Alma Mater Studiorum · Università di Bologna

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Fisica

Estimation of the Adriatic Sea general circulation using Data Assimilation techniques

Tesi di Laurea in Oceanografia e Dinamica della Zona Costiera

Relatore:
Prof.
NADIA PINARDI

Co-Relatore:
Dott.
SRDJAN DOBRICIC

Presentata da:
STEFANO MAFFEI

Prima Sessione
Anno Accademico 2011/2012

Estimation of the Adriatic Sea general circulation using Data Assimilation techniques

Stefano Maffei

June 27, 2012

Abstract

Data Assimilation is a powerful technique that enable us to meld together the information coming from the observations of a natural system and a first guess provided by a prognostic model of the system.
In the presented work we test a 3D-Var Data Assimilation scheme implemented together with a regional OGCM (Ocean General Circulation Model) of the Adriatic Sea: the AREG2 (Adriatic REGional) model.
The target year for our test is the 2006. For this period the model provides daily means of the 2D and 3D fields of physical interest (temperature, salinity, currents, sea surface elevation and others). The data assimilated with the 3D-Var algorithm are in-situ measurements of vertical profiles of temperature and salinity and satellite measured SLA (Sea Level Anomaly).
The quality assessment of the Adriatic Sea Analysis System is based on the misfit calculated comparing the results provided by the model to the assimilated in-situ data and to an independent dataset of sea level observations from tide gauges.
After the quality assessment the effect of the assimilation on the general circulation simulated by the model will be investigated.
The results show that the 3D-Var scheme is effective in enforcing the main aspects of the general circulation and that it is able to correctly spread the information to physical quantities that are not directly observed.

Abstract

La Data Assimilation è una tecnica che permette di unire le informazioni provenienti da osservazioni di un sistema naturale e da una ipotesi a priori fornita da un modello prognostico del sistema stesso.
Nel lavoro presentato verrà testato uno schema di Data Assimilation di tipo 3D-Var implementato insieme ad un modello oceanografico per la circolazione generale (OGCM) nel mare Adriatico: il modello AREG2 (Adriatic REGional).
L'anno analizzato è il 2006. Per questo periodo il modello calcolerà le medie giornaliere di campi 2D e 3D di interesse fisico (temperatura, salinità, correnti, altezza della superficie del mare e altri). I dati assimilati con l'algoritmo 3D-Var saranno misurazioni in-situ di profili verticali di salinità e temperatura e SLA (Sea Level Anomaly) misurata da satellite.
La stima della qualità del sistema di analisi del mare Adriatico si baserà sulle differenze (misfit) calcolate confrontando i risultati forniti dal modello con i dati in-situ assimilati e con un insieme di dati indipendente di altezza del livello del mare misurato da stazioni mareografiche.
Dopo questa stima verrà indagato l'effetto dell'assimilazione sulla circolazione generale simulata dal modello.
I risultati mostreranno che lo schema 3D-Var è efficace nel rinforzare gli aspetti principali della circolazione generale e che è capace di diffondere correttamente le informazioni a quantità fisiche che non sono state direttamente osservate.

Disclaimer

The present thesis has been written in the framework of a work within GNOO (Gruppo Nazionale di Oceanografia Operativa) group. However, due to lack of time, the results exposed have not been validated by the GNOO scientists.
Therefore it must be consider the possibility that the results presented in the following work are not suitable to be used in other scientific projects.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Adriatic Sea

The Adriatic Sea is one of the principal sub-basins of the Mediterranean Sea (figure 1.1). It is a semi-enclosed basin situated between the Italian peninsula and the Balkan regions and it extends in the NNW-SSE direction for about 770 km from the latitude of about 39° N to 45° 50' N and has a mean width of 160 km. It has a surface of about 138.600 km^2 and a total volume of approximately 35.000 km^3.

The Adriatic Sea is surrounded by the mountain chains of the Apennines to the West, the Dinaric Alps to the North-East and the Balkans to the East.

The depth does not exceed 300 m in the northern part but reaches more than 1200 m in the southern area, along the route from Bari to the mouths of Kotor.

The average salinity is 3.8 ‰, with strong differences between the northern, less saline, and the south.

The Adriatic Sea is connected with the Mediterranean Sea (in particular with the Ionian Sea) by the Otranto Channel. This represents the southern boundary of the basin and it is recognized as an important region of exchange of water masses that originate from the Adriatic Sea, the Ionian Sea and the eastern Aegean Sea, influencing the thermohaline circulation of the adjacent basins. The Adriatic Sea is, in fact, a well-known site of dense water formation related to the winter surface heat losses, as well as to the ingression into the basin of Levantine Intermediate Water.

The Adriatic Sea has many features that make its dynamics very special in many respects. Numerical simulation of the circulation of this basin has always been a challenge and it is complicated by the particular bathymetry of the basin, the high variability of the atmospheric forcings (in particular, the winds) on the interannual and seasonal time scales and by the abundant supply of fresh water due to the numerous rivers that flow into the basin. In particular the Adriatic Sea contains both open ocean and shelf processes with a strong interaction between them.

In this chapter we will describe in more details this features and the resulting general circulation.

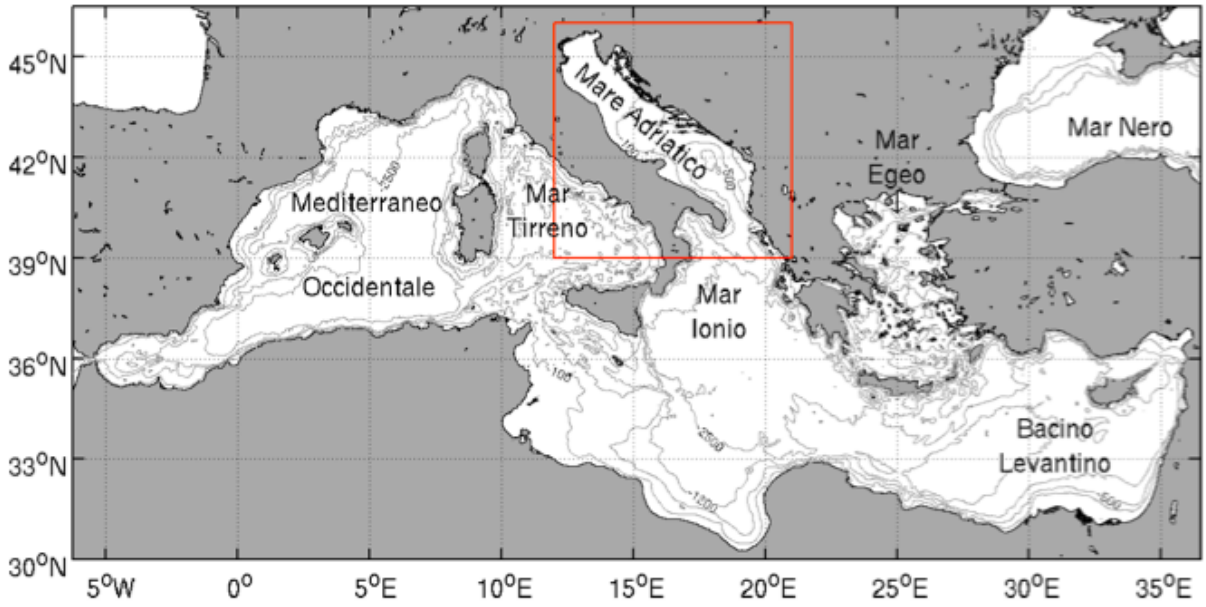


Figure 1.1: *The Mediterranean Sea and It's principal sub-basins, with the Adriatic Sea highlighted in the red box*

1.1.1 Description of the basin and principal forcings

Basin bathymetry and surrounding morphology

First of all we will briefly describe the bathymetry of the Adriatic Sea: it presents strong asymmetry along and across the basin and this is one of the reasons for the numerical modeling of the circulation of the basin being so challenging. A good model of the Adriatic Sea must be able to span from the deep ocean conditions (in the southern part) to the shallow and coastal water environment modeling (especially in the northern part of the basin).
Hence, on the basis of the bottom morphology, the Adriatic Sea is conventionally divided into three sub-basins: the northern, the middle and the southern (figure 1.3). As we shall see later in this work, this is not a purely formal split, as even the major structures of the climatological circulation of the Adriatic Sea, the gyres, are clearly located in the southern, central and northern part of the basin and thus are respectively called North Adriatic (NAd) Gyre, Middle Adriatic (MAd) Gyre and South Adriatic (SAd) Gyre (see figure 1.12).

The northern sub-basin is the region that includes the line of coast between Venice and Trieste, extending to the line connecting Ancona to Zadar. It is characterized by a low average depth (30 m) with a weak bathymetric gradient along the major axis. Therefore, the whole northern Adriatic Sea is considered as a continental shelf area.
The northern Adriatic alone receives about 15% of all the riverine freshwater input of the whole Mediterranean Sea. Moreover in this area of the Adriatic, in winter period, due to

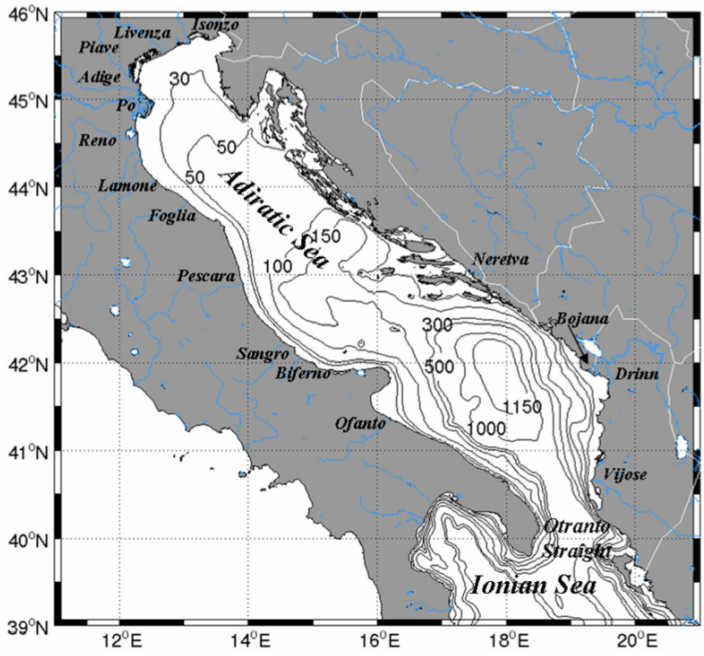


Figure 1.2: *Adriatic Sea domain, bathymetry (solid contour lines), and main rivers discharging into the basin*

Bora the denser water of the Mediterranean is formed($\sigma_t > 29.9$ kg/m^3, [3] [1])
The central sub-basin represents a transition zone between the north and south of the basin. Its average depth is around 140 m and is characterized by the presence of two depressions at its northern end, the Pomo Depressions, reaching a maximum depth of 270 m, and by the presence of a seamount at the center of the sub-basin, which rises up to the depth of 60 meters below the sea level. This sub basin ends with the saddle of Palagruža. The southern sub-basin, which includes the southern Adriatic Pit and extends to the Strait of Otranto, is characterized by a very narrow continental shelf (20-30 km), and a very steep slope that reaches up to 1000 m, while the maximum depth reaches 1299 m. The southern Adriatic comprises about 80% of the total volume of the Adriatic and occupies about 55% of its area.
At its southern end, the Adriatic Sea is connected to the Mediterranean Sea by the Otranto Channel, which is relatively wide and deep: its minimum width and depth are respectively about 72 km and 780 m. The vertical size of the Otranto Channel allows for a relevant water exchange, while its horizontal size, larger than the Rossby radius for the Adriatic Sea, implies a pronounced influence of the Coriolis force on the dynamics of the water masses flowing from and in the Adriatic through the Channel.

The asymmetry across the Adriatic mostly consists of a different morphology of the coastal areas (figure 1.4). The line of the west coast (Italian) is relatively uniform and regular, without islands and with a platform of gentle slope. The eastern coastal area,

[1] σ_t is a function of salinity, temperature and pressure so that we can define the density of seawater as $\rho(S,T,p) = 1000Kg/m^3 + \sigma_t(S,T,p)$

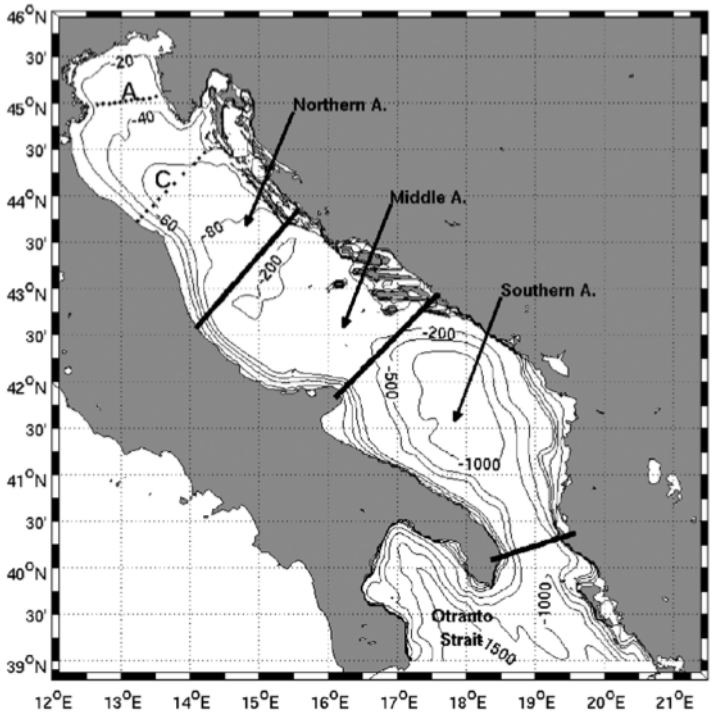


Figure 1.3: *Bathymetry of the Adriatic Sea. Also shown are the three different region defined on the basis of the bottom morphology. From [42]*

instead, is characterized by the numerous islands of Dalmatia, and by a very irregular bottom that rapidly drops to the open sea.
Another source of asymmetry between the two opposite coasts is the different orography at the two sides of the basin (see figure 1.4): while along the eastern side we can notice the presence of relatively high mountain chains just on the right of the coastlines (the Dinaric Alps), the Appennines on the Italian side of the basin are less steep and more distant from the coast.
This contrast between the two main Adriatic coasts is reflected in significant differences between the dynamics of their coastal currents. The simpler structure of the west coast allows the development of a smooth coastal current, the Western Adriatic Current (WACC, Western Adriatic Current), whose spatial variations are due both to the variability of water flow and forcing surface and to the intrinsic instability of the current itself, producing meandering and cell recirculation. In contrast, the many inlets of the east coast forced coastal currents to abrupt reversals, while the surface winds flowing through the channels created by the islands and through the gorges of the mountains contributes to the formation of local jets, which have more complex structures than the meanderings of the opposite coast and strongly influencing the spatial variability of the water currents.

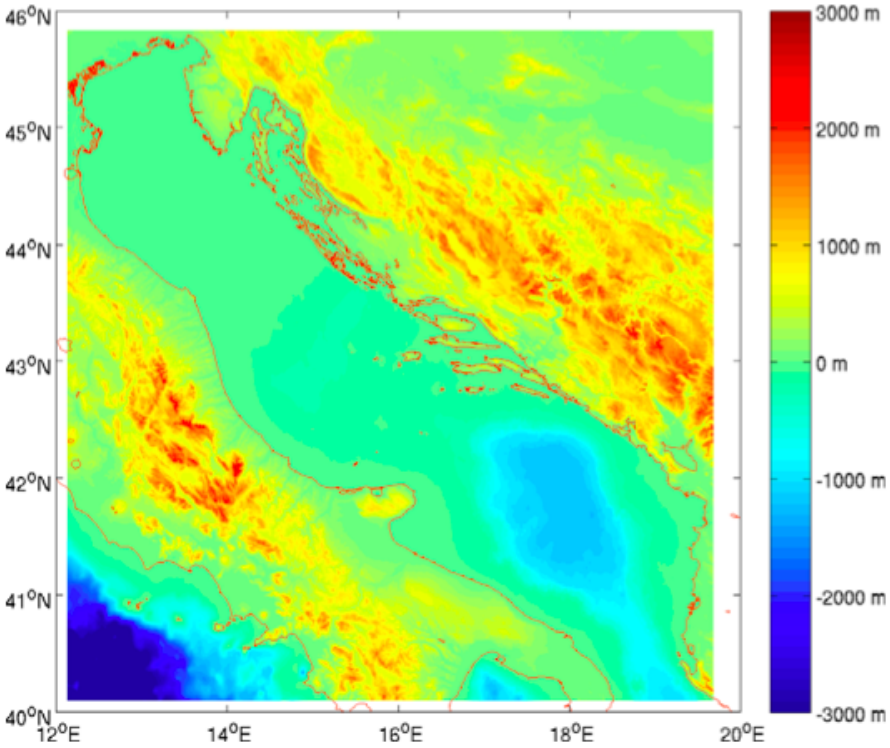


Figure 1.4: *Orography of the lands surrounding the Adriatic Sea.*

Winds in the Adriatic Sea area

The Adriatic Sea, mainly in the northern region, is strongly influenced by atmospheric forcings, due to its geographical location, the orography surrounding the basin and its depth (see figure 1.4). This implies that the oceanographic properties of the Adriatic, the circulation and the distribution of water masses , highly depend on the characteristics of sea-atmosphere exchanges.
The two major wind regimes are Bora and Scirocco (see figure 1.5).

The Bora is a strong flow of continental air blowing on the Adriatic Sea mainly from North-Northeast. The Bora is a katabatic wind, cold and dry forming when a mass of cold, dry air accumulates on the Balkans, and reached the height of the mountain passes, is channeled along the valleys and gorges of the Dinaric Alps forming jets that blow violently on some areas of the Adriatic Sea.
These winds blow with an average speed of 15 m/s reaching peaks of 50 m/s. Generally, events with intensity of at least 15 m/s have a frequency ranging from one day per month (or less) in the summer, up to more than 6 days per month during winter with an average duration between 12 hours and 2 days.
This wind is most common during the cold season, from November to March. The Bora can be very localized and extended only a few miles on the Adriatic Sea; in some cases, the area of the winds extends south to include much of the Adriatic.
The variability of the Bora winds is thought to be one of the principal causes of the

high frequency perturbation in the Adriatic general circulation, especially in the northern sub-basin.

The Scirocco is a tropical wind that blows over the Mediterranean Sea from South-East and generates in the deserts of North Africa and Arabia. It is a wet wind (humidity is collected during the passage across the Ionian Sea) which is channeled along the major axis of the Adriatic basin due to the presence of the Apennines and the Dinaric Alps. The Scirocco tend to blow without a favorite month or season but very evident Scirocco events occur in autumn and winter.
These winds have average intensity of 10 m/s, but can reach intensities greater than 15 m/s, especially in winter, and in spring it can reach even 30 m/s. The Scirocco is connected with the passage of low-pressure systems over the basin, causing the relative increase in sea level in this region due to the inverse barometer effect and to the direct sea level set-up by the wind. This effect is particularly strong in the northern Adriatic and is partly responsible for the effect of "high water" (Acqua Alta) in Venice.

In addition to the Bora and Scirocco described so far, other winds like the Etesian, the Mistral and the marine and land breezes can be locally dominant, especially in summer. In particular the Etesian are formed in consequence of a low pressure zone located over the South-West region of the Asian continent during summer. These winds dominate the oriental Mediterranean during the warm months blowing from North-West and in the Adriatic Sea they represent the large part of the summer winds.

Freshwater inputs and heat fluxes

One of the most important features one must take into account in studying the dynamics of the Adriatic Sea is the abundant input of riverine freshwater from the many rivers that flow in the basin (refer to figure 1.2). In average, the basin receives from its rivers 5700 m^3/s of fresh water ([52], [53]).
The major of these rivers is the Po: alone, it contributes for about the 28 % of the total fresh-water input in the sea (that means, about 1585 m^3/s). Thus the greatest part of the river discharge is concentrated in the northern sub-basin.
Just to give an idea of the orders of magnitude, the second source in order of importance, after the Po, is commonly considered to be the sum of the Albanian rivers and the surrounding drainage areas whose contribution has been estimate to be between 923 m^3/s and 1244 m^3/s.
The flow of the rivers is clearly not constant in the course of the year, depending on the rainfalls and on the melting of the snow. In fact, the total river runoff shows a maximum in autumn (corresponding to the maximum rainfall rate) and in spring (when the snow melting contribution is largest) and it is minimum during the summer months. However there are local deviations from this general trend: in fact the northern rivers shows a minimum during winter.

The river runoff in the Adriatic Sea would imply (in a closed basin) a sea level rise of about 1.17 m/year for the entire basin ([53]). We have to add the river discharge

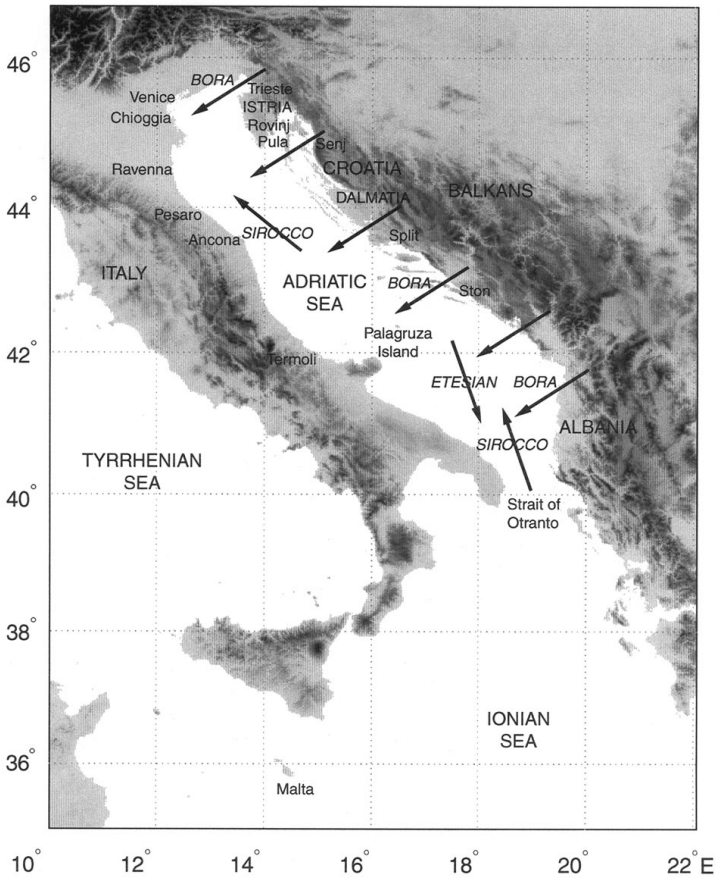


Figure 1.5: *The principal winds that blow over the Adriatic Sea*

contribution (R) to the surface water balance which is the difference between evaporation (E) and precipitation (P). The total water balance is then:

$$\text{W=E-(P+R)} \tag{1.1}$$

Various authors tried to estimate this budget and always found negative values indicating a gain of fresh water for the basin. In particular Raichic ([53]) found a climatological value between -0.65 m/year and -1.10 m/year, essentially due to the river runoff since the evaporation and precipitation contributes balance each other on climatological time scales.
It has been conclused that the Adriatic Sea may be considered as a source of fresh water for the whole Mediterranean Sea where, conversely, the budget W=E-(P+R) is positive (about $+0.55$ m/year, according to [46]). The excess of fresh water is transferred to the Mediterranean through the Strait of Otranto; this mass of water is lighter than salt water, and generates therefore a surface current thus establishing a so called estuarine thermohaline circulation.

The Adriatic Sea (like any other ocean or sea) interacts with the atmosphere also by

the exchange of heat through the sea surface.
The total heat flux Q_T through the air-sea interface is the sum of different contributions,

$$Q_T = Q_S + Q_B + Q_H + Q_E \tag{1.2}$$

where:

- Q_S is the incident short-wavelength solar radiation flux. This always a positive term, indicating a gain of heat of the sea.
- Q_B is the backward long wave radiation. It is (nearly) black body radiation emitted from the surface of land or sea or radiation emitted from the atmosphere after absorption of solar short wave radiation. From the point of view of the sea this is always a negative term, indicating a loss of heat.
- Q_H is the sensibile heat flux that is lost or gained by conduction or convection. It can be either positive or negative, depending on the gradient of temperature at the air-sea interface.
- Q_E is the latent heat flux leaving the sea surface during evaporation processes or an incoming flux during condensation. Like Q_H this term can be either positive or negative.

These fluxes depend on the solar radiation (which can be approximated by the Solar Constant, but is highly variable on various time scales), the cloud cover and the physical state of the atmosphere and sea at their interface, namely the temperature of air and water below and above the sea surface, the velocity of the wind and the surface water current and the humidity.
The evaluation of these terms can be done by using the so-called *Bulk formulas*: semi-empirical relationships between some crucial but easy to obtain parameters and the quantity of interest. More details about bulk formulas used in physical oceanography may be found in [2], [37] and in other fundamental texts on physical oceanography. An explanation of the formulas used to evaluate the terms in equation 1.2 can be found in [4]. Here and other similar articles various scientists evaluated the various terms in equation (1.2) finding that the total balance oscillate from -17 W/m^2 to -22 W/m^2. Negative values indicate that the Adriatic Sea loses heat and this lost is observed to be greater during winter. To maintain the thermal balance of the basin, it has to import heat from the Mediterranean through the Straits of Otranto; we shall see later that this implies the establishment of an antiestuarine thermohaline circulation, opposite to the effects of the freshwater flux.

1.1.2 General circulation in the Adriatic Sea

As already mentioned, the forcing of general circulation in the Adriatic Sea has three major components, perhaps equally important for the overall Adriatic dynamical engine. The first component is river runoff, characterized by the low salinity waters derived mainly

from the Po and Albanian Rivers. The Po runoff, in particular, is an important component of the buoyancy budget in the overall basin. In fact, we have an overall heat loss together with a water gain, in contrast to the overall Mediterranean, which exhibits a water loss.
The second component is the wind and heat forcing at the surface, which produce deep-water masses in the northern and southern Adriatic and forces the circulation to be seasonal.
The third component is the Otranto Channel forcing, which inputs heat and salt in the circulation as a restoring mechanism for the northern heat losses and water gains.

Due to the variability of these forcings the circulation in the basin presents oscillations at various temporal and spatial scales and many aspects of the driving processes are still a matter of debate, being probably a result of baroclinic/barotropic non-linear instabilities in the basin.

Spatial gradients in density are due to the combined effects of salinity and temperature and are the real engine of the circulation in the basin. Thus, in order to understand the circulation in the Adriatic Sea we first have to describe the temperature and salinity fields and profiles in the basin.

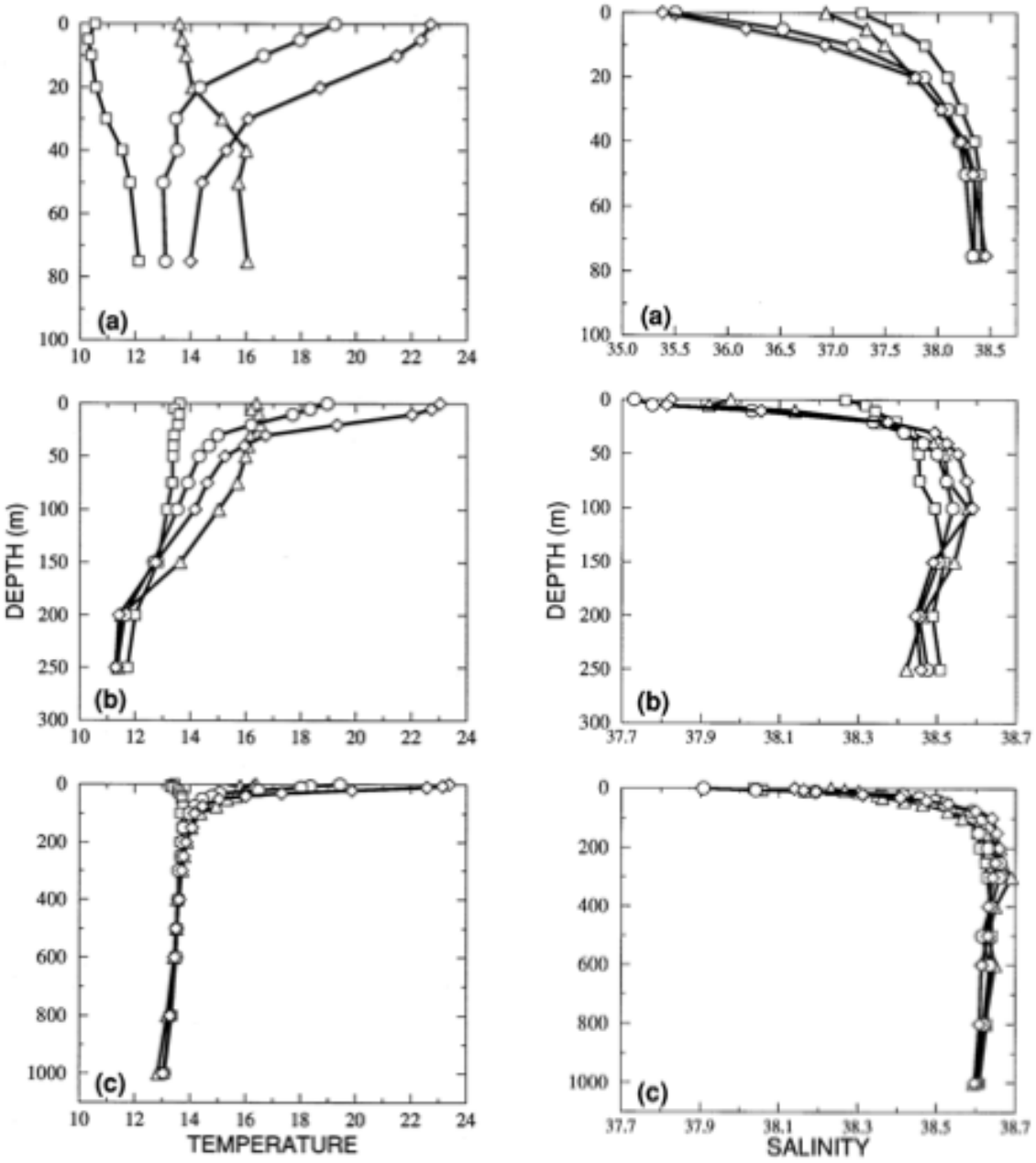
Water mass properties

As already stated, the Adriatic Sea is a basin of fresh-water formation, principally due to its rivers, especially the Po and the Albanian rivers.
Studying a dataset spanning from 1911 to 1983, Artegiani et al ([4] and [4]) were able to draw vertical profiles and horizontal fields for temperature and salinity shown respectively in figures 1.6a,1.6b 1.7, 1.8, 1.9 and 1.10. From these figures seasonal variations are easily identified. Note that these figures are obtained by seasonally averaging the entire dataset analyzed, so they are not able to capture features at sub-seasonal time scales.

Figures 1.6 show the climatological profiles for temperature and salinity. As already mentioned the splitting of the Adriatic Sea into three sub-basins is not only a matter of bathymetry, but is a distinction that can be tracked also in the water properties and (as we will briefly see) in the horizontal structures of the general circulation.
In the northern Adriatic the entire water column exhibits an evident seasonal thermal cycle. A well-developed thermocline is clearly present (from figure 1.6a) in spring and summer down to 30 m depth, whereas a significant cooling begins close to the surface in autumn when the bottom temperature reaches its maximum value, probably due to increased vertical mixing and intrusion of middle Adriatic waters as well as heat absorbed through the surface in summer and diffused downward to the bottom. Only in winter the cooling of the whole water column occurs; in this season temperature generally increases down to the bottom, but the water column stability is preserved due to an associated increase of salinity at depth as clearly visible in the upper panel of figure 1.6b. The freshwater effect is clearly seen in spring and summer due to the increased runoff and the increased water column stratification.

From the vertical profiles of salinity and temperature we can recognize a seasonal layer of northern Adriatic surface water (NAdSW), which corresponds to low salinities and relatively high temperatures of the summer, and a northern Adriatic deep water (NAdDW) layer. On the basis of their results obtained, Artegiani et al define the NAdDW having average characteristics of T = (11.35 ± 1.40) °C, S = (38.30 ± 0.28) psu, and density σ_t > 29.2 kg / m^3 ([4], [42]).
The dense, deep water in the northern Adriatic is formed mainly in winter, when the surface cooling due to heat loss is more efficient also thanks to cold and dry winds blowing during this seasons from North-East (Bora).
In the middle Adriatic the spring–summer thermocline is formed down to a depth of 50 m. In the layer from 50 to 150 m the seasonal temperature changes are still observed. The surface waters are freshened during spring and summer, as in the northern Adriatic, due to river runoff.
Comparing the salinity profiles we can see that below the pycnocline the salinity in the central sub-basin is higher than in the north and is much higher in the southern Adriatic; this is due to the input of salty water from the Otranto Channel. The principal water current responsible for this effect is the modified Levantine Intermediate Water (MLIW), so called because it flows at intermediate depth. The MLIW is considered to be an external forcing, together with the Scirocco, which enforces the input flux through the Otranto Channel.
The Pomo depression are the only regions of the middle Adriatic that are deeper than 150 m and are filled with a deep water mass, which exhibits some seasonal changes: the middle Adriatic deep water (MAdDW). The MAdDW has relatively low average temperature (T = (11.62 ± 0.75) °C) and substantially higher average salinity (S= (38.47 ± 0.15) psu) than the NAdDW but the density can still be defined as $\sigma_t > 29.2$ kg/m^3.
Dense water is formed in the middle Adriatic also thanks to convective processes, but it is thought that the primary source of the dense water here is the NAdDW flowing from north to south, filling the Pomo Depression. The MAdDW can episodically escape the middle Adriatic depressions and flowing south, contributing to the south Adriatic deep waters (SAdDW).

In the southern Adriatic the seasonal thermocline extends down to approximately 75 m. The seasonal cycle of the surface waters is driven by the fresh coastal waters as seen by the decrease in surface salinity in all seasons and the augmented minimum during spring and summer. From 150 m to the bottom, Mediterranean open sea conditions are found since the water is almost homogeneous, with a relatively weak seasonal signal down to 300 m, modulated by MLIW advection and mixing. In this region the MLIW is defined by S > 38.6 psu and T > 13.5 °C, with a layer between 150 and 400 m. The southern Adriatic deep water (SAdDW) again has different average characteristics from NAdDW and MAdDW. It is defined by T = (13.16 ± 0.30) °C and S = (38.61 ± 0.09) psu, corresponding to σ_t = 29.1 kg/m^3. Thus, this water is considerably warmer and saltier compared to NAdDW and MAdDW so that it probably comprises a mixture of MLIW and local surface waters. In fact, in the southern Adriatic the deep water production driven by open-sea like vertical convection processes is more efficient than in the middle sub-basin, and is as important



(a) *Seasonal climatological profiles of temperature (°C) for (a) northern, (b) middle, and (c) southern Adriatic for winter (□), spring (◯), summer (⋄), and autumn (△)..*

(b) *Seasonal climatological profiles of salinity (psu) for (a) northern, (b) middle, and (c) southern Adriatic for winter (□), spring (◯), summer (⋄), and autumn..*

Figure 1.6: *Temperature and salinity profiles for the Adriatic sub-basins*

as the advection of dense water from North.

From figures 1.7 and 1.8 we can now see the seasonal features of the surface temperature and salinity fields. The seasonal temperature excursion is clearly due to the heat flux exchanged with the atmosphere. Note that the horizontal gradient is smoother in winter than in other seasons.
Concerning the salinity surface fields we can note in all seasons strong salinity frontal areas, particularly along the western coast, related to the river runoff. Frontal structures are determined by the strong gradients between the low salinity waters, which are always present along the western side of the Adriatic Sea, and the interior basin salinity field. From spring to summer the relatively fresh waters of the northern Adriatic spread southeastward, intruding into the open sea. Both salinity and temperature fields show large-scale patchiness during the spring–summer seasons, suggesting a higher baroclinic eddy activity.

Similar fields for the surface at 100 m depth are shown in figure 1.9 for temperature and in figure 1.10 for salinity. The temperature fields at 100-m depth shown are similar to those at the surface only in winter when the water column is homogenized, as shown in figure 1.6a. In all the seasons, warmer waters are present on the eastern side of the basin, particularly in autumn, in connection with the development of an eastern current that flow northward (the E-SAd, see figure 1.12), which is known to be correlated to the MLIW intensity.
The salinity fields at 100 m in figure 1.10 do not exhibit any large-scale structure and they are uncorrelated with the surface fields. The ingression of MLIW is evident from spring to autumn from the higher isohalines reaching the middle Adriatic (S $>$ 38.6 psu).

General circulation

From the temperature and salinity fields it is easily possible to calculate the density by using empirical formulas that can be found, for example, in [37]. Then it is possible to derive the baroclinic velocity field by using the dynamic method proposed in [2], or, which is an equivalent formulation, by calculating the *dynamic height* (actually, a dynamic height *anomaly*) relative to some depth. The dynamic height is measured in dynamic meters and is defined as:

$$D(p_1, p_2) = \int_{p_1}^{p_2} \delta(T, S, p) dp$$

where p is the pressure, p_1 and p_2 are two pressure reference levels, δ is the specific volume anomaly, T is the temperature and S is the salinity. Calculating horizontal fields of dynamic height is analogous to a meteorologist's use of a pressure chart, with the direction of flow aligned with the contours and the intensity of flow inversely proportional to the contour spacing. Dynamic heights are preferred over geometric heights in oceanography and meteorology because energy is generally lost or gained when a parcel of fluid moves along a surface of equal geometric height but not when it moves along a surface of equal dynamic height. In meteorology the dynamic height is usually called geopotential height.

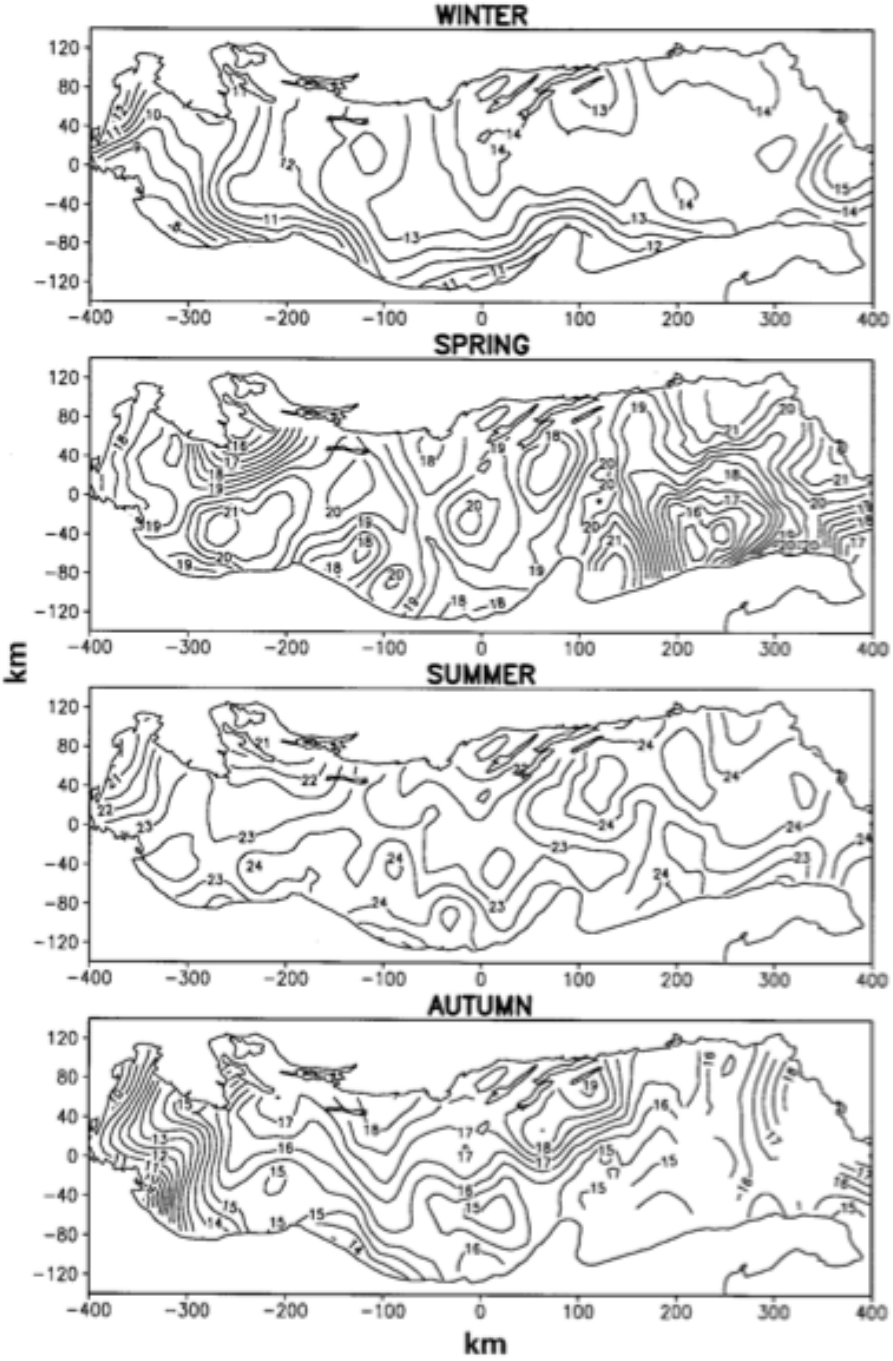


Figure 1.7: *Seasonal temperature (°C) maps at the surface. The contour interval is 0.5°C. The field is plotted for expected error less than 30%. From [5]*

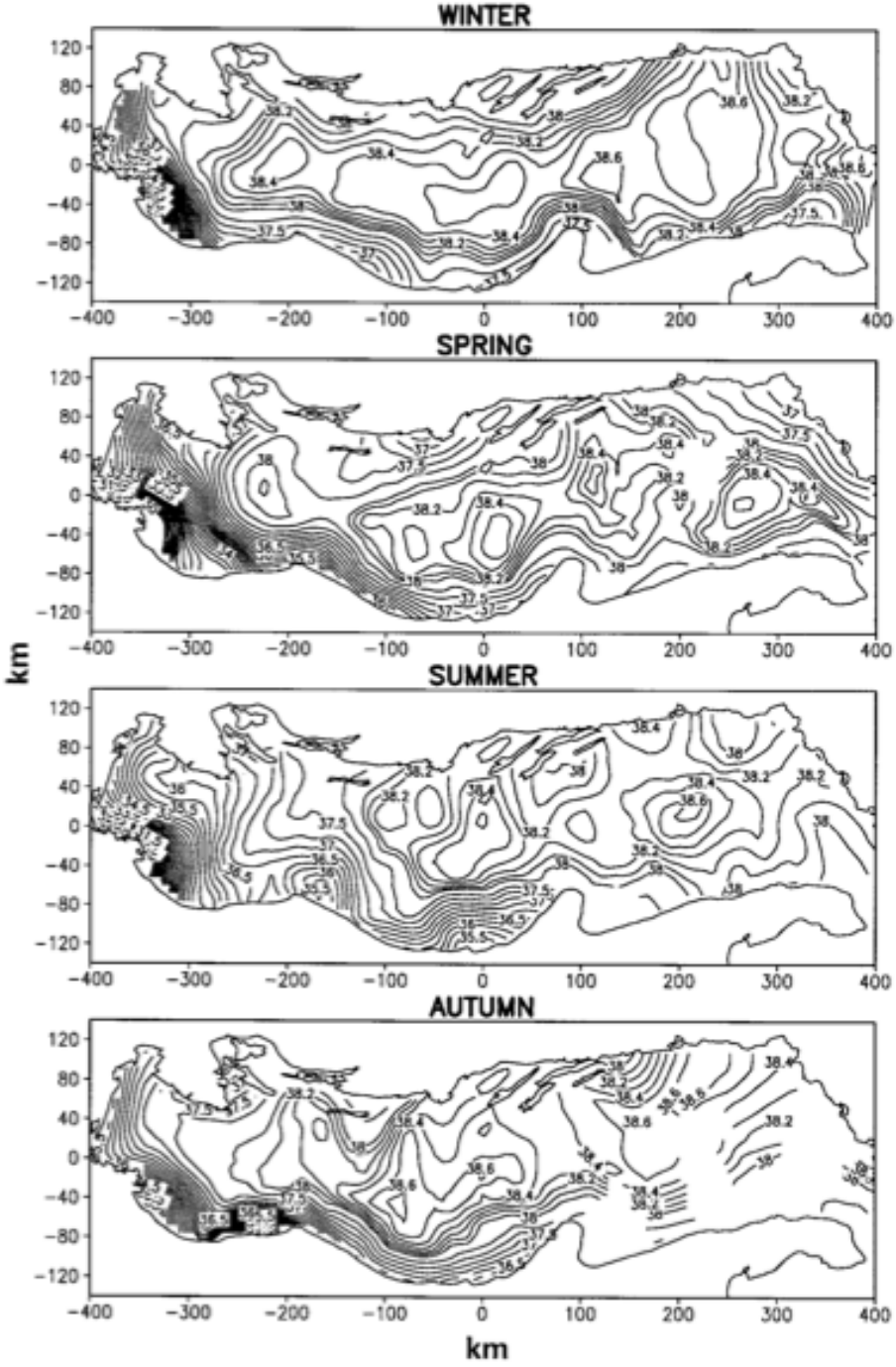


Figure 1.8: *Seasonal salinity (psu) maps at the surface. The contour interval is 0.25 psu for $S < 38$ psu and 0.1 psu for $S > 38$ psu. Contours for $S < 30$ psu are not shown. The field is plotted for expected error less than 30%. From [5]*

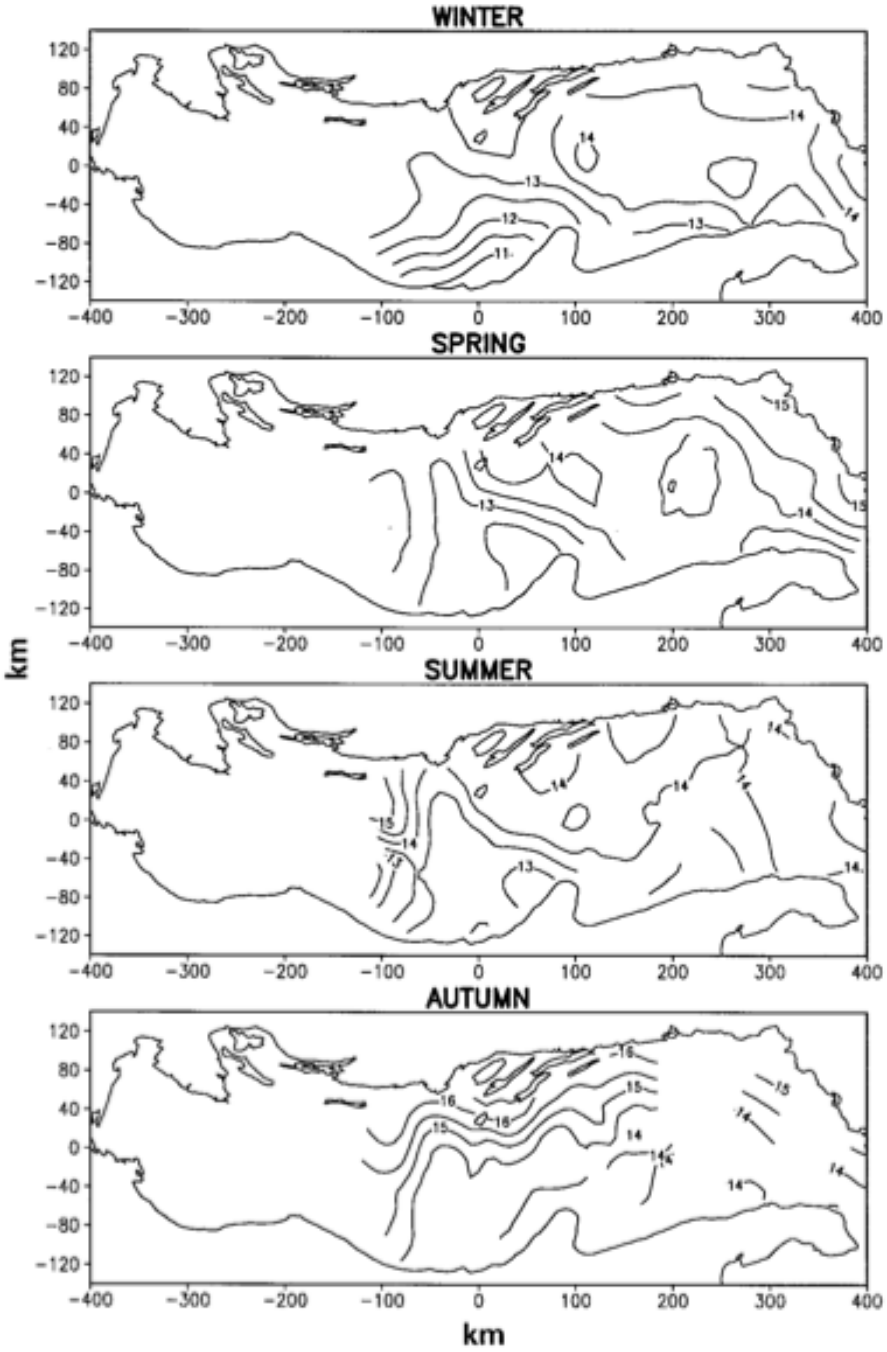


Figure 1.9: *Seasonal temperature (°C) maps at 100-m depth. The contour interval is 0.5°C. The field is plotted for expected error less than 30.%. From [5]*

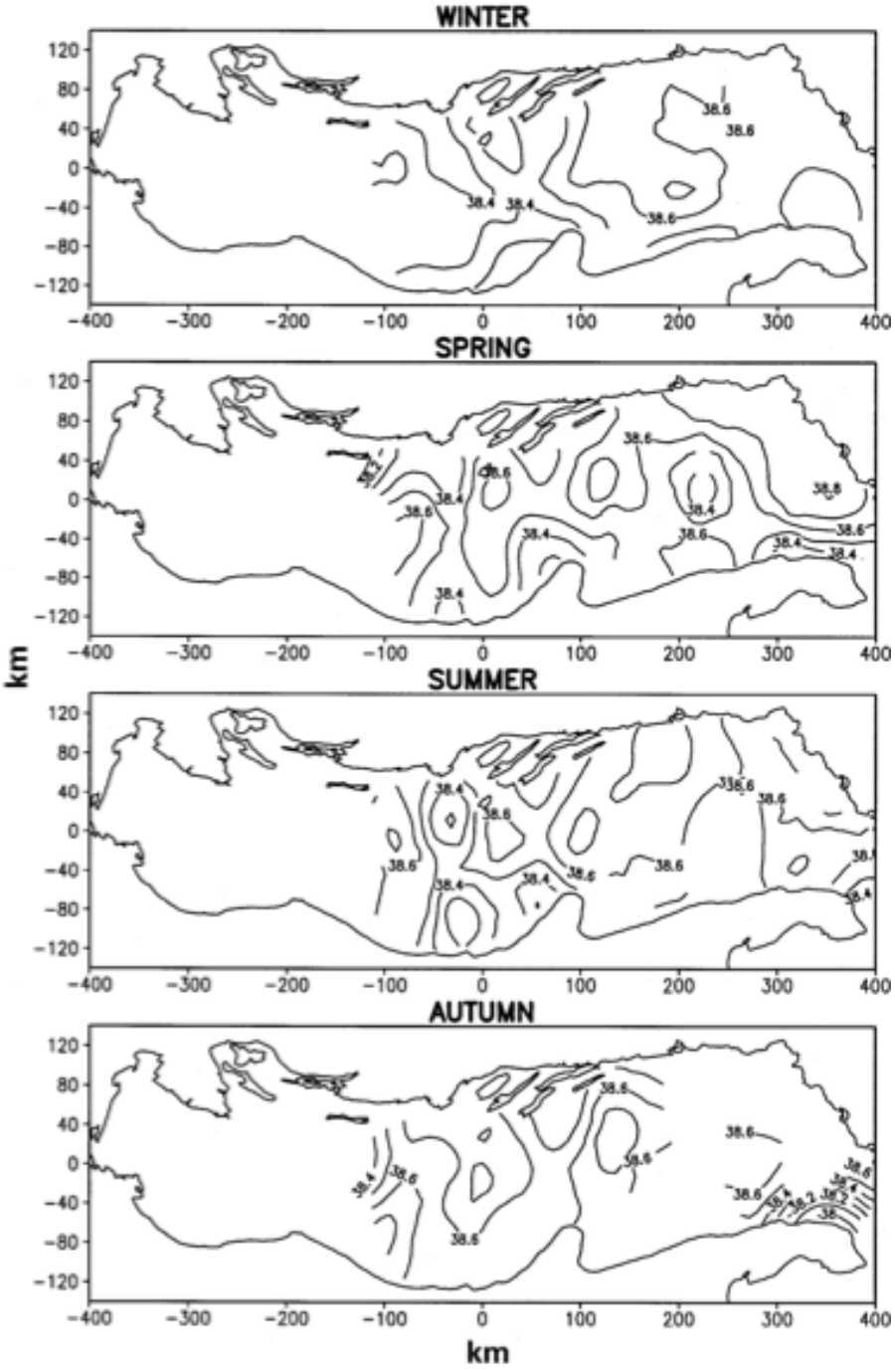


Figure 1.10: *Seasonal salinity (psu) maps at 100-m depth. The contour interval is 0.1 psu. The field is plotted for expected error less than 30%. From [5]*

The average circulation driven by the forcings so far described is cyclonic, with northward currents on the eastern coasts and southward currents on the western coast. A schematic representation is given in figure 1.12 and a seasonal averaged circulation is showed in figure 1.11.
The seasonal character is a very prominent feature of the circulation, which is clearly dominated by the Middle Adriatic and South Adriatic gyres (MAd and SAd gyres), a less prominent North Adriatic (NAd) gyre, an eastern South Adriatic current (E-SAd current), and a western South Adriatic current (W-SAd current). These four features of the circulation strengthen during summer, while they are weaker in spring and almost absent in winter, apart from the SAd gyre that is detectable in any season. During winter, in fact, the field is less energetic due to smoother termohaline gradients (figures 1.7, 1.8) and the circulation is dominated by a smooth flow from South to North along the longitudinal center of the basin. The general circulation is dominated by temperature and salinity compensation effects, which give no resulting density signal. We may speculate that the barotropic, wind-induced transport and circulation is probably a major component of the general circulation during winter.
Summer and autumn show evidence of the intensification of both the surface MAd and SAd gyres. During summer we see segments of W-MAd and W-SAd currents extending as far as Otranto. The W-MAd current persists throughout the autumn when the MAd and SAd gyres reach maximum amplitude and definition. At that time the E-SAd current is also well defined, occupying a large portion of the eastern regions of the basin.The autumn conditions are characterized by maximum spatial coherence in the general circulation structure. In fact, there are three cyclonic gyres (including the NAd gyre), a continuous western Adriatic boundary current, connected between the three sub-basins, and an intense SAd current. This season has maximum MLIW entrance and spreading from Otranto, a well-defined surface mixed layer, and maximum warming of the subsurface layers of the northern Adriatic. The aggregation of the general circulation into large-scale structures could be due both to the stabilization of the water column and to the structure of the external forcing of the circulation. It is believed that the Otranto inflow of MLIW could be a substantial part of that external forcing. The wind driving during autumn also consists of the southeasterly wind, (the Scirocco), which in turn could reinforce the inflow of water at Otranto.

In the second panel of figure 1.12 is shown the schematic pattern of the horizontal circulation at the upper thermocline.
As for the surface, the MAd and SAd gyres are defined from spring to autumn together with the E-SAd current, which is also related to the maximum advection of MLIW in the basin. The winter flow field is weaker, generally cyclonic in the south. In this season again the thermocline flow field seems to be dominated by a broad northward current along the longitudinal center of the basin, with a possible return flow along the shelf areas of the basin.

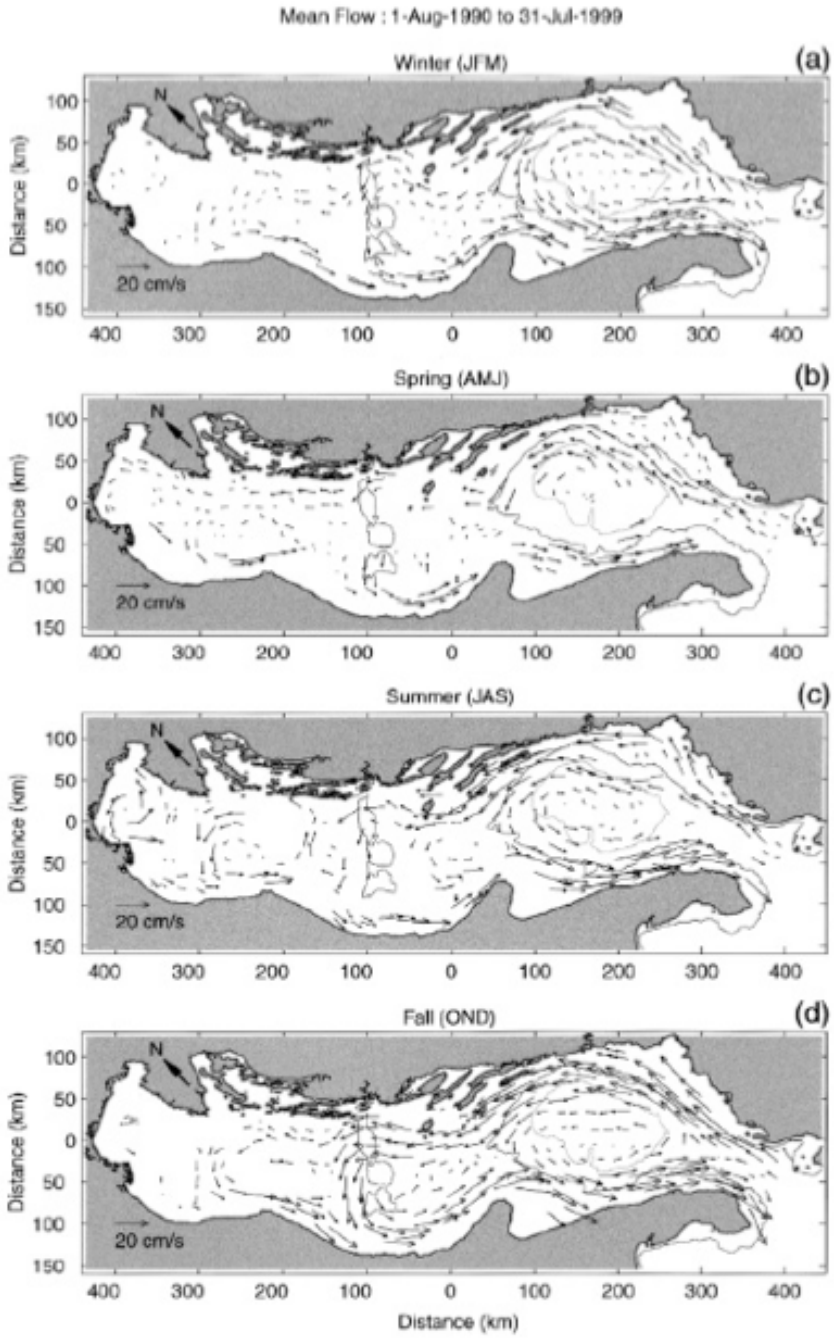


Figure 1.11: *Seasonal maps of average surface circulation as deduced from lagrangian observations from 1990 to 1999: (a) winter, (b) spring, (c) summer, (d) autumn. Adapted from [50]*

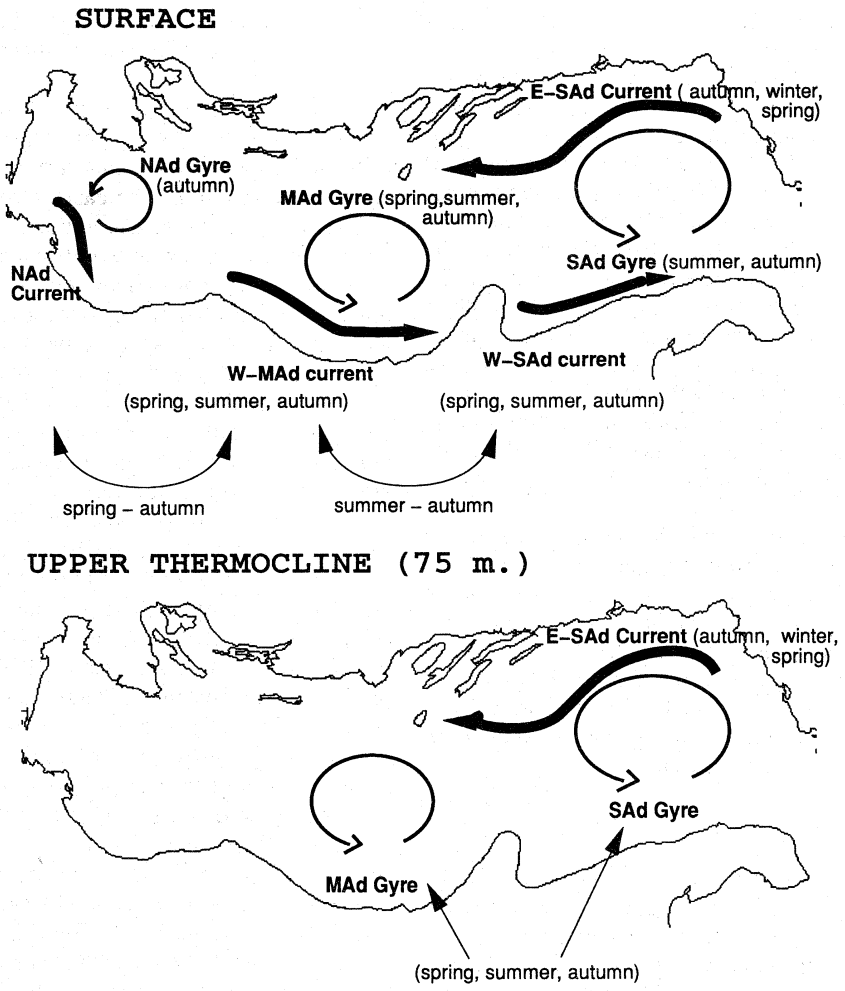


Figure 1.12: *Schematic reproduction of the circulation of the Adriatic Sea. The upper panel shows the surface circulation and the lower panel represent the circulation at a depth of 75 m. From [5]*

The general thermohaline circulation in the Adriatic Sea

Let us focus now on the thermohaline circulation of the basin and on the water exchange through the Otranto Channel with the Mediterranean Sea. This circulation is completely driven by density gradients created by surface heat and freshwater fluxes.
In the Adriatic Sea we already seen that are present two regimes of thermohaline circulation: an *estuarine* circulation and an *antiestuarine* one.
The first consists in an outflow of surface low-density water, that must be balanced by an inflow of more salty, sub-surface water. The fresh-water outflow is due to the southward WACC, freshened by the river water inflow in the basin (especially the Po in the northern Adriatic) that, if not balanced, would produce a considerable sea level rise, as previously explained (see equation (1.1)).
The antiestuarine circulation is due to the flow of dense deep water (NAdDW, MAdDW and SAdDW) flowing south along the bottom of the basin. The resulting outflow current is called Deep Water Outflow Current (DWOC).
The balancing inflow is composed of two currents, both flowing along the eastern side of the Adriatic Sea and thus forming the E-SAd current: the Ionian Surface Water (ISW) flowing near the surface, and the Modified Levantine Intermediate Water (MLIW) flowing at the intermediate depth of about 200 m.

The thermohaline circulation is responsible for keeping the balance of water masses in the Adriatic Sea and is very sensitive to the climatic variations and seasonal oscillations. A schematic view of the thermohaline circulation in the Adriatic Sea is given in figure 1.13.

1.1.3 Perturbations on the general circulation: winds and tides

After a description of the general features of the climatological circulation in the Adriatic Sea we give a brief outlook to the principal high frequency perturbations, namely the winds and the tides.

The effects of the winds on the general circulation

The winds are responsible for perturbations that have a smaller spatial and time scales with respect to the general circulation so far described. These dynamics are generated by the sum of three forces: the stress of wind on the surface, the pressure gradient that is generated along the entire water column and the friction that is exerted at the bottom.
As already discussed, the principal winds that must be considered in studying the dynamics of the Adriatic Sea are Bora and Scirocco.
The effect of the winds are particulary important in the northern sub basin during winter, when the absence of vertical stratification weaken the baroclinic component of the circulation. In particular the Bora episodes are capable to generate temporary cyclonic and anticyclonic eddies and, sometimes, even to reverse the coastal currents along the eastern margins of the basin. This effect is connected to upwelling and thus weakening of the cyclonic circulation, with the development of counter currents like the Istrian Coastal

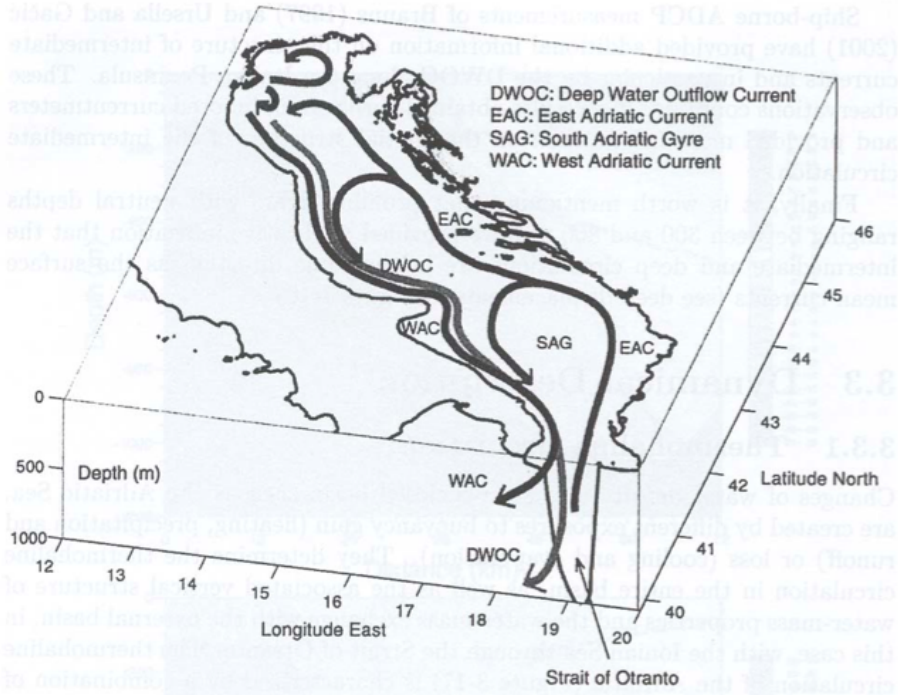


Figure 1.13: *Schematic representation of the thermohaline circulation in the Adriatic Sea. From [50]*

Current (ICC) simulated, for example, by Oddo et al in [42]. This phenomenon also leads to a weakening of the MLIW inflow from south and sometimes it is even able to block it. However the numerical study (and thus the prediction) of the response of the basin to the Bora events is complicated by their temporal variability and the horizontal structure of the wind, highly affected by the interaction with the orography.
On the other hand, the Scirocco, flowing from South to North along the longitudinal extension of the basin, is responsible for an elevation of the sea level that can reach several centimeters, especially in the northern Adriatic. The currents associated to this Scirocco events are directed in the direction of the wind at the sea surface and along the coasts, but sub-surface counter current develop along the central axis of the basin. The northward currents can be strong enough to induce a reversal in the WACC along the Italian coasts. The adjustment after Scirocco events induce seiches and oscillations in the currents that expire after a few days.

There are some evidences that the winds can also be of fundamental importance as forcings of the general circulation.
Prognostic and diagnostic numerical simulation carried by Zavatarelli et al ([62], [63]), suggested the fundamental impact of the winds during winter. In this season, in fact, the absence of stratification due to compensation of temperature and salinity on the density field (especially in the northern adriatic) could weaken the WACC. Instead it was shown that when the baroclinic component (strong in summer and autumn) weakens in winter,

the wind exerts a strong barotropic component to the current.
The same study carried on in the other season clearly shown the importance of the wind on the general circulation in the Adriatic Sea, especially on the seasonal fluctuation in the WACC.

The tides in the Adriatic Sea

The tides can be defined as the periodic fluctuation of the sea level forced by combined gravitational and centrifugal forces of the Earth-Moon and Earth-Sun systems.
In principle, fluids respond to all gravitational fields exerted on them, including that of other planets and celestial bodies of the solar system, but in practice, only the Sun and the Moon induce detectable sea level excursions.
The net force on an oceanic particle is small but can nevertheless give rise to surface elevations which range from near zero to the order of ten meters depending on geographical location.
The total tidal forcing, that can be written in a potential form, is a superposition of many components, each characterized by its period and amplitude. The principal components can be classified by their period in three different categories: semidiurnal, diurnal and long period. The most important components are the principal lunar semidiurnal M2, the lunisolar diurnal K1 and the principal solar semidiurnal S2. A complete description of the tidal potential and of the components can be found in [2], [37] or in any other fundamental text on physical oceanography.

The coastal and shallow water tides are significantly different from the deep sea tides by changes in depth and bottom friction, whose lead to amplitude and phase change of considerable magnitude in the transition zone form the shallow water area to the deep sea are.
In particular, in a semi-enclosed shallow water system in the northern hemisphere (like the Adriatic Sea), a simple solution for the tidal seiche can be found to be a counter-clock wise oscillation propagating from the southern edge toward the north along the eastern coasts and again toward south along the western coasts (see figure 1.14). It is seen that during this cycle the tidal wave elevation appears to rotate counterclockwise about a nodal point near the center of the embayment which is a point of zero elevation and is called the *amphidromic point*. The currents also rotate counter-clockwise about the amphidromic point and have stronger along axis component than cross-bay flow in the case of a narrow basin.
Thus the tidal response of the Adriatic Sea is not a direct consequence of the gravitational attraction exerted by the Sun and Moon, but can be rather explained in terms of a perturbation that is forced by a fluctuation of appropriate frequency and amplitude in the sea level and velocity field at the Strait of Otranto.
This is exactly the starting point of many studies on the Adriatic tides, like the one presented in [24]. In this and other studies has been demonstrated that the response to the astronomical tides of the Adriatic Sea, is another important feature of the general

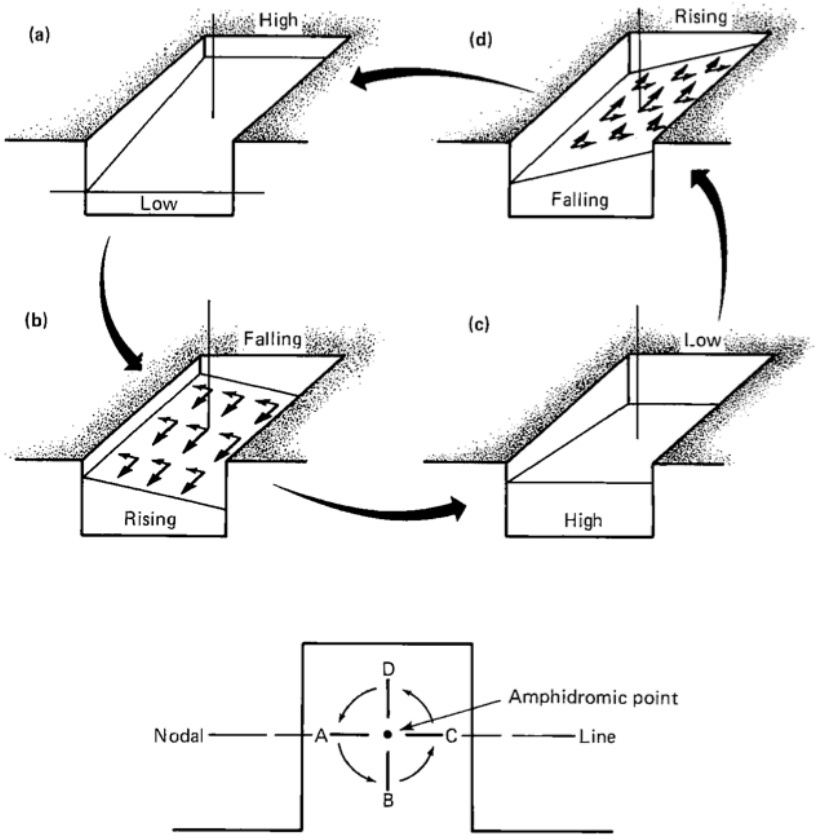


Figure 1.14: *Schematic of rise and fall of tides in an embayment in the Northern Hemisphere. Tidal elevations appear to rotate counterclockwise about the amphidromic point. From [2].*

circulation variability at high frequencies of the basin.

The Adriatic, together with the Strait of Sicily, represents the only area within the Mediterranean where tides have a range of up to more than a meter. These amplitudes occur in the northernmost Adriatic region (the Gulf of Trieste) where the amplitudes of the most energetic frequencies (M2 and K1) reach approximately 27 and 18 cm respectively. As is known from many other studies on barotropic Adriatic tides, the diurnal component frequencies present weak amplitudes in the southern part of the basin and become enhanced as they move northwards, developing amplitude isolines approximately perpendicular to the longitudinal axis of the basin. The semidiurnal frequencies show an amplitude enhancement in the northern part of the basin with the formation of an amphidromic node located in the centre of the basin between Ancona and Zadar.
Both tidal currents and elevation tend to propagate northwestward and counter-clockwise respectively as expected from out previous discussion (see figure 1.14), but this pattern is more evident for the semi-diurnal than for the diurnal components.

Tides in a semi-enclosed shallow water basin like the Adriatic Sea can deeply influence the mixing of temperature and salinity in the water column and induce oscillation in the thermocline. This kind of perturbation were clearly found in the Adriatic although the most remarkable impact, however, was found to be on horizontal advection more than vertical diffusion. In fact, it is demonstrated in [24] that the fundamental structure of the circulation in the basin are deeply influenced by the tides, with an enhanced cyclonicity in the entire basin. The WACC and the E-SAd seem to be much constrained to the coasts and the shape and intensity of the gyres are much more defined by the introduction of the tides in the model used in the experiment.
As expected, the effects of the tides on the circulation appeared to be more evident during periods of limited wind stress.

1.2 Estimating the circulation of the Adriatic Sea

Estimation of both the circulation and of the physical scalar quantities (such as salinity and temperature) is a crucial task of physical oceanography. This is a hard task to accomplish because the behavior of geophysical fluids is non-linear: ocean states that are initially similar often differentiate rapidly over time. As first pointed out by Edward N. Lorenz in his celebrated 1963 paper [40], for such systems, unless the initial state is perfectly known, future states are bound to become unpredictable after a finite period of time, even if computed from exactly known deterministic equations. This is because any initial error with respect to the true state of the system, no matter how small, will necessarily grow at some exponential rate and eventually lead to a predicted state macroscopically different from the true state. In the case of the atmosphere, it is for instance estimated that any initial error will roughly double within 1.5 days, which, given current meteorological data quality and assimilation strategies, translates into the familiar possibility of accurately forecasting the weather up to three days in advance, with some confidence up to a week, but hardly beyond. For the oceans the predictability limit is often considered to be 10 days. The real value actually depends on the process under study, on the system in use and is a function of the initial uncertainties but it is quite of general validity the concept that after 10 days any prediction of the physical state of the ocean is no longer reliable.

The estimation of the future state of the ocean is one of the task of the discipline called *operational oceanography* that can be defined as the activity of systematic and long-term routine measurements of the seas, oceans and atmosphere, and their rapid interpretation and dissemination. These informations must be delivered on a routine basis and with sufficient speed, such that users can act on the information and make decisions before the relevant conditions have changed significantly, or become unpredictable. Thus a compromise has always to be found between speed and quality of collection, interpretation and analysis of data.
The field of interest of operational oceanography is not only the physics of seas and oceans,

but oceanography in a broader sense, including also the biological, chemical and geological aspects. Thus there are a variety of scientific and practical applications for which a realistic estimation of the physical fields is relevant: climate studies, biogeochemical cycles, ecosystem dynamics, sustainable fisheries management, tactical decisions in environmental risk, search and rescue operations are just some of the areas where the marine sciences find their practical application.
Important products derived from operational oceanography are not only *forecasts* of the future state of the ocean as far ahead as possible, but also the so called *nowcasts* and *hindcasts*. The former are descriptions of the present state of the sea, while the seconds are analysis of the past state. The knowledge of the past and present state are not easier to obtain than the future state, because of the non-linear behavior of the system and of an uneven availability of the data, both in time and in space: is not easy to asses the salinity in a certain point of the sea if no one has actually observed it.
Operational science has been applied to meteorology since the '30s while operational oceanography developed between the past and the present centuries. The reason for this delay lies in the greater difficulty in collecting data on a routinely basis in the ocean compared to the atmosphere.

In this section we aim to describe the three components of an operational network. The first two are the data that has to be collected and the numerical model that must be correctly reproduce the dynamics of the system under study. The third component is a way of merge the data into the model in order to obtain the model representation which is most consistent with the observations: the *data assimilation* technique.

1.2.1 Data

There are two ways of collecting data from the oceanic environment: *satellite* data and *in-situ* data. These two type of observation are complementary and both are needed to gain a better insight in the ocean dynamics.

In-situ data are local observations of the properties of the sea or ocean, providing not just information about the physical, biological, chemical state at the surface of the sea but also below. Since many physical quantities are needed to fully describe the present state of the sea, many techniques of in-situ observation have been developed. A full description of all of these techniques is not the subject of this work as entire books exist on the topic: the interested reader may refer to the first chapters of [18]. Here we limit our discussion to the properties and techniques that are relevant for the present work.

Some of the key physical-chemical informations that can be obtained with in-situ techniques are temperature and salinity vertical profiles of the water column, lagrangian or eulerian velocity components and sea surface elevation from coastal tide gauges (which actually requires a satellite observation to distinguish the real sea-level contribution to the datum from the land movements).
Salinity and temperature profiles are commonly obtained by means of XBT, CTD and

bottle casts.

The XBT (eXpendable Bathy- Thermograph) is a small device for obtaining a record of temperature as a function of depth from a moving ship. Temperature is measured with a thermistor within an expendable (from here comes the first part of the name), weighted casing. Depth is determined by a priori knowledge of the rate at which the casing sinks and the time of each recorded data value. A pair of fine copper wires which pay out from both a spool retained on the ship and one dropped with the instrument, provide a data transfer line to the ship for shipboard recording. Eventually, the wire runs out and breaks, and the XBT sinks to the ocean floor. Since the deployment of an XBT does not require the ship to slow down or otherwise interfere with normal operations, XBT's are often deployed from vessels of opportunity, such as cargo ships or ferries, rather than a dedicated research ship where a CTD would normally be used in preference.

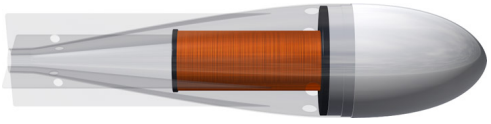


Figure 1.15: *The XBT device*

Sources of errors in an XBT cast lies both in the estimation of depth and in the temperature measurement. The XBT depth errors are known to be functions of depth itself since they depend on an incorrect fall-rate equation but it can be shown ([18]) that with an appropriate calibration it is possible to find an equation such that uncertainty in depth can be neglected compared to the error on temperature measurement. The latter can also be evaluated from the calibration of the instrument, but an error of $0.1 - 0.2$ °C seems to be a quite general characteristic of these instruments.
Since the XBT is expendable it is a good habit to deploy it in deep-sea areas. In the Adriatic Sea measurements with this instrument are made only in the southern sub-basin.

A CTD (Conductivity, Temperature and Depth) profiler is an essential instrument of physical and biological oceanographic field research. The instrument is a cluster of sensors that measure continuous data of depth, salinity, and temperature. Sensors commonly scan at a rate of 24 Hz. Depth measurements are derived from measurement of hydrostatic pressure which can be easily translated in depth by means of well known empirical equations (for example, see [20]) , salinity is measured from electrical conductivity and temperature by means of a high performance resistance thermometer.
Based on technical reports, modern CTDs are accurate to approximately 0.002°C in temperature, 0.005 psu in salinity and $<0.5\%$ of full-scale pressure in depth.
Sensors are arranged inside a metal housing, the metal used for the housing determining the depth to which the CTD can be lowered. CTD devices are usually mounted on a carousel together with twelve to twenty vertically mounted water samplers, the *bottles*. The carousel is also called a rosette (see figure 1.16). Bottles are 10-liters containers that are used to collect samples of sea water at various depths which are then analyzed in the laboratory at the end of the casting procedure. The closing of the sampling bottles is triggered from an onboard computer, thus water samples can be obtained at specific depths, from which biological and chemical parameters can be measured. Of course, the hydrographic bottle should remain well sealed on the upcast in order to preserve the char-

acteristics of the water collected in the container.

Precision on the laboratory measure on salinity can reach $10^{-3} - 10^{-4}$psu, thus often analyses of the water samples are used in the calibration of the CTD sensors.
Other sensors may be added to the rosette cluster, including some that measure chemical or biological parameters, such as dissolved oxygen and chlorophyll fluorescence, the latter being an indication of the concentration of microscopic photosynthetic organisms (phytoplankton) contained in the water.

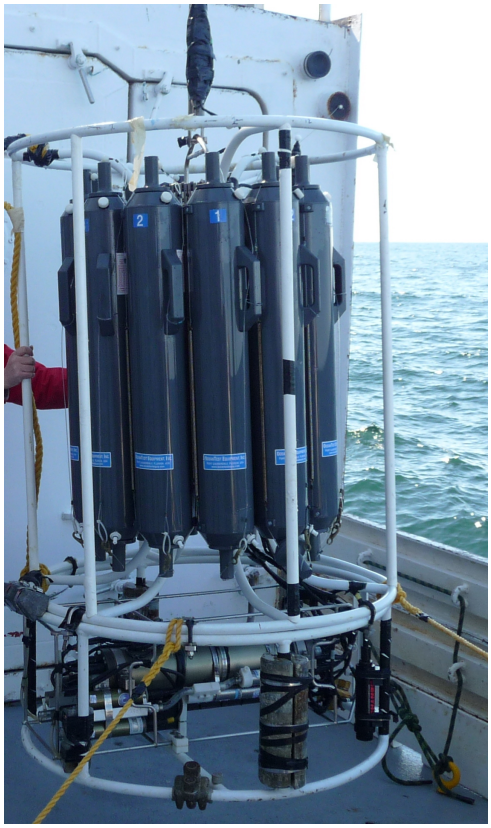


Figure 1.16: *CTD unit with rosette*

One of the most important quantity that can be measured is the sea-level, which is, in fact, one of the oldest form of oceanic observation. As explained later, the satellite measurement of the sea level is crucial in understanding the ocean dynamics. On the other hand, the coastal measurement of in-situ sea-level changes by tide-gauges is essential to study the wind-generated waves and storm surges. In addition to measuring the vertical movement of the coastal land mass, long-term sea-level observations reflect variations in large-scale ocean circulation, surface wind stress, and oceanic volume. Because they provide a global-scale integrated measure of oceanic variability, long-term ($>$ 50 years) sea-level records from the global tide-gauge network provide some of the best information available on global climate change.
Since coastal stations actually measure the movement of the ocean relative to the land, land-based sea-level measurements are referred to as relative sea-level (RSL) measurements. To get the actual sea level, the changes in gauges height relative to the land can be taken into account and deleted from the raw data. This is done by properly tying the instrument to a bench mark height and running the level back to the nearest available geodetic datum or directly with a GPS observation of the station elevation (see figure 1.17).
First tide gauges were floating gauges, measuring the sea-level changes by the observation of the movements of a float. Over time, various and more accurate systems have been developed such as pressure and acoustic gauges. Depending on the instrument's age and on the principle of operation, the uncertainty in the measurement of RLS can vary from some centimeters to a few millimeters.

The usefulness of in-situ data is limited by poor space/time coverage. The geographic distribution of monitoring stations is far from uniform in terms of the spacing between

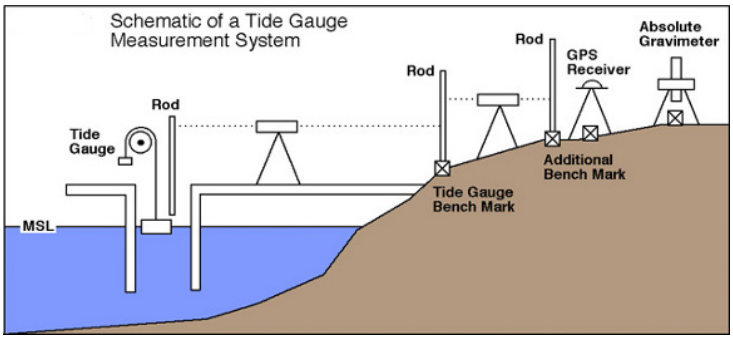


Figure 1.17: *Schematic of tide-gauge station with the gauge, network of benchmarks and advanced geodetic link.*

them. Thus, while the data series collected at each station may themselves consist of evenly and densely spaced measurements in time, the space intervals between stations will be highly irregular. Platforms of opportunity (usually merchant or turist ships) produce uniquely irregular sets of measurements. Most merchant ships repeat the same course with minor adjustments for local weather conditions and season. Added to the seasonal track changes is the nature of the daily sampling procedure. Usually the ship takes measurements at some specified time interval which, due to variations in ship track, ship speed and weather conditions, may be at very different positions from sailing to sailing. Thus, the ship of opportunity data are irregular both in space and time.

Satellite observations, on the other hand, provide global measurements of key surface parameters such as Sea Surface Temperature (SST), sea level, winds, significant wave hight, ocean color (which can be related to bio-chemical properties such as the surface chlorophyll abundance) and, thanks to the newly launched ESA satellite SMOS (figure 1.18), Sea Surface Salinity (SSS).
The principal observations that are used in operational oceanography, principally due to the numbers of satellites and the experience of the oceanographic community, are SLA and SST.
The most important of these techniques (both historically and scientifically) is satellite altimetry since it has unique capabilities for providing global synoptic view of the ocean circulation. It provides measurements of sea surface topography which is an integral of the ocean interior, thus this observations are a strong constrain to the three-dimensional ocean circulation estimation.
The basic principle of satellite altimetry is simple (the reader is referred to [2]): the altimeter measures the distance between the satellite and the ocean surface by determining the time taken by a radar pulse emitted by the satellite and reflected by the sea surface back to space. However, this is not the only measurement made in the process, and a lot of other informations are contained in this observation. First of all the two way travel time is just a relative observation. We have to know the position of the satellite with respect to a reference geometrical surface close to the Earth surface: the reference ellipsoid. Combin-

Figure 1.18: *The new ESA satellite SMOS for measuring sea surface salinity and soil moisture*

ing this two measurements yields an estimation of the sea level relative to the reference ellipsoid. This estimation comprises the geoid (an equipotential of the gravity field that would coincide with the surface of an unperturbed, motionless ocean, see [56]) and the ocean *dynamic topography*, the parameter of interest here. The dynamic topography η is in turn composed of two parts a *mean dynamic topography* $< \eta >$, which is the long term topography due to the sum of the geoid and the climatological ocean dynamics and winds, and a variable part $\eta' = \eta - < \eta >$, easily extracted since no previous knowledge of the geoid height is needed. This quantity is usually called *Sea Level Anomaly*, or SLA.
Altimeter measurements are affected by a large number of errors, mainly due to errors in the determination of the satellite orbit, to propagation of the radar signal in the atmosphere and to atmospheric effects on the sea surface height which are long wavelength phenomena (such the inverse barometer effect), not wanted in our SLA measurement.
A reliable SLA observation requires the uncertainty in the orbit determination and on the sea-satellite distance to be less than 2 cm each.

Great advantages of satellite measurements (not only satellite altimetry) are, first of all, the greatest spatial coverage with respect to the conventional sea-level measurement system. Tide-gauges, in fact, are limited by the need of a fixed platform installation so that the observation of the sea-level changes with these instruments can be only performed from coasts or islands where they can be referenced to the land boundary (see previous discussion and figure 1.17).
An additional advantage is that measurements can be acquired and processes in *near-real time* (2 days), which is a fundamental prerequisite for operational oceanography. As already discussed, the processing of satellite data requires a certain number of correction to be performed on the raw data, such as the correction for atmospheric propagation,

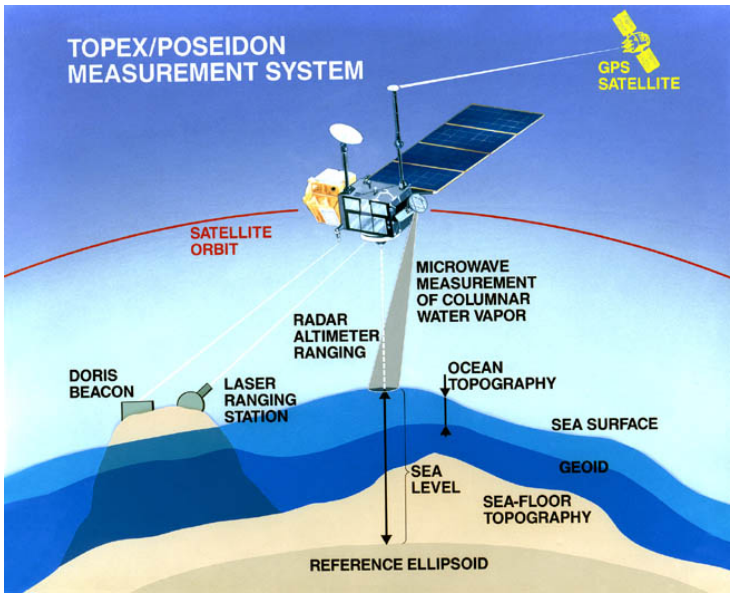


Figure 1.19: *Satellite altimetry measuring principle*

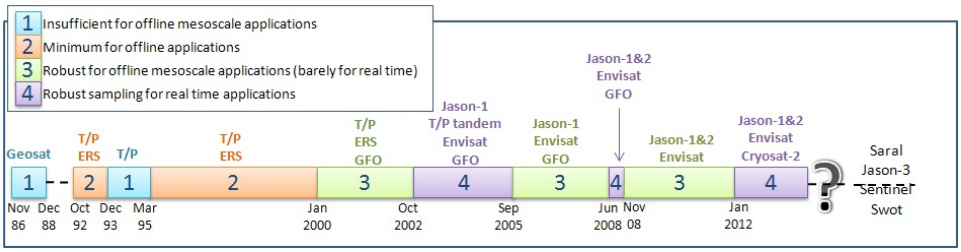


Figure 1.20: *Storyline of satellites available in time related to applications*

the orbit computation, the subtraction of the geoid (if needed), the inverse-barometer correction and others. Clearly this computation requests time and a quick analysis for operational purposes means that approximate correction has to be performed. In particular the orbit computation requires environmental parameters which cannot be obtained in real time. On the contrary, a full time analysis and an accurate, longer computation of the correction leads to a more precise product, which is called the *delayed time* product. There is always a compromise to reach between spatial coverage and temporal resolution: a satellite that revisits the same spot frequently covers fewer points than a satellite with a longer orbital cycle. This point has always been a primary source of discussion among the scientific community, working constantly in designing the best orbit depending on the phenomena of interest. The solution comes from the availability of many satellites that can operate together, combining the force points of the orbit of the single satellites.
Topex/Poseidon-ERS and Jason-Envisat are fine examples of how altimetry satellites can operate together. Topex/Poseidon and Jason-1 follow a repeat cycle of ten days designed to monitor ocean variations, so they pass over the same points fairly frequently but their ground tracks are some 315 kilometers apart at the equator-more than the average span

of an ocean eddy. On the other hand, ERS-2 and Envisat only revisit the same point on the globe every 35 days but the maximum distance between two tracks at the equator is just 80 kilometers.
At least two altimetry satellites are required to map the ocean and monitor its movements precisely, particularly at scales of 100 to 300 kilometres (mesoscale). With four altimetry satellites available (Jason-1+Jason-2+Envisat+Crysaot-2 from 2012, or Jason-1, Envisat or ERS-2, Topex/Poseidon and GFO between 2002-2005), the resolution of sea surface height measurements is greatly enhanced. At least three satellites are needed to observe eddies and mesoscale phenomenon.
Besides, multiple satellites observing the same point allows for cross-calibration.
A schematic view of the satellites for SLA observations available from the first satellites to present days is given in figure 1.20.

A great advantage of the SLA observations over SST is that the latter is highly prone to weather conditions over the seas and oceans. A traditional SST observation, in fact, requires the evaluation of the total energy reaching the satellite in the infrared range of the electromagnetic spectrum. This radiation is reflected by clouds, creating a cool bias in satellite-derived SSTs within cloudy areas. However, passive microwave techniques can accurately measure SST and "see" through clouds.
Besides in infrared remote sensing methodology the radiation emanates from the top ”skin” of the ocean, approximately the top 0.01 mm or less, which may not represent the bulk temperature of the upper meter of ocean due primarily to effects of solar surface heating during the daytime, reflected radiation, as well as sensible heat loss and surface evaporation. All these factors make somewhat difficult to compare satellite data to measurements from buoys or shipboard methods, complicating ground truth efforts.
On the other hand the altimetry observation is carried using microwaves and radio waves that are not influenced by clouds, but are disturbed by the water content of the entire atmosphere. Despite this fact (that require an adequate correction in the data processing phase) the reflected radiation is not stopped by the clouds and the measurement can be performed regardless of weather conditions.
Beside this factors the SLA is the satellite observation that is most used in operational oceanography because of the greater number of dedicated satellites than any other oceanographic measurement. This provides the spatial and temporal coverage that is needed for operational purposes.

There is an obvious complementarity between satellite and in-situ data since they have different contents (namely, vertical structure from in-situ data, surface information from satellites) and very different space-time sampling. In-situ data are also required for calibration and long-term validation of satellite data.

1.2.2 Numerical models

A numerical model can be defined as a set of mathematical instructions that provide an approximate representation of a physical system and of its temporal evolution. Thus an oceanographic model applicable to the Adriatic sea (or one part of it) must contain a representation of the physical laws that dominate the dynamics within the basin, under well defined forcing. Namely the governing equations of an oceanographic model are some approximation, but still non-linear versions, of the Navier-Stokes equation and diffusion equations for temperature (from Fourier's law, or heat equation) and salinity (from Fick's law). Other equations can of course be implemented for other quantities of interests, like other chemical and biological tracers.

Numerical modeling has been applied in the Adriatic Sea since the '70s. The first models, however, did not implement the whole basin due to the high computational restrictions of times. So the spatial domains were quite limited and the models were used essentially for storm surges prediction. Particular attention was posed in the simulation of the response of the sea surface in the northern sub-basin to the winds, like the high-water events in the gulf of Venice.
Subsequently, numerical models of the entire Mediterranean Sea were implemented, by supposing that he Adriatic Sea was one of the components but had not a dedicated system.

An important milestone of the numerical modeling in the oceans and seas came in the late '70s with the development of the Princeton Ocean Model (see the webpage `http://www.aos.princeton.edu/WWWPUBLIC/htdocs.pom/index.html`), or simply the POM, thanks to the work of Blumberg and Mellor. POM is a free-surface, three-dimensional finite-difference numerical model based on the primitive equations with Boussinesq and hydrostatic approximations. One of the peculiarity of the model was the fact that the vertical resolution is made variable by discretizing the depth not with z-levels, but with the so called sigma-layers (σ-layers).

In 1995 Zavatarelli and Mellor applied the POM numerical model to the Mediterranean Sea. This particular implementation had 16 vertical levels and a variable resolution horizontal grid. The simulations pointed out that the Adriatic is more influenced by flows at the air-sea interface rather than by winds, more important in the rest of the Mediterranean. The model was also able to simulate the formation of deep waters in the northern and southern sub-basins.
Subsequently, other, more refined numerical models were implemented in the Mediterranean Sea. In the late '90s, in the framework of the MFS (Mediterranean Forecasting System) Pilot Project (MFSPP, see the website `http://mfstep.bo.ingv.it/WP8/model_des.htm`), Pinardi et al. implemented a three dimensional primitive equation ocean general circulation model (OGCM) based on the GFDL-MOM (Geophysical Fluid Dynamics Laboratory-Modular Ocean Model) designed by Bryan and Cox (see [8] and [10]). In 1990, using ideas of modular programming to increase model flexibility, GFDL released MOM as a successor to the Cox version of the code.

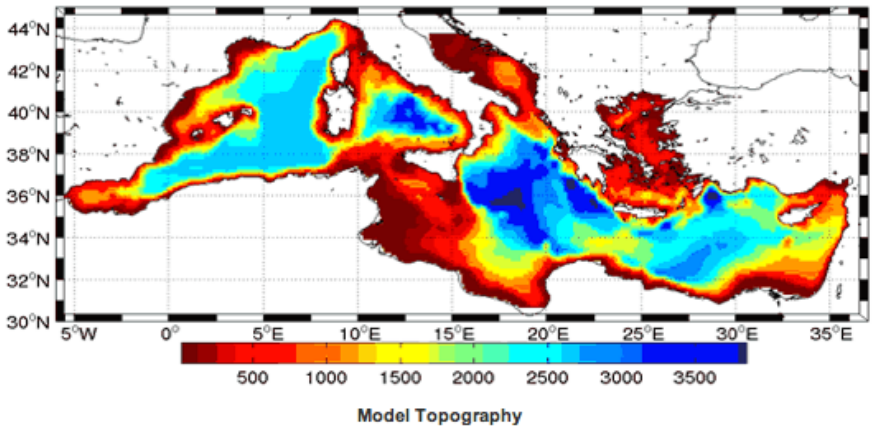


Figure 1.21: *Model domain and bottom topography of the MOM 1.1 model of the Mediterranean Sea. Note the absence of the northern Adriatic area.*

The version of the model used in MFSPP is the MOM1.1 version implemented in the Mediterranean at 1/8° x 1/8° degrees of horizontal resolution (correspondnig to about 12.5 km) and 31 levels in vertical. Results from this model implementation are available in [48] and [13].
In the same project two other models were implemented, specific for the Adriatic Sea:

- The Adriatic Intermediate Model (AIM), a model for the Adriatic general circulation with a horizontal resolution of about 5 km (1/20°) and 21 sigma layer in the vertical. The model extends from the entire northern Adriatic to the Strait of Otranto, where the open boundary conditions are given by the coarser resolution model of the Mediterranean Sea through a technique called *nesting*.
- The North ADriatic Shelf Model (NASM), a model with horizontal resolution of 1.5 km (1/37°) and 11 sigma-layers. The model domain comprises the entire northern sub-basin and is nested with the AIM model at the latitudes of the Istrian Peninsula.

So the three model system has been developed so that a finer resolution model take its open boundary condition from the immediately higher resolution model: the nesting are defined between the OGCM and the AIM and between the AIM and the NASM.
In the Mediterranean ocean model itself must be provide a lateral boundary condition must be added to take into account of the Gibraltar strait, where the water exchange with the Atlantic Ocean has to be defined. The OGCM of the MSFPP was not nested in an Atlantic ocean model, but the model state at the boundary is relaxed to annual mean climatology.
It is also important to note (from figure 1.21) that in the OGCM model the northern Adriatic was not implemented because it was considered a coastal zone to be modeled by the dedicated NASM model.

The OGCM model just described was used for operational purposes at the GNOO (Gruppo Nazionale di Oceanografia Operativa, a section of INGV) till 2004, when it has been replaced by the model OPA 8.1 developed in the framework of the MFSTEP (MFS

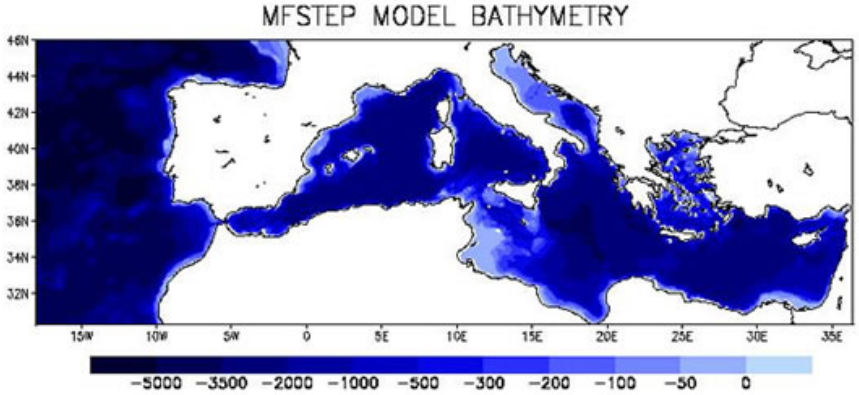


Figure 1.22: *Model domain and bottom topography of the OPA 8.1 model of the Mediterranean Sea.*

Toward Environmental Prediction) project. The model was developed by Institut Pierre Simon Laplace, Laboratoire d'Oceanographie DYnamic et de Climatologie, Paris. A detailed description of the code can be found in [34]. The model is a primitive equation model, the Navier-Stokes equations are used with the approximation of thin-shell, Boussinesq, hydrostatic and incompressible fluid. In the frame of MFSTEP Project the code has been implemented in the Mediterranean at 1/16° x 1/16° horizontal resolution and 72 unevenly spaced vertical levels. The model domain and the bathymetry are shown in figure 1.22. The atlantic box has been enlarged and the northern Adriatic had been included in the model. In the first versions of the model the Atlantic box was still not nested within a coarser resolution model of the Atlantic Ocean circulation, even if the larger atlantic box provided a much more realistic simulation of the fluxes at the Strait of Gibraltar. This version of the model is forced by the ECMWF (European Centre fro Medium range Weather Forecast) near real time forecasts for the atmospheric forcing and has been operatively used for Mediterranean forecasts from 1999 to 2004. The updated version of the model, the OPA 9.0 is currently used for operational purposes at the GNOO, as a component of the MFS (Mediterranean Forecasting System, `http://gnoo.bo.ingv.it/mfs/`). The Atlantic box is nested within the monthly mean climatological fields computed from the daily output of the 1/4 x 1/4 degrees global model developed by the Mercator Ocean Oceanography Centre ([17]). The nesting technique allows a better representation of the inflowing Atlantic water proprieties and therefore implies an improvement in the representation of seasonal oscillations of the Mediterranean volume ([43]). A dedicated sub-model computes air-sea fluxes of momentum and heat from 6-hours operational analysis atmospheric forcing (from ECMWF) in all the Mediterranean basin.

In October 2001 the ADRICOSM (ADRiatic sea integrated COastal areaS and river basin Management system) pilot project was launched with the purpose of implement a near real time monitoring system and a near real time basin-shelf marine forecasting system that now run in operational mode for the Adriatic Sea (`http://gnoo.bo.ingv.`

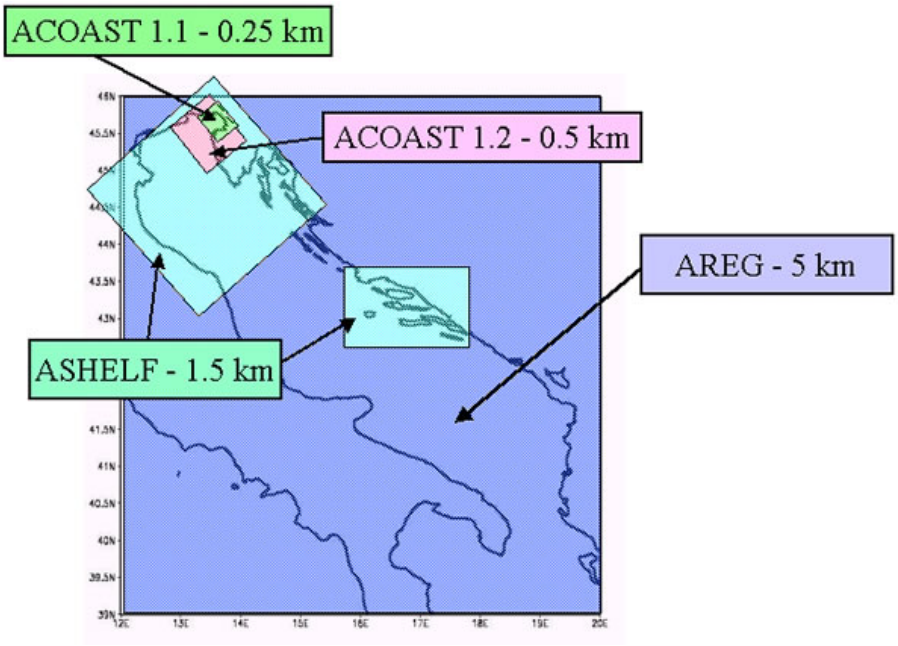


Figure 1.23: *Geographical domains of the models developed in the Adriatic Sea during the ADRICOSM project*

`it/adricosm/`). At the end of the project (in 2005) the following system of models have been implemented:

- The Adriatic REGional model (AREG), a general circulation model for the Adriatic basin that extends up to the Ionian Sea ([61], [42], [44]). AREG has an horizontal resolution of about 5 km (1/20°), while the vertical resolution is defined by 21 sigma levels. The model is an implementation of the Princeton Ocean Model, POM. The weather data (temperature, relative humidity, cloud cover, zonal and meridional components of the wind) with a horizontal resolution of 0.5 ° and temporal frequency of 6 hours, are provided by the European Center for Medium Range Weather Forecast (ECMWF). The values of precipitation come from the average monthly climatological global means by Legates and Wilmott ([29]). The data of rivers flow are derived from monthly averages of climatological [52], with the exception of the river Po. The Po river flow values are inferred from acquiring daily observations at the discharge point of Pontelagoscuro that are then divided on six grid points of the model. At the open boundary located in the Ionian Sea, AREG receives the boundary conditions from the operational ocean model of the Mediterranean Sea (which, at the time, was the 1/8° horizontal resolution model). The forecasting system AFS (Adriatic Forecasting System, `http://gnoo.bo.ingv.it/afs/`) has been operational since April 2003 and issuing 7-day forecasts and simulations each week. As initial condition (January 1, 1999 at 00:00 GMT) were used climatological fields of temperature, salinity and speed of movement produced by simulations of the Adriatic from [61];

- the north Adriatic Shelf Model (ASHELF-1), a model of the Northern Adriatic has a horizontal resolution of 1.5 km Its open boundary cuts the basin along an imaginary line that extends from the southern tip of the Istrian peninsula until the Italian coast , approximatively at a latitude of 43.6 ° N;
- the Adriatic Shelf Model for the coastal area of Split (ASHELF-2), a model whose domain includes the coastal area near Split and includes the islands of Brač and Šolta, the channel between the peninsula Pelješac and the island of Korčula, and the channel between the main island and the island of Hvar ([45]); ASHELF-2 has a horizontal resolution of 1 km. The model includes four rivers: Jadro, Žrovnica, Cetina and Neretva;
- two very high resolution models for the Gulf of Trieste (ACOAST-1 and ACOAST-2). The domain ACOAST-1 has a horizontal resolution of 250 m and is an implementation model circulation MIT (Massachusetts Institute of Technology, Boston), a three-dimensional free surface, finite volume model ([11]). The ACOAST-2 is an implementation of the POM with a horizontal resolution of 0.5 km, and its domain includes the area around the Gulf of Trieste, Northern Adriatic([35]).

The models are hierarchically nested into one another, with a finer resolution model nested into a coarser one.

In recent years many spin-off of the ADRICOSM pilot project has been funded:

- the ADRICOSM-EXTENSION (March2005-December 2006) with the objective of enlarging the ADRICOSM experience to all Adriatic Countries (i.e. Italy, Slovenia, Croatia, Albania, Bosnia-Herzegovina and Serbia-Montenegro) and of continuing the implementation of a state of the art monitoring and forecasting system for the marine coastal areas and the river catchments;
- the ADRICOSM-NERES (October 2004-September 2006) project, which has the aim to improve ecosystem and protected area management practices of the Neretva River Delta. The study of the estuarine area of the river is essential in order to understand the sea-land interactions in the transition-environments.
- the ADRICOSM Pula-Bay project (October 2004-January 2006), which applies the ADRICOSM experience and methodology to the Pula Bay coastal environment.
- the ADRICOSM-STAR project (December 2006-December 2009) which aims at the development and partial implementation of an integrated coastal area and river and urban waters management system that considers both observational and modelling components. The area of investigation is the Bojana river Delta and the whole Montenegro Adriatic coast.

These projects aim to improve the forecasting and management system and education in the Adriatic Sea area, not only for a purely short term, operational application, but also in the broader framework of studying the climate change at regional level.

From June 2007 in the framework of the AFS project implemented during the ADRICOSM project, many improvements have been made to all components of the forecasting

system, namely the observing system, in the data assimilation algorithm and its implementation and in the numerical model. In particular a new implementation of the AREG model has been realized ([25]). The updated model, called AREG2 is characterized by a horizontal grid consisting of 311 points in latitude and longitude in the 287 points corresponding to a horizontal resolution of 2.2 km and a vertical composed of 31 levels non-uniformly distributed coordinates (sigma). The increase of the resolution allows a more detailed representation of the structures that the old model was not able to resolve. The atmospheric forcings always came from ECMWF analyses and forecasts (depending on the mode of the model run, diagnostic or prognostic). The precipitation rate comes from a climatological dataset from [29] and the river runoff is parameterized by the Raicich climatoligy, [52] except for the Po and the Buna/Bojana rivers. The Po river flow values used are daily means observed at the cross section of Pontelagoscuro (as the first implementation of AREG) provided by ARPA SIM Emilia-Romagna, and the Buna/Bojana climatological flow values of Raicich's have been substituted with those ones of UNEP published in [59].
The morphology of the seabed dataset used by the model was extracted from the database NATO DBDB1 with 1 minute resolution. Lateral open boundary conditions are defined through a simple off-line, one-way nesting technique with the general circulation model of the Mediterranean Sea (OGCM) implemented in the MFS.

1.2.3 Data Assimilation

Although the oceanographic models are becoming more and more refined they will never be capable of a perfect, long term forecast of the full state of the oceans.
The reason for that lies both in the approximation made in the equations implemented in the model and (more important) in the fact that we are trying to model a natural, non-linear, dynamical system.

Recall the fact that real systems are intrinsically chaotic. The peculiarity of such a system is the sensitivity to initial condition, a concept explained in figure 1.24. Consider the set Λ, whose points are assumed to represent different possible configurations of the dynamical system we are considering. Now take a point $x \in \Lambda$ that can be seen as a possible initial condition for the system; the evolution of the system in time, starting from a given initial condition x is called $f_t(x)$. We assume there exists a point $y \in V$, being $V \subseteq \Lambda$ and a number $\delta > 0$. A fundamental feature of the dynamical system is that, given the points x and y just defined, given a time $t > 0$ and however small is V, the distance between $f_t(x)$ and $f_t(y)$ is always greater than δ. This char-

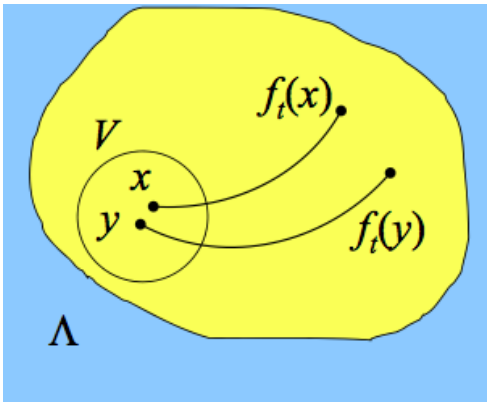


Figure 1.24: *Sensitivity to initial conditions*

acteristic is called sensitivity to initial condition and
means that, even if two initial conditions of a dynamical system are very close, the evolution of the system will, soon or later, differ macroscopically.
So, even if we assume that our model perfectly describes the evolution of a certain natural system, meaning that there are no approximation in the implementation and integration of the governing equations of the system, we will eventually end up with completely unreliable forecasts, because our initial condition, given necessarily by an analysis of observations that are affected by some uncertainties, will never be exactly the real initial state of the system. In practice this lead to the somewhat conventional predictability limit of 7 days in the atmosphere and about 10 days in the oceans that we already mentioned at the beginning of the section.

Summarizing, we can't, and we'll never be able to, produce an error-free forecast of the sea, and these errors will continuously grow in time, until our model state is completely different from the real state. All we can do, is to control the error growth of the model state with the data, which are an approximate measurement of some of the real state parameters.
The proper way to proceed is effectively summarized in a famous sentence of Gauss:

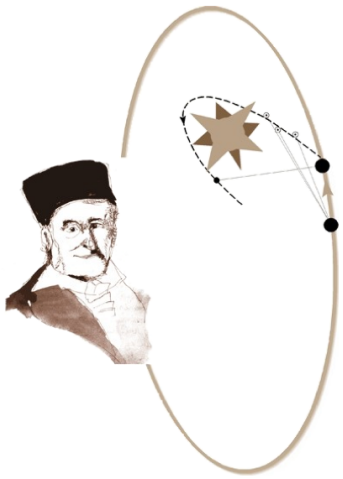


Figure 1.25: *An image of Karl Friedrich Gauss (1777-1855) with a schematic representation of the prediction of a celestial body's orbit*

> '... since our observations are nothing more than approximation to the truth, ... we need ... a suitable combination of all observations to approximate as much as possible the truth. The problem can be only undertaken properly when an approximate knowledge of the orbit has been already attained, which is afterward to be corrected so as to satisfy all of the observations in the most accurate manner possibly'

Following this philosophy he was able to solve a many-body problem of celestial mechanics which is an example of a deterministic chaotic system. In 1801 the Italian astronomer Giuseppe Piazzi discovered the asteroid Ceres, but he could follow its orbit just for a few days before it disappeared behind the moon. Gauss predicted the exact point where it would reappear making use of the newly discovered method of least squares. Ceres appeared at the point indicated by Gauss.
The concept behind the technique used by Gauss is an example of what we need to solve our predictability problem: we have to melt the data into the model. Doing this in a proper way is the aim of *data assimilation* techniques.

The idea is actually really simple: since the forecasting capabilities of the model alone are intrinsically limited by the non-linearities of the governing equations, it has to be somehow corrected for the available observations. This is the fundamental concept expressed

in Gauss's sentence: our understanding of the ocean's state, circulation and variability (as any field of the human knowledge) must proceed following an incremental approach meaning that our initial knowledge about the state of the ocean has to be corrected as new observations became available leading to a new, improved analyzed state of the ocean which can serve as initial condition for initializing a new forecast.
Technically, data assimilation is a procedure in which observations are combined in an *optimal way* with the numerical modeling. The optimal way is to be intended as minimizing the combined errors of the observations and the model simulations. From the mathematical point of view data assimilation could be considered as an inverse problem. That means, given two independent estimates of a the same unknown and its respective errors data assimilation looks for the most probable solution.
The field estimate made by melding data and dynamics by a short dynamical adjustment model run, after assimilating data, is called *nowcast* which is of primary importance to initialize the model in order to produce successful forecast.
Forecasts of increasing accuracy have resulted from refining numerical models and from refining data assimilation techniques together with the availability of new data sets, such as satellite data. Both refinements have been made possible by investment in powerful computer systems, complementing the even larger investment made in the observing systems.

So far we have considered the problem of initializing a forecast, but this is not the only application of data assimilation. DA techniques are used to produce *analysis* and *reanalysis*. The main difference between the two is that usually analysis are carried out in a operational way: day by day or week by week an analysis is produced and the last field is the nowcast from which a forecast is initialized. A reanalysis is the best estimate of the *past* state of the ocean based on complete time series of observations at previous times (in which case the DA algorithm is called a *filter*) or even at successive times (and the DA algorithm is called a *smoother*).
In a re-analysis, the observations collected in past decades are fed into a modern forecasting system that is much more refined than the systems available when most of the observations were made. Oceanic and boundary conditions are reconstructed for each day of the period over which suitable observations exist.
Re-analysis are becoming a tool of increasing importance for climate change studies because they create a new view of the climate of the ocean and its variations by processing a wide variety of observations simultaneously, using the physical laws embodied in the forecast model and knowledge of the typical errors of forecasts and observations to interpret conflicting or indirect observations and fill gaps in observational coverage.

Data Assimilation algorithms and schemes

Depending on the purpose of our study, on our oceanographic model, on the kind and amount of available data and on the computational power we can rely on, we must choose carefully the assimilation algorithm and its practical implementation.

The techniques that are currently available can be classified into two different categories: *sequential* and *variational*.
Methods of the first class look for a solution written in an explicit form starting from a forecast of the state vector and of its error covariance. Then a correction is computed based on the *data residuals*, i.e. the difference between the data and model predicted values for the data. Common algorithms beolonging to this first class are Optimal Interpolation (OI), Kalman Filter (KF) and derivatives like Ensemble KF, Extended KF, Kalman Smoothing. Smoother differs from filter because it uses also formally future observations and then it should be capable to obtain more accurate estimates.
The variational methods, instead minimize a cost function in which properly defined distances (or *misfits*) from a first guess of the model state and from the information provided by the data are evaluated. In this categories we find 3D-Var and 4D-var (three dimensional and four dimensional).

An example of both sequential and variational methods is given in figures 1.26 and 1.27. In these figures the blue line represent the time evolution of a mono-dimensional model, or of just one variable of the model state. The blue dotted line represents the first-guess, or the model forecast without the correction for the observations, given by the red stars. The red bar represent the errors in the observations and the blue bar represent the model error.
In figure 1.26 the sequential approach to data assimilation is explained: Starting from the initial time, the model trajectory follows the initial forecast, and is characterized by a growth of the forecast error. As soon as the first observation is available, the analysis is performed (green bullet), and the associated error decreases (green error bar). The same cycle is repeated anytime an observation is available, with the assimilated trajectory deviating from the initial guess (the dashed line).
In figure 1.27 the correction is performed once, assimilating all the available observations in the time window of interest. After adjustment of the initial condition $\mathbf{x}_0$ (the green bullet on the $t = t_0$ axis) by means of an iterative minimization algorithm, the model trajectory is corrected over the entire time window, in order to provide an optimal fit to the data (in a generalized least squares sense).

Main advantages of the sequential algorithm is that they provide a formal estimate of errors for the data assimilation solution. However, at each step the filter/smoother gain has to be evaluated from the inversion of a combination of covariance matrix, which is the most problematic aspect of every inverse problem theory in geophysical sciences.
Variational approaches seek the solution of a minimization problem performing several evaluations of the cost function, which gives the distance between the model variables and the observations, and its gradient. It is therefore an *optimization* problem, that can be troublesome if a non well suited minimization algorithm is used: compared with variational approaches such as the 3D-Var, sequential algorithms require less initial investment in terms of coding. On the other hand the variational schemes does not require, as in sequential schemes, complex linearized observational operator (projector of the model space

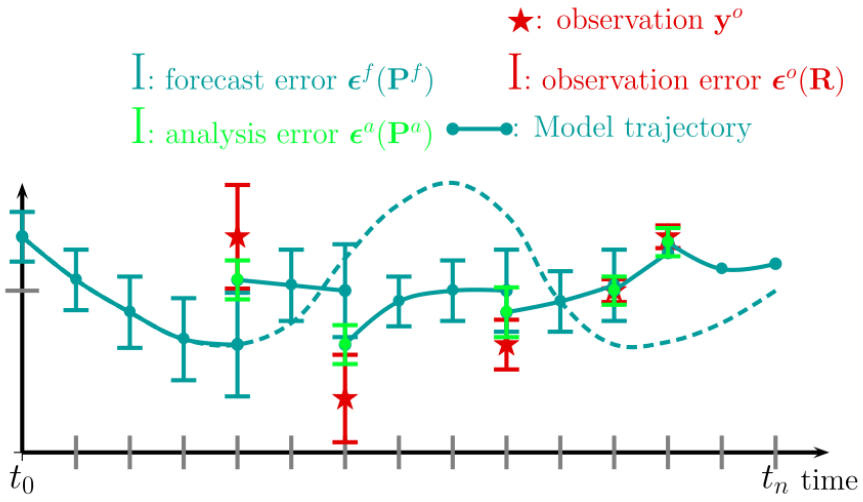


Figure 1.26: *An example of Sequential Data Assimilation. From [21]*

onto the observational space) and are naturally designed to incorporate gradual developments.

Although data assimilation theory is well-known, direct implementation of such assimilation schemes is impractical for most state-of-the-art OGCM because of the computational requirements. Typical state-of-the art general circulation models have millions of prognostic variables. Consequently, the corresponding model's state error covariance matrix, which describes the errors relationships, has over a million squared elements.
In order to solve this problem various approximations of variational and sequential methods have been put forth to simplify their computational implementation. The main problem is therefore finding a compromise between physical and stochastic coherence and computational feasibility, the definition of which may vary depending on the purpose of the system (whether it is or not an operational system).

1.2.4 Forecasting activities in the Adriatic Sea: the Adriatic basin Forecasting System

As already mentioned, within the ADRICOSM project a system for the Adriatic Sea monitoring and forecasting has been implemented: the Adriatic basin Forecasting System (AFS), maintained operational by the scientists of the Operational Oceanography National Group GNOO (Gruppo Nazionale di Oceanografia Operativa) in Bologna. The GNOO is a section of the Italian Geophysical Institute (INGV) and is responsible for the forecasting activities both in the Adriatic Sea and in Mediterranean Sea.
Every day GNOO produces forecasts for the next ten days in the whole Mediterranean and for the next nine days at the sub-regional scale of the Adriatic basin. Temperature, salinity, intensity and direction of currents and non tidal sea level are released from sea surface to the bottom. Wind stress, sea level anomaly and heat flux are computed at

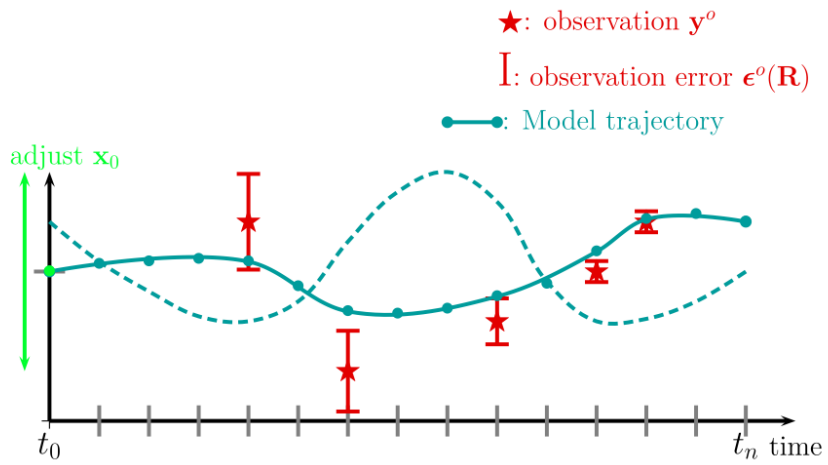


Figure 1.27: *An example of Variational Data Assimilation. From [21]*

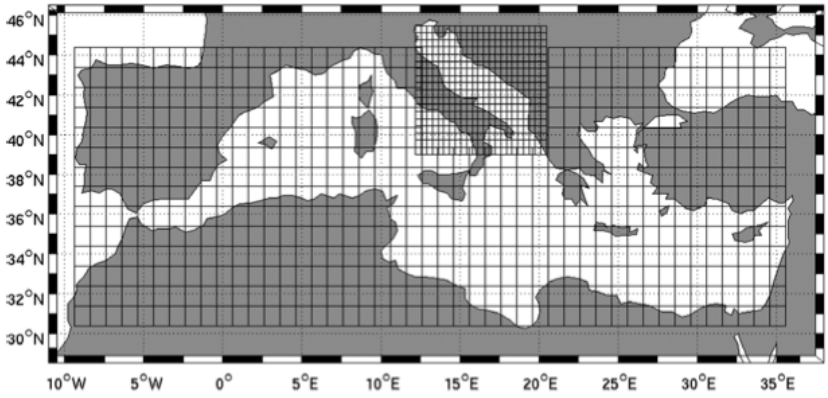


Figure 1.28: *MFS and AFS domain and bounds. The Adriatic Sea model grid is the finer resolution grid superimposed to the coarser resolution grid of the Mediterranean model*

air/sea interface.
Further information, graphical displays of forecasts and analyses can be found on the MFS and AFS website:

- MFS: `http://gnoo.bo.ingv.it/mfs`
- AFS: `http://gnoo.bo.ingv.it/afs`

There were many implementation of the Adriatic forecasting system, which is continuously validated and improved, both in the prognostic model and the assimilation algorithm. In this section we briefly describe the various versions of the system in light of the concepts so far exposed about the numerical ocean modeling and forecasting.

The AFS experience starts with the development of the AREG model during the ADRICOSM project. The model is presented in [42]. In the same paper, the authors also

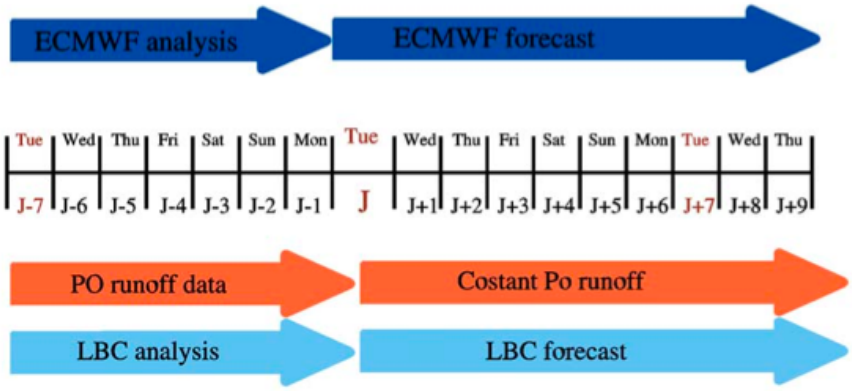


Figure 1.29: *Time line of hindcast-forecast procedure. The arrows indicate the external data collected on Tuesday (J) and used for the model simulations. LBC indicates the large scale model data used for Lateral Open Boundary conditions. From [44].*

exposed a validation experiment for the model's capabilities in reproducing the principal aspects of the Adriatic Sea general circulation. The authors performed an analysis of the 1999-2003 hindacst simulation focusing on the interannual variability of the Adriatic Sea general circulation. No data assimilation is implemented in this study. At the lateral open boundary the model is one way nested with the general circulation 1/8° resolution of the Mediterranean Sea that, at the time, was used in the MFS. A representation of the nesting is given in figure 1.28: the open boundary of AREG is at a latitude of 39°N . The quantities specified at the boundary are daily averaged temperature, salinity and velocity fields, properly interpolated on the finer AREG grid in order to preserver the volume transport across the boundary. The model is forced with ECMWF analyses and lateral boundary condition and Po runoff are given by MFS results for the period analyzed. The validation of the model is based on comparison between the simulations with *in-situ* observations.The model provided its capabilities in simulating the characteristics of the known general circulation and its variability connected to the atmospheric forcing.
The comparison with the observation showed a general overestimation in the vertical and horizontal mixing processes, the main issues in ocean models, and a deficiency connected to the inflow of salty waters from the Ionian Sea. Principal source of errors were also found to be the rivers temperature effects, actually neglected in this implementation of AFS.

The AREG model has been used operatively since 2003 without data assimilation. In [44] the first implementation of the AFS is presented and its hindacst/forecast performances are tested. The model design and implementation is the same presented in [42]: in particular the AFS model is one way nested in the MFS coarser resolution model.
The operational scheme implemented in this version of AFS is showed in figure 1.29. Every week on Tuesday (J) the system produced an analysis of the previous 7 days, starting from noon of the previous Tuesday (J-7) to noon of the current Tuesday. The model was then integrated forward in time, producing a 9-days forecast, spanning the period from 6.00 p.m. of the current Tuesday (J) to Thursday of the successive week (J+9), using the

analysis of the past 7 days as initial condition.
During the forecast cycle, the forcing and lateral boundary conditions (LBC in figure 1.29) were given by the ECMWF and MFS forecasts respectively, while the Po runoff was set equal to the last available observation of the river discharge (see [44] for a justification for this). During the analyses these information were provided by ECMWF and MFS hindcasts and Po past observed runoffs.
In [44] a validation experiment of AFS is presented. The forecast accuracy and the quality of the hindacst have been evaluated for the year 2003 (the intensive data collection period for the ADRICOSM project) from a comparison with both with in-situ and remote observations of SST.

In 2006, Grezio and Pinardi, presented a data assimilation scheme suitable for an AFS implementation. In [23] the authors proposed an OI scheme called SOFA (System for Ocean Forecasting and Analysis), which is a reduced order, optimal interpolation scheme based on the method presented in [12]. This scheme was already implemented in the MFS. The data considered for the assimilation in [23] came from the ADRICOSM observing system, composed of XBT profiles collected along Voluntary Observing Ships (VOS) tracks in the southern part of the basin, and of CTDs profiles collected on four coastal shelf area in the northern and middle Adriatic.
The SOFA algorithm should have become operationally implemented in AFS but it didn't, because further studies showed that the scheme was not sufficiently "robust" from a statistical point of view, even if it had already proven good capabilities in MFS.

Subsequently, the modeling component of the system was updated by Oddo and Guarnieri who first implemented the AREG2 model, already discussed, on the same operational scheme previously used for AREG. The new model had a higher spatial resolution, both in the horizontal and in the vertical where the sigma layer where augmented from 21 to 31. No paper has been published by the GNOO scientists on this step of the evolution of AFS.
Finally thanks to Guarnieri, the AREG2 model reached its present version. The model is now capable of release both hourly results and the usually daily means for temperature, salinity, intensity and direction of currents on the entire model grid and wind stress, sea level anomaly and heat flux computed at air/sea interface. Besides, following the results presented in [24] also the tidal effects on these variables are now implemented in the model.

The forecast/hindcast scheme of the present AFS implementation is schematically shown in figure 1.30. Every day the model produces daily forecasts for the future 9 days (FCST in figure 1.30). RST stand for RESTART, which is the name of the file given as an output from the simulation of the previous day and act as an initial condition for the subsequent simulation.
Once per week the system provides an analysis (simulation) with the best estimates of the physical conditions of the Adriatic Sea for the previous 15 days.
The lateral boundary condition for the analyses is provided by the MFS hindcast, the

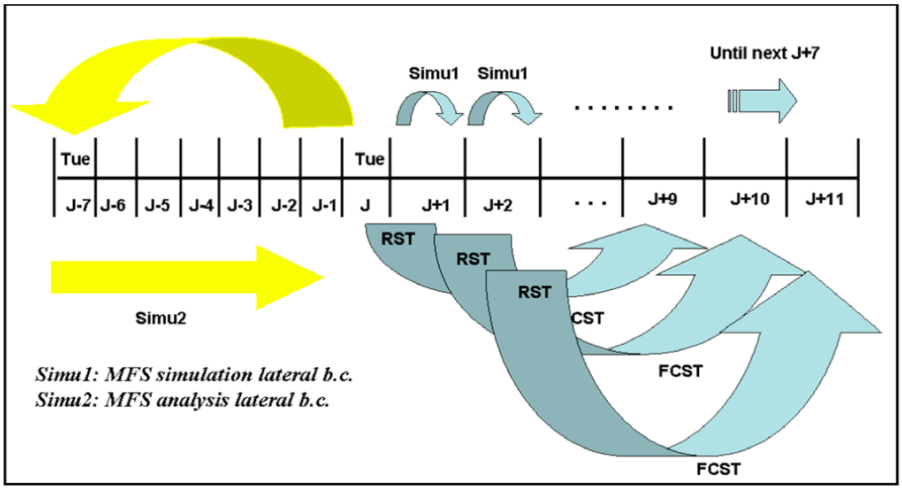


Figure 1.30: *The AFS assimilation algorithm implementation.*

atmospheric forcing is given by the ECMWF analyses fields and the Po runoff is given by the observations. During the forecasts the forcing and lateral boundary condition are given by ECMWF and MFS forecasts, while the Po river runoff is considered as a constant equal to the last available observations for the river discharge.

Data are available in NetCDF format and distributed to research users (with free access) and commercial users (with specific contracts).
A graphical display of the daily forecasts and simulation is freely available at the AFS website.

But the model is still implemented in AFS without a suitable assimilation scheme.
In recent years the attention of GNOO has turned to variational assimilation schemes for operational purposes both in the Mediterranean and in the Adriatic Sea. The present computational power and the experience gained with the MFS and AFS experience should permit the implementation of these algorithms for oceanographic forecasting activities.
At the present days the future evolution of AFS is thought to be the implementation of the 3D-Var scheme presented in [16].
It is in this framework that the present thesis is placedThe results in this thesis are concerned with the previous questions.

1.3 Objectives and structure of the thesis

The implementation of a new assimilation algorithm in an operational oceanographic system must be carried on with great care. Every system has its own peculiarities and an assimilation scheme that has proven to be suitable for another, although similar, environmental system will not necessarily fit in a new forecasting network.
Thus, even if the 3D-Var has already proven good capabilities when implemented in oper-

ational systems (both in metereologic and oceanographic services), the specific algorithm that would be developed for the Adriatic basin Forecasting System has to be carefully validated.

This thesis presents a validation experiment for the possibly future version of AFS: the AREG2 model combined with a 3D-Var assimilation algorithm. Specifically the experiment will be a one year daily assimilation of the available observations for the year 2006 using data that will be realistically used in AFS.
This validation experiment is similar to that presented in [23] for the SOFA assimilation scheme but here we will assimilate in-situ and satellite SLA observations using a different, more evolved assimilation algorithm.
The fact that the target year is well back in the past, allows us to use *delayed time* data, meaning that the processing and quality control of the observation is performed at much higher quality with respect to *near-real time* data that should be used in for operational purposes. This allow us to focus only on the performances of the system, neglecting possible gross errors in the observations due to outliers and erroneous analyses.
The present work is thus an analysis of the first run of the system that would be implemented in the new AFS.

The thesis is organized as follows:
In chapter 2 we present the data that have been collected for the target year 2006 and the dataset selected for the experiment.
In chapter 3 we describe the system set up. We describe the features of the AREG2 numerical model, the governing equations and the specific set up of the model implemented in the Analysis System used for the study. Then we give a more mathematical background of the Data Assimilation techniques focusing on the 3D-Var assimilation scheme and its specific implementation with AREG2, describing the assimilation scheme. The experiment design is also presented.
In chapter 4 we give the results of the one year assimilation cycle of the system. We present an evaluation of the analysis quality and how the dynamics of the Adriatic Sea improved thanks to the assimilation of observations.
In chapter 5 we give our conclusions together with final remarks and future development of the system.

Chapter 2

Observed data set

As previously said the target year for the experiment under discussion is the year 2006. This year has been chosen on the basis of the abundance of in-situ data, the only direct information we have about temperature and salinity of the sea.
Thus, in this study, we assimilate in-situ observations of temperature and salinity profiles and satellite measured SLA. In addition to the assimilated data we will consider the sea level records of the tide gauges, kept outside the assimilation process as an independent test set.

In the rest of the chapter we give more details about the dataset used in the experiment.

2.1 In-situ data

2.1.1 Temperature and salinity profiles

There are many datasets of T-S profiles available for the experiment. Some of these datasets are not specific for the Adriatic Sea, so we considered the profiles which latitude were not under 39° N.
The available datasets are:

SeaDataNet Composed of 1921 T-S profiles from XBT, CTD, and bottle casts. SeaDataNet (SDN) is a pan-european infrastructure that has the objective of propose a standardized system for managing the large and diverse data sets collected by the oceanographic fleets and the automatic observation systems. More informations and data access policy can be found on the SDN website: `http://www.seadatanet.org/`.
In figure 2.1 is shown the spatial distribution of the SDN profiles in the Adriatic Sea for the year 2006.

ENEA The Global Climate Unity of the ENEA (Ente Nazionale per le Energie Alternative) maintains, among the other oceanographic activities, the portal for the dissemination and control of data from ships of opportunity (SOOP and VOS) and

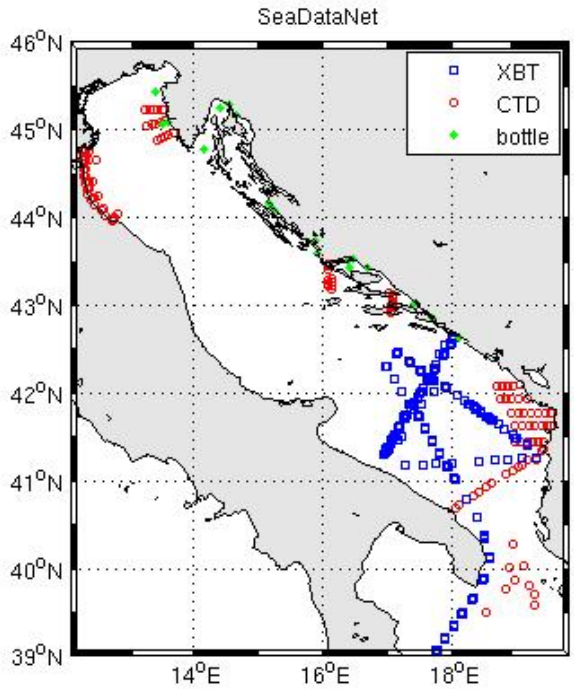


Figure 2.1: *Spatial distribution of the SeaDataNet profiles in 2006*

from coastal buoys. The interest of ENEA, its modeling and observational activities are mainly focused in the Mediterranean and Tyrrhenian Sea.
In this Adriatic Sea data set we have 1306 T-S profiles from XBT and CTD, distributed as shown in figure 2.2.

ARPA The Emilia-Romagna environmental agency ARPA provides, in operational modes, a dataset of the principal biogeochemical and physical quantities of the sea along the Romagna's coasts. The Oceanographic Structure DAPHNE is a Thematic Unit of ARPA and represents an operative node of the operational network for environmental monitoring. This activity takes place at 42 stations up to 20 km from the coast (figure 2.3). Monitoring is carried out by the motor boar Daphne II, which sails twice a week all year round. Daphne II is an oceanographic boat equipped to collect measurements of the marine ecosystem and water quality. This observations, for their nature and temporal resolution, provide a unique dataset for understanding the environmental system and dynamic features of the coastal area of Emilia-Romagna and it is one of the principal coastal datasets of the whole Adriatic basin.
The dataset provided by the DAPHNE platform for the year 2006 is composed of 1383 T-S profiles from CTD casts.

Considering the profiles of this datasets, it was clear that the both the ARPA data and ENEA data were included in the SDN datasets. Besides, among the others SDN provides the best quality control on the data. So we decided to use the SDN dataset for the

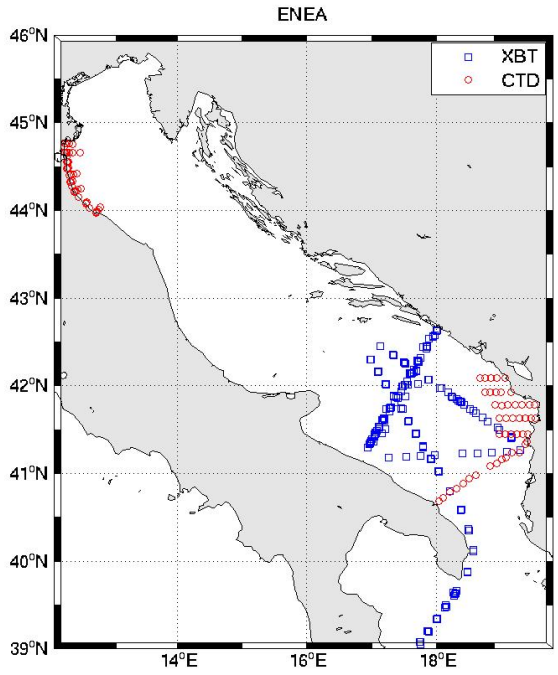


Figure 2.2: *Spatial distribution of the ENEA profiles in 2006*

assimilation experiment.

Now we briefly analyze our dataset. Considering the bathymetry of the basin and the figure 2.1 we can identify profiles that are located in truly coastal areas, especially in the northern and central sub-basin, and other profiles that are taken in open ocean areas, in the southern Adriatic. In particular we can see that the XBT observations are casted where the depth is more than 500 m. This is due to the fact, mentioned in the previous chapter, that an XBT observation should be carried in deep sea environment because of the expendable nature of the probe. This observations are carried by vessel of opportunity (VOS and SOOP) having routes within and outside the Adriatic.
In the southern sub-basin we can also identify CTD casts in open ocean areas and coastal T-S profiles in the eastern side of the Otranto Channel, near the mouth of the Albanian rivers.
This profiles are essential in catching the specific properties of the waters flowing through the Otranto Channel and thus their assimilation allow the model for a better simulation of the therohaline circulation of the basin.
On the other hand, the regional network provide important information about coastal dynamics that the model would hardly be able to reproduce if left in free run mode (that is, without assimilation), due to the high frequency, low wavelength variations that can characterize such environments.

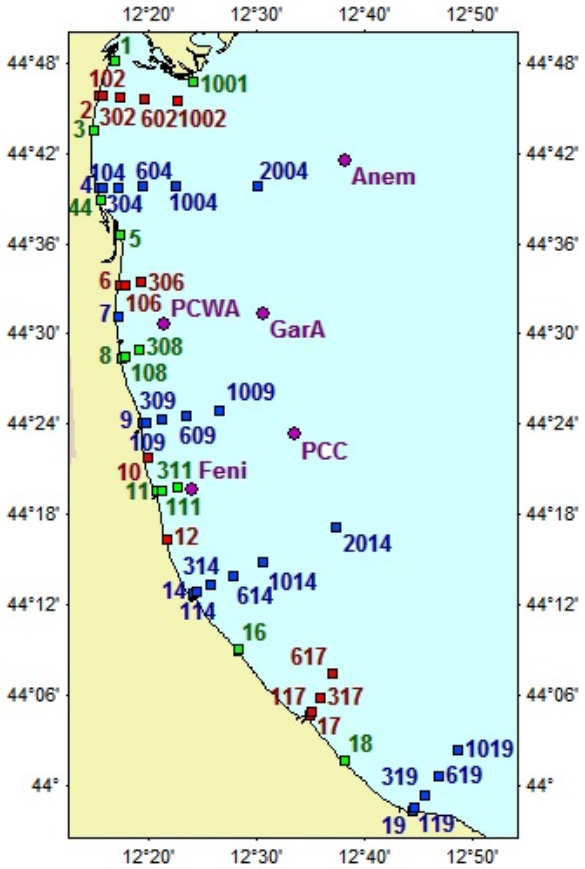


Figure 2.3: *The ARPA-DAPHNE observation points*

On the basis of these considerations and following [23] we can divide the entire dataset into four subsets, as shown in figure 2.4:

- The ARPA region in front of the Emilia-Romagna coasts comprises the data included in the ARPA dataset previously described;
- The Rovinj region includes the local monitoring network of Rovinj and some bottle casts around the Istrian peninsula and in the Kvarner Gulf:
- The Middle Adriatic region comprises the data from the coastal monitoring system of Spit and some bottle casts along the eastern coast of the basin;
- The Southern Adriatic region is the open ocean area of XBT casts and includes some coastal and open sea CTD T-S profiles.

The monthly distribution of the T-S profiles from the SDN dataset in the target year 2006 in figure 2.5. We can note from this figure that the data are numerous and well distributed over the year, with only one month (January) that has less than 100 total ovservations. For the single subsets, the same distribution is plotted in figures 2.6. Even from this figure we can note a good temporal coverage of the data, with few months having

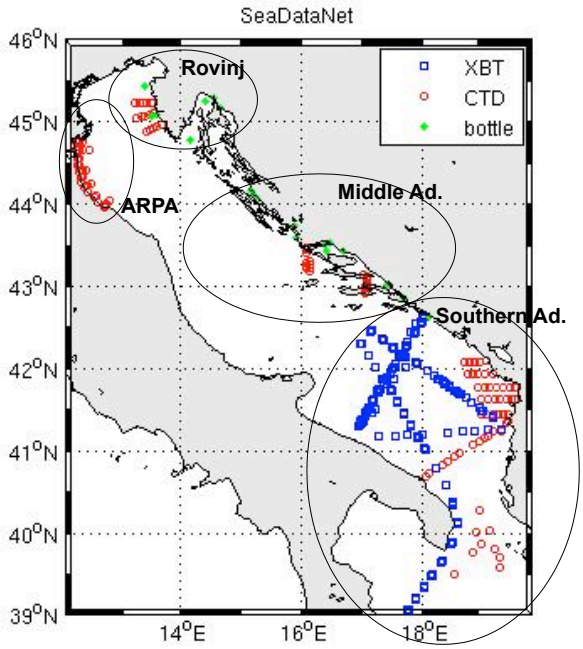


Figure 2.4: *The SDN dataset and the identified subsets*

no observations. We could have already noticed from the considerations on the ARPA, ENEA and SDN datasets that the observations in the ARPA region are more abundant than elsewhere in the Adriatic Sea and the temporal distribution is not far from being constant.
The other subsets are well populated too over the entire year. Therefore we expect that our model to be free from T-S profiles assimilation for no more than few days during our run. Our model will be well constrained by the in-situ observations.

2.1.2 Sea Level

In this work the in-situ sea level observations is given by 8 ISPRA (the national institute for the research and protection of the environment) stations whose names and locations in the Adriatic Sea is given in figure 2.7.
This data are collected in hourly records of sea level, already corrected for the land movements and comprehensive of the tidal elevations.
As already explained, this observations are not assimilated into the model in this experiment, but are used as an independent test set, to be compared with the outputs of the model before and after the assimilation.

The actual data that we will use in our experiment are daily means of these records. There are two reason for that: First of all the output of the model (as explained in the next section) is composed of daily averages of the fields of interest (temperature, salinity, currents, surface flux and surface elevation). So we need the sea level daily observation to

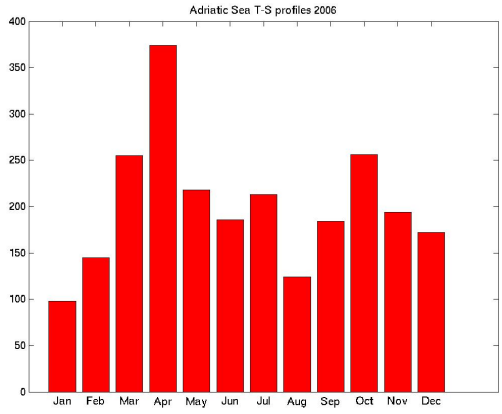


Figure 2.5: *Temporal distribution of the T-S profiles in the whole Adriatic Sea from SDN dataset*

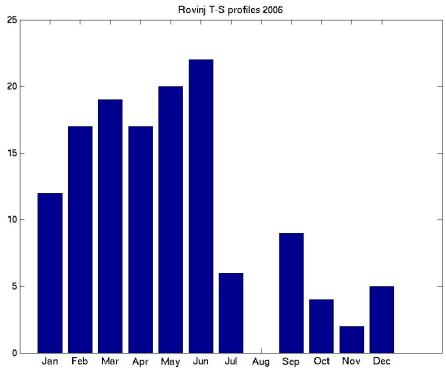
make the right comparison with the model output. In the second place, calculating the daily average will filter out the semidiurnal and diurnal tidal effect. This is also a necessary processing for the right comparison because the model we will use in the experiment does not implement the tidal forcing. This is because we don't want to care about the tides in this experiment, but on the principal features of the general circulation.
Daily means of observed sea level are given in figures 2.8.

2.2 Satellite data

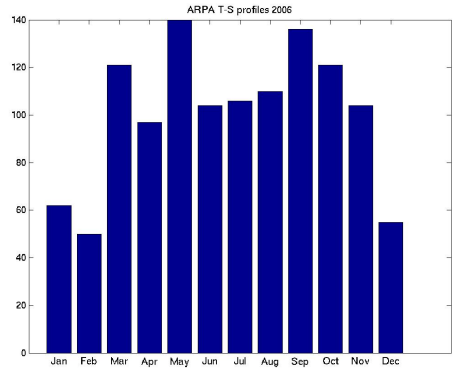
In the previous chapter we already explained that the principal satellite observations that are typically assimilated in operational oceanography are SLA and SST. We also mentioned the advantages that the first techniques has over the latter.

In our experiment we will assimilate SLA observations but not SST due to the poor quality of the data for the target year.

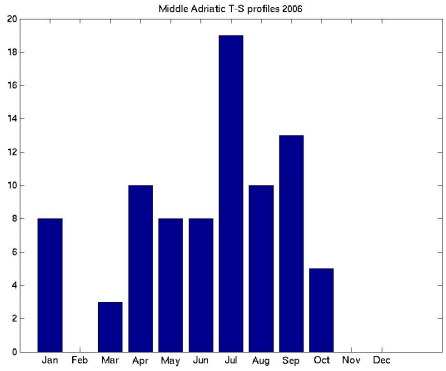
For the target year 2006 the altimetry satellites available are Geosat2, Envisat and Jason1 (see figure 1.20). This data provides a set of observations that is uniformly distributed both in time and space: an average of about 1000 observation points are present every month almost uniformly distributed on the surface of the Adriatic basin.
The data we used in this experiment are not the real-time (RT) data quickly analyzed for operational purposes. We make use of dealy-time (DT) data analyzed in the 2010. Therefore they are the best analyzed data available for the target year in the Adriatic Sea.
More information about this dataset can be found on the website from where they have been downloaded: the AVISO web site `http://www.aviso.oceanobs.com/`.



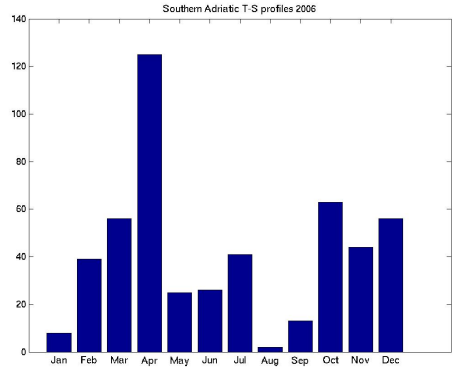
(a) *Rovinj data*



(b) *ARPA data*



(c) *Middle Adriatic data*



(d) *Southern Adriatic data*

Figure 2.6: *Monthly distribution of the subsets of SDN in the Adriatic Sea in 2006. Subsets are defined in figure 2.4*

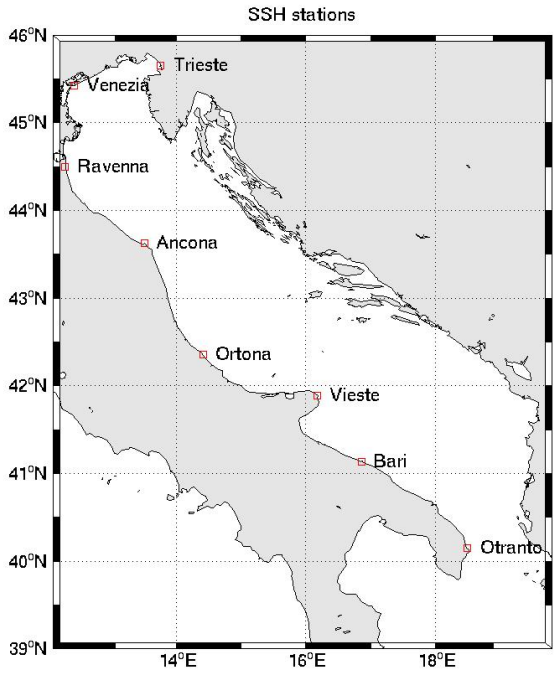
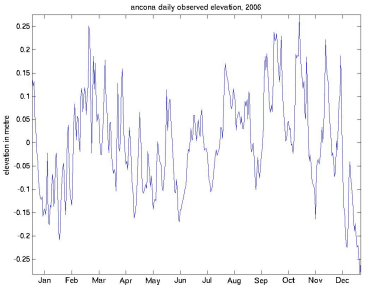
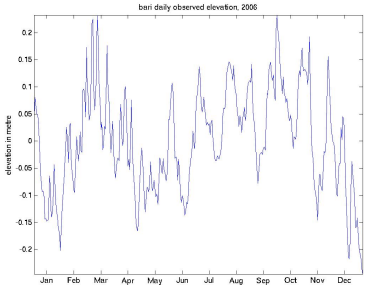


Figure 2.7: *The location of the ISPRA mareographic stations*

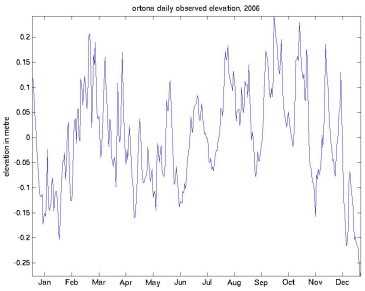
This data set has not been fully analyzed both because of lack of time and because our interest here is mainly focused on the in-situ observations.
It should be noted, again, that in the Adriatic Sea there is no systematic study of the SLA observations. Meaning that no one yet has stated what details of the dynamics in the basin are best observed by this techniques.



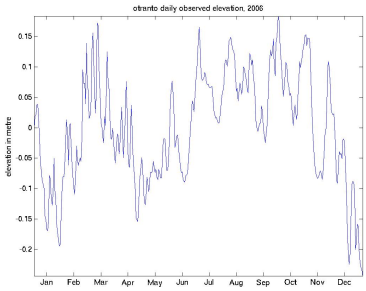
(a) *Ancona*



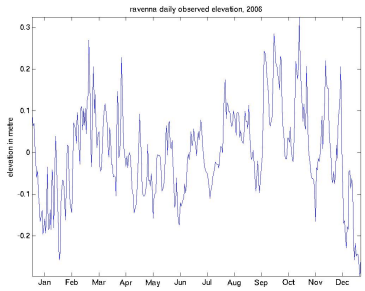
(b) *Bari*



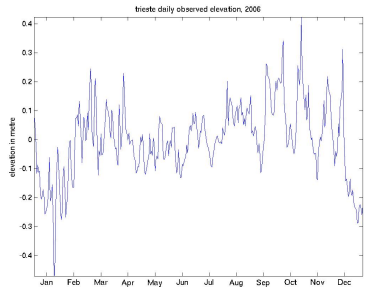
(c) *Ortona*



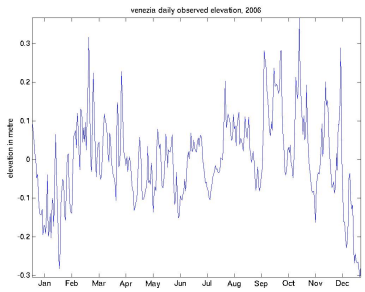
(d) *Otranto*



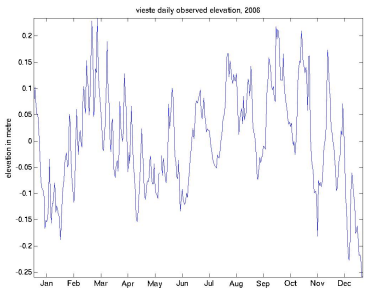
(e) *Ravenna*



(f) *Trieste*



(g) *Venezia*



(h) *Vieste*

Figure 2.8: *Time series of the daily observed sea level from the tide gauges of figure 2.7. This time series are obtained subracting the mean elevation of the year 2006 and averaging over the single days, or the plot would be dominated by high frequency tidal effects.*

Chapter 3

Adriatic Sea Analysis System

In this chapter we describe in more detail the circulation model used in this study (the AREG2). In particular we will give the governing equation integrated in the model and the output of the model run. Then we give a more mathematical background of the DA techniques and we introduce the 3 dimensional variational algorithm (3D-Var). After that we describe how these two components are implemented and finally we present the design of the numerical experiment.

3.1 The numerical model (AREG2)

The AREG2 (Adriatic REGional model, second version) is a circulation model that is based on the Princeton Ocean Model (POM) described in [38] and specifically developed for the Adriatic Sea.
As we already explained, this is not the first version of the model, but the underlying physics and the POM basic concepts remain the same. In particular POM is a free-surface, three-dimensional finite-difference numerical model based on the primitive equations with Boussinesq and hydrostatic approximations.
We shall now briefly describe the AREG2 model.

3.1.1 Model domain and coordinate system

The model domain and the bathymetry of the model are shown in figure 3.1. As we can see the model domain is the entire Adriatic Sea and presents an open boundary line at 39 ° N.
The horizontal grid is a a simple rectilinear coordinate system, so the model horizontal resolution is constant and is about 1/33 ° along the longitudinal axis and 1/45 ° in the latitude.
The vertical dimension, on the other hand, is discretized with a $\sigma - layer$ coordinate meaning that the water column is not subdivided into evenly spaced, parallel z-layers (or $x - y$ planes), but we have a fixed number of vertical levels over the entire domain that therefore have a different spacing according to the $(x - y)$ position.

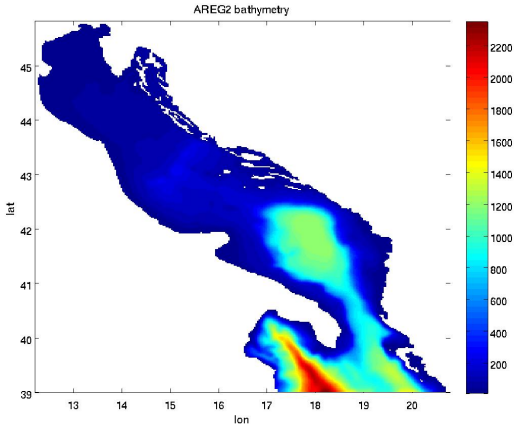


Figure 3.1: *AREG2 model domain and bathymetry*

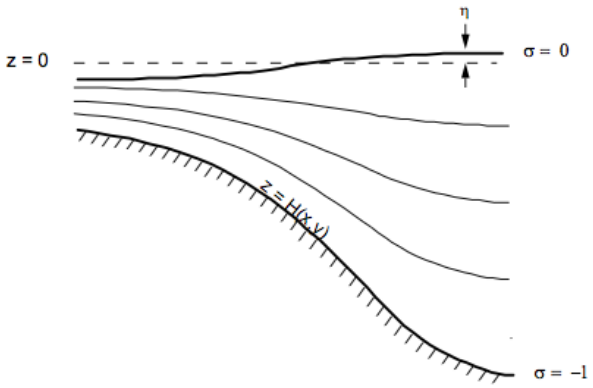


Figure 3.2: *The sigma coordinate system*

The σ-layers, terrain following (or bottom following) coordinate system has ben introduced in [47] and Blumberg and Mellor implemented it in the POM model. The idea is given in figure 3.2: having a variable vertical resolution allow us for a better vertical resolution in shallow water areas and a greater spacing between levels in deep sea regions, which is exactly what is desirable for optimizing the integration of the governing equation.

The original POM σ levels were simply obtained with a linear transformation from the z coordinate:

$$\sigma = \frac{z - H}{H + \eta} \tag{3.1}$$

where $H(x, y)$ is the bathymetry and $\eta(x, y, t)$ is the sea surface. In this way $\sigma = 0$ at the surface and $\sigma = -1$ at the bottom. It is important to note that in the AREG2 model, every point with depth less than 10 m is considered of having a bathymetry of exactly 10 m. So this is the minimum depth resolved by the model.

The actual distribution of the σ levels in AREG2 is not linear but as shown in figure 3.3 it is clear that there exist a finer vertical resolution near the surface and the bottom

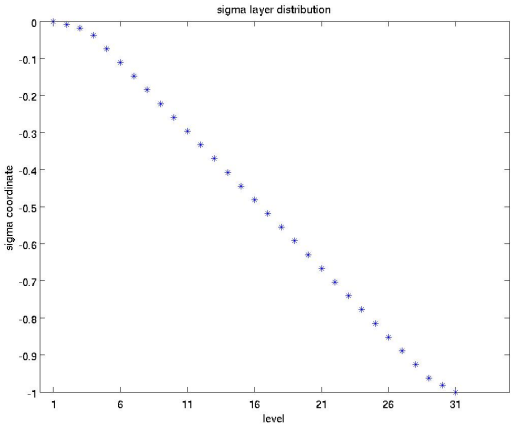


Figure 3.3: *Distribution of the vertical σ levels in AREG2*

(obtained with a logarithmic transformation from the z to the σ coordinate), while the intermediate levels are evenly spaced like the original POM implementation. this choice clearly allows for an improved representation of upper thermocline and bottom boundary layer phenomena

Note from figure 3.2 that the definition of $\eta(x, y, t)$ require the definition of the unperturbed zero level of the sea surface. This is done numerically, by "filling" the model domain with water, in absence of forcing and currents; so the level $\eta = 0$ corresponds to a particular horizontal plane of the model domain.

3.1.2 Governing equations

The governing equation integrated on the three dimensional grid just described are the primitive equations derived from the Navier-Stokes for a rotating fluid subject to the gravity field and the equations for advection-diffusion of salinity and potential temperature (respectively obtained from Fick's law and Fourier's law).
The primitive equations are obtained from the N-S equations through the common simplification of Boussinesq and assuming an incompressible fluid. All the equations are written in rectangular coordinates and contain spatially and temporally varying horizontal eddy viscosity and diffusion coefficients. The model solves the following equations for the ocean velocity $\mathbf{U} = (u, v, w)$ potential temperature θ and salinity S, casted here in rectangular

coordinates for simplicity:

$$\begin{aligned}\frac{\partial(u,v)}{\partial t} + \mathbf{U}\cdot\nabla(u,v) + f(-v,u) = &-\frac{1}{\rho_0}\left(\frac{\partial p}{\partial x},\frac{\partial p}{\partial y}\right) + \\ &+ \nabla_h\cdot(A_M\nabla(u,v)) + \frac{\partial}{\partial z}\left(K_M\frac{\partial(u,v)}{\partial z}\right)\end{aligned} \tag{3.2}$$

$$\nabla\cdot\mathbf{U} \tag{3.3}$$

$$\frac{\partial\theta}{\partial t} + \mathbf{U}\cdot\nabla\theta = \nabla_h\cdot(A_M\nabla_h\theta) + \frac{\partial}{\partial z}\left(K_H\frac{\partial\theta}{\partial z}\right) + \frac{1}{\rho_0 C_p}\frac{\partial I}{\partial z} \tag{3.4}$$

$$\frac{\partial S}{\partial t} + \mathbf{U}\cdot\nabla S = \nabla_h\cdot(A_M\nabla_h S) + \frac{\partial}{\partial z}\left(K_H\frac{\partial S}{\partial z}\right) \tag{3.5}$$

In the above equations p, g and ρ_0 are the pressure, the gravity acceleration on Earth's surface and the reference density value, f is the Coriolis parameter and C_p and α are the specific heat and the thermal expansion coefficient for seawater. A_M is the eddy viscosity coefficient and K_M and K_H are the vertical mixing coefficients for momentum and tracers respectively.
K_M and K_H are calculated with a second-order turbolence closure submodel presented in [39], while the eddy viscosity is parameterized following the scheme of Smagorinsky presented in [57]. The advection part of the hydrodynamics equations are solved through a Monotonic Up-Stream Scheme for Conservation Law (MUSCL, [1]), capable of simulate in a realistic way the horizontal and vertical gradient.
The last term in equation (3.4) is the parameterization of the heat penetration in the water column (already implemented in the MFS, [49]) and the function $I(z)$ is defined as:

$$I(z) = TrQ_s e^{-\lambda z} \tag{3.6}$$

where Qs is the short-waver radiation flux and Tr and λ are the Jerlov's transmission and absorption coefficients, [27], for which we adopted those corresponding to the clear water.
Equations (3.2), (3.3), (3.4) and (3.5) are correlated with the hydrostatic equation:

$$\frac{\partial p}{\partial z} = -\rho(S,\theta,p)g \tag{3.7}$$

where $\rho(S,\theta,p)$ is the density of seawater, which is a function of temperature (here potential temperature), salinity and pressure. This terms may be calculated through well-known empirical state equations: here we use the UNESCO formulas given by Mellor in [36].

Eddy viscosity and vertical mixing

The eddy viscosity A_H is given by the Smagorinsky formula:

$$\begin{aligned}A_M &= Cdxdy\frac{1}{2}\left|\nabla(u,v) + \nabla(u,v)^T\right| \\ &= Cdxdy\frac{1}{2}\sqrt{(\partial u/\partial x)^2 + 1/2\,(\partial v/\partial x + \partial u/\partial y)^2 + (\partial v/\partial y)^2}\end{aligned} \tag{3.8}$$

where dx and dy are the horizontal grid spacing and C is the non dimensional *horcon* parameter. In the POM user guide, [38], the authors suggests for C a value in the range 0.10 - 0.20; in the AREG2 model the horcon is set to 0.05.
Advantages of this formulation of the eddy viscosity are that C is non dimensional and A_M decreases as resolution improves.

Concerning the vertical mixing coefficients, the POM model obtain them from the second order turbulence closure submodel, which is a peculiar feature of the model. The turbulent mixing is parameterized through equation for the turbolent kinetic energy $q^2/2$ and the turbulence length scale l (see [39]). This equations allow for a realistic representation of the turbulent, near-surface and near-bottom layers. The turbulent kinetic energy and turbulent macroscale equations are the following:

$$\begin{aligned}\frac{\partial q^2}{\partial t} + \mathbf{U}\cdot\nabla q^2 = \frac{\partial}{\partial z}\left(K_q\frac{\partial q^2}{\partial z}\right) + \\ +2K_M\left[\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2\right] + \frac{2g}{\rho_0}K_H\frac{\partial\tilde{\rho}}{\partial z} - \frac{2q^3}{B_1 l} + F_q\end{aligned} \tag{3.9}$$

$$\begin{aligned}\frac{\partial q^2 l}{\partial t} + \mathbf{U}\cdot\nabla\left(q^2 l\right) = \frac{\partial}{\partial z}\left(K_q\frac{\partial q^2 l}{\partial z}\right) + \\ +lE_1K_M\left[\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2\right] + \frac{lE_1 g}{\rho_0}K_H\frac{\partial\tilde{\rho}}{\partial z} - \frac{q^3}{B_1}\tilde{w} + F_l\end{aligned} \tag{3.10}$$

where $\tilde{w}$ is the wall proximity function: $\tilde{w} = 1 + E_2\cdot(l/kL)^2$ with $(L)^{-1} = (\eta - z)^{-1} + (H - z)^{-1}$. k is the von Karman's constant and is set to 0.4; E_1, E_2 and B_1 are close constant of the analyzed mode. K_q is the vertical eddy diffusivity of the kinetic energy. $\partial\tilde{\rho}/\partial z = \partial\rho/\partial z - c_s^{-2}\partial p/\partial z$ where c_s is the speed of sound in water. F_q and F_l are horizontal mixing terms for q^2 and $q^2 l$ respectively and are parameterized as follows:

$$F_{q,l} = \frac{\partial}{\partial x}\left[A_H\frac{\partial(q^2, q^2 l)}{\partial x}\right] + \frac{\partial}{\partial y}\left[A_H\frac{\partial(q^2, q^2 l)}{\partial y}\right] \tag{3.11}$$

Mellor and Blumberg in [39] give a solution of (3.9) and (3.10) obtaining:

$$K_q = qlS_q \tag{3.12}$$
$$K_M = qlS_M \tag{3.13}$$
$$K_H = qlS_H \tag{3.14}$$

S_q S_M and S_H themselves are functions of q and l and are analytically evaluated by Mellor and Blumberg in their paper. The solution and the coefficient needed for the computation are also reported in [38].

3.1.3 Boundary conditions

Vertical boundary conditions

We already discussed in the first chapters the air-sea interactions that take place at the sea surface. These interactions result in exchange of momentum, heat and condensation or evaporation trough the surface. Adding the precipitation rates and the river runoff to the evaporation rate gives equation (1.1), while taking into account all the heat fluxes that passed through the air-sea interface gives the heat equation (1.2).
In order to parameterize the air-sea processes that take place at the sea surface, the wind stress, the heat fluxes and the evaporation rates are calculated through bulk formulas by means of the atmospheric forcings and of the sea surface temperature predicted by the model. For example, in [38] we find the following formulas for the computation of sensible and latent heat flux at the surface, also used in the AREG2 model:

$$Q_h = \rho_A C_{pA} C_h |\mathbf{U}_{10} - \mathbf{U}_w|(T_{10} - T_W) \tag{3.15}$$

$$Q_c = \dot{E} L \tag{3.16}$$

where $\dot{E}$ and L are the evaporation flux and latent heat of evaporation and can be expressed as:

$$\dot{E} = \rho_A C_E |\mathbf{U}_{10} - \mathbf{U}_w|(q_W - q_{10}) \tag{3.17}$$

$$L = (2.501 - 0.002 T_W/C) \cdot 10^6 \text{ J Kg}^{-1} \tag{3.18}$$

In these latter formulas ρ_A, C_{pA}, T_{10} are the air denisty, specific heat and temperature of the air at 10 m above the water surface and T_W is the water surface temperature; $\mathbf{U}_1 0$ and $\mathbf{U}_W$ are the velocity vectors of the air at 10 meters above the sea surface and of the water at the surface; q_{10} is the air humidity at 10 m height and q_W is the specific humidity (density of water vapor/density of moist air) taken to be the saturation humidity evaluated at the sea surface temperature; C_E is the moisture transfer coefficient.

The reader is referred to the POM user guide [38] for an explanation of the bulk formulas used in the original implementation of the POM.
These formulas results in the following vertical boundary conditions at the free surface $z = \eta$ for momentum, heat and salinity:

$$\rho_0 K_M \left.\frac{\partial(u,v)}{\partial z}\right|_{z=\eta} = (\tau_{wz}, \tau_{wy}) \tag{3.19}$$

$$\begin{aligned} \rho_0 K_H \left.\frac{\partial \theta}{\partial z}\right|_{z=\eta} = \frac{1}{C_p} [&(1 - Tr) Q_s(C) - Q_B(T_a, \theta_{z=\eta}, C, rh) - \\ &- Q_e(T_a, \theta_{z=\eta}, rh, |\bar{\nu}_w|) - Q_h(T_a, \theta_{z=\eta}, |\bar{\nu}_w|)] \end{aligned} \tag{3.20}$$

$$K_H \left.\frac{\partial S}{\partial z}\right|_{z=\eta} = S_{z=\eta}(E - P - R) \tag{3.21}$$

where Tr is the coefficient of penetration of light into the water according to Jerlov, [27], Q_s is the short wave radiation incident to the sea surface, and is calculated through an

astronomical formula according to Reed, [54] and [55], Q_B is the long wave radiation emitted back by the sea and is computed through the formula proposed by May, [60], while Q_c and Q_h are the latent and sensible heat fluxes respectively, calculated according to the formulas given in the POM user guide,[38] and given in (3.15) and (3.16); τ_{wx} and τ_{wy} are the zonal and meridional components of the wind stress produced on the sea surface, computed according Hellerman and Rosenstein, [26].
As we can see, the flux terms in equation (3.20) depend on the cloud cover C, on the air temperature T_a, on the sea surface temperature $\theta|_{z=\eta}$ and on the relative humidity rh, as already mentioned.
Equation (3.21) gives a surface salinity flux, given by the water balance $E-P-R$, where E is the evaporation, P the precipitation and R is the river runoff, multiplied by the model predicted surface salinity $S_{z=\eta}$. R is a non-zero value only at the estuary grid points and is taken from the climatology of Raicich, [52] for all the rivers except for the Po and the Buna/Bojana. The runoff values of the Po are daily means observed at the cross section of Pontelagoscuro, a few tens of kilometers upstream the river's delta, provided by ARPA SIM Emilia-Romagna. The precipitation, together with the other atmospheric forcings is given by the ECMWF forecasts and analyses, at a horizontal resolution of 0.5 degrees and updated every 6 hours.

To complete the set of the vertical boundary condition at the surface, we must add to (3.19), (3.20) and (3.21) the following condition for the surface vertical velocity, turbulent kinetic energy and turbulent length scale:

$$w|_{z=\eta} = \left(\frac{\partial \eta}{\partial t} + (u,v)\cdot\nabla\eta\right)\Big|_{z=\eta} \tag{3.22}$$

$$q^2 = B_1^{2/3} u_{\tau s}^2 \tag{3.23}$$

$$q^2 l = 0 \tag{3.24}$$

where B_1 is one of the turbulence closure constants and $u_{\tau s} = \sqrt{\frac{|\tau_w|}{\rho_0}}$ is the friction velocity at the sea surface.

At the bottom the boundary condition for the continuity equation (3.3) and hydrostatic approximation (3.7) results in:

$$w|_{z=H} = -\left(u\frac{\partial H}{\partial x} + v\frac{\partial H}{\partial y}\right)\Big|_{z=-H} \tag{3.25}$$

$$\rho_0 K_M \left(\frac{\partial (u,v)}{\partial z}\right)\Big|_{z=-H} = (\tau_{bx}, \tau_{by}) \tag{3.26}$$

where H is the water depth, τ_{bx} and τ_{by} are the zonal and meridional components of the bottom stress, parameterized as:

$$(\tau_{bx}, \tau by) = C_d \rho_0 |\mathbf{U}|(u,v) \tag{3.27}$$

where ρ_0 and $\mathbf{U}$ are always the reference seawater density and the water velocity. C_d is the bottom drag coefficient expressed as:

$$C_d = \max\left\{0.0025; \left[\frac{k}{\ln\frac{\delta_b}{z_0}}\right]^2\right\} \tag{3.28}$$

Here k is the Von Karman constant (equal to 0.40), z_0 is the bottom roughness length scale and δ_b the thickness of the last model level above the bottom.
The remaining boundary conditions at the bottom are:

$$K_H\left(\frac{\partial\theta}{\partial z}, \frac{\partial S}{\partial z}\right) = 0 \tag{3.29}$$

$$q^2 = B_1^{2/3} u_{\tau b}^2 \tag{3.30}$$

$$q^2 l = 0 \tag{3.31}$$

where $u_{\tau b} = \sqrt{\frac{|\tau_b|}{\rho_0}}$ is the friction velocity associated to the bottom stress τ_{bx}, τ_{by}.

The advantage in casting the equations and conditions so far discussed in σ-coordinates is particular evident for the vertical boundary conditions on the velocity field, both at the surface and at the bottom. In the POM coordinate system the component of the velocity field are transformed according to:

$$\mathbf{U} = (u, v, z) \quad \rightarrow \quad \mathbf{U}^* = (u*, v*, \omega) \tag{3.32}$$

where the asterisk denotes the velocity field in the σ-coordinates system. Thus the vertical velocity is substituted with the new velocity component ω in the POM coordinates system, indicating the velocity perpendicular to surfaces with constant σ. Being the sea surface and the bottom such surfaces with respectively $\sigma = 0$ and $\sigma = -1$, we can simply write down a no-flow boundary condition:

$$\omega|_{\sigma=-1} = \omega|_{\sigma=0} = 0 \tag{3.33}$$

Lateral boundary conditions

At the lateral walls of the basins there is no advection of heat and salinity and the non-slip boundary condition are applied for velocity:

$$\mathbf{U}|_{\partial\Omega} = 0 \tag{3.34}$$

$$\left.\frac{\partial S}{\partial \mathbf{n}}\right|_{\partial\Omega} \tag{3.35}$$

$$\left.\frac{\partial \theta}{\partial \mathbf{n}}\right|_{\partial\Omega} \tag{3.36}$$

where $\partial\Omega$ express the later boundary of the basin and $\mathbf{n}$ the normal to the walls.

At the open boundary (OB) at 39 ° N the AREG2 model is nested in the OGCM model of the Mediterranean Sea implemented in MFS. The nesting conditions derives from the formulation by Flather in [19] and generalized by Oddo and Pinardi in [41] and is based on the following procedure:
Take the incompressible continuity equation 3.3 for the nested (AREG2) and the nesting (Mediterranean OGCM) models. Integrating the continuity equation on the entire water column we obtain:

$$\frac{\partial \eta}{\partial t} + \nabla \cdot [(H + \eta)\vec{V}_{BT}] = 0 \tag{3.37}$$

where $\vec{V}_{BT}$ is the horizontal barotropic flow field. defined as:

$$\vec{V}_{BT} = \frac{1}{H + \eta} \int_{-H}^{\eta} (u, v) dz \tag{3.38}$$

The hypothesis that both the nesting and the nested model conserves the mass as written in (3.37) is reasonable. Thus it will be possible to equate the two mass conservation equations obtaining:

$$\nabla \cdot [(H_c + \eta_c \bar{V}_{BTc})] + \frac{\partial \eta_c}{\partial t} = \nabla \cdot [(H_f + \eta_f \bar{V}_{BTf})] + \frac{\partial \eta_f}{\partial t} \tag{3.39}$$

where the subscripts c and f refer to the coarser and the finer resolution model respectively: the MFS-OGCM and AREG2 in our case. Assuming, following [41], that the surface elevation tendency can be expressed with a radiation equation like:

$$\frac{\partial \eta}{\partial t} = -\nabla \cdot (\vec{C}\eta) \tag{3.40}$$

inserting this definition in (3.39) and assuming that for both model C is equal to the linearized free gravity wave speed $\sqrt{gH}$, that $H_c = H_f = H$[1] and that $|\eta << H|$ we obtain the following condition on the barotropic component normal to the boundary:

$$V^n_{BTf} = V^n_{BTc} - \frac{\sqrt{gH}}{H}(\eta_c - \eta_f) \tag{3.41}$$

The radiative equation holds for temperature and salinity too, but in order to avoid for inconsistency between the models a nudging term is added to the equations in writing down the boundary conditions. The full derivation can be found in [41] and leads to the following boundary condition for the temperature and salinity:

$$\frac{\partial(\theta_f, S_f)}{\partial t} + C_x \frac{\partial(\theta_f, S_f)}{\partial x} + C_y \frac{\partial(\theta_f, S_f)}{\partial y} = -\frac{1}{\tau}[(\theta_f, S_f) - (\theta_c, S_c)] \tag{3.42}$$

[1] In practice one has to interpolate the barotropic velocity field of the MFS-OGCM on the finer grid of the AREG2 model. Doing this one then uses $H = H_f$ in equation (3.41). Similar argument hold for equation (3.42) too

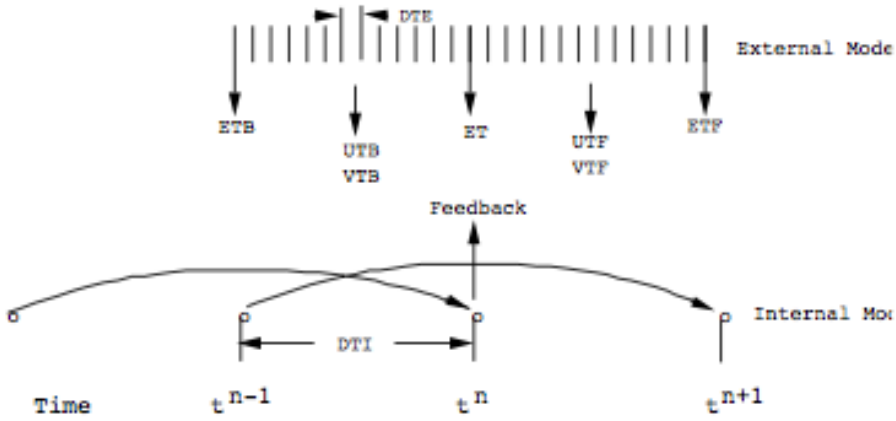


Figure 3.4: *A simplified illustration of the interaction of the External Mode and the Internal Mode of the POM model. The former uses a short time step, dte, whereas the latter uses a long time step, dti. The external mode primarily provides the surface elevation to the internal mode whereas, as symbolized by "Feedback", the internal mode provides intergrals of momentum advection, density integrals and bottom stress to the external mode.*

Here τ is the nudging time scale for the nudging term, C_x and C_y are the horizontal phase speeds. The latter may be calculated from the surrounding interior (in the Adriatic) grid point:

$$C_x = -\partial(\theta_f, S_f)\frac{\partial(\theta_f, S_f)/\partial x}{(\partial(\theta_f, S_f)/\partial x)^2 + (\partial(\theta_f, S_f)/\partial y)^2} \tag{3.43}$$

$$C_y = -\partial(\theta_f, S_f)\frac{\partial(\theta_f, S_f)/\partial y}{(\partial(\theta_f, S_f)/\partial x)^2 + (\partial(\theta_f, S_f)/\partial y)^2} \tag{3.44}$$

3.1.4 Numerical integration

Numerical schemes and grid arrangement

The dynamics following the solution to the governing equation is a superpositions of internal (barotropic) and external (baroclinic) modes. Typical examples of the first are external gravity waves, while examples of the latter are the internal gravity waves, moving slowly than the firsts. In order to optimize the computation procedure in the POM model (and in AREG2 too) the determination of the state vector is separated into two different modes, corresponding to the barotropic and baroclinic modes, having different time steps reflecting the different time scales that characterize the two modes. A schematic representation of the two modes is given in figure 3.4. Both modes solves the respective equations through finite difference numerical techniques.

The external mode has a shorter time step dte and calculates the surface elevation and the vertically integrated velocity components by solving the barotropic equations with a

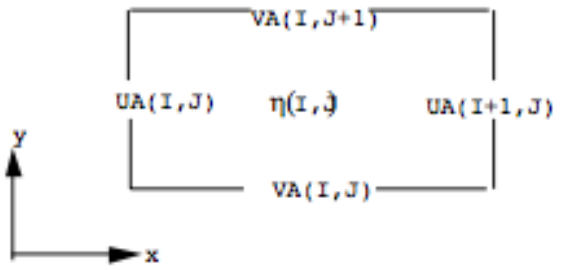


Figure 3.5: *The two dimensional external mode grid of POM. UA and VA are vertically averaged velocity components and η still represent the surface elevation*

leap frog (thus *explicit*) method.
The internal mode has a longer time step dti and solves the baroclinic equations for velocity, temperature, salinity and the turbulent quantities q and l. The calculation of these three dimensional variables is separated into a vertical diffusion numerical scheme and an advection plus horizontal diffusion one. The former is *implicit* (to accommodate small vertical spacing near the surface) whereas the latter is *explicit*.
During the "Feedback" process the external modes is provided with integrals involving the baroclinic forcing and the advective terms along with bottom stress. Their values are held constant during $t^n < t < t^{n+1}$. The external mode solves the barotropic equations until $t = t^{n+1}$ and then the external mode calculates the time average over the time step dti of the external variables. In figure 3.4 etb, utb and vtb are time averages over dti of surface elevation and external velocities at the previous time step; etf, utf and vtf are the same thing for the subsequent time step.
Since vertical integrals of internal and external variables may differ, the internal velocity u has to be adjust so that its vertical integral is the mean of utb and utf. The same adjustment holds for v.
A similar procedure is carried to match the internal and external surface elevation. This point is also quite delicate as an incorrect adjustment may results in numerical instabilities.

The numerical integrations explained above are implemented on an Arakawa C, staggered grid, as in figures 3.5 and 3.6. A three dimensional view of the Arakawa grid for both the internal and external mode is given in figure 3.7.
So, considering the horizontal grid we can see that:

- η, H, θ, S, ρ are evaluated at the center of the cell;
- The three dimensional velocity field and the vertically averaged horizontal velocities are calculated on the edge of the cell.

Remember that while the horizontal grid steps (Δx and Δy in figure 3.7) are constant in the AREG2 domain, the vertical spacing ($\Delta\sigma$ on the same figure) is variable in order to provide a better resolution in shallow water areas and near the vertical boundaries.

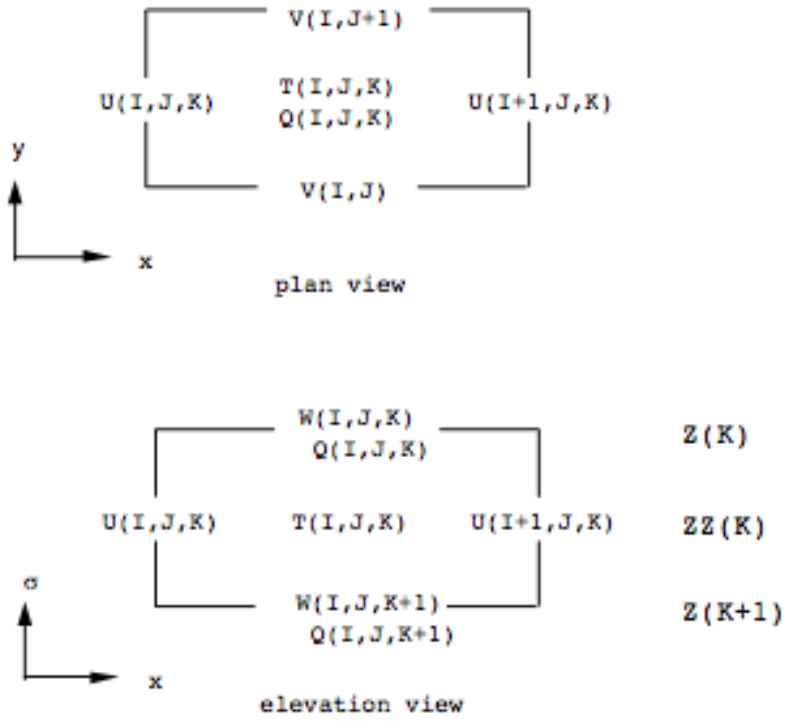


Figure 3.6: *The three dimensional internal mode grid of POM. Q represents* q^2, q^2l,K_M *and* K_h, *T represents* θ, *S and* ρ. *Z and ZZ are the vertical coordinates of the cell vertices and centers. U, V and W are velocity components*

Numerical stabilities

In the leap frog time differencing scheme, the solutions may diverge. In POM this problem is solved by the application of a filter that at every time step smooths the solution:

$$F_s = F^n + \frac{\alpha}{2}(F^{n+1} - 2F^n + F^{n-1}) \tag{3.45}$$

where F is any of the variables calculated by the model, the subscript s indicates the solution after the smoothing, the superscripts denotes the time step and α is a smoothing parameter, usually set to 0.05. After the smoothing F_s is reset to F^{n-1} and F^{n+1} to F^n.

The time step of the internal and external integration scheme is limited by the Courant-Friedrichs-Levy (CFL) computational stability condition. This condition is, for the external mode:

$$\Delta t_E \leq \frac{1}{C_t}\left|\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right|^{-1/2} \tag{3.46}$$

where $C_t = 2(gH)^{1/2} + U_{max}$; U_{max} is the expected, maximum velocity and $\sqrt{gH}$ is the external wave speed. There are other restrictions but in practice the CFL limit is the most

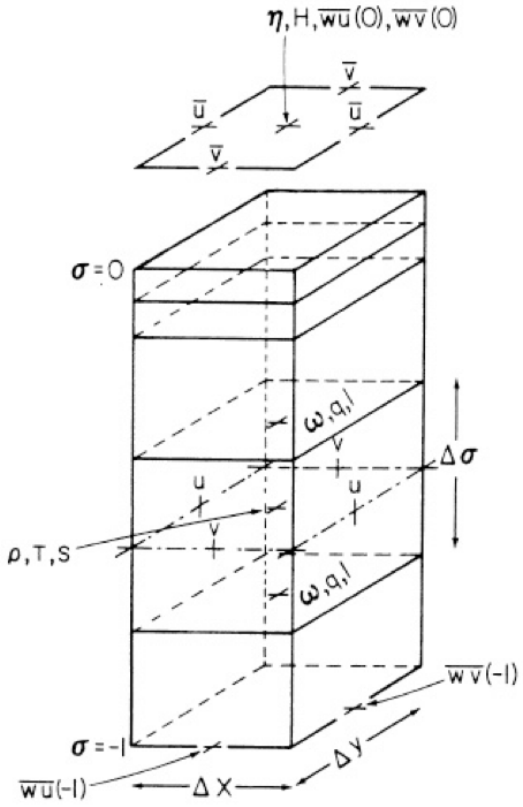


Figure 3.7: *Arakawa C grid, adapted from [6]. Δx and Δy are the horizontal spacings on the grid and $\Delta\sigma$ is the vertical increment. The other symbols have already been presented.*

stringent. The model time step is usually 90 % of this limit.
The internal mode has a much less stringent time step since the fast moving external mode effects have been removed. For the internal scheme the CFL condition is:

$$\Delta t_I \leq \frac{1}{C_T} \left| \frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} \right|^{-1/2} \tag{3.47}$$

with $C_T = 2C + U_{max}$; C is the maximum velocity for internal waves and U_{max} here is the maximum advection velocity.
In the AREG2 model implemented in this study, the ratio $\Delta t_I / \Delta t_E$ is equal to 100.
Other limitations on the internal time step imposed by horizontal diffusion of momentum or scalars and by the Earth rotation can be found in [38] but are less stringent than (3.46)

3.2 The 3D-Var Data Assimilation

3.2.1 Basis of DA

In mathematical terms (following [30] and [7]) let us assume we have a prognostic model of our system, that provides a model state $\mathbf{x}$, a vector with dimension N_x which is a discrete representation of the state of the sea. There are many ways of representing the model state. The fields themselves can be represented as grid-point values (i.e. averages of the fields inside grid boxes), spectral components, EOF values, finite-element decomposition. All these transforms do not change the analysis problem, only its representation. In any case the definition of some pre-defined discrete basis enable us to use simpler matrix methods, rather than methods based on function theory. In our case, we deal with a grid-point representation in which the physical variables ($\mathbf{U}$,θ,S,η and the other variables described in the previous sections) are directly represented in a discrete space.
In dealing with oceanographic numerical models, both space and time are usually discretized, so that we can integrate the governing equation in order to produce the model state vectors $\mathbf{x}_i$ at a a discrete time t_i, being $i \in \{0, 1, 2, \ldots, n\}$. In particular, during a forecast, the model produces the future model states $\mathbf{x}_i$, $i \neq 0$ from the initial condition $\mathbf{x}_0$. We can formally express the action of the model as a possibly non-linear operator M such that:

$$\mathbf{x}(i) = M_{i-1,i}(\mathbf{x}(i-1)) \tag{3.48}$$

We also have a set of observations of certain quantities provided in the model state, represented by a vector $\mathbf{y}_o$ of length N_y; typically $N_y << N_x$. In this sense we have an underdetermined problem.
We assume that the relation between observable values $\mathbf{y}$ and a model state $\mathbf{x}$ at a given time is represented by an explicit (and possibly non-linear) operator H so that

$$\mathbf{y} = H(\mathbf{x}) \tag{3.49}$$

is the estimate for $\mathbf{y}$ given the model state $\mathbf{x}$. The best possible estimate of the observation would be the projection in the observation space of the *true model state* $\mathbf{x}_t$:

$$\mathbf{y}_t = H(\mathbf{x}_t) \tag{3.50}$$

One has to keep in mind that $\mathbf{x}_t$ is not a vector containing the real value of the variables at the points specified by the model, but is the *best* possible representation achievable by the model considered (see [7]).
The *observational* or *forward operator* H is in principle a time dependent, non-linear operator, but often it is considered as a simple interpolation operator.
The problem we are trying to solve is to find the best analyzed state vector $\mathbf{x}_a$ (which will never be exactly equal to $\mathbf{x}_t$) that inverts (3.49) for a given set of observations $\mathbf{y}_o$ allowing for observation errors and other prior information, namely a first guess of the state vector, called the *background* state vector $\mathbf{x}_b$, and the informations about the statistical confidence we put in this *a priori* statement.

The problem can be handled in many ways, but here we will follow the most common, bayesian approach proposed, for example, in [30]. To do this we have to introduce the appropriate probability density functions (PDFs) and write our version of Bayes' theorem:

$$p(\mathbf{x}|\mathbf{y}_o) = \frac{p(\mathbf{x})p(\mathbf{y}_o|\mathbf{x})}{p(\mathbf{y}_o)} \tag{3.51}$$

Equation (3.51) states how our knowledge on the state of the ocean, expressed by a PDF on $\mathbf{x}$, changes after adding the information from the observation $\mathbf{y}_o$. $p(\mathbf{x})$ is the prior distribution: our knowledge based on previous observations. $p(\mathbf{x}|\mathbf{y}_o)$ is the posterior PDF, after adding the information $\mathbf{y}_o$. $p(\mathbf{y}_o|\mathbf{x})$ is the PDF of getting the observation $\mathbf{y}_o$ given the model state $\mathbf{x}$. Note that this is a density function in the $\mathbf{y}$ space, not in $\mathbf{x}$ space. In fact it does note integrate to 1 over $\mathbf{x}$: it is called the *likelihood function* for $\mathbf{x}$. Finally $p(\mathbf{y}_o) = \int p(\mathbf{y}_0|\mathbf{x})p(\mathbf{x})d\mathbf{x}$ is the PDF of the observations, usually regarded as a normalization factor.

We have now to define the shape of the PDFs in equation (3.51).The usual and easiest way to proceed is introducing normal distributions. The common assumption is to use normal distributions. So the prior PDF becomes:

$$p(\mathbf{x}) = N(\mathbf{x}|\mathbf{x}_b, \mathbf{B}) = (2\pi|\mathbf{B}|)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\right] \tag{3.52}$$

where $\mathbf{x}_b$ is the best prior estimate of the vector $\mathbf{x}$ and $\mathbf{B}$ the corresponding covariance matrix.
The likelihood function has to be parametrized considering that a real observation has its instrumental error, in our case a gaussian error around a true value $\mathbf{y}_t$:

$$p(\mathbf{y}_o|\mathbf{y}_t) = N(\mathbf{y}_o|\mathbf{y}_t, \mathbf{E}) = (2\pi|\mathbf{E}|)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y}_o-\mathbf{y}_t)^T\mathbf{E}^{-1}(\mathbf{y}_o-\mathbf{y}_t)\right] \tag{3.53}$$

The expression above states that the observed $\mathbf{y}_o$ is normally distributed around the true observable value $\mathbf{y}_t$ with covariance matrix $\mathbf{E}$, which is a function of the instrument characteristics. Actually the assumption that we can model the instrumental error with a normal distribution deserves some discussion because in this way the presence of observations with large errors can be disastrous for the algorithm: since the gaussian distribution has practically no tails, the presence of an observation far away from the expectation value is considered as a meaningful information rather than an outlier. The consequence is that the entire analysis is shifted towards an outlier having the same weight as other observations, while it shouldn't.
It is the purpose of a good quality control algorithm to recognize gross errors and treat them properly, either by rejection or by incorporating them in a statistical consistent way in the data assimilation algorithm. The latter choice alter the PDF for the observations and are no more normal distributions.
For a discussion about the gross errors handling and the gaussian assumption about observation distribution see also [30], [33], [58] and [51].

Now, the probability of a perfect instrument observing $\mathbf{y}_t$ given the model state $\mathbf{x}$ is assumed to be:

$$p(\mathbf{y}_t|\mathbf{x}) = N(\mathbf{y}_t|\mathbf{x}, \mathbf{F}) = (2\pi|\mathbf{F}|)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y}_t - H(\mathbf{x}))^T \mathbf{F}^{-1}(\mathbf{y}_t - H(\mathbf{x}))\right] \tag{3.54}$$

The latter is also called the the *forward* modeling probability distribution, from where comes the covariance matrix $\mathbf{F}$ which is a function of the model resolution. This distribution contains information about how the true value are represented by the model.
To get the likelihood function we have to convolve the two previous PDFs to get the combined observational PDF:

$$\begin{aligned} p(\mathbf{y}_o|\mathbf{x}) &= \int p(\mathbf{y}_o|\mathbf{y}_t) p(\mathbf{y}_t|\mathbf{x}) d\mathbf{y}_t \\ &= N(\mathbf{y}_o|H(\mathbf{x}), \mathbf{R}) = \\ &= (2\pi|\mathbf{R}|)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y}_o - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x}))\right] \end{aligned} \tag{3.55}$$

where $\mathbf{R} = \mathbf{E} + \mathbf{F}$ is the *representativeness* covariance matrix. The form of the likelihood (3.55) and in particular of the covariance $\mathbf{R}$, tells us that the uncertainty in the representation of the observations $\mathbf{y}_o$ given the model $\mathbf{x}$ won't be zero even if we measured the variable $\mathbf{y}_t$ with a perfect instrument. This is because the contribution of $\mathbf{F}$ in the uncertainty expressed in $\mathbf{R}$: the model will *never* be able to perfectly represent the real state of the ocean so the forward operator will never be able to perfectly simulate the true observation $\mathbf{y}_t$.
Generally speaking the matrix $\mathbf{R}$ is, in most cases, diagonal, or block-diagonal, because there is no reason to assume observation error correlations between independent observing networks, observing platforms or station, and instruments, except some special cases.

Finally, recalling equation (3.51), the posterior PDF is found to be:

$$\begin{aligned} p(\mathbf{x}|\mathbf{y}_o) &= \frac{p(\mathbf{x}) p(\mathbf{y}_o|\mathbf{x})}{p(\mathbf{y}_o)} \\ &\propto N(\mathbf{y}_o|H(\mathbf{x}), \mathbf{R}) N(\mathbf{x}|\mathbf{x}_b, \mathbf{B}) \\ &\propto \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - \frac{1}{2}(\mathbf{y}_o - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x}))\right] \end{aligned} \tag{3.56}$$

The solution we are looking for is the one that maximizes the posterior PDF (3.56). Note that the PDF obtained is *not* in general a gaussian distribution because the observational operator H is not, in general, linear. However it usually makes physical sense to assume that H is linearizable around the background value $\mathbf{x}_b$ and a matrix $\mathbf{H}$ is introduced defined as the first derivatives of H with respect to the components of $\mathbf{x}$, so that:

$$H(\mathbf{x}) \simeq H(\mathbf{x}_b) + \mathbf{H}(\mathbf{x} - \mathbf{x}_b) \tag{3.57}$$

And the posterior PDF (3.56) can be re-written as:

$$p(\mathbf{x}|\mathbf{y}_o) \propto \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - \frac{1}{2}(\mathbf{y}_o - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y}_o - \mathbf{H}\mathbf{x})\right] \tag{3.58}$$

Other typical hypothesis are the following:

- $\mathbf{B}$ and $\mathbf{R}$ are positive definite matices
- The expectation in the background and observation errors is zero, meaning they are unbiased, i.e.: $\mathbf{x}_b \bar{-} \mathbf{x}_t = \mathbf{y}_o - \bar{H}(\mathbf{x}_t) = 0$
- The errors in the background state and in the observations are uncorrelated, i.e. $(\mathbf{x}_b - \mathbf{x}_t)(\bar{\mathbf{y}} - H(\mathbf{x}_t))^T = 0$

For a full discussion on the assumptions usually made in deriving the solution, refer to [7].

If the forward operator H is exactly linear (as in the case of an interpolation operator) equation (3.58) is exactly true and the posterior PDF is also a gaussian. This is easily seen by taking the logarithm of (3.58):

$$\begin{aligned} -ln(p(\mathbf{x}|\mathbf{y}_0)) =& \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \\ &+ \frac{1}{2}(\mathbf{y}_o - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y}_o - \mathbf{H}\mathbf{x}) + constant \end{aligned} \tag{3.59}$$

which is also called the *misfit* function. In the case of a linear forward operator the misfit (3.59) is a quadratic funtion of $\mathbf{x}$, so the posterior PDF is a gaussian that could be written in terms of an a posterior expected value $\mathbf{x}_a$ and a covariance matrix $\mathbf{A}$:

$$p(\mathbf{x}|\mathbf{y}_o) \propto \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{x}_a)^T \mathbf{A}^{-1}(\mathbf{x} - \mathbf{x}_a)\right] \tag{3.60}$$

where

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y}_o - H(\mathbf{x}_b)) \tag{3.61}$$

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T \tag{3.62}$$

where the matrix $\mathbf{K}$, called the *gain* is defined as follows:

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \tag{3.63}$$

So our best analysis is $\mathbf{x}_a$ defined by (3.61) and the covariance matrix is reduced, after adding the information from the observations, from $\mathbf{B}$ to $\mathbf{A}$.
Equations (3.61), (3.62) and (3.63) define the *least squares estimation*, also called the *Best Linear Unbiased Estimator (BLUE)*. This solution has been justified for a gaussian, linear (in H) problem and in the case, the BLUE analysis represent the maximum likelihood estimation for $\mathbf{x}_t$.
In general, non-linearities are accounted for in equation (3.61) and the solution is an optimal estimation in the sense that minimizes the root mean square error between the analysis and the true state. The BLUE analysis is in fact equivalent to the minimization of (3.59), or equivalently, of the following cost function:

$$\begin{aligned} J(\mathbf{x}) =& (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (\mathbf{y}_o - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x})) \\ =& J_b(\mathbf{x}) + J_o(\mathbf{x}) \end{aligned} \tag{3.64}$$

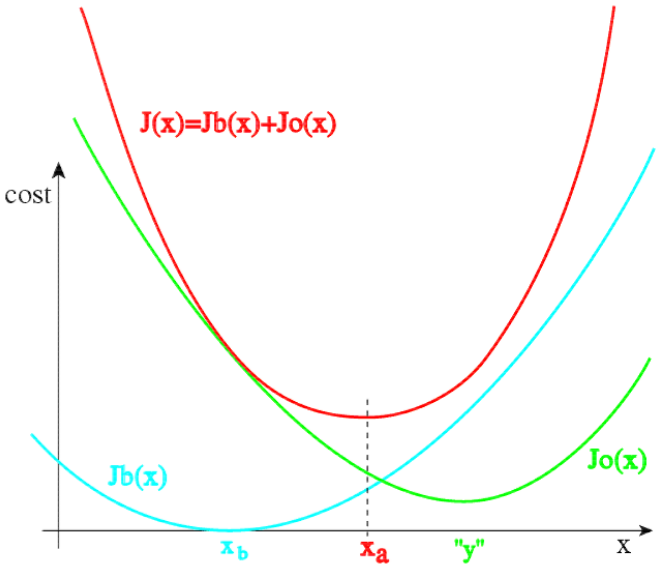


Figure 3.8: *Schematic representation of the variational form of the least-squares analysis, in a scalar system where the observation* $\boldsymbol{y}$ *is in the same space as the model* $\boldsymbol{x}$*: the cost-function terms* J_b *and* J_o *are both convex and tend to "pull" the analysis towards the background* $\boldsymbol{x}_b$ *and the observation* $\boldsymbol{y}$*, respectively (from [7])*

This is a sum of distances defined in the L_2 space from the covariances matrix $\mathbf{B}$ and $\mathbf{R}$ and the *variational optimization problem* equivalent to the BLUE solution is:

$$\begin{aligned} \mathbf{x}_a &= \text{Arg min}[J] \\ J(\mathbf{x}) &= J_b(\mathbf{x}) + J_o(\mathbf{x}) \end{aligned} \tag{3.65}$$

For linearizable H we can solve it explicitly, giving the least squares solution This case is shown in figure (3.8). What we obtained in (3.61) is the best estimate of the state of the ocean given a previous estimate $\mathbf{x}_b$ and observations $\mathbf{y}_o$. Remember that our goal here was to correctly initialize a forecast by giving the best estimate of the present state of the sea, $\mathbf{x}_a$.

3.2.2 A digression on DA techniques

We will now briefly discuss practical implementation and extension of solution (3.61), (3.62). A summarized history of the main data assimilation techniques used in meteorology and oceanography is given in figure 3.9.

Three dimensional DA schemes: OI and 3D-Var

The direct implementation of the BLUE solution is known as the *Optimal Interpolation* (**OI**) solution. To design a practical assimilation scheme we have to calculate the elements

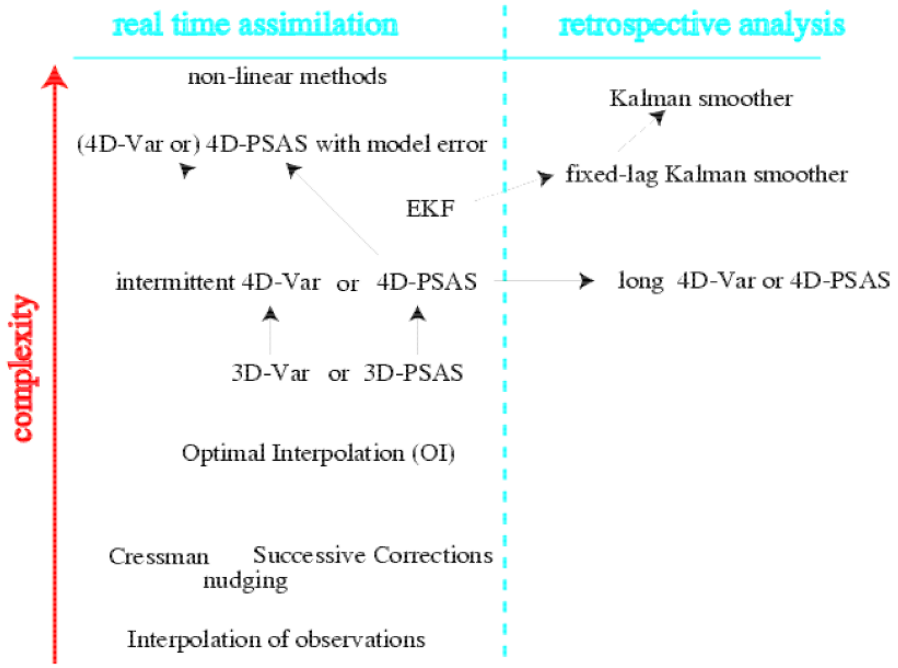


Figure 3.9: *Summarized history of the main DA algorithms used in meteorology and oceanography. The classification is roughly based on the complexity of the method. From [7]*

of the matrix $\mathbf{K}$, the hardest part of which is the specification of the matrix $\mathbf{B}$. Typically some simplification is required in order to reduce the computational cost of the assimilation.
The assimilation algorithm previously proposed for the AFS and tested in [23] is based on a Reduced order OI Method proposed by De Mey and Benkiran in [12]. We now briefly describe the algorithm because of the historical interest it has in the framework of this work.
In the resulting algorithm, also called **SOFA** (System for Ocean Analysis and Forecast), the background error covariance matrix is approximated introducing the *Empirical Orthogonal Functions* (EOFs) and a separation between the vertical and horizontal structure of $\mathbf{B}$. In particular, since we assume that the background error covariances matrix is positive definite, it can be decomposed in eigenvalues and eigenvectors:

$$\mathbf{B} = \mathbf{V}^T \Delta_r \mathbf{V} \tag{3.66}$$

where $\mathbf{V}$ is a matrix whose columns are 3-D eigenvectors, also called 3-D EOFs of $\mathbf{B}$ and Δ is a diagonal matrix whose diagonal elements are the eigenvalues. So, at any time, we can write the decomposition of $\mathbf{B}$ over a set of pre-calculated (based on climatological data, for example) orthonormal vectors $\mathbf{V}$, any of which express a particular mode of the ocean 3-D variability. The simplification proposed in [12] is then to "truncate" the matrix $\mathbf{V}$ to the dominant modes, using an external criterion as well as physical insights, thus

substituting in (3.66) a truncated matrix $\tilde{\mathbf{V}}$ to the full matrix $\mathbf{V}$.
It is possible to proceed further by observing that the ocean large scale state variables can be separated into vertical and horizontal modes. So it is possible to write:

$$\mathbf{B} \simeq \mathbf{S}^T \mathbf{B}\mathbf{r}\mathbf{S} \tag{3.67}$$

where $\mathbf{S}$ is a matrix containing only vertical multivariate eigenvectors or v-EOFs, and $\mathbf{Br}$ contains the eigenvalues and the horizontal covariance structures associated with each v-EOF. This decomposition make $\mathbf{B}$ singular since (3.67) is not a perfect equality. Following the demonstration of [12] and the comments in [23] we may reduce the order of the scheme by retaining only the dominant v-EOFs in $\mathbf{S}$, truncating it to $\tilde{\mathbf{S}}$.This is allowed because in the ocean the representa- tive vertical modes are fewer than the number of vertical levels chosen for the numerical model. The ocean is "low order" in the vertical and this allows the order reduction.
Then, the reduced order Kalman gain for this particular OI scheme is:

$$\mathbf{K}^{ROOI} = \tilde{\mathbf{S}}^{-1} \mathbf{B}\mathbf{r} \widetilde{\mathbf{H}\mathbf{r}}^T (\widetilde{\mathbf{H}}^T \mathbf{B}\mathbf{r} \widetilde{\mathbf{H}}^T + \mathbf{R}\mathbf{r})^{-1} \tag{3.68}$$

where $\widetilde{\mathbf{H}\mathbf{r}} = \mathbf{H}\tilde{\mathbf{S}}$ is the tangent linear observation operator projected onto the vertical modes and $\mathbf{Rr}$ is the reduced order observational error coariance matrix referred to the vertical modes. The analysis is then:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}^{ROOI}(\mathbf{y}_o - H(\mathbf{x}_b)) \tag{3.69}$$

where $\mathbf{x}_a$ is then the improved estimate of the initial condition for a new forecast; $\mathbf{x}_b$ is the previous estimate for $\mathbf{x}_a$ and is given by the previous forecast of the model state; $\mathbf{y}_o$ are the data used in the analysis.
In an operational network the analysis (3.69) (or, in general (3.61)) is used as initial condition for the new forecast. The forecasted state vector will be used as the background state vector for the successive DA process that will make use of the future observations in order to produce a new, optimal initial condition to initialize a successive forecast and so on.

Advantages of the SOFA (and OI in general) assimilation scheme are its simplicity of implementation and its relative small cost if the right assumptions can be made.
Main drawbacks of this kind of methods are in the algebraic cost of evaluating the Kalman gain $\mathbf{K}$.
If the observational operator is non linear the tangent linear hypothesis may not hold anymore and the OI solution is far from optimal. In this case solution the solutions (3.61), (3.69) may be applied iteratively in an empirical attempt to make the analysis more optimal. This is called the *incremental method* and for strongly non linear cases there is no guarantee that the analysis will be a good one. Other method can be used in this case, like the *simulated annealing method* and other specific, stochastic techniques.

Variational algorithms have been specifically designed in order to avoid the main drawbacks of the methods of the OI family, also called *sequential* methods. As we have

seen the discussion that leads us to the BLUE solution (3.61), (3.62) that the optimal solution is the one that minimizes the logarithm of the posterior PDF (3.56) or the cost function (3.64). Thus, the general solution to the assimilation problem can be achieved following two different strategies. The first class seeks to find an explicit optimization of the cost function (3.64) like the least square solution.The second class methods aim to find the vector $\mathbf{x}_a$ that minimizes the cost function by solving a problem like (3.65) and are therefore called variational techniques.
The principle, and the great advantage, of the variational methods is exactly to avoid the computation of the gain $\mathbf{K}$, defined in (3.63) by looking for the analysis as an approximate solution to the equivalent minimization problem defined by the cost function J. The variational version of the OI is the 3D-Var, because no time dependence is explicitly inserted in the algorithm
In practice, the solution is sought iteratively by performing several evaluations of the cost function (3.64):

$$\begin{aligned} J(\mathbf{x}) =& (\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{y}_o - H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x})) \\ =& J_b(\mathbf{x}) + J_o(\mathbf{x}) \end{aligned}$$

of its gradient:

$$\nabla J(\mathbf{x}) = 2\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) - 2\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x})) \tag{3.70}$$

and, depending on the optimization method, of the Hessian:

$$J'' = \nabla\nabla J = 2(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) \tag{3.71}$$

For a discussion on the descent algorithms, the reader may refer to [51].
A schematic representation of an optimization method is given in figure 3.10. This kind of methods require a starting point in the $\mathbf{x}$ space. In practice, the initial point of the minimization, or first guess, is taken equal to the background $\mathbf{x}_b$. This is not compulsory, however, so it is important to distinguish clearly between the terms *background* (which is used in the definition of the cost function) and *first guess* (which is used to initiate the minimization procedure). If the minimization is satisfactory, the analysis will not depend significantly on the choice of first guess, but it will always be sensitive to the background.

The popularity of 3D-Var stems from its conceptual simplicity and from the ease with which complex observation operators can be used, since only the operators and the adjoints of their tangent linear need to be provided.
Another advantage is the ability to enforce external weak (or *penalty*) constraints, such as balance properties, by putting additional terms into the cost function, usually denoted J_c, so that $J = J_b + J_o + J_c$. However, this can make the preconditioning of the minimization problem difficult.
It is easy to see that a block-diagonal $\mathbf{R}$ implies that J_o is a sum of N scalar cost functions $J_{o,i}$, each one defined by a submatrix $\mathbf{R}_i$ and the corresponding H_i and $\mathbf{y}_i$ of the

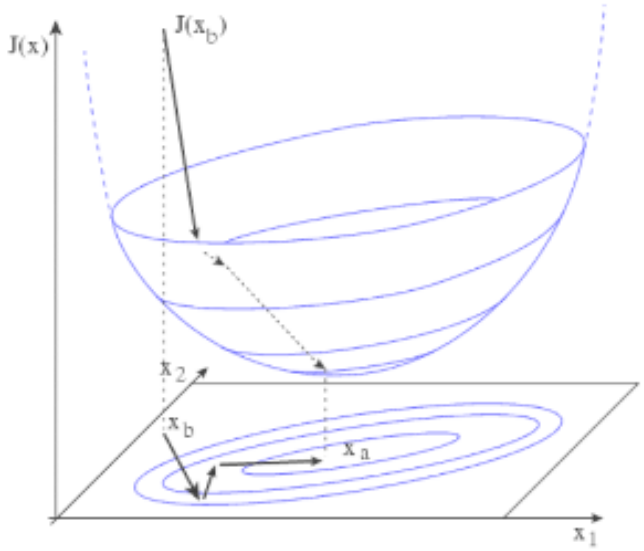


Figure 3.10: *Schematic representation of the variational cost function minimization (here in a two variable model space): the quadratic cost function has the shape of a parabolid with the minimum at the optimal analysis $\boldsymbol{x}_a$. The minimization works by performing several line-searches to move the control variable $\boldsymbol{x}$ to areas where the cost function is smaller, usually by looking at the local slope and sometimes at the curvature of the cost function (from [7])*

observation operators and values:

$$\begin{aligned} J_o &= \sum_{i=1}^{N} J_{o,i} \\ J_{o,i} &= (\mathbf{y}_i - H(\mathbf{x}))^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H(\mathbf{x})) \end{aligned} \tag{3.72}$$

The breakdown of J_o is a useful diagnostic tool of the behaviour of 3D-Var in terms of each observation type: the magnitude of each term measures the misfit between the state $\mathbf{x}$ and the corresponding subset of observations. It can also simplify the coding of the computations of $\mathbf{J}_o$ and its gradient.

For the 3D-Var method too a big difficulty is the need to design a model for $\mathbf{B}$ that properly defines background error covariances for all pairs of variables. In particular it has to be symmetric definite positive and the background error variances must be realistic when expressed in terms of observation parameters, because this is what will determine the weight of the observations in the analysis. In the next section we will describe how this problem is handled in our version of the 3D-Var assimilation algorithm.
Another disadvantage of the variational techniques is that the analysis error is not explicitly calculated, while OI provides the matrix $\mathbf{A}$ without any extra computational cost. For a variational algorithm, dedicated algorithms should be used. for example, the er-

ror covariances of the analysis can be inferred from the matrix of second derivatives, or Hessian, of the cost function thanks to the following result, demonstrated in [7]:

$$\mathbf{A} = \left(\frac{1}{2}J''\right)^{-1} \tag{3.73}$$

If the linearization of the observation operator H can be performed exactly, the cost function is exactly quadratic and J'' does not depend on the value of the analysis: $\mathbf{A}$ can be determined as soon as J is defined, even before the analysis is actually carried out. If the linearization is not exact, $J''(\mathbf{x})$ is not constant: care must be taken in evaluating the Hessian because of, for example, presence discontinuities in the first derivatives of the cost function or saddle points on which $J''(\mathbf{x})$ drop to zero. In practice J must be modified to use the tangent linear of H, which can be acceptable in a close vicinity of $\mathbf{x}_a$.

The four-dimensional analysis problem

The main limitation of the OI and 3D-Var algorithms is that data and background state must refer to the same time: it does not allow for the assimilation of observations distributed in time. In order to do that, a generalization of the methods discussed so far has been developed: the Kalman Filter and the 4D-Var. The equations are the same, provided the observation operators are generalized to include a forecast model that will allow a comparison between the model state and the observations at the appropriate time.

Let's start with the **Kalman Filter**, or KF, which is a well established assimilation techniques, that is usually regarded as a reference in the design of more approximate assimilation algorithms which are being developed nowadays. It is the four dimensional generalization of the least square assimilation methods in the framework of a sequential data assimilation, in which each background is provided by a forecast that starts from the previous analysis.
The same equations that bring us to (3.61) and (3.62) appears in the analysis step of the KF and of its evolutions (*Extended KF*, *Ensemble KF*, *Kalman Smoother*). The notation is the same , except that:

- the background (i.e. forecast) and analysis error covariance matrices are now respectively denoted $\mathbf{P}_f$ and $\mathbf{P}_a$. The background state $\mathbf{x}_b$ is a previous forecast denoted $\mathbf{x}_f$;
- The time index of each quantity is denoted by the index i. The model forecast operator from i dates to $i+1$ is denoted by $M_{i,i+1}$ and the forward operator referred to the data observed at a time i is H_i and the corresponding linearization is $\mathbf{H}(i)$;
- the deviation of the forecast prediction from the true evolution, $M_{i,i+1}(\mathbf{x}_t(i)) - \mathbf{x}_t(i+1)$ is called the model error and we assume that it is not biased and that the model error covariance matrix $\mathbf{Q}(i+1)$ is known;
- the analysis errors $\mathbf{x}_a(i) - \mathbf{x}_t(i)$ and model errors of the subsequent forecast $M_{i,i+1}(\mathbf{x}_t(i)) - \mathbf{x}_t(i+1)$ are assumed to be mutually uncorrelated;

- the variations of the model prediction in the vicinity of the forecast state are assumed to be a linear function of the initial state: for any $\mathbf{x}(i)$ close enough to $\mathbf{x}_a(i)$, $M_{i,i+1}(\mathbf{x}(i)) - M_{i,i+1}(\mathbf{x}_a(i)))\mathbf{M}_{i,i+1}(\mathbf{x}(i)) - \mathbf{x}_a(i+1)$, where $\mathbf{M}$ is a linear operator.

Under the specified hypotheses the optimal way (in the least squares sense) to assimilate sequentially the observations is given by the Kalman filter algorithm defined below by recurrence over the observation times i:

$$\textbf{State forecast :} \quad \mathbf{x}_f(i+1) = M_{i,i+1}(\mathbf{x}_a(i)) \tag{3.74}$$

$$\textbf{Error covariance forecast :} \quad \mathbf{P}_f(i+1) = \mathbf{M}_{i,i+1}\mathbf{P}_a\mathbf{M}_{i,i+1}^T + \mathbf{Q}(i) \tag{3.75}$$

$$\textbf{Kalman gain computation :} \quad \mathbf{K}(i) = \mathbf{P}_f(i)\mathbf{H}^T(i)[\mathbf{H}(i)\mathbf{P}_f(i)\mathbf{H}^T(i) + \mathbf{R}(i)]^{-1} \tag{3.76}$$

$$\textbf{State analysis :} \quad \mathbf{x}_a(i) = \mathbf{x}_f(i) + \mathbf{K}(i)[\mathbf{y}(i) - \mathbf{H}(i)(\mathbf{x}_f(i))] \tag{3.77}$$

$$\textbf{Error covariance of analysis :} \quad \mathbf{P}_a(i) = [\mathbf{I} - \mathbf{K}(i)\mathbf{H}(i)]\mathbf{P}_f(i) \tag{3.78}$$

and the analyses are the sequences of $\mathbf{x}_a(i)$. The proof for this solution can be found in [7].
In the above relations operators $\mathbf{M}$ and $\mathbf{H}$ are always presented as linear operators. If they are not the algorithm written above is called **Extended Kalman Filter** (EKF). Linearization for H are defined in exactly the same way as previously done for the least square solution and $\mathbf{H}$ can be defined as its tangent linear in the vicinity of $\mathbf{x}_b$. Similarly, if M is non-linear, which is the case of most meteorological and oceanographical models, $\mathbf{M}$ can be defined as the *tangent linear forecast model* in the vicinity of $\mathbf{x}_a$, i.e. we assume that for any likely initial state $\mathbf{x}(i)$ (notably $\mathbf{x}_t(i)$):

$$M_{i,i+1}(\mathbf{x}(i)) - M_{i,i+1}(\mathbf{x}_a(i)) \simeq \mathbf{M}_{i,i+1}(\mathbf{x}(i) - \mathbf{x}_a(i)) \tag{3.79}$$

and the realism of this hypothesis must be appreciated using physical arguments.
Note that the linearization of M interacts with the model errors in a possibly complicated way. If non-linearities are important, it may be necessary to include empirical correction terms in the equation, or to use a more general stochastic prediction method such as an ensemble prediction (or *Monte-Carlo*) method, which yields an algorithm known as the **Ensemble Kalman Filter**.
The practical organization of the KF looks like a couple stream of estimations of model states and error covariances (figure 3.11). As can be easily suspected, the computational and storage cost of the FK or EKF analysis is generally high.

The word filter characterizes an assimilation techniques that uses only observation from the past to perform each analysis. An algorithm that uses observations from both past and future is called a *smoother* and is used to analyze a particular time step within the assimilation time window, rather than producing a forecast for subsequent times. There exists a generalization of the KF algorithm for this kind of purposes, therefore called **Kalman smoother**

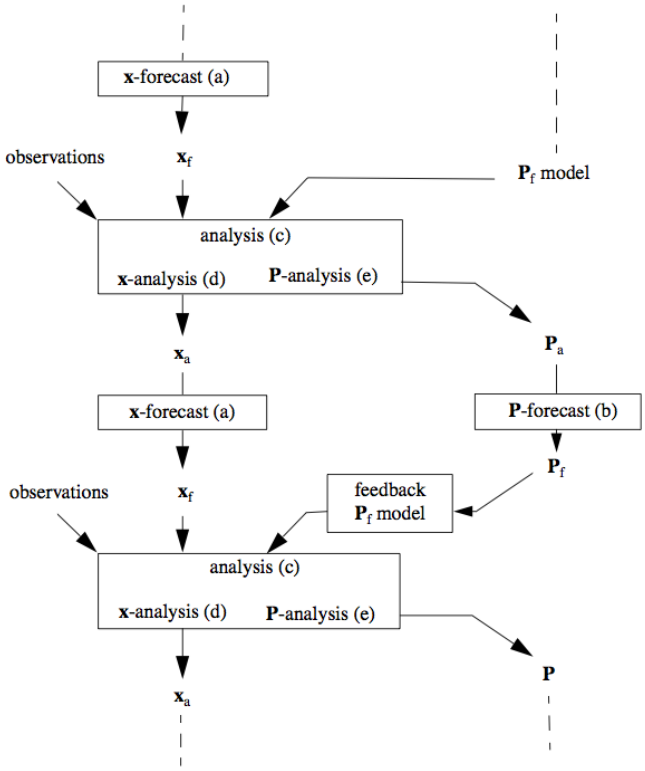


Figure 3.11: *The organization of computations in a KF or EKF assimilation(from [7])*

The variational version of the KF or its Extended version is the **4D-Var** scheme. It is a simple generalization of 3D-Var for observations that are distributed in time.
Over a given time interval, the analysis being at the initial time, and the observations being distributed among n times in the interval, we denote by the subscript i the quantities at any given observation time i. Hence, $\mathbf{y}_i$, $\mathbf{x}_i$ and $\mathbf{x}_t, i$ are the observations, the model and the true states at time i, and $\mathbf{R}_i$ is the error covariance matrix for the observation errors $\mathbf{y}_i - H_i(\mathbf{x}_{t,i})$ The observation operator H_i is linearized ad $\mathbf{H}_i$. The background error covariance matrix $\mathbf{B}$ is only defined at initial time, the time of the background $\mathbf{x}_b$ adn of the analysis $\mathbf{x}_a$.
In its general form, it is defined as the minimization of the following cost function:

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \sum_{i=0}^{N} (\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H_i(\mathbf{x}_i)) \tag{3.80}$$

which can be proven, like in the three-dimensional case detailed previously, to be equivalent to finding the maximum likelihood estimate of the analysis subject to the hypothesis of Gaussian errors.
The first term of (3.80) is clearly the same J_b as the one encountered in the discussion on the 3D-Var, and the remaining part is a four dimensional generalization of J_o.

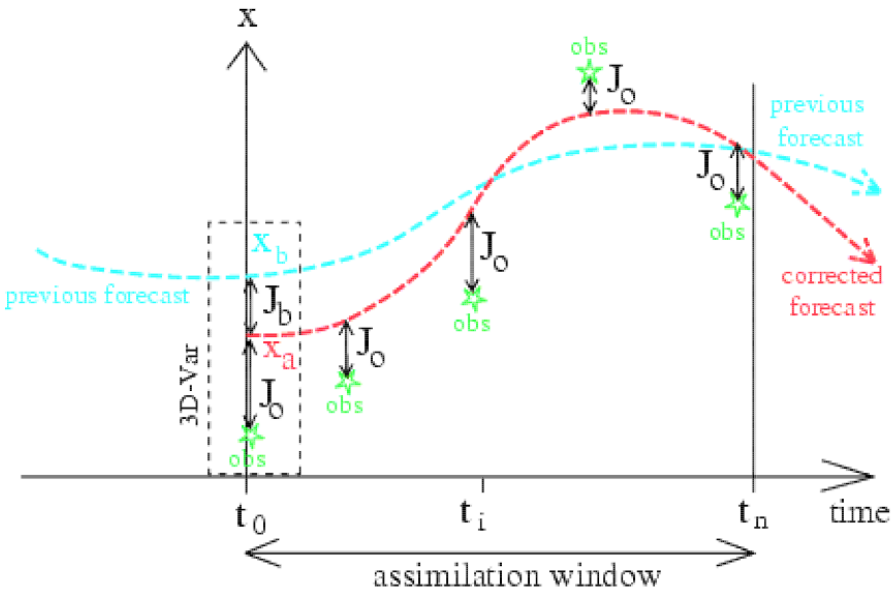


Figure 3.12: *Example of 4D-Var assimilation in a numerical forecasting system. (from [7])*

The 4D-Var analysis, or four-dimensional variational assimilation problem, is by convention defined as the above minimization problem subject to the strong constraint that the sequence of model states $\mathbf{x}_i$ must be a solution of the model equations:

$$\forall i, \quad \mathbf{x}_i = M_{0,i}(\mathbf{x}) \tag{3.81}$$

The optimization problem can be greatly simplified with some hypothesis and simplifications like the tangent linear hypothesis (both for M and for H), and other discussed in [7]. The effect of a 4D-Var assimilation is shown in figure 3.12. The blue line represent the previous forecast given the initial condition $\mathbf{x}_b$. Every time a new observation became available, a 4D-Var is performed using a segment of a previous forecast as background. This updates the model trajectory for subsequent forecast by adjusting the initial condition. At the end of the assimilation the new initial condition $\mathbf{x}_a$ is produced in order to minimize the cost function (3.80). This produce a new, in principle improved forecast for $t > t_n$.

There are many similarities between 4D-VAR and the KF. In particular a theorem exists that states the equivalence of the two algorithms in the case of linear operators, a perfect model and with the same input data. A special property of the 4D-Var analysis in the middle of the time interval is that it uses all the observations simultaneously, not just the ones before the analysis time. It is said that 4D-Var is a *smoothing* algorithm, equivalent to the Kalman smoother algorithm which is a generalization of the Kalman filter, but at a much smaller cost. Other differences between 4D-Var and KF are:

- 4D-VAR can be run for assimilation in a realistic operational framework because it is computationally much cheaper than the KF or EKF;

- 4D-VAR is more optimal than the (linear or extended) KF inside the time interval for optimization because it uses all the observations at once, i.e. it is not sequential, it is a smoother;
- unlike the EKF, 4D-VAR relies on the hypothesis that the model is perfect (i.e $\mathbf{Q} = 0$);
- 4D-VAR can only be run for a finite time interval, especially if the dynamical model is non-linear, whereas the EKF can in principle be run forever;
- 4D-VAR itself does not provide an estimate of $\mathbf{P}_f$, a specific procedure to estimate the quality of the analysis must be applied, which costs as much as running the equivalent EKF.

Despite its elegance, when compared to the sequential (in time) application of a 3D-Var algorithm, the 4D-Var has the following drawbacks:

- it provides the optimal solution in the case that the model is a good (in principle perfect) one. Problems can be expected if the model errors are large. Since the model operator M is typically non-linear and far from perfect, 4D-Var implementation for operational purposes may be hard to achieve;
- it requires the implementation of the operator $\mathbf{M}_i^T$, the so called *adjoint model*. This may be a lot of work if the forecast is complex, slowing down the entire forecasting activity;
- in a real-time system it requires the assimilation to wait for the observations over the whole 4D-Var time interval to be available before the analysis procedure can begin, whereas sequential systems (meaning 3-dimensional techniques implemented in sequence as soon as new observations are available) can process observations shortly after they are available. This can delay the availability of $\mathbf{x}_a$

Because of the above mentioned drawbacks, 4D-Var implementation for ocean forecasting purposes is still a matter of research.

3.3 3D-Var DA in the Adriatic Sea Analysis System

The 3D-Var assimilation scheme we want to test for a future implementation in AFS is taken from a similar algorithm proposed for MFS and presented by Dobricic and Pinardi in [16] and we will give here a full description of it.

Being a 3D-Var assimilation scheme, it iteratively finds the minimum of the cost function (3.64):

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (\mathbf{y}_o - H(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y}_o - H(\mathbf{x}))$$

where the notation is the same as the previous section. The state vector considered in our study contains the following model state variables:

$$\mathbf{x} = [T, S, \eta, u, v]^T \tag{3.82}$$

We now discuss the simplification that made this assimilation scheme suitable for operational purposes.

3.3.1 Cost function simplification

The cost function may be re-written in the following linearized form around the background state:

$$J(\mathbf{x}) = \delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + (\mathbf{H}\delta\mathbf{x} - \mathbf{d})^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \tag{3.83}$$

where $\mathbf{d} = (\mathbf{y}_o - H(\mathbf{x}_b)$ is the misfit, $\mathbf{H}$ is the linearized observational operator evaluated for $\mathbf{x} = \mathbf{x}_b$ and $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$ are the increments (referred to the background state). So we redefined the optimization problem in the field of the increments instead than in the field of the model state and the cost function is now quadratic, a form with a single minimum. Let's now turn to the background covariance matrix $\mathbf{B}$. In order to avoid the inversion of this matrix and preconditioning the minimization problem, it is assumed that $\mathbf{B}$ can be written in the form:

$$\mathbf{B} = \mathbf{V}\mathbf{V}^T \tag{3.84}$$

and, following [31] the cost function is minimized using a new control variable $\mathbf{v}$ defined using the transformation matrix $\mathbf{V}^+$, indicating the generalized inverse of the matrix $\mathbf{V}$:

$$\mathbf{v} = \mathbf{V}^+\delta\mathbf{x} \tag{3.85}$$

this is a transformation form the physical space of the increments in the space of the control variable. The cost function is now written as:

$$J(\mathbf{v}) = \mathbf{v}^T\mathbf{v} + (\mathbf{H}\mathbf{V}\mathbf{v} - \mathbf{d})^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{V}\mathbf{v} - \mathbf{d}) \tag{3.86}$$

So, we end up with an expression for which is no longer necessary to invert the matrix $\mathbf{B}$. The background covariance matrix is only required for the definition of the matrix $\mathbf{V}$, that is the operator for the transformation into the control space. Clearly the information about the background state variability is contained in the operator $\mathbf{V}$.

3.3.2 The transformation matrix factorization

Due to its large size, the matrix $\mathbf{V}$ has to be somehow simplified. The assumption made in [16] is that it is possible to factorize the entire matrix in the following way:

$$\mathbf{V} = \mathbf{V}_D\mathbf{V}_{uv}\mathbf{V}_\eta\mathbf{V}_H\mathbf{V}_V \tag{3.87}$$

where every operator on the right hand side is a linear operator. $\mathbf{V}_V$ transforms coefficients which multiply vertical EOFs into vertical profiles of tem- perature and salinity defined at the model vertical levels, $\mathbf{V}_H$ applies horizontal covariances on fields of temperature and salinity, $\mathbf{V}_\eta$ calculates the sea surface height error covariance from three dimensional fields of temperature and salinity, $\mathbf{V}_{uv}$ calculates velocity from sea surface height, temperature and salinity, and $\mathbf{V}_D$ applies a divergence damping filter on the velocity field.
These operator must contain the same information as the starting matrix $\mathbf{B}$. Let us have a closer look on them.

Vertical covariances

The vertical covariances are represented by multivariate EOFs of surface elevation and vertical profiles of temperature and salinity. Following [15] we can introduce the EOFs in the following way. First of all we must consider that a single set of EOFs is not sufficient for catching the variability of the entire basin. Great variability in water masses structure exists from north to south in the Adriatic Sea. Further geographical variability is introduced by the presence of the rivers and by the difference characteristics of the coastlines, being the Italian coasts smoother that the eastern coasts.
In order to create a more accurate vertical structure of the background-error covariance matrix it has been decided to compute separate sets of EOFs for every point in the basin. In particular, the EOFs are calculated for the combined temperature and salinity statistics in every model point in the basin.
The fields used for the computation of EOFs originate from an interannual simulations that produced a great set of vertical T-S profiles and velocity and elevation fields that are then daily averaged. Every temperature and salinity profile, stream function Ψ (from which the horizontal currents are derived, see [2]) and sea surface elevation is then taken to form a state matrix $\mathbf{X}$:

$$\mathbf{X} = \left(\frac{\delta\eta}{\sigma_\eta}, \frac{\delta\Psi}{\sigma_\Psi} \frac{\delta T_1}{\sigma_T}, \frac{\delta T_2}{\sigma_T}, \dots \frac{\delta T_n}{\sigma_T}, \frac{\delta S_1}{\sigma_S}, \frac{\delta S_2}{\sigma_S}, \dots \frac{\delta S_n}{\sigma_S} \right) \tag{3.88}$$

where σ represent the standard deviation of the corresponding fields, and δ indicates the difference between the daily averaged value and temporal mean for each season. Each vector composing (3.88) is a time series of daily values.
The EOFs are then calculated from a singular value decomposition of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{S}\Lambda^{1/2}\mathbf{V}^T \tag{3.89}$$

where $\mathbf{S}$ is the matrix whose columns are the vertical EOFs that we are looking for.
In order to create EOFs independent of space geometry, i.e. number of levels, thickness of levels and in general the grid spacing, the matrix $\mathbf{X}$ has to be multiplied by a metric factor matrix $\mathbf{g}$ whose diagonal elements are:

$$\mathbf{g} = \left(11 \frac{\Delta z_1}{H_1}, \frac{\Delta z_2}{H_2}, \dots, \frac{\Delta z_n}{H_n}, \frac{\Delta z_1}{H_1}, \frac{\Delta z_2}{H_2}, \dots, \frac{\Delta z_n}{H_n} \right) \tag{3.90}$$

where the Δz are the model layer thicknesses in correspondence of the vertical T-S profiles, and H are the corresponding bathymetry.

The multivariate EOFs so obtained represent vertical covariances of surface elevation, stream function and T-S profile. Therefore, in each evaluation of the cost function the control vector $\mathbf{v}$ contains coefficients which multiply vertical multivariate EOFs, which are then transformed into corrections of tempera- ture, salinity velocity and surface elevation. The vertical transformation operator $\mathbf{V}_V$ has the form:

$$\mathbf{V}_V = \mathbf{S}\Lambda^{1/2} \tag{3.91}$$

It is interesting to note that allowing the EOFs to have a seasonal variability, we may introduce an implicit time dependence in the assimilation scheme. The 3D-Var algorithm is by itself a three dimensional method, therefore the time dependence is not explicitly introduced in the cost function. We may let the EOF, and therefore the covariance matrix $\mathbf{B}$, vary with time, thus allowing the analysis to consider a different background variability depending on the particular season of the analyzed time step.

Horizontal covariances

While the vertical variability is expressed by vertical EOFs, the horizontal covariances are assumed to be Gaussian with constant correlation radius. In oceanographic model, isotropic and Gaussian spatial correlations can be modelled in many ways, being the direct application of Gaussian correlation functions not the cheaper one. Dobricic and Pinardi proposed the application of the recursive filter already introduced in meteorology by Lorenc in [32]. It is conceptually simple, typically requires only a few iterations in order to approximate the Gaussian function, and its application on a horizontal grid can be split into two independent directions:

$$\mathbf{V}_H = \mathbf{W}_y\mathbf{G}_y\mathbf{W}_x\mathbf{G}_x \tag{3.92}$$

where $\mathbf{G}_x$ and $\mathbf{G}_y$ represent the recursive filter operators in directions x and y, whilst $\mathbf{W}_x$ and $\mathbf{W}_y$ are diagonal matrices with normalization coefficients with may account for variable grid resolutions in x and y direction.

While the application of this horizontal covariances operator is not too problematic in the atmosphere, in the ocean the application of the recursive filter is more complicated than in the atmosphere, due to the presence of coastlines.
In order to account for coastal boundaries in a computationally efficient way, we first apply an accurate boundary condition inside operators $\mathbf{G}_x$ and $\mathbf{G}_y$. In each one-dimensional application of the filter the grid is extended over land by adding several imaginary sea points at coastlines (figure 3.13). The number of points is arbitrary. Typically it spans several horizontal correlation radii. Once the grid is sufficiently extended inland it is not necessary to provide any lateral boundary condition, because with the grid extension we extrapolate the solution over land. After the application of the filter the imaginary

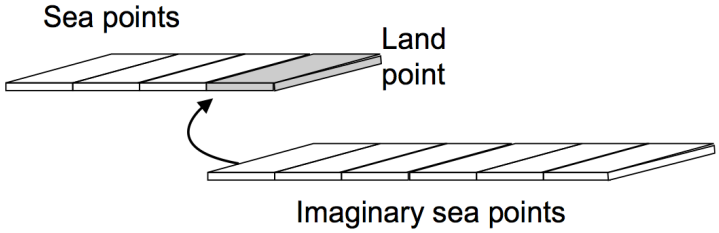


Figure 3.13: *The insertion of imaginary points at the coasts. Points are inserted at each coastal grid point over the land. The recursive filter is calculated on the extended grid with the larger number of points. After the application of the recursive filter imaginary points are removed. From [16]*

points are removed. Therefore, if the number of imaginary points is sufficiently large, the recursive filter does not transfer the information across land.
In one dimension this procedure accurately defines the function at the coastal boundary. However, in two dimensions the solution in each direction is obtained on a different grid, because imaginary sea points are eliminated after each one-dimensional application of the filter. As a result the filter does not provide a symmetric solution which is independent of the order of the applications of one-dimensional filters. Therefore, it is necessary to redefine (3.92) in terms of a symmetric operator. The adopted form for this purpose is the following:

$$\mathbf{V}_H = \frac{1}{2}(\mathbf{W}_y\mathbf{G}_y\mathbf{W}_x\mathbf{G}_x + \mathbf{W}_x\mathbf{G}_x\mathbf{W}_y\mathbf{G}_y) \tag{3.93}$$

This solution doubles the number of computations in each application of the horizontal filter, but it still allows us to apply the recursive filter separately in each direction. Figure 3.14 shows an example of the application of the recursive filter near the coasts using the symmetric operator.

Combining horizontal and vertical covariances operators

It is important to note in equation (3.87) that we first convert from EOF space into physical space, and then model the horizontal covariances. This means that the three-dimensional spatial covariances $\mathbf{V}_S$ are computed by the multiplication:

$$\mathbf{V}_S = \mathbf{V}_H\mathbf{V}_V \tag{3.94}$$

A computationally more efficient approach would be first to apply horizontal covariances in EOF space and then transform the result into physical space: that is, to use the formula:

$$\mathbf{V}_S^* = \mathbf{V}_V\mathbf{V}_H \tag{3.95}$$

In equation (3.94) the computational time for the calculation of horizontal covariances is roughly proportional to the product of the number of vertical levels and the number of state vector variables, whilst in equation (3.95) it is proportional to the relatively small

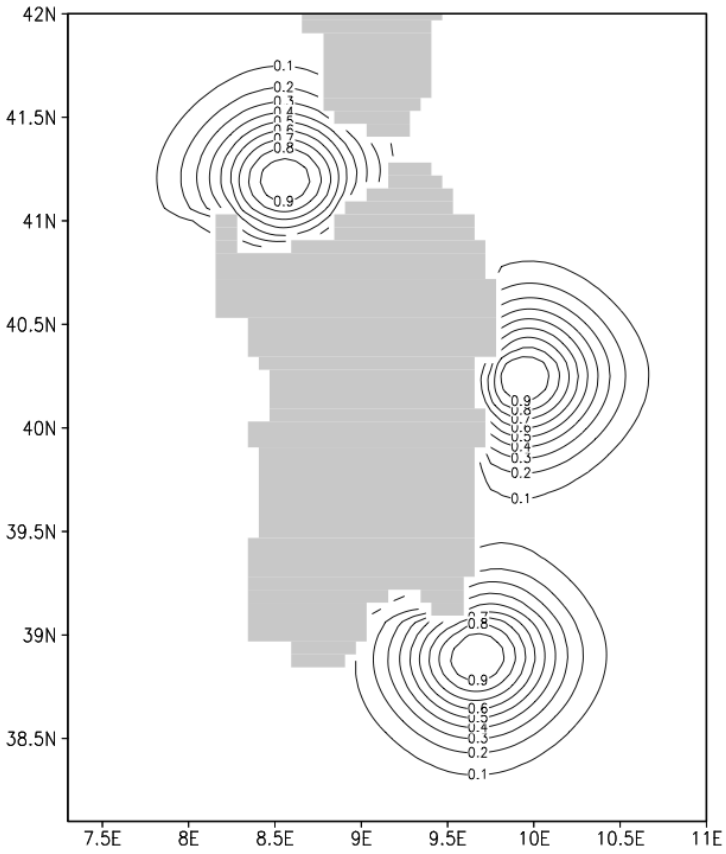


Figure 3.14: *The filtering result for covariances of sea level height (in units of distance) corresponding to 1.0 unit perturbations in three points close to the coast of Sardinia (Western Mediterranean). The contour interval is 0.1 units. The map projection is stretched latitude–longitude. The corre- lation radius is 30 km, and eight iterations of the recursive filter were used (four by the operator* $\boldsymbol{V}_H$ *and four by its transpose). Land points are shaded gray. From [16]*

number of EOFs. It is clear from (3.84) that both approaches result in a matrix $\mathbf{B}$ which is positive definite and symmetric, but they would give the same $\mathbf{B}$ only when vertical EOFs do not vary spatially within the horizontal correlation scales. In our case EOFs are different from one point to another, and the two approaches will produce very different result, with (3.95) introducing a lot of numerical, unwanted noise.
Thus, in order to produce smooth spatial correction, we must apply the most expensive (3.94)

Free surface operator

The dynamical balance between surface elevation perturbations and the corresponding temperature and salinity perturbations may be estimated either statistically, or by applying some kind of the geostrophic constraint. In [16] and sources, it is shown how if we

assimilate only in situ observations of temperature and salinity, the accurate estimation of the corresponding sea level corrections is not an important issue. In that case all unbalanced corrections will be rapidly removed by the fast barotropic adjustment.
On the other hand, it was demonstrated by Dobricic et al. in [14] that, when SLA observations were assimilated, the enforcement of the geostrophic relationship for the sea level in the error covariance matrix had a significant positive impact on the accuracy of the analyses.
Considering the highly variable bottom topography of the Adriatic basin, and results from previous experiment in the Mediterranean Sea, the operator $\mathbf{V}_\eta$ is constructed in order to give the steady state results of a barotropic model forced by density perturbations induced by the insertion of tem- perature and salinity perturbations. The barotropic model equations, discretized in time by the semi-implicit scheme [28], are:

$$\frac{U^{n+1} - U(n-1)}{\Delta t} - fV^n = -gH\frac{\partial \eta^*}{\partial x} - \int_{-H}^{0}\left[\int_{-z}^{0}\frac{\partial(\delta b)}{\partial x}dz'\right]dz + \gamma\nabla^2 U^{n-1} \tag{3.96}$$

$$\frac{V^{n+1} - V(n-1)}{\Delta t} + fU^n = -gH\frac{\partial \eta^*}{\partial y} - \int_{-H}^{0}\left[\int_{-z}^{0}\frac{\partial(\delta b)}{\partial y}dz'\right]dz + \gamma\nabla^2 V^{n-1} \tag{3.97}$$

$$\frac{\eta^{n+1} - \eta(n-1)}{\Delta t} + \left(\frac{\partial U*}{\partial x} + \frac{\partial V*}{\partial y}\right) = 0 \tag{3.98}$$

where U and V are vertically integrated velocity components, f is the Coriolis parameter, g the gravity acceleration, H the bottom depth, η the surface elevation, $\delta b = g(\delta\rho/\rho_0)$ the buoyancy anomaly, and γ is the coefficient of horizontal viscosity. Horizontal viscosity terms are added in order to speed-up the adjustment. The density perturbation $\delta\rho$ is estimated by the linear equation:

$$\delta\rho = \alpha\delta T - \beta\delta S \tag{3.99}$$

where α and β are expansion coefficients for temperature and salinity. Model equations are discretized spatially using the fully staggered C grid. The non-linear advection terms and the bottom friction are neglected, because in the first tests they appeared to be insignificant. The superscripts indicate the time step relative to n , and the superscript "*" indicates the weighted average between two time steps:

$$A^* = aA^{n+1} + (1-a)A^{n-1} \tag{3.100}$$

where $0.5 \leq a \leq 1$ in order to have numerically stable solutions. Coriolis terms are solved by the leapfrog scheme, and horizontal viscosity terms by the forward Euler scheme. Numerically, equations (3.96), (3.97), (3.98) are solved by forming an elliptic equation for η^{n+1} which is solved by successive corrections, and then by substituting the solution for η^{n+1} into (3.96), (3.97).
In equation (3.100) higher values for a filter more the short gravity waves and provide

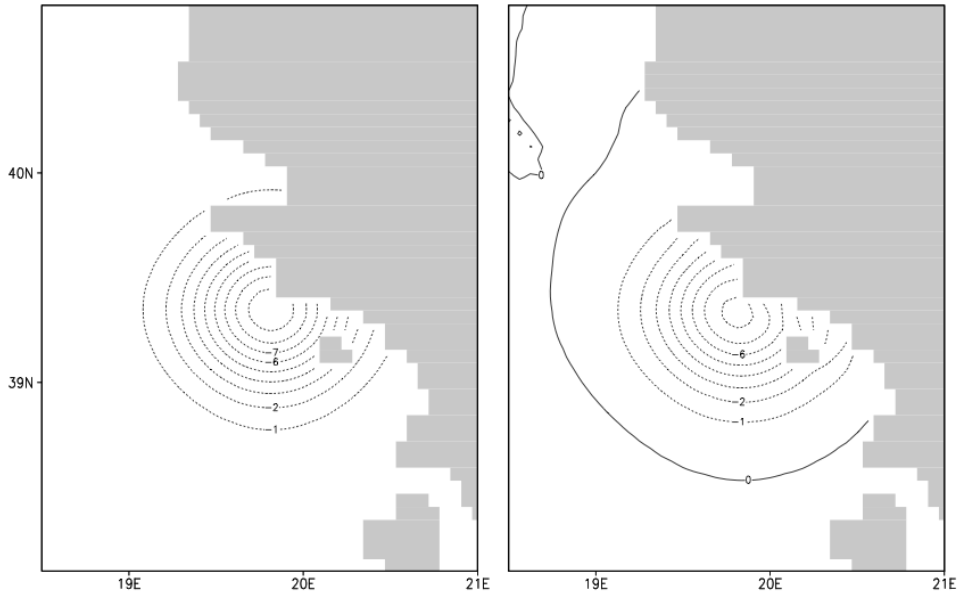


Figure 3.15: *Impact of the application of the sea surface height operator to the estimate of sea level correction (cm) in a case of an Argo temperature and salinity profile close to the coast of Greece in the Northern Ionian Sea. Left panel shows correction obtained with EOFs obtained by the method described in [14] and the right panel with the barotropic model described in this study. From [16]*

a faster convergence to the steady state. However, initial tests of the model showed that sometimes for $a = 1$ the steady state solution was significantly different from the solution obtained for $a = 0.5$. Therefore, as a compromise between the efficiency and accuracy, the model uses $a = 0.8$. This choice provides a steady state solution which is practically identical to the solution with $a = 0.5$, and converges relatively rapidly. It was found that the stationary solution is obtained as a daily average already after only a few days of the integration (3.96), (3.97), (3.98). Expansion coefficients α and β are assumed to be constant and are estimated in advance as mean coefficients calculated from background fields of temperature and salinity.

After forming corrections in temperature and salinity, we can use the barotropic model to estimate the corrections in sea level. Figure 3.15 shows the effect of the $\mathbf{V}_\eta$ operator.

Velocity operator

The $\mathbf{V}_{uv}$ operator calculates the velocity corrections by splitting the velocity field into the barotropic and baroclinic components. The barotropic velocity correction is obtained by solving (3.96), (3.97), (3.98) whilst the baroclinic parts of the velocity components u_{bc} and v_{bc} are estimated from the usual dynamic height formula, assuming geostrophic balance:

$$fv_{bc} = \int_{-z}^{0} \frac{\partial(\delta b)}{\partial x} dz - \frac{1}{H} \int_{-H}^{0} \left[\int_{-z}^{0} \frac{\partial(\delta b)}{\partial x} dz' \right] dz \tag{3.101}$$

$$fu_{bc} = -\int_{-z}^{0} \frac{\partial(\delta b)}{\partial y} dz + \frac{1}{H} \int_{-H}^{0} \left[\int_{-z}^{0} \frac{\partial(\delta b)}{\partial y} dz' \right] dz \tag{3.102}$$

The advantage of the geostrophic assumption is that it re- quires only a small computational effort, but it is not valid at the Equator and may produce velocity vectors orthogonal to the coast. There is no problem about that in this study, where we want to apply this 3D-Var in the Adriatic Sea.

Divergence damping operator

In the absence of coastlines the velocity operator $\mathbf{V}_{uv}$ imposes a dynamical balance between corrections in velocity, sea surface height and mass fields. The velocity field has only the rotational part for the baroclinic component. However, we need to enforce the zero boundary condition for the total velocity component perpendicular to the coast. As a result, the baroclinic velocity component perpendicular to the coast can suddenly change from its geostrophic value to zero, and the divergence component of the velocity field may become unrealistically large. Therefore, the divergence damping operator $\mathbf{V}_D$ in (3.87) will damp velocity divergence near coasts. It filters the divergence part of the velocity field, and maintains the vorticity part unchanged. It could be especially important to apply the divergence damping operator already inside the 3D-VAR analysis scheme in order to form a more precise estimate of velocity perturbations when assimilating observations close to the coast.
The divergence damping is achieved by the successive application of a filter in the form:

$$\delta\mathbf{v}^{n+1} = \delta\mathbf{v}^{n} + \kappa\nabla D^{n} \tag{3.103}$$

where $\delta\mathbf{v} = (\delta u, \delta v)^T$ is the total velocity increment at each vertical level, superscripts indicate the iteration step number, κ is the filter coefficient, and D is the horizontal divergence of the velocity increments.

The effect of the filter is illustrated in figure 3.16. The major impact is at the coastline, where the filter changes the direction of the velocity which initially had an unrealistic direction perpendicular to the boundary. Away from the coast, the impact of the filter is practically negligible. The computational cost of the divergence damping filter is relatively high, because it is applied on each vertical model level. When there are no velocity increments close to the coast the divergence damping operator practically provides a smoother initial condition for the forecast.

3.3.3 Numerical implementation

Most of the computational time necessary to minimize the 3D-VAR cost function is spent in the transformation of increments from the control space to the physical space by the linear operator $\mathbf{V}$ and its transpose. Therefore, the computational cost of the 3D-VAR is approximately proportional to the size of the control vector $\mathbf{v}$. Furthermore it also depends on the complexity of the linear operators in (3.87). It is important to note that

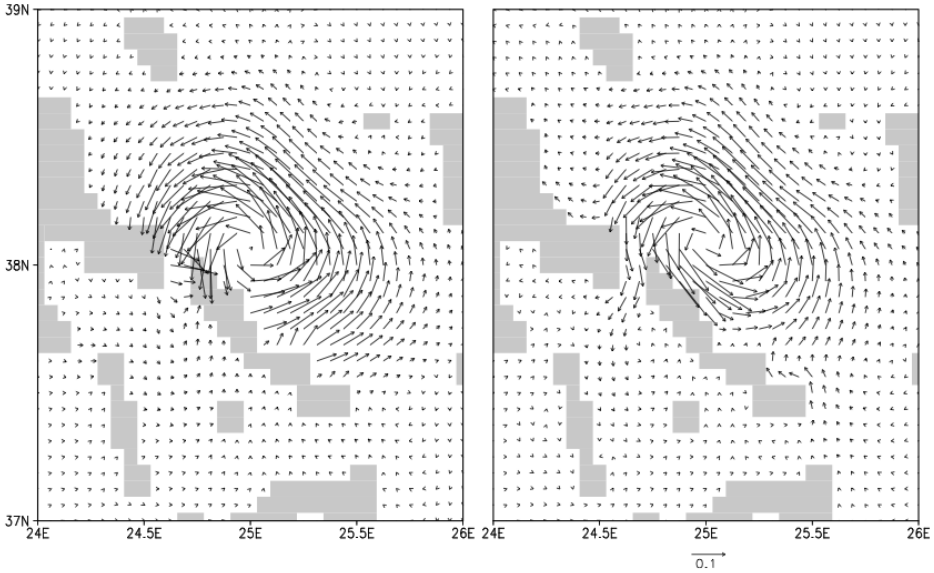


Figure 3.16: *Impact of the divergence damping operator on corrections of surface velocity (ms^{-1}) corresponding to observations of temperature and salinity profiles by an Argo float in the Aegean Sea. The left panel shows corrections produced by the analysis without the application of the divergence damping filter, and the right panel shows filtered corrections. Grid points corresponding to the land are shaded gray. From [16]*

the reduction in computational time obtained by using vertical EOFs in the 3D- VAR is larger than in the optimal interpolation scheme. In addition to reducing the number of operations on the state vector inside the minimizer code, the use of EOFs in the 3D-VAR reduces the size of the control vector $\mathbf{v}$ and therefore may reduce the number of iterations of the minimizer.

To further decrease the computational costs of the algorithm, the minimization is performed by the application of a *multigrid* approach. The multigrid method is well known for the solution of partial differential equations. In the case of the 3D-Var optimization, it consist of splitting the procedure in two main steps. First the optimization is performed on a grid that has a lower spatial resolution than the full grid in the physical space and then the solution is interpolated on the finer resolution grid for a further minimization. The practical implementation of this method is done by two loops: an outer loops that is responsible for the interpolation from the coarser grid to the finer resolution one, and an inner loop that minimizes the cost function on each grid.
This particular implementation has many advantages. First of all we can, we can decide, with some specific criteria, that the quick minimization on the coarse resolution grid is already sufficiently accurate, and therefore prevent the algorithm from performing the minimization on the finer grid.
An additional trick in order to save computational time is to use further simplifications of the linear operators in the first optimization and then using the results as a first guess

for a more precise minimization on the finer grid, using the full linear operators. This method is particular attractive when dealing with the highly expensive operators $\mathbf{V}_\eta$ and $\mathbf{V}_D$. In particular, the computational cost of the divergence damping filter is relatively high, because it is applied on each vertical model level. Therefore, this operator is applied only in the last iteration of the outer loop of the numerical minimization.

The optimization is practically performed using a quasi-Newton L-BFGS minimizer proposed in [9]. Iterations of the minimizer are stopped when the absolute value of the gradient of the cost function becomes small relative to its initial value (typically one percent of the initial value). The calculation of the gradient of the cost function requires the adjoint of the operators $\mathbf{V}$ and $\mathbf{H}$. The adjoint is formulated by hand-coding the transpose of each subroutine containing operators $\mathbf{V}$ and $\mathbf{H}$, following rules given by [22].

3.4 Analysis System set up and numerical experiments

We have so far discussed the three components of our Analysis System for the Adriatic Sea: the observing network, the numerical model and the Data Assimilation scheme. We now describe how these components are merged together in this work and the numerical experiment we performed using this system.

3.4.1 Analysis System set-up

Observed Dataset

As already discussed we will perform our analysis for the target year 2006. This year was chosen based on the in-situ data spatial and temporal coverage. This choice is justified by the consideration made in the second chapter.
We consider here satellite data and in-situ data.

The satellite dataset is entirely assimilated in our study and is composed of the SLA observations from the satellites available in the 2006: Geosat 2 (g2), Jason 1 (j1) and Envisat (en).
The data are delayed time (DT): the last analysis was performed in the year 2010. Therefore we have the best possible satellite data available at the time and we can focus our attention on the effect of in-situ data assimilation.

The in-situ dataset we want to assimilate in the model is the 2006 SeaDataNet (SDN) dataset shown in figures 2.4, 2.5 and 2.6. It is composed of 1921 vertical profiles of XBT, CTD and bottles.

There is another in-situ dataset, composed of the ISRPA tide-gauges stations shown in figure 2.7. This data are not assimilated in the model and will be considered as an independent test-set.

The Numerical Model

The numerical model used for the study is the Adriatic REGional model AREG2. The results that we analyzed are daily means of the variables contained in the model state $\mathbf{x}$. The model is in principle able to produce hourly results of the desired variables but we didn't consider this possibility for two reasons: first of all because of the computational cost it would have running the model with an hourly time step, and second, more important, because we want here to simulate the behavior of this system in the framework of a future implementation in AFS, that produces daily means for operational purposes.
So we focus our attention on the daily means, for the entire 2006, of the state of the sea. In particular we are interested in: temperature, salinity, sea level, currents. For these quantities the model calculates one spatial field per day which is a daily mean centered around midnight (00:00 am) of the considered day.

The initial condition for the entire analysis is given by the restart of the AFS analyses.

DA algorithm and implementation

The assimilation algorithm we want to test for this is the 3D-Var described in the previous section and already proposed for the Mediterranean Sea in [16].

The core of the DA algorithm is the minimization of the cost function $J(\mathbf{x})$ (3.64). The matrix $\mathbf{B}$ is calculated using the methods given in the previous section, while the matrix $\mathbf{R}$ is a diagonal matrix which elements are given by the uncertainties associated to each observation. For simplicity, observations from the same kind of instrument are considered to have the same observational error. In particular, the uncertainties on the CTD and bottles data are 0.05 °C on temperature and 0.02 psu on salinity. On the XBT observations the uncertainty on the temperature is considered to be 0.5 °C. Finally, the error associated with the SLA data is 0.005 m.
These numbers are arbitrary chosen by the scientists of the GNOO, but we can justify them on the basis of what we saw in the first chapter about the oceanographic observations. The uncertainties on CTD, XBT, bottle devices and SLA observations are given there and are taken from [18]. Looking at the uncertainties given there we can see that they are always smaller than those considered in the matrix $\mathbf{R}$ concerning the in-situ data, thus allowing for a smaller confidence on the quality of the observations. As it has to be, we can see that temperature uncertainty from XBT is higher than the same observation from a CTD or a bottle device.
On the other hand the error chosen for SLA observation (0.005 m) is small compared to the single observation error of a Real Time SLA data. This is clearly justified by the fact that we are using DT data, much better analyzed than RT data.

The theoretical usage of the 3D-Var algorithm should be as follows. Every day the background state $\mathbf{x}_b$ is given by the forecast of the day before. Then the L_2 distance between the model state variable $\mathbf{x}$ and the background should be calculated at the same time of the available observations. This temporal coherence is expressed by the FGAT

concept: First Guess at Appropriate Time. But our model produces daily averages of the fields of interest, meaning one field per day for every variable in $\mathbf{x}$. Moreover, these field are evaluated on a finite resolution grid. In other words, no observation would never be directly compared with the values furnished by the model.
Spatially, the problem is resolved simply by linearly interpolating the model output on the observation points.
The conceptual problem arises when dealing with the temporal displacement between the background state (midnight) and the observations, which are possibly distributed over the entire day. Here we simply ignore this fact and every observation is compared to the daily average of the corresponding day provided by the model. This is equivalent of assuming that every observation have been made at the same instant: midnight of the considered day.
A more acceptable solution would be using the model for calculating hourly means of the physical fields, but this would require not indifferent computational time and storage costs to simulate a whole year.

Another simplification made for saving computational time is that the misfits are not used to correct the field of the same day. This would require two model run instead of one: the first providing the background state, and the second for producing the analysis $\mathbf{x}_a$.
An acceptable approximation consists in using the correction for the day i as the input for the run of the day $i + 1$. This is exactly what we do in our experiments.

3.4.2 Experiments design

The design of the experiments is simple. We will run the model twice. The first time we are not assimilating anything: as in AFS the model is forced by the river runoff, the water exchange with the Mediterranean and the atmospheric forcings, but no data are assimilated. This first experiment will therefore be referred as the "noassim" experiment.
In the second experiment we perform the daily assimilations as described above. So the model is both forced as in the first experiment and "guided" by the information provided by the assimilation of available observations. The second experiment will be named the "assim" experiment.
Clearly we expect the model fields of the second experiment to be much more accurate than those provided in the first experiment.

In the following chapter we will compare the outputs of the two experiments in order to evaluate the quality of the system analysis. Then we will show the effect of the assimilation on the general circulation description capabilities of the AREG2 model

Chapter 4

The estimate of the Adriatic Sea circulation

In this chapter we present the results of the two experiment explained at the end of the previous chapter. The assessment of the analysis quality of the model with the assimilation scheme with respect to the model without assimilation will be carried on in two steps, corresponding to the two sections of this chapter.
In the first section we will evaluate the analysis quality by comparing the observations collected for the experiment (and described in the second chapter) to the results of the two experiment.
In the second section we will show how the assimilation changes the general circulation pattern and the properties of the seawater in the basin.

4.1 Evaluation of the AS analysis quality

In order to evaluate the accuracy of the results from the two experiment we can first of all compare them to the observations. The results presented in this section are obtained in the following way. First of all both the "noassim" and the "assim" results have to be interpolated on the observations points in order to calculate the *misfits*, that is, the differences between the model and the observations.
The interpolation depends whether the quantity of interest is a two dimensional field (like the sea level elevation) or a volume property of the model (like temperature, salinity and currents). In the first case the required procedure is an interpolation on the horizontal, in order to bring the model field of interest on the observation point. In particular we make use of a bilinear interpolation on the horizontal plane (figure 4.1). Care must be taken in the case of an observation close to the coasts, because in this case one or more of the Q points of figure 4.1 may be land points. In this case the field value on the land points is taken as the average value of the remaining sea points. In the case of a two dimensional field (like the sea level or the SST) the horizontal interpolation gives us the desired scalar value comparable to the observation carried at the point P.

In the case of a three dimensional field we have to interpolate both on the vertical and

on the horizontal. So we first obtain a vertical vector corresponding to a vertical profile of the model quantity take in correspondence of the horizontal position of the vertical observation. The vertical model profile is then linearly interpolated on the depth corresponding to the observation depths.
In any case we now have the observed quantity ϕ_i^o and the model quantity defined on the observation point (or points) ϕ_i^m. The quantity ϕ can be temperature, salinity, sea level or any other quantity of interest ;the superscript m can refer both to the model without the assimilation and the model with the 3D-Var; the subscript i represent a particular observation.

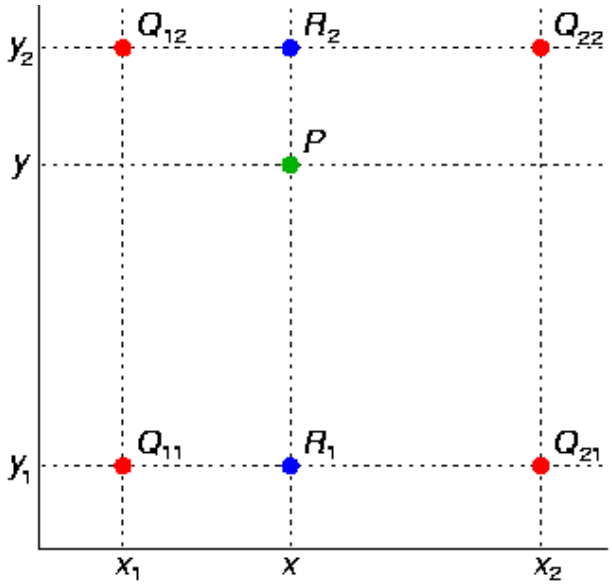


Figure 4.1: *Schematic representation of the bilinear interpolation. The quantities Q_{ij} are the model quantities specified on the (i, j) point of the AREG2 grid. P is the observation point*

Now, the i-th misfit is defined as:

$$\Delta\phi_i = \phi_i^m - \phi_i^o \tag{4.1}$$

From the misfits from N available observations we can calculate the *bias* and the *root mean square error* (RMSE) defined as:

$$bias = \frac{1}{N}\sum_{i=1}^{N}\Delta\phi_i \tag{4.2}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\Delta\phi_i} \tag{4.3}$$

Now we use this two quantities to express quantitatively the improvement of the model results after the assimilation of the data. The computation is presented both for the assimilated and the independent observations.

4.1.1 Assimilated data

Here we estimate the AS analysis quality with respect to the assimilated data. This statistical index is not a check of the assimilation and model performance by independent data but it can show the improvement of the model solution due to the assimilation of data.
We are showing the comparison of the two experiments and the in-situ T-S data.
The same comparison for the SLA data is not available due to lack of time.

T-S profiles

The results of the comparison with vertical T-S profiles from XBT-CTD-bottle is shown here keeping in mind the subdivision into sub-regional observation areas (shown in figure 2.4) and the splitting of the whole year in four seasons: winter (January-April), spring (May-June), summer (July-October), fall (November-December). Therefore the reader has to keep in mind the results from figures 2.5, 2.6 and 2.4. This definition of the four seasons is also found by Artegiani et al in [4] to be reflected in the water masses properties.

The presentation of the results not in a single solution is essential in order to well understand the reasons of the biases and RMSEs that are found from the analysis. Otherwise it would not be possible to control the assimilation effect.

Winter Results for winter 2006 are shown in tables 4.1, 4.2. The seasonal dataset is shown in figure 4.2.
This season is characterized by a fairly good data coverage in all the four region defined for the SDN dataset, except for the first two month of the year in the middle and southern Ad. observations. In the middle Ad. in particular, we have only bottle casts. However, in the southern Ad. we have both XBT and CTD observations extending below the 40° N latitude; thus, both salinity and temperature are assimilated in the southern region in this season.

Watching at the tables we can see that there is a general improvement in both temperature and salinity profiles reproduction after the assimilation. In particular we have that the salinity is much better reproduced in the "assim" experiment. This fact will be a recurrent feature for the other seasons too.

Watching the single lines of the tables 4.1 and 4.2 we can give the following comments:

full dataset On this line we see the things that we have already suggested about the different improvement in temperature and salinity reproduction. Apparently the temperature profiles of the "noassim" experiment are even closer to the observation than that of the "assim" experiment. However we can not judge the entire assimilation quality from this result: we have to look at the other lines first.

XBT The XBT data are present in the southern Adriatic and clearly they provide information only on temperature. This result seem to be in line with the first line of the two tables. See also the comments about the line of the Southern Ad.

CTD In this case we have an apparent improvement of the temperature profiles bias but a slight degradation in a RMS sense. In salinity we have a significant improvement.

bottle Bottle data seem to be much better reproduced after the assimilation. The bias, in particular is greatly reduced, indicating that the whole profiles get closer to the observations.

ARPA In front of the Emilia-Romagna coasts, we have a reduced bias in temperature but a greater RMSE. In the salinity we don't seem to have a reduced misfit but it is interesting to notice that the bias in salinity has an opposite sign in the two experiments: in the "noassim" results the modeled salinity was systematically smaller than the observed profiles, while the "assim" implementation brings the modeled salinity profiles to be systematically greater than the observations.

Rovinj Here we recognize an improvement both in temperature and salinity. We can comment this fact by saying that this is a season of good data coverage in the Rovinj region (see also figure 2.6a)

Middle Ad. The temperature has a smaller bias and RMSE in the "assim" experiment, with an opposite sign in the bias. The salinity get better concerning the absolute value of the bias, with a change of sign after the assimilation, but is slightly worse reproduced in a RMSE sense. Remember that the Eastern coast of the middle Ad. is a region of strong river runoff, being the river distributed along the whole coast. It is therefore hard for the model to correctly reproduce the great variability in salinity, even with an assimilation algorithm and good observations.

Southern Ad. In the southern region we have almost the same results for both experiment, with the temperature slightly worse reproduce, both concerning the bias and the RMSE.

Remember that this region is special because of the water exchange with the saltier and warmer Mediterranean Sea. Many data could be required to a correct reproduction of the complex dynamics involved in this region and the assimilation algorithm may produce noisy profiles in the attempt of correctly meld them into the model. We can in fact anticipate that, even if the results from the southern Ad. doesn't seem to indicate an improvement of the T-S reproduction after the assimilation, the remaining season are much more problematic than this. The cause of this may be regarded as a lack of both temperature and salinity observations in the other seasons.

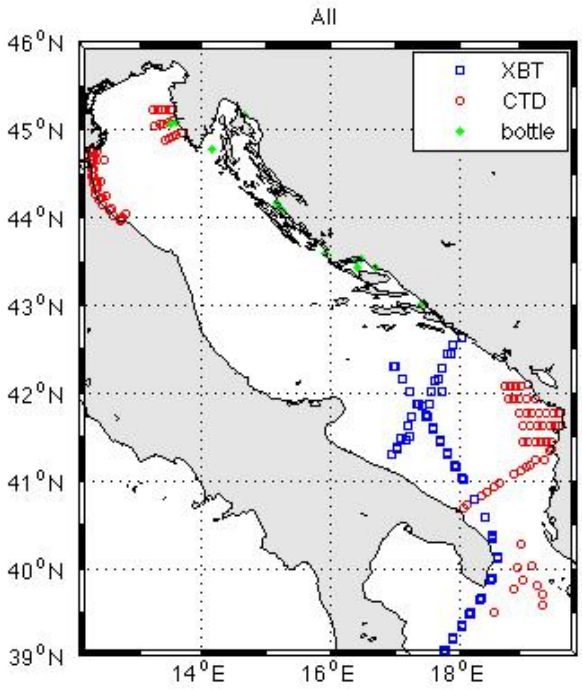


Figure 4.2: *Temperature-salinity profiles available in the Adriatic Sea for winter 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.4	0.6	−0.2	1.1
XBT	−0.4	0.5		
CTD	−0.8	1.1	−0.16	1.1
bottle	−0.3	1.0	−0.04	0.5
ARPA	−1.1	1.5	−0.18	1.9
Rovinj	−1.0	1.3	−0.4	0.4
Middle Ad.	−0.3	1.0	−0.13	0.5
Southern Ad.	−0.4	0.5	−0.05	0.4

Table 4.1: *bias and RMSE between data and model without assimilation for the T-S profiles of winter 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.5	0.8	−0.01	1.0
XBT	−0.5	0.6		
CTD	−0.2	1.3	−0.01	1.0
bottle	−0.002	0.9	−0.007	0.5
ARPA	−0.02	2.3	0.2	1.8
Rovinj	−0.5	0.9	−0.3	0.3
Middle Ad.	0.002	0.9	0.11	0.6
Southern Ad.	−0.5	0.6	−0.05	0.4

Table 4.2: *bias and RMSE between data and model with assimilation for the T-S profiles of winter 2006*

Spring Spatial data coverage for spring is given in figure 4.3 and bias and RMSE results are reported in tables 4.3 for the "noassim" experiment and 4.4 for the "assim" experiment. In this season we can see a poor spatial coverage of the data in the middle and southern Adriatic. In the middle sub-basin, as in winter, we have only bottle casts. The southern Adriatic is covered only with XBT casts (except for few bottle casts) in this season. This fact is reflected in the last lines of the two tables below indicating, as previously anticipated, that temperature alone is not sufficient for the assimilation in order to bring the vertical profiles closer to the observations in this region.
In reading the following comments, remember that this season, together with fall, has only two months, instead of four (as in winter and summer).
Furthermore remember that the assimilation produce a correction of the full state vector based on the overall evaluation of the cost function. The cost function express a combination of the distance from the data and from the background state and this analysis is only focused on the vertical observed T-S profiles: the fact that the vertical temperature and salinity are not always closer to the observations after the assimilation does not have to be interpreted as an overall degradation of the solution with respect to the "noassim" result. This will be clear in the next section, where we will show the temperature, salinity and circulation fields for the two experiments.

full dataset We can not conclude much from this first line. It is however interesting to note how the assimilation produced an opposite bias in salinity: the model without assimilation is systematically underestimating the vertical salinity while the assimilation is bringing the modeled profiles to be in general more salty than the observed profiles.

XBT The XBT profiles reproduction seem to be slightly worse after the assimilation. A possible explanation could be that the assimilation might require more data than the provided observations to better constrain the vertical temperature in this region.

CTD In this case we have a slightly improvement of the bias of the solution, with a change of sign from "noassim" experiment to "assim". In this case too, the model without assimilation tends to produce fresher and colder profiles than the observations while the assimilation make them saltier and warmer than the data. A possible explanation for this is given in the next section and is tied to the tendency of the model without the assimilation to smooth the temperature and salinity gradients that are instead important in the dynamics of the seas.

bottle We do not have any clear improvement or degradation of the T-S profiles from the bottle casts, except for the temperature bias, that is bigger in the "assim" experiment and opposite in sign when compared to the "noassim" experiment.

ARPA In this case too the solution after the assimilation does not seem to be much better or worse than the results provided by the model in the "noassim" implementation. A more biased temperature can be however noticed.

Rovinj In this case we have a better reproduction of the temperature profiles after the assimilation (both from the bias and RMSE estimation) but a worse reproduction of the salinity profiles.

Middle Ad. Salinity profiles seem to be not differently reproduced from the two experiments (at least for what concerns this analysis). Temperature seems to have a greater bias and a smaller RMSE after the assimilation. This could be interpreted as a bigger systematic error but a reduction of the quadratic distance to the observed profiles, which is in general indicating that at least the shape of the modeled profile is more similar tot the observed. Although we would DA techniques to reduce both bias and RMSE, it is not always the case, as shown in this (and other) example.

Southern Ad. The observed profiles in this region seem to be not well reproduced after the assimilation of the data in the model. This could be due to the lack of open sea salinity measurement from CTD campaigns, as previously noticed and to the overall smaller number of observations with respect to the other seasons.

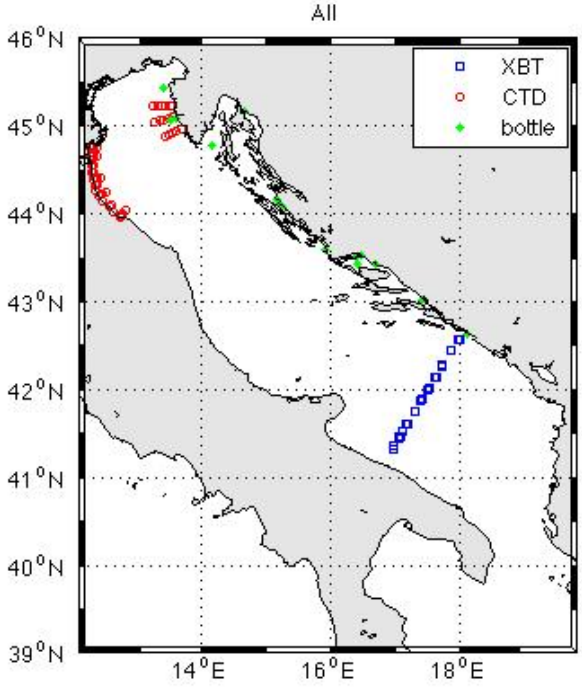


Figure 4.3: *Temperature-salinity profiles available in the Adriatic Sea for spring 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.4	0.9	−0.2	1.4
XBT	−0.5	0.6		
CTD	−0.3	1.6	−0.2	1.4
bottle	−0.04	1.3	0.4	0.8
ARPA	0.02	1.6	−0.4	1.7
Rovinj	−0.8	1.6	0.06	0.4
Middle Ad.	−0.03	1.3	0.7	1.1
Southern Ad.	−0.4	0.6	0.005	0.3

Table 4.3: *bias and RMSE between data and model without assimilation for the T-S profiles of spring 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.5	1.0	0.2	1.3
XBT	−0.6	0.7		
CTD	0.2	1.8	0.2	1.3
bottle	0.2	1.2	0.4	0.8
ARPA	0.4	1.9	0.2	1.6
Rovinj	−0.08	1.5	0.13	0.5
Middle Ad.	0.16	1.2	0.6	1.1
Southern Ad.	−0.6	0.7	0.3	0.4

Table 4.4: *bias and RMSE between data and model with assimilation for the T-S profiles of spring 2006*

Summer Spatial data coverage for summer is given in figure 4.4 and bias and RMSE results are reported in tables 4.5 for the "noassim" experiment and 4.6 for the "assim" experiment.
Note the lack of CTD observations in the Rovinj zone and the absence of CTD casts in the southern Adriatic, where we have a modest number of data in summer especially in August (see figure 2.6d). This is clearly be reflected in the T-S modeled profiles.
Instead, in the middle Ad. we can notice the presence of CTD observations together with bottles.

full dataset First of all we can notice the greater bias in temperature in the second table. Note however that this difference is only at the order of the second decimal digit, corresponding to $O(0.01C)$, thus not a big difference. As usual we have to watch the single sub-region to understand what is the actual effect of the assimilation. However we can notice that the bias in salinity has the same absolute value in the two tables but opposite sign. In this case too seems that the modeled S profiles before the assimilation are fresher than the observations suggests.

XBT We can notice a slightly higher bias in temperature bias and a slightly reduced RMSE after the assimilation, meaning that at least the noise along the profile seems to be reduced in the second experiment.

CTD We notice a general degradation of the temperature profile reconstruction in the second table with respect to the first one. The salinity shows opposite bias and quite similar RMSE.

bottle Here we have an overall improvement of the solution after the assimilation: bias and RMSE for both T and S profiles are smaller in the second table. There is also an opposite sign in the salinity bias in the two table with the "assim" bias being negative and the "noassim" experiment suggesting a positive bias in the salinity profiles.

ARPA Temperature bias and RMSE are greater in the second table than in the first one, indicating that the observed T profile are not well reconstructed after the assimilation. Modeled salinity profiles, on the other hand, seems to be closer to the observations after the assimilation. In this case too we notice the reversed sign in the salinity bias.

Rovinj In this season in the Rovinj area we have only bottle casts. And we seem to have a problem because no temperature comparison has been made. Remember that from bottles both temperature and salinity may be measured (together with other physical and biogeochemical quantities). What could happened is that these temperature profiles did not passed the quality control of the Sea Data Net analysis and are therefore not available for this study.
In salinity there is a greater bias and RMSE.

Middle Ad. There is no update of the temperature bias and RMSE in from the first to the second experiment, meaning that at least the local solution is not getting worse.
In salinity we have a smaller bias and RMSE.
These good results can be due to the presence of a great number of data and from

the presence of both bottle and CTD casts in this region for this months.

Southern Ad. There is an overall increase in the bias, both for temperature and salinity in the second table. This is probably due to the non-optimal data coverage in this season. However the RMSE seems to be unchanged, with the temperature getting slightly better after the assimilation.

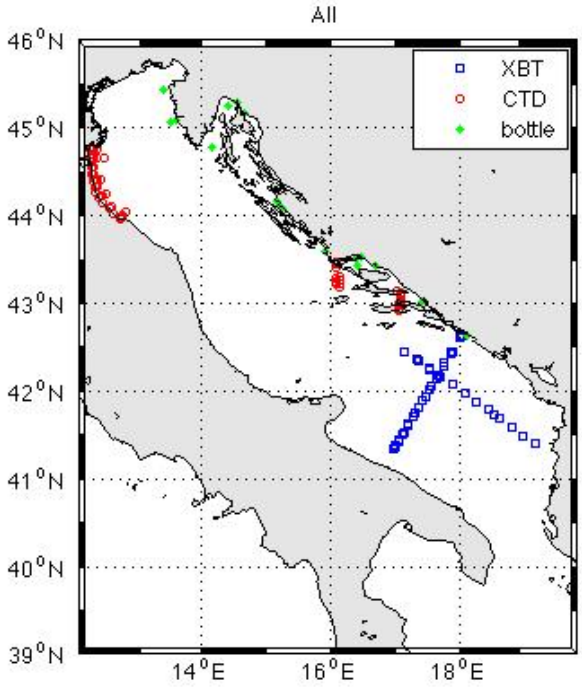


Figure 4.4: *Temperature-salinity profiles available in the Adriatic Sea for summer 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.03	1.1	−0.4	1.6
XBT	−0.08	1.1		
CTD	0.2	1.3	−0.4	1.6
bottle	−0.5	1.6	0.4	0.6
ARPA	−0.01	1.1	−0.6	1.7
Rovinj			0.2	0.4
Middle Ad.	1.0	1.8	0.4	1.0
Southern Ad.	−0.07	1.1	0.07	0.4

Table 4.5: *bias and RMSE between data and model without assimilation for the T-S profiles of summer 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.09	1.1	0.4	1.5
XBT	−0.17	1.0		
CTD	0.3	1.7	0.4	1.5
bottle	−0.4	1.6	−0.3	0.5
ARPA	0.12	1.6	0.5	1.6
Rovinj			0.3	0.5
Middle Ad.	1.0	1.8	0.3	0.9
Southern Ad.	−0.17	1.0	0.15	0.4

Table 4.6: *bias and RMSE between data and model with assimilation for the T-S profiles of summer 2006*

Fall Spatial data coverage for summer is given in figure 4.5 and bias and RMSE results are reported in tables 4.7 for the "noassim" experiment and 4.8 for the "assim" experiment. Notice the lack of CTD data in the Rovinj area and the total absence of data in the middle Ad. (also anticipated in figure 2.6c).
In this season too, we do not have CTD casts in the southern Adriatic.

full dataset The overall quality assessment is not greatly unchanged, except for the reversed sign in the bias of salinity

XBT In this dataset there is only i slight increasing in temperature bias. However this change is hardly significant being observed only on the second decimal.
The XBT casts are only present in the southern sub-basin

CTD Notice that in this season this row is equal to the ARPA row, being the CTD present only in front of the Emilia-Romagna coasts.
In this row we can see a general improvement of the modeled profiles after the assimilation, except for an increased RMSE in temperature. The bias salinity has an opposite sign, in addition to the smaller magnitude.

bottle In this season too we have a problem with bottle temperature. The cause again may be the quality control on the dataset.
The salinity shows a slightly smaller bias after the assimilation, indicating that the observed salinity profiles are resolved equally well in the two model implementations.
Notice that the bottle observations in this seasons are all concentrated in the Rovinj zone.

ARPA See the comments on the CTD data.

Rovinj see the comments on the bottle data.

Middle Ad. No observations are available in this region.

Southern Ad. See the comment on the XBT data.

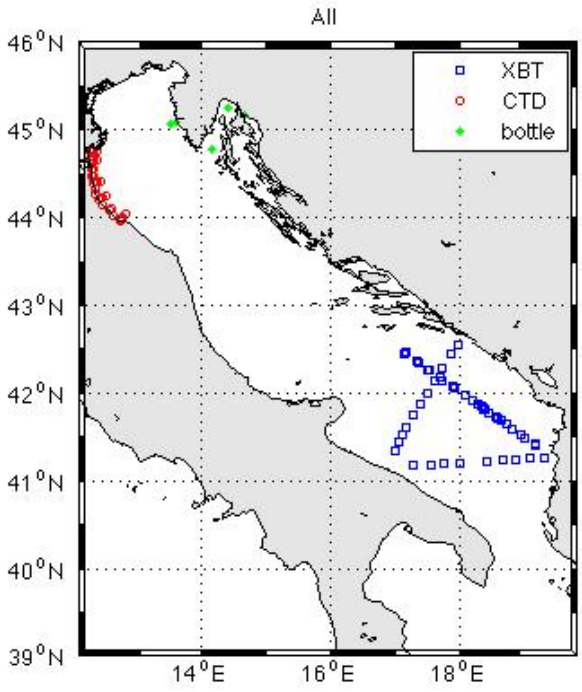


Figure 4.5: *Temperature-salinity profiles available in the Adriatic Sea for fall 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.17	0.8	−0.5	1.3
XBT	−0.12	0.7		
CTD	−1.3	1.8	−0.6	1.3
bottle			−0.15	0.4
ARPA	−1.3	1.8	−0.6	1.3
Rovinj			−0.15	0.4
Middle Ad.				
Southern Ad.	−0.12	0.7		

Table 4.7: *bias and RMSE between data and model without assimilation for the T-S profiles of fall 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.14	0.8	0.4	1.2
XBT	−0.14	0.7		
CTD	−0.07	2.2	0.4	1.3
bottle			−0.12	0.4
ARPA	−0.07	2.2	0.4	1.3
Rovinj			−0.12	0.4
Middle Ad.				
Southern Ad.	−0.14	0.7		

Table 4.8: *bias and RMSE between data and model with assimilation for the T-S profiles of fall 2006*

Entire 2006 Here we present the summarized results for the entire year. The dataset is shown in figure 4.6, which is clearly equivalent to figure 2.4, and the quality assessment base don the observed profiles is given in table 4.9 for the "noassim" experiment and in table 4.10 for the "assim" experiment.
In this section we clearly find results that contains information that are a sum of the ones presented for the single seasons so the comments that can be made are quite similiar. Keep in mind that the analysis of the full year results is a good summary, but detailed information for the various season can be more informative in estimating the critical and optimal aspects of the assimilation of the data in the Adriatic Sea.

full dataset The overall quality assessment based on the observed T-S profiles shows good results. The only exception is represented by the RMSE in temperature, slightly greater in the second experiment.

XBT In this case we can only notice a little difference in the bias of the temperature profiles. Remember that the XBT are concentrated in the southern sub-basin, so this results reflect the upgrades in the observed temperature profiles reconstruction in the southern Adriatic

CTD CTD profiles are generally better resolve by the model with the 3D-Var. Notice the opposite sign in the salinity bias, also found in previous results.

bottle On this row we find a very good improvements after the assimilation. This can be due to the high quality of the bottle observations and therefore to their good impact on the 3D-Var effect.

ARPA The ARPA region is the one with the best data coverage in this study. Therefore we can see good results also for this observed profiles. There only is a little degradation of the temperature RMSE, but this is on the second significant digit, so it is not necessarily a significant difference in the results.

Rovinj In this case we have a very good quality assessment with no parameter getting worse after the assimilation.

Middle Ad. We have little changes in the middle Adriatic. This seems also to be a very good result.

Southern Ad. Same comment as the Middle Ad.

So the results based on the analysis of the full year seems to be encouraging, but we again repeat that for a complete insight of the action of the 3D-Var and on the details of the quality assessment, the reader is referred to the comments of the single seasons observations.

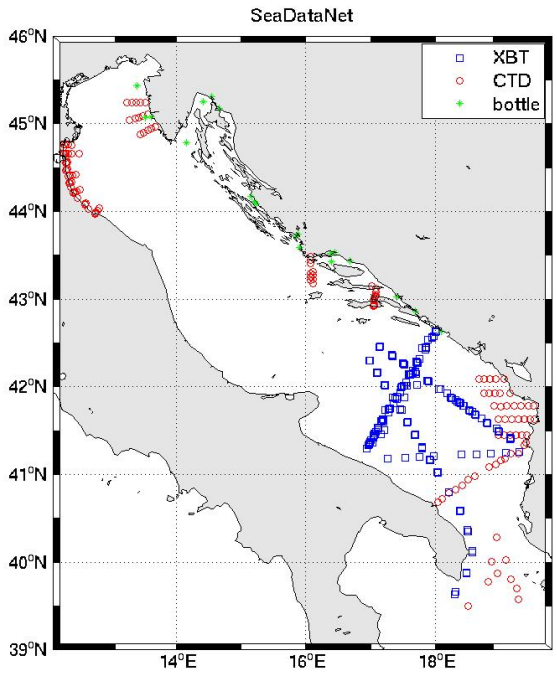


Figure 4.6: *Temperature-salinity profiles available in the Adriatic Sea for the entire 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.3	0.8	−0.3	1.3
XBT	−0.3	0.7		
CTD	−0.5	1.3	−0.3	1.3
bottle	−0.3	1.4	0.2	0.6
ARPA	−0.5	1.4	−0.4	1.7
Rovinj	−0.9	1.4	−0.2	0.4
Middle Ad.	0.9	1.7	0.4	0.9
Southern Ad.	−0.3	0.7	−0.05	0.4

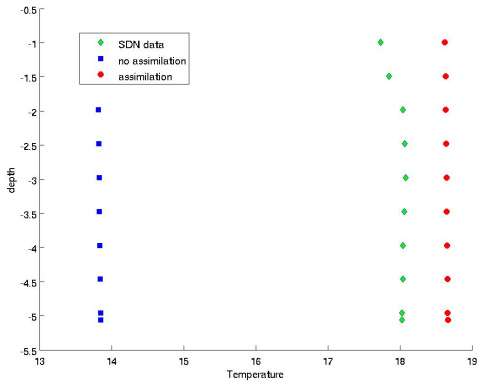
Table 4.9: *bias and RMSE between data and model without assimilation for the T-S profiles for the entire year 2006*

dataset	bias T	RMSE T	bias S	RMSE S
full dataset	−0.3	0.9	−0.16	1.2
XBT	−0.4	0.7		
CTD	0.018	1.6	0.16	1.2
bottle	−0.12	1.3	0.2	0.6
ARPA	−0.11	1.9	0.3	1.6
Rovinj	−0.4	1.1	−0.11	0.4
Middle Ad.	0.9	1.8	0.3	0.9
Southern Ad.	−0.4	0.7	−0.05	0.4

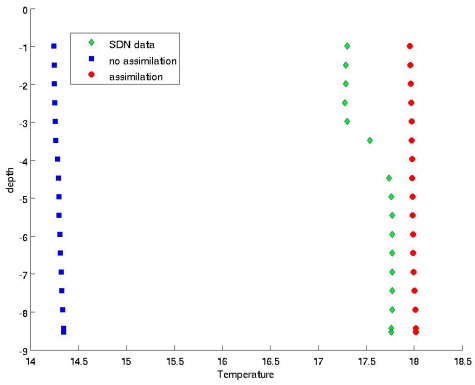
Table 4.10: *bias and RMSE between data and model with assimilation for the T-S profiles for the entire year 2006*

Punctual effect of the assimilation on T-S profiles

The following figures show the typical action of the DA on the model. In all the figures 4.7 , 4.8 we have the modeled (in blue and red) and observed (the green diamonds) T-S profiles. The blue squares represent the modeled profile provided by the model without the 3D-Var ("noassim" experiment) and the red circles are the points of the vertical profile modeled by the "assim" model implementation, meaning after the correction of the 3D-Var. The three profile in each figure clearly refer to the observed profile: the blue and the red dots are the models profiles interpolated on the observations points. Remember that the model output is a daily one, so the observation is not necessarily referred to the same time as the model (this problem is explained in the third chapter).



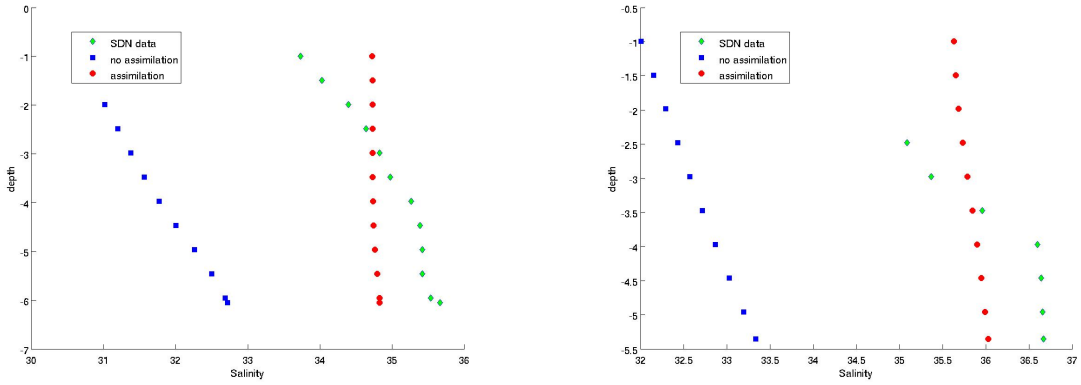
(a) *ARPA CTD temperature profile*
Day: 20061123
Longitude: 12.34
Latitude: 44.33.



(b) *ARPA CTD temperature profile*
Day: 20060127
Longitude: 12.36
Latitude: 44.41.

Figure 4.7: *Effect of the assimilation on two temperature profiles*

In any case we can clearly distinguish the improvement due to the 3D-Var: the profile is brought closer to the observations. The red profile blue profile is far away from the observations when compared to the red profile.
Another interesting thing to notice is a clue of what we stated before about the models smoothing the gradients, this time on the vertical dimension. Consider for example figure 4.7b we can clearly see that the observations have a raise in temperature at about 3-4 meters depth. This is correct since in winter we expect the temperature profiles having this typical pattern. In the considered region (ARPA) this effect could also be due to the freshwater of the Po, flowing south along the western coasts of the basin. In any case the both models have not this characteristic. The same thing can be tracked in figures 4.8.
This is an intrinsic limitation of the oceanographic models used today. Even with a DA scheme implemented together with the model, it always tends to diffuse to much, resulting in a generalized excessively smoother vertical profiles and horizontal field.
This results, for example, in great difficulties of the numerical models in recreating steep



(a) *ARPA CTD salinity profile*
Day: 20061023
Longitude: 12.29
Latitude: 44.61.

(b) *ARPA CTD salinity profile*
Day: 20060705
Longitude: 12.29
Latitude: 44.76.

Figure 4.8: *Effect of the assimilation on two salinity profiles*

thermocline or well marked salinity fronts, sometimes even if the data is melded in the model by a DA scheme.

4.1.2 Sea level observations from tide gauges

Now we apply the same analysis that we used for the T-S profiles to the tide gauges observations. The calculation must however be carried on with a slightly different methodology. In the previous chapter we said that the sea level measurement must be referred to some reference level and that this level is not the same for the model and for the observation. Therefore we can not compare directly the tide gauge data and the variable "SSH" (Sea Surface Height) calculated by the model. In the case of a satellite SLA observation the problem would be the same. Thus we have to remove the mean level from each sea level time series.
The full procedure is the sequent.
The observations are already arranged in a time series. At any observation time the scalar quantity measuring the sea level is measured. These observations are usually at a hourly frequencies, so we first have to daily average the tide gauges data in order to obtain a daily time series. From these time series we calculate the average level of the period analyzed and subtract it to the daily observation in order to get an unbiased series (i.e. it has zero mean).
To obtain the modeled time series of the same quantity we first have to interpolate on the horizontal the modeled sea level (which is a two dimensional field) on the tide gauges coordinates. Thus we get a modeled daily time series in correspondence of every station point. Finally we remove the average sea level from these time series too and we get the modeled, unbiased time series that can be compared to the observations with equations

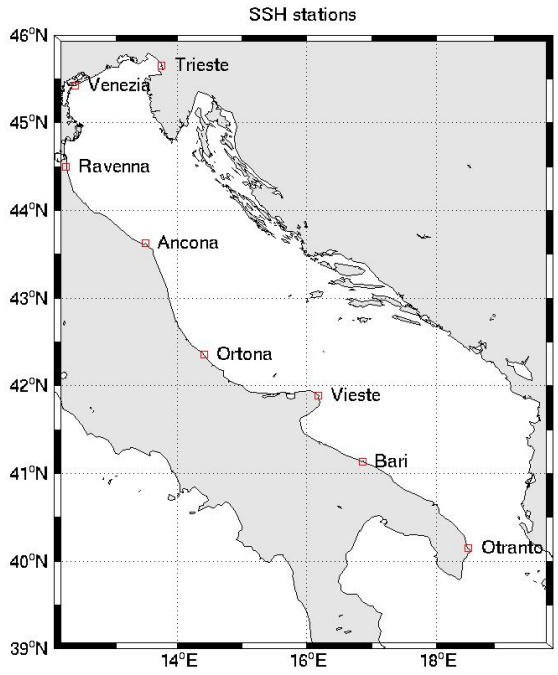


Figure 4.9: *The location of the ISPRA mareographic stations*

(4.1) and (4.3). (4.2) is not used to calculate and compare the biases because we are considering unbiased time series, so we are not interested in this quantity.
Remember that the tide gauges time series are not assimilated with the 3D-Var, so they are an independent test set for the Analysis System presented in this thesis. Simply speaking, if an improvement in the "assim" experiment is present in the comparison with these observations, it could be considered as a real upgrade of the analysis quality with respect to the "noassim" implementation.

The RMSE calculated from the misfits between the results from the two experiment and the observations are given in tables 4.11-4.15. These results are presented in a similar fashion as for the T-S vertical profiles: the single seasons are presented in tables 4.11,-4.14 and the whole year analysis is reported in table 4.15. In each table, the first line represent an overall quality assessment for all stations while the following rows represent the RMSEs updates for the single stations.
Remember that the geographic distribution of the stations is given in figure 2.7 and reproposed in figure 4.9 for simplicity.

Note that the dimension of the results reported in the tables and figures is meters, so the RMSE are at the centimetric level, which is a typical results from the comparison of the model with the observed sea level elevation.

We now make some considerations about the RMSEs reported in the tables. Keep

always in mind that for a full discussion about the sea level we should consider also the SLA observations, that are assimilated in the 3D-Var. This is not done in this thesis and in general there is no systematic study of the SLA observations in the Adriatic Sea at the present time. Another difficulty in commenting the tables is that the proposed results refer to a coastal observation. Here the SLA is in general a difficult observation and we are not sure about the correction coming from the assimilation of the direct assimilation of the satellite observations. Here the information for the correction given by the 3D-Var may come from (as it actually may always be) the spreading of the information from the T-S assimilated profiles and diffused to the whole state vector based on the shape of the matrix **B**. So it is not always easy to state the reason for a certain time series to be better reproduced or not by the "assim" experiment.

In any case, the results from the comparison with the tide gauges observations seems to be quite good, showing that the RMSE for the "assim" experiment is generally equal or lower than the "noassim" columns.
In particular we can make the following observations.

Winter Results from winter are given in table 4.11. From this table seems that in winter there is not a great improving or degradation of the modeled sea level with respect to the observations, except for the Ortona stations, that shows a smaller error in the "assim" experiment. We can be satisfied from the winter results.

Spring Results from spring are given in table 4.12. This seems to be one of the best season, both from the first line and from the single stations rows. For no station the quantities in the second column is greater than the numbers in the first one, while there is an improvement in the modeled time series for the Vieste station.

Summer From table 4.13 seems that in this season no improvement has been made in the solution. For some stations the RMSE is also greater in the second experiment.

Fall In table 4.14 again, we have an overall improvement of the solution, expressed in the RMSE estimator. In particular at Ravenna and at Vieste we have a smaller RMSE in the second experiment.

Whole 2006 Results for a whole year analysis are reported in table 4.15. As noted for the T-S profiles analysis, this is a summary table but is not as informative as the previous ones. So we must comment this last table carefully. However we can see that in general the second experiment is not doing much better or worse in recreating the station time series. In Ravenna we have a slight degradation of the sea level solution, but this, again must not be considered as a degradation of the overall solution.

Two examples of the temporal evolution of the misfits calculated between the modeled and observed time series is given in figures 4.10 - 4.19 for the Ancona and Bari stations. In these figures the blue and green lines represent the misfits between the observations

and the "noassim" and "assim" results respectively.
From the figures we can get more detailed information about the effect of the assimilation. First of all we can agin notice the order of magnitude of this kind of misfit. It is at the centimetric level, as usual for the present time ocean modeling capabilites.
The first two images are the misfits for the whole year. It is clear from this figures that the differences form the two experiment is not macroscopical; we can rather notice smaller departures especially located in the "peaks" of the graphs. Initially, these figures could be expected to show the gradual growing of the errors in the "noassim" experiment, while the "assim" results should have remain more controlled due to the assimilation. Instead there is no clear trend in any of the two solutions: the oscillate around a zero value.

The subsequent figures (4.12-4.19) are a zoom of the annual misfits evolution. Actually from these figure we can notice a more detailed behavior of the solution. As already anticipated, main differences can be found in correspondence of the relative maximum and minimum of the graph. In figures 4.14, 4.15 in particular, we may notice that in both time series there is clearly a period during which the "noassim" misfits are greater than those for "assim". The period is the first half of May. The finding of this kind of pattern in both graph may be encouraging but one can question about the significancy of these differences in the two misfits, their order of magnitude being in the sub-centimetric range.

Concluding we may say (considering in particular the RMSE results) that the tide gauges observations are equally well reproduced by the two model implementation, meaning that at least, the 3D-Var is not degrading the solution.

Station	*noassim* RMSE	*assim* RMSE
All stations	0.08	0.08
Ancona	0.08	0.08
Bari	0.07	0.07
Ortona	0.08	0.07
Otranto	0.06	0.06
Ravenna	0.09	0.09
Trieste	0.10	0.10
Venezia	0.09	0.09
Vieste	0.07	0.07

Table 4.11: *RMSE (expressed in meters) between model and tide gauges observations for winter 2006*

Station	*noassim* RMSE	*assim* RMSE
All stations	0.05	0.04
Ancona	0.05	0.05
Bari	0.04	0.04
Ortona	0.05	0.05
Otranto	0.04	0.04
Ravenna	0.05	0.05
Trieste	0.05	0.05
Venezia	0.05	0.05
Vieste	0.05	0.04

Table 4.12: *RMSE (expressed in meters) between model and tide gauges observations for spring 2006*

Station	*noassim* RMSE	*assim* RMSE
All stations	0.06	0.06
Ancona	0.07	0.07
Bari	0.05	0.05
Ortona	0.06	0.06
Otranto	0.03	0.04
Ravenna	0.07	0.08
Trieste	0.07	0.08
Venezia	0.07	0.08
Vieste	0.06	0.06

Table 4.13: *RMSE (expressed in meters) between model and tide gauges observations for summer 2006*

Station	*noassim* RMSE	*assim* RMSE
All stations	0.11	0.11
Ancona	0.11	0.11
Bari	0.09	0.09
Ortona	0.10	0.10
Otranto	0.09	0.09
Ravenna	0.13	0.12
Trieste	0.12	0.12
Venezia	0.12	0.12
Vieste	0.10	0.09

Table 4.14: *RMSE (expressed in meters) between model and tide gauges observations for fall 2006*

Station	*noassim* RMSE	*assim* RMSE
All stations	0.08	0.08
Ancona	0.09	0.09
Bari	0.07	0.07
Ortona	0.08	0.08
Otranto	0.07	0.07
Ravenna	0.09	0.10
Trieste	0.10	0.10
Venezia	0.09	0.09
Vieste	0.08	0.08

Table 4.15: *RMSE (expressed in meters) between model and tide gauges observations for the full 2006*

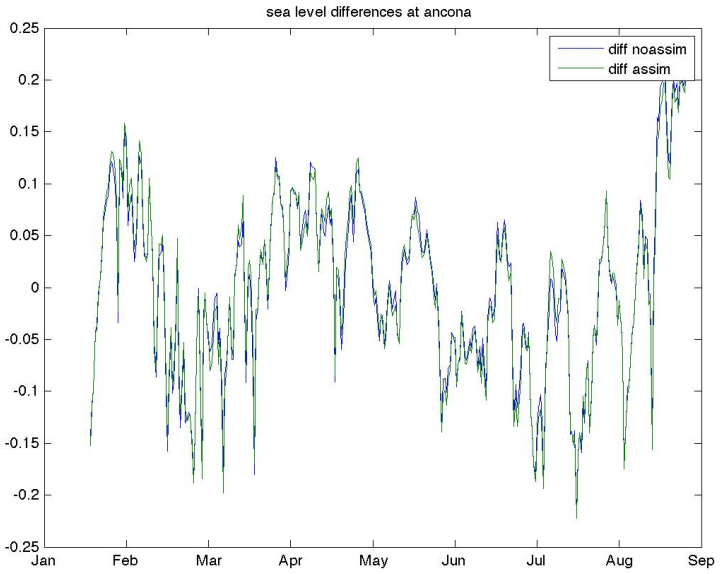


Figure 4.10: *Daily differences (in meters) in sea level between the two models and the observations at the Ancona station for the whole 2006*

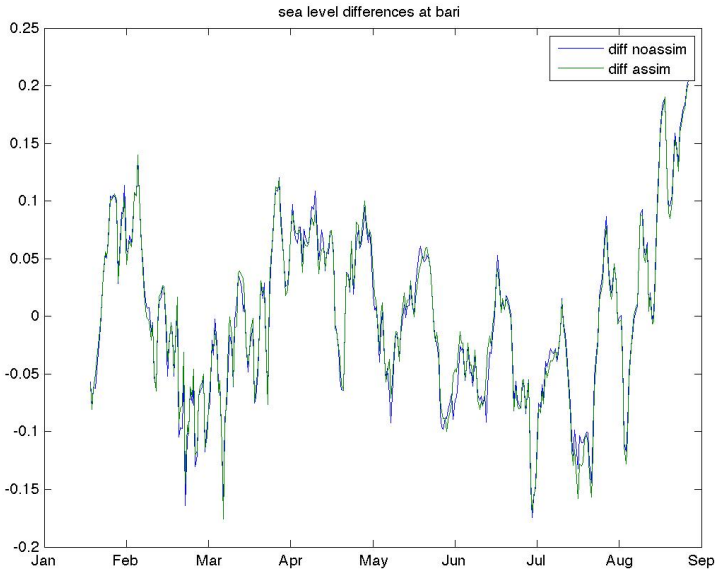


Figure 4.11: *Daily differences in sea level between the two models and the observations at the Bari station for the whole 2006*

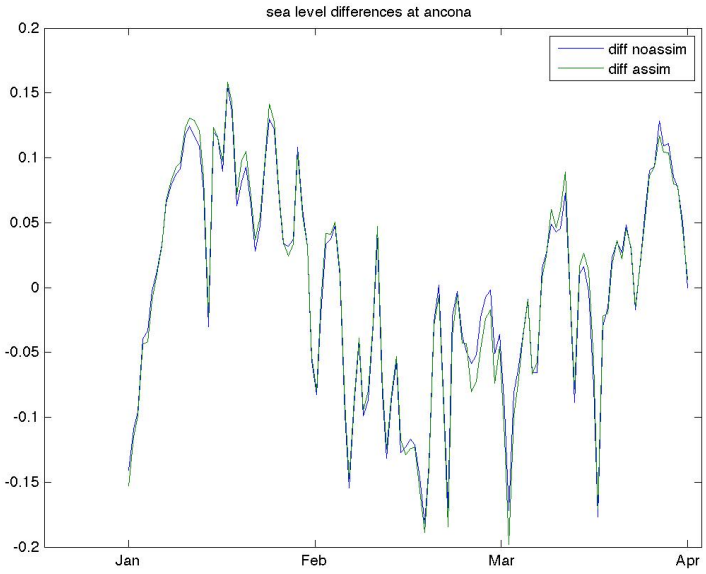


Figure 4.12: *Daily differences in sea level between the two models and the observations at the Ancona station for winter 2006*

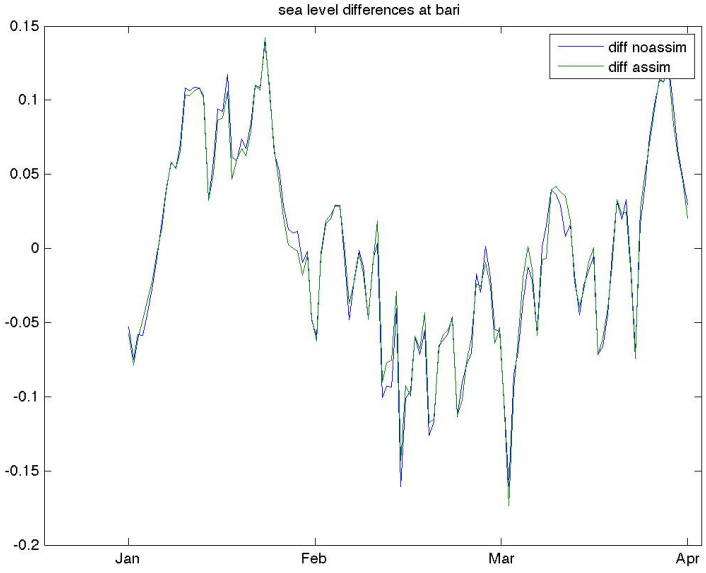


Figure 4.13: *Daily differences in sea level between the two models and the observations at the Bari station for winter 2006*

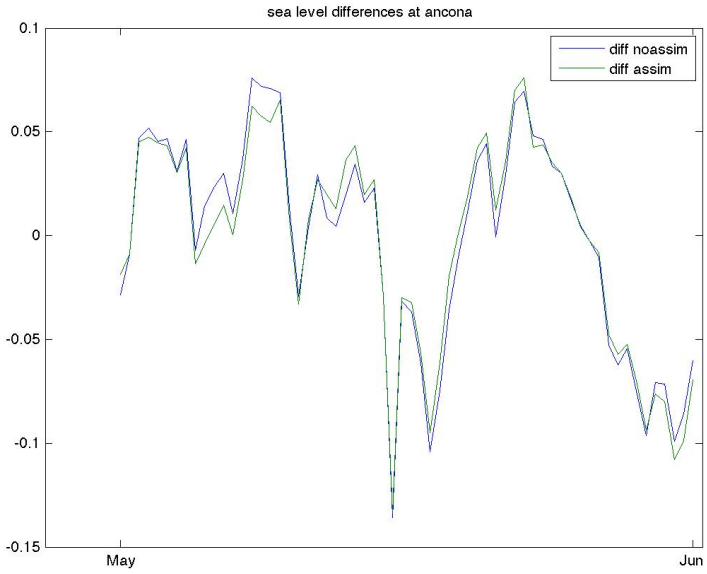


Figure 4.14: *Daily differences in sea level between the two models and the observations at the Ancona station for spring 2006*

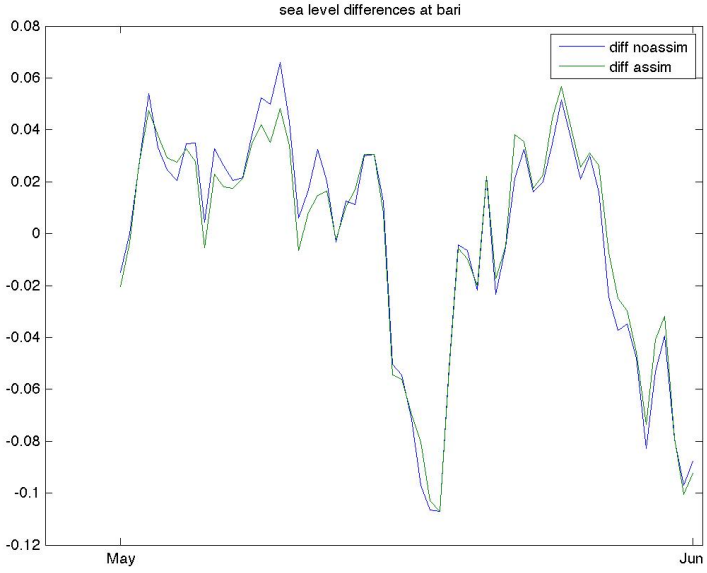


Figure 4.15: *Daily differences in sea level between the two models and the observations at the Bari station for spring 2006*

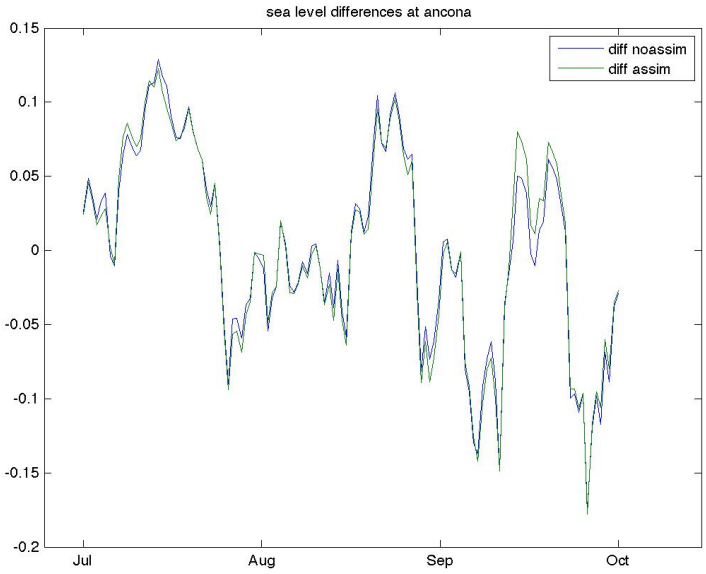


Figure 4.16: *Daily differences in sea level between the two models and the observations at the Ancona station for summer 2006*

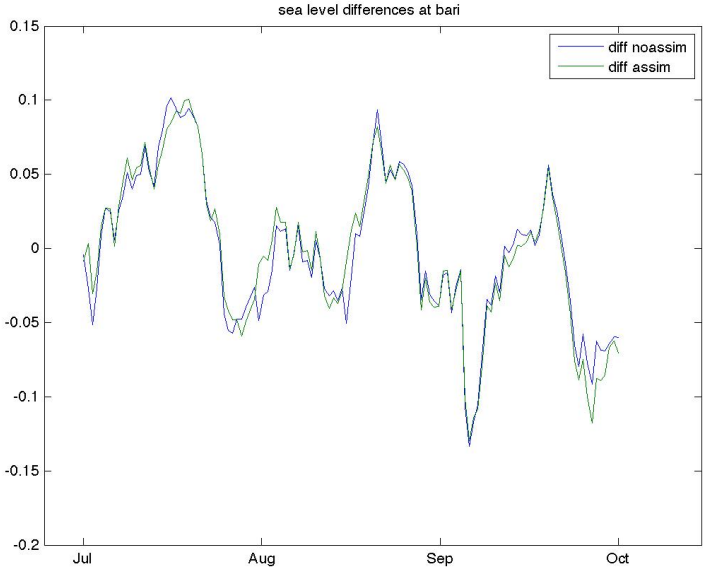


Figure 4.17: *Daily differences in sea level between the two models and the observations at the Bari station for summer 2006*

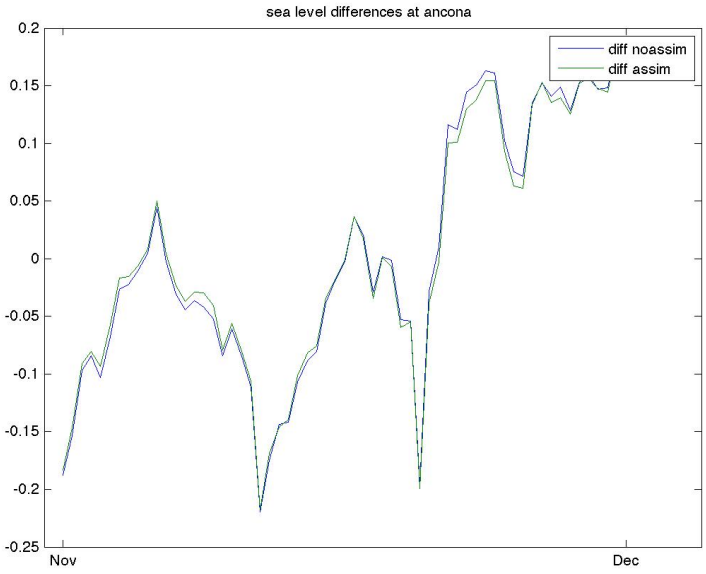


Figure 4.18: *Daily differences in sea level between the two models and the observations at the Ancona station for fall 2006*

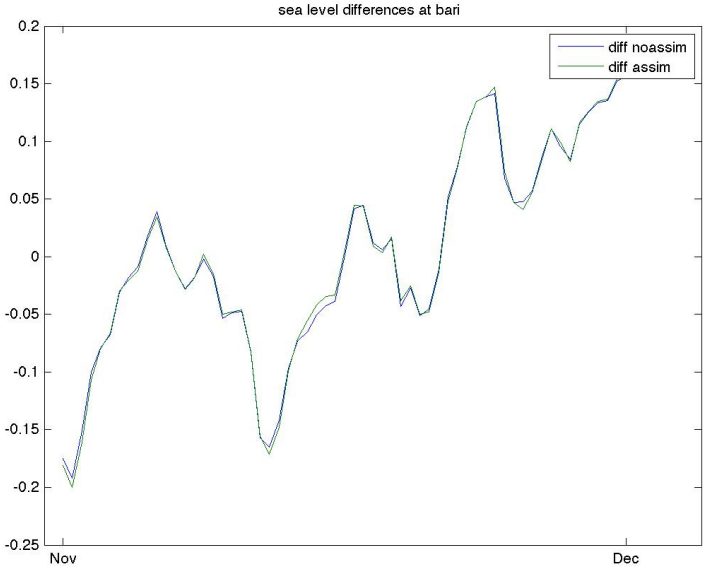


Figure 4.19: *Daily differences in sea level between the two models and the observations at the Bari station for fall 2006*

4.2 Assessment of the results

In this final section we will show same examples of the effect that the assimilation has on the horizontal temperature, salinity and surface elevation fields and therefore on the horizontal general circulation of the Adriatic Sea. This assessment is carried on in a quite qualitative way, showing snapshots of the fields for the "noassim" and "assim" model, trying to highlight the major differences by watching them.
More quantitative methods are possible, but here we don't make use of them, both for lack of time and because actually, the study of the fields, having a phenomenological knowledge of the dynamic in the Adriatic Sea is already an informative method for comparing the two implemented models. The following discussion are therefore made in light of the consideration made in the first chapter about the general features of the circulation in the Adriatic Sea. In particular the reader is referred to the figure 1.12 and the corresponding discussion.

We now present examples of the surface horizontal fields (at 2 m depth) and then the same fields (with the obvious exception of the sea level) at 75 m depth.

4.2.1 Surface salinity

Examples of surface salinity field are shown in figures 4.20 - 4.23 for March and May.
Consider for the moment figures 4.22 and 4.23, referring to march. What we wanted to show in this figures is the typical effect of a correct assimilation on the surface fields. First of all we may note in the figure referred to the "noassim" experiment that the AREG2 model alone is reproducing the freshwater spreading from the rivers mouths into the Adriatic Sea in a dynamically consistent way. Note the fresher areas in correspondence of the albanian rivers (in particular in front of the Bojana) and the massive presence of freshwater in the northern Adriatic, mainly due to the Po river and the spreading of these waters southward along the Italian coasts. This features are all in line with the general phenomenological considerations made in the first chapter.
Now turn out attention to figure 4.23. It is clear that in this figures the same general salinity pattern is present, but with more prominent features and stronger gradients. The freshwater rivers on the Albanian coasts is now spreading more into the Adriatic, indicating a slower diffusion of the salinity than the one suggested by the model alone. This concept has been already suggested in the previous sections: the model alone always tends to diffuse to much.
More interesting is the way the Po freshwater is extending from the discharging area into toward the open sea and South. The pattern here is more "irregular", suggesting some sort of henanced eddy activity. Besides there is a low salinity zone in front of the Emilia-Romagna's coasts that was totally absent in the "noassim" figure. This feature has probably been captured by the ARPA observing system and inserted into the model by the 3D-Var.
So in general we see a salinity field that is far less smooth than the one suggested by the

model alone. Similar considerations for the salinity fields in other month can be made. We will show how this fact is reflected in more energetic circulation.

Consider now figures 4.20 and 4.21. From the "noassim" figure we may make similar considerations as form figure 4.22. What is different is the "assim" figure. In this case we have an anomalously high salinity area in the northern Adriatic. This pattern too is probably due to the ARPA observations, but we are not confident that this is indicating a natural phenomena. It is more probably indicating that the 3D-Var is assimilating not well checked observations.
However it is interesting the fact that this feature is present from March to May in salinity and temperature field both at the surface and at 75 m (see the subsequent figures). This is interesting because no observations are present at 75 m in the northern Adriatic: this effect is probably due to the 3D-Var, that is spreading the information from the top towards the bottom, from one field to another, embedding it into the model dynamics. This is what we expect from a correctly designed Data Assimilation algorithm.
So we can say that, even if the information that the 3D-Var is using is not a correct one it is correctly diffusing it to the variables of the state vector not corresponding to observed quantities.

4.2.2 Surface temperature

Now we turn our attention to the surface temperature fields, plotted in figures 4.24 and 4.25 for the date 20060315, in figures 4.26 and 4.27 for the 20060415, in 4.28 and 4.29 for 20060715 and in 4.30 and 4.31 for the date 20061015.

Consider these last two figures for the moment. In the "noassim" figure we can see a typical surface salinity field for this season, showing, in particular a cold water currents along the Eastern basin of the coast and due to the input of the colder Mediterranean waters that are driven northward by the E-SAd coastal current.
On the opposite side of the basin and in the Northern Adriatic we found warmer waters, due to the summer warming of the surface of the Adriatic Sea.
In the "assim" figure, the same features are present, but, again, we found them to be more "extreme": the warm waters in the North are more confined along the coasts and many local minima or maxima that are present in figure 4.28, here are more pronounced. Note for example the high temperature lobe centered at about 42° N , 18.5°E, in the Southern Adriatic. This lobe was already reproduced by the model without assimilation, but here we found it bigger and warmer. As we will see, this lobe will also influence the surface circulation (figures 4.36 and 4.37).

In figure 4.24 the pattern is almost opposite as the one of figure 4.28. In this season, cold waters are concentrated in the Northern Ad. and along the Italian coasts while in the South Ad. and along the South Eastern coasts there is the spreading of the Mediterranean, warmer water. In this season, in fact, the Adriatic Sea is losing heat from the surface, therefore getting colder in the Northern part. Also in this figure is well distinguishable

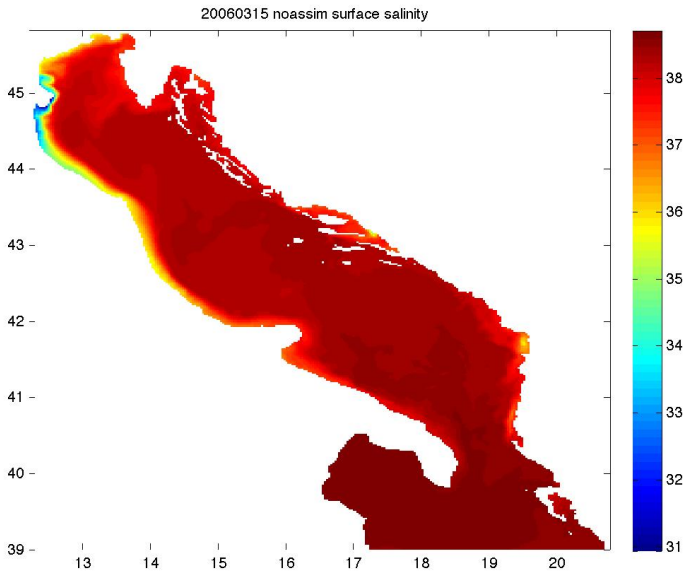


Figure 4.20: *Sea surface salinity from the "noassim" model on date 20060315. Color scale is expressed in PSU*

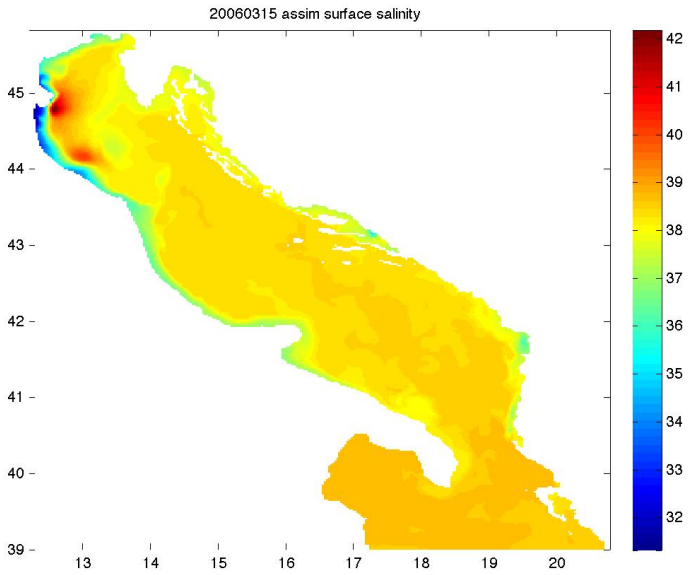


Figure 4.21: *Sea surface salinity from the "assim" model on date 20060315. Color scale is expressed in PSU*

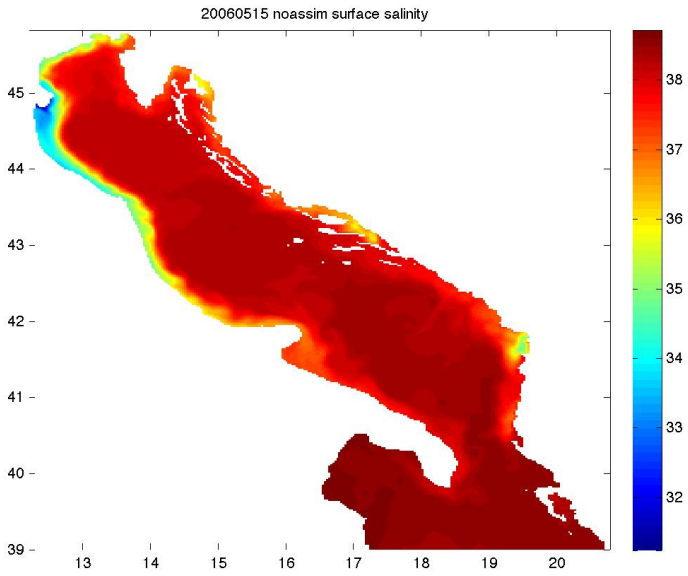


Figure 4.22: *Sea surface salinity from the "noassim" model on date 20060515. Color scale is expressed in PSU*

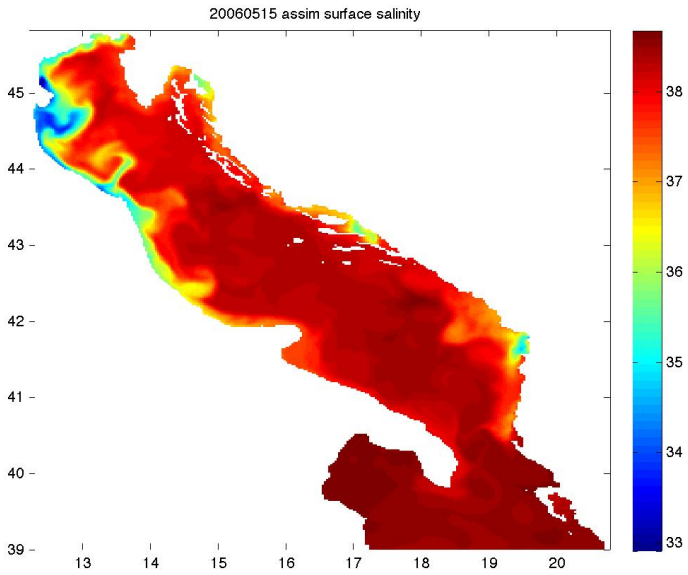


Figure 4.23: *Sea surface salinity from the "assim" model on date 20060515. Color scale is expressed in PSU*

the southern flow of North Adriatic (the Western Ad. coastal current) water along the Italian coasts, extending to the deep south of the basin.
In figure 4.25 we found the situation given for the same day from the model with the data assimilated by the 3D-Var. The general characteristics are the same of the previous figure, indicating the same known phenomenology. As already discussed, there is a strange high salinity-high temperature anomaly around the ARPA region. Although it is not clear wether it is a physical signal or an observation error, it is an information somehow consistent with the salinity anomaly for the same period. The same feature is present in figure 4.27: the high temperature anomaly is present till May in the analyzed fields.
A part from this anomaly, we again found greater horizontal gradients in the field calculated by the model with the assimilation.

In figure 4.26 we have the same general features of the figure 4.24, with the cold waters of the Northern Adriatic spreading southward. Clearly the winter is still getting the heat out of the sea, lowering more the surface temperatures. The cold water flow along the Italian coasts is clearly present down to the Gargano Peninsula, suggesting a strong western Ad. current and a situation similar to that shown in the upper panel of figure 1.11.
The corresponding "assim" temperature field has the same feature along the Italian coasts and in general a more irregular pattern in the rest of the Adriatic Sea. A part from the Northern temperature anomaly already discussed, we can notice the presence of many high temperature lobes in the middle and southern Adriatic. This rougher field again suggests the presence of a more energetic surface circulation in the "assim" experiment than in the "noassim" one. We will confirm this fact in the next figures.

Finally, consider now figures 4.30 and 4.31. Here we can clearly see the more smoothness of the "noassim" field, that, even with a finer scale than the "assim" solution, does not show the same spatial variability. Note in particular the minimum in the central Adriatic (centered at about 43°N, 16°E): it is present in both figures, but in the model without assimilation it is reproduced with a much greater spatial extension and a less pronounced amplitude. This will have consequences on the reproduction of the Middle Adriatic gyre.

4.2.3 Surface elevation and horizontal circulation

Finally we now present the results for the surface circulation of the Adriatic Sea. These results are shown in figures 4.32 and 4.33 for the day 20060315, 4.34 and 4.33 for the 20060315, 4.36 and 4.37 for the 20060715 and 4.38 and 4.39 for the 20061015.
In these figures we have a schematic representation of the currents given by the arrows, superimposed to a sea surface elevation plot.
In discussing these plots, we have to keep in mind what we saw for the corresponding temperature fields.

Let's start with the 20060315 (figures 4.32 and 4.33). We Already discussed the surface temperature for the same date, founding, as usual, a more rough field after the assimilation. Considering the surface elevation we can see a quite similar situation. In the "assim"

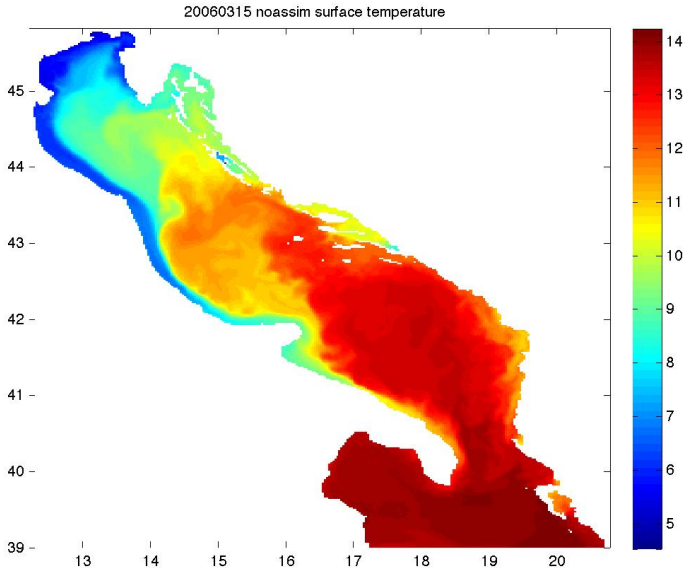


Figure 4.24: *Sea surface temperature from the "noassim" model on date 20060315. Color scale is expressed in °C*

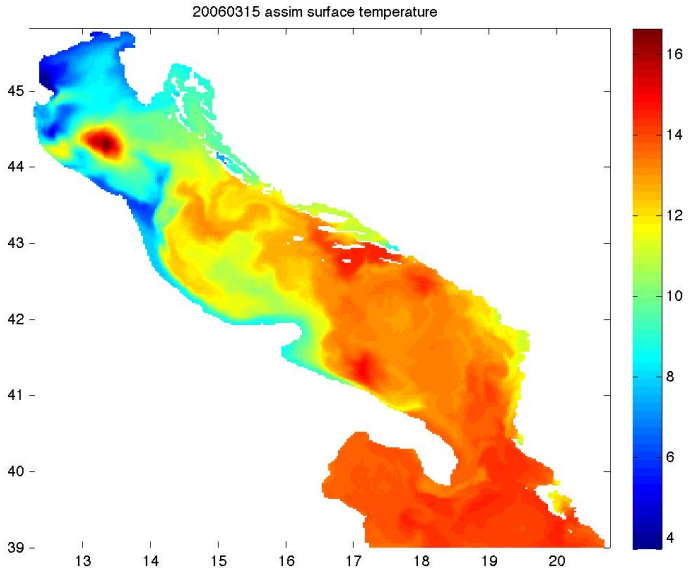


Figure 4.25: *Sea surface temperature from the "assim" model on date 20060315. Color scale is expressed in °C*

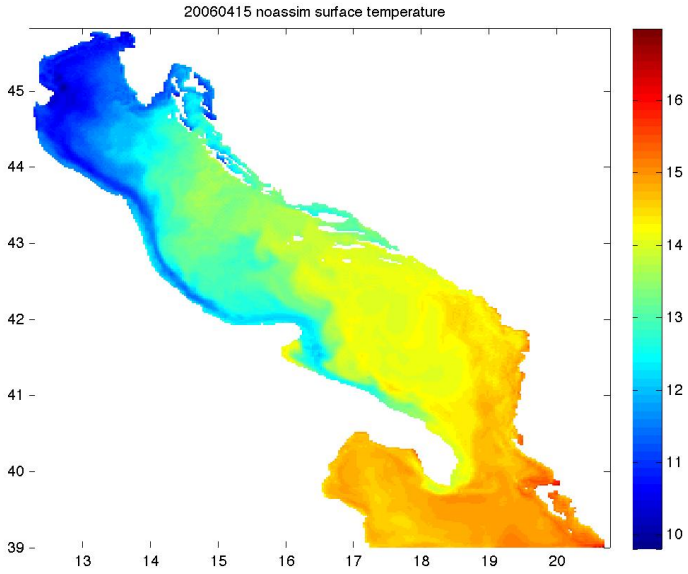


Figure 4.26: *Sea surface temperature from the "noassim" model on date 20060415. Color scale is expressed in °C*

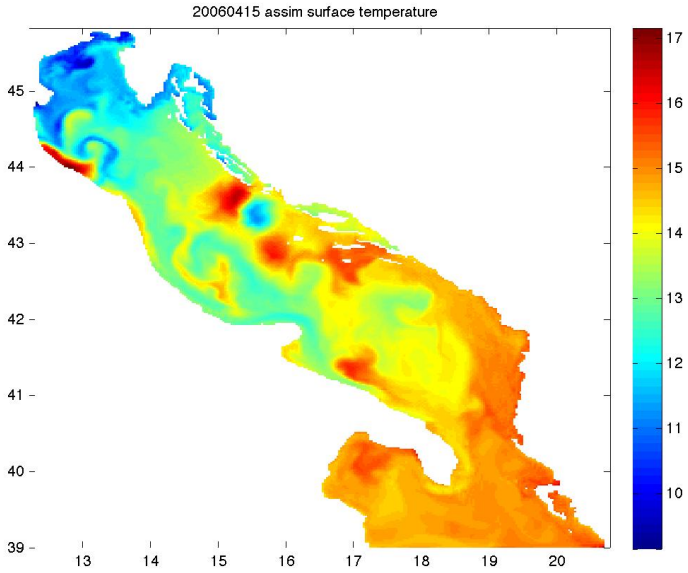


Figure 4.27: *Sea surface temperature from the "assim" model on date 20060415. Color scale is expressed in °C*

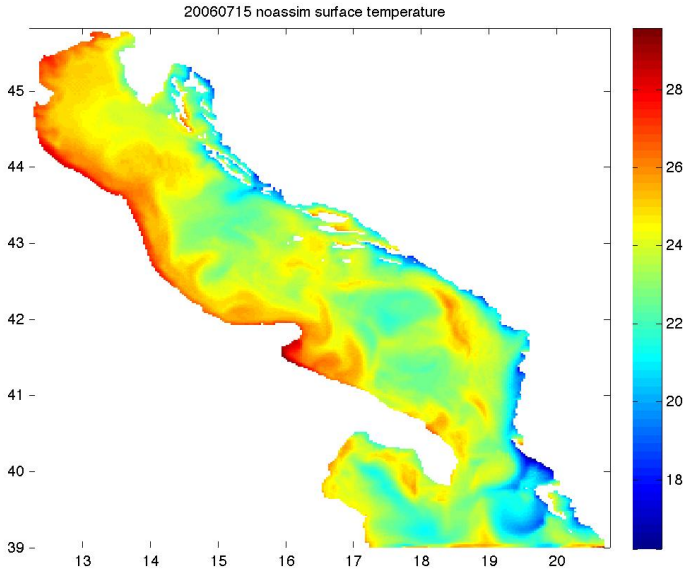


Figure 4.28: *Sea surface temperature from the "noassim" model on date 20060715. Color scale is expressed in °C*

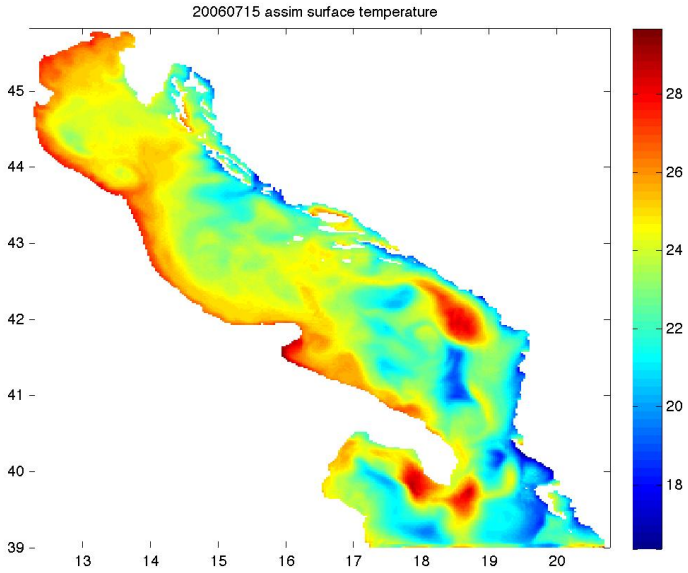


Figure 4.29: *Sea surface temperature from the "assim" model on date 20060715. Color scale is expressed in °C*

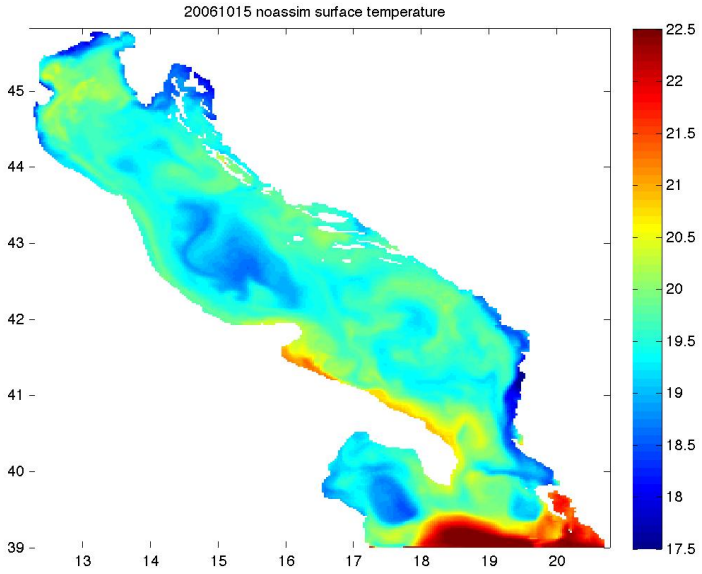


Figure 4.30: *Sea surface temperature from the "noassim" model on date 20061015. Color scale is expressed in °C*

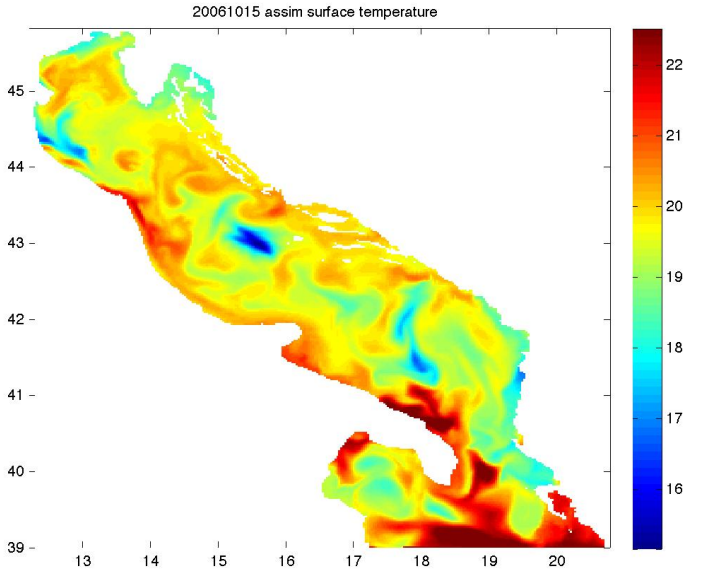


Figure 4.31: *Sea surface temperature from the "assim" model on date 20061015. Color scale is expressed in °C*

elevation we found deepest and narrower minima in the northern Ad. with respect to the "noassim" field. Besides we have highest maximum in the middle Ad. and a more irregular situation in the southern sub-basin.
The consequences on the surface currents are clear, and easily explained. Steepest gradients in the surface temperature and elevation fields results in stronger currents. In the northern Adriatic we have, centered at about 44°N, 14°E, a single extended well defined anticyclonic structure. In the "assim" figure we find that at the same location the minima are now two, deepest and with stronger currents around the central points. The warm water lobe found in that area in figure 4.25 has certainly a role in the splitting of the gyre of figure 4.32.
Right south of this structure we can find in the "assim" figure an anticyclonic gyre around a local maximum in sea level that was not well defined in the "noassim" figure.
In any case there is physical consistency between the lowering of the minimum and the henancement of the cyclonic currents and, on the opposite, between an higher maxima and a stronger anticyclonic gyre around it. The assimilation is therefore able to maintain this basic feature of the oceanic circulation.
In the southern sub-basin we note a stronger and more irregular current field in the "assim" model, as a consequence of the lack of a defined structure both in temperature and in sea surface elevation as opposite to the "noassim" model, that exhibits create a quite strong anticyclonic gyre, evidently corrected by the assimilation of the data.

Concerning the month of April (figures 4.34 and 4.33) we again have a much smoother sea surface field before the assimilation and a much less energetic current field. We can however distinguish, in the "noassim" figure, well defined Western and Eastern coastal currents, together with a middle Adriatic cyclonic structure and a southern anticyclonic core. After the assimilation we can see that this last structure is almost deleted, while the coastal currents and the middle Ad. gyre are enforced. We may suspect that the strong gradients in the temperature field due to the presence of two warm lobes and a colder one (at a latitude of 42-44°N along the eastern coasts of the basin) play an important role in the circulation pattern that has developed after the assimilaton.

On the 20060715 (figures 4.36 and 4.37) we have an interesting situation.
Along the eastern coasts of the Adriatic Sea the model without assimilation produced a strong coastal southward current that reached the Gargano peninsula. This current is split in the "assim" experiment in a northern Ad. coastal current and a middle Ad. coastal current this splitting is due to a strongly enforced MAd gyre and the development of a northward current in the Northern basin after the assimilation.
In the southern part we have the growing of strong local cyclonic gyres, connected with the increase in surface elevation and to the warm and cold lobes from the temperature field of figure 4.29. As a result, the pattern of the "noassim" circulation in the southern Ad. is similar to that produced by the "assim" model, with the latter exaggerating the features produces by the first one.
We may notice that in both experiments there is the development of a strong eastern coastal current in the southern Ad. but no northward western coastal current is devel-

oped in this period. Therefore, the SAd gyre is not present in these simulations.

Finally consider the month of October. Remember that this should be the period of stronger baroclinic activity in the basin, with the presence of all the typical features of figure 1.12. In figure 4.38 the model without the assimilation is developing a weak NAd gyre and a MAd gyre, a well developed eastern coastal current flowing southward, but there is no SAd gyre and the western coastal current is flowing south.
After the assimilation we have stronger NAd and MAd gyre and the ddisappearance of the high level anomaly in the sea level in the southern Adriatic. Therefore a SAd gyre is formed and a northward current on the western side of the basin is developed as a consequence of the growth of the gyre.
Therefore the assimilation enforced the typical features of the general circulation of the basin.

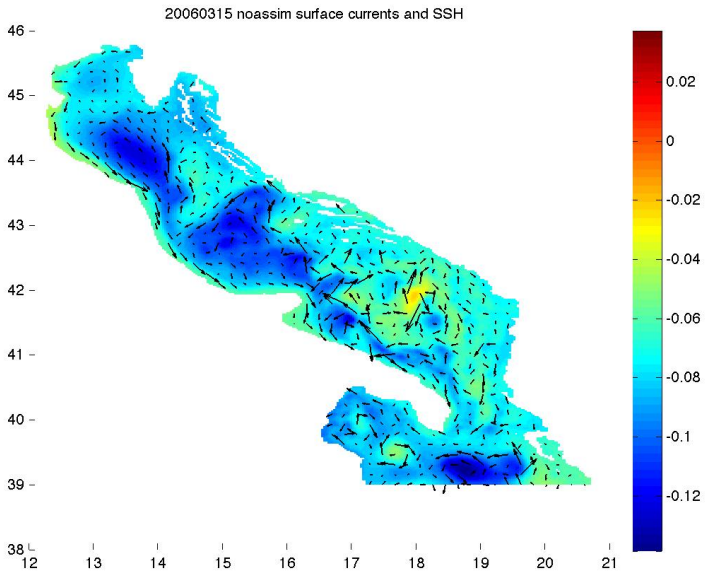


Figure 4.32: *Sea surface elevation and currents from the "noassim" model on date 20060315. Color scale is expressed in meters. Maximum velocity modulus is 0.5 m/s*

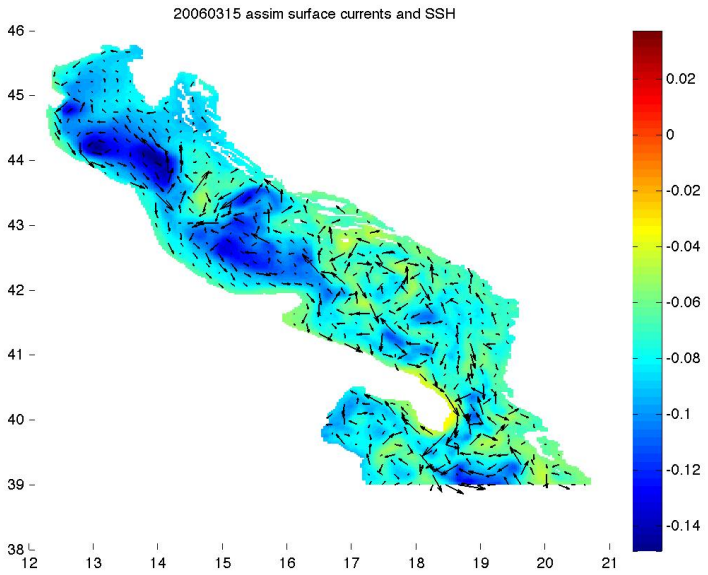


Figure 4.33: *Sea surface elevation and currents from the "noassim" model on date 20060315. Color scale is expressed in meters. Maximum velocity modulus is 0.5 m/s*

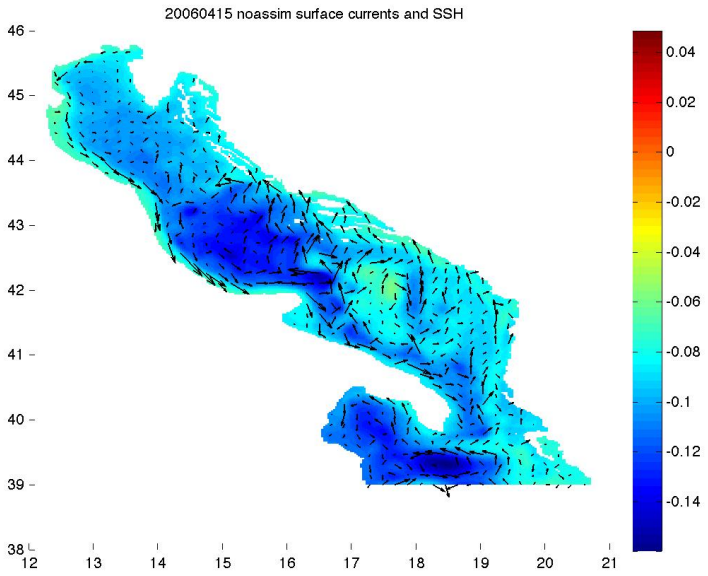


Figure 4.34: *Sea surface elevation and currents from the "noassim" model on date 20060415. Color scale is expressed in meters. Maximum velocity modulus is 0.4 m/s*

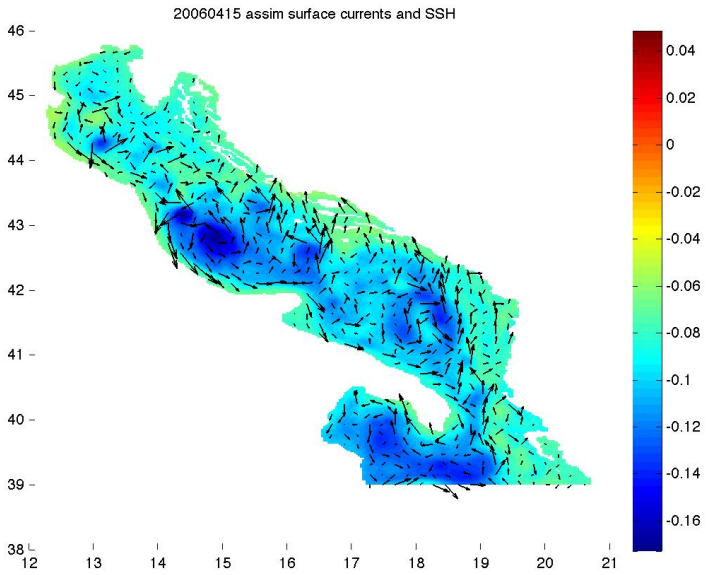


Figure 4.35: *Sea surface elevation and currents from the "noassim" model on date 20060415. Color scale is expressed in meters. Maximum velocity modulus is 0.4 m/s*

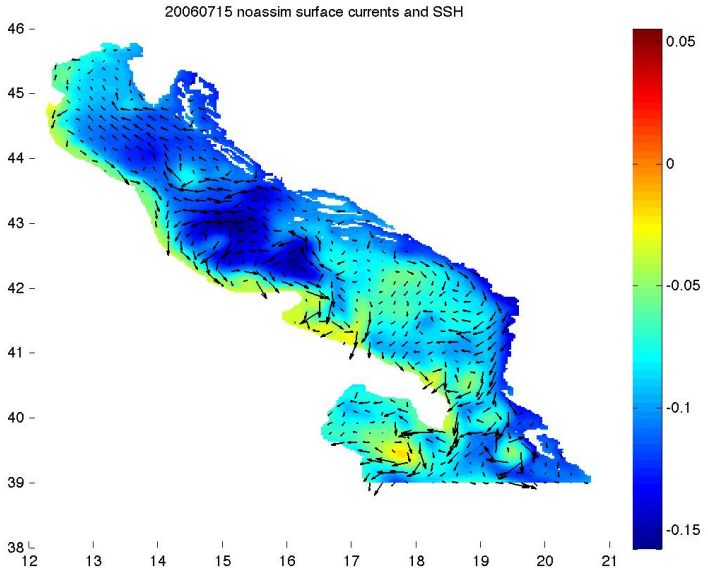


Figure 4.36: *Sea surface elevation and currents from the "noassim" model on date 20060715. Color scale is expressed in meters. Maximum velocity modulus is 0.6 m/s*

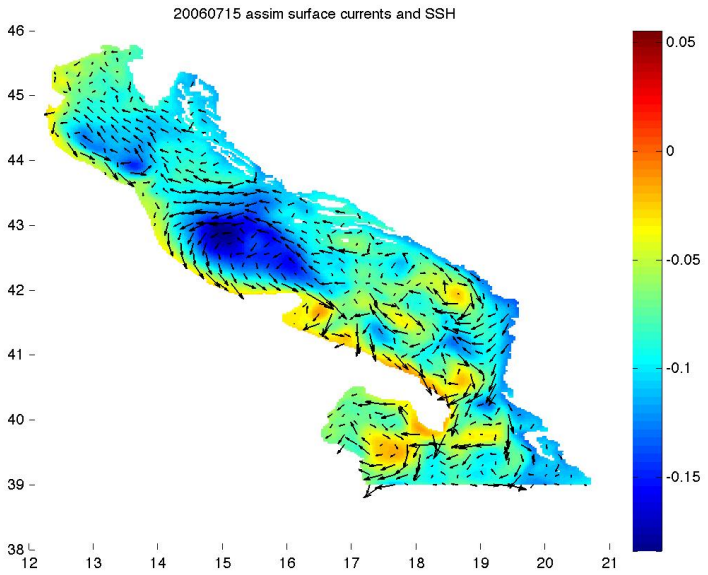


Figure 4.37: *Sea surface elevation and currents from the "noassim" model on date 20060715. Color scale is expressed in meters.*

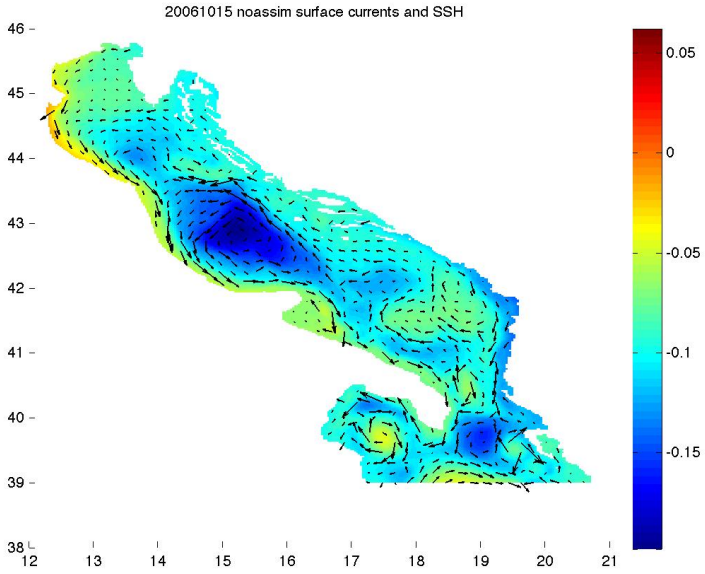


Figure 4.38: *Sea surface elevation and currents from the "noassim" model on date 20061015. Color scale is expressed in meters. Maximum velocity modulus is 0.5 m/s*

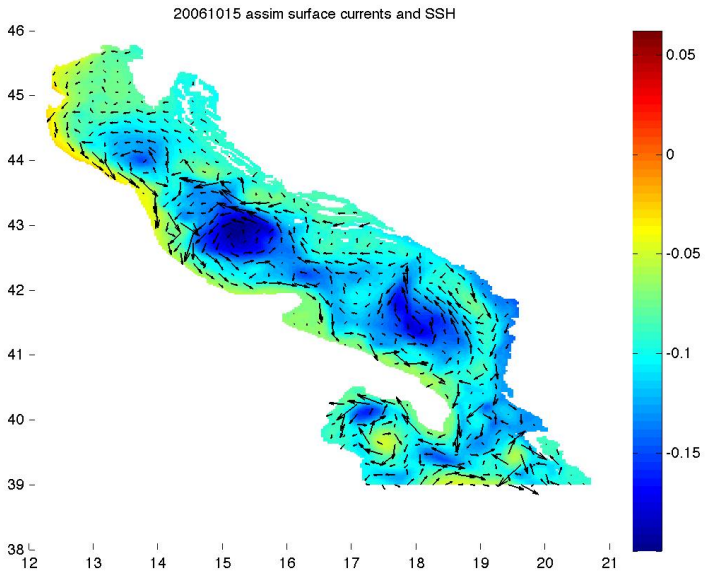


Figure 4.39: *Sea surface elevation and currents from the "noassim" model on date 20061015. Color scale is expressed in meters. Maximum velocity modulus is 0.5 m/s*

4.2.4 Horizontal salinity at 75 m depth

We now briefly discuss the horizontal field at 75 m depth, starting from salinity.
We already discussed the high T-S anomaly in the surface fields. This anomaly is roughly present from March to June at the surface in both fields.
Here the same anomaly is seen in the salinity field as shown in the figures 4.40 and 4.41 for middle may. As stated before, we are not saying that this could be a real feature of the circulation of that period: it is probably due to an error in the quality control of the data. But what is interesting is that the field presented here is not directly observed at the latitude where the anomaly is present. Therefore what is bringing this information at 75 m is the 3D-Var that is spreading an information taken at the surface to the other variables of the model state vector, therefore creating a new model state that is dynamically consistent with the original information.

The discussed anomaly is however found in other period too, therefore suggesting that the same information is sometimes present in the data (probably ARPA data).

Other deep salinity fields may be discussed in the same fashion of the surface fields: the assimilation generally tends to roughening the salinity, thus generating steeper gradients.

4.2.5 Temperature and velocity fields at 75 m

Finally we can consider the deep temperature and current fields.
We can for example consider the date of 20060415 for which the two experiments results are given in figure 4.42 and 4.43.
In the first image we can notice the presence of a weak MAd gyre in correspondence of a low temperature area.
After the assimilation the entire temperature field is rougher, with enhanced maximum and minima. Moreover the deep part of the MAd gyre is now well defined and there also is the formation of a strong eastern Ad. current flowing northward.

In other examples we can find a similar pattern when considering the two experiments, concluding that in the deep Adriatic too, the 3D-Var contributes to enforced the general features of the horizontal fields, therefore giving more energy to the general circulation.

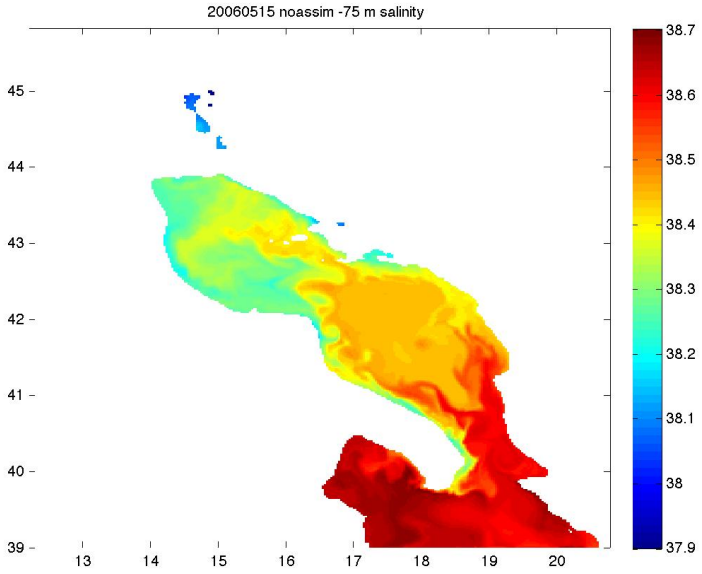


Figure 4.40: *Horizontal salinity at 75 m depth from the "noassim" model on date 20060515. Color scale is expressed in PSU*

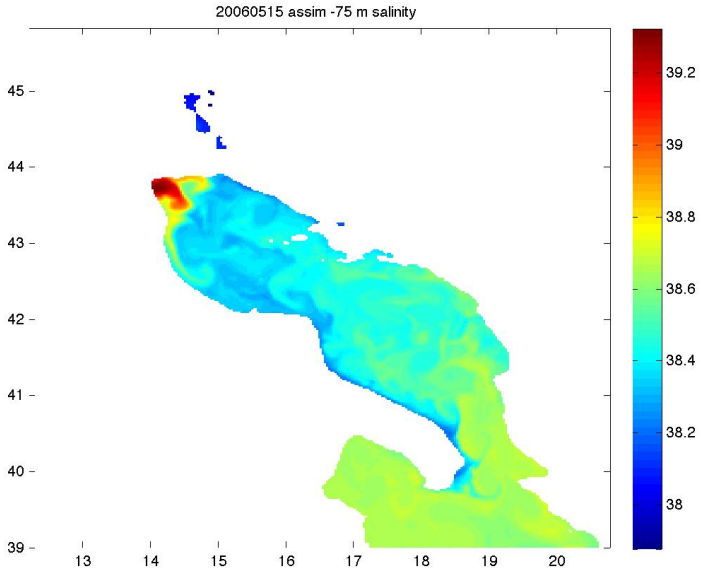


Figure 4.41: *Horizontal salinity at 75 m depth from the "assim" model on date 20060515. Color scale is expressed in PSU*

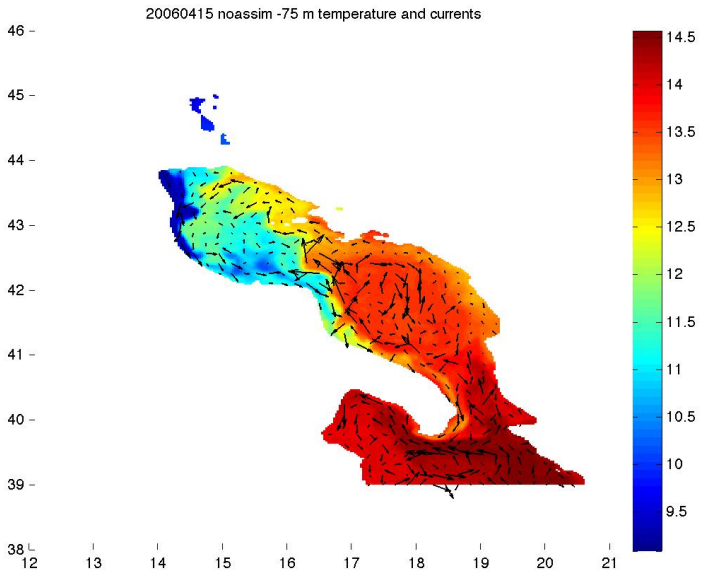


Figure 4.42: *Horizontal temperature and currents at 75 m depth from the "noassim" model on date 20060415. Color scale is expressed in °C. Maximum velocity modulus is 0.3 m/s*

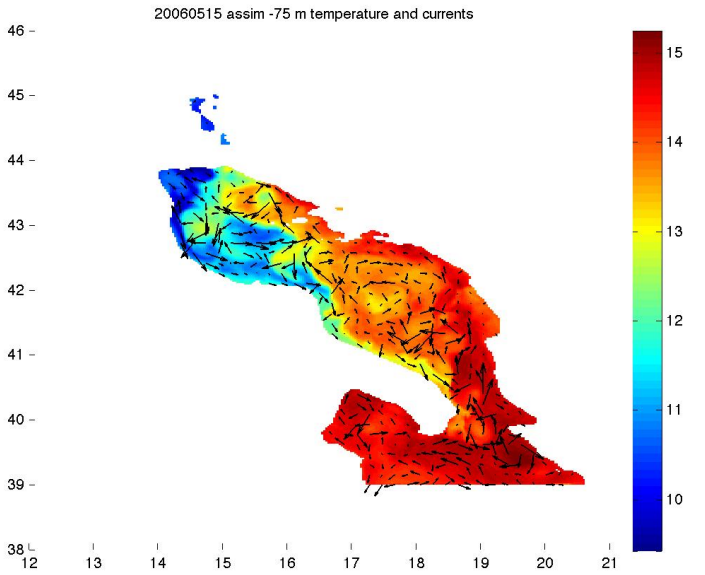


Figure 4.43: *Horizontal temperature and currents at 75 m depth from the "noassim" model on date 20060415. Color scale is expressed in °C. Maximum velocity modulus is 0.3 m/s*

Chapter 5

Summary and Conclusions

In the present work we tested a 3D-Var assimilation algorithm in a realistic implementation of the Adriatic Sea Forecasting System. Therefore the model implemented in the study is the AREG2 model currently working operatively for the forecasting activities at GNOO (INGV), delivering daily averages of the 2D and 3D field of physical interest in the Adriatic Sea.
In particular we applied the 3D-Var assimilation scheme proposed in [16] which include several techniques for solving the assimilation problem in an efficient way, on operational time scales.
The detailed implementation of the analysis system used in this study is described in chapter 3.

The data we assimilated in the model are both in-situ and satellite data of the target year 2006. In situ data come from the SeaDataNet dataset and comprises XBT, CTD and bottle vertical profiles of temperature and salinity. Satellite observations are SLA observations measured from the three satellites available in 2006: Geosat2, Envisat and Jason1.
Aside from these assimilated observations we kept out of the process an independent set of sea level observations from coastal tide gauges.
A general description about the measuring devices can be found in the first chapter while the details about the data sets is given in chapter 2.

The experiments that we made for evaluating the impact of the assimilation on the model output are simple. Both experiments consist in a model run of the full year 2006, but in one experiment we assimilate all the available SLA and in-situ observations and in the other we don't assimilate anything. Therefore the two experiment are called "noassim" and "assim".

The results from the two experiments are presented and analyzed in chapter 4. We first calculate bias and RMSE between the in-situ T-S profiles and the two model results. This comparison has not statistical relevance as a test of the assimilation algorithm, since we are comparing the result of the 3D-Var method with the data used in the algorithm. This

comparison, however, is useful to understand the effect of the DA on the vertical profiles of the two experiments. This comparison shows that the overall agreement between the modeled vertical profile obtained by interpolation on the observation point and the data has improved after the assimilation, especially for what concerns the salinity profiles. The regional and seasonal details of this comparison are given in chapter 4.
Then we made the same analysis with the tide gauges observations. This comparison is a more objective one in estimating the quality of the Analysis System because the data used for the calculation of the misfit are not used in the assimilation process. The misfits are then used to calculate RMSEs for the two experiments. This test provided suggested that the model with the assimilation reproduces the sea level time series with uncertainties with the same order of magnitude as the model without the assimilation. Again, the details about the seasonal results for each stations is given in chapter 4.

Finally we compared the horizontal fields of salinity, temperature, sea level and circulation of the two experiment, in order to give a visual interpretation of the effects of the assimilation. The comparison is done at the surface and at 75 m depth.
What emerged is that the assimilation has the wanted effect on the analyzed fields. In particular the 3D-Var recreates the horizontal gradients that the model alone is not able to simulate because of its natural tendency of unrealistically diffusing salinity and temperature. This results in a more energetic circulation pattern and in the presence of features in the "assim" results that are not always present in the results of the "noassim" experiment. A part from this, thanks to what probably is an error in the observations, we could conclude that the 3D-Var assimilated the data in the correct way: given an observation of a certain parameters of the state of the sea, the assimilation algorithm should correct the whole state vector in order to embed the data into the modeled fields in a statistically and dynamically consistent way. Therefore an observation of (say) a temperature profile at a certain point, should generate a correction of the salinity and velocity field, even if there are no direct observation of these quantities.

Therefore we may make the following conclusions:

- The 3D-Var assimilation results in a reduced bias and RMSE for in- situ observations, especially for the salinity;
- The 3D-Var alter significantly the surface circulation pattern. The typical feature of the Adriatic Surface circulation are generally better defined after the assimilation
- On the basis of the surface temperature and salinity field and of the underwater salinity we can state that, as expected, the model tend to smooth the surface gradients. After the assimilation a more energetic field is revealed

Furhter development

Here we have tested the 3D-Var algorithm and founded that it could be, in principle, provide better estimates of the future state of the sea.

But other test are advised, before a full operational implementation in AFS could be thought.
For example, the 3D-Var could be tested on other sets of observations, not used in this work, like satellite measured Sea Surface Temperature (SST) or in-situ observations from floats and drifters. The 3D-Var flexibility makes it an ideal model for assimilating any kind of observations.

Furthermore the 3D-Var assimilation scheme itself could be improved in time. For example, a 4D-Var could be implemented and tested. This would arise completely different computational problems in which some of the approximation that are usually done in order to save same computational cost (like the tangent linear hypothesis for the operators M and H) could be completely wrong in the sea and therefore resulting in uncorrected analyses. This problem is particular serious in high resolution ocean models, like the regional or sub-regional models (like the AREG2) and therefore more refined, and expensive methods may be required in the optimization process.

Based on the observation made in the first chapter other two aspects that could be improved in order to get better forecasts from the system are the numerical model (or better, of the *physics* implemented by the model) and the observing system.

Bibliography

[1] Estubier A. and M. Lévy. *Quel schéma numérique pour le transport d'organismes biologiques par la circulation océanique*. Note Techniquesdu Pôle de modélisation, Institut Pierre-Simon Laplace; 81 pp. 2000.

[2] J R Apel. *Principles of Ocean Physics*. Vol. 38. International Geophysics Series. Academic Press, 1987.

[3] A. Artegiani, R. Azzolini, and E. Salusti. "On the dense water in the Adriatic Sea". In: *Ocean. Acta.* 12 (1989), pp. 151–160.

[4] A. Artegiani et al. "The Adriatic Sea general circulation. Part I: Air Sea interaction and water mass structure". In: *J. Phys. Oceanogr.* 27 (1997), pp. 1492–1514.

[5] A. Artegiani et al. "The Adriatic Sea General Circulation. Part II: Baroclinic circulation structure". In: *J. Phys. Oceanogr.* 27 (1997), pp. 1515–1532.

[6] A. F. Blumberg and G. L. Mellor. "A description of a three-dimensional ocean circulation model". In: *Three-dimensional Coastal Ocean Circulation Models, Coastal Estaurine Sci.* 4. Amer. Geophys. Union. Washington: Heaps, N. S., 1987, pp. 1–16.

[7] F. Bouttier and P. Courtier. *Data assimilation concepts and methods*. Lecture notes. ECMWF, 1999.

[8] K. Bryan. "A numerical method for the study of the circulation of the world ocean". In: *Journal of computational Physics* 4 (1969), pp. 347–376.

[9] R. H. Byrd et al. "A limited memory algorithm for bound constrained optimization". In: *SIAM journal on Scientific Computing* 16 (1995), pp. 1190–1208.

[10] M. D. Cox. *A primitive equation, 3-dimensional model of the ocean*. Tech. rep. 1. GFDL Ocean Group, 1984.

[11] A. Crise, S. Querin, and V. Malačič. "A strong bora event in the Gulf of Trieste: a numerical study of wind driven circulation in stratified conditions with a pre-operational model". In: *Acta Adriatica* 47 (suppl) (2006), pp. 185–206.

[12] P. De Mey and M. Benkiran. "A multivariate Reduced-order Optimal Interpolation Method and its application to the Mediterranean Basin-scale Circulation, In Ocean Forecasting". In: *Ocean Forecasting*. Ed. by Nadia Pinardi and J. Woods. 2002, pp. 281–305.

[13] E. Demirov et al. “Assimilation scheme of the Mediterranean Forecasting System: operational implementation”. In: *Annales Geophysicae* 21 (2003), pp. 189–204.

[14] S. Dobricic et al. “Daily oceanographic analyses by the Mediterrranean basin scale assimilation system”. In: *Ocean Sciences* 3 (2007), pp. 149–157.

[15] S. Dobricic et al. “Mediterranean Forecasting System: An improved assimilation scheme for sea-level anomaly and its validation”. In: *Q. J. R. Meteorol. Soc.* 131 (2005), pp. 3627–3642.

[16] Srdjan Dobricic and Nadia Pinardi. “An oceanographic three- dimensional variational data assimilation scheme”. In: *Ocean Modelling* 22 (2008), pp. 89–105.

[17] M. Drevillon et al. “The GODAE/Mercator-Ocean global ocean forecasting system: results, applications and prospects”. In: *Journal of Operational Oceanography* 1 (2008), pp. 51–57.

[18] William J. Emery and Richard E. Thomson. *Data Analysis Methods in Physical Oceanography*. second and revised edition. Elsevier, 2001.

[19] R. A. Flather. “A tidal model of the northwest European continental shelf,” in: *Mém. Soc. R. Sci. Liège* 6(X) (1976), pp. 141–164.

[20] Nick P. Fofonoff and Robert C. Millard. “Algorithms for computation of fundamental properties of seawater”. In: *Unesco Tech. Pap. in Mar. Sci.* 44 (1983), p. 53.

[21] A. Fournier et al. “An Introduction to Data Assimilation and Predictability in Geomagnetism”. In: *Space Sci. Rev.* 155 (2010), pp. 247–291.

[22] R. Giering and T. Kaminski. “Recipies for adjoint code construction”. In: *ACM Transaction on Mathematical Software* 24 (1998), pp. 437–474.

[23] A. Grezio and N. Pinardi. “Data assimilation of temperature and salinity profiles in the Adriatic Sea regional model”. In: *Acta Adriatica* 47 (suppl) (2006), pp. 149–168.

[24] A. Guarnieri et al. “Modelling the baroclinic circulation with tidal components in the Adriatic Sea”. In: *Journal of Geophysical Research* 14 (2012).

[25] A. Guarnieri et al. “The Adriatic Basin Forecasting System new model and system developement. Coastal to Global Operational Oceanography: Achievements and Challenges”. In: *Proceeding of 5th EuroGOOS Conference*. Exeter, UK: Daln, H., Fleming, N., and Petersson, S., 2010.

[26] S. Hellerman and M. Rosenstein. “Normal monthly wind stress over the world ocean with error estimates”. In: *J. Phys. Oceanogr.* 13 (1983), pp. 1093–1104.

[27] N. G. Jerlov. *Marine Optics*. Vol. 14. Elsevier Oceanography Series. New York: Elsevier Sci., 1976.

[28] M. Kwizak and A. J. Robert. “A semi-implicit scheme for grid point atmospheric models of the primitive equations”. In: *Montlhy Weather Review* 99 (1971), pp. 32–36.

[29] D. R. Legates and C. J. Wilmott. "Mean seasonal and spatial variability in a gauge corrected global precipitation". In: *International Journal of Climatology* 10 (1990), pp. 121–127.

[30] A. C. Lorenc. "Analysis methods for numerical weather prediction". In: *Quarr. J. R. Met. Soc.* 112 (1986), pp. 1177–1194.

[31] A. C. Lorenc. "Development of an operational variational assimiliation scheme". In: *Journal of the Meteorological Society of Japan* 75 (1997), pp. 339–346.

[32] A. C. Lorenc. "Iterative analysis using covariance functions and filters". In: *Q. J. R. Meteorol. Soc.* 118 (1992), pp. 569–591.

[33] A. C. Lorenc and O. Hammon. "Objective quality control of observations using Bayesian methods - Theory, and practical implementation". In: *Quarr. J. R. Met. Soc.* 114 (1988), pp. 515–543.

[34] G. Madec et al. *Ocean General Circulation Model reference manual*. Note du Pole de modelisazion. Institut Pierre-Simon Laplace (IPSL). France,

[35] V. Malačič and B. Petelin. "Numerical modeling of the winter circulation of the Gulf of Trieste (northern Adriatic)". In: *Acta Adriatica* 47 (suppl) (2006), pp. 207–217.

[36] G. L. Mellor. "An equation of state for numerical models of ocean and estuaries". In: *J. Atmos. Ocean. Technol.* 8 (1991), pp. 609–611.

[37] G. L. Mellor. *Introduction to Physical Oceanography*. Springer, 1996.

[38] G. L. Mellor. *USERS GUIDE for A THREE-DIMENSIONAL, PRIMITIVE EQUATION, NUMERICAL OCEAN MODEL*. Program in Atmospheric and Oceanic Sciences. Princeton University. Princeton,

[39] G. L. Mellor and T. Yamada. "Development of a turbulence closure model for geophysical fluid problems." In: *Rev. Geophys. SPace Phys.* 20 (1982), pp. 851–875.

[40] Lorenz E. N. "Deterministic nonperiodic flow". In: *Journal of Atmospheric Sciences* 20(2) (1963), pp. 130–141.

[41] P. Oddo and N. Pinardi. "Lateral Open Boundary Conditions for Nested Limited Area Models: Process selective approach." In: *Ocean Modelling* 20 (2008), pp. 134–156.

[42] P. Oddo, N. Pinardi, and M. Zavatarelli. "A numerical study of the interannual variability of the Adriatic Sea (2000–2002)". In: *Sci. Total Environ.* 353 (1-3) (2005), pp. 39–56.

[43] P. Oddo et al. "A nested Atlantic-Mediterranean Sea general circulation model for operational forecasting". In: *Ocean Sci. Discuss.* 6 (2009), pp. 1093–1127.

[44] P. Oddo et al. "The Adriatic Basin forecasting system". In: *Acta Adriatica* 47 (suppl) (2006), pp. 169–184.

[45] M. Orlič et al. "Nested modeling of the East Adriatic coastal waters". In: *Acta Adriatica* 47 (suppl) (2006), pp. 219–245.

[46] I.M. Ovchinnikov. "On the water balance of the Mediterranean Sea." In: *Oceanology* 14 (1974), pp. 198–202.

[47] N. A. Phillips. "A coordinate system having some special advantages for numerical forecasting". In: *J. Metoeorol.* 14 (1957), pp. 184–185.

[48] N. Pinardi and E. Masetti. "Variability of the large scale general circulation of the Mediterranean Sea from observations and modelling: a review". In: *Palaeogeography, Palaeoclimatology, Palaeoecology* 158 (2000), pp. 153–173.

[49] N. Pinardi et al. "The Mediterranean ocean forecasting system: first phase of implementation (1998-2001)". In: *Ann. Geophys.* 21 (2003), pp. 3–20.

[50] P. M. Poulain. "Adriatic Sea surface circulation as derived from drifter data between 1990 and 1999". In: *J. Mar. Syst.* 29 (2001), pp. 3–32.

[51] W. H. Press et al. *Numerical Recipes: The Art of Scientific Computing.* 3rd Edition. Cambridge University Press, 2007.

[52] F. Raichic. *Note on flow rates of the Adriatic rivers.* Tech. rep. Istituto Talassografico Sperimentale di Trieste, 1994.

[53] F. Raichic. "On fresh water balance of the Adriatic Sea." In: *J. Mar. Syst.* 9 (1996), pp. 305–319.

[54] R. K. Reed. *An evaluation of formulas for estimating clear sky insulation over the ocean.* Technical report. NOAA-ERL 352-PMEL, 1975.

[55] R. K. Reed. "On estimating insolation over the ocean". In: *J. Phys. Oceanogr.* 1 (1977), pp. 854–871.

[56] Gunter Seeber. *Space Geodesy.* Second and revised edition. de Gruyter, 2003.

[57] J. Smagorinsky. "Some historical remarks on the use of nonlinear viscosities". In: ed. by Cambridge University Press. New York: Galperin, B. and Orszag, S. A., 1993, pp. 3–36.

[58] Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation.* SIAM, 2005.

[59] UNEP. *Mediterranean Action Plan (MAP), Implications of Climate Change for the Albanian Coast.* Technical Reports Series No. 98. 1996.

[60] May P. W. *A brief explanation of Mediterranean heat and momentum flux calculations NORDA Code 322.* Miss. 7 Nav. Oceanogr. Atmos. Res. Lab. Stennis Space Center, 1986.

[61] M. Zavatarelli and N. Pinardi. "The Adriatic Sea modelling system: a nested approach". In: *Ann. Geophys.* 21 (2003), pp. 345–364.

[62] M. Zavatarelli, N. Pinardi, and A. Maggiore. *Winter density compensation in the northern Adriatic Sea. A numerica model process study.* Ecosystem Research Report No. 32, EUR 18834, European Commission, Brussels: 141-155. 1999.

[63] M. Zavatarelli et al. “Diagnostic and prognostic model studies of the Adriatic Sea circulation. Seasonal variability”. In: *Journal of Geophysical Research* (2002).