Alma Mater Studiorum · University of Bologna

Department of Physics and Astronomy
Department of Biological, Geological and Environmental Sciences

Master Degree in Science of Climate

Multivariate analysis of sea level and thermohaline characteristics in the Mediterranean Sea

Supervisor:
Prof. Nadia Pinardi

Co-supervisor:
Prof. Paolo Oddo
Dr. Mario Adani

Submitted by:
Rebecca Zaja

Academic Year 2023/2024

Abstract

The precise estimation of three-dimensional oceanic parameters, such as temperature and salinity, presents significant challenges. The coarse sampling of in situ observations restricts our ability to directly analyze the three dimensional thermohaline structure of the oceans. Satellite remote sensing offers extensive spatial coverage, but it is limited to surface-level data. To mitigate this limitation, reanalyses merge model outputs with both in situ and satellite observations. This research intends to study how to improve reanalyses by identifying optimal techniques for extracting subsurface structural information from satellite data in a simpler setting than the full data assimilation and numerical model system.

The primary emphasis of this research is on the assimilation of sea level anomaly data. The methodology adopted is based on the framework established by Ezer and Mellor (1991) and Adani et al. (2011). In the first approach, the problem is addressed by employing correlation factors that establish a connection between anomalies in sea surface elevation and anomalies in subsurface temperature and salinity profiles. These correlation factors are derived from time averages, making the selection of an appropriate time averaging interval crucial. This study explores various time intervals to determine which one yields the most favorable outcomes.

The second approach by Adani et al. (2011) utilizes Empirical Orthogonal Functions (EOFs) to derive the correlation factors. Specifically, bi- or tri-variate EOFs are applied to surface elevation data and vertical profiles of temperature and/or salinity. This strategy enables the identification of a vertical pattern that accounts for a significant portion of the variability in the temperature and salinity profiles.

The various methodologies have been applied to investigate the reconstructions of salinity and temperature profiles for the years 2019-2020 across four distinct locations in the Mediterranean Sea, namely the Alboran, Tyrrhenian, Adriatic, and Levantine Seas, as well as to examine the structures of two different mesoscale eddies, specifically anticyclonic and cyclonic. Initially, sea level anomalies derived from reanalysis data are utilized to mitigate multiple sources of error arising from data inconsistencies. Subsequently, sea level anomalies from L4 satellite data are employed, yielding results that are comparable to those obtained from model data reconstructions.

The reconstruction of tri-variate EOFs, utilizing monthly correlation coefficients derived solely from data spanning 2019 to 2020, demonstrates optimal results. This finding implies that the multi-variate statistics and a small temporal range connected to our target reconstructions is the optimal choice. For the future, extending this approach to the use of Neural Networks could help to uncover the non-linear couplings between sea level and thermohaline subsurface structures using the longer time series of data available.

Contents

Chapter 1

Introduction

The role of oceans in shaping human life is remarkably significant. A considerable number of individuals inhabit coastal regions, with fishing being a primary livelihood for many. Additionally, oceans function as critical heat and carbon reservoirs for the Earth, thus contributing significantly to climate regulation and stability. As such, it is crucial to obtain accurate estimates of oceanic parameters, particularly temperature and salinity. However, the vast expanse of the ocean and the constraints associated with in situ observations hinder our ability to directly assess the thermohaline structure of marine environments. While remote sensing observations can offer extensive coverage, the transmission of electromagnetic waves through seawater is limited, restricting our data acquisition to the ocean's surface. To overcome this limitation, reanalyses, which combine model outputs with both in situ and satellite observations, can be utilized. This study seeks to improve the efficacy of reanalyses by identifying optimal methodologies for extracting subsurface structural information from satellite data. This chapter serves as a concise introduction, emphasizing the concept of reanalyses, the specific area of investigation, and the equations that dictate the dynamics and vertical structure of eddies. In section 1.1, a detailed examination of the three components of reanalyses is presented, with a specific focus on the uncertainties inherent in both models and observations. Section 1.2 delves into the Mediterranean Sea, the primary area of interest, concentrating on its thermohaline characteristics and the diverse scales of oceanic phenomena observable in this region. Section 1.3 presents an overview of mesoscale dynamics and eddies, as the estimation of their structure is an aspect of this research. Lastly, section 1.4 delineates the aims of the thesis, providing a comprehensive overview of the content to be explored in subsequent chapters.

1.1 Reanalyses

Ocean reanalyses represent comprehensive reconstructions of the oceanic state, achieved through the integration of ocean models constrained by atmospheric surface forces and informed by ocean observations via data assimilation methods [1]. This approach generates four-dimensional fields that provide the most reliable estimates of the ocean's condition, serving as a vital resource for understanding past and elucidating the underlying processes that should be captured by numerical models.
These computations were originally designed to track and analyze climate change, yet they also enable the investigation of critical signals and processes that remain

partially unobservable, such as subsurface and bottom circulation. Furthermore, they can be utilized as initial and boundary conditions in modeling efforts [2].
The generation of reanalyses is achieved through the combination of a model, empirical observations, and data assimilation approaches. The next sections will deliver a brief introduction to all three elements.

1.1.1 Ocean models and cascade of uncertainty

As previously indicated, a model constitutes the primary component of a reanalysis. Ocean models are numerical frameworks that depict the properties of the ocean and its circulation dynamics [3]. These models have significant applications, such as analyzing ocean currents, understanding the dynamic interactions with the atmosphere, studying sea ice behavior, and assessing land runoff. They are also instrumental in the study of transport of biogeochemical materials, predicting climate variations due to both natural and anthropogenic factors, examining pollution dispersal, and managing fisheries and other aspects of the biosphere [4].
The global ocean's physical and dynamical characteristics are mathematically described through equations that represent the conservation of momentum, energy, and mass for a viscous, gravity-forced fluid that flows in a rotating reference system. They are typically represented by the Navier-Stokes equations or their approximations, so by a set of non-linear partial differential equation. These equations, often referred to as the equations of motion, provide a framework for understanding the temporal changes in temperature, salinity, and ocean currents. Additionally, ocean models require suitable boundary and initial conditions that delineate the interactions between the ocean, atmosphere, land, and sea ice, as well as the ocean's state when the model is initiate [5]. The solution of these equations cannot be analytical, so a discretization is needed to permit computers to solve them and it is performed using numerical methods aimed at solving partial differential equations. This discretization, along with the selection of boundary and initial conditions, constitutes a significant source of uncertainty in ocean modeling.
Actually the uncertainties present in ocean modeling can be attributed to several factors. Firstly, numerical discretizations can lead to miscalculations, which are contingent upon the specific numerical scheme employed and are ultimately unavoidable. Second, the initialization of models and the assignment of parameter values rely on ocean data; however, the availability of raw measurements is often limited in terms of both coverage and accuracy, and these measurements are frequently processed to focus on information within a specific scale window, leading to inaccuracies in initial conditions and parameters. The third source of uncertainty is related to the approximate nature of models that describe the interactions between the ocean and the Earth system, which results in imprecise ocean boundary conditions. The fourth source is linked to the physical approximations done to reduce computational costs, where explicit calculations are restricted to a narrow range of spatial and temporal scales. Consequently, influences from scales outside this window are either disregarded, parameterized, or incorporated at the boundaries, introducing further errors. Finally, uncertainties also stem from the incomplete understanding of processes within the scale window, which results in approximate representations or additional parameterizations. [6].
It is also essential to understand that these sources of uncertainty are interconnected

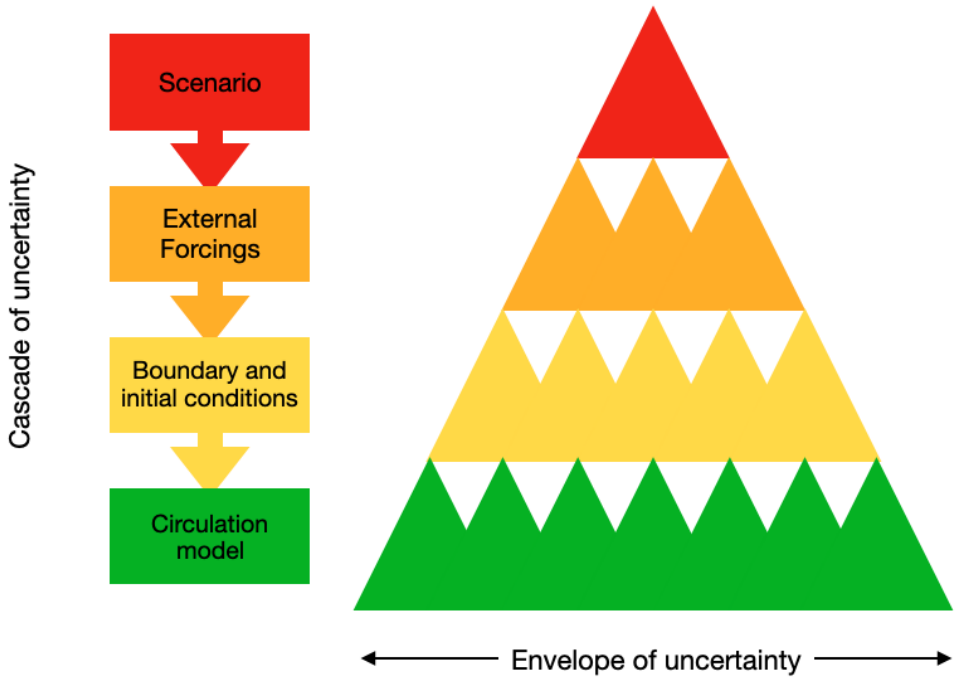


Figure 1.1: Schematics of the cascade of uncertainty. Figure redrawn from [7]

rather than independent. Changes in one source can significantly impact others, leading to what is termed a cascade of uncertainty [7]. In this context, a change in the considered scenario results in alterations in the oceanic forcing, which subsequently impacts the boundary and initial conditions, ultimately leading to modifications in the circulation model itself. As the process unfolds, the range (or envelope) of uncertainty broadens at each stage, thereby amplifying the errors produced by the model as can be seen in figure 1.1.

1.1.2 Ocean observations and satellite oceanography

Ocean observations represent the second critical element in the process of reanalysis. A robust ocean observational system typically includes a combination of in-situ measurements and remote sensing methodologies. The former refers to data collected through sensors installed on vessels, buoys, moorings, autonomous system and coastal monitoring stations. These measurements track variations in oceanic properties over time and depth at designated locations, providing direct assessments of these properties [8]. The efficacy of in-situ data however is constrained by inadequate spatial and temporal coverage, as well as challenges in accessing remote areas.

In contrast, satellite data offers the advantage of providing consistent, repeated, and global measurements of critical oceanic parameters over extended periods [9]. Satellite oceanography is the process of gathering information from a distance, through the utilization of remote sensors mounted on satellites and aircraft, which capture and record energy that is either reflected or emitted [10]. Notably, all sensors utilized in ocean-observing satellites rely on electromagnetic radiation to monitor the ocean's surface. Specifically, when an electromagnetic signal of a certain type emanates from

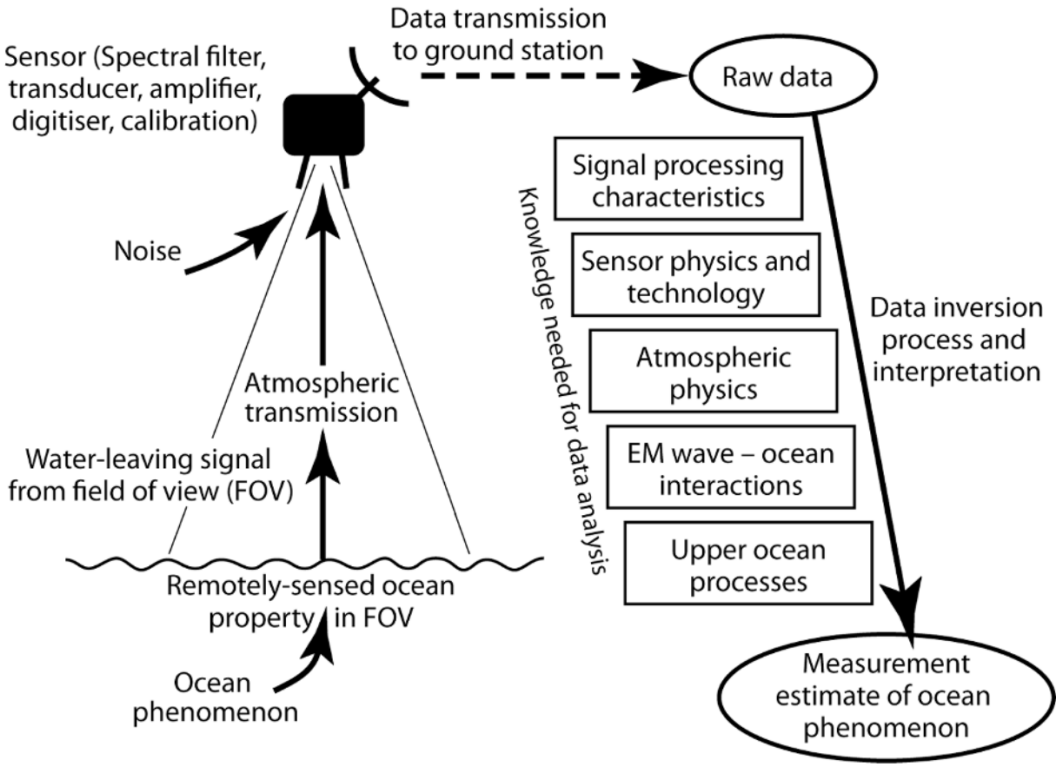


Figure 1.2: Schematic of a passive sensor and of information flow in satellite oceanography. Figure reproduced from [11]

a surface, it traverses the atmosphere, where it may undergo alterations and may also accumulate noise. Consequently, the ability of a sensor to accurately measure particular oceanic properties and its effectiveness in penetrating the atmosphere is heavily influenced by the segment of the electromagnetic spectrum it employs. As a matter of facts atmospheric transmittance varies with wavelength; for a substantial portion of the spectrum, the atmosphere is opaque, rendering it unsuitable for remote sensing. However, there are several windows in the spectrum where most radiation can penetrate, particularly in the visible, infrared, and microwave regions [11].

After traversing the atmosphere, the signal is detected by the sensor, which recognizes specific features of the radiation and translates each measurement into a digital signal for encoding and sending to the ground as illustrated in figure 1.2. Once the digital data have been received at the ground station the task remains to recover useful oceanographic information from a long string of numbers. Once the ground station receives the digital data, the objective is to extract meaningful oceanographic information from a lengthy series of numbers. This task requires reversing the information flow, where each phase must be inverted.

Remote sensors are also subject to classification. While various classification methods are available, the most widely adopted is based on the illumination source they utilize. Sensors that rely on natural solar energy are classified as passive sensors, while those that produce their own energy are termed active sensors. Predominantly, passive systems function within the visible, infrared, thermal infrared, and microwave segments of the electromagnetic spectrum, while most active sensors op-

erate within the microwave range, enabling them to penetrate the atmosphere under a variety of conditions [10]. Altimeters, whose data are the center of this work and which will be addressed in section 2.1.2, are among the most significant types of active sensors.
Observations, much like models, are inherently subject to uncertainty. One primary source of error is associated with sampling variability, which stems from the finite duration of measurements taken from an environmental process that is essentially continuous. This is both spatial and temporal, as the data collected may not fully capture the range of environmental conditions due to the specific time and space intervals chosen for observation [12].
For satellite observations a further source of error arises from the instruments used. Modern ocean sensors typically generate an electrical signal that corresponds to a specific ocean parameter. This signal is subsequently transformed into a numerical representation of that parameter. However, the values obtained from the signal are not solely determined by the ocean parameter itself; they are also influenced by various external factors. These influences include unavoidable instabilities and inhomogeneities present in the water surrounding the sensor during in-situ measurements, as well as atmospheric conditions affecting remote sensors. Additionally, the technical characteristics of the sensor, such as instrument noise and bias, further complicate the accuracy of the readings. As a result, even under optimal measurement conditions and in the absence of instrument malfunctions, it remains impossible to ascertain the extent to which the signal value diverges from the true value of the ocean parameter being measured. Therefore, the actual value of the sensor signal is characterized by uncertainty [13].
Furthermore, the uncertainty associated with these measurements is not constant, unlike that found in laboratory data. Indeed, field measurements differ fundamentally from laboratory assessments; in-situ and remote measurements are characterized by their unique spatial and temporal contexts and their transient nature. In the field of ocean sciences, where access to environmental data is frequently constrained by the availability of ships or adverse weather conditions, researchers often rely on single sensor data. This reliance, coupled with the absence of replication and limited insights into data quality, exacerbates the uncertainty surrounding the findings [13].

1.1.3 Data assimilation

The last essential element for the development of reanalyses is a data assimilation framework. Data assimilation is the process of integrating prior information from numerical model simulations and observed data in order to achieve the most accurate description of a dynamic system [14]. Since both the model and observations are incomplete and inaccurate, relying solely on one of them is not feasible. Thus we need to fit the model state to the observations in an optimal way [15].
Considering the errors in both the model and observations to be random, it is reasonable to conclude that they can be accurately described by probability density functions. This leads to the conclusion that the optimal framework for addressing the data assimilation challenge is founded on the Bayesian approach and Bayes theorem. According to this theorem the probability distribution, which serves as the most accurate depiction of the system's state, called the posterior, is a point wise multiplication of the probability distribution of our prior knowledge, derived

from the numerical model, with the probability distribution of the observations given each possible state of the model [16]. Consider the sequences of system states and observations within the interval $[t_0, t_K]$ as $\mathbf{x}_{0:K} = \{\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_K\}$ and $\mathbf{y}_{1:K} = \{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_K\}$ respectively. Then Bayes's theorem reads as

$$p(\mathbf{x}_{0:K}|\mathbf{y}_{1:K}) = \frac{p(\mathbf{y}_{1:K}|\mathbf{x}_{0:K})p(\mathbf{x}_{0:K})}{p(\mathbf{y}_{1:K})} \tag{1.1}$$

where:

- $p(\mathbf{x}_{0:K}|\mathbf{y}_{1:K})$ is the conditional probability (also known as *posterior* distribution) that the model states $\mathbf{x}_{0:K}$ are true given that the observations $\mathbf{y}_{1:K}$ are true;
- $p(\mathbf{x}_{0:K})$ denotes the probability that the model is true, independent of any observations. It is also known as *prior* knowledge, since it is the information provided by the model before incorporating new observations;
- $p(\mathbf{y}_{1:K}|\mathbf{x}_{0:K})$ is the conditional probability (also known as the *likelihood*) of the observations being true under the assumption that $\mathbf{x}_{0:K}$ corresponds to the true state;
- $p(\mathbf{y}_{1:K})$ signifies the probability of a series of observations being true, independent of the state of the model. It is also called *marginal* distribution and, since it does not depend on the model, can be considered a normalization factor.

Upon deriving a probability distribution function from this process, the next step is to determine the optimal estimator for the system's state. In accordance with [17], two primary options, namely the *mean* and *mode* of the distribution, represent two conventional approaches to state estimation. Estimators focused on the mean are referred to as minimum squared error estimators, as the mean always serves as the minimum squared error estimate. Conversely, estimators targeting the mode, which signifies the peak of the distribution and the most probable state, are known as maximum a posteriori estimators. In the former case, data assimilation involves minimization of the variability or uncertainty of the model error, while in the latter case, it entails minimizing a suitable cost function. A classic example of the minimum variance method is the Kalman filter, whereas variational methods are a common example of maximum a posteriori methods.

1.1.4 Importance of accurate reanalyses

Reanalyses plays a crucial role in the investigation of climate and oceanic systems. A fundamental aspect of reanalyses is its ability to bridge the gaps in observational data across both space and time, resulting in the creation of "maps without gaps." The reliability of these estimates is influenced by the robustness of the model employed, the presence of observations for additional variables and the quality of the data assimilation scheme used. Consequently, reanalyses can effectively fill in historical climate maps by evaluating the most likely conditions, even in areas where observational data for the variable of interest is limited [18].
Reanalyses serves as a vital tool in the development and evaluation of policies at various levels, including global, European, and national, particularly in the context

of climate change adaptation and the assessment of climate risks and resilience. As a matter of facts contemporary reanalyses serve as a valuable source of evidence for elucidating recent temperature trends, investigating fine-scale extremes across both spatial and temporal dimensions, and analyzing climatic changes over the past century to millennium. Furthermore, reanalyses can enhance the post-processing of climate model outputs and inform impact models [19].

An additional domain for the application of reanalyses is in the formulation of forecasting models that utilize machine learning techniques. These data-driven models possess the capability to autonomously discern spatial-temporal relationships from extensive oceanographic datasets, thereby effectively identifying the dynamics of oceanic changes without the necessity of prior physical knowledge. The development of efficient data-driven ocean forecasting models continues to pose a significant research challenge. Unlike the atmospheric domain, where various elements are interconnected, the ocean is segmented into several relatively autonomous regions characterized by distinct water mass properties, largely due to the presence of continents and islands. This segmentation complicates the automatic learning of internal change patterns and the interrelations among different oceanic regions [20]. Consequently, the availability of a high-quality training dataset is imperative. In this context, reanalyses data serve as a valuable resource for training machine learning algorithms, as they provide the most accurate representation of the ocean's state.

The success of these applications relies heavily on the quality of reanalyses. Therefore, it is essential to engage in ongoing development and enhancement of these reanalyses, a goal that this study tries to achieve.

1.2 Area of interest: Mediterranean Sea

Situated in the midlatitudes, the Mediterranean Sea is a semi-enclosed sea that stretches from 5°E to 36°W and from 32°N to 46°N, featuring an average depth of around 1500 meters. This sea is interconnected with the Atlantic Ocean through the Strait of Gibraltar, which allows for the interchange of waters, and it also connects to the Black Sea via the Dardanelles and the Bosphorus Strait.

The Mediterranean Sea serves as an invaluable laboratory basin for the study of general circulation patterns. It is noteworthy that numerous processes integral to the global ocean's circulation are replicated within the Mediterranean, either in the same manner or through analogous mechanisms. However the smaller dimensions of the Mediterranean facilitate logistical aspects of research, resulting in a reduced need for resources [21]. What renders the Mediterranean Sea particularly advantageous is its function as a miniature ocean, characterized by significantly shorter temporal and spatial scales compared to the global ocean. Furthermore, the relatively high salinity of the Mediterranean outflow at the Gibraltar Strait is known to influence global circulation patterns. Thus, even with its semi-enclosed characteristics, shifts in the Mediterranean Sea's circulation can have a direct impact on the global ocean [22].

1.2.1 Circulation and thermohaline structure

The dynamics of Mediterranean circulation and its thermohaline structure are governed by a range of forcings. Among these, the exchanges occurring at the Strait of Gibraltar and the interactions between the atmosphere and the sea, which involve momentum, heat, and freshwater, are paramount. Additionally, contributions from river inflows and exchanges with the Black Sea also play a significant role in shaping these dynamics [21, 23].
The circulation of the Mediterranean Sea, influenced by these various forcings, exhibits an antiestuarine characteristic, indicating a net deficit of freshwater due to evaporation surpassing the combined effects of precipitation and river runoff. To counterbalance this deficit and sustain the overall water volume, a compensatory flow is necessary, which occurs via the Strait of Gibraltar. This compensatory flow consists of a surface inflow that is part of a bidirectional flow system. The second component involves the outflow of denser Mediterranean water, which moves toward the Atlantic Ocean as a deep undercurrent. The dynamics of these two flows are directly related to the density differences between Mediterranean and Atlantic waters [23].
The thermohaline properties commonly found in the Mediterranean Sea's water column are significantly influenced by driving forces, too. As a matter of facts the ocean's vertical structure, or stratification, emerges from the mixing of various water masses that have been advected to a specific location as a result of these forces. A water mass is defined as "a body of water with common formation history, having its origin in a particular region of the ocean" [24]. During its formation, a water mass develops distinct characteristics of temperature and salinity. Notably, temperature and salinity are considered conservative properties, subject to change only at the sea surface through air-sea interactions. Consequently, as long as a water mass remains isolated from the surface, these properties will undergo slow gradual changes due to mixing with neighboring water masses, yet they will largely retain their original characteristics, allowing for the identification of the water mass.
To accurately describe the thermohaline structure of the Mediterranean Sea, it is essential to investigate the various water masses present. This analysis will follow the framework provided in *Oceanography of the Mediterranean Sea* [23]. At mid-latitudes, the vertical structure of the water column can be segmented into three main layers: the mixed layer, which is located near the surface (100-150 meters) and is characterized by homogeneously mixed water properties; the thermocline, where there is a significant and rapid decline in temperature with increasing depth; and the abyssal layer. The Mediterranean's three principal water masses are distributed across these layers, with *Atlantic Water* occupying both the mixed layer and the thermocline, while *Intermediate Water* and *Deep Water* are found at different depths within the abyssal layer, as shown in figure 1.3.
Atlantic Water (AW) refers to the water mass entering from the Strait of Gibraltar. As it flows, AW disperses throughout the basin, contributing to the formation of gyres, mesoscale eddies, and various currents. Throughout its transit across the basin, AW undergoes continuous alterations due to interactions with the atmosphere and its mixing with other water masses. This process results in a gradual increase in temperature and salinity along a west-east gradient; however, the core of AW can still be identified throughout the Mediterranean Sea as a near-surface salinity minimum. Generally, the temperature and salinity ranges of AW are between 13

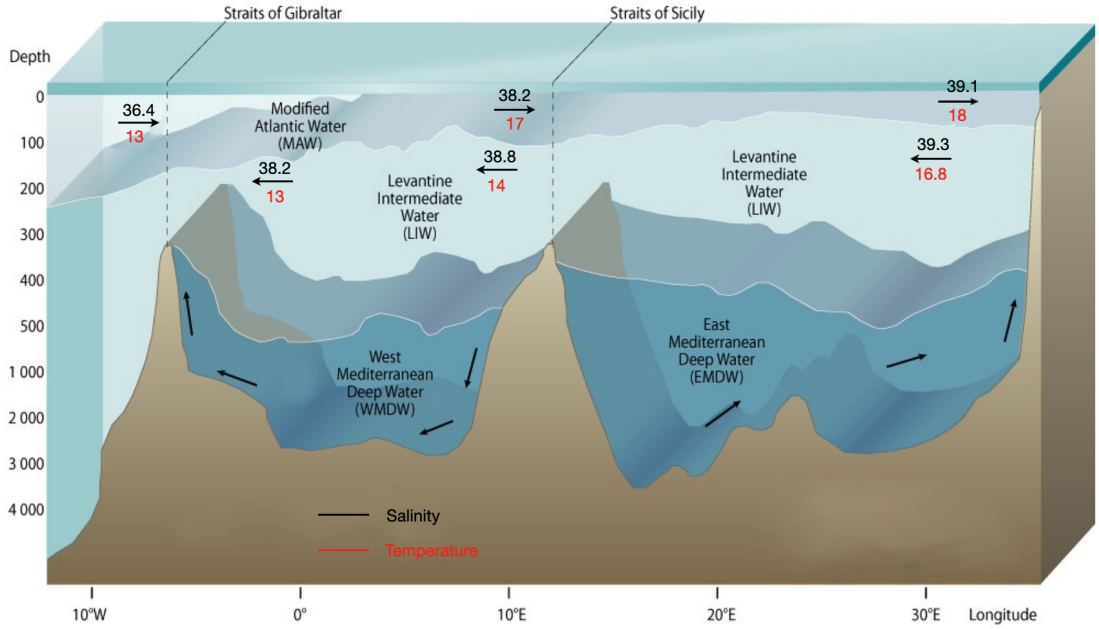


Figure 1.3: Vertical distribution of water masses in Mediterranean Sea. Redrawn from [25]

to 17°C and 36.4 to 38.2 psu in the western Mediterranean, while in the eastern Mediterranean, these ranges are from 15 to 18°C and 38.5 to 39.1 psu.

The defining feature of *Intermediate Water* (IW) is the presence of a subsurface salinity maximum typically found at depths ranging from 250 to 600 meters, which is often accompanied by a relative maximum in temperature. This water mass is primarily generated in the Levantine and Cretan seas, regions characterized by elevated salinity and temperature levels. As IW migrates westward, it gradually loses its extremes, resulting in a decrease in both temperature and salinity. Consequently, a gradient is observed along the east-west axis, with temperature and salinity values in the Eastern Mediterranean ranging from 13.8 to 16.8°C and 38.8 to 39.3 psu, respectively, while the Western Mediterranean exhibits values between 13 to 14°C and 38.2 to 38.7 psu. IW constitutes the majority of the Mediterranean outflow through the Strait of Gibraltar, which, due to its limited depth, restricts the passage of deeper water masses.

Deep Water (DW) represents the most dense water mass, which exhibits a slight decline in temperature and salinity as depth increases. This water mass is formed at various sites along the northern coastlines, primarily due to the influence of cold and dry winter winds. The specific properties of Deep Water can exhibit slight variations depending on the formation site. For instance, the Western Mediterranean Deep Water is produced in the Gulf of Lion, with an average temperature of 12.8°C and an average salinity of 38.45 psu [26]. In contrast, the Eastern Mediterranean Deep Water is categorized into two distinct types: Adriatic Deep Water, which has a temperature of 13°C and a salinity of 38.6 psu [27], and Aegean Deep Water, with a temperature of 14.25°C and a salinity of 39.1 psu [28]. It is important to note that these values are general estimates and may fluctuate over time. Furthermore, the bathymetric constraints of the Sicily Channel, which separates the Western and Eastern Mediterranean Seas, ensure that these two deep water masses remain

distinct from one another.

1.2.2 Temporal and spatial scales

The overall circulation in the Mediterranean Sea is characterized by a complex structure, characterized by three primary and interrelated spatial scales: the basin scale, the sub-basin scale, and the mesoscale. This complexity, along with the various scales, is a result of numerous driving forces, significant topographic and coastal effects, and intrinsic dynamical processes [29].
The circulation at the basin scale encompasses the dynamics of the various water masses and the associated overturning circulation. This extensive circulation system spans the entire basin and is characterized by both zonal and meridional components. The zonal circulation consists of three distinct structures. The first structure is a shallow clockwise cell, which corresponds to the eastward flow of AW at the surface, while IW, particularly that formed in the Levantine basin, flows westward at greater depths. Beneath this layer, two additional cells are identified: one located in the western Mediterranean and the other in the eastern region. The western cell is a straightforward counterclockwise circulation, whereas the eastern cell exhibits a more complex, multi-centered clockwise circulation. The meridional circulation varies depending on the specific area of the Mediterranean under consideration. In the western Mediterranean, a clockwise cell with several maxima exists from approximately 100 meters to 2000 meters depth, alongside a counterclockwise cell in the deeper regions of the Algerian basin. In contrast, the eastern Mediterranean features a counterclockwise surface cell within the upper 250 meters, while a clockwise multi-centered overturning cell extends from 250 meters to 1000 meters. Below 1000 meters, a deep counterclockwise Eulerian cell is present in the meridional region of the Ionian Sea [30].
At the sub-basin scale, one can identify various structures, such as free and boundary currents, as well as jets that undergo bifurcation, in addition to both permanent and transient cyclonic and anticyclonic gyres. A notable example is found in the Alboran Sea, where the AW flows in an anti-cyclonic manner in the western part of the basin and cyclonically in the eastern part. The Algerian Current serves as a significant current in this area, generating meanders, commonly termed 'coastal eddies,' due to its unstable nature. The eastern basin is characterized by energetic sub-basin scale features, including jets and gyres, which exhibit substantial variability in their shape, location, and strength. This variability is particularly pronounced in the permanent sub-basin gyres, as well as in the occurrence of transient and aperiodic eddies, jets, and filaments [29]. The classification of eddies in both basins as mesoscale phenomena is dependent on their dimensions and the temporal scale of their existence.
Small yet energetic mesoscale eddies are particularly prevalent in the Mediterranean Sea. In the western basin, significant mesoscale activities manifest as instabilities along coastal currents, resulting in the generation of mesoscale eddies that may either traverse the basin or interact with the prevailing currents. For instance, along the Algerian Current, both cyclonic and anticyclonic eddies form and develop over several months, gradually drifting eastward. Typically, the anticyclonic eddies increase in size and detach from the coastline. In the Levantine basin, one can also observe energetic mesoscale eddies in the open ocean [29].

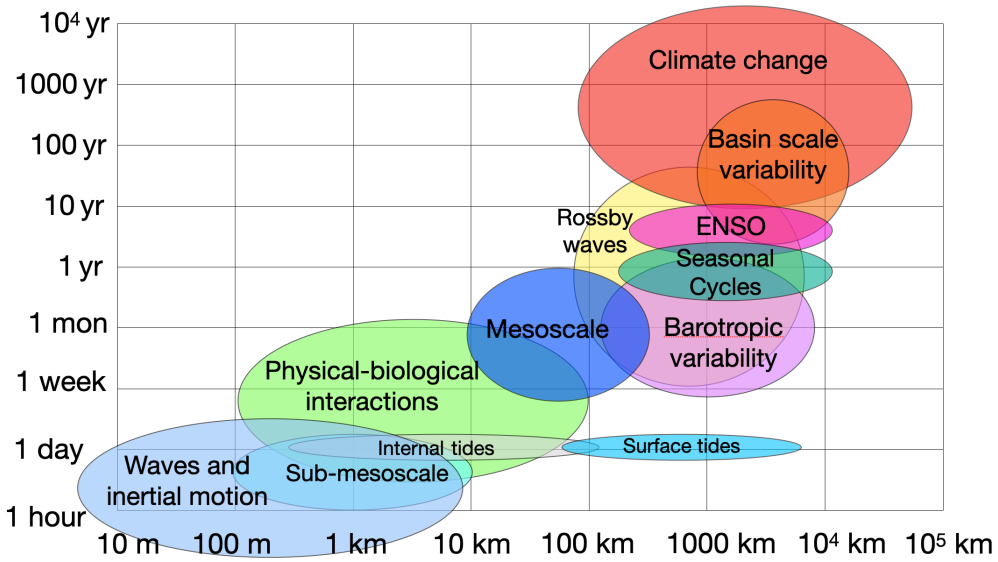


Figure 1.4: Time and space scales of physical oceanographic phenomena. Figure redrawn from [31]

It is a well-established principle that usually phenomena characterized by short spatial scales are generally associated with short timescales, whereas phenomena with long spatial scales are linked to longer timescales [32]. This principle suggests that the extensive range of spatial scales corresponds to a similarly extensive range of timescales, as depicted in figure 1.4. A clear illustration of this can be found in the sea surface temperature of the Mediterranean Sea, which is governed by heat fluxes that can exhibit considerable daily variability, especially in the presence of strong winds and cold, dry air masses [23].The primary scales of variability for sea surface temperature include seasonal variability, high-frequency variability, and interannual variability, with seasonal variability being the most dominant, although high-frequency variability also plays a significant role [33]. Consequently, it is crucial to account for various temporal scales when analyzing the thermohaline properties of the Mediterranean.

1.3 Mesoscale and eddies

1.3.1 Equations of motion

In accordance with the discussion in section 1.1.1, the mathematical representation of oceanic physical and dynamical features is achieved through the Navier-Stokes equations or some approximation of them. By implementing a series of approximations to the Navier-Stokes equations, one can derive the primitive equations [34, 35]. The necessary approximations include the hydrostatic approximation which asserts that the pressure at any point in the ocean is determined by the weight of the fluid above; the Boussinesq approximation, which indicates that density variations are relatively small compared to the mean density and are relevant only in the buoyancy term; the incompressible approximation, which states that the density of incompressible fluids remains constant along the flow over time; and the Reynolds averages approximation, which allows for the separation of each variable

into a mean and a fluctuating component. The primitive equations derived from these approximations are

$$\frac{\partial u}{\partial t} + \overrightarrow{u} \cdot \nabla u - 2\Omega \sin\theta v = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} + \nu\nabla^2 u + \nabla_H \cdot (A_H \nabla_H u) + \frac{\partial}{\partial z}\left(A_v \frac{\partial u}{\partial z}\right) \quad (1.2)$$

$$\frac{\partial v}{\partial t} + \overrightarrow{u} \cdot \nabla v + 2\Omega \sin\theta u = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} + \nu\nabla^2 v + \nabla_H \cdot (A_H \nabla_H v) + \frac{\partial}{\partial z}\left(A_v \frac{\partial v}{\partial z}\right) \quad (1.3)$$

$$\frac{1}{\rho}\frac{\partial p}{\partial z} = -g \quad (1.4)$$

$$\nabla \cdot \overrightarrow{u} = 0 \quad (1.5)$$

where ∇^2 is the three-dimensional Laplacian operator, Ω is the angular velocity of the Earth, θ represents latitude, $\overrightarrow{u} = (u, v, w)$ is the three-dimensional velocity vector, p signifies pressure, g denotes gravitational acceleration, ρ indicates density, ν refers to kinematic viscosity and A_H and A_v are called the horizontal and vertical turbulent viscosity coefficients respectively.
The execution of a spatial and temporal scale analysis enables the distinction of motion within various dynamical regimes. This is accomplished through the establishment of the Rossby number, expressed as:

$$\epsilon = \frac{U}{Lf} \quad (1.6)$$

In this equation, U denotes the magnitude of the horizontal velocity field, L signifies the horizontal scale of the phenomenon, and $f = 2\Omega \sin\theta$ represents the Coriolis parameter, which approximates to $\sim 10^{-4}$ at mid-latitudes. The mesoscale regime is characterized by a Rossby number value of less than one (< 1), permitting the application of a perturbative approach to expand the variables u, v, and p into a series of components, each multiplied by ascending powers of ϵ. This process yields a new set of equations, each corresponding to a distinct dynamical mode linked to a specific power of ϵ. The zero-order equations reduce to the *geostrophic balance*, while the first-order equations correspond to the *quasi-geostrophic balance*.

1.3.2 Eddy dynamics

The phenomenon of ocean mesoscale eddies can be described as the ocean's equivalent of weather systems, with horizontal dimensions of under 100 kilometers and timescales that approximate one month [36]. These eddies typically possess distinct physical and chemical properties, such as variations in temperature, salinity, and carbon concentration, compared to their adjacent waters, enabling the transport of these characteristics as they navigate through the ocean [37].
The circulation characteristics of eddies can be derived from the principles of geostrophic balance. The geostrophic equations are

$$f v_g = \frac{1}{\rho_0}\frac{\partial p}{\partial x} \quad (1.7)$$

$$f u_g = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} \quad (1.8)$$

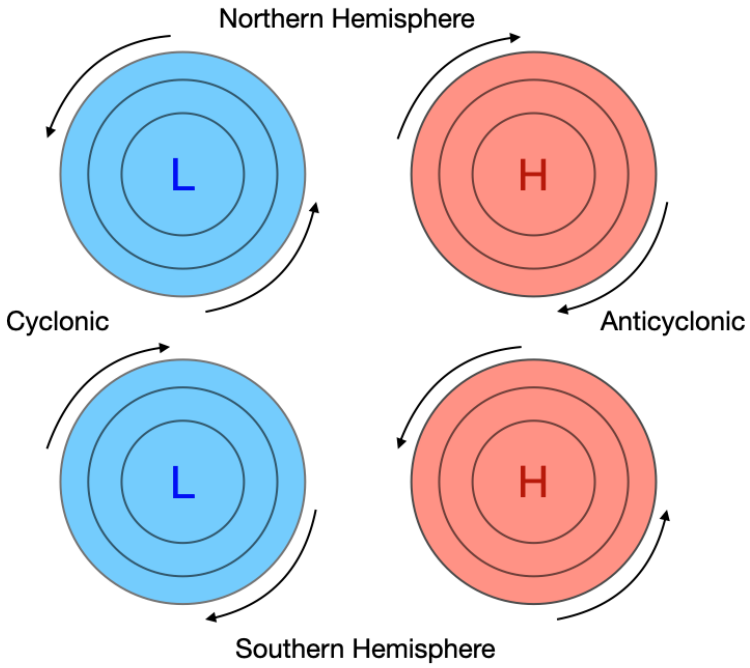


Figure 1.5: Diagram of eddy sense of rotation depending on the hemisphere. L stands for low sea level center, while H stands for high sea level center. Figure redrawn from [34]

They illustrate a balance between the Coriolis force on the left and the horizontal pressure gradient force on the right. The velocity field associated with geostrophic flow is influenced by the pressure gradient, with its orientation being perpendicular to the isolines of pressure. Specifically, currents tend to circulate in a nearly circular pattern around areas of elevated or reduced pressure (or sea level), thereby forming an eddy. The direction of rotation is determined by the sign of f, which varies according to the selected hemisphere of the Earth, as illustrated in figure 1.5.
Conducting a scale analysis of the equations 1.8 and 1.7 yields the following relationships

$$fU = \frac{1}{\rho_0}\frac{\Delta p}{L} \tag{1.9}$$

Through the examination of scale analysis for equation 1.4, we can also establish the following equation:

$$\frac{1}{\rho_0}\frac{\Delta p}{H} = g$$

From this, we can derive that $\frac{\Delta p}{\rho_0} = gH$. Substituting this finding into 1.9 yields:

$$fU = \frac{gH}{L} \tag{1.10}$$

which can be rearranged to yield $U = \frac{gH}{fL}$. By substituting this expression into 1.6, we derive:

$$\epsilon = \frac{gH}{f^2L^2} = \left(\frac{\lambda}{L}\right) < 1 \tag{1.11}$$

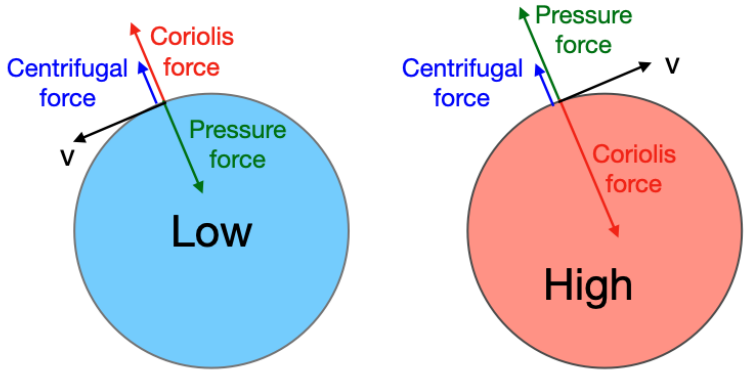


Figure 1.6: Balances between pressure gradient force, Coriolis force and centrifugal force in Northern Hemisphere eddies. Figure redrawn from [38]

where $\lambda = \frac{\sqrt{gH}}{f}$ represents the internal deformation radius. Consequently, a small Rossby number, which is essential for the formation of such eddies, arises from the condition that the horizontal scale is significantly larger than the deformation radius [38].
When the horizontal scales approximate the deformation radius, resulting in a Rossby number near unity, the centrifugal force becomes a significant factor in the force balance. In the vicinity of a low sea level center, this force acts outward, partially countering the inward pressure gradient force. Consequently, the Coriolis force must decrease, and since the Coriolis parameter f a constant, the rotational velocity must also diminish, as illustrated in figure 1.6. In contrast, around a high sea level center, the pressure gradient force acts outward, similar to the centrifugal force, necessitating that the Coriolis force counterbalance both forces. This results in an increase in magnitude and, consequently, an increase in rotational velocity [38]. Such dynamics are characteristic of smaller eddies.
The deformation radius, which is characteristic of mid-latitude regions in the global ocean, typically ranges from 30 to 50 kilometers. Conversely, the Mediterranean Sea presents a considerably smaller deformation radius, which varies from 5 to 12 kilometers depending on the season and specific locations within the sea [31].

1.3.3 Eddy properties and vertical structure

Eddy formations are observed around centers of both low and high sea levels. Specifically, centers of low sea level are associated with low sea surface pressure and cyclonic circulation, whereas high sea level centers correspond to anticyclonic circulation. In this context, cyclonic circulation refers to a rotation that aligns with the Earth's spin — clockwise in the Southern Hemisphere and counterclockwise in the Northern Hemisphere — while anticyclonic circulation exhibits the opposite rotation. Consequently he identification of local extremes in sea level anomalies enables the detection and monitoring of both cyclonic and anticyclonic eddies [39].
The vertical characteristics of these eddies are intrinsically related to the effects produced by these extremes within the water column. To analyze these properties, we consider the hydrostatic equation 1.4. As stated before this equation asserts

that the pressure at any depth in the ocean is determined by the force per unit area resulting from the weight of the fluid above that depth, and integrating it from the ocean's bottom to its surface η we get

$$\eta\rho_0 g + \int_{-H}^{0} \rho(z) g dz = p(-H) \tag{1.12}$$

In this context, η represents the sea level relative to a designated reference point at 0, while H denotes the ocean's depth. The term $\rho(z)$ signifies the density of the water column at a given depth, $p(-H)$ indicates the pressure exerted at the ocean's bottom [40]. In this context we are not considering atmospheric pressure, which is usually assumed as a constant. When satellite observations indicate a variation in sea level, denoted as $\Delta\eta$, and considering the stipulation that the pressure at the ocean floor remains unchanged, i.e., $\Delta p(-H) = 0$, we derive the following equation:

$$\Delta\eta\rho_0 g + \int_{-H}^{0} \Delta\rho(z) g dz = 0 \tag{1.13}$$

This implies that any alteration in the weight of the water column must balance the observed change in surface level as recorded by the altimeter.
This balance requires an adiabatic and frictionless adjustment of water parcels in space [41], ensuring that their temperature and salinity characteristics remain unchanged. The most straightforward form of this reconfiguration involves a vertical displacement of the water. Each water parcel within a specific water column is displaced vertically by an identical amount, denoted as Δh. This uniform displacement applies to all isopycnals, thereby conserving the quantity of water with specific properties within the column. However, vertical displacement alone cannot maintain the conservation of water properties across all isopycnals. IIn cases of low sea surface pressure, the water column is elevated by a distance Δh, , which involves the introduction of denser water from the bottom and the removal of lighter surface water. This process raises colder, saltier water to the surface, resulting in the uplift of both the thermocline and the halocline. Conversely, when a center of high sea surface pressure is present, the opposite effect occurs [42]. This means that cyclonic eddies generate negative temperature anomalies and/or positive salinity anomalies, leading to positive density anomalies, while anticyclonic eddies produce the reverse effects. Consequently, the thermocline exhibits a slope that is opposite to that of the sea surface [43], and anticyclonic eddies manifest as depressions in surfaces of constant temperature, whereas cyclonic eddies appear as "domes" of deep water that rise toward the surface. The sea surface above experiences an upward deflection that contrasts with the downward deflection of the underlying density layers [44], as shown in figure 1.7.
Anomalies associated with eddies are not restricted to their immediate area; they can extend horizontally beyond 1.5 times the radius of the eddy. Furthermore, it has been observed that anticyclones typically reach greater depths than cyclones on average. The vertical structure of these eddies also varies; anticyclones are characterized by a smaller radius at depth compared to their surface radius, while cyclones tend to exhibit a more cylindrical vertical profile [42].
Throughout the lifespan of an eddy, its properties may undergo significant changes: at the onset, both the surface radii and amplitudes are minimal, followed by a phase where they stabilize at their peak size, and finally, a rapid decrease in radius occurs just before the eddy collapses [31, 42].

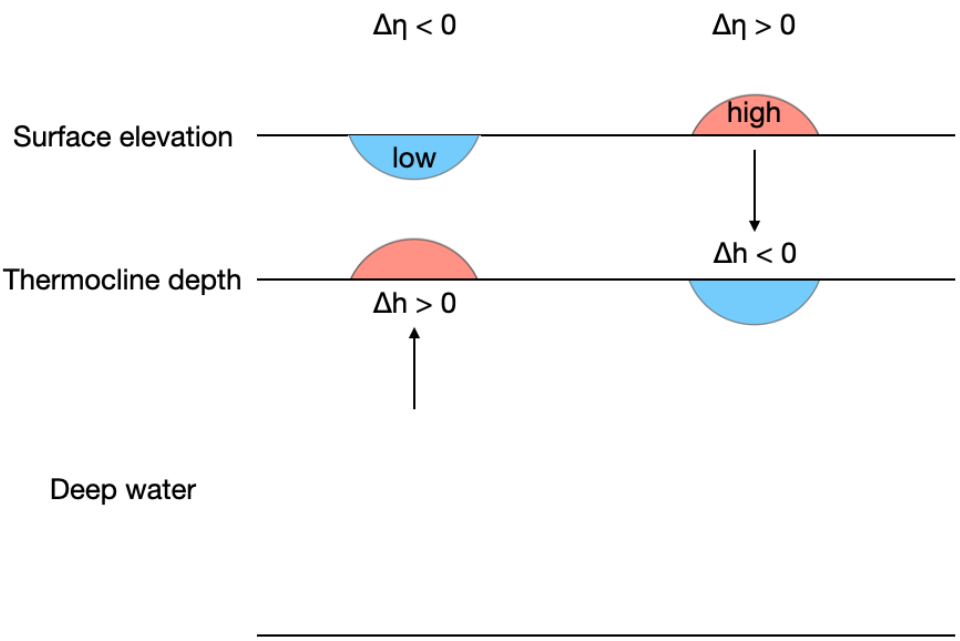


Figure 1.7: Schematic of the vertical displacements that correlates the surface elevation with the thermocline displacements. Figure redrawn from [40]

1.4 Thesis objectives

This study aims to explore the interrelationship between sea level anomaly, temperature, and salinity, with the goal of identifying the functions that most effectively correlate these three variables. This analysis is intended as a preliminary step to enhance reanalysis data, as the assimilation framework for one variable (sea level) requires the extraction of relevant information regarding the other two variables.
To achieve this correlation, a method originally introduced by Mellor and Ezer ([45], [46]) is employed. This approach utilizes correlation factors to connect the three variables. These correlation factors are time-independent but are derived from time-averaged data, and they may vary on a monthly or daily basis. Consequently, selecting an optimal time averaging interval and an appropriate number of correlation factors is crucial. The determination of this optimal time interval constitutes the first aim of this study.
A new method for the evaluation of correlation factors is subsequently introduced, drawing inspiration from the work of Ezer and Mellor. This method incorporates the application of Empirical Orthogonal Functions, utilized in both bivariate and trivariate forms as introduced by Adani et al. [47]. A comparative analysis of these methodologies is essential to determine the efficacy of the various correlation factors and to ascertain which correlation factors yield superior performance. To achieve this, we initially utilize the different methods to extrapolate temperature and salinity profiles based on sea surface anomalies derived from reanalysis data, thereby evaluating the quality of the proposed method. Additionally, to gain insights into the practical application of correlation factors, we extend our analysis to include

sea level anomalies obtained from satellite observations, similarly to what done by Guinehut et al. [48] .
Both applications — utilizing sea level anomalies from reanalysis and satellite data — are employed to reconstruct vertical profiles at four specific locations within the Mediterranean, as well as to analyze the vertical structure of mesoscale eddies. This latter analysis is particularly significant, given that mesoscale eddies exhibit unique vertical characteristics and play a critical role in oceanic circulation. Thus, evaluating the effectiveness of the various methods in reconstructing these vertical features is of critical importance.
The datasets utilized in this study are detailed in chapter 2, which provides an overview of satellite altimetry and Mediterranean reanalyses data. Chapter 3 outlines the methodology proposed by Mellor and Ezer, along with its enhancement through the application of bivariate and trivariate empirical orthogonal functions (EOFs). Additionally, this chapter outlines the preprocessing procedures applied to the data. In chapter 4, the efficacy of various time intervals and methodologies is evaluated using sea surface anomalies derived from reanalysis. Lastly, chapter 5 explores the implementation of the diverse techniques utilizing satellite data, while the conclusions drawn from the research are summarized in chapter 6.

Chapter 2

Satellite and reanalyses data

The primary aim of this research is to examine different techniques for reconstructing temperature and salinity profiles using satellite altimetry data. As such, a thorough representation of this data is necessary, which is addressed in this chapter. In section 2.1, an overview of satellite altimetry and sea level is presented, with particular attention given to the data used. Subsequently, section 2.2 describes the reanalyses data that is utilized for evaluating correlation factors and serves as a validation set. This section also describes the model and the assimilation scheme that are employed in constructing this specific type of reanalyses, along with a description of the dataset used.

2.1 Sea level and satellite Altimetry

Sea surface height refers to the elevation of the ocean's surface relative to a reference ellipsoid. Fluctuations in this height arise from changes in the mass of the underlying land, such as differences in rock density or mantle thickness, as well as variations in the mass of the water itself. The former changes occur at a slow pace and can be regarded as static, while the latter changes are more rapid and constitute the dynamic aspect of sea surface height. Consequently, measurements of sea level capture this dynamic aspect. They hold significant importance owing to a distinctive characteristic they possess. While the majority of ocean variables captured by satellite are limited to the surface and do not directly correlate with subsurface conditions, altimetric measurements of sea surface height represent a notable exception. Specifically, the variations in sea surface height above the ocean floor are directly influenced by changes in the mass or density of the water column [49].

2.1.1 Sea level and subsurface properties

To investigate tendencies in sea level, we can begin with the hydrostatic balance as done in [50] described by the equation

$$dp = -g\rho dz \tag{2.1}$$

where p represents pressure, g is gravitational acceleration and ρ is the density of the water. By integrating this equation from the ocean floor at $z = -H(x, y)$ to the water surface at $z = -\eta(x, y, t)$, similarly to what we have done in section 1.3.3, the

bottom pressure equation can be derived:

$$p_b = p_a + g \int_{-H}^{\eta} \rho dz \tag{2.2}$$

Here, p_b refers to the pressure at the ocean bottom, while p_a indicates the atmospheric pressure exerted at the sea surface, which is this time included. This relationship illustrates that the bottom pressure is the cumulative effect of the surface pressure and the weight of the seawater column above the ocean floor. By differentiating equation 2.2 with respect to time, we derive the following expression:

$$\frac{\partial(p_b - p_a)}{\partial t} = g\rho(\eta)\frac{\partial \eta}{\partial t} + g \int_{-H}^{\eta} \frac{\partial \rho}{\partial t} dz \tag{2.3}$$

In this equation, $\rho(\eta) = \rho(z = \eta)$ signifies the density at the free surface of the ocean [51].Rearranging this equation as done in [52] yields a diagnostic expression that characterizes the tendency of sea level:

$$\frac{\partial \eta}{\partial t} = \frac{1}{g\rho(\eta)} \frac{\partial(p_b - p_a)}{\partial t} - \frac{1}{\rho(\eta)} \int_{-H}^{\eta} \frac{\partial \rho}{\partial t} dz \tag{2.4}$$

The first component, $\frac{1}{g\rho(\eta)} \frac{\partial(p_b - p_a)}{\partial t}$ reflects the variations in ocean mass. Specifically, by integrating 2.2, we derive the expression

$$\frac{\partial(p_b - p_a)}{\partial t} = g \frac{\partial}{\partial t} \Big(\int_{-H}^{\eta} \rho dz \Big)$$

This equation delineates the mass budget per horizontal area within a column of seawater. It illustrates that the mass of the column per unit horizontal area is influenced by the convergence of mass transported laterally by ocean currents, as well as by mass exchanges occurring at the ocean's surface, including evaporation, precipitation, sea ice dynamics, and river discharge. As a matter of facts

$$\frac{\partial}{\partial t} \Big(\int_{-H}^{\eta} \rho dz \Big) = -\nabla U + Q_w \tag{2.5}$$

where $U = \int_{-H}^{\eta} \rho u dz$ is the horizontally transported mass and Q_w is the mass crossing the free surface. The second element of 2.4, $\frac{1}{\rho(\eta)} \int_{-H}^{\eta} \frac{\partial \rho}{\partial t} dz$, represents changes in ocean density.
An alternative method to convey the behavior of sea level is through the continuity equation, as done in [53]. By assuming that the density can be expressed as $\rho = \rho_0 + \rho'$ the principle of mass conservation is formulated as follows

$$\nabla \cdot \mathbf{u} = -\frac{1}{\rho(\eta)} \frac{D\rho'}{Dt} \tag{2.6}$$

where $\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla$ and $\mathbf{u} = (u, v, w)$ is the three dimensional velocity. To derive the sea level equation, it is necessary to perform a vertical integration of this equation between the sea level $h(x, y, t)$ and the bathymetric depth $-H(x, y)$ while applying the kinematic and dynamic boundary conditions at the two vertical limits of the fluid

$$w_{z=\eta} = \frac{D\eta}{Dt}\Big|_{z=\eta} + Q_w$$

$$w_{z=-H} = -\mathbf{u}|_{z=-H} \cdot \nabla H$$

The integration leads us to the following expression

$$\frac{\partial \eta}{\partial t} = -\nabla[(H+\eta)\overline{\mathbf{u}}] - Q_w - \frac{1}{\rho(\eta)} \int_{-H}^{\eta} \frac{D\rho'}{Dt} dz \tag{2.7}$$

where $\overline{\mathbf{u}} = \frac{1}{H+\eta} \int_{-H}^{\eta} \mathbf{u} dz$ represents the barotropic velocity field. This formulation clearly delineates the two components associated with mass tendency: the first term, which arises from the horizontal transport of mass through convergence or divergence ($\nabla[(H+\eta)\overline{\mathbf{u}}]$, and the second term, which accounts for the mass that crosses the ocean's free surface (Q_w). Together, these terms constitute the sea level tendency in incompressible models. The final term, which is influenced by density variations $\frac{1}{\rho(\eta)} \int_{-H}^{\eta} \frac{D\rho'}{Dt} dz$ Together, these terms constitute the sea level tendency in incompressible models. The final term, which is influenced by density variations.
Given that the sea level trend can be partitioned into two primary components, it follows that the overall sea level can be effectively reconstructed by aggregating the incompressible contributions with those from steric effects. This can be articulated in the following equation:

$$\eta = \eta_i + \eta_s \tag{2.8}$$

where η_i signifies the incompressible portion and η_s indicates the steric portion. As said variations in steric sea level are primarily driven by changes in the density of water, which is influenced by temperature and salinity. Consequently, also the overall steric effect can be decomposed into two distinct components: the thermosteric component and the halosteric component. This allows for the representation of the steric sea surface height anomaly as follows:

$$\eta_s = \eta_{sT} + \eta_{sS} = -\frac{1}{\rho(\eta)} \Big(\int \rho(T, S^*) dz + \int \rho(T^*, S) dz \Big) \tag{2.9}$$

where T^* and S^* represent the reference values of temperature and salinity [54]. Given the significant influence of temperature and salinity on the steric sea level, this relationship can be effectively utilized to extrapolate temperature and salinity profiles, as demonstrated in this study. Specifically, sea level anomalies serve as a basis for deriving temperature and salinity anomalies through the application of correlation factors. Subsequently, a climatological dataset is incorporated to accurately reconstruct the profiles. The correlation factors are derived by integrating sea level anomalies obtained from satellite observations or reanalyses with temperature and salinity anomalies sourced from reanalyses data. Detailed information regarding sea level altimetry data can be found in section 2.1.2, while the salinity and temperature data are discussed in section 2.2.

2.1.2 Satellite altimetry

A satellite altimeter is a type of microwave radar that generates an echo. This echo moves through the atmosphere and is reflected back to the radar by the surface below. The radar then calculates the time interval between the emission and reception of the echo, along with its return strength and shape, to determine the distance from the satellite to the sub-satellite point illuminated by the radar. The distance measurement is crucial for deriving the topography of the sea surface. This process

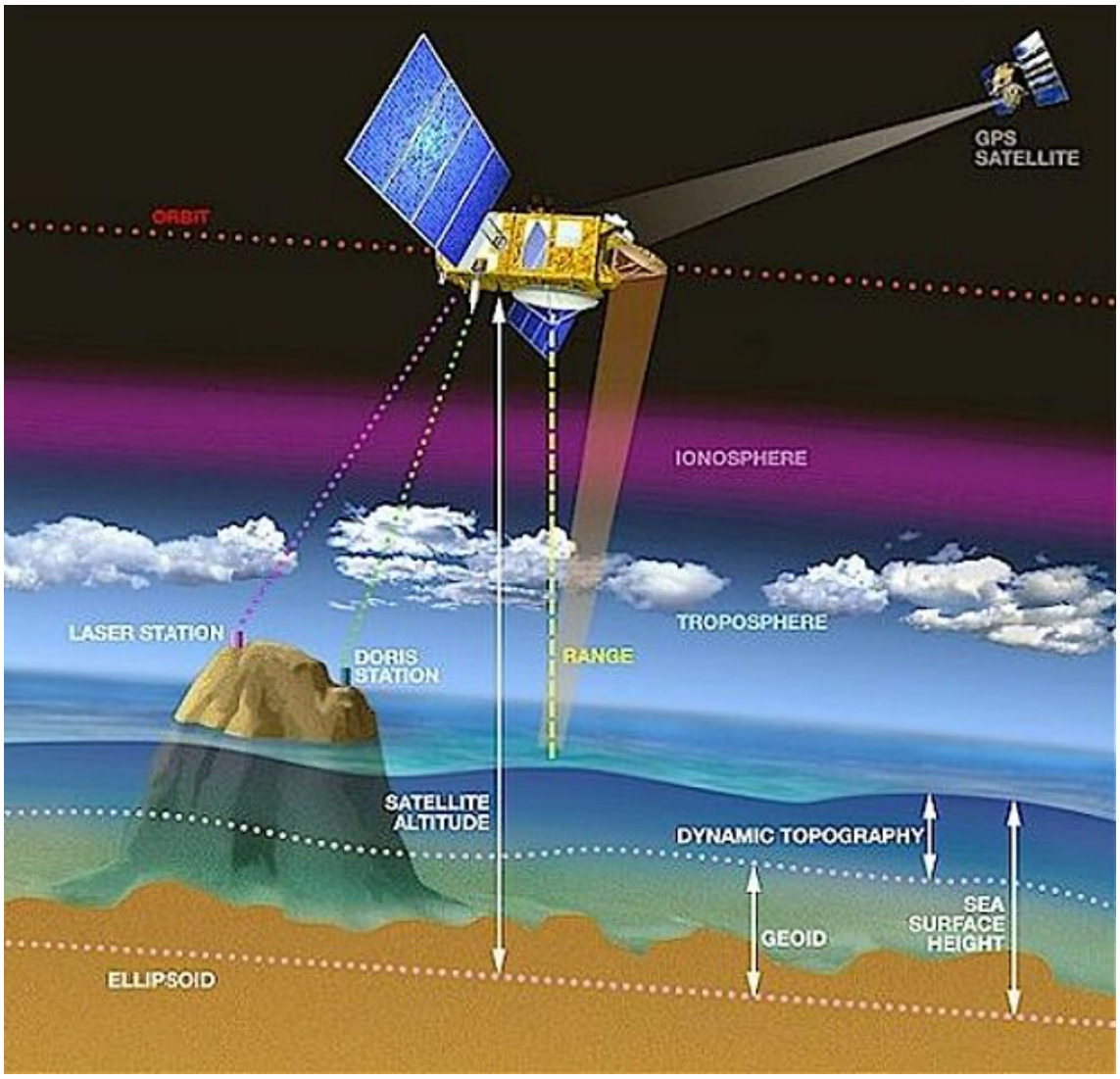


Figure 2.1: Components of the sea surface height as observed by satellite altimeter. Figure reproduced from [56]

also allows for the assessment of significant wave height and the modulus of sea surface wind speed [55].
The technique employed in radar altimetry encompasses two distinct geometric measurements: firstly, the distance from the satellite to the sea surface (range R), as previously mentioned, and secondly, the precise determination of the satellite's position (satellite altitude H_{sat}) relative to a fixed Earth coordinate system (ellipsoid) 2.1. The resultant difference between these two measurements yields a precise assessment of sea surface topography [57]:

$$\eta = H_{sat} - R \tag{2.10}$$

Similar to other remote sensing techniques, altimetry requires the inversion of the information flow to accurately interpret the data collected. A classification system is available for satellite altimetry products, based on the degree of processing they have undergone [11, 58]:

- *Level 0*: raw data received from satellite, in standard binary form;

- *Level 1*: data with sensor calibration, which converts it into an estimate of the electromagnetic properties that the sensor is designed to measure;

- *Level 2*: geolocated and geophysical calibrated data with atmospheric correction applied. Geolocation refers to the inclusion of geographical coordinates corresponding to the measurement location, although these may not be depicted on a map. The geophysical calibration process adjusts the data to estimate specific ocean variables, while atmospheric correction addresses the impacts of atmospheric absorption or attenuation;

- *Level 3*: mapped data derived from a single sensor. This data is presented in a geographically meaningful format, complete with latitude and longitude coordinates. In some instances, these images are composites, merging multiple datasets from the same sensor taken at various times and locations. The use of a single sensor can result in gaps due to the use of a single sensor;

- *Level 4*: mapped data on a comprehensive grid, free of gaps, achieved through the integration of other satellite sources or by filling gaps via interpolation.

2.1.3 Sea level anomaly from satellite

Sea level anomaly (*SLA*) refers to the deviation of the altimeter signal from a temporal average. Analyzing these anomalies enables researchers to enhance their comprehension of ocean circulation dynamics and refine predictions of climatic phenomena [59].
The dataset utilized in this study is the daily *SLA* Level 4 data product provided by the CLS (*Collecte Localisation Satellite*) group. This data is generated through optimal interpolation, which integrates Level 3 along-track measurements from various altimeter missions and is presented on a uniform grid with a resolution of 0.125°. The data is accessible via the *Copernicus Marine Service* (CMEMS) platform [60].
To derive the measurement of *SLA*, the initial step involves acquiring the value of η from satellite observations 2.1. This measurement can be understood as comprising several components:

$$\eta = \eta_g + \bar{\eta} + \delta\eta \tag{2.11}$$

In this equation, η_g represents the geoid height, $\bar{\eta}$ denotes the Mean Dynamic Topography, and $\delta\eta$ signifies the *SLA*. The geoid height refers to the elevation of the geoid relative to the reference ellipsoid. The geoid itself is a gravity equipotential surface that would coincide with the ocean surface if it were at rest, unaffected by currents and solely influenced by the gravitational field. The Mean Dynamic Topography represents the average sea surface height (SSH) above the geoid over a defined time period. Both the geoid height and Mean Dynamic Topography are treated as time-invariant. Therefore, to calculate the *SLA*, one simply needs to subtract a climatological mean, which is an average derived from a long-term dataset, from the measured value of η

$$\delta\eta = \eta - \eta_{clm} \tag{2.12}$$

where η_{clm} is a mean over the period 2012-1993 [60].

2.2 Reanalyses data

The reanalyses utilized in this study to obtain correlation function and as a validation set are derived from an Ocean General Circulation Model, which is founded on the NEMO framework (Nucleus for European Modelling of the Ocean) and from a data assimilation scheme, known as OceanVar.

2.2.1 Ocean General Circulation Model

NEMO framework encompasses various ocean-related engines that address ocean dynamics and thermodynamics, sea-ice dynamics and thermodynamics, as well as biogeochemistry. It serves as a versatile instrument for investigating the ocean and its interactions with other elements of the Earth's climate system across a broad spectrum of spatial and temporal scales [61].
This model resolves the primitive equations utilizing a non-linear free surface formulation and employs time-varying vertical z-star coordinates. It operates with a uniform horizontal resolution of 0.042° and incorporates 141 vertically uneven levels. The model is forced by momentum, water and heat fluxes interactively which are computed interactively via bulk formulae, drawing on operational analysis and forecast data from the European Centre for Medium-Range Weather Forecasts (ECMWF) at a 6-hour frequency and a horizontal resolution of 0.125° and on the model predicted surface temperatures. The calculation of surface water flux is based on the net effect of evaporation, precipitation, and runoff. Evaporation is derived from latent heat flux, while precipitation data is provided by ECMWF, aligning with the temporal resolution of other atmospheric forcing fields. Furthermore, the model incorporates 39 rivers as volume inputs from monthly mean datasets, and the Dardanelles Strait is closed but modeled as a net volume input through a river-like parameterization. A quadratic bottom drag coefficient is utilized at the seabed, with topographical data sourced from the General Bathymetric Chart of the Oceans (GEBCO) [62].

2.2.2 Variational methods and 3D-VAR

OceanVar represents a variational approach, specifically a 3D-VAR method. This section aims to clarify the derivation of a variational equation for the optimal analysis from the broadly applicable Bayesian equation 1.1, as discussed in [63]. According to equation 1.1, the posterior probability can be expressed as being proportional to the product of the prior probability and the likelihood. To simplify our analysis, we will disregard the time indices and represent the complete series with $\mathbf{x}$ and $\mathbf{y}$, resulting in the following formulation:

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \tag{2.13}$$

When considering a "perfect" model, it is possible to conclude that the prior probability is determined exclusively by the model's initial state, which can be articulated in terms of deviations from a known background $\mathbf{x}_b$

$$p(\mathbf{x}) = p_b(\mathbf{x} - \mathbf{x}_b) \tag{2.14}$$

Regarding the likelihood component, it is understood to encapsulate the observational error, which can be decomposed into two distinct parts: the first part pertains

to the measurement error, while the second part arises from the transition from the model to the observation via a non-linear observational operator H. This operator links the model state vector to the observations and can be represented as $p_f(y_k - H(x_k))$, where the index k and the absence of bold notation indicate a specific element within the time series. By considering all possible states along with the measurement error, we can express the relationship as follows:

$$p(\mathbf{y}|\mathbf{x}) = \int p_o(y_l - y_k)p_f(y_l - H(x_k))dy_l \tag{2.15}$$

where the integration encompasses all elements within the time series. Consequently, we arrive at:

$$p(\mathbf{x}|\mathbf{y}) \propto \Big[\int p_o(y_l - y_k)p_f(y_l - H(x_k))dy_l\Big]p_b(\mathbf{x} - \mathbf{x}_b) \tag{2.16}$$

To further our investigation, it is important to clarify that in the framework of 3D-VAR, we specify the probability distributions p_o, p_f and p_b, which are conventionally assumed to be multi-dimensional Gaussian functions

$$p_b(\mathbf{x} - \mathbf{x}_b) \propto \exp\Big[-\frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)\Big] \tag{2.17}$$

$$p_b(\mathbf{y} - \mathbf{y}_t) \propto \exp\Big[-\frac{1}{2}(\mathbf{y} - \mathbf{y}_t)^T\mathbf{O}^{-1}(\mathbf{y} - \mathbf{y}_t)\Big] \tag{2.18}$$

$$p_b(\mathbf{y} - H(\mathbf{x})) \propto \exp\Big[-\frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T\mathbf{F}^{-1}(\mathbf{y} - H(\mathbf{x}))\Big] \tag{2.19}$$

where $\mathbf{B}$, $\mathbf{O}$ and $\mathbf{F}$ are covariance matrices representing the error of the three components and the subscript t indicates the true state. Bayes probability function becomes

$$p(\mathbf{x}|\mathbf{y}) \propto \exp\Big[-\frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T(\mathbf{O}+\mathbf{F})^{-1}(\mathbf{y} - H(\mathbf{x})) - \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)\Big]$$

Moreover calling $O + F = R$ it can be rewritten as

$$p(\mathbf{x}|\mathbf{y}) \propto \exp\Big[-\frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x})) - \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)\Big] \tag{2.20}$$

Variational methods targets the mode of the system so they mean to maximize this probability. Maximizing it is equivalent to minimize $-\ln(p(\mathbf{x}|\mathbf{y}))$ which becomes known as cost function

$$J = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x})) \tag{2.21}$$

Within this equation, the symbol $\mathbf{x}$ is utilized to represent the analysis state vector, whereas $\mathbf{x}_b$ is indicative of the background state vector. The term $\mathbf{B}$ refers to the background error covariance matrix, and $\mathbf{R}$ signifies the observational error covariance matrix. Furthermore, H functions as the non-linear observational operator that establishes a connection between the model state vector and the observations. The OceanVar methodology is based on this equation, and further details can be found in 2.2.3.

2.2.3 OceanVar

OceanVar scheme is detailed in [64] and [65]. It operates through an iterative process aimed at minimizing the cost function described before in equation 2.21.
The equation however undergoes linearization about the background state, and the variable is substituted with a control variable labeled $\mathbf{v}$ to eliminate the necessity of inverting $\mathbf{B}$. As a result, the cost function is reformulated as:

$$J = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}(\mathbf{d} - \mathbf{HVv})^T\mathbf{R}^{-1}(\mathbf{d} - \mathbf{HVv}) \tag{2.22}$$

In this context the vector $\mathbf{v}$ is defined in such a manner that the relationship $\mathbf{x}-\mathbf{x}_b = \mathbf{Vv}$ is satisfied. To do so it is assumed that the matrix $\mathbf{B}$ can be represented as $\mathbf{B} = \mathbf{VV}^T$, thus employing the transformation matrix $\mathbf{V}$. The quantity $\mathbf{d}$, known as the misfit, is defined by the equation $\mathbf{d} = \mathbf{y} - H(\mathbf{x}_b)$, which captures the deviations of the observations. Finally, the operator $\mathbf{H}$ is the linearized observational operator assessed at $\mathbf{x} = \mathbf{x}_b$ [64].
Owing to its considerable size, the transformation matrix V is represented at each iteration of minimization as a series of linear operators. It is decomposed into multiple components that include balance operators ($\mathbf{V}_\eta$), horizontal correlations ($\mathbf{V}_h$), and vertical covariances ($\mathbf{V}_v$), articulated as $\mathbf{V} = \mathbf{V}_\eta\mathbf{V}_h\mathbf{V}_v$. The operator $\mathbf{V}_\eta$ serves as the sea level operator, converting increments in temperature and salinity into corresponding sea level increments, which are essential for the assimilation of altimetric data. Horizontal correlations are modeled using a four-iteration first-order recursive filter, applied sequentially in both zonal and meridional directions. Vertical covariances are characterized by a set of tri-variate Empirical Orthogonal Functions (EOFs) that represent salinity and temperature at the full vertical resolution of the model, derived from anomalies relative to the long-term mean [65].
The assimilated information includes satellite-derived Sea Level Anomaly (SLA), which considers the atmospheric pressure effects, in addition to vertical profiles of temperature and salinity collected from Argo, XBT, and glider systems. Moreover, objectively analyzed Sea Surface Temperature (SST) fields are applied for the correction of surface heat fluxes [62].

2.2.4 Temperature, salinity and sea surface height data

For the purpose of evaluating correlation factors and for validation of results, temperature, salinity and sea surface height data have been sourced from the Mediterranean reanalyses products previously outlined, which are generated by the Centro euro-Mediterraneo sui Cambiamenti Climatici (CMCC) and provided through the Copernicus Marine Service (CMEMS) platform [66]. The dataset consists of three-dimensional daily measurements from 1993 to 2020, characterized by a resolution of 0.042° and 141 unevenly spaced levels. Since this resolution exceeds that of the satellite altimetry data, discussed in 2.1.3, a linear interpolation was conducted to achieve consistency in resolution.

Chapter 3

Methods and preprocessing

In Chapter 2, the attributes of the satellite and reanalysis data were examined. This chapter focuses on the process of combining these datasets to extract profiles of temperature and salinity, as well as the necessary data preprocessing steps involved in this process. Understanding the statistical relationships among temperature, salinity, and sea level allows the estimation of two of these variables when third is measured. This study employs two distinct methodologies to achieve this: the approach developed by Ezer and Mellor [45] and the application of Adani et al. [47] EOFs. Section 3.1 provides an overview of the method proposed by Mellor and Ezer, which enables the inference of these profiles. The concept of Empirical Orthogonal Functions is introduced in section 3.2, offering an alternative approach for reconstructing temperature and salinity profiles. Section 3.4 addresses the different types of time averages employed and the necessity of their application. Lastly section 3.3 underscores the diverse preprocessing strategies that are essential for the effective implementation of the methods outlined.

3.1 Mellor and Ezer's Method

As highlighted in section 2.1.1, a direct relationship exists between sea level measurements and subsurface conditions. In their research from 1991 [45] and 1994 [46], Mellor and Ezer employed this link to derive temperature and salinity profiles from altimetry data. The core of their methodology for extending surface data into subsurface fields is based on these linear regression

$$\delta T(x, y, z, t) = F^T(x, y, z)\delta\eta(x, y, t) \tag{3.1}$$

$$\delta S(x, y, z, t) = F^S(x, y, z)\delta\eta(x, y, t) \tag{3.2}$$

Within this framework, F^T and F^S are established functions known as correlation factors, which are invariant over time. The anomalies $\delta\eta(x, y, t)$ are characterized as the instantaneous deviations from their respective time-averaged values. Additionally, $\delta T(x, y, z, t)$ and $\delta S(x, y, z, t)$ denote the instantaneous anomalies of reconstructed temperatures and salinities, respectively, to which a time-average must be incorporated in order to accurately reconstruct the true profiles of temperature and salinity. The coordinates (x, y) correspond to latitude and longitude, while z indicates depth. Referring to each grid point in the model as i, the correlation factors

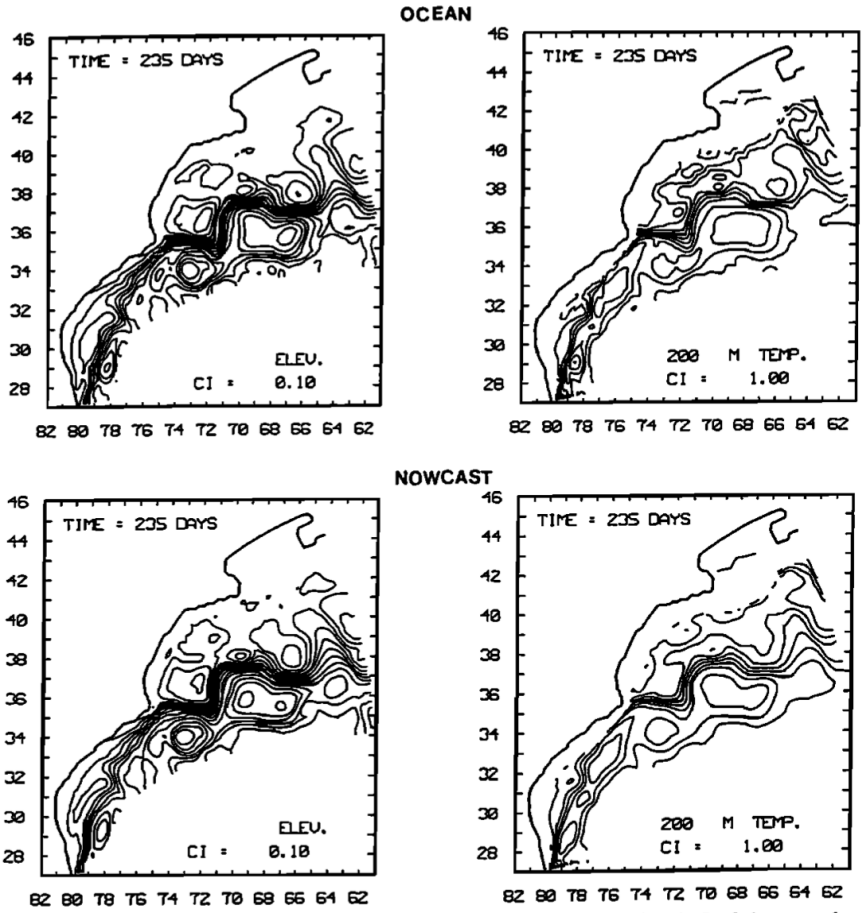


Figure 3.1: Surface elevation (left) and temperature at 200m (right) of the control ocean (top) and of the nowcast (bottom). Figure from [45]

at these points can be defined mathematically as:

$$F_i^T = \frac{\overline{\delta T_i \delta \eta_i}}{\overline{\delta \eta_i^2}} \tag{3.3}$$

$$F_i^S = \frac{\overline{\delta S_i \delta \eta_i}}{\overline{\delta \eta_i^2}} \tag{3.4}$$

The overbar notation in equations 3.3 and 3.4 signifies a temporal average. This research calculates correlation factors and its sensitivity to the temporal averages. Section 3.4 offers a detailed explanation of the time averages utilized and the specific periods for which the correlation factors are derived.
Mellor and Ezer assess the accuracy of their equations by conducting a 10-day diagnostic run of their model (nowcast), which yields dynamically adjusted velocity fields and a revised elevation field. Upon comparing the elevation fields and 200m temperature fields with the corresponding control fields as shown in figure 3.1, they discovered that the global (area-averaged) root mean squared error between the control elevation and the nowcast elevation is approximately 0.04 m, representing a 15% error relative to the unassimilated ocean data for the same day. Subsequently, they incorporated this approach into their assimilation scheme, which is founded on an optimal interpolation method [45].

3.2 Empirical Orthogonal Functions

As outlined in section 2.2.3, the assimilation methodology employed in the reanalyses of the Mediterranean Sea derives vertical covariances from trivariate Empirical Orthogonal Functions as described in Adani et al (2011) [47]. A primary objective of this thesis is to utilize this method for the direct inference of temperature and salinity profiles and compare it to the Mellor and Ezer (1991) [45] method. To facilitate a comprehensive understanding of the process, an overview of Empirical Orthogonal Functions and their evaluation is provided in 3.2.1, followed by their application within this thesis in 3.2.2.

3.2.1 Overview of EOFs

Introduced by Lorenz in 1956 [67], Empirical Orthogonal Function (EOF) analysis is a technique utilized to identify a series of natural or empirical orthogonal functions that can be linearly combined to represent a given dataset [68]. This analytical approach specifically decomposes the temporal variance of the dataset into orthogonal spatial patterns, referred to as empirical eigenvectors. When these eigenvectors are arranged in order, each one accounts for the maximum possible remaining variance in the dataset [69]. Each eigenvector is associated with a time series that illustrates the temporal variation in the amplitude of the pattern, known as the principal component (PC). Collectively, each eigenvector and its corresponding time series are termed a mode, which can be analyzed for potential physical significance [68]. By combining together all modes, one can reconstruct the original dataset.
From a mathematical perspective, Empirical Orthogonal Functions (EOFs) can be interpreted as the eigenvectors of the covariance matrix of the data, where the matrix elements represent anomalies. These anomalies are defined as the deviations of the observed or modeled variables from their long-term mean [69].
To derive EOFs, we construct a state vector that incorporates all anomalies of the L variables, measured at M locations and across N sequential time intervals. The matrix that represents the anomaly of a single variable, denoted as $\mathbf{v}$, is expressed as follows:

$$\mathbf{v} = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1N} \\ v_{21} & v_{22} & \dots & v_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ v_{M1} & v_{M2} & \dots & v_{MM} \end{bmatrix}$$

The matrix that represents the entire state vector is determined by the number of variables considered. For example, if we examine two variables, $\mathbf{v}$ and $\mathbf{w}$, the matrix will be of size $2M \times N$. It can be represented mathematically as:

$$\mathbf{X} = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \tag{3.5}$$

where $\mathbf{v}$ and $\mathbf{w}$ are defined as previously indicated. The covariance matrix for which eigenvectors are to be determined is formulated as:

$$\mathbf{C} = \frac{1}{N}\mathbf{X}\mathbf{X}^T \tag{3.6}$$

The process of finding eigenvalues and eigenvectors involves solving the eigenvalue problem, which can be articulated as:

$$\mathbf{C}\overrightarrow{e_m} = \lambda_m \overrightarrow{e_m} \tag{3.7}$$

In this context, $\overrightarrow{e}$ denotes the eigenvector (also referred to as the EOF), while λ signifies the associated eigenvalue.
The process of solving the eigenvalue problem is generally complex. An alternative approach to computing empirical orthogonal functions (EOFs) involves the application of singular value decomposition (SVD). This mathematical technique in linear algebra asserts that any matrix $\mathbf{A}$ of dimensions $M \times N$ and rank r can be expressed as a product of matrices in the following form:

$$\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

In this representation, $\mathbf{U}$ is an orthogonal matrix of dimension $M \times r$ that contains the left singular vectors, $\mathbf{\Sigma}$ is a diagonal matrix of dimension $r\times$ filled with singular values and $\mathbf{V}$ is an orthogonal matrix of dimension $r \times N$ that includes the right singular vectors. It can be asserted that the left singular vectors correspond to the patterns of Empirical Orthogonal Functions (EOFs), while the right singular vectors relate to principal components, and the singular values represent the square root of the associated eigenvalues. . This methodology allows for the extraction of all components required for EOF analysis, offering enhanced numerical stability and, in many instances, significantly improved computational speed [70].

3.2.2 Application of EOFs

In this research, as done in Adani et al. (2011) [47], vertical EOFs are utilized to derive the covariance matrix, which plays a crucial role in the formulation of new weights to infer temperature and salinity profiles.
As outlined in section 3.2.1, EOFs are derived from a state vector that encapsulates the anomalies of specific variables in relation to a defined mean. These anomalies require normalization to ensure that all quantities within the state vector are dimensionless and appropriately weighted by their standard deviations, facilitating comparability. The variables examined in this investigation include temperature, salinity, and sea level. Both temperature and salinity are structured as vectors with dimensions $D \times N$, where D corresponds to the number of levels associated with the maximum depth at the specific location under consideration, and N indicates the number of measurements taken, which depends on the temporal window considered. Normalization of these variables is performed using a standard deviation, denoted as σ_T and σ_S, which is independent of depth and calculated as the average of the standard deviations across all levels

$$\delta\mathbf{T}_{norm} = \begin{bmatrix} \frac{\delta\mathbf{T}_1}{\sigma_T} \\ \frac{\delta\mathbf{T}_2}{\sigma_T} \\ \vdots \\ \frac{\delta\mathbf{T}_D}{\sigma_T} \end{bmatrix} \qquad \delta\mathbf{S}_{norm} = \begin{bmatrix} \frac{\delta\mathbf{S}_1}{\sigma_S} \\ \frac{\delta\mathbf{S}_2}{\sigma_S} \\ \vdots \\ \frac{\delta\mathbf{S}_D}{\sigma_S} \end{bmatrix}$$

In this context, the vectors $\delta\mathbf{T}_i$ and $\delta\mathbf{S}_i$ are characterized by a dimension of N, which corresponds to the time series data of temperature and salinity anomalies at

the level indexed by i. Likewise, the vector $\delta\boldsymbol{\eta}$ undergoes normalization with respect to its standard deviation, σ_η, leading to the formulation:

$$\delta\boldsymbol{\eta}_{norm} = \frac{\delta\boldsymbol{\eta}}{\sigma_\eta}$$

A secondary normalization process has been applied to temperature and salinity data. To ensure that the EOFs are independent of spatial geometry—specifically the number of levels and their respective thicknesses—we have introduced a factor denoted as $\mathbf{g}$ [71]. The elements of this factor reflect the relative contribution of each layer compared to the total water column, represented by the following equation:

$$\mathbf{g} = [\frac{\Delta z_1}{H}, \frac{\Delta z_2}{H}, \ldots, \frac{\Delta z_D}{H}]$$

Thickness Independent vector anomalies are derived by multiplying the previously normalized anomalies by the factor $\mathbf{g}$. We will refer to these as $\delta\mathbf{T}^{TI}_{norm}$ and $\delta\mathbf{S}^{TI}_{norm}$. The introduction of this normalization is justified by the expectation that thicker layers will have a relatively greater influence on surface elevation errors compared to thinner layers. Therefore, the multiplication of temperature and salinity anomalies by the layer thickness yields a representation that more accurately reflects the underlying physical relationships between the temperature and salinity fields and the changes observed in sea surface elevation [71].
The normalized anomaly vectors are initially employed to calculate bi-variate Empirical Orthogonal Functions. These EOFs are derived from the Singular Value Decomposition of a state vector that consists solely of temperature and sea level anomalies or salinity and sea level anomalies. The state vectors can be represented as follows:

$$\mathbf{X}^T = \begin{bmatrix} \delta\mathbf{T}^{TI}_{norm} \\ \delta\boldsymbol{\eta}_{norm} \end{bmatrix} \qquad \mathbf{X}^S = \begin{bmatrix} \delta\mathbf{S}^{TI}_{norm} \\ \delta\boldsymbol{\eta}_{norm} \end{bmatrix} \tag{3.8}$$

Specifically, the anomalies associated with temperature and salinity are represented as vectors of dimension $D \times N$, while the sea level anomalies are characterized by a single value at each time point, resulting in vectors of dimension N. Consequently, each state vector is structured as a matrix of dimension $(D+1) \times N$.
Subsequently, similar to the approach utilized in OceanVar, tri-variate EOFs have been employed [72]. The state vector has been consolidated into a singular representation for both variables, expressed as

$$\mathbf{X} = \begin{bmatrix} \delta\mathbf{T}^{TI}_{norm} \\ \delta\mathbf{S}^{TI}_{norm} \\ \delta\boldsymbol{\eta}_{norm} \end{bmatrix} \tag{3.9}$$

In this context, the state vector is characterized by a matrix of dimensions $(2D + 1) \times N$. To enhance clarity, the expressions $\delta\mathbf{T}^{TI}_{norm}$, $\delta\mathbf{S}^{TI}_{norm}$, and $\delta\boldsymbol{\eta}_{norm}$ will be simplified to $\delta\mathbf{T}$, $\delta\mathbf{S}$, and $\delta\boldsymbol{\eta}$.
The determination of the necessary number of Empirical Orthogonal Functions we have opted to evaluate the extent to which they explain for the variance within the dataset. The contribution of each EOF to the overall variance is quantified by its corresponding eigenvalue. The proportion of variance explained is expressed as follows:

$$PVE = \frac{\lambda_m}{\sum_{i=1}^{M} \lambda_i} * 100 \tag{3.10}$$

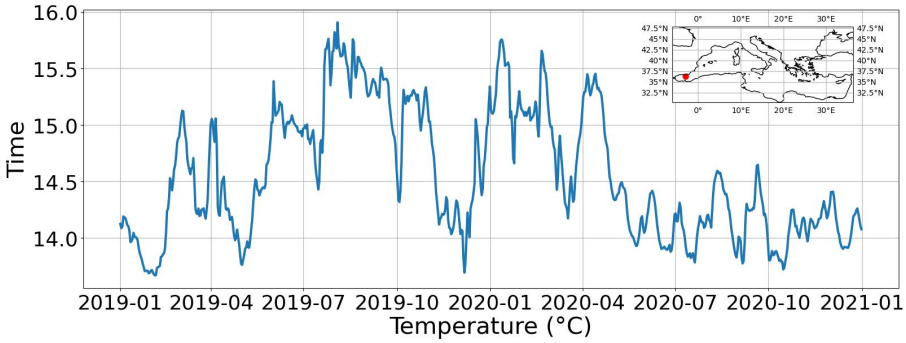


Figure 3.2: Time series of temperature reanalysis data at a depth of 105 meters in a specific location in the Alboran Sea, identified by the coordinates (-2.9375 E, 36.3125 N) over the time period spanning from 2019 to 2020.

where M denotes the total count of EOFs. Typically, a significant fraction of the total variance can be captured by a limited number of modes, while the remaining variance is attributed to minor features, smaller-scale variations, and noise [69]. This investigation adopts a threshold of at least 99.7% of the variance consistently explained, which necessitates the application of 15 EOFs in the bivariate scenario and 25 EOFs in the trivariate scenario.
Upon completing the computation and selection of EOFs, the covariance matrix is derived as follows:

$$\mathbf{C} = \mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^T \tag{3.11}$$

Here, $\mathbf{U}$ signifies the matrix that contains the EOFs, and $\boldsymbol{\Sigma}$ is the diagonal matrix comprising the eigenvalues. In the context of bi-variate EOFs, the covariance matrix is constructed as follows:

$$\mathbf{C}^T = \begin{bmatrix} \delta\mathbf{T}\delta\mathbf{T} & \delta\boldsymbol{\eta}\delta\mathbf{T} \\ \delta\mathbf{T}\delta\boldsymbol{\eta} & \delta\boldsymbol{\eta}\delta\boldsymbol{\eta} \end{bmatrix} \qquad \mathbf{C}^S = \begin{bmatrix} \delta\mathbf{S}\delta\mathbf{S} & \delta\boldsymbol{\eta}\delta\mathbf{S} \\ \delta\mathbf{S}\delta\boldsymbol{\eta} & \delta\boldsymbol{\eta}\delta\boldsymbol{\eta} \end{bmatrix} \tag{3.12}$$

In this context, $\delta\mathbf{T}\delta\mathbf{T}$ is used to express the covariance of temperature with itself, whereas $\delta\mathbf{S}\delta\mathbf{S}$ pertains to the covariance of salinity with itself. Furthermore, the covariances involving temperature and salinity with sea level are represented by $\delta\boldsymbol{\eta}\delta\mathbf{T}$, $\delta\mathbf{T}\delta\boldsymbol{\eta}$, $\delta\boldsymbol{\eta}\delta\mathbf{S}$, and $\delta\mathbf{S}\delta\boldsymbol{\eta}$. Lastly, $\delta\boldsymbol{\eta}\delta\boldsymbol{\eta}$ reflects the covariance of sea level with itself. For trivariate EOFs the covariance matrix is

$$\mathbf{C} = \begin{bmatrix} \delta\mathbf{T}\delta\mathbf{T} & \delta\mathbf{S}\delta\mathbf{T} & \delta\boldsymbol{\eta}\delta\mathbf{T} \\ \delta\mathbf{T}\delta\mathbf{S} & \delta\mathbf{S}\delta\mathbf{S} & \delta\boldsymbol{\eta}\delta\mathbf{S} \\ \delta\mathbf{T}\delta\boldsymbol{\eta} & \delta\mathbf{S}\delta\boldsymbol{\eta} & \delta\boldsymbol{\eta}\delta\boldsymbol{\eta} \end{bmatrix} \tag{3.13}$$

with the same meaning as before.
The specific elements we aim to extract are the covariances between temperature and sea level, as well as between salinity and sea level. These covariances are extracted from the full covariance matrix. They are subsequently applied in equations 3.3 and 3.4 as the numerator, rather than being derived solely from the product of anomalies.

3.3 Preprocessing of data

As emphasized in chapter 2.2.4, the primary data utilized in this study consist of satellite altimetry data along with interpolated temperature, salinity, and sea

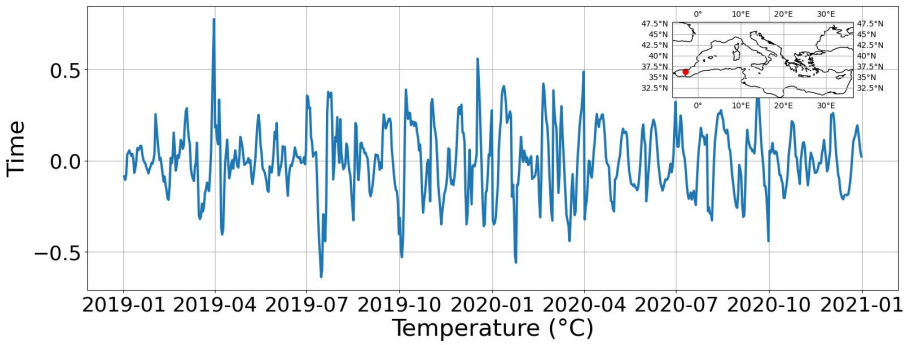


Figure 3.3: Time series of detrended temperature anomalies (obtained with the subtraction of a monthly climatology) at a depth of 105 meters in a specific location in the Alboran Sea, identified by the coordinates (-2.9375 E, 36.3125 N) over the time period spanning from 2019 to 2020.

surface height data derived from reanalysis. While satellite altimetry provides sea level anomalies directly, the other three variables require computation to extract their respective anomalies. As a matter of facts these variables are presented as basic time series, as shown in figure 3.2. To extract the necessary anomalies, a climatological average is calculated and then deducted from the complete dataset, as illustrated in the following equations:

$$\delta T = T - T_{clm}$$

$$\delta S = S - S_{clm}$$

$$\delta \eta = \eta - \eta_{clm}$$

The climatological average can be established on either a daily or monthly basis, which is crucial for analyzing the sensitivity of correlation factors to temporal averages. Consequently, both daily and monthly correlation factors are examined, which means that to ensure consistency, it is essential to consider the anomalies that arise from the subtraction of both climatological averages.
This dual approach to climatology allows for the identification of two distinct trends in the anomalies, which are depicted in figure 3.4. The first is a long-term linear trend characterized by rising temperatures, salinity, and sea levels, which are indicative of long term variability, and is evident when either daily or monthly climatology is subtracted. The second type of trend can manifest within a single month, where temperatures may fluctuate either upward or downward based on the specific month, and is only observable when the monthly climatology is subtracted. To ensure the efficacy of our methods, it is essential to eliminate both trends. This can be achieved by fitting a linear polynomial to the anomalies to compute the linear trend, which is subsequently subtracted from the anomalies. The long-term trend fitting encompasses the entire timeseries, denoted as δT_T, δS_T, and $\delta \eta_T$. Conversely, the monthly trend is calculated on a month-by-month basis, resulting in trends labeled δT_{Tm}, δS_{Tm}, and $\delta \eta_{Tm}$. Following this, the first contribution is consistently eliminated, whereas the second contribution is discarded solely when a monthly climatology has been subtracted. For instance, in the case of anomalies derived from the subtraction of a monthly average, the process is as follows

$$\delta T_{det} = \delta T - \delta T_T - \delta T_{Tm}$$

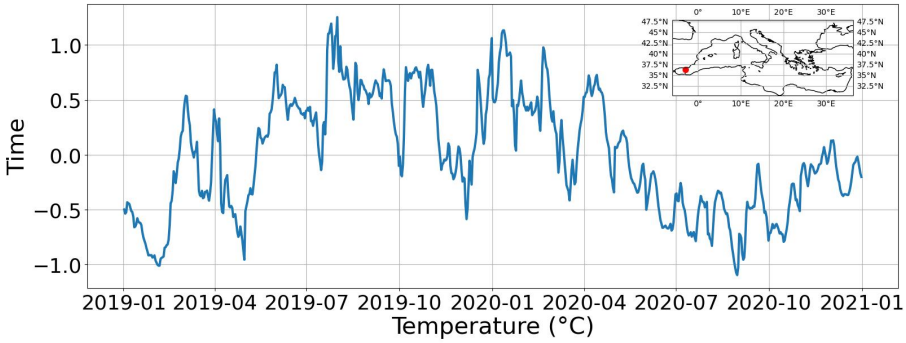


Figure 3.4: Time series of temperature anomalies (obtained with the subtraction of a monthly climatology) at a depth of 105 meters in a specific location in the Alboran Sea, identified by the coordinates (-2.9375 E, 36.3125 N) over the time period spanning from 2019 to 2020.

$$\delta S_{det} = \delta S - \delta S_T - \delta S_{Tm}$$

$$\delta \eta_{det} = \delta \eta - \delta \eta_T - \delta \eta_{Tm}$$

Upon the elimination of the appropriate contributions, the resulting detrended anomalies display oscillatory behavior centered around a mean of zero. This is depicted in figure 3.3, which illustrates the anomalies derived from the subtraction of a monthly average.

3.4 Different time averages

This study seeks to identify the optimal time averages for obtaining the most accurate reconstructions of salinity and temperature profiles for the two years 2019-2020. This objective can be approached by considering both the duration of the periods utilized in calculating the averages and the number of correlation factors employed, which are associated with the time intervals selected for their computation (whether in months or days).
In both scenarios, it is essential to emphasize the importance of consistency. Specifically, once a particular average is selected for climatological assessment, it is imperative that the same average is applied when evaluating the correlation factors.
As indicated in section 3.1, correlation factors are computed on both monthly and daily intervals. Comparing these factors is crucial due to the significant variability in the timescales of oceanic processes, as illustrated in figure 1.4. While daily weights are intended to reflect rapid fluctuations, they may occasionally interfere with slower phenomena, such as mesoscale eddies, which operate on timescales ranging from weeks to years [32].
Monthly correlation factors are established such that each month of the year is associated with a distinct factor. This implies that climatological data also exhibit distinct values for each month. To assess these factors, we first develop a function that allows us to filter through all dates in our dataset, selecting only those that correspond to a specific month, which are in number N. The data corresponding to the identified month is then compiled into a new dataset, which serves as the basis for applying equations 3.1 and 3.2. This process is analogous to executing the

following procedure at each grid point i:

$$F_i^T = \frac{\frac{\sum_{j=1}^{N} \delta T_{ij} \delta \eta_{ij}}{N}}{\frac{\sum_{j=1}^{N} \delta \eta_{ij}^2}{N}} \tag{3.14}$$

In this analysis, we have aggregated the covariance for all days corresponding to a specific month and subsequently divided this sum by the total number of days within that month. The methodology for calculating weights derived from Empirical Orthogonal Functions follows a similar approach, involving the construction of a new dataset that exclusively includes days from the designated month. This newly formed dataset will serve as the basis for generating the state vector for bi- or tri-variate EOFs, upon which singular value decomposition will be performed.

$$\mathbf{X}^T = \begin{bmatrix} \delta \mathbf{T}_{month} \\ \delta \boldsymbol{\eta}_{month} \end{bmatrix} \qquad \mathbf{X}^S = \begin{bmatrix} \delta \mathbf{S}_{month} \\ \delta \boldsymbol{\eta}_{month} \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} \delta \mathbf{T}_{month} \\ \delta \mathbf{S}_{month} \\ \delta \boldsymbol{\eta}_{month} \end{bmatrix} \tag{3.15}$$

The covariance matrices obtained through equation 3.11 are specifically associated with individual months. Consequently, it is crucial to choose the appropriate matrix when reconstructing temperature and salinity for a particular day, ensuring that the month corresponding to the desired date is taken into account.
The daily correlation factors indicate that the analysis is conducted not on a monthly basis, but rather on a daily one. This results in the generation of 366 distinct weights instead of the usual 12. In this context, daily climatology, which varies for each day, is subtracted in equations section 3.3 and section 3.3. The function responsible for selecting the appropriate dataset now examines all available dates and isolates those that correspond to the specific day of the year (e.g., January 1st, January 2nd, etc.). Although the number of selected days is smaller than that of the days in a month, denoted as n, the methodology remains consistent for both the Mellor and Ezer approaches as well as the EOFs techniques

$$F_i^T = \frac{\frac{\sum_{j=1}^{n} \delta T_{ij} \delta \eta_{ij}}{n}}{\frac{\sum_{j=1}^{n} \delta \eta_{ij}^2}{n}}$$

$$\mathbf{X}^T = \begin{bmatrix} \delta \mathbf{T}_{day} \\ \delta \boldsymbol{\eta}_{day} \end{bmatrix} \qquad \mathbf{X}^S = \begin{bmatrix} \delta \mathbf{S}_{day} \\ \delta \boldsymbol{\eta}_{day} \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} \delta \mathbf{T}_{day} \\ \delta \mathbf{S}_{day} \\ \delta \boldsymbol{\eta}_{day} \end{bmatrix}$$

The construction of monthly and daily datasets relies on two distinct timeframes. The first timeframe encompasses a comprehensive dataset spanning 28 years, from 1993 to 2020. This extensive duration allows for the assessment of both monthly and daily correlation factors, as it provides a sufficient number of data points for daily averages. The climatological analysis conducted during this evaluation is based on the entirety of the 28-year period. The second timeframe is limited to the two years of interest, specifically 2019 and 2020. Due to the brevity of this period, daily weights cannot be calculated, as only two data points are available. Consequently, only monthly correlation factors have been analyzed, with the climatology for this assessment also derived from the two-year span to maintain consistency. The utilization of both timeframes is crucial; the shorter period allows for the identification of anomalies or changes specific to those years, while the longer period is more effective in revealing trends and broader climatic shifts.

Chapter 4

Assessement of methods

The algorithms detailed in chapter 3 are utilized in this one to extract temperature and salinity profiles for the years 2019-2020. We will illustrate their functionality in a synthetic environment, where sea level anomaly is also calculated from reanalysis data. This setup facilitates a comprehensive analysis of the performance of the various time averages and methods.
The process of retrieving temperature profiles from four distinct locations in the Mediterranean Sea is outlined in section 4.1, emphasizing the time periods selected for correlation factors and the effectiveness of the different techniques employed. Additionally, section 4.2 will address the reconstruction of mesoscale eddies.

4.1 Profiles reconstruction

4.1.1 Locations under analysis

The algorithms designed for the retrieval of temperature and salinity profiles, as outlined in chapter 3, are initially applied to four distinct locations within the Mediterranean Sea, as depticted in figure, 4.1 to evaluate their efficacy in reproducing individual profiles:

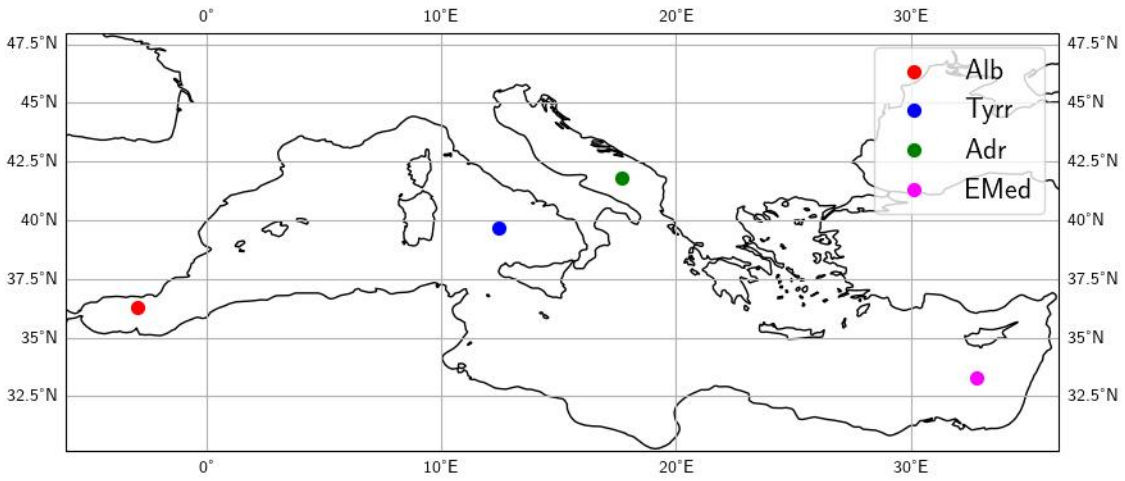


Figure 4.1: The four points used in this analysis

- *Alb* is situated near the Strait of Gibraltar in the Alboran Sea, an area characterized by significant water exchange with the Atlantic Ocean;
- *Tyrr* is located in the Tyrrhenian Sea, recognized as the deepest and most isolated basin in the western Mediterranean;
- *Adr* is found in the South Adriatic Sea, a region where Adriatic Deep Water is generated;
- *EMed* is positioned near the eastern Mediterranean coast, specifically in the Levantine Sea.

These four locations exhibit distinct thermohaline structures, primarily influenced by the circulation patterns of water masses. Thus, examining their reconstruction is valuable for providing insights into the effectiveness of our algorithms. The key characteristics of the water masses present in each area are subsequently outlined.
The Alboran Sea exhibits a notable minimum salinity near its surface, a phenomenon attributed to the influx of Atlantic Water originating from the Strait of Gibraltar. Beneath this surface layer, there exists a zone where the water characteristics suggest a blend of both Atlantic and Mediterranean waters. At depths ranging from approximately 300 to 400 meters, one can observe the presence of Levantine Intermediate Water (LIW), which is distinguished by a peak in salinity. Below the LIW layer, the deep Mediterranean waters extend from depths of about 500-600 meters down to the seabed [73].
The Tyrrhenian Sea is characterized by the presence of various water masses that extend from the surface to the deeper layers of the basin. Notably, a significant influx of surface waters of Atlantic origin (AW) is typically found within the upper 100-200 meters of the water column. These waters are distinguished by a subsurface salinity minimum. At deeper levels, one can observe the presence of salty Levantine Intermediate Water (LIW) that enters from the Eastern Mediterranean via the Sicily Strait. This is indicated by an absolute maximum in salinity, which coincides with a relative maximum in temperature at depths of approximately 400-500 meters. The deepest strata are predominantly occupied by Western Mediterranean Deep Water (WMDW), which is characterized by lower temperature and salinity. The relatively weak dynamics in these deep layers, combined with the presence of gyre structures, promote mixing between LIW and WMDW, resulting in the creation of a unique water mass referred to as Tyrrhenian Deep Water, situated below 1000 meters [74, 75].
The Adriatic Sea is among the limited regions where deep water formation occurs. From depths of 150 meters to the ocean floor, the water exhibits conditions typical of the open Mediterranean Sea, resulting in a nearly uniform water column. The Levantine Intermediate Water (LIW) is present in a layer that extends from 150 to 400 meters. The southern Adriatic deep water (SAdDW) possesses unique average characteristics, with a temperature of about 13.1°C and a salinity of approximately 38.6 psu. This water is significantly warmer and saltier than the deep water found in other areas of the basin, indicating that it likely represents a mixture of LIW and local surface waters [76].
Within the Levantine Sea, Atlantic Water (AW) is detectable in the 50–100 meter layer, exhibiting the characteristic subsurface salinity minimum. The upper 50 meters are dominated by Levantine Surface Water, a saline water mass that originates

in the Levantine Basin and contributes to the formation of Levantine Intermediate Water (LIW). The LIW, marked by a subsurface salinity maximum, is clearly discernible between 100 and 400 meters in the Levantine Basin, with the Eastern Mediterranean Deep Water situated below this layer [77].

4.1.2 Reconstruction error statistics

For each location across the four basins, mean profiles and standard deviations have been calculated. In order to identify the optimal method, we will utilize various statistical error indicators, such as mean bias, standard deviation error, and unbiased root mean square error [78]. The presentation will be limited to the results from January, given that the outcomes for the remaining months are largely analogous.
Mean bias is a statistical measure that reflects the consistent differences between reconstructed datasets and reanalysis over a particular period. It can be represented mathematically as:

$$MB = \overline{r} - \overline{R} \tag{4.1}$$

In this formulation, $\overline{r}$ refers to the January mean of the reconstructed temperature profiles, whereas $\overline{R}$ refers to the January mean of the reanalysis profiles.
Standard deviation error assesses the extent to which the reconstruction effectively represents the amplitude of variability in temperature and salinity. This concept is articulated through the equation

$$SDE = \sigma_r - \sigma_R \tag{4.2}$$

in which σ_r indicates the standard deviation of the January reconstructions, and σ_R signifies the standard deviation obtained from January reanalysis.
Lastly the unbiased root mean square error (uRMSE) is mathematically expressed as follows:

$$uRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(r_i - \overline{r}) - (R_i - \overline{R})]^2} \tag{4.3}$$

In this equation, N represents the total number of observations in the January time series, while r_i and R_i denote the $i-th$ values of the reconstructed and reanalysis variables, respectively. The uRMSE serves as the unbiased counterpart to the root mean square error (RMSE), which quantifies the average of the squared differences between the reconstructed and reanalysis values. However, the RMSE can be influenced by biases present in either the mean or the amplitude of the fluctuations in the reconstructed data. The uRMSE addresses this issue by eliminating the mean bias from the calculations. Consequently, equation 4.3 can be reformulated as:

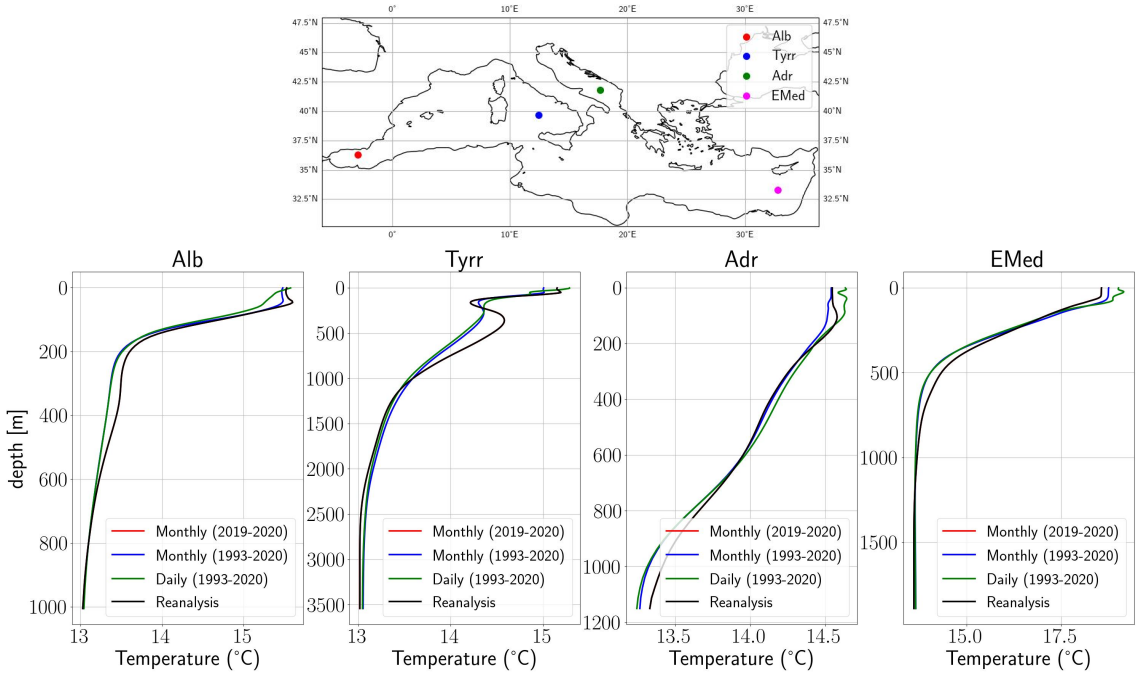
$$uRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[r_i - R_i - (\overline{r} - \overline{R})]^2}$$

This reformulation highlights the adjustment made for the mean bias.

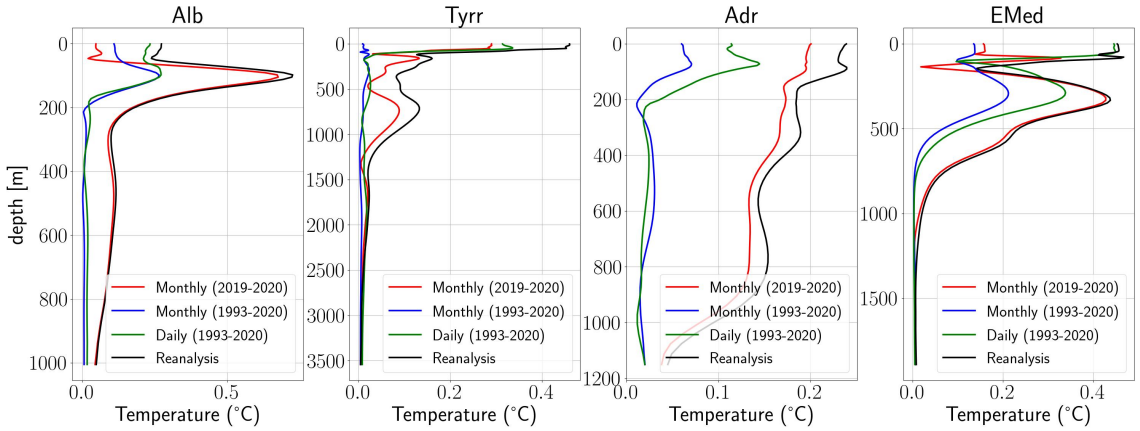
4.1.3 Comparison of reconstruction methods with different periods of time statistics

This section will focus on comparing the outcomes derived from three distinct time periods utilized for reconstructing salinity and temperature profiles for the years

2019-2020. For clarity, the reconstruction employing monthly correlation factors averaged over the 2019-2020 period will be referred to as MS (monthly short). The reconstruction based on monthly correlation factors from 1993 to 2020 will be designated as ML (monthly long), while the reconstruction utilizing daily correlation factors will be labeled as D. Given that the results across the three methodologies — Ezer and Mellor, bi-variate EOFs, and tri-variate EOFs — exhibit similarities, we will exclusively present the findings derived from the tri-variate EOFs.



(a) Mean

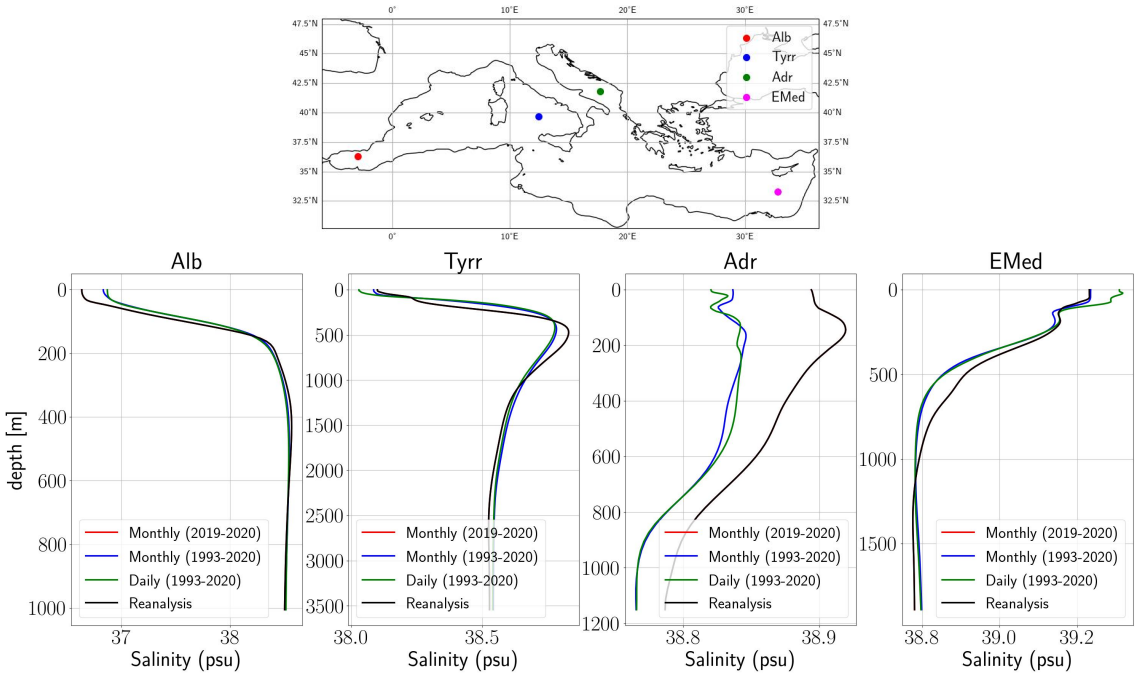


(b) Standard Deviation

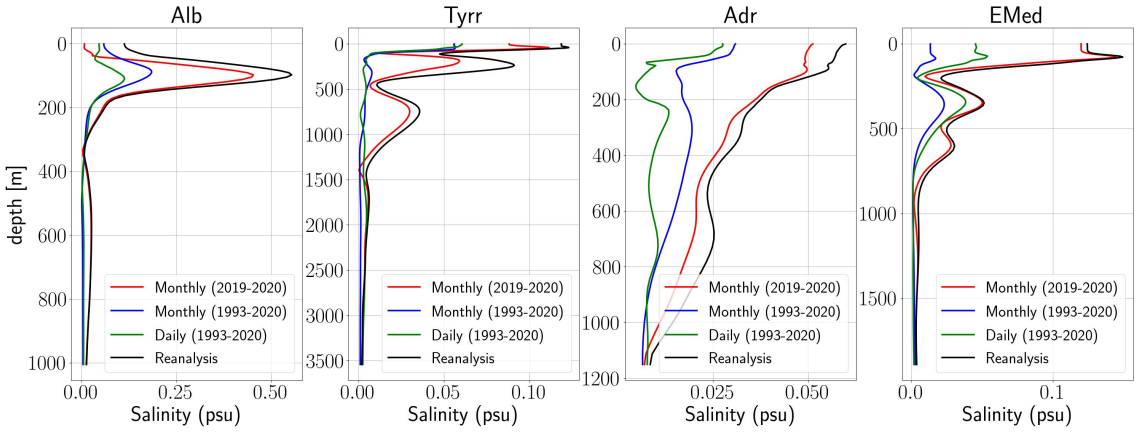
Figure 4.2: **January mean and standard deviation of temperature using tri-variate Empirical Orthogonal Functions.** Red profiles results from monthly correlation factors on period 2019- 2020, blue monthly correlation factors on period 1993 to 2020, green daily correlation factors.

The analysis of figures 4.2a and 4.3a reveals that the reconstructed means of temperature and salinity exhibit the characteristics outlined in section 4.1.1. In the case of *Alb*, a subsurface minimum in salinity is observed, followed by a layer exhibiting mixed properties up to a depth of 200 meters. At 400 meters, the salinity maximum

characteristic of LIW is also evident. The presence of AW is also clearly observable in the *Tyrr* profile, characterized by a subsurface salinity minimum, while the LIW is identified by a salinity maximum at a depth of 500 meters, accompanied by slightly elevated temperatures just above this depth. The characteristics of the *Adr* profile are akin to those previously described, though it exhibits slightly higher values for both temperature and salinity (13.3° instead of 13.1° and 38.76 instead of 38.6 psu). In this profile, LIW is observed at depths of approximately 100 to 150



(a) Mean



(b) Standard Deviation

Figure 4.3: **January mean and standard deviation of salinity using trivariate Empirical Orthogonal Functions.** Red profiles results from monthly correlation factors on period 2019- 2020, blue monthly correlation factors on period 1993 to 2020, green daily correlation factors.

meters, where the salinity maximum occurs. For the *EMed* profile, a minimum in salinity is succeeded by a maximum, representing AW and LIW, respectively. The primary insight derived from this study is that the reconstruction produced by the MS (illustrated by the red line) closely mirrors the reanalysis profile (shown as the

black line). Notably, at each of the four points analyzed, the two reconstructions appear virtually identical.

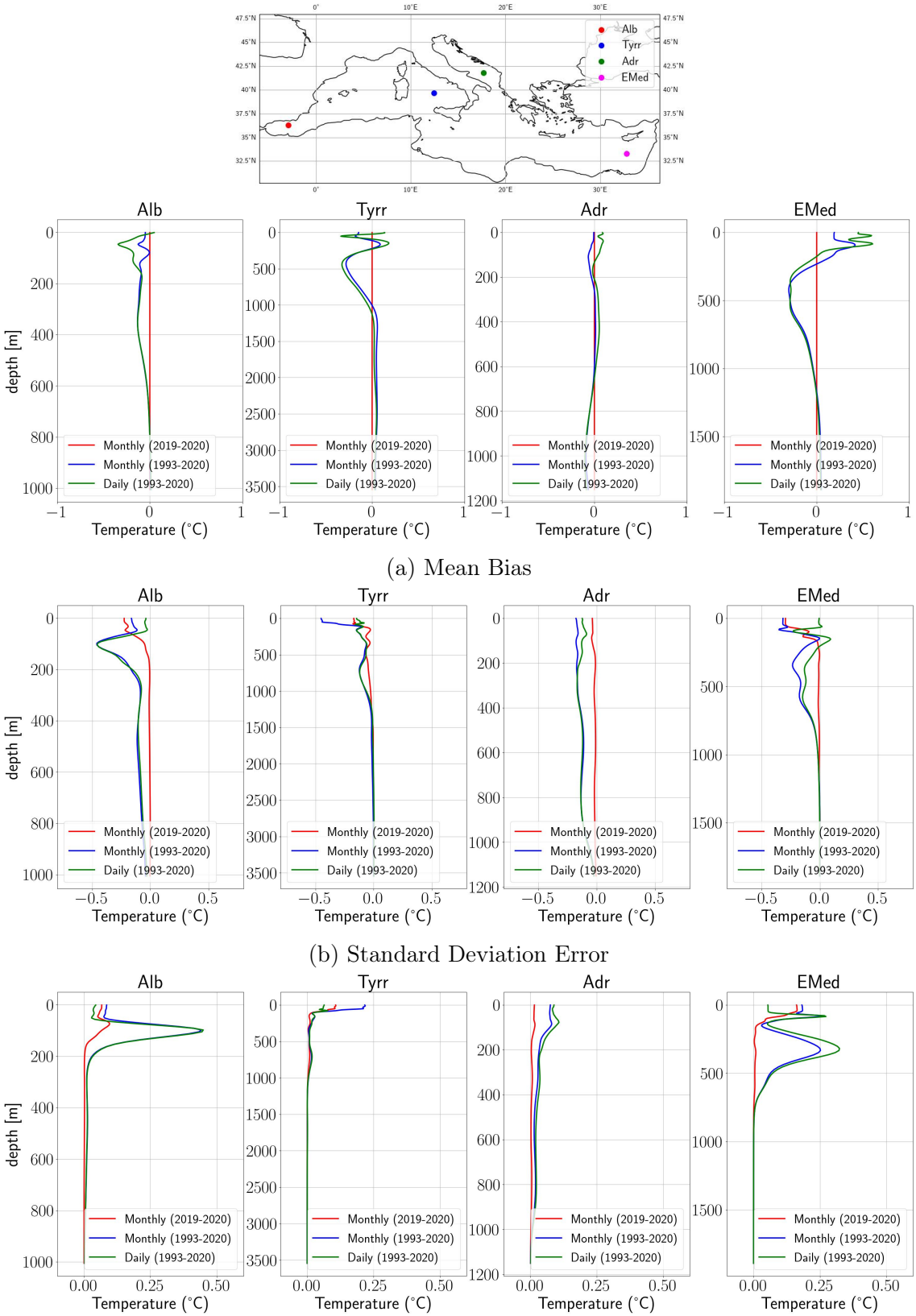
The same conclusion can be drawn concerning standard deviation. Specifically, figure 4.2b allows for a comparison of the January standard deviation across the four selected points. In all instances, the MS reconstruction of temperature standard deviation demonstrates a closer alignment with the reanalysis standard deviation, except at the surface level, where it tends to underestimate variability. At the surface, the D reconstruction for *Alb*, *Tyrr*, and *EMed* consistently outperforms, remaining very close to the reanalysis values. However, as one descends through the layers, this reconstruction increasingly underestimates variability, while the MS reconstruction of standard deviation continues to closely follow the reanalysis variability. Additionally, for temperature standard deviation in *Adr*, it is evident that the D reconstruction approaches the reanalysis more closely at the surface compared to the ML reconstruction, indicating that the application of daily correlation factors enhances performance, despite both being evaluated over the same time period.

The standard deviation of salinity, illustrated in figure 4.3b, does not exhibit the previously mentioned pattern. It is evident that the standard deviation is consistently more accurately represented by the MS method, even at the surface level. The other two time intervals fail to capture the peaks of variability observed in regions such as *Tyrr* and *EMed*. This suggests that a reconstruction strategy that prioritizes data from the specific period of interest is superior in representing the variability characteristic of that time.

The exploration of these attributes is most effectively achieved through the study of statistical errror metrics. Figures 4.4a and 4.5a illustrate that the mean bias for the MS reconstruction consistently remains significantly lower than that of the other reconstructions, approaching zero. This observation indicates that, on average, the MS reconstruction aligns more closely with the reanalysis, as previously noted. Furthermore, the other two reconstructions exhibit a high degree of similarity to one another. This suggests that the average reconstruction's accuracy may be more influenced by the time period utilized for reconstruction rather than the quantity of correlation factors employed. Notably, the two reconstructions over the extended periods D and ML display their most pronounced differences at the surface level, indicating that the application of daily factors can significantly impact the reconstruction of those layers.

This phenomenon becomes significantly more evident when examining the SDE related to temperature, as illustrated in figure 4.4b. As previously mentioned, the variability reconstructed using daily factors appears to outperform all other methods at the surface level. This observation likely stems from the fact that surface temperature variability operates on shorter temporal scales compared to deeper layers. Specifically, surface temperature fluctuations are highly influenced by atmospheric conditions such as solar radiation and wind patterns, which can change rapidly. Consequently, employing a set of correlation factors that represent each day individually allows for a more accurate capture of this rapid variability. Additionally, the graph reveals that the STD tends to diverge more significantly from the reanalysis data at the surface and within the first 500 meters of depth. This indicates that, despite the most effective reconstruction methods, a portion of the variability in this layer remains unaccounted for.

This phenomenon is clearly illustrated by the $uRMSE$ for temperature, as depicted

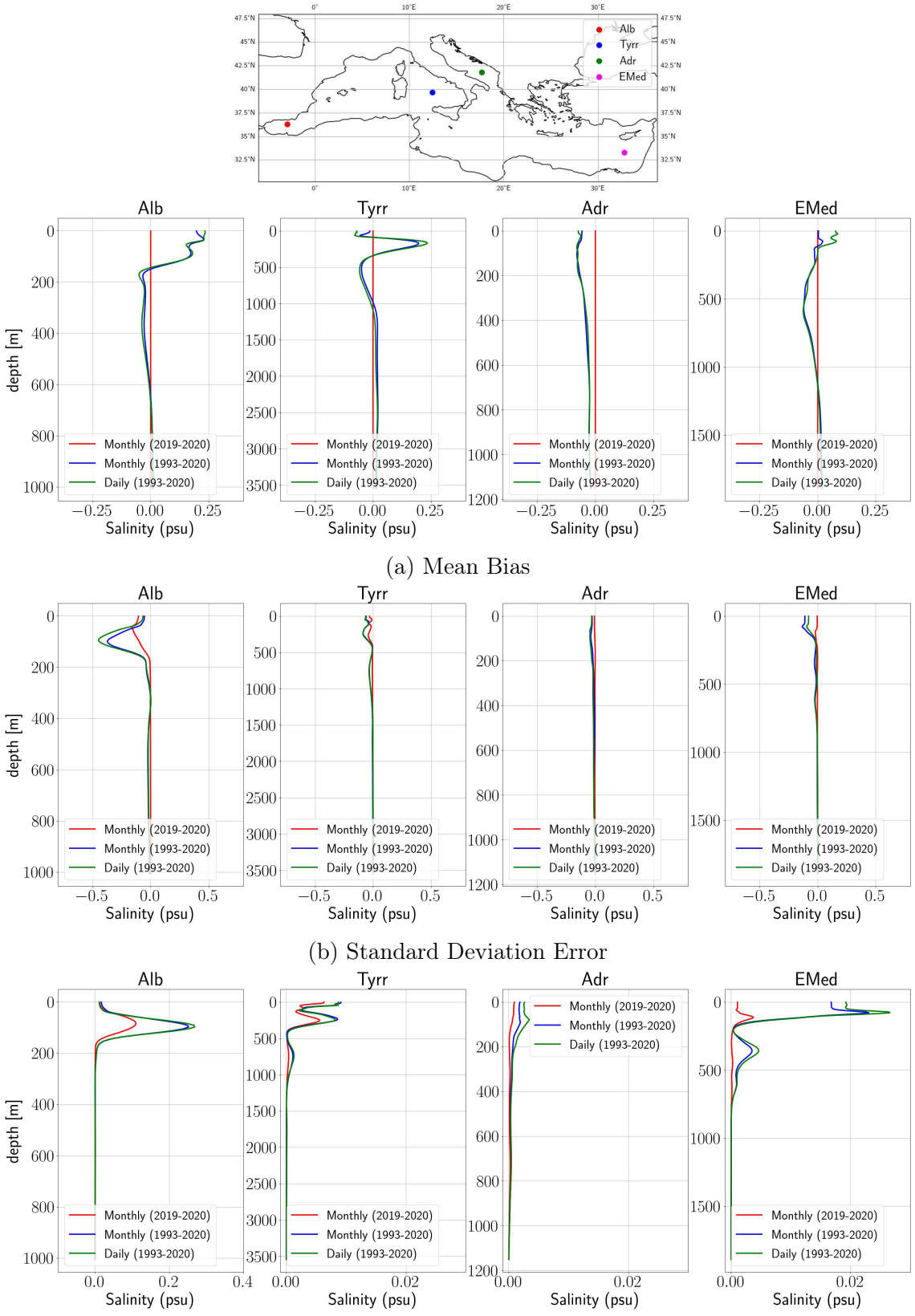


(a) Mean Bias

(b) Standard Deviation Error

(c) Unbiased Root Mean Square Error

Figure 4.4: **January error statistics of temperature reconstructions using tri-variate Empirical Orthogonal functions.** Red profiles results from monthly correlation factors on period 2019- 2020, blue monthly correlation factors on period 1993 to 2020, green daily correlation factors.



(a) Mean Bias

(b) Standard Deviation Error

(c) Unbiased Root Mean Square Error

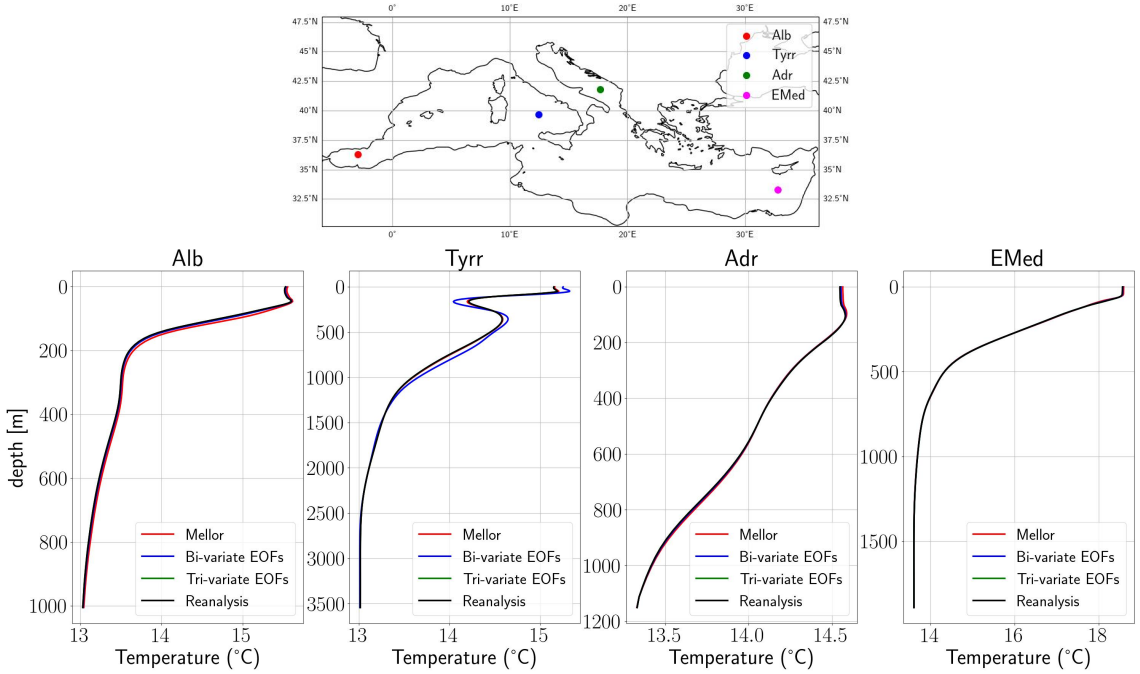
Figure 4.5: **January error statistics of salinity reconstructions using trivariate Empirical Orthogonal Functions.** Red profiles results from monthly correlation factors on period 2019- 2020, blue monthly correlation factors on period 1993 to 2020, green daily correlation factors.

in figure 4.4c. Notably, the values remain consistently low; however, the highest values are observed at the surface, even in the context of the most accurate reconstructions. Similarly to mean bias and standard deviation error, the unbiased root mean square error is minimized in the MS reconstruction. For instance, this reconstruction significantly decreases the error in the subsurface layers of *Alb* and *EMed*. The sole exception occurs, as stated before, within the upper 50 meters or less for *Alb*, *Tyrr*, and *EMed*.
The metrics SDE and $uRMSE$ for salinity are both lower than those for temperature, as depicted in figures 4.5b and 4.5c. However, these error metrics do not demonstrate the beneficial characteristics of D reconstruction at the surface. Notably, both the SDE and $uRMSE$ values are superior for the MS reconstruction compared to other methods. This discrepancy may be attributed to the influence of processes affecting sea surface salinity, such as river runoff, which typically operate over extended timescales, with seasonal variability being a predominant factor [79].
From the statistical analysis, it is evident that the reconstruction utilizing monthly correlation factors, assessed over a timeframe closely aligned with the target reconstruction period, yields the most favorable results across all metrics. The sole exception pertains to the surface layer of the reconstructed temperature, where daily variability prevails, indicating that the application of daily correlation factors is more effective in capturing this specific variability.

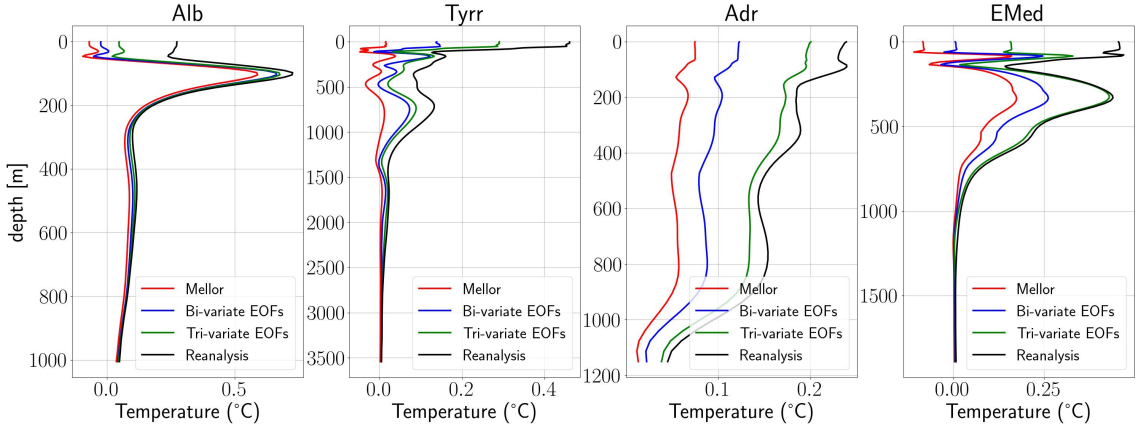
4.1.4 Comparison between methods

This section will provide a comparative analysis of the three distinct methods. For this purpose, we have selected the profiles reconstructed using the monthly correlation factors from the period 2019-2020, which represents the most accurate reconstruction, as indicated in 4.1.3. Henceforth, the reconstructions derived from the Ezer and Mellor method will be referred to as EM, those obtained through bi-variate EOFs will be labeled BE, and the reconstructions achieved via tri-variate EOFs will be designated as TE.
The analysis presented in figures 4.6a and 4.7a indicates that the January mean reconstruction closely aligns with the reanalysis across all three methodologies. The sole exception is observed at the *Tyrr* sites, where the mean reconstruction from BE exhibits a slight divergence from the reanalysis. This phenomenon can be explained by the foundational principles of our methodology, which seeks to infer temperature and salinity profiles from sea level data. The relationship is predicated on the notion that sea level serves as an integral of density, which is a function of both temperature and salinity 2.1.1. Generally, one of these variables tends to dominate; however, in cases of pronounced stratification, both variables can significantly influence density. Therefore, relying solely on one variable for the calculation of Empirical Orthogonal Functions (EOFs) may result in incomplete representation of sea level. The incorporation of both variables in a tri-variate framework resolves this issue.
The results of the standard deviation reconstructions are illustrated in figures 4.6b and 4.7b. Generally, the TE reconstruction demonstrates superior performance, often aligning so closely with the reanalysis data that it becomes indistinguishable, particularly in the case of *EMed*. Conversely, the Ezer and Mellor methods consistently exhibit a tendency to underestimate the standard deviation, a trend that is particularly evident in figures 4.8b and 4.9b. It is evident that all methods generally

underestimate the standard deviation, especially at the surface level. Among them, the TE reconstruction stands out as the most effective, achieving nearly perfect reconstruction of the standard deviation in depth while exhibiting a slight underestimation at the surface.



(a) Mean

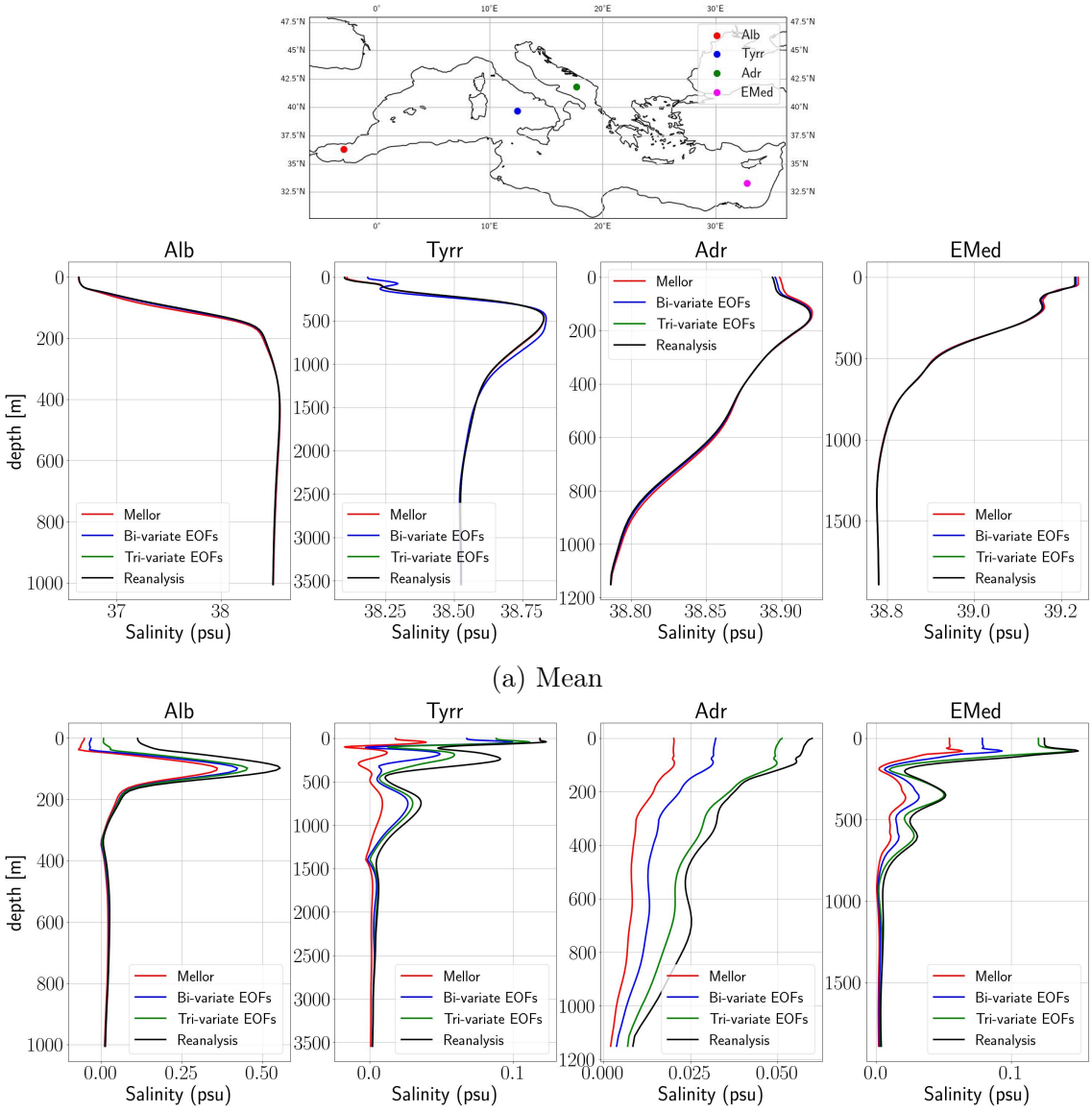


(b) Standard Deviation

Figure 4.6: **January mean and standard deviation of temperature using monthly correlation factors with mean on period 2019-2020.** Red profiles have been obtained using Mellor and Ezer plain method. Blue profiles have been obtained using bi-variate EOFs method. Green profiles have been obtained using tri-variate EOFs method.

A deeper analysis of the mean can be observed studying mean bias in figures 4.8a and 4.9a. Initially, the reconstructions appear nearly identical; however, a deeper investigation uncovers notable differences. The TE reconstruction shows a mean bias that is nearly zero, a characteristic not shared by the other two methods. In particular, the Ezer and Mellor techniques tend to yield the least favorable results. This is evident across all four points, especially in the surface and subsurface layers,

and is also apparent at the bottom for *Alb* and *Adr*. While the differences are not particularly large, they are distinctly observable. Conversely, *Tyrr* stands out as an exception, with the BE method exhibiting the worst mean reconstruction.

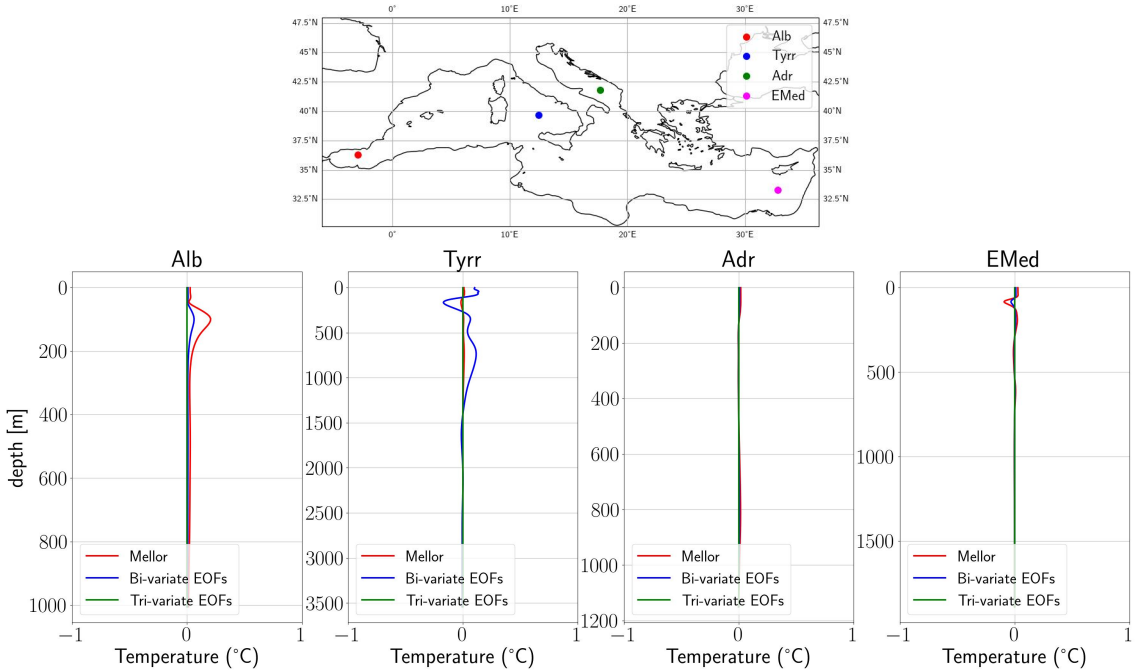


(a) Mean

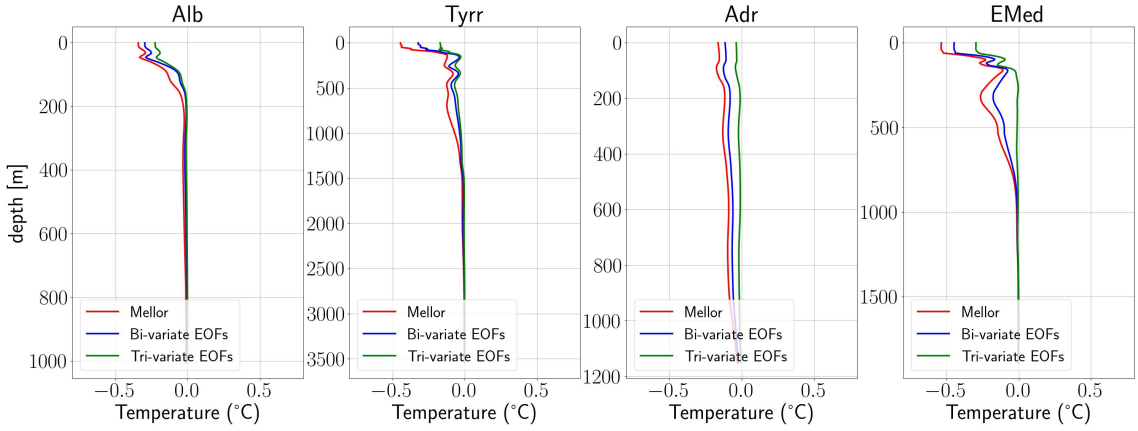
(b) Standard Deviation

Figure 4.7: **January mean and standard deviation of salinity using monthly correlation factors with mean on period 2019-2020.** Red profiles have been obtained using Mellor and Ezer plain method. Blue profiles have been obtained using bi-variate EOFs method. Green profiles have been obtained using tri-variate EOFs method.

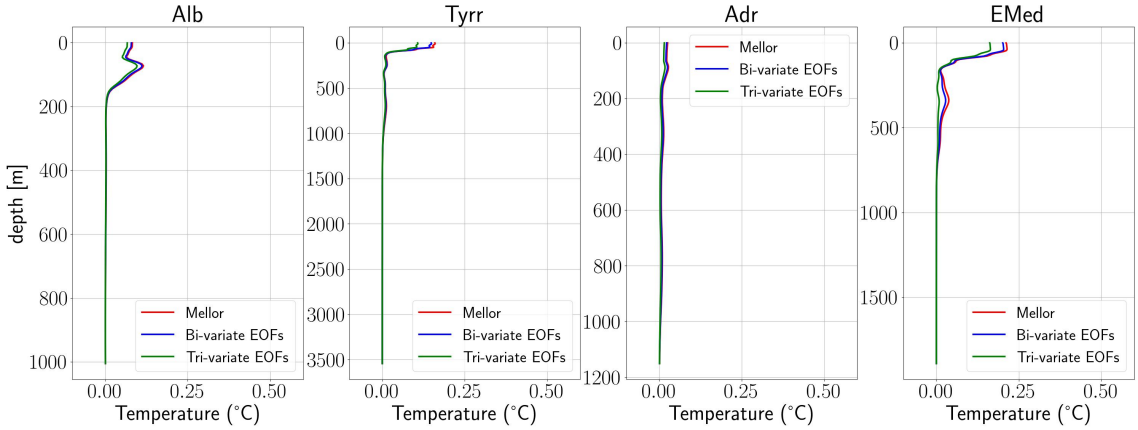
The outstanding efficacy of TE reconstruction is clearly demonstrated when examining the $uRMSE$. Figures 4.8c and 4.9c reveal that the error associated with all methods is predominantly low. In fact, it is nearly zero, with maximum values of 0.1 °C and 0.015 psu observed in the subsurface layers. At deeper levels, the error approaches insignificance. Nonetheless, TE consistently exhibits the lowest $uRMSE$, highlighting its exceptional performance relative to the other methods.



(a) Mean Bias

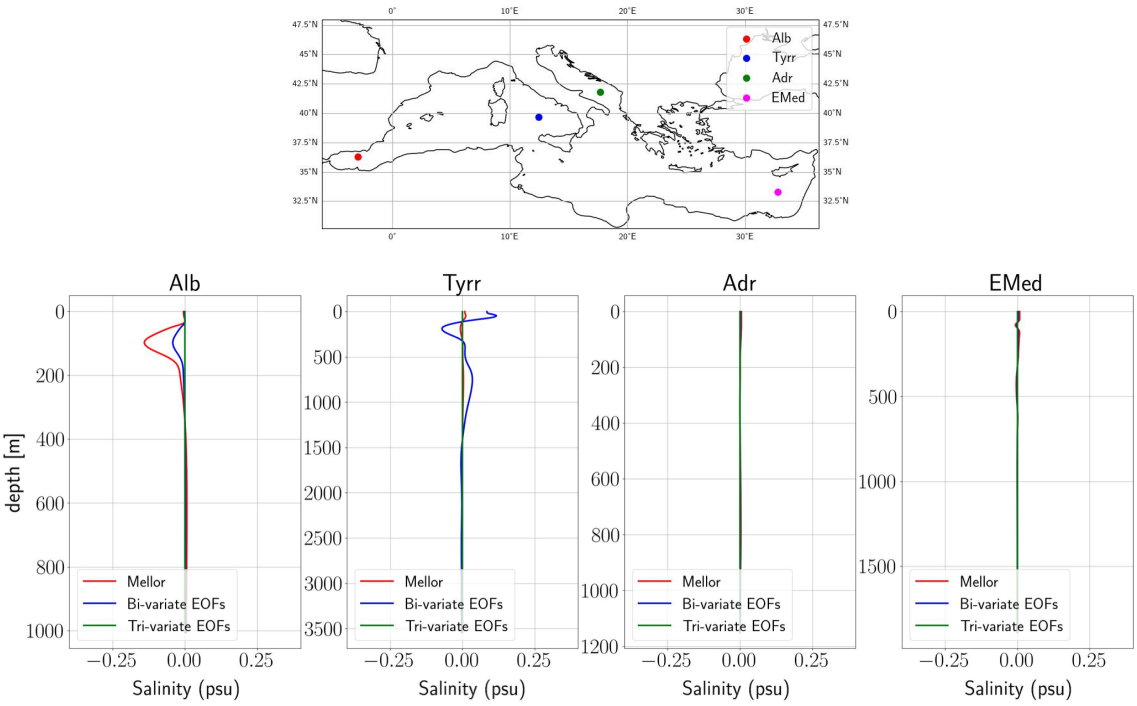


(b) Standard Deviation Error

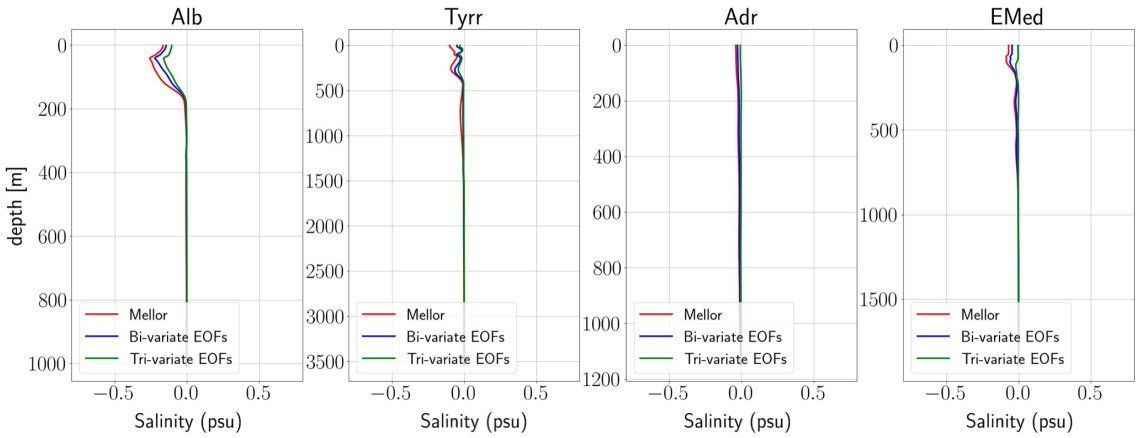


(c) Unbiased Root Mean Square Error

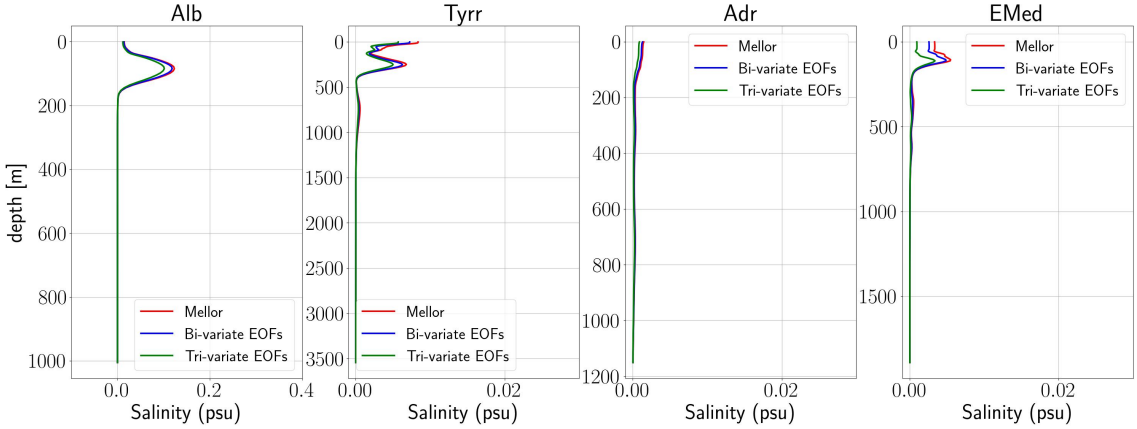
Figure 4.8: **January error statistics of temperature using monthly correlation factors with mean on period 2019-2020.** Red profiles have been obtained using Mellor and Ezer plain method. Blue profiles have been obtained using bi-variate EOFs method. Green profiles have been obtained using tri-variate EOFs method.



(a) Mean Bias



(b) Standard Deviation Error



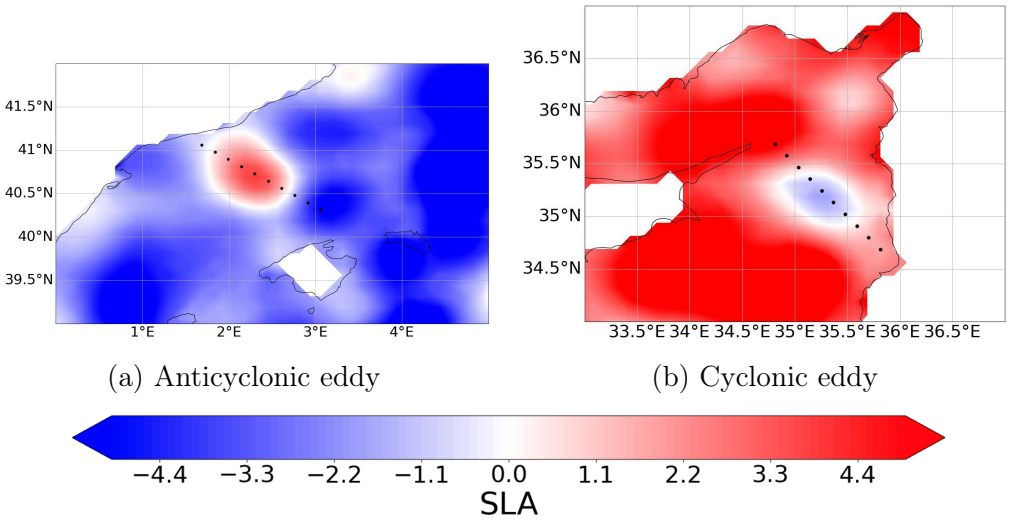
(c) Unbiased Root Mean Square Error

Figure 4.9: **January error statistics of salinity using monthly correlation factors with mean on period 2019-2020.** Red profiles have been obtained using Mellor and Ezer plain method. Blue profiles have been obtained using bi-variate EOFs method. Green profiles have been obtained using tri-variate EOFs method.

4.2 Eddies reconstruction

As emphasized in section 1.3, oceanic mesoscale characteristics can be equated to the "weather" of the marine environment. This analogy suggests that salinity and temperature characteristics fluctuate daily due to these features. Consequently, investigating their reconstruction is crucial for evaluating the efficacy of our methodologies. A significant aspect of this analysis will be the focus on comparing the performance of various reconstruction methods when applied to the reconstruction of a single day, as opposed to averaging over multiple profiles as done before. Furthermore, as elaborated in section 1.3.3, the vertical configuration of eddies is distinctly characterized, allowing us to assess the accuracy of our algorithms in reconstructing this peculiar structure. All of these components will facilitate our ability to determine the optimal reconstruction approach from the three presented.
To locate eddies for investigation, altimetry data serves as a valuable resource, given that the sea surface experiences upward or downward deflections corresponding to anticyclonic and cyclonic eddies, respectively as underlined in section 1.3.3. Consequently a basic technique for their detection is the examination of sea level anomaly images. However, this method presents some challenges. Specifically, the resolution constraints of satellite altimetry products can pose significant limitations. When the deformation radius is small, the resolution may be insufficient to capture the complete spectrum of mesoscale dynamics. This issue is particularly pronounced in the Mediterranean Sea, where the deformation radius is notably low, potentially complicating the study of these eddies [39, 31].
Despite the existing challenges, two transient eddies have been identified and analyzed, as illustrated in 4.10. These cases of study include:

- An anticyclonic eddy, approximately 100 km in radius, located off the coast of Spain on 24 March 2019;
- A cyclonic eddy, with a radius of around 75 km, situated near Cyprus on 30 May 2020.



(a) Anticyclonic eddy (b) Cyclonic eddy

Figure 4.10: Eddies under study. The anticyclonic eddy is found near Spain Coast on 24 March 2019, while the cyclonic one near Cyprus on 30 May 2020. Black dots represents points taken into consideration.

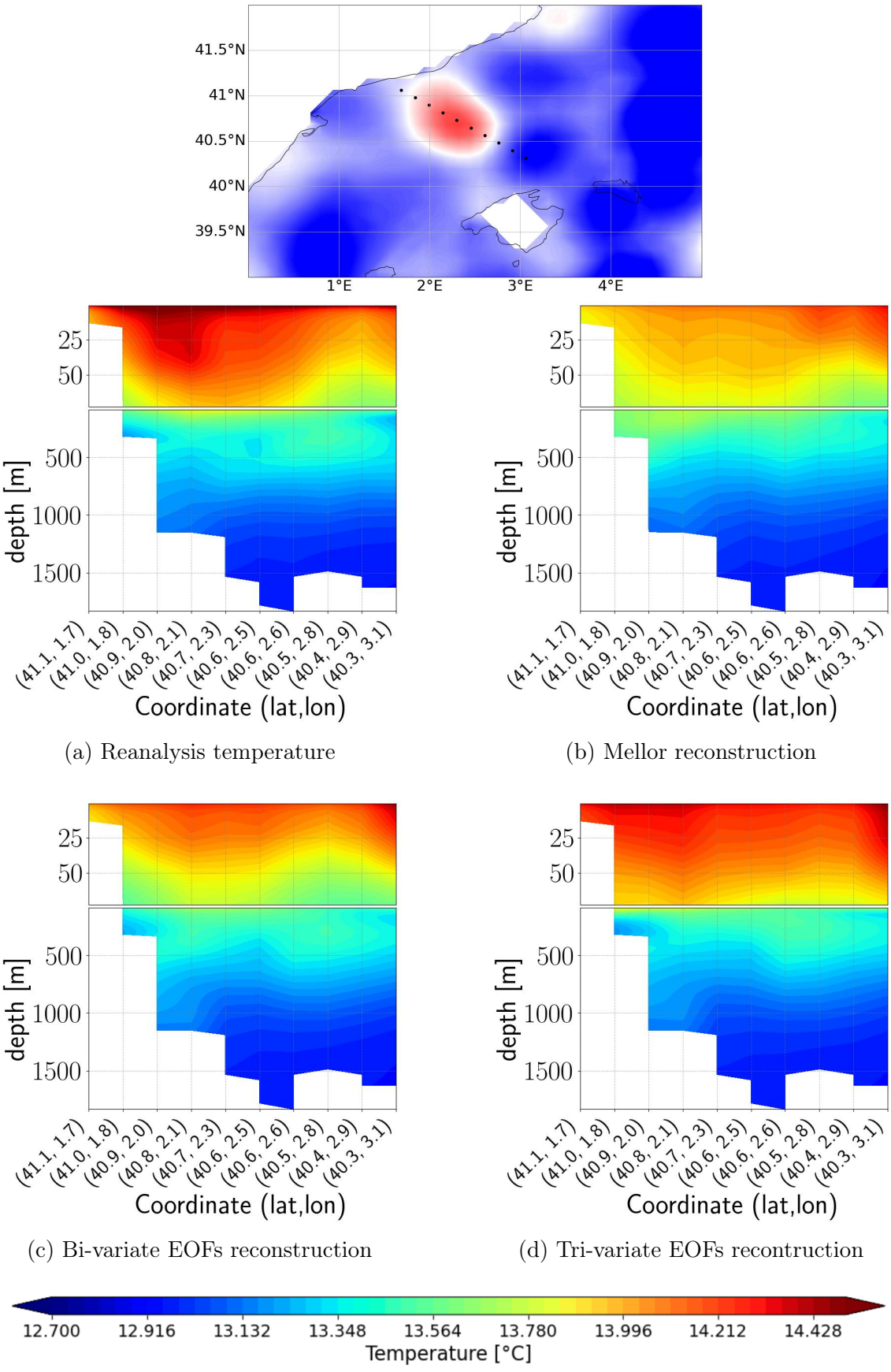
Both of these eddies have radii that surpass the deformation radius characteristic of the Mediterranean Sea. While smaller eddies are present, their identification is limited due to the low resolution of altimetry data. The contrasting characteristics of these eddies are particularly valuable, as they allow for a thorough investigation into the various properties of eddy dynamics.
In this investigation into the reconstruction of eddies, we will illustrate the precise reconstruction of the eddy for the selected day of analysis. Furthermore, we will evaluate the percentage discrepancies between the reconstructed field and the reanalysis field, which can be expressed mathematically as

$$\Delta = \frac{r - R}{R} * 100 \tag{4.4}$$

with r indicating the reconstructed field and R signifying the reanalysis field.

4.2.1 Anticyclonic eddy near Spain

The first eddy exhibits an anticyclonic nature, which is prominently indicated by a positive sea level anomaly that is surrounded by negative anomalies. This characteristic is clearly illustrated in the reanalysis plots of temperature and salinity, as shown in 4.11a and 4.12a. A notable feature is the depression observed in the isolines of temperature and salinity, suggesting that warmer and fresher surface waters are being displaced to deeper layers. The influence of the eddy extends slightly beyond its radius, though the extent is minimal. Additionally, it displays a typical structure where the radius at depth is smaller than that at the surface. The salinity field further indicates two regions of fresher water at the surface, which are distinctly marked by two minima.
The three temperature reconstructions are illustrated in figure 4.11, with their corresponding percentage differences from the reanalysis depicted in figures 4.13a, 4.13b, and 4.13c. Both the Mellor and Ezer reconstruction and the bivariate EOFs exhibit a similar pattern of curved isolines in the same regions identified in the reanalysis. However, the Mellor and Ezer reconstruction appears to underestimate temperature values. Additionally, the depth of displacement in these reconstructions is less pronounced compared to the reanalysis. In contrast, the trivariate EOFs reconstruction displays a markedly different profile. It reveals a less distinct deepening of isolines, which, while present, lacks smoothness. Nevertheless, this method achieves a greater depth of displacement and reconstructed temperature values that are more aligned with the original data. Notably, an evident warm pool is identified in the same region as the reanalysis, specifically between 41 and 40.6°N, a feature absent in the other two reconstructions. As evidenced by the difference figures, these characteristics contribute to a more accurate temperature estimate for the trivariate EOF method in the central domain, indicating that while it may not effectively reconstruct the isoline shapes, it excels in capturing the temperature field values.
The reconstructions of salinity are illustrated in figure 4.12, while the percentage differences from reanalysis are depicted in figures 4.13d, 4.13e, and 4.13f. Notably, all reconstructions indicate a tendency to overestimate surface water salinity compared to actual values. Additionally, none of the reconstructions adequately represent the left pole at the surface. In the TE reconstruction, the right pole appears to be accurately positioned at the surface, however the fresh water bulb penetrates to



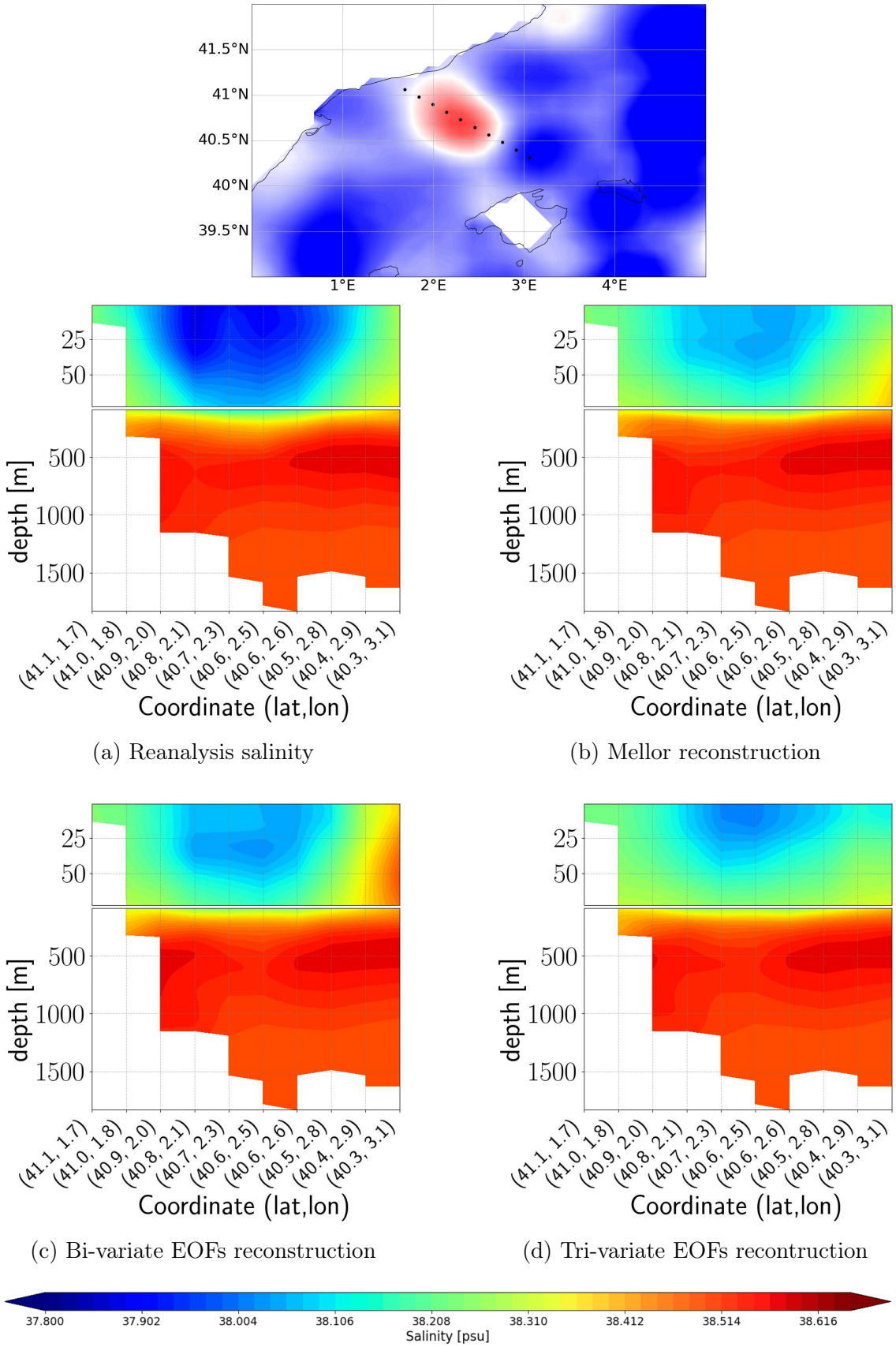
(a) Reanalysis temperature

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 4.11: Temperature of the transect crossing the eddy near Spain coast on 24 March 2019



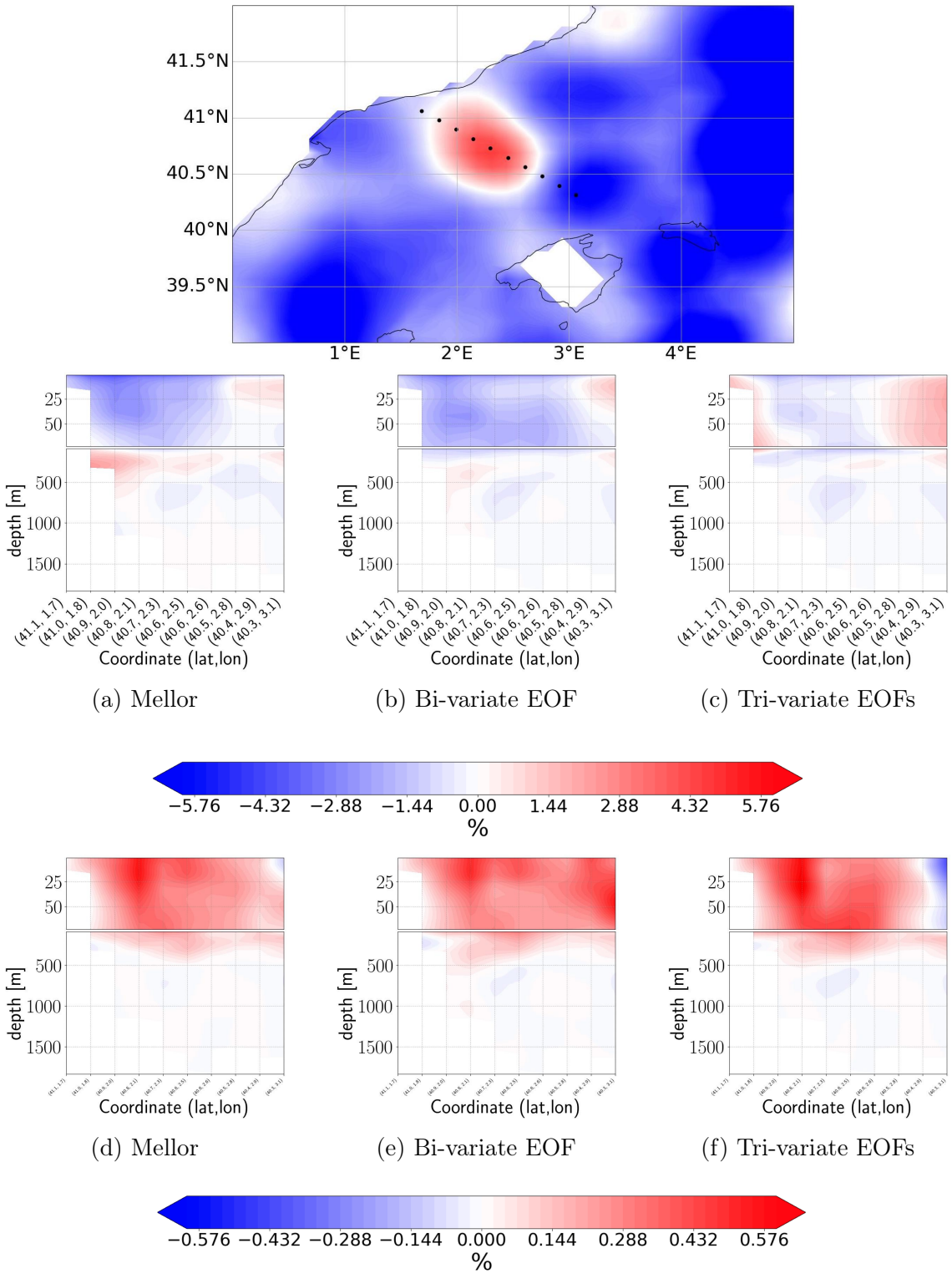
(a) Reanalysis salinity

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 4.12: Salinity of the transect crossing the eddy near Spain coast on 24 March 2019



(a) Mellor (b) Bi-variate EOF (c) Tri-variate EOFs

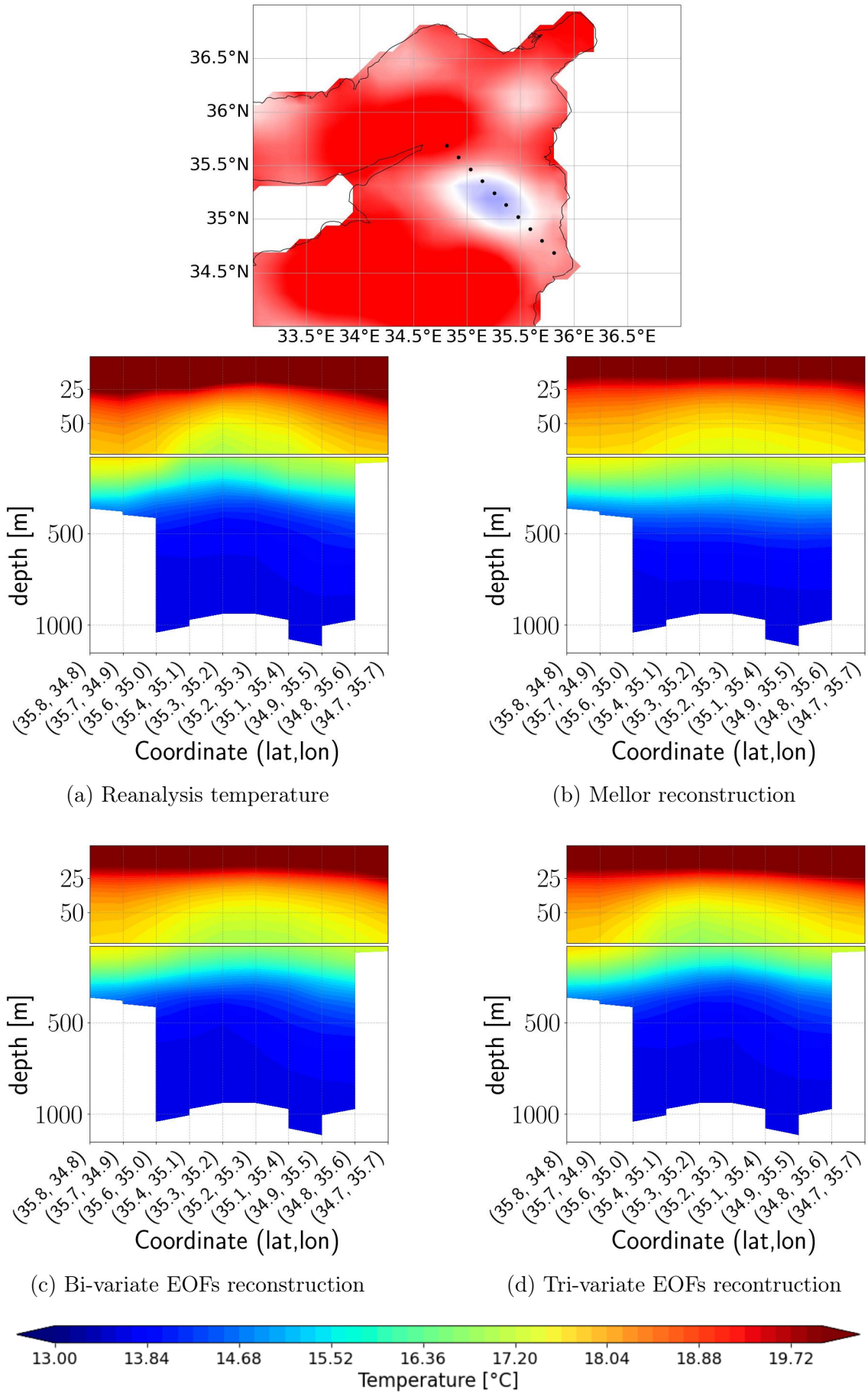
(d) Mellor (e) Bi-variate EOF (f) Tri-variate EOFs

Figure 4.13: **Percentage differences between reanalysis and reconstructions.** 4.13a, 4.13b and 4.13c are relative to temperature reconstructions, while 4.13d, 4.13e and 4.13f are relative to salinity reconstructions

a significantly lesser depth than the other two reconstructions, and its lateral extension on the right is limited. For the BE reconstruction, there is a pronounced overestimation of salinity on the right side of the eddies, resulting in much saltier water than what is indicated by reanalysis. This observation aligns with the understanding that salinity exhibits a weaker correlation with sea level than temperature [80], suggesting that the correlation factors alone are insufficient to capture all the desired characteristics.

4.2.2 Cyclonic eddy near Cyprus

The second eddy is characterized as a cyclonic type, which is evident from the negative sea level anomaly that is encircled by positive anomalies. This cyclonic nature is clearly illustrated in the reanalysis plots of temperature and salinity, as shown in figures 4.14a and 4.15a. A prominent dome is observable in the isolines of temperature and salinity, indicating the elevation of colder water from the bottom. The properties of this eddy extend well beyond its radius, presenting a more cylindrical profile when compared to its anticyclonic counterpart.
The temperature reconstructions are illustrated in figure 4.14, while the percentage differences from reanalysis are depicted in figures 4.16a, 4.16b, and 4.16c. All reconstructions effectively elevate the temperature isolines. Notably, the two EOF reconstructions also enhance the depth isolines, whereas the Mellor and Ezer reconstructions exhibit flatter depth isolines. Additionally, the subsurface isolines are not elevated to the same extent as observed in the other two reconstructions. This observation holds true for the bi-variate EOF reconstruction as well, with the TE reconstruction being the only one that appears to replicate a peak in the isolines towards the right side of the domain. This is particularly evident when examining the differences. Specifically, the EM reconstruction tends to overestimate the temperature in the central region of the domain, indicating insufficient cold water is being brought up from the depths. The other two reconstructions display similar behavior, with the TE reconstruction performing slightly better on the right side of the domain. Furthermore, a previously unrecognized characteristic is now apparent: the surface temperature is consistently underestimated, with the tri-variate reconstruction exhibiting the least degree of underestimation.
Figure 4.15 presents the salinity reconstructions, while figures 4.16d, 4.16e, and 4.16f illustrate the percentage differences from reanalysis. A clear dome of isolines is evident across all reconstructions. Additionally, a saltier region is observed at the surface above the core of the dome, slightly shifted to the right of the center. This displacement is uniquely observed in the TE reconstruction. Moreover, both EM and BE reconstructions tend to broaden the cylinder of saltier water in the domain's center, whereas the TE reconstruction appears to narrow it. This trend is also reflected in the percentage differences, with the TE reconstruction underestimating salinity at the domain's periphery, while the EM and BE reconstructions show an overestimation.



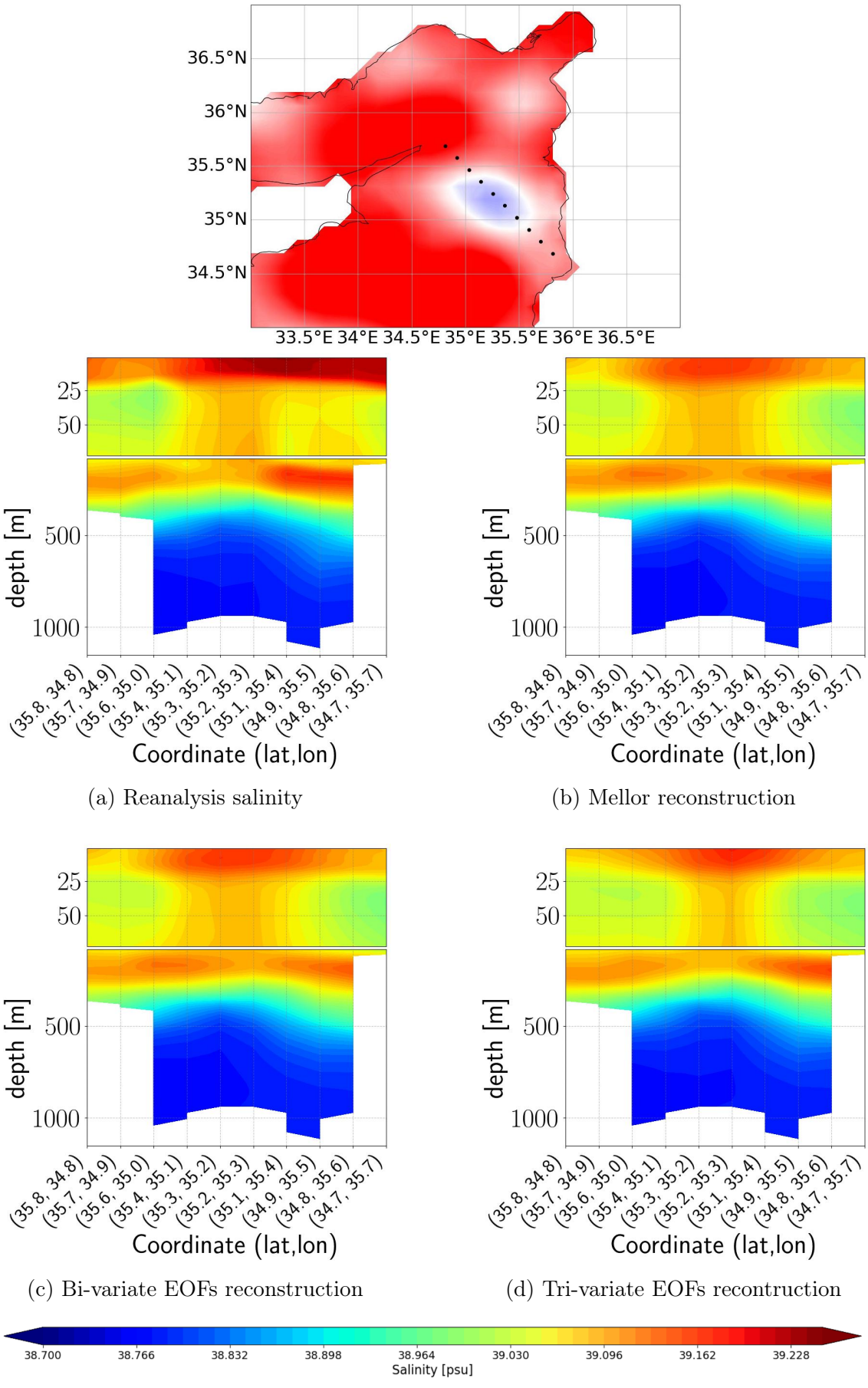
(a) Reanalysis temperature

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 4.14: Temperature of the transect crossing the eddy near Cyprus on 30 May 2020



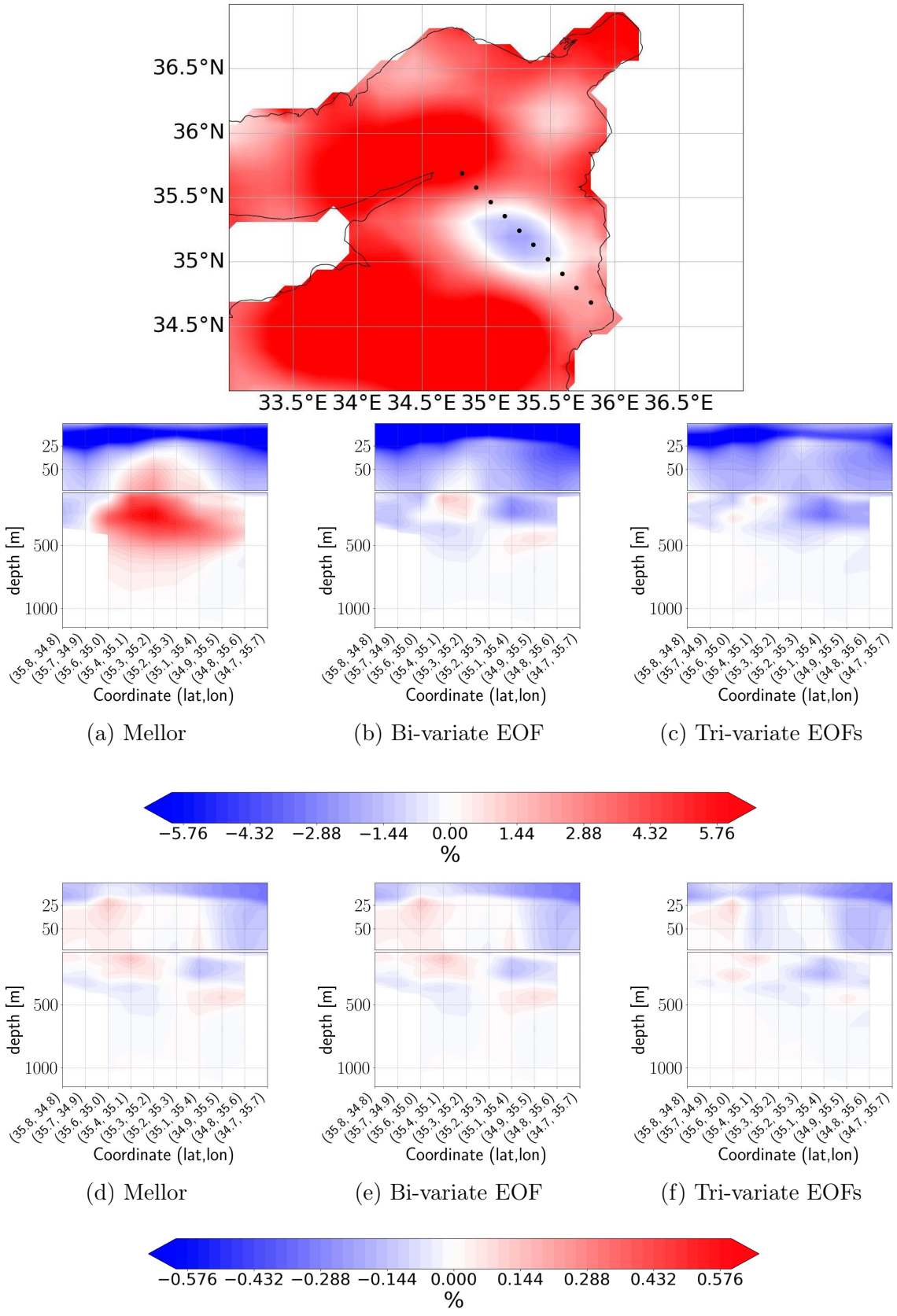
(a) Reanalysis salinity

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 4.15: Salinity of the transect crossing the eddy near Cyprus on 30 May 2020



(a) Mellor (b) Bi-variate EOF (c) Tri-variate EOFs

(d) Mellor (e) Bi-variate EOF (f) Tri-variate EOFs

Figure 4.16: **Percentage differences between reanalysis and reconstructions.** 4.16a, 4.16b and 4.16c are relative to temperature reconstructions, while 4.16d, 4.16e and 4.16f are relative to salinity reconstructions

Chapter 5

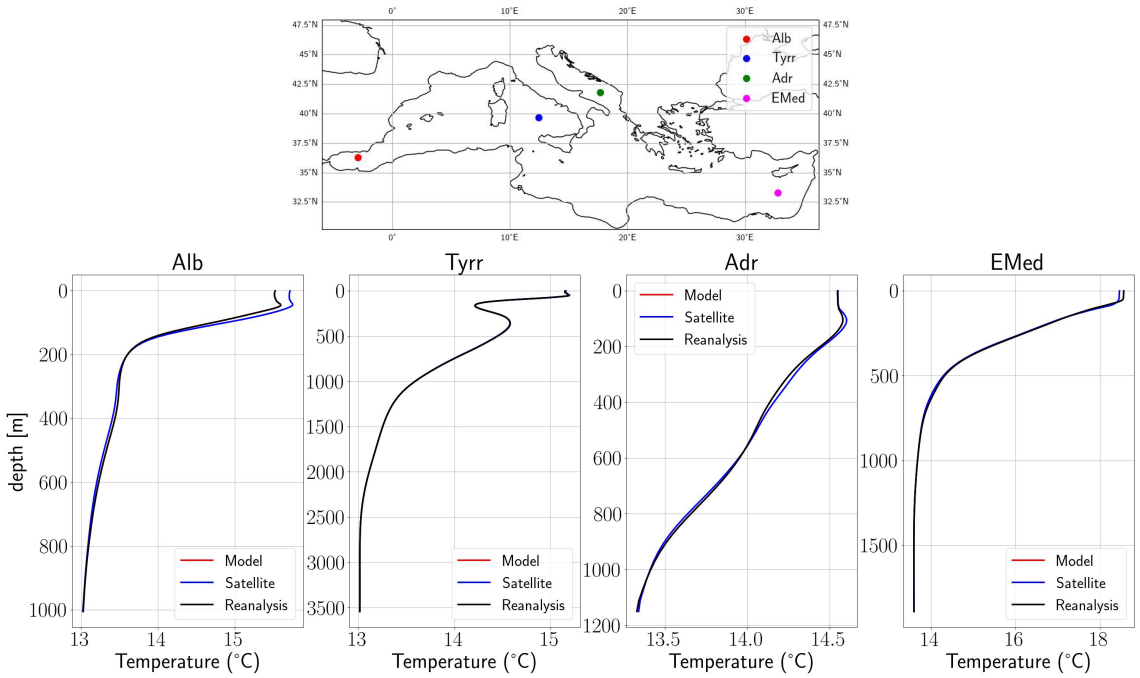
Reconstructions using satellite data

In Chapter 4, we examined the quality of the subsurface reconstruction of temperature and salinity profiles with the methods developed in the previous sections. We will now use sea level anomaly data obtained from satellite altimetry, L4 product. Consistent with the previous chapter, Section 5.1 will detail the reconstruction of temperature and salinity profiles, offering a comparative analysis between reconstructions based on reanalysis and those derived from L4 satellite data. Section 5.2 will focus on the reconstruction of the vertical structures of eddies.

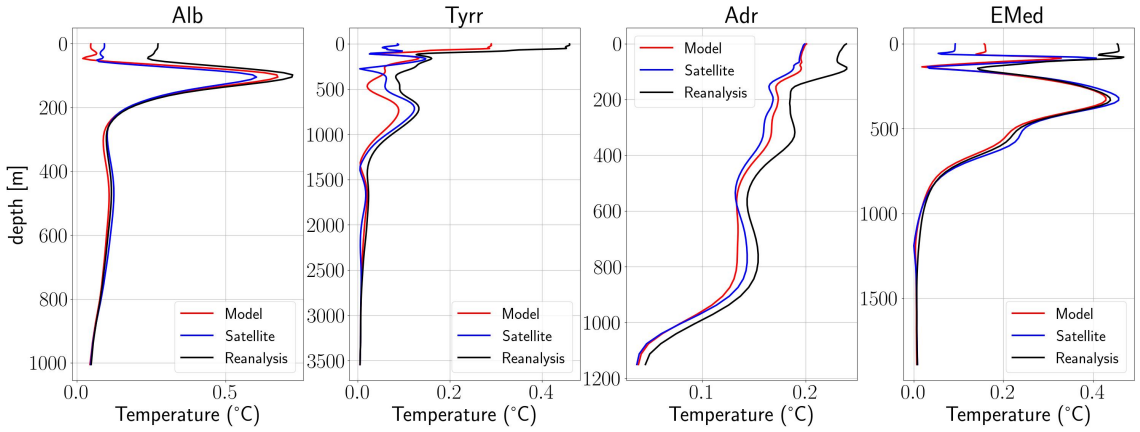
5.1 Profile reconstructions

In section 4.1, we emphasized that the optimal reconstruction is achieved through the application of monthly correlation factors derived from tri-variate EOFs, utilizing data from the period spanning 2019 to 2020 for the averages. It is important to note throughout this discussion that the salinity reconstructions yield results identical to those of temperature reconstructions and, therefore, are not presented separately.
The choice to employ satellite-derived sea level anomaly (SLA) data in place of reanalysis data can introduce a level of inconsistency into our reconstructions. In fact, reanalysis assimilates along track altimetry data and offers its own reconstruction of sea level anomaly, which differs from the process used to obtain L4 satellite data. The latter is generated through an objective analysis of various satellite altimeters, as discussed in section 2.1.2. Therefore, the correlation factors obtained using the same algorithms as those in the previous chapter, when applied to L4 satellite data, could lead to a different subsurface thermal and haline structure, due to these different sea level estimates. The extrapolation of this structure from an altimetric product and the reanalysis is analysed in this chapter.
Figure 5.1a presents a comparison of the two reconstructions for mean temperature, while figure 5.2a illustrates the mean bias. It is evident that the reconstruction from L4 data is different from the reanalysis one, but it is consistent with it. The discrepancy between the reanalysis and the satellite reconstruction is particularly noticeable at the surface, where the bias may reach as high as 0.2°C in the Alboran Sea. This is probably due to the fact that the upper layer temperature and salinity is the least correlated part of the subsurface thermohaline structure to the sea level. Thus the satellite altimetry signal can be very different from the reanalysis one where the upper layer structures are corrected also by the numerical model dynamics. Another reason for the difference is probably connected to the number of vertical

modes used in the reconstruction that is relative low (25 for the trivariate EOFs). The reconstruction of the standard deviation presents a more complex scenario, as illustrated in figures 5.1b and 5.2b. There is considerable variability across the different study sites, with the most significant differences typically occurring in the surface and upper ocean layers. An exception to this trend is the region identified



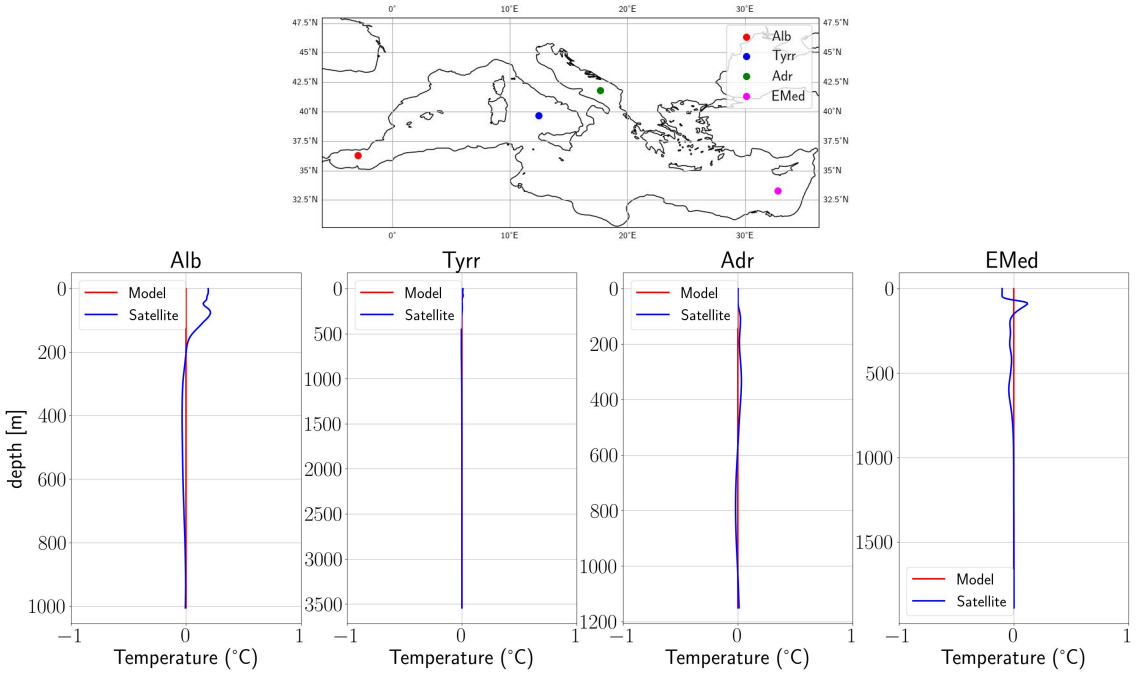
(a) Mean



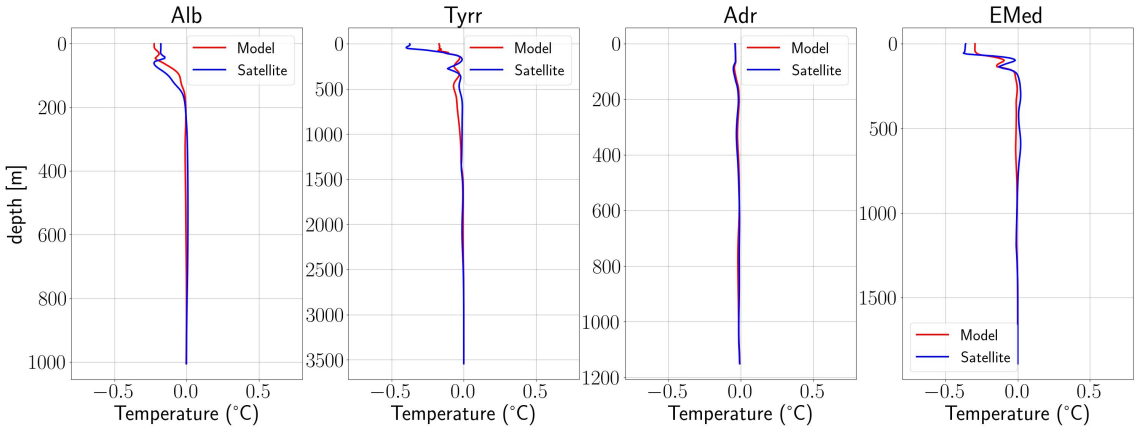
(b) Standard Deviation

Figure 5.1: **January mean and standard deviation of temperature using SLA from reanalysis and from satellite.** Red profiles results from monthly correlation factors on period 2019- 2020 obtained with SLA from reanalysis, blue profiles from monthly correlation factors on period 2019 to 2020 obtained with SLA from satellite.

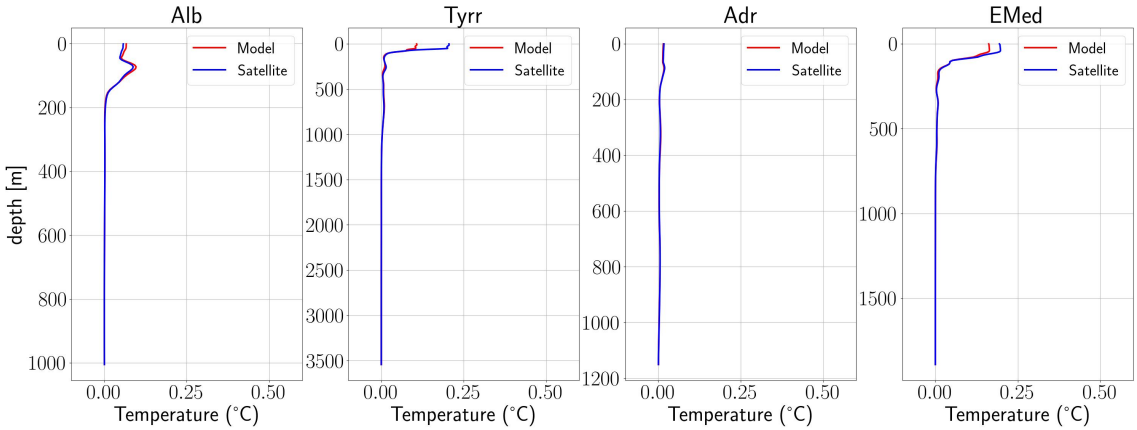
as *Alb*. Given its strong interactions with the Atlantic Ocean, the dynamics in this area pose considerable modeling challenges, suggesting that satellite data may provide a more accurate representation of actual conditions.
A comparable challenge in interpretation arises when examining the $uRMSE$, as illustrated in figures 5.2c. Notably, this metric exhibits significant variability from



(a) Mean Bias



(b) Standard Deviation Error



(c) Unbiased Root Mean Square Error

Figure 5.2: **January error statistics of temperature using SLA from reanalysis and from satellite.** Red profiles results from monthly correlation factors on period 2019- 2020 obtained with SLA from reanalysis, blue profiles from monthly correlation factors on period 2019 to 2020 obtained with SLA from satellite.

one point to another. However overall, the reconstructions form satellite and reanalysis estimates are consistent using the same extrapolation method and the $uRMSE$ is in both cases very small.

5.2 Eddy reconstruction

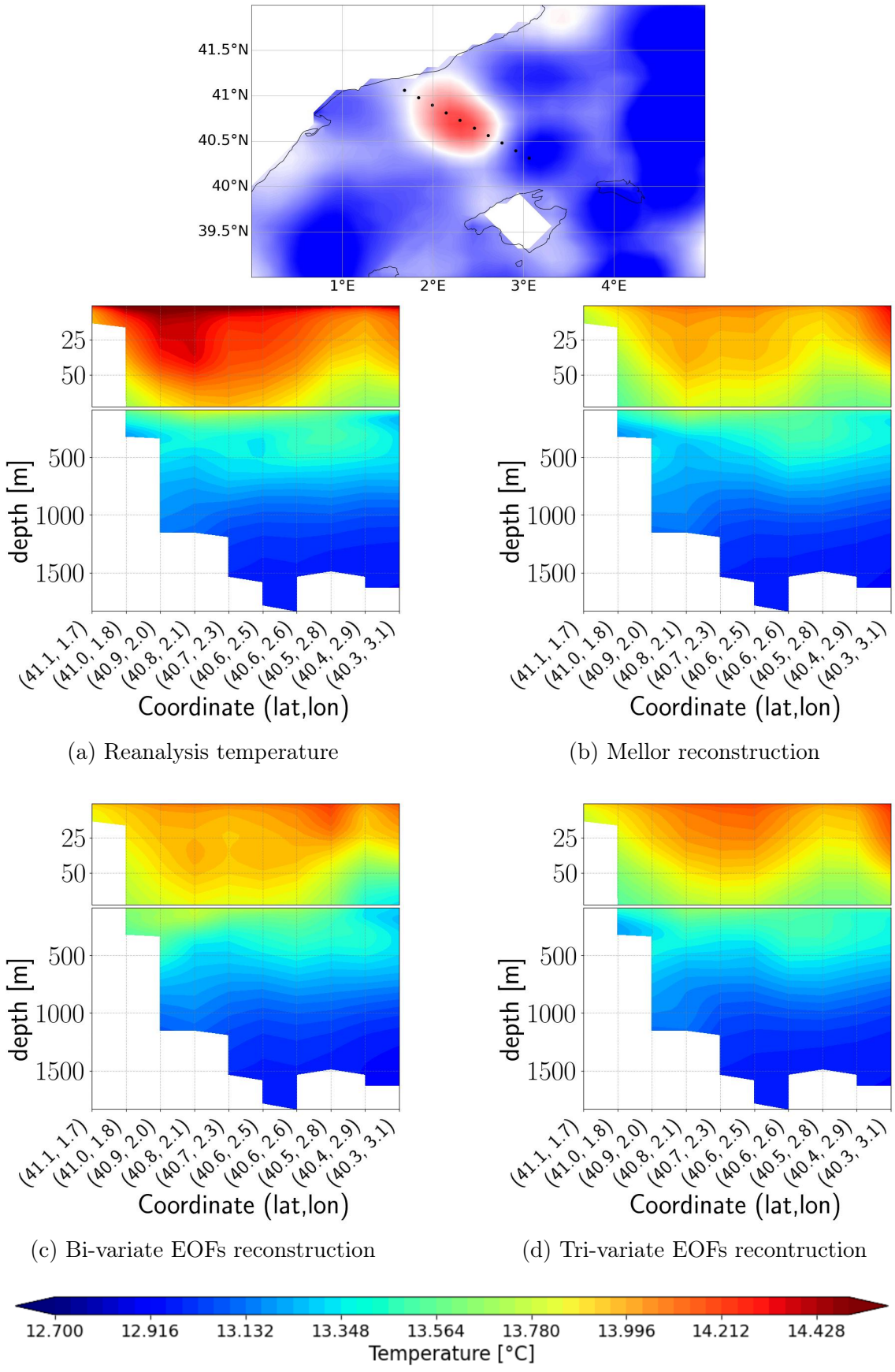
The instantaneous reconstruction of eddies from altimetric L4 products is a challenging task. Notably, both temperature and salinity reanalysis reconstructions reveal a range of advantages and disadvantages associated with different extrapolation techniques. We will now extend the investigation by utilizing satellite L4 sea level anomalies.
This analysis will illustrate the reconstruction of the eddy on the designated day and the percentage discrepancies between the reconstructed field based on satellite SLA and the field reconstructed using reanalysis SLA. The calculation for this difference is expressed as follows:

$$\Delta = \frac{r_s - r_m}{r_m} * 100 \tag{5.1}$$

where r_m is the reanalysis reconstructed field and r_s the satellite reconstructed field.

5.2.1 Anticyclonic eddy near Spain

The three temperature reconstructions for the anticyclonic eddies reveal differences between the three methods that were not evident before. Each reconstruction distinctly illustrates the deepening of isotherms. Nevertheless, the reconstructed temperatures are notably lower, particularly at the center of the eddy. Ezer and Mellor's reconstruction, illustrated in figure 5.3b, appears to have a more limited horizontal and vertical extent. Specifically, the right side of the eddy demonstrates a shallower profile. This observation is further supported by the differences, as depicted in figure 5.5a. The findings indicate that satellite altimetry possibly consists of a weaker eddy, reduced in horizontal extent.
In the context of bi-variate EOF (BE) reconstruction, as illustrated in figure 5.3c, a similar behaviour is observed, characterized by a lower amplitude temperature field. However, the eddy appears to have deepened in this reconstruction. This finding is further validated by the differences illustrated in 5.5b. It is also apparent that the reconstructed temperature is declining on the right side of the eddy, suggesting that, similar to the Mellor and Ezer's reconstruction method, the eddy size is smaller. The reduction in the eddy's dimensions is particularly clear in the tri-variate temperature reconstruction (ME) shown in figure 5.3d, where a general temperature decrease is also observed. This decline is especially marked by the vanishing of the warm pool located on the left side of the eddy. Nevertheless, this reconstruction offers a significant advantage, as the isolines distinctly curve downward.
The salinity reconstructions appear to be significantly different than in the reanalysis reconstructions, as illustrated in figure 5.4. Notably, all satellite L4 reconstructions exhibit a consistent lower minimum of salinity in the subsurface. Additionally, the low salinity structure is predominantly confined to the surface layer, with minimal presence at greater depths. The most pronounced discrepancies between L4 reconstructions and reanalysis ones are observed in the tri-variate reconstruction, where the salinity minimum in the eddy is barely detectable. This is particularly evident



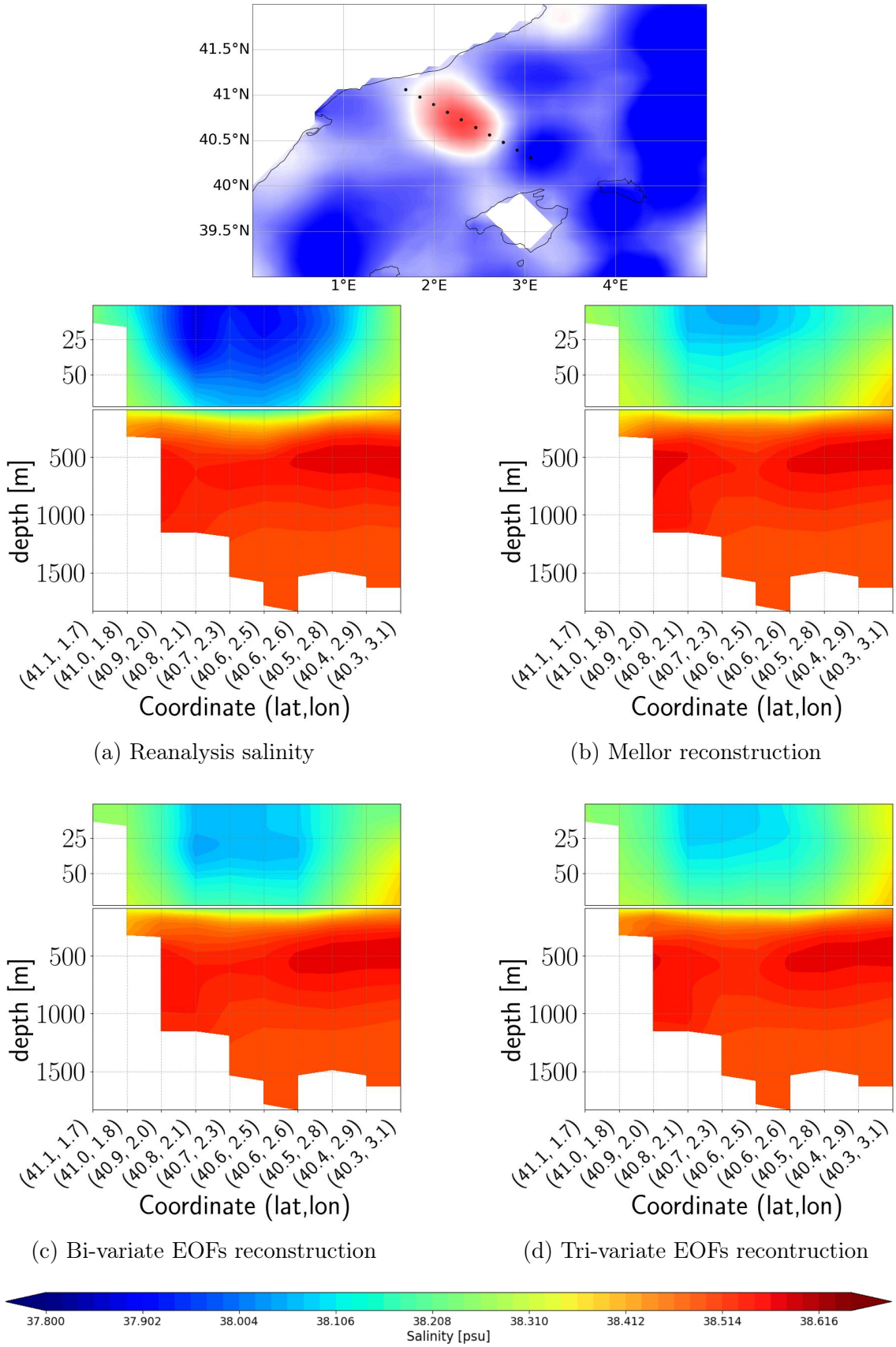
(a) Reanalysis temperature

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 5.3: Temperature of the transect crossing the eddy near Spain coast on 24 March 2019



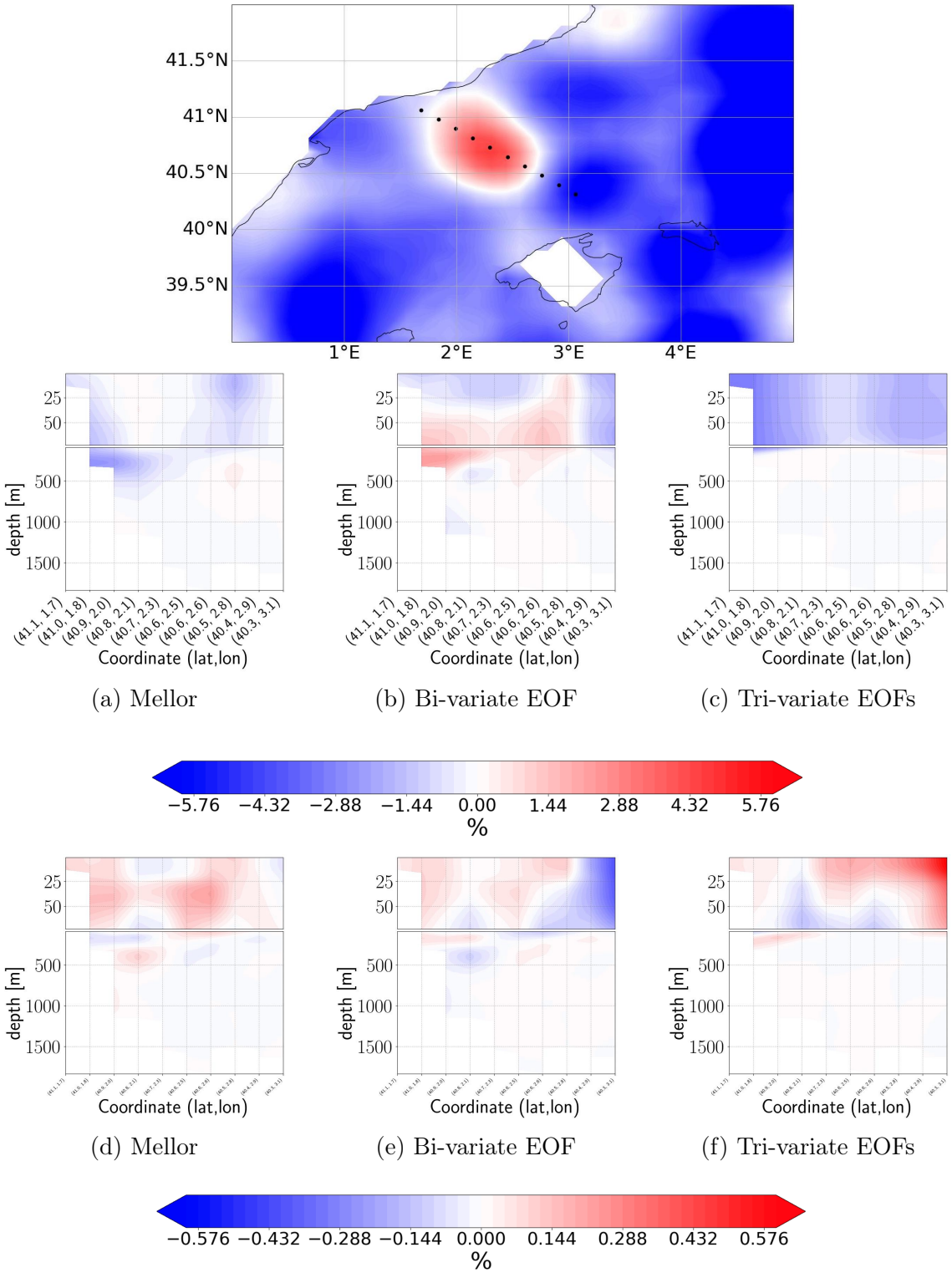
(a) Reanalysis salinity

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 5.4: Salinity of the transect crossing the eddy near Spain coast on 24 March 2019



(a) Mellor (b) Bi-variate EOF (c) Tri-variate EOFs

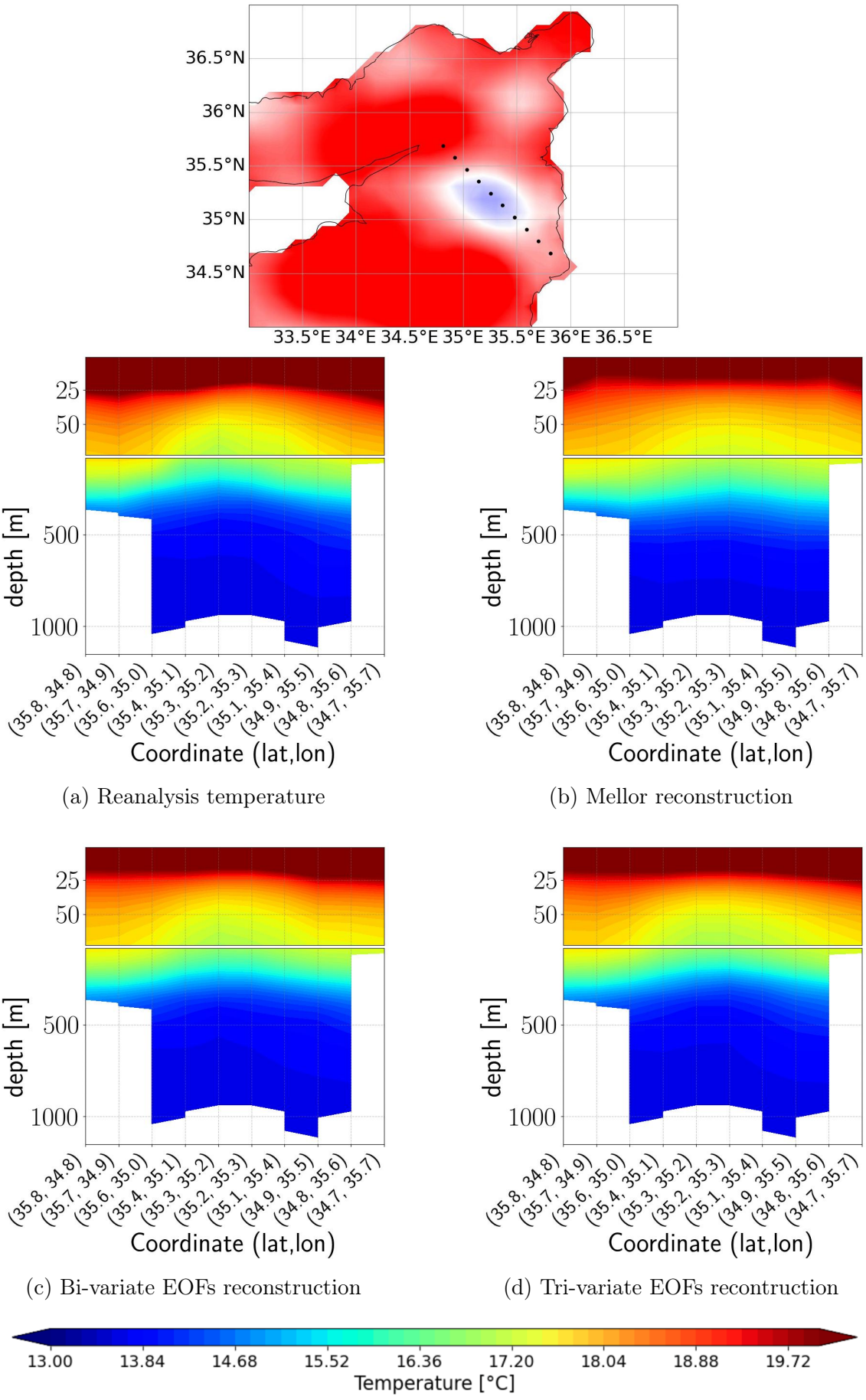
(d) Mellor (e) Bi-variate EOF (f) Tri-variate EOFs

Figure 5.5: **Percentage differences between reconstructions with SLA from reanalysis and reconstructions with SLA from satellite.** 5.5a, 5.5b and 5.5c are relative to temperature reconstructions, while 5.5d, 5.5e and 5.5f are relative to salinity reconstructions

in figure 5.5f, which reveals a larger salinity throughout the whole eddy. Conversely, the other two reconstructions appear somewhat more aligned with the reanalysis ones, although they still exhibit a notable overestimation of salinity. The only instance of salinity underestimation relative to the model data reconstruction occurs on the right side of the BE reconstruction. It is worth noting that this area was previously marked by considerable overestimation, thus representing a step forward in accuracy.

5.2.2 Cyclonic eddy near Cyprus

An examination of the basic temperature reconstructions presented in figure 5.6 reveals several different characteristics of the reconstruction by the three methods. One notable observation with respect to the anticylconic case is that the EM reconstruction appear to have higher skill in reproducing the uplift of the isotherms. The analysis of 5.8a, 5.8b, and 5.8c, confirms that the EM reconstruction is quite similar to the BE amd TE reconstructions. In the case of the BE reconstruction, it is evident that the eastern section of the eddy overestimate temperature when compared to the reanalysis reconstruction. However, it is important to note that the reanalysis reconstruction itself has larger uplift of the isotherms, suggesting that the reconstruction from the satellite provides a different representation. A similar conclusion can be drawn for the TE reconstruction, which closely resembles the reanalysis one.
In examining the salinity reconstructions presented in figure 5.7, several key elements emerge. The EM and BE reconstructions demonstrate superior performance compared to the model reconstruction. Notably, the cylindrical area of saltier water at the center of the domain has contracted, and the existence of a pole of saltier water at the surface, positioned to the right of the center, is distinctly apparent in both reconstructions. Conversely, the TE reconstruction shows minimal variation compared to previous assessments, maintaining its status as the most effective in capturing the configuration of the deeper isolines, which are elevated and exhibit a peak to the right of the center, consistent with reanalysis data.
Overall, it can be concluded that satellite reconstructions are somewhat more effective for this cyclonic eddy. This may be attributed to the inherent complexity of modeling such eddies, suggesting that the incorporation of additional satellite data enhances the reconstruction process. The tri-variate reconstruction, however, appears to be less influenced by this enhancement.



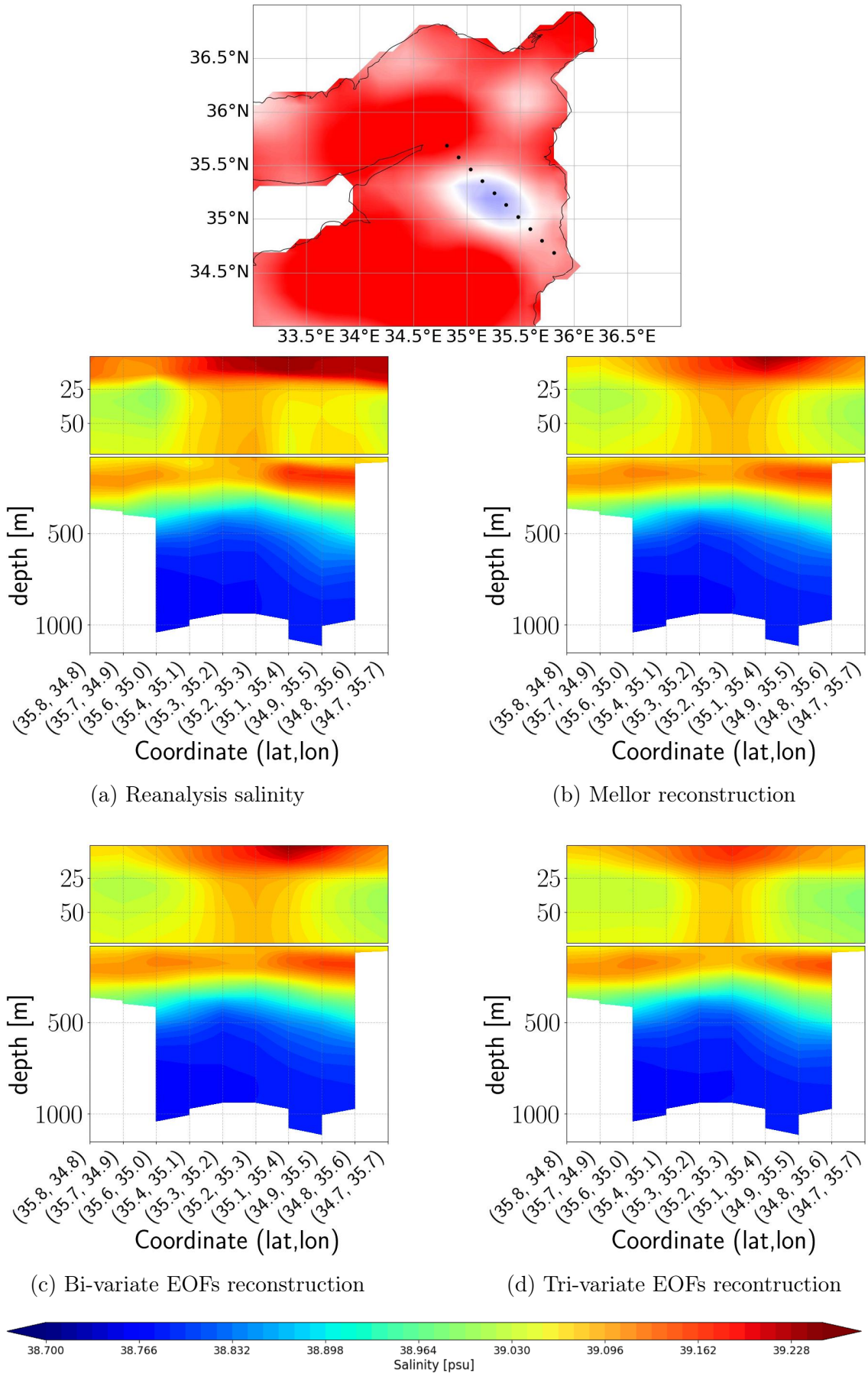
(a) Reanalysis temperature

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 5.6: Temperature of the transect crossing the eddy near Cyprus on 30 May 2020



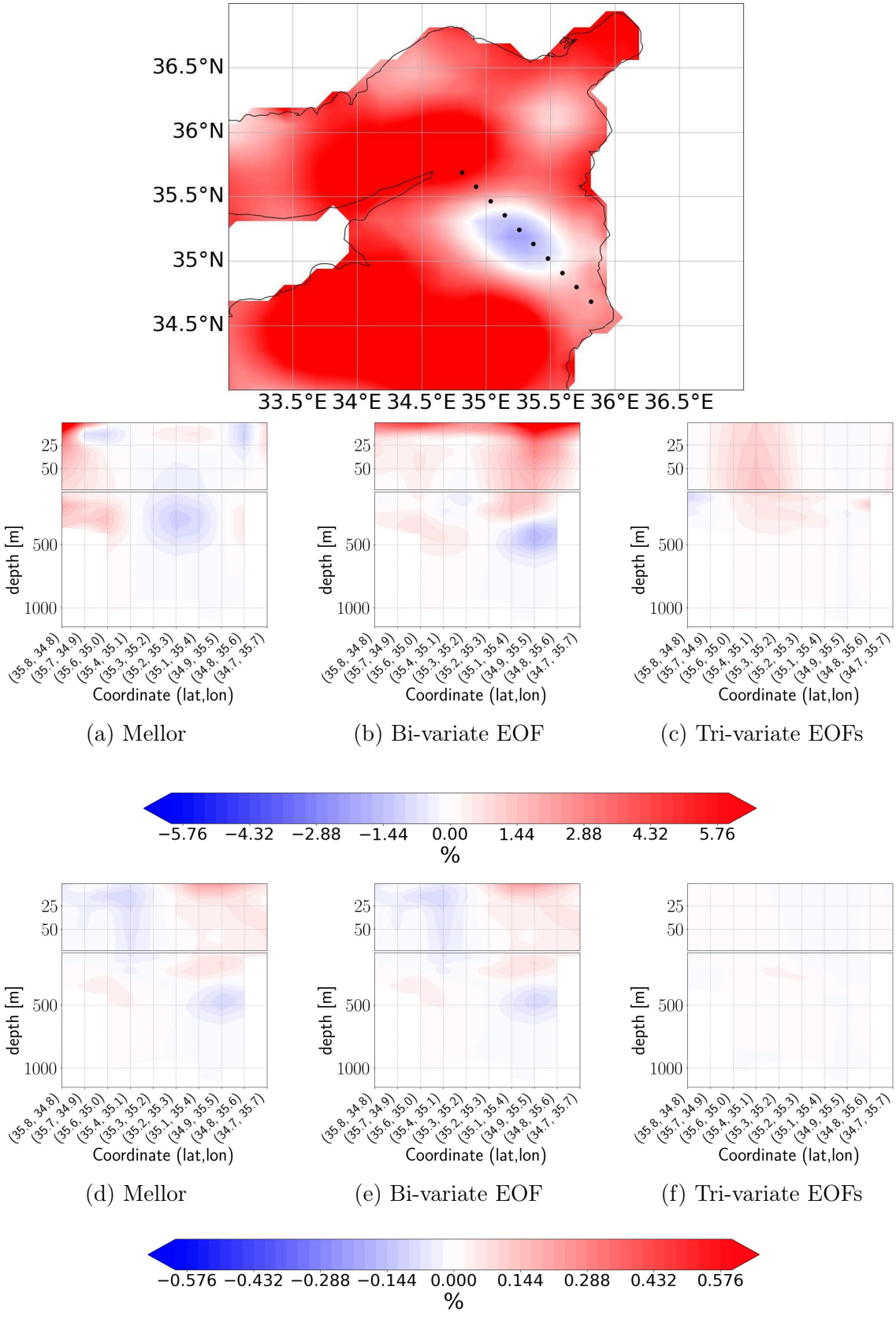
(a) Reanalysis salinity

(b) Mellor reconstruction

(c) Bi-variate EOFs reconstruction

(d) Tri-variate EOFs recontruction

Figure 5.7: Salinity of the transect crossing the eddy near Cyprus on 30 May 2020



(a) Mellor (b) Bi-variate EOF (c) Tri-variate EOFs

(d) Mellor (e) Bi-variate EOF (f) Tri-variate EOFs

Figure 5.8: **Percentage differences between reconstructions with SLA from reanalysis and reconstructions with SLA from satellite.** 5.8a, 5.8b and 5.8c are relative to temperature reconstructions, while 5.8d, 5.8e and 5.8f are relative to salinity reconstructions

Chapter 6

Conclusions and future investigation

This study has proposed several approaches for assessing the subsurface structure of the ocean using sea level anomalies. The algorithms developed aim to evaluate the quality of subsurface reconstructions using Ezer and Mellor's algorithm ([45], [46]) and Adani et al. [47] bi-variate or tri-variate EOFs. Furthermore, different averaging periods have been tested to determine the optimal methodology. To maintain consistency, we have utilized reanalysis data for temperature, salinity, and sea level, thereby estimating the absolute error in the reconstruction of reanalysis sea level. Subsequently, we utilized satellite altimeter data for sea level measurements and applied identical subsurface extrapolation techniques to derive the thermohaline structures, thereby inferring a relative error through comparison with the subsurface data from the reanalysis reconstructions.

The various methods were initially implemented at four distinct locations within the Mediterranean Sea, specifically in the Alboran, Tyrrhenian, Adriatic, and Levantine basins. Profiles from the years 2019 and 2020 were reconstructed, and several statistical error indicators — including mean bias, standard deviation error and unbiased root mean square error — were utilized to evaluate the quality of the reconstructions. The algorithms yielded consistent results across these four locations, revealing that reconstructions utilizing monthly correlation factors, evaluated against averages from the corresponding period, yielded the most accurate representations of temperature and salinity structures. However, the reconstruction of surface temperature presented a unique challenge; due to the increased variability of this layer, employing a daily variable extrapolation method enhanced the reconstruction quality. Moreover the tri-variate EOFs reconstruction, which integrates the effects of temperature and salinity on sea level, exhibited the best performance. This is consistent with the principle that sea level is influenced by both variables, as it is fundamentally linked to the integral of density and the combined temperature and salinity effects in the equation of state for density.

The algorithms utilizing correlation factors from the months of 2019-2020 have been employed to analyze two specific case studies of eddies: an anticyclonic eddy located in the Balearic Sea and a cyclonic eddy situated near Cyprus. This analysis aims to evaluate the effectiveness of the different algorithms in reconstructing the vertical structure of these eddies on a selected date. For the anticyclonic eddy, the algorithms are expected to demonstrate a downward deflection of temperature and salinity isolines, whereas for the cyclonic eddy, an upward deflection is anticipated. All reconstructions effectively represent this structural characteristics, although there

is a tendency to underestimate the slope of the isohalines and isotherms. The tri-variate EOFs reconstruction for temperature stands out as the best method, the one that effectively represents the upsloping and downsloping of isotherms. Conversely, the salinity reconstruction presents a notable exception, as the performances of Ezer and Mellor's and bi-variate EOFs reconstructions tend to outperform that of the tri-variate EOFs. This may be due to the relatively weaker correlation of salinity with sea level compared to temperature; thus, the addition of temperature may overshadow some of the variance, leading to reduced correlation factors between salinity and sea level, which adversely affects the reconstruction. Another limitation could be deriving from the limited number of trivariate EOFs used (only 25) which do not allow to reconstruct the high frequency structure of the salinity, especially in the Levantine basin.
Finally, the profiles of temperature and salinity have been reconstructed using sea level anomaly L4 satellite data. This approach was used in the past to extrapolate the temperature and salinity structure of the world ocean using only information from data [48] and here we did it using the extrapolation methods from reanalysis. In general, L4 satellite data changes the reconstruction with respect to the reanalysis but it is consistent.
The results of this investigation underscore the importance of employing a tri-variate EOFs reconstruction techniques that simultaneously considers both variables and their covariance with the sea level for effectively reconstructing the ocean's thermohaline structure from sea level. Furthermore, capturing the variability and changes within a specific period of interest yields superior results compared to employing an extended time series. This advantage arises because the algorithms used are based on straightforward statistical analyses that do not discern patterns from the time series data. Consequently, a more temporally focused dataset allows for a more accurate representation of temperature and salinity characteristics. In the future, the use of Artificial Neural Networks (ANNs), which are capable of fully harnessing the pattern information embedded in time series and can also accommodate nonlinear variability [81] should be tried. The incorporation of ANNs could lead to significant advancements in the inference of subsurface structures from satellite observations, helping to refine the error covariance matrix for the data assimilation algorithm used in reanalyses.

Acknowledgements

I want to express my sincere gratitude to my supervisor, Prof. Nadia Pinardi. She has been an incredible mentor, always there to encourage me to do my best and to spread her passion for science and oceanography.
I would also like to express my gratitude to my co-supervisors, Prof. Paolo Oddo, for his ability to always ignite the correct spark that have allowed me to proceed in my research, and Dr. Mario Adani for his patience and hawk-eye in finding errors in my code.
I hope this is the beginning of a new incredible adventure in the world of oceanography!

Bibliography

[1] Andrea Storto, Aida Alvera-Azcárate, Magdalena A. Balmaseda, Alexander Barth, Matthieu Chevallier, Francois Counillon, Catia M. Domingues, Marie Drevillon, Yann Drillet, Gaël Forget, Gilles Garric, Keith Haines, Fabrice Hernandez, Doroteaciro Iovino, Laura C. Jackson, Jean-Michel Lellouche, Simona Masina, Michael Mayer, Peter R. Oke, Stephen G. Penny, K. Andrew Peterson, Chunxue Yang, and Hao Zuo. "Ocean Reanalyses: Recent Advances and Unsolved Challenges". In: *Frontiers in Marine Science* (2019). DOI: 10.3389/fmars.2019.00418. URL: https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2019.00418.

[2] Romain Escudier, Emanuela Clementi, Andrea Cipollone, Jenny Pistoia, Massimiliano Drudi, Alessandro Grandi, Vladislav Lyubartsev, Rita Lecci, Ali Aydogdu, Damiano Delrosso, Mohamed Omar, Simona Masina, Giovanni Coppini, and Nadia Pinardi. "A High Resolution Reanalysis for the Mediterranean Sea". In: *Frontiers in Earth Science* (2021). DOI: 10.3389/feart.2021.702285. URL: https://www.frontiersin.org/journals/earth-science/articles/10.3389/feart.2021.702285.

[3] NOAA National Centers for Environmental Information. *Ocean Models*. Accessed: 04-09-2024. URL: https://www.ncei.noaa.gov/products/weather-climate-models/ocean.

[4] James C. McWilliams. "Oceanic General Circulation Models". In: *Ocean Modeling and Parameterization*. Ed. by Eric P. Chassignet and Jacques Verron. Springer Netherlands, 1998, pp. 1–44. DOI: 10.1007/978-94-011-5096-5_1. URL: https://doi.org/10.1007/978-94-011-5096-5_1.

[5] Jin-Ho Yoon and Po-Lun Ma. "Oceanic General Circulation Models ocean/oceanic general circulation models (OGCM)". In: *Encyclopedia of Sustainability Science and Technology*. Springer, 2012, pp. 7365–7381. DOI: 10.1007/978-1-4419-0851-3_356.

[6] Pierre F.J. Lermusiaux, Ching-Sang Chiu, Glen Gawarkiewicz, Phil Abbot, Allan R. Robinson, Robert N. Miller, Patrick J. Haley, Wayne G. Leslie, Sharanya J. Majumdar, Alex Pang, and Francois Lekien. "Quantifying Uncertainties in Ocean Predictions". In: *Oceanography* (2006). DOI: 10.5670/oceanog.2006.93.

[7] Robert L. Wilby and Suraje Dessai. "Robust adaptation to climate change". In: *Weather* (2010). DOI: 10.1002/wea.543. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/wea.543.

[8] R. Venkatesan, Amit Tandon, Debasis Sengupta, and K. N. Navaneeth. "Recent Trends in Ocean Observations". In: *Observing the Oceans in Real Time*. Ed. by R. Venkatesan, Amit Tandon, Eric D'Asaro, and M. A. Atmanand. Springer International Publishing, 2018, pp. 3–13. DOI: 10.1007/978-3-319-66493-4_1. URL: https://doi.org/10.1007/978-3-319-66493-4_1.

[9] Pierre-Yves Le Traon. "Satellite Oceanography for Ocean Forecasting". In: *Ocean Forecasting: Conceptual Basis and Applications*. Ed. by Nadia Pinardi and John Woods. Springer Berlin, Heidelberg, 2002, pp. 19–36. DOI: 10.1007/978-3-662-22648-3_2. URL: https://doi.org/10.1007/978-3-662-22648-3_2.

[10] NASA EarthData. *What is remote sensing?* Accessed: 06-08-2024. 2023. URL: https://www.earthdata.nasa.gov/learn/backgrounders/remote-sensing#data-processing-interpretation-and-analysis.

[11] Ian S. Robinson. *Discovering the Ocean from Space*. Springer Berlin, Heidelberg, 2010. DOI: 10.1007/978-3-540-68322-3. URL: https://doi.org/10.1007/978-3-540-68322-3.

[12] Elzbieta M. Bitner-Gregersen and Oistein Hagen. "Uncertainties in data for the offshore environment". In: *Structural Safety* (1990). DOI: 10.1016/0167-4730(90)90010-M. URL: https://www.sciencedirect.com/science/article/pii/016747309090010M.

[13] Christoph Waldmann, Philipp Fischer, Steffen Seitz, Manuela Köllner, Jens-Georg Fischer, Markus Bergenthal, Holger Brix, Stefan Weinreben, and Robert Huber. "A methodology to uncertainty quantification of essential ocean variables". In: *Frontiers in Marine Science* (2022). DOI: 10.3389/fmars.2022.1002153. URL: https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2022.1002153.

[14] Geir Evensen, Femke C. Vossepoel, and Peter Jan van Leeuwen. *Data Assimilation Fundamentals*. Springer Cham, 2022. DOI: 10.1007/978-3-030-96709-3. URL: https://doi.org/10.1007/978-3-030-96709-3.

[15] Mark Asch, Marc Bocquet, and Maëlle Nodet. *Data Assimilation*. Society for Industrial and Applied Mathematics, 2016. DOI: 10.1137/1.9781611974546.

[16] Peter Jan Van Leeuwen, Yuan Cheng, and Sebastian Reich. *Nonlinear Data Assimilation*. Springer Cham, 2015. DOI: 10.1007/978-3-319-18347-3. URL: https://doi.org/10.1007/978-3-319-18347-3.

[17] Alberto Carrassi, Marc Bocquet, Laurent Bertino, and Geir Evensen. "Data assimilation in the geosciences: An overview of methods, issues, and perspectives". In: *WIREs Climate Change* 9.5 (2018), e535. DOI: 10.1002/wcc.535. URL: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcc.535.

[18] Copernicus Climate Change Service. *Reanalysis Q&As*. Accessed: 10-09-2024. 2020. URL: https://climate.copernicus.eu/reanalysis-qas.

[19] Intergovernmental Panel on Climate Change (IPCC). "Framing, Context, and Methods". In: *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2023, pp. 147–286.

[20] Xiang Wang, Renzhi Wang, Ningzi Hu, Pinqiang Wang, Peng Huo, Guihua Wang, Huizan Wang, Senzhang Wang, Junxing Zhu, Jianbo Xu, Jun Yin, Senliang Bao, Ciqiang Luo, Ziqing Zu, Yi Han, Weimin Zhang, Kaijun Ren, Kefeng Deng, and Junqiang Song. "XiHe: A Data-Driven Model for Global Ocean Eddy-Resolving Forecasting". In: (2024). DOI: 10.48550/arXiv.2402.02995. URL: https://arxiv.org/abs/2402.02995.

[21] Allan R. Robinson and Maryam Golnaraghi. "The Physical and Dynamical Oceanography of the Mediterranean Sea". In: *Ocean Processes in Climate Dynamics: Global and Mediterranean Examples*. Ed. by Paola Malanotte-Rizzoli and Allan R. Robinson. Springer Netherlands, 1994, pp. 255–306. DOI: 10.1007/978-94-011-0870-6_12. URL: https://doi.org/10.1007/978-94-011-0870-6_12.

[22] Jacopo Chiggiato, Katrin Schroeder, Baptiste Mourre, Elda Miramontes, Piero Lionello, Marta Marcos, Nadia Pinardi, Evan Mason, Marta Álvarez, and Fabio Trincardi. "Introduction". In: *Oceanography of the Mediterranean Sea*. Ed. by Katrin Schroeder and Jacopo Chiggiato. Elsevier, 2023, pp. 1–11. DOI: 10.1016/B978-0-12-823692-5.00002-9.

[23] Katrin Schroeder, Toste Tanhua, Jacopo Chiggiato, Dimitris Velaoras, Simon A. Josey, Jesús García Lafuente, and Manuel Vargas-Yáñez. "The forcings of the Mediterranean Sea and the physical properties of its water masses". In: *Oceanography of the Mediterranean Sea*. Ed. by Katrin Schroeder and Jacopo Chiggiato. Elsevier, 2023, pp. 93–123. DOI: 10.1016/B978-0-12-823692-5.00005-4.

[24] Matthias Tomczak. "Some historical, theoretical and applied aspects of quantitative water mass analysis." In: *Journal of Marine Research* (1999).

[25] GRID-Arendal. *Mediterranean Sea water masses: vertical distribution*. Accessed: 07-09-2024. 2013. URL: https://www.grida.no/resources/5885.

[26] John Marshall and Friedrich Schott. "Open-ocean convection: Observations, theory, and models". In: *Reviews of Geophysics* (1999). DOI: https://doi.org/10.1029/98RG02739.

[27] Miroslav Gačić, Alexander Lascaratos, Beniamino B. Manca, and Anneta Mantziafou. "Adriatic Deep Water and Interaction with the Eastern Mediterranean Sea". In: *Physical Oceanography of the Adriatic Sea: Past, Present and Future*. Ed. by Benoit Cushman-Roisin, Miroslav Gačić, Pierre-Marie Poulain, and Antonio Artegiani. Springer Netherlands, 2001, pp. 111–142. DOI: 10.1007/978-94-015-9819-4_4. URL: https://doi.org/10.1007/978-94-015-9819-4_4.

[28] Alexander Theocharis, Efstathios Balopoulos, Soterios Kioroglou, Harilaos Kontoyiannis, and Athanassia Iona. "A synthesis of the circulation and hydrography of the South Aegean Sea and the Straits of the Cretan Arc (March 1994–January 1995)". In: *Progress in Oceanography* (1999). DOI: https://doi.org/10.1016/S0079-6611(99)00041-5.

[29] A.R. Robinson, Wayne Leslie, Alexander Theocharis, and A. Lascaratos. "Mediterranean Sea Circulation". In: *Ocean currents* (2001). DOI: 10.1016/B978-012374473-9.00376-3.

[30] Nadia Pinardi, Paola Cessi, Federica Borile, and Christopher L. P. Wolfe. "The Mediterranean Sea Overturning Circulation". In: *Journal of Physical Oceanography* (2019). DOI: 10.1175/JPO-D-18-0254.1. URL: https://journals.ametsoc.org/view/journals/phoc/49/7/jpo-d-18-0254.1.xml.

[31] Evan Mason, Bàrbara Barceló-Llull, Antonio Sánchez-Román, Daniel Rodríguez-Tarry, Eugenio Cutolo, Antoine Delepoulle, Simón Ruiz, and Ananda Pascual. "Chapter 8 - Fronts, eddies and mesoscale circulation in the Mediterranean Sea". In: *Oceanography of the Mediterranean Sea*. Ed. by Katrin Schroeder and Jacopo Chiggiato. Elsevier, 2023, pp. 263–287. DOI: https://doi.org/10.1016/B978-0-12-823692-5.00003-0.

[32] Lynne D. Talley, George L. Pickard, William J. Emery, and James H. Swift. "Chapter 2 - Ocean Dimensions, Shapes, and Bottom Materials". In: *Descriptive Physical Oceanography (Sixth Edition)*. Ed. by Lynne D. Talley, George L. Pickard, William J. Emery, and James H. Swift. Academic Press, 2011, pp. 7–27. DOI: 10.1016/B978-0-7506-4552-2.10002-2. URL: https://www.sciencedirect.com/science/article/pii/B9780750645522100022.

[33] Valeria Di Biagio, Gianpiero Cossarini, Stefano Salon, Paolo Lazzari, Stefano Querin, Gianmaria Sannino, and Cosimo Solidoro. "Temporal scales of variability in the Mediterranean Sea ecosystem: Insight from a coupled model". In: *Journal of Marine Systems* (2019). DOI: https://doi.org/10.1016/j.jmarsys.2019.05.002.

[34] Nadia Pinardi. "Physical Oceanography - Notes for Master Degree program in Earth System Science". 2022.

[35] Paolo Oddo. "Note Ocean Dynamics". Notes for Master Degree program in Science of Climate. 2022.

[36] Geophysical Fluid Dynamics Laboratory - NOAA. *Ocean mesoscale eddies*. Accessed: 09-09-2024. URL: https://www.gfdl.noaa.gov/ocean-mesoscale-eddies/.

[37] Antonio Bonaduce, Andrea Cipollone, Johnny Johannessen, Joanna Staneva, Roshin Raj, and Ali Aydoğdu. "Ocean Mesoscale Variability: A Case Study on the Mediterranean Sea From a Re-Analysis Perspective". In: *Frontiers in Earth Science* (2021). DOI: 10.3389/feart.2021.724879.

[38] Benoit Cushman-Roisin and Jean-Marie Beckers. "Fronts, Jets and Vortices". In: *Introduction to Geophysical Fluid Dynamics*. Academic Press, 2011, pp. 543–575. DOI: 10.1016/B978-0-12-088759-0.00018-3.

[39] Dudley B. Chelton, Michael G. Schlax, and Roger M. Samelson. "Global observations of nonlinear mesoscale eddies". In: *Progress in Oceanography* (2011). DOI: https://doi.org/10.1016/j.pocean.2011.01.002.

[40] Keith Haines. "Assimilation of Satellite Altimetry in Ocean Models". In: *Ocean Forecasting: Conceptual Basis and Applications*. Ed. by Nadia Pinardi and John Woods. Springer Berlin Heidelberg, 2002, pp. 117–130. DOI: 10.1007/978-3-662-22648-3_7. URL: https://doi.org/10.1007/978-3-662-22648-3_7.

[41] Mike Cooper and Keith Haines. "Altimetric assimilation with water property conservation". In: *Journal of Geophysical Research: Oceans* (1996). DOI: https://doi.org/10.1029/95JC02902. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95JC02902.

[42] Romain Escudier, Lionel Renault, Ananda Pascual, Pierre Brasseur, Dudley Chelton, and Jonathan Beuvier. "Eddy properties in the Western Mediterranean Sea from satellite altimetry and a numerical simulation". In: *Journal of Geophysical Research: Oceans* (2016). DOI: https://doi.org/10.1002/2015JC011371. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JC011371.

[43] Matthias Tomczak and J. Stuart Godfrey. "CHAPTER 3 - The Coriolis force, geostrophy, Rossby waves and the westward intensification". In: *Regional Oceanography*. Ed. by Matthias Tomczak and J. Stuart Godfrey. Pergamon, 1994, pp. 29–37. DOI: https://doi.org/10.1016/B978-0-08-041021-0.50007-2.

[44] Peter B. Rhines. "Mesoscale Eddies". In: *Encyclopedia of Ocean Sciences (Third Edition)*. Ed. by J. Kirk Cochran, Henry J. Bokuniewicz, and Patricia L. Yager. Third Edition. Academic Press, 2019, pp. 115–127. DOI: https://doi.org/10.1016/B978-0-12-409548-9.11642-2.

[45] George Mellor and Tal Ezer. "A Gulf Stream model and an altimetry assimilation scheme". In: *Journal of Geophysical Research* (1991). DOI: 10.1029/91JC00383.

[46] Tal Ezer and George Mellor. "Continuous Assimilation of Geosat Altimeter Data into a Three-Dimensional Primitive Equation Gulf Stream Model". In: *Journal of Physical Oceanography - J PHYS OCEANOGR* (1994). DOI: 10.1175/1520-0485(1994)024<0832:CAOGAD>2.0.CO;2.

[47] Mario Adani, Srdjan Dobricic, and Nadia Pinardi. "Quality Assessment of a 1985–2007 Mediterranean Sea Reanalysis". In: *Journal of Atmospheric and Oceanic Technology* (2011). DOI: 10.1175/2010JTECHO798.1. URL: https://journals.ametsoc.org/view/journals/atot/28/4/2010jtecho798_1.xml.

[48] S. Guinehut, Anne-Lise Dhomps, Larnicol Gilles, and Pierre-Yves Traon. "High Resolution 3-D temperature and salinity fields derived from in situ and satellite observations". In: *Ocean Science Discussions* (2012). DOI: 10.5194/osd-9-1313-2012.

[49] Dean Roemmich, Philip Woodworth, Svetlana Jevrejeva, Sarah Purkey, Matthias Lankhorst, Uwe Send, and Nikolai Maximenko. "In Situ Observations Needed to Complement, Validate, and Interpret Satellite Altimetry". In: *Satellite Altimetry over Oceans and Land Surfaces*. Ed. by Anny Cazenave Detlef Stammer. CRC Press, 2017, pp. 113–148. DOI: 10.1201/9781315151779-3.

[50] Stephen Griffies. "Ocean primitive equations and sea level equations - Notes for the school Advanced School on Earth System Modelling". 2016.

[51] Detlef Stammer and Stephen Griffies. "Ocean Modeling and Data Assimilation in the Context of Satellite Altimetry". In: *Satellite Altimetry over Oceans and Land Surfaces*. CRC Press, 2017, pp. 547–580. DOI: 10.1201/9781315151779-17.

[52] A.E. Gill and P.P. Niller. "The theory of the seasonal variability in the ocean". In: *Deep Sea Research and Oceanographic Abstracts* (1973). DOI: https://doi.org/10.1016/0011-7471(73)90049-1.

[53] N. Pinardi, A. Bonaduce, A. Navarra, S. Dobricic, and P. Oddo. "The Mean Sea Level Equation and Its Application to the Mediterranean Sea". In: *Journal of Climate* (2014). DOI: 10.1175/JCLI-D-13-00139.1. URL: https://journals.ametsoc.org/view/journals/clim/27/1/jcli-d-13-00139.1.xml.

[54] R. Nerem, Michaël Ablain, Anny Cazenave, John Church, and Eric Leuliette. "A 25-Year Satellite Altimetry-Based Global Mean Sea Level Record". In: *Satellite Altimetry over Oceans and Land Surfaces*. CRC Press, 2017, pp. 187–210. DOI: 10.1201/9781315151779-5.

[55] Philippe Escudier, Alexandre Couhert, Flavien Mercier, A. Mallet, Pierre Thibaut, Ngan Tran, Laïba Amarouche, Bruno Picard, Loren Carrere, Dibarboure Gerald, Michaël Ablain, Jacques Richard, Nathalie Steunou, Pierre Dubois, Marie-Helene Rio, and Dorandeu Joël. "Satellite Radar Altimetry". In: *Satellite Altimetry over Oceans and Land Surfaces*. Ed. by Anny Cazenave Detlef Stammer. CRC Press, 2017, pp. 1–70. DOI: 10.1201/9781315151779-1.

[56] Aviso+ satellite altimetry data. *How altimetry works*. Accessed: 09-08-2024. URL: https://www.aviso.altimetry.fr/en/techniques/altimetry/principle/basic-principle.html.

[57] Margaret Srinivasan and Vardis Tsontos. "Satellite Altimetry for Ocean and Coastal Applications: A Review". In: *Remote Sensing* (2023). DOI: 10.3390/rs15163939. URL: https://www.mdpi.com/2072-4292/15/16/3939.

[58] Ian S. Robinson. *Measuring the Oceans from Space*. Springer Berlin, Heidelberg, 2004.

[59] NOAA coastwatch. *Detecting Sea Level Anomalies with Satellites*. Accessed: 09-08-2024. 2021. URL: https://coastwatch.noaa.gov/cwn/news/2021-09-23/detecting-sea-level-anomalies-satellites.html.

[60] E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). *European Seas Gridded L 4 Sea Surface Heights And Derived Variables Reprocessed 1993 Ongoing*. Accessed: 05-01-2024. DOI: 10.48670/moi-00141.

[61] Madec G. and the NEMO team. *NEMO ocean engine. Note du Pôle de modélisation*. Institut Pierre-Simon Laplace (IPSL), 2008.

[62] Emanuela Clementi, Jenny Pistoia, Damiano Delrosso, Gelsomina Mattia, Claudia Fratianni, Andrea Storto, Stefania Ciliberti, Benedicte Lemieux-Doudon, Elisa Fenu, Simona Simoncelli, Massimiliano Drudi, Alessandro Grandi, Davide Padeletti, Pierluigi Pietro, and Pinardi Nadia. "A 1/24° resolution Mediterranean physical analysis and forecasting system for the Coperincus Marine Environment Monitoring Service". In: *Extended abstract to the 8th EuroGOOS Conference, Bergen*. 2018.

[63] A. C. Lorenc. "Analysis methods for numerical weather prediction". In: *Quarterly Journal of the Royal Meteorological Society* (1986). DOI: https://doi.org/10.1002/qj.49711247414. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711247414.

[64] Srdjan Dobricic and Nadia Pinardi. "An oceanographic three-dimensional variational data assimilation scheme". In: *Ocean Modelling* (2008). DOI: 10.1016/j.ocemod.2008.01.004. URL: https://www.sciencedirect.com/science/article/pii/S1463500308000176.

[65] Andrea Storto, Simona Masina, and Antonio Navarra. "Evaluation of the CMCC eddy-permitting global ocean physical reanalysis system (C-GLORS, 1982–2012) and its assimilation components". In: *Quarterly Journal of the Royal Meteorological Society* (2016). DOI: 10.1002/qj.2673. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2673.

[66] E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). *Mediterranean Sea Physics Reanalysis*. Accessed: 06-01-2024. DOI: 10.48670/moi-00141.

[67] E.N. Lorenz. "Empirical Orthogonal Functions and Statistical Weather Prediction". In: *Scientific Report No1, Statistical Forecasting Project* (1956).

[68] Douglas Martinson. "Empirical Orthogonal Function (EOF) Analysis". In: *Quantitative Methods of Data Analysis for the Physical Sciences and Engineering*. Cambridge University Press, 2018, pp. 495–543. DOI: 10.1017/9781139342568.020.

[69] Jose P. Peixoto and Abraham H. Oort. *Physics of Climate*. American Institute of Physics Melville, 1992.

[70] Korres Gerasimos, Pinardi Nadia, and A. Lascaratos. "The Ocean Response to Low-Frequency Interannual Atmospheric Variability in the Mediterranean Sea. Part II: Empirical Orthogonal Functions Analysis". In: *Journal of Climate - J CLIMATE* (2000). DOI: 10.1175/1520-0442(2000)013<0732:TORTLF>2.0.CO;2.

[71] S. Dobricic, N. Pinardi, M. Adani, A. Bonazzi, C. Fratianni, and M. Tonani. "Mediterranean Forecasting System: An improved assimilation scheme for sea-level anomaly and its validation". In: *Quarterly Journal of the Royal Meteorological Society* (2005). DOI: 10.1256/qj.05.100. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.100.

[72] Jenny Pistoia, Emanuela Clementi, Damiano Delrosso, Gelsomina Mattia, Claudia Fratianni, Massimiliano Drudi, Alessandro Grandi, Davide Padeletti, Pierluigi Pietro, Andrea Storto, and Pinardi Nadia. "Last improvements in the data assimilation scheme for the Mediterranean Analysis and Forecast system of the Copernicus Marine Service". In: *Operational Oceanography serving Sustainable Marine Development*. 2018.

[73] J.C. Gascard and C. Richez. "Water masses and circulation in the Western Alboran sea and in the Straits of Gibraltar". In: *Progress in Oceanography* (1985). DOI: https://doi.org/10.1016/0079-6611(85)90031-X.

[74] M. Astraldi and G. P. Gasparini. "The Seasonal Characteristics of the Circulation in the Tyrrhenian Sea". In: *Seasonal and Interannual Variability of the Western Mediterranean Sea*. American Geophysical Union (AGU), 1994, pp. 115–134. DOI: 10.1029/CE046p0115. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/CE046p0115.

[75] Giuseppe Aulicino, Yuri Cotroneo, Teodosio Lacava, Giancanio Sileo, Giannetta Fusco, Ryan Carlon, Valeria Satriano, Nicola Pergola, Valerio Tramutoli, and Giorgio Budillon. "Results of the first Wave Glider experiment in the southern Tyrrhenian Sea". In: *Advances in Oceanography and Limnology* (2016). DOI: 10.4081/aiol.2016.5682. URL: https://www.pagepressjournals.org/aiol/article/view/5682.

[76] A. Artegiani, E. Paschini, A. Russo, D. Bregant, F. Raicich, and N. Pinardi. "The Adriatic Sea General Circulation. Part I: Air–Sea Interactions and Water Mass Structure". In: *Journal of Physical Oceanography* (1997). DOI: 10.1175/1520-0485(1997)027<1492:TASGCP>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/phoc/27/8/1520-0485_1997_027_1492_tasgcp_2.0.co_2.xml.

[77] Milena Menna, Riccardo Gerin, Giulio Notarstefano, Elena Mauri, Antonio Bussani, Massimo Pacciaroni, and Pierre-Marie Poulain. "On the Circulation and Thermohaline Properties of the Eastern Mediterranean Sea". In: *Frontiers in Marine Science* (2021). DOI: 10.3389/fmars.2021.671469. URL: https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2021.671469.

[78] Oddo P., Falchetti S., Viola S., Pennucci G., Storto A., Borrione I., Giorli G., Cozzani E., Russo A., and Tollefsen C. "Evaluation of different Maritime rapid environmental assessment procedures with a focus on acoustic performance". In: *The Journal of the Acoustical Society of America* (2022). DOI: 10.1121/10.0014805.

[79] Frederick M. Bingham and Tong Lee. "Space and time scales of sea surface salinity and freshwater forcing variability in the global ocean (60°S–60°N)". In: *Journal of Geophysical Research: Oceans* (2017). DOI: 10.1002/2016JC012216. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JC012216.

[80] Junlin Ran, Nengfang Chao, Lianzhe Yue, Gang Chen, Zhengtao Wang, Tangting Wu, and Chengchun Li. "Quantifying the contribution of temperature, salinity, and climate change to sea level rise in the Pacific Ocean: 2005-2019". In: *Frontiers in Marine Science* (2023). DOI: 10.3389/fmars.2023.1200883. URL: https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2023.1200883.

[81] Kai Mao, Chang Liu, Shaoqing Zhang, and Feng Gao. "Reconstructing Ocean Subsurface Temperature and Salinity from Sea Surface Information Based on Dual Path Convolutional Neural Networks". In: *Journal of Marine Science and Engineering* (2023). DOI: 10.3390/jmse11051030. URL: https://www.mdpi.com/2077-1312/11/5/1030.