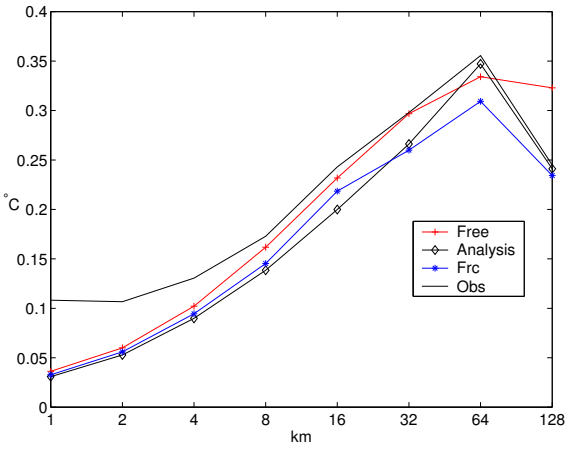

Figure 11.6: SST standard deviation at each of the spatial scales, for observations, `Free run`, `Analysis` and `Forecast`, averaged over the nine weeks of the experiment.

large scales are easier to forecast, as they are more stable in time, thus more predictable. The assimilation of satellite data corrects the model big trends, but highly variable, small-scale processes are more difficult to control. We can look at this more carefully.

Error relative to scale

An error measure that takes into account the variability of each spatial scale is desirable. A more talkative way to look at the RMS error is by normalising by this intrinsic variability. For this, the Sum Square Error (SSE) is used rather than the RMS. SSE is the squared sum of the differences between model and observations:

$$SSE(n) = \sum_{n=1}^{N} \left(x_n^f - x_n^o\right)^2 \tag{11.1}$$

The SSE can be divided by the observations variability obtained at this scale. This weighted error measure allows us to examine the contribution of each scale to the total error, without the effect of the increasing variability. We can call this error the *error relative to scale*.

$$\text{error relative to scale}(n) = \frac{\sum_{n=1}^{N} \left(x_n^f - x_n^o\right)^2}{S_o} \tag{11.2}$$

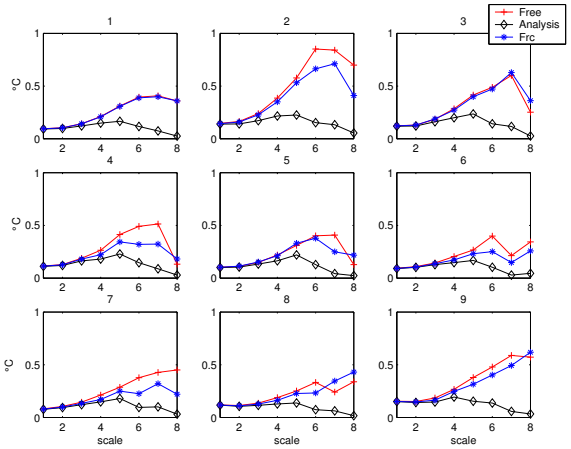

Figure 11.7: SST RMS at each of the spatial scales, for `Free run`, `Analysis`, and `Forecast` for each of the nine weeks of the experiment (weeks are specified at the top of each figure).

with S_o the variability of the observations. The result is presented in Figure 11.8 for the temporal evolution of the *error relative to scale*, and in Figure 11.9 for the time averaged error. Smaller scales have a higher relative error than larger scales, just the opposite as with RMS error. Again, the `Analysis` step has a big impact on scales 3 to 8 (from 4 to 128 km), but at scales 1 and 2 the `Analysis` decreases the skill of the model. Due to this effect, the `Forecast` has a higher *error relative to scale* than the `Free run` from 1 to 8 km, (scales 1 to 4) and the opposite happens for larger scales.

We have seen in both the SD and the *error relative to scale* that the assimilation has a positive impact on large scales, decreasing the error presented by the `Free run`. However, small scales do not benefit from the assimilation in the same way, as they are degraded in the `Forecast`. This degradation of the `Forecast` skill will be further studied in this work.

Anomaly Correlation Coefficient

The correlation is an error measure that it is not affected by the inherent variability to each scale. The Anomaly Correlation Coefficient (ACC) will be again used in this section. The distribution of ACC along the nine weeks of the experiment and the time averaged ACC are shown in Figures 11.10 and 11.11. We can see that weeks 6, 7 and 8 have a good behaviour, and that the other weeks the ACC is worse, a behaviour that was also seen in the quantitative error assessment in chapter 10. The time averaged ACC (Figure 11.11) shows better correlation at small scales, although large scales do not present a very low ACC.

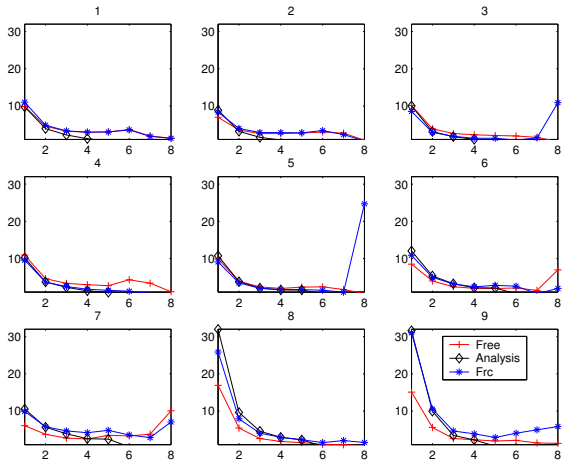

Figure 11.8: SST error relative to scale at each of the spatial scales, for `Free run`, `Analysis` and `Forecast`, for each of the nine weeks of the experiment (weeks are specified at the top of each figure).

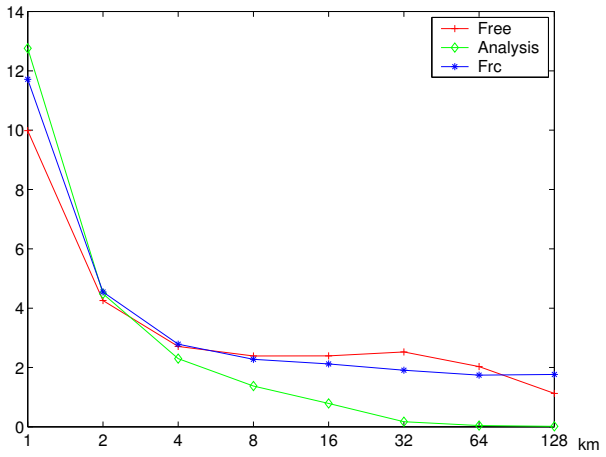

Figure 11.9: SST error relative to scale at each of the spatial scales, for `Free run`, `Analysis` and `Forecast`, averaged over the nine weeks of the experiment.

Wavelet Coefficients

The identification of the processes represented by the wavelet decomposition at each spatial scale can help us to understand the results. The detail coefficients for the four first scales (from 1 to 8 km) are shown in Figures 11.12 to 11.23 for the `Free run`, `Forecast` and observations, respectively. We can see that the observations and the model present the highest differences at small scales. Observations field presents a random distribution at scale 1 (Figure 11.14) while the model (both `Free run` and `Forecast`, in Figures 11.12 and 11.13 respectively) present coherent fields with some structures already visible. We can see, for example, how the Northern Current deviates southward in the `Free run`, and other weak signals mainly situated near the NC. Scale 1 represents a resolution of 1 km, and there is no coherent features present at this small scale in the observations (it is mainly noise). On the other hand, the model always presents correlated structures, even at the smallest scales, so they are very different from the observations in the distribution of SST details.

The second resolution scale shows us again small coherence in the observations field (Figure 11.17), although we can already identify the WCC and the NC in the detail coefficients. `Free run` is shown in Figure 11.15, and the deviation of the NC is again the major feature that can be identified. At this scale, the signal of the NC is already visible from the north of the basin, where it is formed. In the `Forecast` (Figure 11.16), the NC signal is not as strong as in the `Free run`. The cyclonic structure of the SST can be

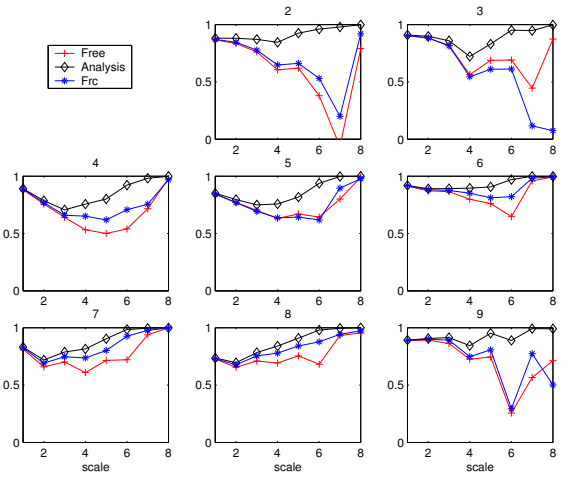

Figure 11.10: SST ACC at each of the spatial scales, for `Free run`, `Analysis`, and `Forecast`, for each of the nine weeks of the experiment (weeks are specified at the top of each figure).

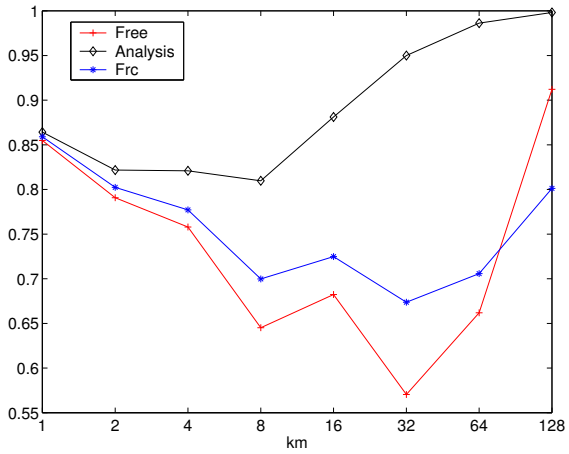

Figure 11.11: SST ACC at each of the spatial scales, for `Free run`, `Analysis` and `Forecast`, averaged over the nine weeks of the experiment.

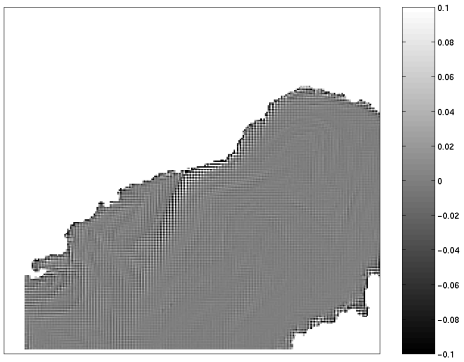

Figure 11.12: `Free run` coefficients at scale 1.

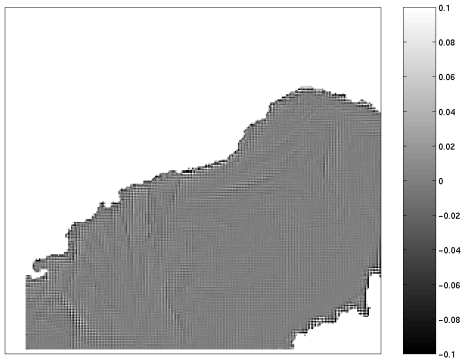

Figure 11.13: `Forecast` coefficients at scale 1.

already detected at this scale. In the `Forecast` we detect a signal at the southeast of the basin that is not present in the observations, and that can be also due to a deviation of the NC. The proximity to the model boundary can affect the path of the NC, leading to some perturbations as the observed in the `Free run` and `Forecast`.

Scale three (in Figures 11.18, 11.19 and 11.20, for `Free run`, `Forecast` and observations respectively), shows clearly the cyclonic circulation path in the three data sets. The three fields are similar now, as the large features are more easily representable.

We have seen that model smaller scales differ more from the observations that at scale 4, for example in terms of standard deviation. This is what was also obtained when computing the *error relative to scale*, which is bigger for small scales. The larger the scale, the more the structures are well represented. However, the ACC is slightly better for small scales, so that the trend of small scales appear easier to forecast.

Taylor Diagram

Wavelet decomposition is a useful tool to study the model results at different resolution levels. However, the results are complex to analyse, because there are many factors that influence each scale, and a different behaviour for each scale and error measure. For example, why the *error relarive to scale* is high for small scales and the ACC is good?, or what is the impact of assimilation at each scale?. Taylor diagram (Taylor, 2001) can be very useful in this situation, because it groups all error measures used in this chapter into one graphic. We have represented the results obtained above for `Forecast` and `Free run` on a Taylor diagram. It includes the SD, the correlation and the centred RMS (see chapter 5 for a detailed explanation about the construction of a Taylor diagram).

Figure 11.24 contains the Taylor diagram for `Free run`, `Forecast` and `Analysis` for the eight scales. We can see in this graphic the position of each scale relative to the observations that have been normalised by its standard deviation. Observations are thus at the x-axis and have a standard deviation of 1. `Forecast` and `Free run` SD have been also normalised by the observations standard deviation. Small scales have better correlations and a SD smaller than observations. The larger the scale, the worse the correlation and the best the SD.

This graphic shows us also the impact of the assimilation on the model, because the evolution from the `Free run` to the `Forecast` at each scale is highlighted. The effect of the assimilation is higher at large scales, where there is a higher difference between `Free run` and `Forecast`. Small scales show smaller displacement from the `Free run` to the `Forecast`. The assimilation has the effect of decreasing the SD of the model. The diffusion imposed by the model is thus too high, in both the `Free run` and `Forecast` as both have the same behaviour from large to small scales. However, the assimilation

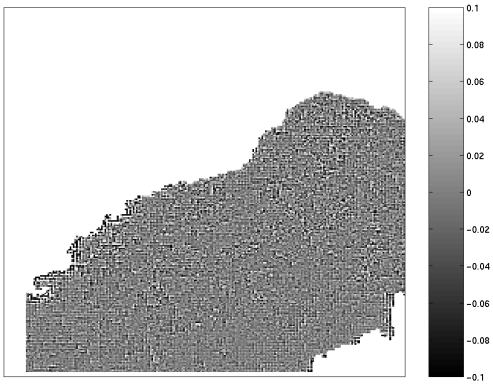

Figure 11.14: Observation coefficients at scale 1.

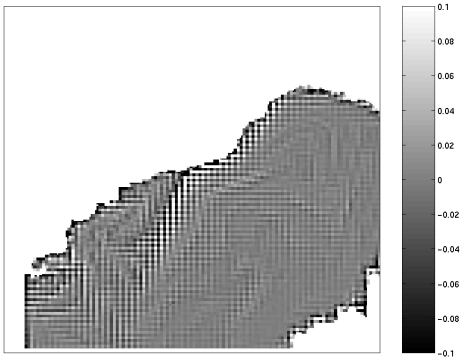

Figure 11.15: `Free run` coefficients at scale 2.

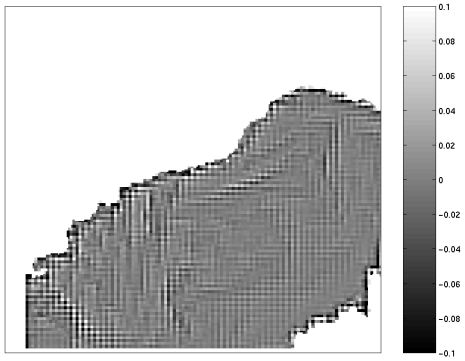

Figure 11.16: `Forecast` coefficients at scale 2.

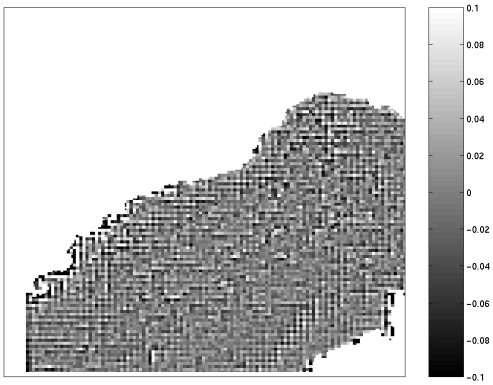

Figure 11.17: Observation coefficients at scale 2.

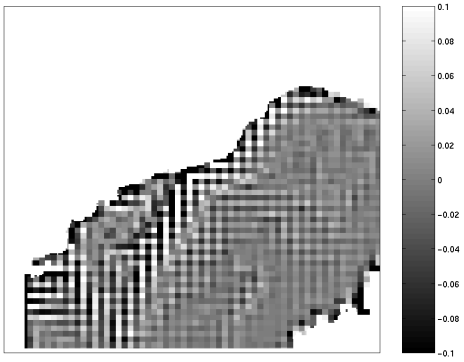

Figure 11.18: `Free run` coefficients at scale 3.

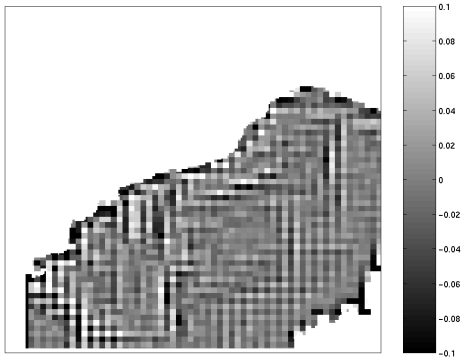

Figure 11.19: `Forecast` coefficients at scale 3.

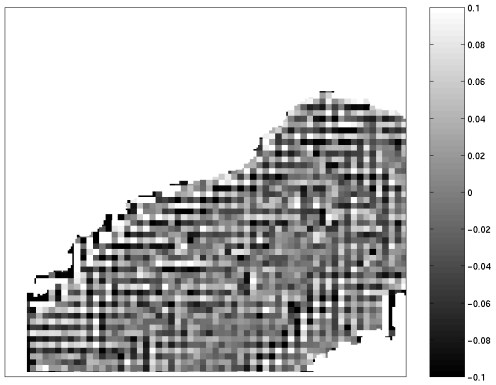

Figure 11.20: Observation coefficients at scale 3.

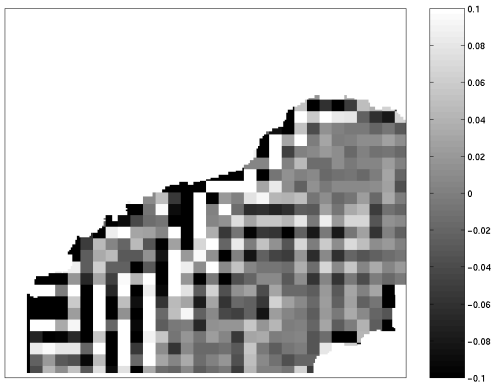

Figure 11.21: `Free run` coefficients at scale 4.

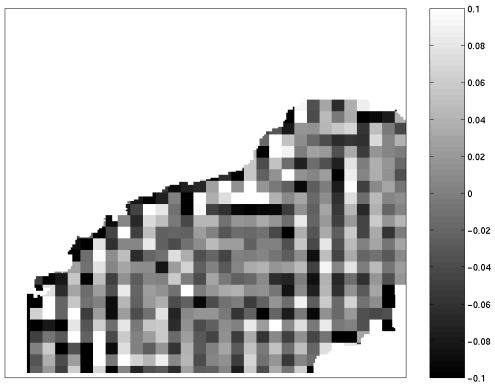

Figure 11.22: `Forecast` coefficients at scale 4.

accentuates the effect of the diffusion, as `Forecast` has a smaller SD at each scale. We can also see in Figure 11.24 that the `Analysis` (in green) has a positive effect on the correlation at large scales and only a slight effect at small scales.

Summary

Wavelets have allowed us to take an insight into the model skill at different resolution levels, rather than the more general study done in the previous chapter. We have seen that the small scales are more difficult to forecast, as they present higher errors. The assimilation has a good impact on large scales, but small scales skill is decreased by it. We have also seen that the model imposes a big diffusion, that is mainly felt in small scales.

11.2 EOF analysis

The next step to assess the goodness of the model in a spatio-temporal approach is to look at the dominant EOF modes. In this case, the whole Mediterranean model grid is used. EOFs are not affected by irregular boundaries, because only sea points are retained for the construction of matrix $\mathbf{X}$ (Equation 8.1), which is an advantage in this case because we are dealing with irregular land-sea boundaries. Recall that for the computation of EOF no missing data can be present in the analysed matrix. In this section, the Pathfinder v5 data reconstructed in section 8.4 are used for the comparison with model results.

Model and Pathfinder v5 data have different resolutions, so the number of points representing the Mediterranean Sea is different for each of these matrices. Pathfinder v5 has 139188 sea points and the Mediterranean grid model has 4155 sea points. Both matrices have 60 time instants. For the direct comparison between the EOFs obtained with Pathfinder and the EOFs obtained with the model results, the ratio between the number of points must be considered. The relationship between the number of points and the singular values ($\mathbf{S}$) of the EOFs based on the observations and model for this case is:

$$\frac{\mathbf{S}_{\text{model}}}{\mathbf{S}_{\text{obs}}} \sim \frac{\sqrt{N_{\text{model}}}}{\sqrt{N_{\text{obs}}}} \sim 0.17 \tag{11.3}$$

The right EOFs $\mathbf{U}$ are then multiplied by the singular values $\mathbf{S}$ when displayed to take into consideration the difference in the number of points. For the computation of the EOFs, the two-month time average is subtracted at each grid point of the matrices. The time average differs slightly from one matrix to another, and they were presented in

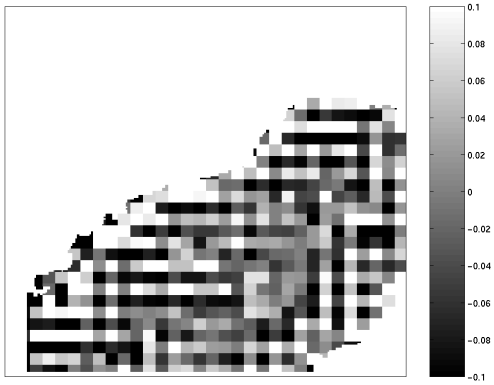

Figure 11.23: Observation coefficients at scale 4.

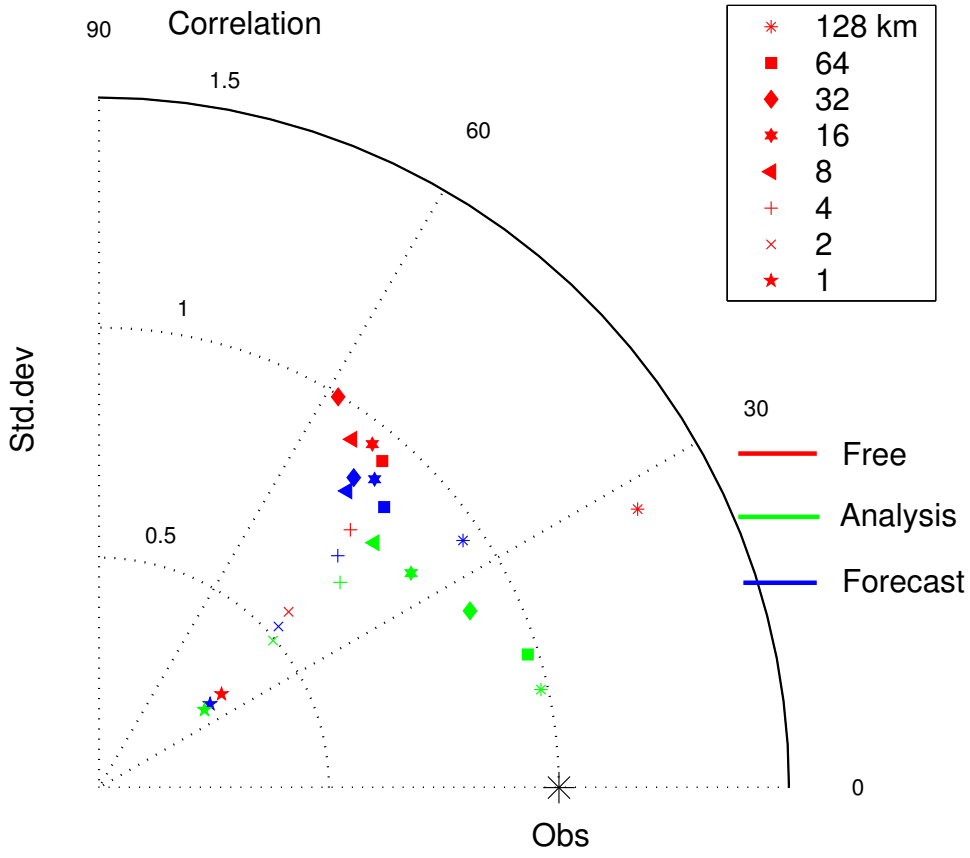

Figure 11.24: Taylor diagram for `Free run`, `Analysis` and `Forecast` for each resolution level.

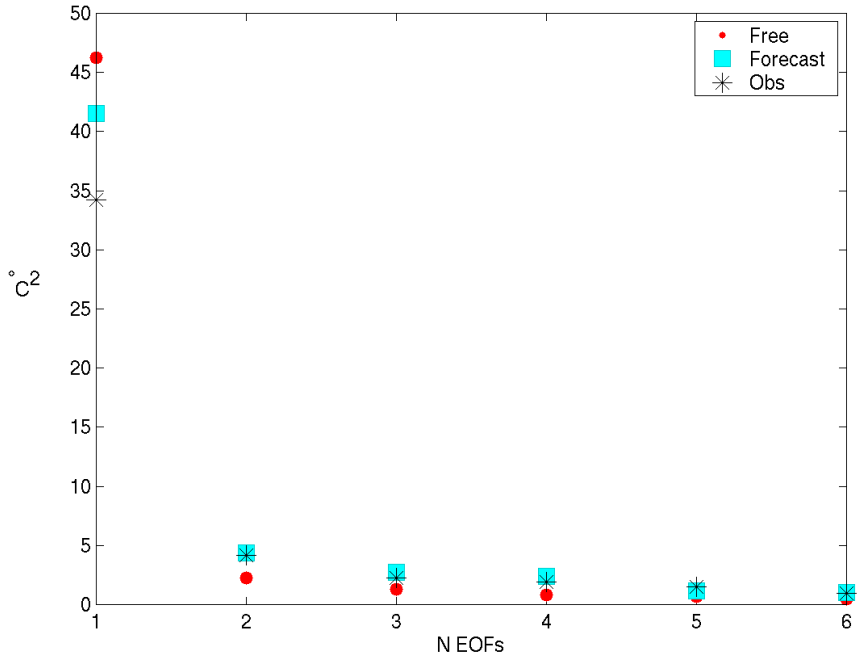

Figure 11.25: Variance of `Free run`, `Forecast` and observations for each EOF mode

chapter 9 (Figures 9.2 and 9.4).

The first six EOFs have been calculated for each of the studied matrices (`Free run`, `Forecast` and observations). The variance explained by each mode is shown in Figure 11.25. The two most important modes account for a large fraction of the total variance explained by the six first EOFs. Thus, only the two first modes will be kept for the study. The third and fourth modes variance is very similar, and their interpretation would add no essential information.

First EOF mode

Figures 11.26 to 11.28 show the most dominant spatial modes for `Free run`, `Forecast` and Pathfinder v5 respectively. The temporal evolution of the first EOF mode for the three data sets is shown in Figure 11.29.

The first temporal mode (Figure 11.29) presents an increasing trend, with negative values in July and positive values in August. This first temporal mode accounts thus for the monthly variance of the data. Model and observations present a very similar path for this temporal mode. The spatial structure of the first EOF mode (Figures 11.26 to 11.28) is very similar for the three data sets. General features are represented, as well as

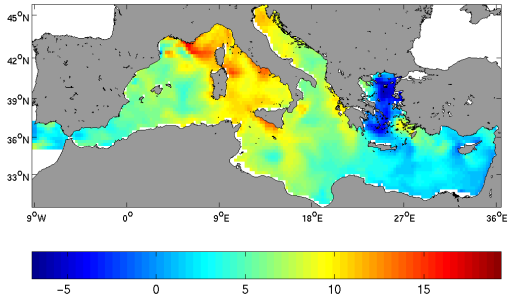

Figure 11.26: First spatial mode for the `Free run`

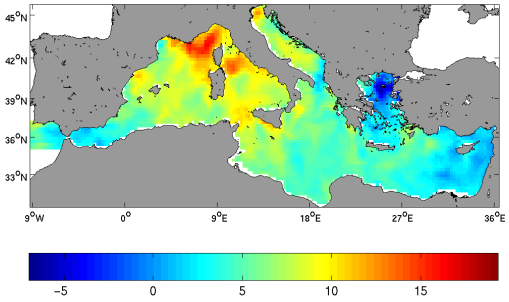

Figure 11.27: First spatial mode for the `Forecast`

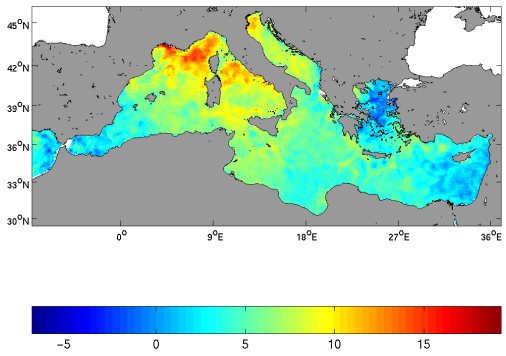

Figure 11.28: First spatial mode for the observations

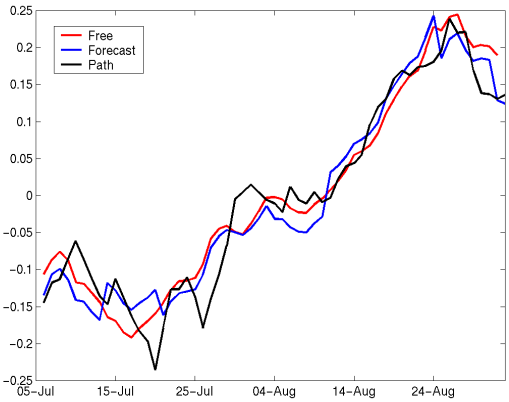

Figure 11.29: First temporal mode for `Free run`, `Forecast` and observations

smaller aspects of the Mediterranean SST distribution. In this first spatial mode, nearly the whole Mediterranean basin has positive values, except the Aegean Sea, near the Bosphore Strait, with negative values. The highest positive values are found in the Gulf of Lions, Ligurian Sea and in the Northern Adriatic. The Gulf of Lions and the Northern Adriatic are influenced by river runoff of the Rhone River and the Po River, respectively. Po River waters are warmer than the surrounding open sea waters in summer (*e.g.* Gacić *et al.*, 1997) so it is the Po River runoff influence that causes the warming of these waters. The rhone River water has been also found to be warmer than the surrounding waters in the Gulf of Lions, at least for the studied year (*e.g.* Younes *et al.*, 2003).

A West-East temperature gradient covering the entire Mediterranean can be seen in the first spatial mode, with decreasing positive values from west to east. Negative values are reached near the Bosphore Strait. Water coming from the Black Sea shows a gentle temperature decrease from July to August, in an opposite trend than the rest of the Mediterranean Sea.

In general, the first spatial mode of the model results and the observed data is very similar, with the intensity of the observed spatial EOF higher than the model. The temporal amplitudes are very similar for all the three data sets. The trend of the basin first spatial mode is a general warming, with a minimum at the beginning of July and a gently temperature decrease at the end of August. The amplitude minimum found in July is slightly shifted in the model respect to the observations. The RMS error between the model and the observations is 0.04°C for both `Free run` and `Forecast`. If the first temporal mode of the model is forwarded in time 2 days, the RMS error is minimised to 0.035°C, indicating a slight shift between model and observations. The maximum in August is reached at the same time by the model and the observations. The first EOF mode accounts nearly for the total variance, so we can say that almost all the SST variability of these data sets is due to the summer warming.

Second EOF mode

The second spatial mode for `Free run`, `Forecast` and observations is presented in Figures 11.30 to 11.32, and the respective temporal mode is shown in Figure 11.33.

The second spatial mode presents more dissimilarities between the model results and Pathfinder v5 data. The temporal mode, however, has a similar evolution for the three data sets. We find negative values in July, which progressively increase towards zero. The first two weeks of August have positive values. Then, the amplitudes decrease again and they evolve around zero until the end of August. Minimum and maximum values are slightly shifted in the model with regard to Pathfinder v5 observations, but their amplitude is similar.

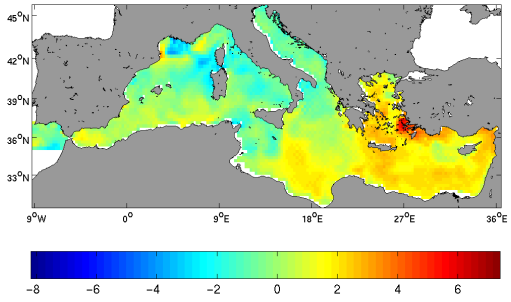

Figure 11.30: Second spatial mode for the `Free run`

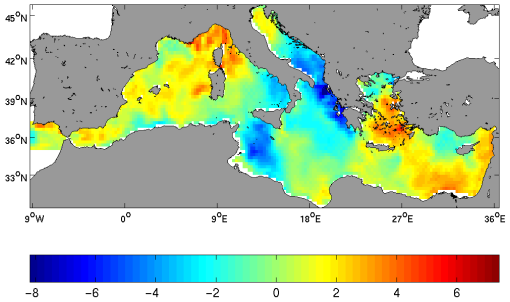

Figure 11.31: Second spatial mode for the `Forecast`

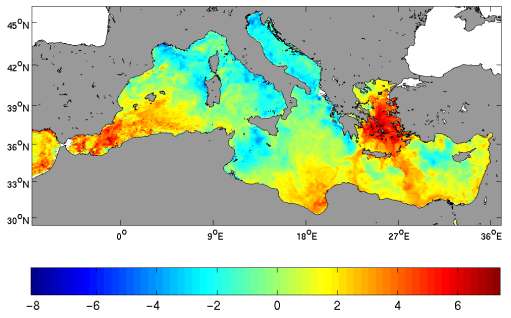

Figure 11.32: Second spatial mode for the observations

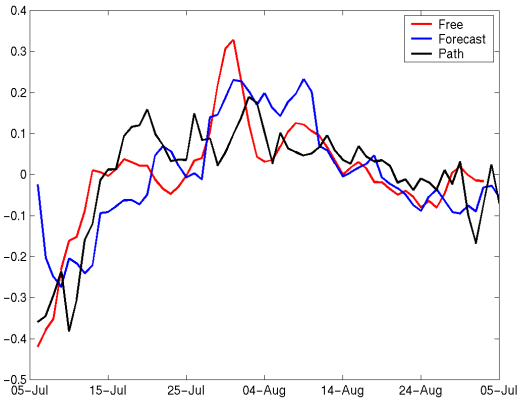

Figure 11.33: Second temporal mode for `Free run`, `Forecast` and observations

We can see three different zones in the second spatial mode for the `Forecast` and the observations. The Western Mediterranean has almost everywhere positive values, the Central Mediterranean has negative values and the Eastern Mediterranean has again positive values. The `Free run` has an east-west division, with positive values at the Eastern Mediterranean and very weak values, almost zero, in the Western Mediterranean.

The first and second spatial modes for the observations are clearly complementary: negative zones in the first spatial mode are positive in the second one, and vice versa. This behaviour is also noticed in the `Forecast`, althought less clearly, and in the `Free run` to a less extent. The second mode is thus a modulation of the first one, and this modulation is mainly driven by the second temporal mode. For example, the positive values of the Aegean Sea present in the first mode, which gently become negative from July to August, are first weakened by the negative values of the second mode. At the end of the record the negative values of the first mode are reinforced by the second mode. Similar explanations can be depicted out for other zones of the Mediterranean Sea.

There are some divergences between `Forecast` and observations second spatial mode. In the Western Mediterranean, the Atlantic water input has a bigger impact on observations than in the `Forecast`. In addition, the `Forecast` presents high values near the Ligurian Sea, a feature that is not present in the observations. In the Central Mediterranean, the `Forecast` has intense negative values, not present in the observations. Eastern Mediterranean spatial structure is very similar for both data sets.

11.3 Summary

In this chapter we have applied two spatio-temporal techniques to the model results in order to assess its skill by taking into account the high variability in space and time that characterises ocean domains.

First, a wavelet decomposition has been applied to the model and observations in the Ligurian Sea. The model and the observations have been decomposed into eight different scales, ranging from 1 to 128 km of resolution. At each of these scales, the traditional error measures that are used for skill assessment have been applied. This has allowed us to identify the model errors and to relate them to a specific physical scale.

Large scales have present a high variability, so that error measures based on the amplitude of the model and observations are influenced by large scales. An error measure that takes into account the variability inherent to each scale has been used. We can see that for this error measure, the *error relative to scale*, small scales present a higher error than large scales. In addition, the data assimilation leads to a decrease of error at large scales, but small scales do not benefit from the assimilation the same

way as large scales do. Another effect that has been pointed out is that the standard deviation of the model is smaller than for observations, mainly at small scales. The decrease in standard deviation can be caused by a too strong diffusion on the model. The representation of these error measures in a Taylor diagram has helped us to interpret them.

The use of wavelets on an irregular boundary domain is not straight forward. The presence of land points lead to perturbed wavelet coefficients, and thus the use of those coefficients is also affected by this boundary problem. In this work we have done a boundary treatment, where the sea values are extended towards land points. The transition between sea and land points is made progresively, instead of presenting a big step. There are other works assessing the problem of wavelets on irregular domains, and these should be the future guidelines to improve the application of wavelet analysis on oceanographic forecast verification.

The EOF analyses have shown that the main cycle in this data set is the summer warming. The two first modes represent the major part of the data sets variance. First and second modes are complementary, the second mode being a modulation to the first one. We have seen that there is a time delay of 2 days between the model and the observations in the temporal evolution of the first mode. This time shift can be due to a phasing between the observations and the forcings, because the modle is in advance respect to observations. If the cause is a late reaction of the model, it will present a shift forward in time.

The model and observations first and second modes present a good spatial and temporal coherence, with only a slightly shift in the first temporal mode. The `Forecast` shows better concordance with the observations than the `Free run`, mainly in the second spatial mode.

A spatio-temporal analysis of the data has allowed us to study the model results in a comprehensive way. After the qualitative and quantitative error assessment, some questions remained unanswered, as how accurately the spatial distribution of the model represents the observed one, or how the assimilation affects the model. The spatio-temporal analysis has also shown some specific problems related with the small scales, and we have seen that the assimilation effect is mainly felt at large scales.

Chapter 12

SOFT predictors results

The assimilation results of SOFT predicted SST fields are compared to the assimilation of observed SST data. Now the skill of the SOFT systems is analysed. These two sets are compared:

- Model with observed SST assimilation, only in the Ligurian Sea, called AssimSST in this document.
- Model with assimilation of SOFT predicted SST fields, in the Ligurian Sea, called AssimSOFT in this document.

The use of other model results will make the comparison with SOFT systems very difficult, as the assimilation of SST and SSH in the whole Mediterranean. We should thus use similar configurations in this section.

First, the classical verification measures are used, in order to establish the skill of the SOFT systems related to AssimSST. The observations used here to establish the skill are non-dependent, because they are assimilated to the model, so the results may be interpreted carefully. Then, the comparison with Sirena cruise data is done, where the vertical structure of the water column is compared to observations, and the position of the Mixed Layer Depth is finally studied. Finally, a multiresolution decomposition of the model runs and the observations is done using Wavelet transforms. The distribution of the error into different spatial scales is thus studied.

12.0.1 Classical Error Measures

The results of AssimSOFT are very good, compared to AssimSST. This last has a better skill, but the advantage of the SOFT systems lies in their early availability. In Figures **??** to 12.4 we can see the different errors of AssimSOFT and AssimSST.

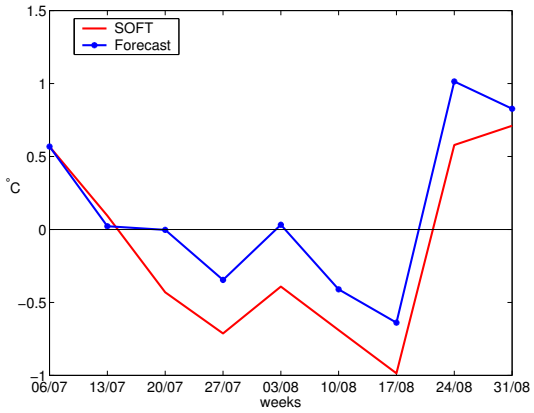

Figure 12.1: Bias

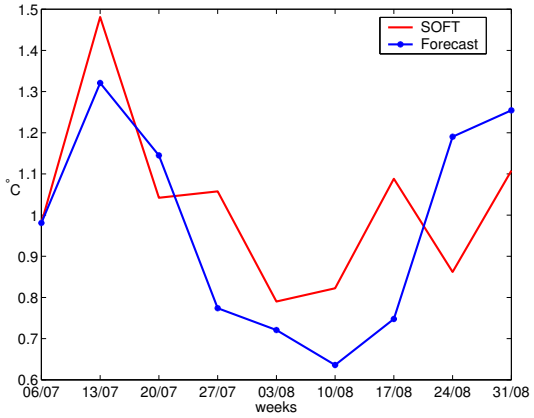

Figure 12.2: Root Mean Square Error

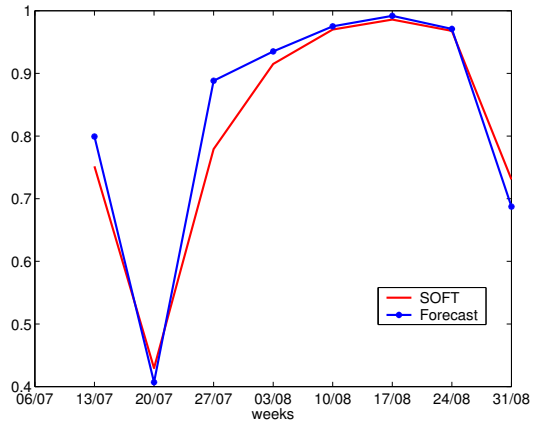

Figure 12.3: Anomaly Correlation Coefficient

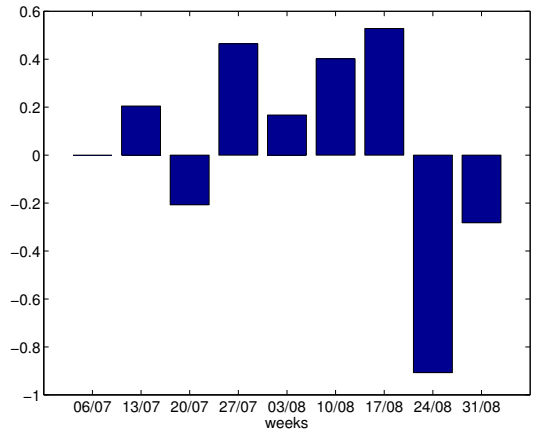

Figure 12.4: Mean Square Error Skill Score

In Figure **??** the Mean Absolute Error (MAE) is presented. AssimSOFT has a higher error overall, except for weeks 8 and 9 (24 and 31 August), where AssimSST presents a big error. This is observed generally for AssimSST, and is probably related to the atmospheric forcing. Another kind of error measure is however preferred when studying the behaviour of a model. The Root Mean Square (RMS) Error is more sensitive to large forecast errors than MAE is, due to the square in its expression.

The RMS error is presented in Figure 12.2. We can see that AssimSST is better that AssimSOFT, with again the exception of weeks 8 and 9 (24 and 31 August). As the forcing is the same for both runs, the only possibility is that the assimilated SST images are perturbing the model at those weeks. The SOFT predictors, although derived from the same SST images, do not contain any more the perturbing features. That is, the SOFT predictors, as they filter out the noise from the SST images, perform a better forecast than the direct SST assimilation. Both AssimSST and AssimSOFT present an important error of accuracy the second week. The second week presents a high RMS error for both AssimSST and SOFT.

In Figure 12.1 we can see the bias between both forecasts and the observations. It is noticeable that the bias of the second week is very little, almost zero, for both cases. The error of this week is then due only to the accuracy of the forecast, maybe due to the presence of outliers (points too warm and too hot, that cancel themselves mutually, so that the bias is zero). This will be studied more closely later. An abrupt change in bias noticed from week 7 to week 8 (17 and 24 August). AssimSST and AssimSOFT are colder and colder than the observations until week seven. At week eight and nine ((24 and 31 August), this tendency is inversed and they pass to be warmer than the observations, making a total change of nearly 2°C in two weeks. This change may be reflecting the source of the problem in weeks 8 and 9, although AssimSST and AssimSOFT should be acting differently, as we have seen in the other error measures.

The next Figure (Figure 12.3) presents the Anomaly Correlation Coefficient. Here the second week presents a good score, but surprisingly the third week seems to be very inaccurate. In addition, weeks 8 and 9 present a good ACC. The error in these weeks is mainly due to a magnitude misfit between the forecast and the observations, but the patterns present in the observations are good represented.

Finally, the Mean Square Error Skill Score is presented (Figure 12.4). In this Figure, the improvement of the SST assimilation over the SOFT predicted fields assimilation is shown. As depicted out from the previous error measures, AssimSST is better that AssimSOFT except for weeks 3, 8 and 9, where SOFT beats the SST assimilation.

In Table 12.4 the mean errors for the SST assimilation and the SOFT assimilation are summarised. It can be seen that the mean errors are very close for both experiments. MAE and bias are better in AssimSST, but RMS or ACC are basically equal for both

Table 12.1: Mean errors for observed SST data assimilation and SOFT assimilation

Measure	AssimSST	SOFT
MAE	$0.87°C^2$	$0.85°C^2$
RMSE	1.02°C	1.026°C
Bias	0.005 °C	-0.1396°C
ACC	0.7256	0.7256

experiments.

The problem presented at weeks 8 and 9 is due to a late warming of the sea surface temperature in both AssimSST and AssimSOFT, as can be seen in the evolution of the mean SST of the Ligurian Sea over time (see Figure 12.5). Both experiments accurately represent the minimum of temperature of week 2 (13 July), but the maximum observed at week 7 has moved to week 8 for AssimSST and AssimSOFT. The reason for this late warming may be in the atmospheric forcings, which present this shift. In Figure 12.6 we can see the mean air temperature 2m above the sea surface of the Ligurian Sea. The SST is also included, as well as the moments where assimilation of SST is done. The maximum temperature for the air temperature is reached at the eight assimilation cycle, whereas the SST has already begin to decrease. There is still one open question: if the shift in temperature is due to atmospheric forcing, so that both AssimSST and SOFT present a warming at week 8, why SOFT does not has a bigger RMS error at these time?. The bias is similar for AssimSST and SOFT, but recall from Figure 12.2 that the RMS for SOFT at week 8 is smaller than for AssimSST.

12.0.2 Comparison with Sirena Cruise data

The Sirena data are not assimilated into the model, so they can be used as independent data. This is useful for verification purposes, because the skill assessment of a model with non-independent data can be very ambiguous. The distribution of the Sirena Cruise in the Ligurian Sea was shown in Figure 6.1, and they correspond to the last 8 days of simulation, from 22 to 30 August 2000.

The RMS error between the forecast (observed SST and SOFT) and the Sirena observations at each of these stations is presented in Figures 12.7 and 12.8. The error is integrated over the whole profile, being 383 m the minimum depth and 476 m the maximum (see Figure 6.1. The mean error for the assimilation of observed SST is 1.14 °C, and for the assimilation of SOFT predictors is 0.9 °C. However, the first 100 meters in the ocean are much more variable than deep waters, so the error at depth is usually

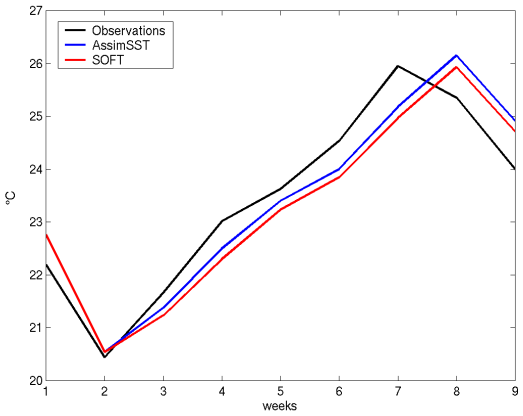

Figure 12.5: Mean SST for observations, AssimSST and AssimSOFT

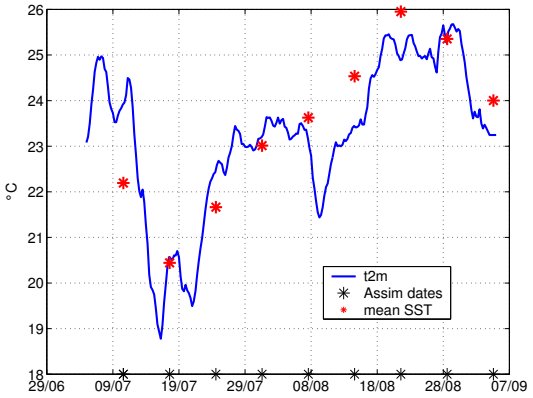

Figure 12.6: Air temperature evolution in the Ligurian Sea. Red asterisks show the temperature of the assimilated SST and black asterisks (at the time axis) show the moments where the assimilation is done.

very little. For this reason, in Figures 12.9 and 12.10 the error has been only computed in the first 100 m of the water column. The distribution is very similar for both assimilation experiments, being the smaller errors in the north of the Ligurian basin. The mean RMS error when only considering the first 100 m of the water column is 2.4°C for the assimilation of observed SST and 1.9 °C for the SOFT predictors assimilation.

In Figure 12.11 we can see the vertical distribution of the RMS error for both assimilation experiments, integrated over all Sirena Profiles. As can be seen, the maximum error is located around 20 m depth, the thermocline depth. The assimilation of observed SST presents a bigger error, nearly 1°C bigger than the SOFT predictor assimilation.

The Mixed Layer Depth (MLD) has been calculated for AssimSST, AssimSOFT and the Sirena profiles. The study of the MLD allows knowing if the error in the vertical is due to a general misfit between observations and forecast or if they are due to a shift in the vertical structure of the forecast profiles. Once the depth of the Mixed Layer is known, the error in depth can be calculated separately from the error in temperature at the MLD. In Figures 12.12 and 12.13 the mean fitted profile is compared to the mean real profile (mean over the 51 initial profiles). The fitted profile matches very well the real profile, except at the surface, where the real profile is warmer. However, as the parameter of interest is the MLD, this difference at the surface does not affect this study. In the same figures, in red, the residual of the fit, which is smaller than 1°C. In Tables 12.2 and 12.3 the RMS errors in depth and in temperature for AssimSST and AssimSOFT are presented. As expected, the misfit in depth is very large, being the model MLD results deeper than the observations. However, when looking at the temperature of the

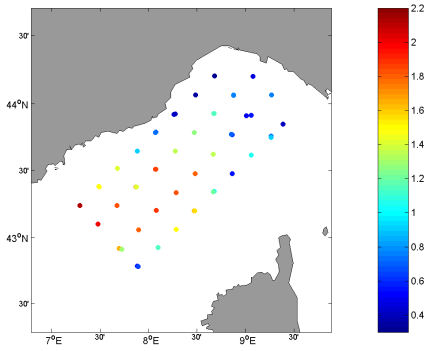

Figure 12.7: Distribution of SST assimilation RMS in the Ligurian Sea

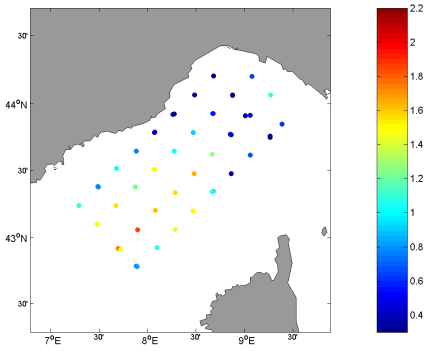

Figure 12.8: Distribution of SOFT RMS in the Ligurian Sea

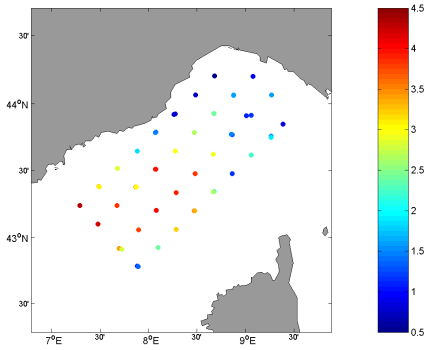

Figure 12.9: Distribution of SST assimilation RMS in the Ligurian Sea for the first 100m of the water column

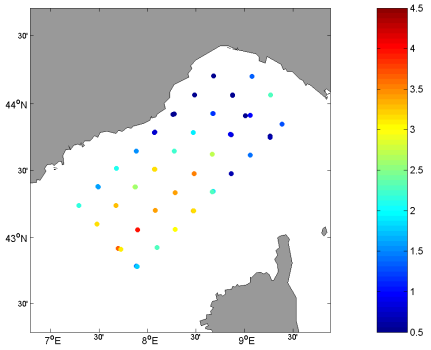

Figure 12.10: Distribution of SOFT RMS in the Ligurian Sea for the first 100m of the water column

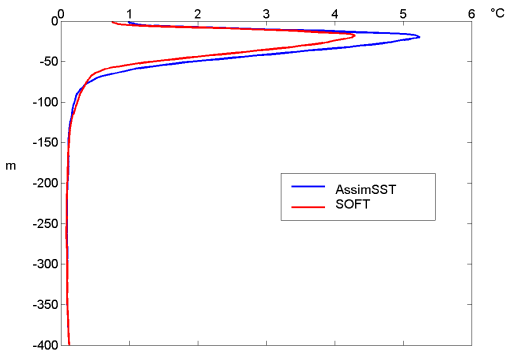

Figure 12.11: Vertical distribution of the RMS error

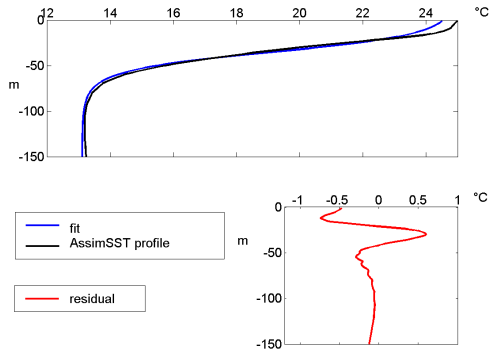

Figure 12.12: Mean profile and fit for AssimSST. In red, the residual of the fit.

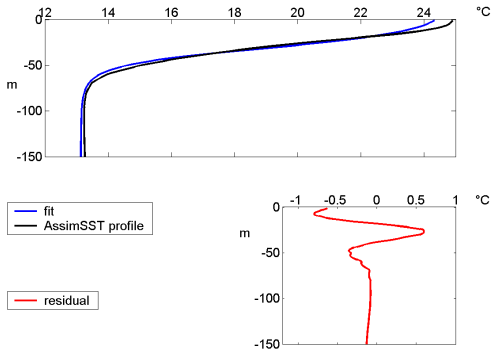

Figure 12.13: Mean profile and fit for AssimSOFT. In red, the residual of the fit.

Table 12.2: RMS error and bias in the depth representation of the Mixed Layer

	AssimSST	Assim SOFT
RMS	16.17 m	12.74 m
bias	-14.28 m	-9.39 m

Table 12.3: RMS error and bias in the temperature representation of the Mixed Layer

	AssimSST	Assim SOFT
RMS	0.76°C	0.787 °C
bias	-0.26°C	-0.22 °C

MLD, we can see that the error is much more little than what we obtained when looking directly to the profiles error (Figure 12.11).

Multiresolution Analysis

Two-dimensional Wavelet Transforms have been used to decompose the SST (of the model and the observations) in order to look at different spatial scales the differences and similarities between the forecast and the observations. The SST obtained by the assimilation of observed SST and the SST obtained from the assimilation of SOFT predictors are decomposed thus into eight scales, ranging from 2 to 256 km of spatial resolution. The decomposition depends on the initial resolution of the data, and in the size of the initial matrix. The figures included in this section present thus the results for these 8 spatial scales. However, the study and interpretation is centred in the 6 first scales, from 2 to 64 km, because the decreasing number of points obtained by the wavelet decomposition makes the interpretation of those big scales very difficult. The reason why all scales are presented is the conservation property of the Wavelet Transform. The integral of the variability over all scales plus the amplitude of the coarsest scale gives the total variability of the initial data (see the chapter on Wavelet Transforms for a more

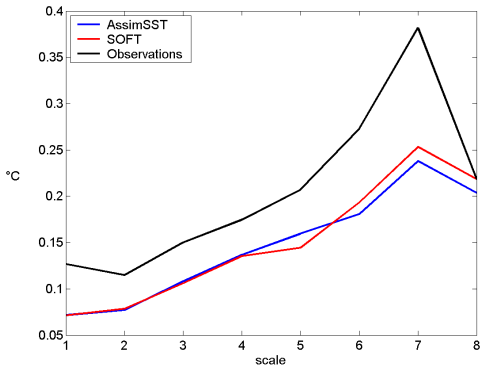

Figure 12.14: Distribution of standard deviation at different spatial scales

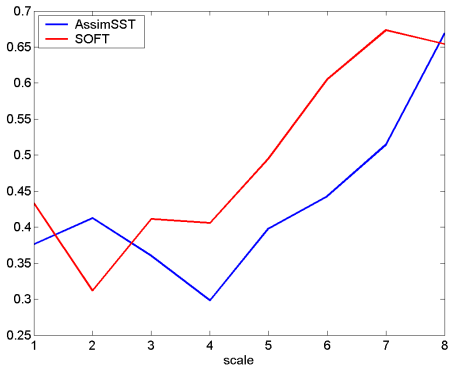

Figure 12.15: Energy ratio of the different spatial scales

detailed description). The choice of 64 km as the last scale to interpret is because the Northern Current, the most important current in the Ligurian Sea, has a mean width of about 30 km, so the scale 6 will contain certainly any change in the behaviour of this Northern Current.

In the following figures, the contribution of each spatial scale to the overall error is presented. First, the spatial mean error has been calculated for each scale to see the evolution in time and scale of the error. To avoid too many plots the error mean over time has been calculated, so that the information about each scale is retained.

In Figure 12.14 the standard deviation of the observations, AssimSST and AssimSOFT is presented. The variability of the observations is higher for all scales, but the assimilation experiments seem to follow the changes in increasing variability in the same amount of the observations. Figure 12.15 presents the energy ratio for AssimSST and AssimSOFT. The nearer to zero, the nearer the variability of the analysed field is to observations. AssimSST is closer to observations than AssimSOFT, in terms of variability. The mean energy value is 0.43 for AssimSST and 0.5 for AssimSOFT.

The RMS error is given in Figure 12.16. The distribution of the RMS is as expected after the standard deviation figure, increasing with the increasing resolution. AssimSOFT presents overall a smaller error than AssimSST. ACC, for example (Figure 12.17) shows that meso-scales (scales 4 to 6, 16 to 64 km resolution) are the more difficult to represent in terms of pattern match.

The MSE has the advantage that a percentage score can be calculated. In Figure 12.18 and in Table 12.4 the percentage each scale contributes to the total error is shown. For small scales, both AssimSST and SOFT contribute in a similar amount to the total error. At scale 5, differences begin to appear, that increase at bigger scales.

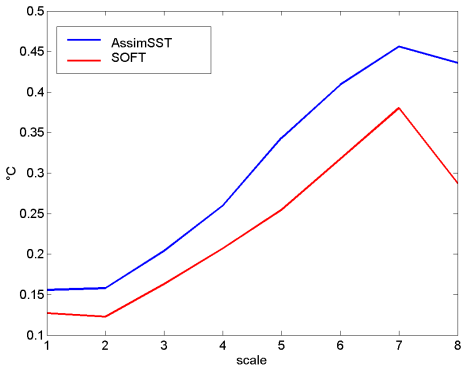

Figure 12.16: Distribution of RMS error at different spatial scales

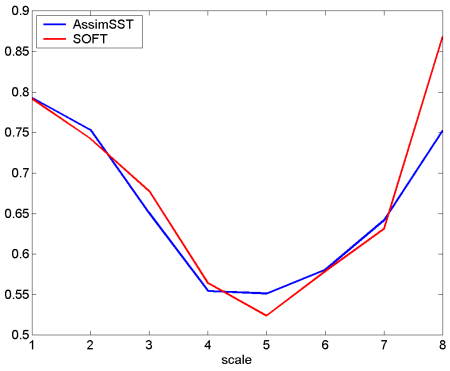

Figure 12.17: Anomaly Correlation Coefficient of the different spatial scales

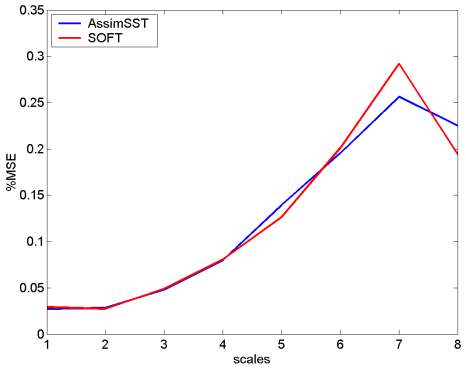

Figure 12.18: Percentage of contribution for each spatial scale to MSE

Table 12.4: Percentage of contribution to ACC and MSE for each of the spatial scales

scale		1	2	3	4	5	6	7	8
km		2	4	8	16	32	64	128	256
AssimSST	%MSE	0.0274	0.0284	0.0480	0.0798	0.1395	0.1954	0.2564	0.2252
AssimSOFT	%MSE	0.0293	0.0274	0.0491	0.0808	0.1263	0.2009	0.2920	0.1942

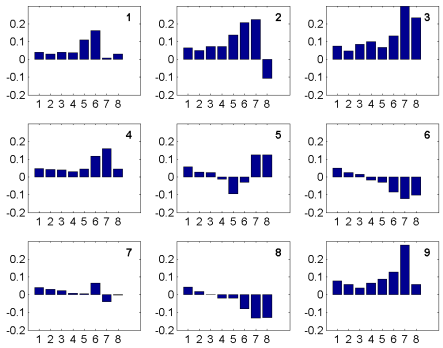

Figure 12.19: Time evolution of the difference in standard deviation between observations and AssimSST, for each scale.

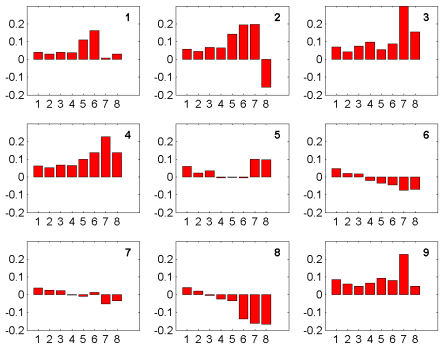

Figure 12.20: Time evolution of the difference in standard deviation between observations and SOFT, for each scale.

Making temporal means have the disadvantage of losing the information about what happens in weeks 8 and 9. For this, we have looked to the difference in standard deviation between observations and AssimSST (Figure 12.19) and between observations and SOFT (Figure 12.20), for each scale, evolving in time. The difference in standard deviation follows the scheme of Figure 12.14, with the variability of the observations bigger than the standard deviation of the assimilation experiments. This difference increases with the scale. However, for weeks 6, 7 and 8, the opposite trend is detected: the standard deviation is bigger for AssimSST is bigger that for the observations. The same behaviour is observed for SOFT. This difference is more evident at week 8, where the shift between air temperature and sea surface temperature has been detected.

12.0.3 Summary

The results of an experiment where observed SST is assimilated to a three-dimensional hydrodynamic model have been compared to the results of SOFT predicted SST fields assimilation. The overall skill shows that both approaches perform in a similar way, although the assimilation of SOFT predicted fields present slightly smaller errors.

The main problems present in AssimSST are also present in SOFT, as a shift of the sea surface temperature between model results and the observed SST on 24 August. This shift is clearly seen in the bias, where both AssimSST and SOFT behaves in a similar way, but the RMS error evolution is different for each experiment. The RMS accounts for errors in accuracy, and this makes the difference with bias. The ACC is good on 24 and 31 August.

The study of the Mixed Layer Depth (MLD) reveals that the model profiles are

'pushed' to the bottom if we compare them to Sirena profiles. As a result, the depth of the mixed layer is about 15 meters deeper than the observations, but the temperature at the MLD is well represented. This deepening effect is due to the effect of the assimilation of SST into the model. The difference between Sirena observations and the model has a spatial component, being the RMS error bigger in the southern part of the Ligurian basin.

Finally, a multiresolution decomposition has been realised to see the behaviour of the model at different spatial scales. The overall trend is that bigger scales have bigger errors. For the ACC, the meso-scale (about 30 km) presents the worst score. Again, when looking to the evolution in time, 24 August has a different behaviour than the rest of the assimilation cycles.

The skill of SOFT predicted fields results is very similar to the skill of observed SST assimilation. The advantage of using SOFT predictors lies in their availability, which allows the forecasts to assimilate information about the future state of the ocean.

Part IV

Summary and Conclusions

Chapter 13

Summary

The quality of a 3D hydrodynamic, primitive equation model of the Mediterranean Sea was evaluated. The resolution of this model is 1/4°. Two grid refinements have been made, to the Liguro-Provençal basin (with 1/20° resolution) and to the Ligurian Sea, where a resolution of 1/60° is reached. The coupling between these models is done by a two-way nesting approach.

An assimilation experiment with a duration of nine weeks has been realised. During this period, Sea Surface Temperature (SST) and Sea Level Anomaly (SLA) data have been assimilated via a Singular Evolutive Extended Kalman filter (SEEK). The assimilation has been done once a week, so a total of nine assimilation cycles have been done. SST from the DLR and SLA from TOPEX/Poseidon+ERS-2+GFO are assimilated into the model.

In this work, we have presented the quality assessment of this model. The aim of this work is to establish the impact of the assimilation on the model performances. The comparison of the model with observations has allowed us to establish the quality of the model results. In addition, the comparison of a free run of the model (*i.e.*, a model run without data assimilation) with observations has been used to assess the impact of the assimilation alone. An additional test was carried out, to quantify the accuracy of the model when SOFT statistical predictors of the Ligurian Sea SST are assimilated, instead of observed SST.

The verification of the mentioned models has been done in three different ways. (i) First, a qualitative assessment of the model skills has been made. This study has allowed us to have a first overview of the model performances. (ii) In a second part, a quantitative error study has been realised. In this part, the classical error measures has been applied to the model. (iii) Finally, the spatio-temporal characteristics of the model have been studied, in order to realise a complete study of the model behaviour, taking into account the complex spatial distribution of the model variables. In this chapter, the main results of the model verification are summarised.

Qualitative error assessment

The mean temperature, salinity and velocity fields of the entire Mediterranean Sea have been studied, at the surface and a depth of 100 m. The distribution of these variables has been compared to MEDAR/Medatlas climatology and to Pathfinder v5 for the SST. The main observed features are present in the model, and we have seen that the temperature is closer to observations and climatology than the salinity. This is due to the beneficial impact of the SST assimilation. The grid refinement has also a beneficial impact on the model results, as the Western Mediterranean has a more realistic distribution of these variables. There are also some discrepancies between model and observations, as the reversal of the Eastern Corsican Current or the deviation to the north of the Algerian Current. These features, however, have been occasionally observed in the Mediterranean Sea and are reproduced by other models.

TS diagrams of the Eastern and Western basins have been used to identify the presence or absence of water masses. There is a good agreement between the model and the climatology at the Western basin. TS diagrams are slightly different for the model and climatology in the Eastern basin. This is again due to the nesting.

Finally, the transports of the main currents present in the Ligurian Sea have been calculated. We have compared the Eastern Corsican Current, Western Corsican Current and Northern Current transports to values fomr the literature. The Ligurian Sea model has been chosen because it has the finest grid and represents more accurately the small variations of these currents. The mean transports are realistic, although the ECC has weaker values than the observed ones. This is due to the reversal in the surface current in the Strait of Corsica.

Quantitative Error Assessment

The quantitative error assessment has been done in different steps. First, the model results are directly compared to the assimilated observations, to establish the impact of assimilation on the model. Afterwards, non-assimilated observations have been used to assess the model skill. The use of assimilated observations leads to a wrong perception of the model quality, as they can artificially ameliorate its skill.

In average, the `Forecast` performs better than the `Free run`. However, due to a time delay between the assimilated SST and the atmospheric forcings, the `Forecast` presents a bad skill during two last weeks of the experiment, making the overall `Forecast` error only slightly better than the `Free run` one. This error is identified as an accuracy and bias error. The Anomaly Correlation Coefficient, which accounts for the similarity between the observed and `Forecast` anomaly patterns regardless of bias, is good for

these weeks. This highlights the fact that the mismatch during those weeks is due to a time delay between the atmospheric forcings and the assimilated SST that causes a high bias in the `Forecast`, but the model trend continues to follow the observations. The improvement of the `Forecast` SLA with respect to the `Free run` is smaller than the improvement experimented for the SST. The `Forecast` performs better than the `Free run`, but the improvement is very small.

The difference between the state after assimilation and the forecast of the next week is high, indicating that the `Forecast` drifts away from the observations during the week. The comparison of the model with Pathfinder v5 SST, although they are not independent (they are an AVHRR product, as the assimilated DLR SST) has allowed us to establish the evolution of the model SST between two assimilation cycles. We have seen that the influence of the assimilation lasts between two and four days, but the fifth or sixth day the error `Forecast` increases. This indicates that the observations should be assimilated more frequently to have a stronger control over the model error.

When the comparison between the model and *in situ* profiles has been done, we have observed an increase in the error mainly at the Mixed Layer Depth (MLD), for both `Free run` and `Forecast`. The MLD is a highly variable zone, and thus difficult to model. However, the `Free run` presents smaller error than the `Forecast` at the MLD. We have made a study about the positioning and temperature of the MLD, and the results have shown that the `Free run` and `Forecast` mixed layer is deeper than the observations, although the error in temperature of both runs is very small. The deep position of the mixed layer may be caused by a too high vertical mixing in the model, as both runs are affected by this problem. However, the `Forecast` mixed layer is deeper than the `Free run`. The assimilation of SST and SLA causes thus a change in the water column structure, making it deeper. This mixed layer deepening is caused by the assimilation of warm SST, which forces the model to warmer temperatures at the surface, and modifies the water column structure.

Spatio-temporal analysis

The final part of the verification process has consisted in a spatio-temporal analysis of the model results and observations. Instead of applying the error measures directly to the model to study its spatial distribution, we have first realised a wavelet decomposition of the model and observations. This decomposition has allowed us to work with eight different spatial scales, with different resolutions ranging from 1 to 128 km. Two-dimensional wavelets decompose a field into orthogonal components, so we can examine field closeness at different resolution levels.

The error assessment at each of those spatial scales has shown that small scales (from 1 to 8 km) have higher errors relatively to the inherent variability of the scale. However, the

correlation for these small scales is better than the correlation of large scales (from 16 to 64 km resolution). The standard deviation has also been studied, and it has shown a clear increasing trend from small to large scales. All these measures have been summarised into a single diagram, called the Taylor diagram, where we have plotted the standard deviations of the `Forecast` and `Free run` (normalised by the standard deviation of the observations), the Anomaly Correlation Coefficient and the RMS. The representation of these quantities on a single diagram has allowed us to see the impact of the assimilation on the model. The assimilation strongly corrects the patterns and amplitude of large scales (increase in the ACC and decrease of RMS), but it does not correct the ACC or the RMS at small scales. Large scales thus benefit from the assimilation more than small scales. Another feature that we have seen in the diagram is that the standard deviation of the model (with and without assimilation) is too low for small scales, compared to the observed standard deviation. This fact can be caused by a too high diffusion in the model.

Reconstruction of satellite images

For some of the studies realised in this work, we have used the Pathfinder v5 SST data set. This data set contains missing data, due to clouds or other reasons that make them fail the Pathfinder v5 quality tests. However, sometimes it is desirable to have complete fields (as in the qualitative error assessment) and even necessary (as for the EOF decomposition in the spatio-temporal analysis). In this work, we have applied an EOF-based reconstruction method called DINEOF that allows us to obtain complete fields in a fast and accurate way. A Lanczos method has been implemented for the computation of the EOFs used for the reconstruction. The Lanczos method allows us to optimise the computational time of the EOF decomposition.

This method has first been tested on a data set in the Adriatic Sea, where several validation tests have been applied to establish the quality of the results. First, we have compared the obtained reconstruction to *in situ* data from the MEDAR/Medatlas database. This test has led to accurate results, with an RMS error between the reconstruction and *in situ* data smaller than 1°C. We have also tested DINEOF with different amounts of missing data, artificially adding clouds to a subset with no missing data. In this way we have constructed three data sets having 40%, 60% and 80% of missing data respectively. These fields have been reconstructed with DINEOF and then compared to the original images.

DINEOF has been also compared to the reconstruction made by Optimal Interpolation (OI), a classical method for data reconstruction. The Harvard Ocean Descriptive Predictive System (HOPS) OI package has been used. The sub-optimal version is used in order to minimise the computational cost. The results obtained by OI have been directly compared to *in situ* data, giving rise to higher error than DINEOF. However, if the most dominant EOFs of the OI matrix are computed and then the amplitudes are fitted to the

initial data, we can have an improvement of the OI results with no important additional computational time. The fitted OI version lead to more accurate results (less than 0.9°C), very similar to the errors obtained by DINEOF. A clear advantage of DINEOF over OI is the computational cost, which is about 40.5 hours for OI and only 1.5 hours for DINEOF. In many situations, such as an operational frame, this time difference is very important.

After the reconstruction of the Adriatic data set, a set of Pathfinder v5 SST has also been also reconstructed. These images cover the whole Mediterranean during the two months of the assimilation experiment. Pathfinder v5 data are provided with a pixel-by pixel quality flag classification, allowing the user to choose the quality level of the data. In this work, all pixels with a quality flag smaller than 5 (from a total of 7) where rejected. The final data set had almost 40% of missing data. We have reconstructed these data with DINEOF. A comparison with MEDAR/Medatlas *in situ* data has also been made, and the total error between those data was in agreement with the errors obtained with the Adriatic data set. The reconstructed Pathfinder v5 data set has then been used in this work, for the quality error assessment and for the computation of EOFs in the spatio-temporal analysis. Such study could not have been made with an incomplete data set.

SOFT systems

An additional test has been carried out, where the assimilation of SOFT statistical predictors of the SST in the Ligurian Sea has been made. The results of this experiment have passed the same verification test than the real SST assimilation. SOFT predictors assimilation lead to robust results, where the evolution of the error is very similar to that obtained with the assimilation of the real SST. We have seen that the SOFT predictors gave even better results than the assimilation of real SST at the end of the experiment. This result is due to a problem in the assimilation of real SST and not really because the SOFT assimilation is better. However, this example shows that the assimilation of statistical predictors is robust and less influenced by noisy SST.

The main advantage of using SOFT statistical predictors in a real forecast frame is their early availability and their accuracy. As they are derived from recent SST data, they can take into account peculiarities of the season where the assimilation is done. The use of a climatology to force the model in a real forecast frame cannot take into account specific events that can modify the distribution of the SST.

The assimilation of SST and SLA gives robust results, apart from some specific problems at some moments of the experiment, as the time delay between the assimilated SST and the atmospheric forcings at the eight and ninth weeks. We have seen that the model with assimilation improves the results

obtained with the free run. However, some problems should be solved, as the low impact of the assimilation on the small scale features of the model, or the high vertical mixing imposed to the model that deepens the mixed layer.

Chapter 14

Conclusions and Perspectives

We can extract the following conclusions from this work:

Forecast Verification

In this work, we have realised the verification of different model configurations (model with and without assimilation, and model with assimilation of statistical predictors). From the beginning of this work we have carefully **planned the strategy to follow in the verification process**. The strategy is divided into three main parts that follow one another as a logical sequence of error assessment. Our knowledge about the model strengths and weaknesses has evolved as the error analysis has been made.

A verification plan should be homogeneously applied to all ocean systems, in order to facilitate their comparison. Among others, the definition of some standard metrics, and the application of spatio-temporal techniques for the study of a model evolution in time and space are necessary for the complete verification of a model.

Mixed Layer Depth

The mixed layer is the most variable part of the water column. A high percentage of a model's error is often located in this surface layer, principally due to a bad representation of its width or its depth. If the error in the positioning of the mixed layer depth and the temperature error are considered separately, we can determine whether one or the other reason causes the error. **A good detection of the mixed layer depth is thus desirable**.

There are some methods that deal with the detection of the mixed layer depth (*e.g.* Thomson and Fine, 2003). In this work, we have applied a fit of the vertical profiles to a sigmoid curve. The detection of the mixed layer depth is easily determined with this method, and **the error can be decomposed into position and temperature**

error. In all the cases studied in this work, a major part of the error was caused by a bad positioning of the mixed layer depth.

Wavelet Analysis

It is very difficult to study the skill distribution in space, because in field forecast verification the correlation between adjacent points should be taken into account to interpret the error measures. In addition, the use of classical error measures does not take into account the complexity of the compared systems, and often reduces the evolution in space and time of two complex fields -model and observations- into one single number.

The use of wavelets has allowed us to study the model at different spatial scales with increasing resolution, leading to interesting results. Wavelets are a powerful tool for analysing purposes. Work is still ongoing in the frame of wavelets and field forecast verification. It should be desirable to **design an error measure that takes into account the wavelet properties** to optimise the interpretation of the decomposed fields. In this work, we have used several classical error measures, and an error measure that takes into account the intrinsic variability of each spatial scale. This last measure has been very useful, as it has brought up some interesting characteristics of the model small scales. The representation of these measures into a single diagram (Taylor, 2001) has also helped in the interpretation of the wavelet decomposition.

Wavelets have a major problem. They are initially constructed to deal with regular domains. In this work we have used irregular boundary domains, and **a boundary treatment has become necessary to avoid perturbations** of the wavelet coefficients. We have used a very simple boundary treatment, which consists in an extrapolation of sea values towards land points, in a decreasing trend to zero. These data have been helpful to the application of the wavelet decomposition. There are other ways to avoid problems with irregular boundaries, as the second generation wavelets developed by Sweldens (1997) or the polynomial wavelet regression by CITE hee-seok. The application of both approaches to oceanic domains should be further investigated.

Reconstruction of satellite images

In this work we have applied an EOF-based method for the reconstruction of oceanographic data sets. The method, called DINEOF, has been tested on two data sets in the Mediterranean Sea and in the Adriatic Sea. The results have been validated with *in situ* data. The first data set has been successfully used in other parts of this work, as in the spatio-temporal analysis.

DINEOF has two major advantages over other classical methods used for data reconstruction. First, there is **no need for** *a priori* **knowledge about the statistics**

of the data, as the correlation length. Second, the method is much **faster** than optimal interpolation schemes, and the **accuracy** of the results is very similar.

The use of DINEOF for the reconstruction of missing data can be of a great interest, especially when the time is a constraint, as in operational oceanography.

Appendix A

Quality flags in Pathfinder SST images.

The Pathfinder v5 data are provided into three different fields: the SST data itself, the quality flags and the number of observations. Quality flags are classified from 0 (the worst) to 7 (the best), but there is no linear relation between those levels (that is, a quality flag of 4 does not mean that these data are twice as good as the data classified with 2). The quality flags are given pixel-by-pixel, and they are the result of a series of 'pass' or 'fails' of different quality tests. These tests are specified below:

1. Brightness temperature: for AVHRR channels 3, 4 and 5, the brightness temperature must be between -10°C and 35°C. This test aims to identify cold pixels related to high cloud tops.

2. Cloud test: to flag a pixel as a cloud, a series of validation tests with *in situ* data are applied. These tets vary with the year and the satellite.

3. Uniformity test 1: the brightness temperature of channels 3 and 4 is again examined. Pixels are took 3 by 3, and the difference between maximum and minimum temperature in this set must be smaller than 0.7°C. This test is based in the assumption that the brightness temperature should be homogeneous at small scales under cloud free conditions. Contamination by small clouds may affect the brightness temperature gradient.

4. Uniformity test 2: this test is the same as the precedent, but with a smaller constraint in the temperature gradient: it must be of 1.2°C maximum.

5. Zenith angle test 1: satellite angle must be smaller than 45°. At higher zenith angles, the radiation emitted by the ocean must go through a longer atmospheric distance before reaching the satellite, and will be therefore attenuated. The limitation of the zenith angle decreases this effect, but some geographical coverage is lost.

6. Zenith angle 2: the limitation of the zenith angle is limited to 55°. With this test there is thus a smaller loss of geographical coverage.

7. Reference test: pixels are compared to a reference field, computed from a three-weeks Reynolds climatology (Reynolds and Smith, 1994). The difference between them must be smaller than 2°C. Coastal pixels, or pixels in large SST gradients may be biased with this test, due to the fact that the Reynolds climatology has a limited coverage in time and space.

8. Stray sunlight test: stray sunlight contamination of pixels occurs primarily on the sun side of the scan line during the austral summer at high southern latitudes viewed under large zenith angles during nominal nighttime. This test identifies these pixels (but there are no applicable to our case)

9. Edge test: pixels on the first or last scan lines, and on the first or last rows on a scan line do not pass this test.

10. Glint test: glint index must be smaller than 0.05 sr^{-1}. Glint indexes are calculated using Cox and Munk (1954) formulation.

A pixel is automatically set to a 0 quality if it does not pass the brightness temperature test, the uniformity test, the zenith angle test 2 and the stray sunlight test. The remaining pixels are classified into the seven quality levels following various combinations of the mentioned tests. This procedure is summarised in Figure A.1. A complete explanation about the Pathfinder products can be found in Kilpatrick *et al.* (2001).

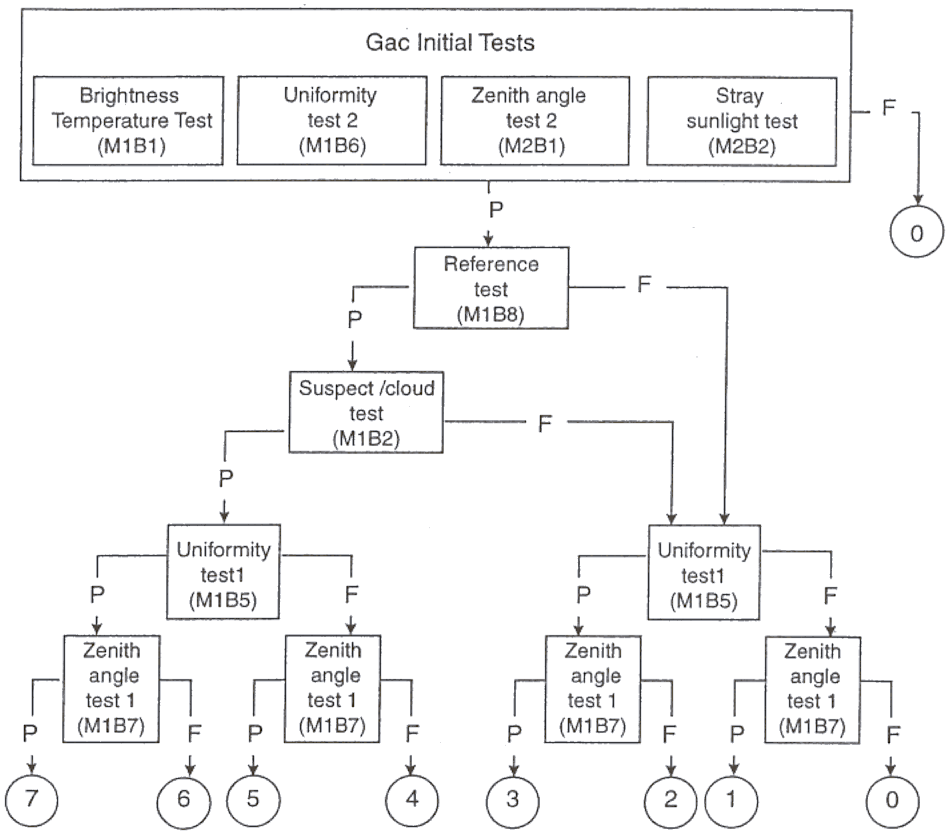

Figure A.1: Tree diagram for pixel-by-pixel quality flag decision. Obtained from (Kilpatrick *et al.*, 2001).

Appendix B

Multiresolution Analysis Properties

Multiscale analysis, as conceived by Mallat (1989) can be explained as follows: A multiresolution analysis consists of a sequence of successive approximation spaces V_j. These closed subspaces satisfy the following properties:

Property 1: A space corresponding to some resolution contains all the information about the space at lower resolution:

$$V_m \subset V_{m+1} ... \; \forall m \in \mathbb{Z} \tag{B.1}$$

Property 2: as the resolution increases the approximated function converges to the original function, and as the resolution decreases the approximated function contain less and less information:

$$\overline{\bigcup_{m=-\infty}^{\infty} V_m} = L^2(\mathbb{R}) \tag{B.2}$$

$$\bigcap_{m=-\infty}^{\infty} V_m = 0 \tag{B.3}$$

Property 3: all spaces are scaled versions of one space:

$$f(x) \in V_m \text{ if } f(2x) \in V_{m+1} \forall m \in \mathbb{Z} \tag{B.4}$$

Property 4: the space is invariant with respect to "integer translations" of a function:

$$f(x) \in V_m \Leftrightarrow f(x - k/2^m) \in V_m \; \forall k \in \mathbb{Z} \tag{B.5}$$

Property 5: translation (by integers) of the approximation of the function is equal to the approximation of the translation of the function (by integers).

The basic principle of multiresolution analysis is that whenever a collection of closed subspaces satisfies Properties 1 to 5, then there exists an orthonormal wavelet basis $\{\psi_{m,n}; m, n \in \mathbb{Z}\}$ of $L^2(\mathbb{R})$, $\psi_{m,n}(x) = 2^{-m/2}\psi(2^{-m}x - n)$, such that, for all f in $L^2(\mathbb{R})$:

$$P_m f = P_{m+1} f + \sum_{k \in \mathbb{Z}} (f, \psi_{m,n})\, \psi_{m,n} \tag{B.6}$$

P_m denotes the orthogonal projection operator onto V_m. For every $m \in \mathbb{Z}$, O_m is the orthogonal complement of V_m in V_{m+1}:

$$V_{m+1} = V_m \bigoplus O_m \tag{B.7}$$

The set of closed subspaces characterize the behaviour of a function at different resolutions. For example, V_m describes functions at 2^m samples per unit length. The representation of functions in these subspaces V_m are obtained through an orthogonal projection by constructing an orthonormal basis for these subspaces. The orthogonal projection of a function onto the subspace V_m corresponds to its approximation at resolution m, *i.e.*, approximation using 2^m sampling points per unit length. Therefore, by successively traversing through the projections of $f(x)$ on the spaces V_m we obtain multiscale representation of the function $f(x)$.

Given the nested structure of V_m (Property 1), it is possible to construct a function $\phi(x)$ in V_0 such that $\{\phi(x-n)\}_{n \in \mathbb{Z}}$ is an orthonormal basis of V_0. The function $\phi(x)$ is called a scale function and satisfies $\int \phi(x)dt = 1$. Given O_0, the orthogonal complement of V_0 in V_1, it is possible to find a function $\psi(x)$, based on $\phi(x)$, such that $\{\psi(x-n)\}_{n \in \mathbb{Z}}$ is an orthonormal basis of O_0. The function $\phi(x)$ is orthogonal to its integer translates, and $\psi(x)$ is orthogonal to its integer translates and dilates, *i.e.* $\phi(x) \perp \phi(x-n)\ \forall n$ and $\psi(x) \perp \psi(2^m x - n)\ \forall m$. The function $\phi(x)$ is such that if its integer translates constitute an orthogonal basis of V_m then the integer translates of $\phi(2x)$ form an orthogonal basis of V_{m+1}. Using the recursive definition of Equation B.7 along with Property 1 and the orthogonality of $\psi(x)$ with its integer translates and dyadic dilates, it can be shown that the dilates on the dyadic sequence and translates on integers of $\psi(x)$, *i.e.* $\left\{2^{m/2}\psi(2^m x - n)\right\}_{(m,n) \in \mathbb{Z}^2}$, form an orthonormal basis of $L^2(\mathbb{R})$, or equivalently, $L^2(\mathbb{R}) = \bigoplus_{m \in \mathbb{Z}} O_m$. The function $\psi(x)$ is called an orthonormal wavelet.

List of Figures

List of Tables

Part V

References

Bibliography

Albérola, C. and C. Millot, 2003: Circulation in the French mediterranean coastal zone near Marseilles: the influence of wind and the Northern Current. *Continental Shelf Research*, **23**, 587–610.

Albérola, C., C. Millot, and J. Font, 1995: On the seasonal and mesoscale variabilities of the Northern Current during the 'PRIMO-0 experiment in the western Mediterranean Sea. *Oceanologica Acta*, **18**, 163–192.

Alvarez, A., 2003: Performance of Satellite-Based Ocean Forecasting (SOFT) Systems: A study in the Adriatic Sea. *Journal of Atmospheric and Oceanic Technology*, **20,5**, 717–729.

Álvarez, A., C. Lopez, M. Riera, E. Hernandez-Garcia, and J. Tintoré, 2000: Forecasting the SST space-time variability of the Alboran Sea with genetic algorithms. *Geophysical Research Letters*, **27**, 2709–2712.

Alvarez, A., A. Orfila, and J. Sellshop, 2003: Satellite based forecasting of sea surface temperature in the Tuscan Archipelago. *International Journal of Remote Sensing*, **24**, 2237–2251.

Alvarez, A., A. Orfila, and J. Tintoré, 2004: Real-time forecasting at weekly timescales of the SST and SLA of the Ligurian Sea with a satellite-based ocean forecasting (SOFT) system. *Journal of Geophysical Research*, **109**.

Alvarez, A. and J. Tintoré, 2001: DARWIN-An evolutionary program for nonlinear modeling of chaotic time series. *Computer Physics Communications*, **136**, 334–349.

Alvera-Azcárate, A., A. Barth, Z. Ben Bouallègue, L. Vandenbulcke, M. Rixen, and J.-M. Beckers, 2004a: Forecast Verification of a 3D model of the Ligurian Sea. the use of Discrete Wavelet Transforms and EOFs in the Skill Assessment of Spatial Forecasts. *Journal of Marine Systems.*, submitted.

Alvera-Azcárate, A., A. Barth, M. Rixen, and J.-M. Beckers, 2004b: Reconstruction of incomplete oceanographic data sets using Empirical Orthogonal Functions. Application to the Adriatic Sea. *Ocean Modelling.*, accepted.

Artegiani, A., D. Bregant, E. Paschini, N. Pinardi, F. Raicich, and A. Russo, 1997a: The Adriatic Sea general circulation. Part I: Air-Sea interactions and water mass structure. *Journal of Physical Oceanography*, **27**, 1492–1514.

— 1997b: The Adriatic Sea general circulation. Part II: Baroclinic circulation structure. *Journal of Physical Oceanography*, **27**, 1515–1532.

Astraldi, M., S. Balopoulos, J. Candela, J. Font, M. Gacić, G. P. Gasparini, B. Manca, A. Theocharis, and J. Tintoré, 1999: The role of straits and channels in understanding the characteristics of Mediterranean circulation. *Progress in Oceanography*, **44**, 65–108.

Astraldi, M. and G. P. Gasparini, 1992: The seasonal characteristics of the circulation in the north Mediterranean basin and their relationship with the atmospheric-climatic conditions. *Journal of Geophysical Research*, **97**, 9531–9540.

Astraldi, M., G. P. Gasparini, A. Vetrano, and S. Vignudelli, 2002: Hydrographic characteristics and interannual variability of water masses in the central Mediterranean:a sensitivity test for long-term changes in the Mediterranean Sea. *Deep-Sea Research I*, **49**, 661–680.

Ayoub, N., P. Y. Le Traon, and P. De Mey, 1998: A description of the Mediterranean surface variable circulation from combined ERS-1 and TOPEX/POSEIDON altimetric data. *Journal of Marine Systems*, **18**, 3–40.

Bahurel, P., E. Dombrowsky, J.-M. Lellouche, and the Mercator Project Team: 2004, Mercator Ocean Monitoring and Forecasting System, near-real-time assimilation of satellite and in situ data in different operational ocean models. *Marine Environmental Monitoring and Predictions*, 37th Liège Colloquium, Liège, oral presentation.

Bahurel, P., P. D. Mey, T. D. Prada, E. Dombrowsky, P. Josse, C. L. Provost, P.-Y. L. Traon, A. Piacentini, and L. Siefridt: 2001, *AVISO Altimetry Newsletter*, CNES, Toulouse, France, chapter MERCATOR, forecasting global ocean. 8, 14–16.

Barth, A., A. Alvera-Azcárate, J.-M. Beckers, M. Rixen, and L. Vandenbulcke, 2004: Multigrid state vector for data assimilation in a two-way nested model of the Ligurian Sea. *Journal of Marine Systems*, submitted.

Barth, A., A. Alvera-Azcárate, M. Rixen, and J.-M. Beckers, 2003: Two-way nested model of mesoscale circulation features in the Ligurian Sea. *Progress in Oceanography*, in press.

Beckers, J.-M., 1991: Application of a 3D model to the western mediterranean. *Journal of Marine Sciences*, **1**, 315–332.

Beckers, J.-M., P. Brasseur, and J. C. J. Nihoul, 1997: Circulation of the western Mediterranean: from global to regional scales. *Deep-Sea Research II*, **44**, 531–549.

Beckers, J.-M. and M. Rixen, 2003: EOF calculations and data filling from incomplete oceanographic data sets. *Journal of Atmospheric and Oceanic Technology*, **20**, 1839–1856.

Beckers, J.-M., M. Rixen, P. Brasseur, J.-M. Brankart, A. Elmoussaoui, M. Crépon, C. Herbaut, F. Martel, F. Van den Berghe, L. Mortier, A. Lascaratos, P. Drakopoulos, P. Korres, N. Pinardi, E. Masetti, S. Castellari, P. Carini, J. Tintore, A. Alvarez, S. Monserrat, D. Parrilla, R. Vautard, and S. Speich, 2002: Model intercomparison in the Mediterranean. The MedMEx simulations of the seasonal cycle. *Journal of Marine Systems*, **33–34**, 215–251.

Bennani, M. and T. Braconnier, 1994: Stopping criteria for eigensolvers. *Technical Report TR/PA/94/22, CERFACS, Toulouse, France*, 1–20.

Bennett, A. F., 2002: *Inverse modeling of the ocean and the atmosphere*. Cambridge, 234 pp.

Benzohra, M. and C. Millot, 1995: Characteristics and circulation of the surface and intermediate water masses off Algeria. *Deep-Sea Research I*, **42,(10)**, 1803–1830.

Bergamasco, A. and M. Gacić, 1996: Baroclinic response of the Adriatic Sea to an episode of Bora wind. *Journal of Physical Oceanography*, **26**, 1354–1369.

Bergamasco, A., T. Oguz, and P. Malanotte-Rizzoli, 1999: Modelling dense water mass formation and winter circulation in the northern and central Adriatic Sea. *Journal of Marine Systems*, **20**, 279–300.

Bethoux, J. P. and B. Gentili, 1999: Functioning of the mediterranean Sea: past and present changes related to freshwater input and climate changes. *Journal of Marine Systems*, **20**, 33–47.

Bleck, R., 2002: An Oceanic Circulation Model Framed in Hybrid Isopycnic-Cartesian Coordinates. *Ocean Modelling*, **1**, 55–88.

Borzelli, G., G. Manzella, S. Marullo, and R. Santoleri, 1999: Observations of coastal filaments in the Adriatic Sea. *Journal of Marine Systems*, **20**, 187–203.

Bouzinac, C., J. Font, and Johannessen, 2003: Annual cycles of sea level and sea surface temperature in the western Mediterranean Sea. *Journal of Geophysical Research*, **108**.

Bouzinac, C., J. Font, and C. Millot, 1999: Hydrology and currents observed in the channel of Sardinia during the PRIMO-1 experiment from November 1993 to October 1994. *Journal of Marine Systems*, **20**, 333–355.

Brankart, J. M. and P. Brasseur, 1996: Optimal analysis of in situ data in the Western Mediterranean using statistics and cross validation. *Journal of Atmospheric and Oceanic Technology*, **16**, 477–491.

— 1998: The general circulation in the Mediterranean Sea: a climatological approach. *Journal of Marine Systems*, **18**, 41–70.

Brankart, J. M., C. E. Testut, P. Brasseur, and J. Verron, 2003: Implementation of a multivariate data assimilation scheme for isopycnic coordinate ocean modes: application to a 1993-96 hindcast of the North Atlantic Ocean circulation. *Journal of Geophysical Research*.

Brasseur, P., J. Ballabera, and J. Verron, 1999: Assimilation of altimetric observations in a primitive equationmodel of the Gulf stream using a Singular Evolutive Extended Kalman filter. *Journal of Marine Systems*, **22**, 269–294.

Brasseur, P., J.-M. Beckers, J. M. Brankart, and R. Schoenauen, 1996: Seasonal temperature and salinity fields in the Mediterranean Sea: Climatological analyses of a historical data set. *Deep-Sea Research I*, **43**, 159–192.

Briggs, W. M. and R. A. Levine, 1997: Wavelets and field forecast verification. *Monthly Weather Review*, **125**, 1329–1341.

Brooks, H. E. and C. A. Doswell, 1996: A comparison of measures-oriented and distributions-oriented approaches to forecast verification. *Weather and Forecasting*, **10**, 288–303.

Bryden, H. L. and T. K. Kinder, 1991: Recent progresses in strait dynamics. *Reviews of Geophysics (Suppl)*, 617–631.

Buizza, R., 2001: Accuracy and potencial economic value of categorical and probabilistic forecasts of discrete events. *Monthly Weather Review*, **129**, 2329–2345.

Buongiorno Nardelli, B. and E. Salusti, 2000: On dense water formation criteria and their application to the Mediterranean Sea. *Deep-Sea Research I*, **47**, 193–221.

Carter, E. F. and A. R. Robinson, 1987: Analysis Models for the Estimation of Oceanic Fields. *Journal of Atmospheric and Oceanic Technology*, **4**, 49–74.

Casati, B., G. Ross, and D. B. Stephenson, 2004: A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorological Application*, accepted.

Castellari, S., N. Pinardi, and K. Leaman, 1998: A model study of air-sea interactions in the Mediterranean Sea. *Journal of Marine Systems*, **18**, 89–114.

Chapa, S. R., B. V. Rao, and G. S. S. Durga-Prasad, 1998: Application of Wavelet Transform to Meteosat-Derived Cold Cloud Index Data over South America. *Monthly Weather Review*, **126**, 2466–2481.

Chatelin, F., 1993: *Eigenvalues of matrices*. Wiley, 382 pp.

Chu, P. C., T. Hsing-Chia, C. P. Chang, and J. M. Chen, 1997: South China warm pool detected in spring from the Navy's Master Oceanographic Observational Data Set (MOODS). *Journal of Geophysical Research*, **102, C7**, 15761–15771.

Cook, K., 1999: Model skill assessment resources. *Western Region Technical Attachment*, **99-15**.

Cox, C. and W. Munk, 1954: Measurements of the roughness of the sea surface from photographs of the Sun's glitter. *Journal of Optics Soc. Am.*, **44**, 838–850.

Crosnier, L., C. Le Provost, P. Bahurel, M. Bell, L. Bertino, F. Blanc, E. Chassignet, G. Coppini, D. E., F. C., H. Hulburt, P.-Y. Le Traon, K. A. Lisaeter, C. L. Liu, and N. Pinardi: 2004, An ocean forecast systems intercomparison within the MERSEA project : Methodology and Results. *Geophysical Research Abstracts*, volume 6, reference number 03276.

Cushman-Roisin, B., M. Gacić, P. M. Poulain, and A. Artegiani, 2001: *Physical Oceanography of the Adriatic Sea*. Kluwer Academic Publishers, 304 pp.

da Silva, J. C. B., S. A. Ermakov, I. S. Robinson, D. R. G. Jeans, and S. V. Kijashko, 1998: Role of surface films in ERS SAR signatures of internal waves on the shelf. 1. Short-period internal waves. *Journal of Geophysical Research*, **103, C4**, 8009–8031.

Daubechies, I., 1992: *Ten Lectures on Wavelets*. SIAM, 357 pp., 357 pp.

Daubechies, I., I. Guskov, P. Schröder, and W. Swelden, 1999: Wavelets on irregular point sets. *Philosophical transactions of the Royal Society of London. A*, **357**, 2397–2413.

Daubechies, I. and W. Sweldens, 1998: Factoring wavelet transforms into lifting steps. *Journal of Fourier Analysis and Applications*, **4**, 245–267.

Davis, R. E., 1985: Objective Mapping by Least Squares Fitting. *Journal of Geophysical Research*, **90 (C3)**, 4773–4777.

de Mey, P. and A. R. Robinson, 1987: Assimilation of altimeter eddy fields in a limited-area quasi-geostrophic model. *Journal of Physical Oceanography*, **17**, 2280–2293.

Denis, B., R. Laprise, and D. Caya, 2003: Sensitivity of a regional climate model to the resolution of the lateral boundary conditions. *Climate Dynamics*, **20**, 107–126.

Dousset, B., J. Firing, P. Flament, H. Jackson, C. Lumpkin, C. Maroni, E. Nacini, P. M. Poulain, S. Pouliquen, M. Sawyer, and D. Young, 1998: Adriatic sea surface temperature images from the NOAA-AVHRR, 1995. *Département d'Océanographie Spatiale IFREMER, Brest*, **CD-ROM**.

Ebert, E. E. and J. L. McBride, 2000: Verification of precipitation in weather systems: determination of systematic errors. *Journal of Hydrology*, **239**, 179–202.

Emery, W. J. and R. E. Thomson, 1998: *Data Analysis Methods in Physical Oceanography*. Pergamon, 634 pp.

Everson, R., P. Cornillon, L. Sirovich, and A. Webber, 1997: An empirical eigenfunction analysis of Sea Surface Temperatures in the western North Atlantic. *Journal of Physical Oceanography*, **27**, 468–479.

Farge, M., G. Pellegrino, and K. Schneider, 2001: Coherent Vortex Extraction in 3D Turbulent Flows Using Orthogonal Wavelets. *Physical Review Letters*, **87**, 054501–1–054501–4.

Farge, M., K. Schneider, and N. Kevlahan, 1999: Non-gaussianity and coherent vortex simulation for two-dimensional turbulence using an adaptive orthogonal wavelet basis. *Physics of Fluids*, **11**, 2187–2201.

Fieguth, P., D. Menemenlis, T. Ho, A. Willsky, and C. Wunsch, 1998: Mapping mediterranean altimeter data with a multiresolution optimal interpolation algorithm. *Journal of Atmospheric and Oceanic Technology*, **15**, 535–546.

Font, J., E. Garcia-Ladona, and E. Gorriz, 1995: The seasonality of mesoscale motion in the Northern Current of the western Mediterranean: several years of evidence. *Oceanologica Acta*, **18**, 207–219.

Fournier, A., 2000: Introduction to Orthonormal Wavelet Analysis with Shift Invariance: Application to Observed Atmospheric Blocking Spatial Structure. *Journal of the Atmospheric Sciences*, **57**, 3856–3880.

— 2002: Atmospheric Energetics in the Wavelet Domain. Part I: Governing Equations and Interpretation for Idealized Flows. *Journal of the Atmospheric Sciences*, **59**, 1182–1197.

— 2003a: Atmospheric Energetics in the Wavelet Domain. Part II: Time-Averaged Observed Atmospheric Blocking. *Journal of the Atmospheric Sciences*, **60**, 319–338.

— 2003b: Instantaneous wavelet energetic transfers between atmospheric blocking and local eddies. *Journal of Climate*, in press.

Fuda, J. L., C. Millot, I. Taupier-Letage, U. Send, and J. M. Bocognano, 2000: XBT monitoring of a meridian section across the western Mediterranean Sea. *Deep-Sea Research I*, **47**, 2191–2218.

Fukumori, I. and C. Wunsch, 1991: Efficient representation of the North Atlantic hydrographic and chemical distributions. *Progress in Oceanography*, **27**, 111–195.

Gacić, M., S. Marullo, R. Santoleri, and A. Bergamasco, 1997: Analysis of the seasonal and inetrannual variability of the sea surface temperature field in the Adriatic Sea from AVHRR data (1984-1992). *Journal of Geophysical Research*, **102**, 22937–22946.

Gamage, N. and W. Blumen, 1993: Comparative Analysis of Low-Level Cold Fronts: Wavelet, Fourier and Empirical Orthogonal Function Decompositions. *Monthly Weather Review*, **121**, 2867–2878.

Gambis, D., 1992: Wavelet transform analysis of the length of the day and the El Niño/Southern Oscillation variations at intraseasonal and interannual time scales. *Annales Geophysicae*, **10**, 429–437.

Gasparini, G. P., G. Zodiatis, M. Astraldi, C. Galli, and S. Sparnocchia, 1999: Winter intermediate water lenses in the Ligurian Sea. *Journal of Marine Systems*, **20**, 319–332.

Ghil, M., M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, and P. Varadi, F. Yiou, 2002: Advanced Spectral Methods for Climatic Time Series. *Reviews of Geophysics*, **40**, 1–41.

Gomis, D., S. Ruiz, and M. A. Pedder, 2001: Diagnostic analysis of the 3D ageostrophic circulation from a multivariate spatial interpolation of CTD and ADCP data. *Deep Sea Research Part I*, **48**, 269–295.

He, R., R. H. Weisberg, H. Zhang, F. Muller-Karger, and H. R. W., 2003: A cloud-free, satellite-derived, sea surface temperature analysis for the West Florida Shelf. *Geophysical Research Letters*, **30**.

Hoffman, R. N., Z. Liu, J.-F. Louis, and C. Grasotti, 1995: Distorsion representation of forecast errors. *Monthly Weather Review*, **123**, 2758–2770.

Hopkins, T. S., 1999: The thermohaline forcing of the Gibraltar exchange. *Journal of Marine Systems*, **20**, 1–31.

Houseago-Stokes, R., 2000: Using optimal interpolation and EOF analysis on North Atlantic satellite data. *International WOCE Newsletter*, **38**, 26–28.

Jawerth, B. and W. Sweldens, 1994: An overview of wavelet based multiresolution analyses. *SIAM Review*, **36**, 377–412.

Jolliffe, I. T., 2002: *Principal Component Analysis*. Springer, 487 pp.

Jolliffe, I. T. and D. B. Stephenson, 2003: *Forecast Verification. A practitioner's Guide in Atmospheric Science*. Wiley, 240 pp.

Kantha, L. H. and C. A. Clayson, 2000: *Numerical models of Oceans and Oceanic Processes*. Academic Press, 940 pp.

Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, B. Blumenthal, and B. Rajagolapan, 1998: Analyses of Global Sea Surface Temperature 1856-1991. *Journal of Geophysical Research*, **103**, 18576–18589.

Kaplan, A., Y. Kushnir, M. Cane, and M. Blumenthal, 1997: Reduced Space Optimal Analysis for Historical Datasets: 136 Years of Atlantic Sea Surface Temperatures. *Journal of Geophysical Research*, **102**, 27835–27860.

Kearns, E. J., J. A. Hanafin, R. H. Evans, P. J. Minnett, and O. B. Brown, 2000: An Independent Assessment of Pathfinder AVHRR Sea Surface Temperature Accuracy Using the Marine Atmosphere Emitted Radiance Interferometer (MAERI). *Bulletin of the American Meteorological Society*, **81**, 1525–1536.

Kilpatrick, K. A., G. P. Podestá, and R. Evans, 2001: Overview of the NOAA/NASA avanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *Journal of Geophysical Research*, **106**, 9179–9197.

Kim, K. Y. and Q. Wu, 2000: Optimal Detection using cyclostationary EOFs. *Journal of Climate*, **13**, 938–950.

Klein, B., W. Roether, B. B. Manca, D. Bregant, V. Beitzel, V. Kovacevic, and A. Luchetta, 1999: The large deep water transient in the Eastern Mediterranean. *Deep-Sea Research I*, **46**, 371–414.

Kumar, P. and E. Foufoula-Georgiou, 1993a: A multicomponent decomposition of spatial rainfal fields. 1. Segregation of large- and small-scale features using wavelet transforms. *Water Resources Research*, **29, n8**, 2515–2532.

— 1993b: A multicomponent decomposition of spatial rainfal fields. 2. Self-similarity in fluctuations. *Water Resources Research*, **29, n8**, 2533–2544.

— 1997: Wavelet analysis for geophysical applications. *Reviews of Geophysics*, **35, n4**, 385–412.

Larnicol, G., N. Ayoub, and P. Y. Le Traon, 2002: Major changes in the Mediterranean Sea level variability from 7 years of TOPEX/Poseidon andERS-1/2 data. *Journal of Marine Systems*, **33-34**, 63–89.

Larnicol, G., P. Y. Le Traon, N. Ayoub, and P. De Mey, 1995: Mean sea level and surface circulation variability of the Mediterranean Sea from 2 years of topex/poseidon altimetry. *Journal of Geophysical Research*, **100, C12**, 25163–25177.

Lascaratos, A., W. Roether, N. Nittis, and B. Klein, 1999: Recent changes in deep water formation and spreading in the eastern Mediterranean Sea: a review. *Progress in Oceanography*, **44**, 5–36.

Lau, K. M. and H. Weng, 1995: Climate signal detection using wavelet transform: how to make a time series sing. *Bulletin of the American Meteorological Society*, **76, 12**, 2391–2402.

Le Traon, P. Y. and G. Didarboure, 1999: Mesoscale Mapping Capabilities of Multiple-Satellite Altimeter Missions. *Journal of Atmospheric and Oceanic Technology*, **16**, 1208–1223.

Le Traon, P. Y., G. Didarboure, and N. Ducet, 2001: Use of a High-Resolution Model to Analyze the Mapping Capabilities of Multiple-Altimeter Missions. *Journal of Atmospheric and Oceanic Technology*, **18**, 1277–1288.

Le Traon, P. Y., F. Nadal, and N. Ducet, 1998: An Improved Mapping Method of Multisatellite Altimeter Data. *Journal of Atmospheric and Oceanic Technology*, **15**, 522–534.

Lehoucq, R. B., D. C. Sorensen, and C. Yang, 1997: ARPACK user's guide: solution of large scale eigenvalue problems with implicity restarted Arnoldi methods, 1–152. URL `http://www.caam.rice.edu/software/ARPACK/`

Lermusiaux, P. F. J., 2001: Evolving the subspace of the three-dimensional multiscale ocean variability: Massachusetts bay. *Journal of Marine Systems*, **29**, 385–422.

— 2002: On the Mapping of Multivariate Geophysical Fields: Sensitivities to Size, Scales, and Dynamics. *Journal of Atmospheric and Oceanic Technology*, **19**, 1602–1637.

Lermusiaux, P. F. J. and A. R. Robinson, 2001: Features of dominant mesoscale variability, circulation patterns and dynamics in the Strait of Sicily. *Deep-Sea Research I*, **48**, 1953–1997.

Liang, X. S. and A. R. Robinson, 2004a: Localized Multiscale Energy and Vorticity Analysis. I. Fundamentals. *Dynamics of Atmospheres and Oceans*, submitted.

— 2004b: Localized Multiscale Energy and Vorticity Analysis. II. Instability theory and validation. *Dynamics of Atmospheres and Oceans*, submitted.

— 2004c: A study of the Iceland-Faroe Frontal variability using the Multiscale Energy and Vorticity Analysis. *Journal of Physical Oceanography*, submitted.

Liu, A. K., Y. S. Chang, M. K. Hsu, and N. K. Liang, 1998: Evolution of nonlinear internal waves in the East and South China seas. *Journal of Geophysical Research*, **103, C4**, 7995–8008.

Liu, P. C., 2000: Wave grouping characteristics in nearshore Great Lakes. *Ocean Engineering*, **27**, 1221–1230.

Liu, P. C. and G. S. Miller, 1996: Wavelet transforms and ocean current data analysis. *Journal of Atmospheric and Oceanic Technology*, **13**, 1090–1099.

Livezey, R. E. and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Eview*, **111**, 46–59.

Malanote-Rizzoli, P., B. B. Manca, M. Ribera d'Alcalà, A. Theocharis, A. Bergamasco, D. Bregant, G. Budillon, G. Civitarese, D. Georgopoulos, A. Michelato, E. Sansone, P. Scarazzato, and E. Souvermezoglou, 1997: A synthesis of the Ionian Sea hydrography, circulation and water mass pathways during POEM-Phase I. *Progress in Oceanography*, **39**, 153–204.

Malanote-Rizzoli, P., B. B. Manca, M. Ribera d'Alcalà, A. Theocharis, S. Brenner, G. Budillon, and E. Ozsoy, 1999: The Eastern Mediterranean in the 80s and in the 90s: the big transition in the intermediate and deep circulations. *Dynamics of Atmospheres and Oceans*, **29**, 365–395.

Mallat, S., 1989: A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis ans Machine Intelligence*, **11, n7**, 674–693.

— 1998: *A wavelet tour of signal processing*. Academic Press, 637 pp., 637 pp.

Manca, B. B., V. Kovačević, M. Gačić, and D. Viezzoli, 2002: Dense water formation in the Southern Adriatic Sea and spreading into the Ionian Sea in the period 1997-1999. *Journal of Marine Systems*, **33-34**, 133–154.

Martin, A. P., 2003: Phytoplankton patchiness: the role of lateral stirring and mixing. *Progress in Oceanography*, **57**, 125–174.

Marullo, S., R. Santoleri, P. Malanote-Rizzoli, and A. Bergamasco, 1999: The sea surface temperature field in the Eastern Mediterranean from advanced very high resolution radiometer (AVHRR) data Part I. Seasonal variability. *Journal of Marine Systems*, **20**, 63–81.

Mason, S. J., 1982: A model for assessment of weather forecast. *Australian Meteorological Magazine*, **30, n4**, 291–303.

Mauri, E. and P. M. Poulain, 2001: Northern Adriatic Sea surface circulation and temperature/pigment fields in September and October 1997. *Journal of Marine Systems*, **29**, 51–67.

MEDAR-Group, 2002: MEDATLAS/2002 database. Mediterranean and Black Sea database of temperature salinity and bio-chemical parameters. Climatological Atlas. *Ifremer edition*, **4 CD-ROM**.

Meyers, S. D., B. G. Kelly, and J. J. O'Brien, 1993: An introduction to wavelet analysis in oceanography and meteorology: with application to the dispersion of Yanai waves. *Monthly Weather Review*, **121**, 2858–2866.

Meyers, S. D. and J. J. O'Brien, 1994: Spatial and temporal 26-day SST variations in the equatorial Indian Ocean using wavelet analysis. *Geophysical Research Letters*, **21**, 777–780.

Miller, A. J., P.-M. Poulain, A. R. Robinson, H. G. Arango, and W. G. Leslie, 1995: Quantitative skill of quasi-geostrophic forecasts of a baroclinically unstable Iceland-Faroe Front. *Journal of Geophysical Research*, **100, C6**, 10833–10849.

Millot, C., 1999: Circulation in the Western Mediterranean Sea. *Journal of Marine Systems*, **20**, 423–442.

Millot, C., M. Benzohra, and I. Taupier-Letage, 1997: Circulation off Algeria inferred from the Médiprod-5 current meters. *Deep-Sea Research I*, **44**, 1467–1495.

Molcard, A., N. Pinardi, M. Iskandarani, and D. B. Haidvogel, 2002: Wind driven general circulation of the Mediterranean Sea simulated with a Spectral Element Ocean Model. *Dynamics of Atmospheres and Oceans*, **35**, 97–130.

Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly Weather Review*, **116**, 2417–2424.

— 1993: What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather and Forecasting*, **8**, 281–293.

— 1996: General Decompositions of MSE-Based Skill Scores: Measures of Some Basic Aspects of Forecast Quality. *Monthly Weather Review*, **124**, 2353–2369.

Murphy, A. H. and E. S. Epstein, 1989: Skill Scores and Correlation Coefficients in Model Verification. *Monthly Weather Review*, **117**, 572–581.

Murphy, A. H. and R. L. Winkler, 1987: A General Framework for Forecast Verification. *Monthly Weather Review*, **115**, 1330–1338.

— 1992: Diagnostic verification of probability forecast. *International Journal of Forecasting*, **7**, 435–455.

Naveau, P. and H.-S. Oh, 2004: Polynomial wavelet regression for images with irregular boundaries. *IEEE Transactions on Image Processing*, **13**, 773– 781.

North, G. R., T. L. Bell, and R. F. Calahan, 1982: Sampling Errors in the Estimation of Empirical Orthogonal Functions. *Monthly Weather Review*, **110**, 699–706.

Oh, H.-S., P. Naveau, and G. Lee, 2001: Polynomial boundary treatment for wavelet regression. *Biometrika*, **88**, 291–298.

Orlić, M., M. Gačcić, and P. E. La Violette, 1992: The currents and circulation of the Adriatic Sea. *Oceanologica Acta*, **15,2**, 109–124.

Ostrovskii, A. G., 1995: Signatures of stirring and mixing in the Japan Sea surface temperature patterns in autumn 1993 and spring 1994. *Geophysical Research Letters*, **22**, 2357–2360.

Pham, D. T., J. Verron, and R. M. C., 1998: A singular evolutive extended kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, **16**, 323–340.

Pierini, S. and A. Simioli, 1998: A wind-driven circulation model of the Tyrrenhian Sea area. *Journal of Marine Systems*, **18**, 161–178.

Pinardi, N., F. Auclair, C. Cesarini, E. Demirov, S. F. Umani, M. Giani, G. Montanari, P. Oddo, M. Tonani, and M. Zavaratelli: 2002, *Toward Marine Environmental Predictions in the Mediterranean Sea Coastal Areas: A Monitoring Approach*, Springer, chapter 17. 472.

Pinardi, N. and E. Masetti, 2000: Variability of the large scale general circulation of the Mediterranean Sea from observing and modelling: a review. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **158**, 153–173.

Pinot, J.-M., J. Tintoré, and D. Gomis, 1995: Multivariate analysis of the surface circulation in the Balearic Sea. *Progress in Oceanography*, **36**, 343–376.

Potts, J. M., C. K. Folland, I. T. Jolliffe, and D. Sexton, 1996: Revised "LEPS" Scores for Assessing Climate Model Simulations and Long-Range Forecasts. *Journal of Climate*, **9**, 34–53.

Poulain, P. M., 1999: Drifter observations of surface circulation in the Adriatic Sea between December 1994 and March 1996. *Journal of Marine Systems*, **20**, 231–253.

— 2001: Adriatic Sea circulation as derived from drifter data between 1990 and 1999. *Journal of Marine Systems*, **29**, 3–32.

Poulos, S. E., P. G. Drakopoulos, and M. B. Collins, 1997: Seasonal variability in sea surface oceanographic conditions in the Aegean Sea (Eastern Mediterranean): an overview. *Journal of Marine Systems*, **13**, 225–244.

Raick, C., A. Alvera-Azcárate, A. Barth, J.-M. Brankart, K. Soetaert, and M. Grégoire, 2004: Application of a SEEK filter to a 1d biogeochemical model of the Ligurian sea: twin experiments and real in situ data assimilation. *Journal of Marine Science*, submitted.

Rao, Y. R. and C. R. Murthy, 2001: Comparative Analysis of Spectral, Empirical Orthogonal Function, and Wavelet Decompositions for an Uwelling Event. *Journal of Atmospheric and Oceanic Technology*, **18**, 704–710.

Reynolds, R. W. and T. Smith, 1994: Improved Global Sea Surface Temperature Analyses Using Optimum Interpolation. *Journal of Climate*, **7**, 929–948.

Rixen, M., J.-M. Beckers, J.-M. Brankart, and P. Brasseur, 2001: A numerically efficient data analysis method with error map generation. *Ocean Modelling*, **2**, 45–60.

Robinson, A. R. and W. G. Leslie, 1985: Estimation and Prediction of Oceanic Eddy Fields. *Progress in Oceanography*, **14**, 485–510.

Robinson, A. R., J. Sellschopp, A. Warn-Varnas, W. G. Leslie, C. J. Lozano, P. J. Haley Jr., L. A. Anderson, and P. F. J. Lermusiaux, 1999: The Atlantic Ionian Stream. *Journal of Marine Systems*, **20**, 129–156.

Roebber, P. and L. F. Bosart, 1996: The Complex Relationship between Skill and Forecast Value: A Real-World Analysis. *Weather and Forecasting*, **11**, 544–559.

Roussenov, V., E. Stanev, V. Artale, and N. Pinardi, 1995: A seasonal model of the Mediterranean Sea general circulation. *Journal of Geophysical Research*, **100**, 13515–13538.

Sammari, C., C. Millot, and L. Prieur, 1995: Aspects of the seasonal and mesoscale variabilities of the Northern Current in the Western Mediterranean Sea inferred from the prolig-2 and pros-6 experiments. *Deep-Sea Research I*, **42,6**, 893–917.

Saunders, R. W. and K. T. Kriebel, 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing*, **9**, 123–150.

Send, U., J. Font, G. Krahmann, C. Millot, M. Rhein, and J. Tintoré, 1999: Recent advances in observing the physical oceanography of the western Mediteranean Sea. *Progress in Oceanography*, **44**, 37–64.

Shen, S. S. P., G. R. North, and K.-Y. Kim, 1994: Spectral Approach to Optimal Estimation of the Global Average Temperature. *Journal of Climate*, 1999–2007.

Smith, T. M., R. W. Reynolds, R. E. Levezey, and D. C. Stokes, 1996: Reconstruction of Historical Sea Surface Temperatures Using Empirical Orthogonal Functions. *Journal of Climate*, **9**, 1403–1420.

Smith, W. H. F. and D. T. Sandwell, 1997: Global sea floor topography from satellite altimetry and ship depth soundings. *Science*, **277**, 1956–1962.

Soetaert, K., P. M. J. Hermana, J. J. Middelburg, C. Heip, C. L. Smith, P. Tett, and K. Wild-Allen, 2001: Numerical modelling ofthe shelfbreak ecosystem: reproducing benthic and pelagic measurements. *Deep-Sea Research*, **48**, 3141–3177.

Sparnocchia, S., G. P. Gasparini, M. Astraldi, M. Borghini, and P. Pistek, 1999: Dynamics and mixing of the Eastern Mediterranean outflow in the Tyrrhenian basin. *Journal of Marine Systems*, **20**, 301–317.

Sparnocchia, S., P. Picco, G. M. R. Manzella, A. Ribotti, S. Copello, and P. Brasey, 1995: Intermediate water formation in the Ligurian Sea. *Oceanologica Acta*, **18**, 151–162.

Sparnocchia, S., N. Pinardi, and E. Demirov, 2003: Multivariate Empirical Orthogonal Function analysis of the upper thermocline structure of the Mediterranean Sea from observations and model simulations. *Annales Geophysicae*, **21**, 167–187.

Spedding, G. R., F. K. Browand, N. E. Huang, and S. R. Long, 1993: A 2-D complex wavelet analysis of an unsteady wind-generated surface wave field. *Dynamics of Atmospheres and Oceans*, **20**, 55–77.

Strass, V. H., 1992: Chlorophyll patchiness caused by mesoscale upwelling at fronts. *Deep-Sea Research*, **39**, 75–96.

Sweldens, W., 1997: The lifting scheme: A construction of second generation wavelets. *SIAM Journal of Mathematical Analysis*, **29**, 511–546.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, **106, D7**, 7183–7192.

Teti, J. G. and H. N. Kritikos: 1992, Weyl-Heisenberg and wavelet coherent frames for SAR ocean image filtering. *IEEE Geoscience and Remote Sensing Symposium*, volume 2, 1318–1320.

Theocharis, A., B. Klein, K. Nittis, and W. Roether, 2002: Evolution and status of the Eastern Mediterranean Transient (1997-1999). *Journal of Marine Systems*, **33-34**, 91–116.

Thomson, R. and I. V. Fine, 2003: Estimating Mixed Layer Depth from Oceanic Profile Data. *Journal of Atmospheric and Oceanic Technology*, **20**, 319–329.

Torrence, C. and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, **79, 1**, 61–78.

Toumazou, V. and J. F. Cretaux, 2001: Using a lanczos eigensolver in the computation of Empirical Orthogonal Functions. *Monthly Weather Review*, **125,5**, 1243–1250.

Tusseau, M. H. and J. M. Mouchel: 1994, Nitrogen inputs to the Gulf of Lions via the Rhône river. *Water Pollution Research Reports, Proceedings of the EROS 2000 Workshop, Hamburg*, J. Martin and H. Barth, eds., commission of the European Communities.

Vidakovic, B. and P. Muller, 1994: Wavelets for kids. *AMS*.

Vignudelli, S., G. P. Gasparini, M. Astraldi, and M. Schiano, 1999: A possible influence of the North Atlantic Oscillation on the circulation of the Western Mediterranean Sea. *Geophysical Research Letters*, **26**, 623–626.

Viudez, A., J. Tintoré, and R. Haney, 1996: Circulation in the Alboran Sea as determined by quasi-synoptic hydrographic observations. Part I: three dimensional structure of the two Anticyclonic Gyres. *Journal of Physical Oceanography*, **26,5**, 684–705.

von Storch, H. and W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.

Wang, J. and C. S. Chern, 2000: The Wavelet Empirical Orthogonal Function and its Application to the Analysis of Interdal Tides. *Journal of Atmospheric and Oceanic Technology*, **17**, 1403–1420.

Weare, B. C. and J. S. Nasstrom, 1982: Examples of Extended Empirical Orthogonal Function Analysis. *Monthly Weather Review*, **110**, 481–485.

Weng, H. and K. M. Lau, 1994: Wavelets, period doubling, and time-frequency location with application to organization of convection over the Tropical Western Pacific. *Journal of Atmospheric Sciences*, **51, 17**, 2523–2541.

Wick, G. A., W. J. Emery, and P. Schluessel, 1992: A Comprehensive Comparison Between Satellite-Measured Skin and Multichannel Sea Surface Temperature. *Journal of Geophysical Research*, **97**, 5569–5595.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Yiou, P., D. Sornette, and M. Ghil, 2000: Data-Adaptive Wavelets and Multi-Scale Singular-Spectrum Analysis. *Physica D.*, **142**, 254–290.

Younes, W. A. N., N. Bensoussan, J.-C. Romano, D. Arlac, and M.-G. Lafont, 2003: Seasonal and interannual variations (1996-2000) of the coastal waters east of the Rhone River mouth as indicated by the SORCOM series. *Oceanologica Acta*, **26**, 311–321.

Zepeda-Arce, J., E. Foufoula-Georgiou, and K. K. Droegmeier, 2000: Space-time rainfall organization and its role in validating quantitative precipitation forecasts. *Journal of Geophysical Research*, **105, D8**, 10129–10146.

Zervakis, V., K. Nittis, L. Perivoliotis, and C. Tziavos, 2002: A comparison of model predictions to observations of seasonal variability and circulation in the Eastern Mediterranean. *The Global Atmosphere and Ocean System*, **8**, 141–162.